frontiers | Research Topics

Data governance in African health research: ELSI challenges and solutions

Edited by
Donrich Thaldar, Simisola Oluwatoyin Akintola and Bonginkosi Shozi

Published in
Frontiers in Pharmacology
Frontiers in Genetics

ISSN 1664-8714
ISBN 978-2-8325-6179-9
DOI 10.3389/978-2-8325-6179-9

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Data governance in African health research: ELSI challenges and solutions

Topic editors

Donrich Thaldar — University of KwaZulu-Natal, South Africa

Simisola Oluwatoyin Akintola — University of Ibadan, Nigeria

Bonginkosi Shozi — University of California, San Diego, United States

Citation

Thaldar, D., Akintola, S. O., Shozi, B., eds. (2025). *Data governance in African health research: ELSI challenges and solutions*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-6179-9

Table of contents

frontiers | Frontiers in Pharmacology

TYPE Editorial
PUBLISHED 11 December 2024
DOI 10.3389/fphar.2024.1525434

OPEN ACCESS

EDITED AND REVIEWED BY
Dov Greenbaum,
Yale University, United States

*CORRESPONDENCE
Donrich Thaldar,
thaldard@ukzn.ac.za

RECEIVED 09 November 2024
ACCEPTED 25 November 2024
PUBLISHED 11 December 2024

CITATION
Thaldar D (2024) Editorial: Data governance in African health research: ELSI challenges and solutions.
Front. Pharmacol. 15:1525434.
doi: 10.3389/fphar.2024.1525434

Editorial: Data governance in African health research: ELSI challenges and solutions

Donrich Thaldar*

University of KwaZulu-Natal, Durban, South Africa

KEYWORDS

consent, health research, data governance, open science, common heritage of humanity, genomic sovereignty, pseudonymised data, AI

Editorial on the Research Topic
Data governance in African health research: ELSI challenges and solutions

This Research Topic, *Data Governance in African Health Research: ELSI Challenges and Solutions*, brings together analyses that address the emerging legal, ethical, and social issues surrounding data governance in African health research. As health research in Africa increasingly utilises digital data, artificial intelligence (AI), and genomic technologies, these articles explore the path forward, offering practical and legal insights into how Africa's unique challenges can be addressed. Together, these contributions set out a forward-looking vision, guiding data governance toward a framework that respects participant rights, aligns with African regulatory environments, and adapts to evolving technological and ethical demands.

Legal and ethical frameworks for consent: empowering research participants

At the heart of data governance lies the principle of consent, which is both a legal requirement and an ethical commitment to participant autonomy. In *Introducing Dynamic Consent for Improved Trust and Privacy in Research Involving Human Biological Material and Associated Data in South Africa*, Prinsen advocates for dynamic consent, a model that aligns with South Africa's Protection of Personal Information Act 4 of 2013 (POPIA) by granting participants ongoing control over their data. This approach recognises the dynamic nature of health data use, positioning participants as active decision-makers, which strengthens both legal compliance and public trust.

Naidoo's *Open Optimism as an "Embodied-Health" Ethic for the Information Era* offers a complementary vision, presenting an "embodied" ethical approach to health data that challenges traditional divides in health governance. By valuing openness and participant engagement, Naidoo highlights the role of ethics in reinforcing participant agency and enhancing transparency in African health research. Together, Prinsen and Naidoo's contributions envision an African health research environment where ethical frameworks are responsive, participatory, and empower individuals while upholding shared values.

Legal instruments and data management: establishing robust governance for health research

A clear legal and data management framework is crucial for protecting privacy and upholding ethical standards in health research. This Research Topic includes several articles addressing these aspects, presenting actionable insights for researchers, institutions, and policymakers. *The Anatomy of a Data Transfer Agreement for Health Research* by Swales et al. provides a comprehensive guide to creating data transfer agreements (DTAs) that align with data protection legislation, such as POPIA. This article underscores the importance of detailed DTA provisions that protect both data privacy and the legal interests of institutions when sharing sensitive research data. The research presented in this article formed the foundation for a freely accessible DTA template that was developed for the South African research community (Swales et al., 2023; Thaldar et al., 2024a; Thaldar et al., 2024b).

Building on the theme of legal compliance, *A Data Management Plan for the NESHIE Observational Study* by Strydom et al. offers a template for constructing data management plans (DMPs) that address the lifecycle of sensitive health data, including security, storage, and access considerations. This template is especially relevant for studies in low- and middle-income countries (LMICs), highlighting ways to balance the complexity of data management with rigorous compliance measures. By adapting this DMP framework, institutions can establish clear standards that both protect participant privacy and meet legal obligations.

Adding further depth to the discussion, *Forcing a Square into a Circle: Why South Africa's Draft Revised Material Transfer Agreement is Not Fit for Purpose* by Esselaar et al. critiques South Africa's National Health Research Ethics Council's (NHREC) draft revision of South Africa's standard material transfer agreement (MTA). The authors draw on the foundational work done by Thaldar et al. (2022) that explored the various legal dimensions of genetic data under South African law—including privacy, ownership, and intellectual property rights—but go further by advocating for a decolonial approach to health research governance, urging the NHREC to empower local research institutions by acknowledging their ownership of the data that they collect and generate.

Open science: transparency and access in genomic research

Open science principles in genomic research promote accessible scientific knowledge and equitable benefit-sharing. In *A Pathway to Strengthening Open Science: Comments on the Draft South African Ethics in Health Research Guidelines*, Gooden critiques South Africa's NHREC draft South African Ethics in Health Research Guidelines, advocating for ethics guidelines that incorporate open science principles and promote African-centric approaches to enhance transparency and legal compliance in African health research. Gooden's recommendations align with *Open Science and Human Genetic Data: Recommendations on South Africa's Draft National Open Science Policy*, where Thaldar et al. underscore the importance of the right to freedom of scientific research, the legal difference between human and non-human genetic data, and data ownership. Importantly, open science does not require data to become public property. Instead, data can remain private property, allowing data originators to benefit while fostering responsible sharing.

Common heritage vs. genomic sovereignty: competing frameworks in genomic research

In *The Human Genome as the Common Heritage of Humanity* (Kabata and Thaldar) and *Regulating Human Genomic Research in Africa: Why a Human Rights Approach Is a More Promising Conceptual Framework than Genomic Sovereignty* (Kabata and Thaldar) examine two approaches to human genomic data as forms of public property, highlighting their practical and ethical implications.

The "common heritage" model views the human genome—often represented by the human reference genome—as a shared asset that belongs to all of humanity. This concept, grounded in international human rights, aims to protect genomic data from privatisation by framing it as an international public good, freely accessible for scientific advancement and collaboration. The focus is on global inclusivity, with genomic resources managed for the collective benefit of humanity.

In contrast, the genomic sovereignty model shifts from a global perspective to a national or community-based one, claiming that genomic data is the exclusive property of specific groups or nations. This approach, driven by concerns over resource exploitation and national interests, empowers countries or population groups to assert control over their genetic resources, restricting external access to protect local interests. The genomic sovereignty model, however, has been criticised for limiting international collaboration.

Kabata and Thaldar propose that a human rights-based framework offers a more balanced and ethical pathway. This approach respects individuals' rights to benefit from scientific advancements while allowing for private ownership of genomic data.

Foundational concepts in data law: considering pseudonymised data

In *Does Data Protection Law in South Africa Apply to Pseudonymised Data?*, Thaldar examines whether pseudonymised datasets fall under POPIA in South Africa, arguing that identifiability—and therefore POPIA's applicability—depends on the specific context of the party handling the data. By interpreting POPIA's exclusions clause and research exception through established South African legal principles, Thaldar concludes that identifiability should be assessed contextually: A dataset remains personal information for a provider retaining both the pseudonymised and linking datasets, but becomes non-personal for a recipient without access to linking data. This approach balances privacy protection with data-sharing flexibility, enabling responsible, context-sensitive data management in health research. Thaldar's insights are particularly valuable as South African institutions navigate complex privacy demands, highlighting the need for legal clarity in an evolving research environment.

AI as the new frontier: defining agency and liability in health research

AI is advancing rapidly in health research, presenting new legal and ethical challenges. *Mapping the Regulatory Landscape of AI in Healthcare in Africa* by Townsend et al. survey AI regulations across 12 African countries, identifying regulatory gaps and recommending cohesive frameworks that can support ethical AI adoption. Townsend et al. highlight that Africa must develop robust AI governance to balance innovation with protection for participants, establishing a regulatory foundation for the continent's AI future.

Liability for Harm Caused by AI in Healthcare: An Overview of the Core Legal Concepts by Bottomley and Thaldar, explores liability Research Topic specific to AI in healthcare, considering legal approaches such as strict liability and the principal–agent relationship. These frameworks aim to clarify accountability in cases where AI systems cause harm, underscoring the need for legal structures that can address AI's unique risks.

Adding depth to AI governance, Naidoo's *What Does It Mean to Be an Agent?* proposes a practical framework for assessing AI agency, focusing on empirical characteristics rather than abstract notions like consciousness. This grading system provides a structured, adaptable model for regulating AI, considering both legal accountability and suitability for specific research contexts. Naidoo's *The Open Ontology and Information Society* further frames AI governance within a broad ethical and legal structure, proposing a qualitative analysis of information to inform regulatory approaches. These articles lay a groundwork for legally and ethically responsible AI governance in African healthcare, ensuring that AI serves the public good within a framework of accountability and participant protection.

Building an African framework for health research governance

Together, these articles provide a comprehensive guide to advancing data governance in African health research. Each contribution demonstrates a commitment to addressing Africa's unique challenges, from privacy laws and consent frameworks to policy guidance and AI governance. This Research Topic offers insights that will help shape Africa's health research landscape in a way that respects individual rights, supports responsible innovation, and aligns with evolving ethical and legal standards. The future of data governance in African health research is a complex, rapidly evolving field, but with a foundation rooted in law and ethics, African researchers and policymakers are well-positioned to navigate it with confidence.

Author contributions

DT: Writing–original draft, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The author acknowledges the support by the US National Institute of Mental Health and the US National Institutes of Health (award number U01MH127690).

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author declared that he was an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author declares that this editorial was prepared with iterative drafting assistance from ChatGPT, an AI language model developed by OpenAI.

Publisher's note

Author disclaimer

The content of this essay is solely the author's responsibility and does not necessarily represent the official views of the US National Institute of Mental Health or the US National Institutes of Health.

References

Swales, L., Botes, M., Donnelly, D., and Thaldar, D. (2023). Towards a data transfer agreement for the South African research community: the empowerment approach. *S. Afr. J. Bioeth. Law*. 16 (1), 13–18. doi:10.7196/SAJBL.2023.v16i1.827

Thaldar, D., Botes, M., Swales, L., and Esselaar, P. (2024a). A data transfer agreement template for South Africa v1. *Zenodo* 1 (1.1). doi:10.5281/zenodo.11355144

Thaldar, D., Botes, M., Swales, L., and Esselaar, P. (2024b). Enhancing data governance in collaborative research: introducing SA DTA 1.1. *S. Afr. J. Bioeth. Law*. 17 (2), e2300. doi:10.7196/SAJBL.2024.v17i2.2300

Thaldar, D. W., Townsend, B. A., Donnelly, D. L., Botes, M., Gooden, A., van Harmelen, J., et al. (2022). The multidimensional legal nature of personal genomic sequence data: a South African perspective. *Front. Genet.* 13, 997595–11. doi:10.3389/fgene.2022.997595

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 30 June 2023
DOI 10.3389/fgene.2023.1208606

OPEN ACCESS

EDITED BY
Alessandro Blasimme,
ETH Zürich, Switzerland

REVIEWED BY
Judit Sándor,
Central European University, Hungary
Bartha Maria Knoppers,
McGill University, Canada

*CORRESPONDENCE
Faith Kabata,
faithkabata@gmail.com

RECEIVED 19 April 2023
ACCEPTED 16 May 2023
PUBLISHED 30 June 2023

CITATION
Kabata F and Thaldar D (2023), Regulating human genomic research in Africa: why a human rights approach is a more promising conceptual framework than genomic sovereignty.
Front. Genet. 14:1208606.
doi: 10.3389/fgene.2023.1208606

Regulating human genomic research in Africa: why a human rights approach is a more promising conceptual framework than genomic sovereignty

Faith Kabata[1]* and Donrich Thaldar[1,2]

[1]School of Law, University of KwaZulu-Natal, Durban, South Africa, [2]Petrie-Flom Center for Health Law Policy, Biotechnology and Bioethics, Harvard Law School, Cambridge, MA, United States

This article revisits the debate on the regulation of human genomic research, with a focus on Africa. The article comprehensively examines the concept of genomic sovereignty, which was invoked mainly in the global South as a conceptual framework for state regulation of human genomic research. It demonstrates that genomic sovereignty has no utility value in human genomic research as it violates the rights of individuals and researchers. By analysing Mexico's regulatory approach based on genomic sovereignty and a divergent regulatory approach, viz Finland's human genomic research framework, we show that a human rights approach is more promising as it aligns with the state obligations under the right of everyone to participate in and benefit from scientific progress and its applications in international human rights law. We conclude by recommending that African states should anchor regulation of human genomic research on a human rights framework based on the right to science.

KEYWORDS

human genomics, genomic sovereignty, human rights, regulation, right to science, Africa

1 Introduction

Is state regulation of access to and use of genomic material the appropriate governance framework for human genomic research? This article offers insights in the debate on regulation of human genomic research. It does this by examining the concept of genomic sovereignty to enquire if the concept has utility value in human genomic research, particularly in relation to Africa. The short answer is "no." The longer answer, and specifically why the concept has no utility value, however, leads to important questions on: the concept, its underpinnings, assumptions and weaknesses; other divergent state approaches in regulation of human genomic research; and questions on where state approaches to the regulation of human genomic research should focus their attention.

To situate the discussion, the article revisits deliberations of the UNESCO International Bioethics Committee during the drafting of the Universal Declaration on the Human Genome and Human Rights. A member of the Committee stated (UNESCO International Bioethics Committee, 1995):

> "We are now proposing to include the human genome in the common heritage of humanity. Legally speaking, this would be a historic and revolutionary measure, fraught with implications and attended by many consequences that would be beneficial to humanity."

Illustratively, the Universal Declaration on the Human Genome and Human Rights refers to the genome as "heritage of humankind," thus seemingly opening the possibility of the common heritage framework of governance (UNESCO, 1997). In response to this, a number of countries in the global South invoked sovereign claims over the genetic material of their citizens as a perceived way of protecting it from foreign exploitation by the global North. These claims were embedded in the concept of genomic sovereignty, which assumed a political and scientific agenda and was incorporated into domestic legislation or guidelines in some countries. Importantly, the concept put into sharp focus the role of the state in regulating access to and use of genomic resources, an issue that remains relevant for Africa.

The article flows as follows: Part 2 draws on an array of genomic sovereignty literature to map out its argumentative logic, and identify its themes and conceptual weaknesses. Part 3 examines Africa's academic engagement with genomic sovereignty, the conceptual underpinnings, and critically assesses the concept. Part 4 reviews Finland's approach to human genomic research, its philosophical underpinnings, and conducts a comparative analysis of both the Mexican and Finnish approaches. Part 5 addresses the question of where approaches to regulation of human genomic research should turn by exploring state obligations in the right to science and specifically in human genomic research. Part 6 summarises the article by offering concluding thoughts.

The article uses the terms developing and developed countries and global South and global North interchangeably, as they are used in the literature referenced. The article also acknowledges that the definition of genomic sovereignty in the literature relied on the terms *control* and *ownership* of genetic resources interchangeably, which is legally problematic.

2 Unearthing genomic sovereignty

In the context of global research, genomic sovereignty has been referred to as the ability of a nation, people or state to own and regulate access to and use samples, data and knowledge on human genes (Slabbert and Pepper, 2010).

The term was coined by Mexican scientists, politicians and policy makers as a biopolitical concept describing political sovereignty in genome mapping and was aimed at protecting national genomics in Mexico (Marìn-Schwartz, 2011). The concept was typified by establishment of the National Institute of Genomic Medicine (INMEGEN) in 2004; the mapping of the "Mexican genome" by the INMEGEN between 2004 and 2009; and the framing of the policy agenda into legislation to protect Mexico's "genomic sovereignty" in 2008 (Vasquez and García-Deister, 2019). The legislation referred to as Mexico's policy on genomic sovereignty instituted amendments in the General Health Law (Marìn-Schwartz, 2011).

The law was designed to regulate everything in the human genome in Mexico. Accordingly, it restricted the movement of biological samples outside Mexico for population genomics studies without express authority from the Secretary of Health and attached penalties of 15 years imprisonment and imposition of fines for unauthorised movement (Marìn-Schwartz, 2011). In addition, it implicitly addressed intellectual property in two ways: first, if genetic material was taken outside Mexico without authorisation no intellectual property claims would be recognised; and second, if there were no benefits to Mexico, intellectual property claims would not be recognised (Marìn-Schwartz, 2011). Significantly, according to the proponents, the law was not meant to impede research, but rather to spur international research collaboration through a permit system (Marìn-Schwartz, 2011). The import of the law was that the state sought to control access to organ, tissue or human components of living or dead persons. The most incisive criticism of the law on genomic sovereignty is that it was anchored on the uniqueness of the Mexican genome, which it assumed could be uniquely identified and policed internationally (Marìn-Schwartz, 2011). This discussion is fully taken up in part 2.2.

Beyond Mexico, the concept of genomic sovereignty also had policy underpinnings in India and Thailand and it generated some academic interest in South Africa. In India, genomic sovereignty was similarly conceived as a policy and scientific regime to prevent unapproved movement of genomic material and data outside India based on the need to secure state investment in genomic research and to ensure that local researchers benefit from their discoveries (Sèguin et al., 2008). The basic premise of India's genomic sovereignty was its population, which was viewed as a resource because of its large size and its uniqueness given the practice of multi-generational endogamy and the existence of proper genealogical records (Sèguin et al., 2008). The genomic sovereignty agenda was thus typified by review of Guidelines for Exchange of Human Biological Material for Biomedical Research Purposes, making it mandatory to obtain government permission to export human biological material (Hardy, 2011).

In Thailand, genomic sovereignty, though not as explicit as in Mexico and India, was similarly conceptualised as a policy agenda to protect the "Thailand genome," specifically DNA samples, from export (Sèguin et al., 2008). While legislative action was not undertaken, there was debate among researchers on the need to strengthen existing guidelines into law in order to limit export of Thai DNA samples (Sèguin et al., 2008).

2.1 Conceptual underpinnings of genomic sovereignty

From the foregoing, the concept of genomic sovereignty is rooted in post-colonial discourses of dispossession. The concept thus did not arise unexpectedly but should be regarded as an extension of the continuing North/South tension over dispossession and foreign exploitation of national resources. Illustratively, Marìn-Schwartz (2011) indicates that while the concept is traced to Mexico, it appears to have been adapted

from arguments in international fora such as UNESCO in relation to indigenous peoples' genetic heritage and tribal knowledge. While *sovereignty* has always been invoked by states, groups or peoples claiming control of natural resources, in this context it was bundled together with the catchword *genomic* to refer to the new national resource, and hence genomic sovereignty.

The argumentative logic that underpins genomic sovereignty is that in the genomic era, national population genomes can be mapped and controlled. Proponents of the concept thus conflate population with territory and as states ordinarily exercise control over national resources in their territory, then they argue that states can assert control over the population genome in their territory as a national resource (Marìn-Schwartz, 2011). This logic is premised on the idea that the human genome has commercial and symbolic value, and hence commercially as a source of economic revenue and symbolically as a source of national pride and identity (De Vries and Pepper, 2012).

Drawing from the logic of national resources and the history of colonial dispossession, genomic sovereignty is grounded in the inter-linked themes of: economic gain, national heritage and patrimony (Hardy, 2011). The economic gain argument is anchored on bio-value to be derived from the genomic revolution (Vasquez and García-Deister, 2019). For instance, Hardy (2011) captures this sentiment in relation to India: "If oil in Alaska can be shared by everybody, (the) Indian genome of India can be shared by everybody. But the fact is that (the) oil of Alaska is not shared by everybody". Strikingly, this comparison of the Indian genome to oil depicts the view of the genome as a national resource and the economic value attached to it, and hence the need to protect it from foreign exploitation.

Similarly, in Mexico, Siqueiros-Garcìa et al. (2013) point out that debate on the genomic sovereignty law drew from past foreign exploitation of petroleum, archaeological resources and biodiversity. Genomic sovereignty was thus to ensure that the Mexican genome was analysed by Mexicans and for Mexicans.

Related to the economic argument is the theme of national heritage, which is premised on the symbolic and cultural value of the human genome. The theme is anchored on the idea that national populations are biologically distinct from other populations, and hence national populations are branded as biological units (Marin-Schwartz and Mendez, 2012). National heritage is thus linked to self-determination and national building for developing countries, with the genetic make-up of the population viewed as a national resource whose exploitation is a nation-building project to deliver specific health outcomes for the national population and ensure participation in the global knowledge-based economy (Benjamin, 2009; Vasquez and García-Deister, 2019). The utility of the nation heritage theme achieved public support for genomic research for two reasons. First, to justify heavy financial investment in genomics research in light of more immediate public health concerns (Benjamin, 2009; Vasquez and García-Deister, 2018); and, second, to secure public uptake which is critical as a source of biological samples (Hardy, 2011).

Finally, the theme of patrimony, which closely mirrors national heritage, suggests protection from bio-exploitation. The central idea is that the national genome can be defined, separated from other populations, and the state can assert sovereignty over the national genome based on patrimonial doctrines (Marin-Schwartz and Mendez, 2012). In this sense, genomic sovereignty meant the duty to protect the genome of populations based on the notion of property. Illustratively, Mexico's genomic sovereignty law described the Mexican genome as a public good, a sovereign resource, which implied that the Mexican government could police and control it internationally (Marìn-Schwartz, 2011). In addition, the patrimony theme alludes to protective control, captured with the expression "genomics by Mexicans, in Mexico and for Mexicans" (Vasquez and García-Deister, 2018).

2.2 Gaps in the concept of genomic sovereignty

The concept of genomic sovereignty as discussed above provokes several questions: Is the claim of a national genome that can be uniquely identified and over which states can assert sovereignty feasible in genomic research? What exactly does genomic sovereignty relate to? Is it biological samples, data, or both? Are genomic sovereignty laws and regulations enforceable?

The Human Genome Project determined that 99.9% of human DNA is similar, with only a 0.1% variation. This unique pattern of variation across populations is at the heart of genomic research. This leads to the question of how the population of interest should be constituted, and how should the population with the unique pattern of variation be mapped? Genomic sovereignty is built on the assumption that the population of the state–based on shared national identity–constitutes a unique genetic mixture, distinct from other nations, which the state then seeks to assert control over (Benjamin, 2009).

A good starting point for analysis on this is to adopt Marin-Schwartz's and Mendez's observation in relation to Mexico (Marin-Schwartz and Mendez, 2012):

> "It is technically feasible to speak of sovereignty when we speak of the individual genome, which is unique; but to speak of sovereignty over the genome of a whole population is pretty difficult. We cannot speak of a unique Mexican make-up, when we are talking of shifting percentages of DNA fragments which are shared by humanity and various populations across the world."

The picture that emerges from Marin-Schwartz's observation is that a nation state's genetic make-up cannot be mapped, defined and separated from that of other world populations. This point finds support in De Vries and Pepper (2012) who also note that, scientifically, genomic information for groups or populations in a country is not unique or distinct. The non-existence of a unique nation state's genetic make-up raises questions that are at the core of the concept of genomic sovereignty.

Turning to the related question of whether the state can assert sovereign control over the genetic make-up of its population, Marìn-Schwartz (2011) observes that the sovereignty claims made by the state as policing genetic information, revoking intellectual property rights, and surveillance over Mexican samples were impractical. First, there is the issue of diasporic populations as law based on sovereignty articulations is territorial, while in the context of genomics, populations are fluid and are found outside state boundaries. Secondly, there is the nature of genomic research in

which genetic information flows in the international information networks as part of large-scale transnational data sets which do not conform to the territory of the nation state. Thirdly, the individual and collective property rights dimension that arise make property and patrimonial doctrines unsuited to regulate population genomics (Marìn-Schwartz, 2011).

In the same strand of arguments, Vasquez and García-Deister (2019) allude to the impracticability of sovereignty claims in the context of genetic information flow by pointing out that, despite the genomic sovereignty law, Mexican samples and DNA were analysed and reduced into cloud data which flowed internationally without being confined to the political boundaries of the state.

On the question of whether genomic sovereignty relates to biological samples, data, or both, the text of Mexico's genomic sovereignty law was silent on data. However, Marìn-Schwartz (2011) asserts that the proponents of genomic sovereignty viewed data as also protected from export without government approval. He observes that the proponents of Mexico's genomic sovereignty variously indicated that ". . . what is being protected is knowledge about genes . . .". On the contrary, Siqueiros-Garcìa et al. (2013) are categorical in their assertion that data is beyond the reach of genomic sovereignty. In addition, Rojas-Martìnez (2015) states that data is out of the scope of reach of Mexico's genomic sovereignty law. Similarly, in India, the Guidelines for the Exchange of Human Biological Material for Research Purposes, which are at the heart of the genomic sovereignty agenda, control the export of human biological material (Ministry of Health and Family Welfare [India], 1997), while in Thailand, the regulations control export of human DNA samples (Sèguin et al., 2008).

It should be accepted that biological samples are the locus of data and information. Even then, two strands of argument defeat genomic sovereignty claims over control of genomic data. First, prevailing state practice among research communities is to openly share genomic data (Contreras and Knoppers, 2018). Marìn-Schwartz (2011) alludes to this in his analysis on the impracticability of Mexico's genomic sovereignty policy, noting that Mexico is part of the international open access network to which it contributes. Secondly, and related to this, genomic research favours large data sets which circulate in the international system without being tied to nation states, and thus national data becomes less valuable. For example, Vasquez and García-Deister (2019) in their evaluation of Mexico's genomic sovereignty find that after publication of data on the Mexican genome mapping, international collaborations shifted the research to large data sets resulting in de-centring of the Mexican genome in favour of Latino genomic data.

Concerning enforcement of genomic sovereignty laws, it focuses on the question of whether the concept is workable. This discussion on the practicality of the concept is fully taken up in part 3.3. Instructively, in the three countries, the laws, regulations and guidelines that underpinned the genomic sovereignty agenda were not enforced. In Mexico, a number of shortcomings in the law informed this outcome. First, the non-existence of the unique Mexican genome, which questions what was to be protected as Mexican "uniqueness" (Marìn-Schwartz, 2011). Secondly, Siqueiros-Garcìa et al. (2013) point to lack of institutional and administrative procedures to implement the law. This view is supported by Marìn-Schwartz (2011) who, based on participant observation, points out that even if the law implicitly purported to control intellectual property in Mexican genomic research, there were no mechanisms put in place. Thirdly, there is the design of the law, in that it sought to regulate population genomics, which is fluid, while articulations of state sovereignty traditionally regulate fixed objects (Marìn-Schwartz, 2011). Similarly, in India, lack of administrative and institutional support hindered the enforcement of guidelines protecting exportation of human samples without government approval (Hardy, 2011).

3 African academics' engagement with genomic sovereignty

3.1 Overview

The literature on genomic sovereignty by African academics is dominated by the views of a South African academic, Pepper, and his collaborators. Accordingly, this section explores in a chronological fashion the development of the opinions of Pepper and his various collaborators on the topic.

Spurred by developments in Mexico, the discourse on genomic sovereignty gained international traction around 2010 (Marìn-Schwartz, 2011). At this time, Pepper published his first article on the topic. Slabbert and Pepper (2010) titled their article "A room of their own: Legal *lacunae* regarding genomic sovereignty in South Africa" and captured prevailing sentiments in the global South on the need for these countries to protect their genomic resources from exploitation by the global North. They linked genomic sovereignty to access to and benefit sharing in genomic research, particularly when genetic material originated from South Africa. They defined genomic sovereignty as "the capacity of a people, a country or a nation to own, to control both access to and use of samples, data and knowledge concerning or emanating from genomic material" which aptly captured the protection from foreign exploitation discourse. In this regard, they highlighted the need for laws to regulate individual data and the export of biological samples from South Africa.

Two years later, De Vries and Pepper (2012), in an article titled "Genomic sovereignty and the African promise: Mining the African genome for the benefit of Africa" explored whether the concept can protect genomic resources in the global South from exploitation by the global North. Pointedly, by 2012, Mexico's genomic sovereignty policy had failed. In this article, De Vries and Pepper took a decidedly more critical view of genomic sovereignty. The authors acknowledged the appeal of the concept of genomic sovereignty in the African context, but pointed out its conceptual limitations. They identified the limitations as: lack of clarity on whom the final authority on access to and use of genomic material rests and the role of the individual donor; inability of states to represent the interests of the populations within their borders equally, including the contestations by indigenous peoples on representation of their interests; existence of ethnic groups across geographical state boundaries; the assumption that the population of a state is a unique biological unit; and the transnational nature of genomic data. Based on these limitations, the authors argued that genomic sovereignty is inadequate on its own to resolve the problems of inequality and unfair distribution of benefits in African genomic research.

Furthermore, in 2017, while discussing the exporting of DNA, Pepper alluded to the need to strike a balance between prevention of exploitation and promotion of innovation. In the discussion on

ownership of DNA, he mentioned the concept of genomic sovereignty as referring to the "need to regulate ownership of human genetic resources" (Pepper, 2017).

Beyond academic research, policy proposals that have alluded to genomic sovereignty in South Africa have similarly been associated with Pepper's involvement in the development of such proposals. For example, in 2011, the National Biotechnology Advisory Committee, of which Pepper was a member, published a statement on genomic sovereignty calling for public debate in South Africa on regulation and monitoring of human genomic material (National Biotechnology Advisory Council [South Africa], 2011). Most recently, in 2018, the Academy of Science of South Africa (ASSAf) published a report by a consensus study group, led by Pepper, on human genetics and genomics in South Africa. This report explicitly advanced the notion of an individual's DNA and genomic data as being "natural resources," to be managed by the state similarly to water or mineral resources (Academy of Science of South Africa (ASSAf), 2018).

3.2 Conceptual underpinnings of genomic sovereignty by African academics

What then underpins the concept of genomic sovereignty as espoused by the African academics discussed above? A textual analysis of Pepper and his research collaborators' academic work yields repetition of the concepts: inequality; exploitation in export of samples; ownership of genomic resources; and unfair distribution of benefits. Based on this, it is therefore plausible to argue that the conceptual underpinnings of genomic sovereignty by African academics are rooted in the history of colonial dispossession, as in Mexico. De Vries and Pepper (2012) allude to this by linking genomic sovereignty to concerns of "revival of colonialist relations between Africa and the western world" in genomic research. In the same strand of argument, these authors state that genomic sovereignty offers a "conceptual framework" for genomic research by regulating ownership of genomic material and samples (De Vries and Pepper, 2012).

Pepper and fellow researchers couch the concerns on dispossession as inequality between local and international researchers, arguing that local researchers do not benefit from international research collaborations. According to these authors, exploitation in the export of samples is comparable to exploitation of natural resources such as oil and minerals, and can be addressed if the concept of genomic sovereignty is enshrined into law.

De Vries and Pepper (2012), while discussing the conceptual limitations of genomic sovereignty, also point out that the concept of genomic sovereignty is inadequate on its own in achieving equity and justice in genomic research. Rather, they highlight the need to develop governance tools to ensure fair distribution of benefits among researchers and populations. Pepper's work on the export of DNA in South Africa revisits genomic sovereignty in the context of regulation of ownership of human genetic samples to guard against exploitation in research, and points out that it remains debated (Pepper, 2017).

3.3 Critical assessment of genomic sovereignty

A significant weakness in the literature discussed above is a failure to consider human rights. Most prominently, as pointed out by Thaldar et al. (2019), positing an individual's genomic material and data as "natural resources"—similar to minerals and water—can be deemed offensive to individual dignity. In addition, the literature discussed above fails to recognise and deal with the right to freedom of scientific research, which is protected as a fundamental human right in some African countries, such as South Africa, Kenya, Morocco and Zimbabwe (Thaldar and Steytler, 2021). This failure to consider human rights raises the question of whether a legislative or policy move towards genomic sovereignty would withstand constitutional scrutiny.

Moreover, in the case of genomic data, one also needs to consider informational privacy rights. Various African countries have enacted data protection legislation. These include some of Africa's most populous countries, such as Kenya, Nigeria, South Africa and Tanzania. Genomic data will typically fall within the scope of these statutes. Also, it is unlikely that genomic data can be de-identified or anonymised (depending on the terminology used in the specific jurisdiction) in order to escape the applicability of these statutes. The informational privacy rights protected in these statutes are therefore likely to apply to genomic data. How does this effect genomic sovereignty? Informational privacy rights belong to *individuals*, and aim to protect *individual* privacy interests. This stands in contrast with genomic sovereignty, which aims to promote *collective* ethnic group interests or state interests. It is not clear from the literature discussed above how proponents of genomic sovereignty propose to solve this philosophical dilemma.

Furthermore, to the extent that genomic sovereignty is understood as entailing *ownership* of genomic material and data by the state, such version of genomic sovereignty would amount to the nationalisation—and *expropriation*—of property that is currently *privately* owned. For example, in South Africa, genomic material is currently owned by the research institution to which such material is donated by a research participant (Thaldar and Shozi, 2022), and genomic data, once sequenced and saved as a digital object, can also be privately owned—likely by the research institution that performed the sequencing (Thaldar et al., 2022). Therefore, if *private* ownership of genomic material and data is replaced with *state* ownership, it means that such genomic material and data are expropriated, which in turn triggers legal protections of property rights. At the very least, the state would have to offer financial compensation to the private owners. The property law dimension of genomic material and data is a legal fact that cannot be ignored or wished away.

4 Finland's state approach to human genomic research

4.1 Finland's human genomic research infrastructure

This section discusses Finland's human genomic research framework as follows: biobanking infrastructure; framework for availability and utilisation of genomic data, including the National Genome Strategy; proposed genome centre and Genome Act; and the Finngen project.

Finland's framework on human genomic research traces back to 2006 when the Ministry of Social Affairs and Health established a

working group to develop a law to regulate biobank operations in Finland (Salokannel et al., 2019). The Finnish Biobank Act entered into force in September 2013 and sets its objectives as supporting research that uses human biological samples, promoting openness in the use of the samples, and securing protection of privacy and self-determination when processing the samples. It regulates all types of biobanks and biological samples and information associated with the samples. The scope of the Act covers: establishment and operation of biobanks; collection of biobank samples and information attached to the samples; storage and processing of samples; rights of registered individuals to protect their privacy; and registers for biobanking (Ministry of Social Affairs and Health [Finland], 2013). The biobanks own the samples and are regarded as common resources, to which researchers have access. The Act allows broad consent for future research and secondary use of samples and the linking of personal data with the biobank information (Ministry of Social Affairs and Health [Finland], 2013). As of 2022, there are 11 registered biobanks in Finland, of which ten are public and one is private. All biobanks must obtain a licence from the Finnish Medicines Agency, Fimea (Ministry of Social Affairs and Health [Finland], 2013).

In 2015, Finland launched the National Genome Strategy which sets measures for incorporation of genomic data in the Finnish healthcare system by 2020 (Ministry of Social Affairs and Health [Finland], 2015). The Strategy is primarily focused on data collection and utilisation and the establishment of a single entity for management of genomic data. A key feature of the Strategy is the establishment of a new public authority, the genome centre, to promote responsible and equal use of genomic data. The genome centre will be mandated to: set up a national reference database of genomes; operate as a service point for research agreements, contracts and commercialisation; promote ethical practices in the use of genomic data by planning and implementing consents; facilitate networking and international collaboration; and initiate and stimulate public debate on utilisation of genomic data. The genome centre is envisioned as a permanent entity established through legislation, the Genome Act. The main objective of the Genome Act is to facilitate responsible, equal and secure processing of genomic data for the benefit of citizens (Ministry of Social Affairs and Health [Finland], 2015). As of 2022, the Genome Act was still being drafted.

In 2017, Finland launched the Finngen study, a national public–private research project to collect and analyse genome and health data from 500,000 participants of the Finnish biobanks by 2023. The study is a partnership between universities, hospitals, biobanks, the National Institute for Health and Welfare, international pharmaceutical companies and the Finns. As of December 2022, the total number of participants was 342,499 (FINNGEN, 2017). The aim of Finngen is to build a data resource that combines nationwide biobank data, national healthcare data, and genome data.

Finland's strengths which have enabled the establishment of the above discussed research infrastructure are: high standard and universal healthcare; uniform treatment practices; national health registers; history of genetic research; and a population that is willing to participate in genomic research (Ministry of Social Affairs and Health [Finland], 2015).

4.2 Philosophical underpinnings of Finland's human genomic research framework

Finland's human genomic research framework is founded on individual (donor) sovereignty. Notably, Finland operates a welfare public healthcare system for all residents. Writing on data donation and exercise of sovereignty, Hummel et al. (2019) argue that individual data sovereignty has both negative and positive dimensions. The negative dimension of individual sovereignty connotes the power to exclude others from personal data, while the positive dimension includes the power to decide where your data goes and how it is to be used. They compare individual sovereignty to classical state sovereignty and posit that state sovereignty has both external and internal dimensions, in which the external dimension connotes no external inference, while the internal sovereignty means the state has the power to govern within its territory as it wills. Similarly, individual data sovereignty would mean the ability to exclude others from personal data, and the ability to operate within the informational self-determination sphere to pursue certain aims and goals with one's data. In addition, in the broad context of sovereignty, they view power as the enabler of the exercise of sovereignty. In the context of individual data sovereignty, power means control over one's individual data, that is, where it goes, who can access it, and what it is used for. Viewed from this perspective, individuals can exercise personal data sovereignty in its positive and negative dimensions (Hummel et al., 2019).

Further, Hummel et al. (2019) argue that individual data sovereignty can be facilitated in three ways: through consent, representation, and organisational level constraints. On consent, the authors state that informed consent would entail a balance between research participation and respect for the self-determination of the individual donor. In this respect they highlight mechanisms that allow for opting out of biobanks and withdrawing from research projects based on evolving preferences. In relation to representation, they point to representation of an individual donor's will in the research governance processes to further their interests. Finally, on organisational-level constraints, they argue for supervisory oversight of institutions involved in data collection and processing through impartial licensing schemes and state legislation to ensure protection of the rights of individual donors.

Returning to Finland's human genomic research framework, notably, the legal and policy documents expressly describe their objectives or goals as enabling individual control over their own genomic data. Illustratively, the Biobank Act sets out its objectives as promoting openness in the use of human biological samples and to ensure protection of privacy and self-determination in processing the samples (Ministry of Social Affairs and Health [Finland], 2013). Similarly, the Genome Act will establish the genome centre as a national reference database, and has, as one of its principles, the ability of individuals to control use of their genomic data (Tervo, 2021). The National Genome Strategy identifies the need for legislation to guarantee individual rights to control, manage and monitor own genomic data (Ministry of Social Affairs and Health [Finland], 2015).

The research framework also contains features of individual sovereignty in relation to control over use and management of their data. For instance, the Biobank Act provides for access to

information for registered individuals. This allows them, upon request, to receive information on their samples such as whether their samples are stored in the biobank and the criteria, who can receive samples taken from them, who can access the samples, and information on transferring the samples from the biobank (Ministry of Social Affairs and Health [Finland], 2013).

In line with the three mechanisms to facilitate exercise of individual sovereignty discussed by Hummel et al. (2019) on consent, the Finnish research framework governing laws allow individuals to exercise control through issuing of consent to research participation. The Biobank Act's provisions allow individuals to voluntarily consent to the use of their samples and data, to impose restrictions while issuing consent, and to cancel their consent at any time without any penalties (Ministry of Social Affairs and Health [Finland], 2013). In addition, the Act allows for change of consent or prohibition of the use of samples at any stage of the research process (Ministry of Social Affairs and Health [Finland], 2013). In this regard, it is arguable that individual data sovereignty is exercised, as the individual manages and controls how, for what, and by whom their samples are used.

In relation to organisational-level controls, the Finnish biobanks are regulated and monitored at the national level. First, prior to establishment, biobanks must obtain a positive statement from the National Committee on Medical Research Ethics confirming that the activities of the biobank comply with the protection of privacy and self-determination requirements (Ministry of Social Affairs and Health [Finland], 2013). Secondly, biobanks are supervised at the national level by Fimea to ensure that they maintain transparency in their biobanking activities (Ministry of Social Affairs and Health [Finland], 2023). Thirdly, biobanks are required to appoint a custodian whose duties are cast as obligations owed to the individual sample/data donor (Ministry of Social Affairs and Health [Finland], 2013). These organisational-level constraints enable individual data sovereignty as individuals retain control and management of how their samples and data are used.

4.3 A tale of two approaches to regulation of human genomic research: Mexico versus Finland

This section compares the above state approaches in the regulation of human genomic research. It is likely that there are more approaches and the choice of these two is random. Yet, both Mexico and Finland represent archetypes for two approaches to human genomic research that states have taken in the past decade in terms of the differences that characterise them and the seemingly unexpected similarities.

The general stance of Mexico's approach may be described as follows. Human genomic resources are akin to other natural resources, and hence the need for the state to exercise a protective barrier to prevent foreign exploitation. The state's role is thus viewed as to control and regulate access to and use of genomic resources. This position is actualised through establishment of the research institution INMEGEN, mapping of the national genome, and incorporation of the concept of genomic sovereignty in law.

Conversely, Finland's approach is that genomic data can be used to improve the health outcomes of the Finnish people. The state's role is thus to put in place the requisite framework for genomic data collection and utilisation. This position is actualised through the establishment of a legal framework on biobanks, mapping of the national genome, and establishment of the research framework through the National Genome Strategy which proposes a national genome centre and the genome law.

From the two approaches, this article identifies the following themes as the basis of comparison: philosophical underpinnings of the approaches; research infrastructure; and ownership of genomic data. As already demonstrated, Mexico and Finland have stark differences in their philosophical approaches. The Mexican approach is premised on genomic sovereignty, understood as state control of access to and use of genomic resources. This philosophy is rooted in the belief that genomic resources are national resources over which the state can exercise a protective barrier from foreign exploitation. In practice, this philosophy was embedded in the genomic sovereignty law which restricted export of human biological samples without state approval and the INMEGEN whose roles included surveillance of the samples. In the end, Mexico's approach, which was underpinned by genomic sovereignty, proved impractical as it hindered international collaborations, hence violating the rights of both researchers and citizens generally. In addition, there was the inability of the state to exercise control due to the nature of population genomics, which defies territorial-based articulations of sovereignty. Finland's research framework is premised on individual sovereignty, in which the individual donor of the samples and data exercises sovereignty by controlling and managing how their genomic data is used. The philosophy is rooted in human rights, in which individuals exercise informational self-determination. It is embedded in the legal framework on Biobanks and in the National Genome Strategy and its proposed national genome centre and genome, by vesting in individuals the power to control and manage use of their data.

On the research infrastructure, there are striking similarities. Both Mexico and Finland have established an institutional framework for genomic research, the INMEGEN in Mexico and the proposed national genome centre in Finland. In addition, both have laws to govern human genomic research: in Mexico, the genomic sovereignty law, while in Finland there is the Biobank Act and the proposed genome law. However, the point of departure is on the roles of the institutions and the objectives of the laws, which reflect the philosophy that underpins the overall state approach. In Mexico, the INMEGEN had the role of centralising and controlling genomic research. Similarly, the objective of the genomic sovereignty law was to restrict movement of samples without state approval. On the contrary, in Finland, the proposed national genome centre will facilitate international collaboration by setting up the national reference database of genomes with the necessary links to international databases and will provide centralised services for research projects and agreements. It will also plan and implement management of consents based on the individual right to decide. The genome law will enable individuals to control, manage and monitor the use of their genomic data.

Finally, how do the two state approaches deal with the issue of ownership, which in Finland is not deemed problematic, but which at the international level remains unsettled? In Finland, the general proposition is that genomic data are owned by the research institutions and biobanks. However, given that the research

framework is underpinned by individual sovereignty, as discussed previously, it is reasonable to argue that the ownership is more nuanced and can perhaps be described as *custodianship*, given that significant control vests in individual donors. The Mexican proposition is the opposite. The issue of ownership of genomic resources undergirds the research framework. Genomic resources are viewed as national resources which can be subjected to exclusive state control. This proposition on ownership explains the violation of the rights of researchers and the complete disregard of the rights of individuals in Mexico's research framework, which, in many states, would not pass constitutional scrutiny.

The foregoing discussion of the Mexican and Finnish approaches to human genomic research demonstrates that state approaches premised on genomic sovereignty are unworkable. First, this approach is not aligned to state obligations under international human rights law, specifically, the right to academic freedom for researchers, the right to privacy and self-determination for individuals, and the right to science. Secondly, the nature of genomic research defies state control premised on the idea of sovereignty, since sovereignty is territorial while polulations are fluid. In addition, the nature of genomic research is transnational, and hence obsession with national-level data is misplaced as national-level data on its own has limited utility value. Conclusively then, state approaches based on genomic sovereignty should be abandoned.

In the next section, the article explores obligations of the state under the right to science and in the specific context of human genomic research, in an effort to map out proposals on where state approaches should focus.

5 The right of everyone to enjoy the benefits of scientific progress and its applications in international human rights law

5.1 State obligations flowing from the right to science

The right of everyone to enjoy the benefits of scientific progress and its applications (the right to science) finds textual expression in the Universal Declaration of Human Rights (UN General Assembly, 1948) (UDHR) and in the International Covenant for Economic, Social and Cultural Rights (UN General Assembly, 1966) (ICESCR). Although not framed in identical wording, there are arguments in favour of construing the right as encompassing the right of everyone to access and contribute to knowledge and information and the right of everyone to benefit from scientific applications (Boggio and Romano, 2018; Yotova and Knoppers, 2020; Mancisidor, 2021). In addition, the right is formulated as part of cultural rights, and state practice has similarly evolved to treat the right as part of cultural rights (UN Human Rights Council, 2009). Arguments made in support of the right as a cultural right posit that both science and culture involve production of knowledge, innovation and creativity which support the full development of the person (Shaheed and Mazibrada, 2021).

Unlike other socio-economic rights, the right to science has not received much scholarly attention or in state implementation and has thus been said to be characterised by stunted development compared to other rights. Yotova and Knoppers (2020) citing Schabas refer to the right as the "sleeping beauty of human rights," and Donders (2011), while discussing the reawakening of the right, refers to the right as recently "having its dust blown off". Mancisidor (2021) notes that the textual positioning of the right at the end of both the UDHR and the ICESCR and its characterisation as a cultural right have contributed to its neglect. Even then, advances in science and technological innovation in the past decade have foregrounded the right. For instance, in the specific context of human genomic research there is scholarly work invoking the right in relation to genomic human research. Yotova and Knoppers (2020) have reviewed state practice on the right and argued that the right to benefit from science and its applications supports genomic data sharing. Elsewhere, Knoppers et al. (2014) anchor their proposal on an international code of conduct for sharing genomics and clinical data on the right to benefit from science and its applications. Below is a brief discussion on the right to science that maps out the state obligations as a prelude to the next section on state obligations in human genomic research.

A plain reading of the right to science as formulated in the ICESCR reveals positive obligations requiring states to "recognize the right of everyone to enjoy the benefits of scientific progress and its applications." The UN Committee on Economic, Social and Cultural Rights (2020) elaborated on the elements of the right and the ensuing state obligations. It lists the elements of the right as: availability, accessibility, quality, and acceptability. *Availability* connotes a requirement of scientific progress and the protection and dissemination of scientific knowledge. Consequently, states have an obligation for conservation, development and diffusion of science by putting in place research infrastructure, funding research, promoting open science, and making accessible the findings and data of publicly funded research (UN Committee on Economic, Social and Cultural Rights, 2020). *Development* has been interpreted to mean state support for science while *diffusion* refers to equitable distribution of the benefits and applications of science, and *conservation* requires sustainable science that caters for present and future generations (Frick and Dang, 2021). *Accessibility* addresses the right of every person to access scientific progress and its applications without discrimination. States are thus to guarantee equal access to the applications of science to all, information on risks and benefits of science, and opportunity for all to participate in scientific progress (UN Committee on Economic, Social and Cultural Rights, 2020). The American Association for the Advancement of Science views accessibility as "a continuum of access" with the public on one end of the spectrum and scientists on the other (Frick and Dang, 2021). Quality relates to both creation of scientific knowledge and access to the benefits and applications of science. To this end, states are required to ensure ethical and responsible development of science and that only certified science is available to the general public (UN Committee on Economic, Social and Cultural Rights, 2020). An example of state failure to adhere to the element of quality in the application of science was the alleged United Arab Emirates use of faulty algorithms in diagnosing tuberculosis in its immigration procedures, which resulted in denial of work permits for migrants (Frick and Dang, 2021). *Acceptability* means respect for cultural diversity and pluralism in that the right to participate in science and enjoy benefits of science and its

applications should be implemented in a manner that accords with specific cultural and social contexts. *Acceptability* also connotes ethical standards including respect for dignity, privacy and autonomy of individuals (UN Committee on Economic, Social and Cultural Rights, 2020). Respect for cultural diversity can be achieved by having community advisory boards in research projects and multidisciplinary and plural ethical review boards.

While the right to science is to be achieved progressively and subject to availability of resources, akin to other rights in the ICESCR, the right imposes obligations of a general nature on states which are of immediate implementation. To this end, the right to science requires states not to take retrogressive measures that impede enjoyment of the right. Such measures include imposition of policies that impede conservation, development and diffusion of science, the imposition of barriers to citizen participation, and adoption of laws and policies that prevent international collaborations (UN Committee on Economic, Social and Cultural Rights, 2020). In addition, states have an immediate obligation to eliminate discrimination in terms of participation in scientific progress and enjoyment of scientific benefits and its applications. Consequently, states should eradicate discrimination in the formulation and implementation of policies on the right to participate and benefit from science and its applications (UN Committee on Economic, Social and Cultural Rights, 2020).

The specific state obligations under the triptych typology are obligation to respect, to prevent or ensure, and to fulfil. Under the obligation to respect, states and their agencies should desist from interfering with the right to participate in and to enjoy the benefits of science and its applications. States therefore should not misinform the public on science and scientific research which could have the effect of eroding public trust in science, create obstacles for international collaboration among scientists, and arbitrarily limit internet access which could impede access to and the dissemination of scientific knowledge. The obligation to protect requires states to prevent violation of the right to science by non-state actors, including individuals and multinational corporations. The measures states should take to protect include: preventing non-state actors from applying discriminatory criteria in scientific research; ensuring and guaranteeing ethical standards for persons in scientific research; and protecting individuals in their familiar, social and cultural contexts when their right to science is violated. The UN Committee on Economic, Social and Cultural Rights (2020) recognises the dominance of private enterprises in the right to science and requires states to establish a legal framework that imposes a duty of human rights' due diligence on multinational corporations. In addition, states have extraterritorial obligations to ensure that multinational corporations within their control do not violate the right to science when acting abroad. On the obligation to fulfil, states should put in place the requisite infrastructure—legal, institutional, financial and administrative—for the right to science. This includes: facilitating participation in international cooperation programmes, facilitating access to the internet, funding research and making scientific knowledge broadly available (UN Committee on Economic, Social and Cultural Rights, 2020).

Under Article 15(4) which refers to the gains of international cooperation in the right to science, states have an obligation to promote and facilitate scientific researchers to participate in the "international scientific and technical community" and to freely share data (UN Committee on Economic, Social and Cultural Rights, 2020). In addition, in recognition of the differences among states in science and technology, the ICESCR imposes special obligations on wealthier states to assist less wealthy states (UN Committee on Economic, Social and Cultural Rights, 1990).

5.2 State obligations in respect of human genomic research

Consistent with the idea of the interdependence and interconnectedness of human rights, this section examines state obligations in human genomic research from the prism of the right to science while reading in the right to health. The section draws from the approaches of Mexico and Finland in human genomic research for contextualisation.

The chronicles of the Mexican and Finnish approaches to human genomic research offer insights on its nature. It follows that: it is driven by large datasets that are tied to other data; it is characterised by transnational and public–private collaboration specifically that nationally bound genomic data has limited utility; it requires research infrastructure; and individual data and the inherent ethical issues are crucial. What then are the state obligations for the realisation of the right to science in this context?

At the outset, although ICESCR rights lend themselves to progressive implementation, as discussed above, states have immediate obligations. First, states are prohibited from taking retrogressive measures in relation to human genomic research. Retrogressive measures include removal of policies that are necessary to support scientific research and legal and policy changes that hinder international collaborations in science. Revisiting genomic sovereignty in Mexico, the result of this political and scientific policy agenda was that the law on genomic sovereignty hindered international collaboration as scientists could not share biological samples. While retrogressive measures may be permitted under the ICESCR, for instance in the case of natural disasters and severe recessions, they must be necessary and proportionate. As discussed, the philosophical underpinnings of genomic sovereignty are a postcolonial dispossession discourse which would not support a reading of it as a permissible measure. Mann et al. (2021) provide some direction on the evaluation of retrogressive measures under the ICESCR rights, such as a country's level of development; economic recession; whether the country is involved in international conflict; and whether the country has sought assistance. Therefore, the concept of genomic sovereignty negates the right to science as it constitutes an unjustifiable retrogressive measure. The core state obligations require the state to take measures that enhance development, diffusion and the conservation of science. There are valuable lessons from Finland's approach in which the state has facilitated research infrastructure, such as the national genome centre. This will create a national reference database for genomes with international links, thus allowing international collaborations in line with the obligation to support science.

The second set of state obligations that is immediate and not subject to progressive measures is non-discrimination. The formulation of the right to science uses the term *everyone*. Significantly, invocation of the term *everyone* in relation to participation in science means that not only researchers but also

the public must participate in science and scientific progress without discrimination. States are thus directed to ensure that marginalised segments of the population can participate in genomic research and also enjoy the benefits and applications of scientific progress.

Flowing from non-discrimination in the right to science is the requirement of participation. On the framework for participation, Mann et al. (2021) point out two areas in which everyone can participate: in active dialogue between scientists and the public to create public trust and promote citizen science; and in data donation. The question then arises on state obligations in this regard. The element of accessibility in the right to science obligates states to put in place measures for everyone to participate in science. In addition, under the obligation to fulfil, the state should put in place legal, institutional and budgetary infrastructure that allows for participation of both citizens and scientists. Furthermore, in relation to data donation, states have an obligation to make data available through adoption of policies that encourage citizens to participate in research. Drawing from Finland's approach, as discussed earlier, the philosophical underpinning of Finland's genomic research is individual data sovereignty, which encourages data donation as citizens control and manage their data in genomic research. This proposition finds support in Finland's legislation and policies, for instance the Biobank Act, the proposed Genome Act and the National Genome Strategy, which ensure that individuals control and manage their data.

Moreover, on transnational data sharing, the element of availability imposes a duty on states to support science. In relation to human genomic research, this implies sharing of data. Yotova and Knoppers (2020) also argue that the obligation of states to diffuse scientific knowledge would require states to share genomic data from a global public goods approach. As pointed out earlier, the Mexican concept of genomic sovereignty restricted transnational sharing of data of the Mexican genome and ultimately inhibited development and diffusion of science given that national genomic datasets have little utility value for genomic research.

On the centrality of individual and the inherent ethical issues, the key state obligations can be derived from the element of acceptability and the state's specific obligation to protect. The element of acceptability requires states to ensure that scientific research incorporates ethical standards that respect privacy, autonomy and dignity of the individual, as well as minimisation of harm and maximisation of benefits. Relatedly, the obligation to protect requires states to prevent violation of rights by non-state actors through adoption of legislation and judicial remedies in instances of violation. To that extent, the approach taken by states in human genomic research should ensure that individual rights are respected and upheld. As demonstrated, the rights in issue are individual privacy and informational self-determination and human dignity. At the outset, it is also apparent that state approaches to human genomic research anchored on genomic sovereignty violate individual rights to human dignity by regarding individual personal data as a national resource and also violate privacy and informational self-determination by disregarding the rights of individuals in relation to their individual data. Taking lessons from Finland's approach, the nature of human genomic research demands that human rights be at the core of the research infrastructure due to the centrality of the individual from the perspective of genomic data and application of the outcomes of scientific progress.

Finally, on the research infrastructure, the obligation to fulfil imposes a positive duty on states to actively facilitate the advancement of science. As noted, human genomic research is reliant on research infrastructure such as institutional frameworks for collecting and processing samples and data and the supporting legal and policy framework, including international linkages and collaborations. States therefore have an obligation to allocate funding for research infrastructure. A key lesson from the Mexican and Finnish approaches is the philosophical underpinning that anchors the research infrastructure. As demonstrated, philosophical underpinnings that view genomic resources as national resources are inclined to establish research infrastructure that hinders the right to science, as in the case of Mexico and genomic sovereignty.

6 Conclusion

This article sought to enquire whether genomic sovereignty is the appropriate governance framework for access to and use of genomic resources. Debate on the governance framework for the human genome remains unsettled 25 years after the adoption of the Universal Declaration on the Human Genome and Human Rights. While the Declaration appears to favour the common heritage of mankind governance framework, at the time of adoption a number of states indicated the need for further deliberations. Twenty-five years on, developments in this regard have remained stunted. However, in this time, state practice in human genomic research has evolved to embody divergent approaches, in some instances in response to the proposed international governance framework and in others motivated by the desire to participate in the bio-economy heralded by the Human Genome Project. This article analysed the divergent approaches by two states, Mexico and Finland. While not conclusively settling the issue of the most appropriate governance framework, the article brings out insights on which approach best supports human genomic research and which approach best enables a state to participate in the genomic knowledge economy. A key finding is that given the centrality of the individual in human genomic research, state-centred approaches—anchored on the notion of state appropriation of genomic resources—such as that embodied by genomic sovereignty, cannot withstand human rights scrutiny and should be abandoned.

Turning to the issue of human rights, the article argues that even in the absence of a conclusive position at the political level on the governance framework, states have legally binding obligations in human genomic research under international human rights law. The ICESCR guarantees everyone the right to science and imposes certain binding obligations on states. Significantly, the ICESCR enjoys near universal ratification and no state has placed a reservation on the right to science (UN Office of the High Commissioner for Human Rights, 2023). To this extent, state parties have binding obligations regardless of the absence of a conclusive international governance framework on human

genomic research. The article once again demonstrates that state-centred approaches to human genomic research violate the rights of individuals and those of researchers, hence running foul of binding ICESCR obligations. In addition, the article maps the state obligations under the right to science—pointing to where state approaches to human genomic research should focus.

Returning to the broader and vexing question of the appropriate governance framework, individual personality rights, state sovereignty and the common heritage concepts have always seemed at odds in governance of human genomic research and it often appears as if one should trump the other. This article, for a start, points to the possibility of marrying these competing regulatory approaches. Undoubtedly, the state remains the primary site of accountability for protection of human rights. In that regard states have a role to play in governance of human genomic research by implementing the right to science and ensuring the protection of the other rights at play. For the individual and their personality rights, the Finnish approach demonstrates the possibility of individuals taking control in the use of their genomic data and samples.

The take-home lesson for African states is that genomic sovereignty has no utility value in human genomic research and thus should be abandoned as the philosophical basis for governing human genomic research, regardless of its superficial appeal in light of Africa's history of colonisation and dispossession. Human rights offer a more promising approach. First, African states have binding obligations under the right to science—notwithstanding availability of resources. Also, taking into account the indivisibility of rights, states have obligations under the right to health to facilitate enjoyment of the highest attainable standard of health. As we have argued, African states should implement their human rights obligations in the context of human genomic research by putting in place the necessary research infrastructure, including ethical and legal frameworks to protect individual rights, and by facilitating international collaborations to foster research.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, and further inquiries can be directed to the corresponding author.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and have approved it for publication.

Funding

Work on this article was supported by the U.S. National Institute of Mental Health and the U.S. National Institutes of Health (Award Number U01MH127690) under the Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) program.

Acknowledgments

We are grateful to Bonginkosi Shozi, Freddy Mnyongani, Peter Munyi, Claude Kamau, and Johanna Rahnasto for their insightful comments while developing the article. Any errors are our own.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Author disclaimer

The content of this article is solely our responsibility and does not necessarily represent the official views of the U.S. National Institute of Mental Health or the U.S. National Institutes of Health.

References

Academy of Science of South Africa (ASSAf) (2018). *Human genetics and genomics: Ethical, legal and social implications in South Africa*. Pretoria: Academy of Science of South Africa Available at: http://dx.doi.org/10.17159/assaf.2018/0033

Benjamin, R. (2009). *A Lab of Their Own*: Genomic sovereignty as postcolonial science policy. *Policy Soc.* 28 (4), 341–355. doi:10.1016/j.polsoc.2009.09.007

Boggio, A., and Romano, C. (2018). "Freedom of research and the right to science: From theory to advocacy," in *The freedom of scientific research*. Editors S. Giordano, J. Harris, and L. Piccirillo (Manchester: Manchester University Press), 162–175.

Contreras, J. L., and Knoppers, B. M. (2018). The genomic commons. *Annu. Rev. Genomics Hum. Genet.* 19 (1), 429–453. doi:10.1146/annurev-genom-083117-021552

De Vries, J., and Pepper, M. S. (2012). Genomic sovereignty and the African promise: Mining the African genome for the benefit of Africa. *J. Med. Ethics* 38, 474–478. doi:10.1136/medethics-2011-100448

Donders, Y. (2011). The right to enjoy the benefits of scientific progress: In search of state obligations in relation to health. *Med. Health Care Philos.* 14 (4), 371–381. doi:10.1007/s11019-011-9327-y

FINNGEN (2017). Finngen. Available at: https://www.finngen.fi/en.

Frick, M., and Dang, G. (2021). "The right to science: A practical tool for advancing global health equity and promoting the human rights of people with tuberculosis," in *The right to science: Then and now*. Editors H. Porsdam and S. P. Mann (Cambridge: Cambridge University Press), 246–267.

Hardy, B.-J. (2011). *Bridging the genomics gap: The role of large-scale genotyping projects in the developing world and the importance of genomic sovereignty*. Toronto: University of Toronto.

Hummel, P., Braun, M., and Dabrock, P. (2019). "Data donation as exercises of sovereignty," in *Ethics of medical data donation*. Editors L. Floridi and J. Krutzinna (Cham: Springer), 23–54.

Knoppers, B. M., Harris, J., Budin-Ljøsne, I., and Dove, E. S. (2014). A human rights approach to an international code of conduct for genomic and clinical data sharing. *Hum. Genet.* 133 (7), 895–903. doi:10.1007/s00439-014-1432-6

Mancisidor, M. (2021). "The dawning of a right," in *Right to science: Then and now*. Editors H. Porsdam and S. P. Mann (Cambridge: Cambridge University Press), 17–32.

Mann, S. P., Donders, Y., and Porsdam, H. (2021). "The right to science in practice: A proposed test in four stages," in *The right to science: Then and now*. Editors H. Porsdam and S. P. Mann (Cambridge: Cambridge University Press), 231–245.

Marin-Schwartz, E., and Mendez, A. A. (2012). The law of genomic sovereignty and the protection of "Mexican genetic patrimony". *Med. Law* 31, 283–294.

Marìn-Schwartz, E. (2011). *Genomic sovereignty and the "Mexican genome"*. Exeter: University of Exeter.

Ministry of Health and Family Welfare [India] (1997). *Guidelines for Exchange of Human Biological Material for Biomedical Research Purposes*. Available at: https://main.icmr.nic.in/content/download-guidelines-transfer-biological-material.

Ministry of Social Affairs and Health [Finland] (2013). *Biobanking Act*. Available at: https://www.finlex.fi/en/laki/kaannokset/2012/en20120688.pdf.

Ministry of Social Affairs and Health [Finland] (2015). National Genome Strategy. Available at: https://frontlinegenomics.com/world-of-genomics-finland/.

Ministry of Social Affairs and Health [Finland] (2023). Supervision of biobanks. Available at: https://thl.fi/en/web/thl-biobank/biobank-info/supervision-of-biobanks.

National Biotechnology Advisory Council [South Africa] (2011). *NBAC position statement on genomic sovereignty in South Africa*. Available at: https://www.naci.org.za/index.php/nbac-position-statement-on-genomic-sovereignty-in-south-africa/.

Pepper, M. S. (2017). Exporting DNA: Striking a balance between preventing exploitation and promoting innovation. *S. Afr. Med. J.* 107 (2), 106–107. doi:10.7196/SAMJ.2017.v107i2.12122

Rojas-Martìnez, A. (2015). Confidentiality and data sharing: Vulnerabilities of the Mexican Genomics Sovereignty Act. *J. Community Genet.* 6, 313–319. doi:10.1007/s12687-015-0233-5

Salokannel, M., Tarkkala, H., and Snell, K. (2019). Legacy samples in Finnish biobanks: Social and legal issues related to the transfer of old sample collections into biobanks. *Hum. Genet.* 138 (11), 1287–1299. doi:10.1007/s00439-019-02070-0

Sèguin, B., Hardy, B.-J., Singer, P. A., and Daar, A. S. (2008). Genomic medicine and developing countries: Creating a room of their own. *Nat. Rev. Genet.* 9, 487–493. doi:10.1038/nrg2379

Shaheed, F., and Mazibrada, A. (2021). "On the right to science as a cultural human right," in *The right to science: Then and now*. Editors H. Porsdam and S. P. Mann (Cambridge: Cambridge University Press), 107–123.

Siqueiros-Garcìa, J. M., Oliva-Sanchez, P. M., and Saruwatari-Zavala, G. (2013). Genomic sovereignty or the enemy within. *Acta Bioethica* 19, 269–273. doi:10.4067/s1726-569x2013000200011

Slabbert, M. N., and Pepper, M. S. (2010). A room of our own?: Legal *lacunae* relating to genomic sovereignty in South Africa. *THRHR* 73, 435.

Tervo, S. (2021). *The Genome Act: Short overview of challenges and relevant solutions*. Helsinki, Finland: Ministry of Social Affairs and Health.

Thaldar, D., Kinderlerer, J., and Soni, S. (2019). An optimistic vision for biosciences in South Africa: A response to the ASSAf report on human genetics and genomics. *S. Afr. J. Sci.* 115 (7), 1. doi:10.17159/sajs.2019/6146

Thaldar, D., and Shozi, B. (2022). The legal status of human biological material used for research. *S. Afr. J. Law* 138 (4), 881–907. doi:10.47348/salj/v138/i4a9

Thaldar, D., and Steytler, M. (2021). Time for Cinderella to go to the ball: Reflections on the right to freedom of scientific research. *S. Afr. Law J.* 138 (2), 260–288. doi:10.47348/salj/v138/i2a2

Thaldar, D. W., Townsend, B. A., Donnelly, D. L., Botes, M., Gooden, A., Van Harmelen, J., et al. (2022). The multi-dimensional legal nature of personal genomic sequence data: A perspective from South Africa. *Front. Genet.* 13, 1–11. doi:10.3389/fgene.2022.997595

UN Human Rights Council (2009). *Resolution 10/23: Independent expert in the field on cultural rights*. Available at: https://ap.ohchr.org/documents/E/HRC/resolutions/A_HRC_RES_10_23.pdf.

UN Committee on Economic, Social and Cultural Rights (1990). *General Comment No. 3: The nature of state parties' obligations (art. 2, para. 1 of the Covenant)*. Available at: https://www.refworld.org/docid/4538838e10.html.

UNESCO International Bioethics Committee (1995). *Report of the second session proceedings*. Paris: UNESCO.

UNESCO (1997). *Universal Declaration on the Human Genome and Human Rights*. Available at: https://www.ohchr.org/en/instruments-mechanisms/instruments/universal-declaration-human-genome-and-human-rights.

UN General Assembly. *Universal Declaration on Human Rights*. Available at: https://www.un.org/en/about-us/universal-declaration-of-human-rights.

UN Committee on Economic, Social and Cultural Rights (2020). *General Comment No. 25 on economic, social and cultural rights (article 15 (1) (b), (2), (3) and (4) of the International Covenant on Economic, Social and Cultural Rights)*. Available at: https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2020-article-15-science-and.

UN Office of the High Commissioner for Human Rights (2023). *Ratification status for International Covenant for Economic, Social and Cultural Rights*. Available at: https://indicators.ohchr.org.

UN General Assembly (1966). International Covenant on Civil and Political Rights. Available at: https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights.

Vasquez, E., and García-Deister, V. (2018). "In pursuit of genomic justice: Sovereignty, inclusion and innovation in Mexico," in *Routledge Handbook on the Politics of Global Health*. Editors R. Parker and J. García (London: Routledge), 421–432.

Vasquez, E., and García-Deister, V. (2019). Mexican samples, Latino DNA: The trajectory of a national genome in transnational science. *Engaging Sci. Technol. Soc.* 5, 109.

Yotova, R., and Knoppers, B. M. (2020). The right to benefit from science and its implications for genomic data sharing. *Eur. J. Int. Law* 31 (2), 665–691. doi:10.1093/ejil/chaa028

frontiers | Frontiers in Pharmacology

TYPE Systematic Review
PUBLISHED 24 August 2023
DOI 10.3389/fphar.2023.1214422

OPEN ACCESS

EDITED BY
Alessandro Blasimme,
ETH Zürich, Switzerland

REVIEWED BY
Alfredo Vellido,
Universitat Politecnica de Catalunya, Spain
Josivania Silva Farias,
University of Brasilia, Brazil

*CORRESPONDENCE
Beverley Alice Townsend,
bev.townsend@york.ac.uk

RECEIVED 29 April 2023
ACCEPTED 18 July 2023
PUBLISHED 24 August 2023

CITATION
Townsend BA, Sihlahla I, Naidoo M, Naidoo S, Donnelly D-L and Thaldar DW (2023), Mapping the regulatory landscape of AI in healthcare in Africa.
Front. Pharmacol. 14:1214422.
doi: 10.3389/fphar.2023.1214422

Mapping the regulatory landscape of AI in healthcare in Africa

Beverley Alice Townsend [1,2]*, Irvine Sihlahla [2], Meshandren Naidoo [2], Shiniel Naidoo [2], Dusty-Lee Donnelly [2] and Donrich Willem Thaldar [2]

[1]York Law School, University of York, York, United Kingdom, [2]School of Law, University of KwaZulu-Natal, Durban, South Africa

Introduction: Artificial intelligence (AI)-enhanced technology has seen unprecedented expansion in the recent past. This growth brings with it huge opportunities for the positive transformation of the economy, business, healthcare, and society. However, a critical question is whether, and to what extent, regulatory measures and mechanisms have been implemented to safeguard its design, development, and deployment. This paper offers a scoping exercise that maps the regulatory landscape of AI in healthcare (including health research) in certain African countries.

Methods: This research is conducted across 12 African countries: Botswana, Cameroon, The Gambia, Ghana, Kenya, Malawi, Nigeria, Rwanda, South Africa, Tanzania, Uganda, and Zimbabwe. As limited specific AI legislation is found in these African countries, and because AI is informed by ancillary regulatory frameworks, we include data protection, digital health, consumer protection, and intellectual property in our research. A scoping review method was applied with a manual search of digital libraries with search terms customised for each repository consisting of core search terms for the various topics, including, among others, "law," "regulation," "artificial intelligence," "data protection," "intellectual property," and "digital health".

Results and discussion: Analysis of the data demonstrated that while in the African countries under investigation there is no sui generis AI regulation, recent developments were found in areas that inform AI adoption, including in digital health, data protection, consumer protection, and intellectual property. Our findings highlight the fragmentation of the African AI regulatory landscape and illustrate the importance of continued AI regulatory development to ensure that Africa is well positioned for future AI adoption in health.

KEYWORDS

artificial intelligence, AI, Africa, regulation, landscape, healthcare

Introduction

Artificial intelligence (AI) is a pivotal player in the emergence of the Fourth Industrial Revolution ('4IR'). Although no harmonised definition of AI exists, we take the broadly functionalist perspective that AI is to enable a machine or a mechanical device to function or behave in a manner that would be called intelligent were a human to behave in that manner (McCarthy et al., 2006). AI-enhanced technologies have recently expanded in scale, scope, and complexity, including a diverse range of applications globally (Aitken et al., 2022). One sector where AI

holds much promise is in its ability to revolutionise and drive healthcare. AI-based technology deployment is hugely advantageous in enhancing connectivity, facilitating the flow of health information, and in providing healthcare services and delivery. The provision of healthcare in Africa faces many challenges, including a shortage of healthcare resources, an increased burden of disease, a large proportion of the population living in rural areas, and a lack of education and primary healthcare–to name a few. Significant advantage may be harnessed by AI application: *inter alia*, in extending healthcare access, by contributing to early disease detection and prevention, supporting diagnostics and drug development, in disease surveillance and tracking, in public health monitoring, healthcare management and clinical decision-making, and in health research more generally (Topol, 2019).

While the disrupting power of such technologies brings with it unprecedented opportunities for the transformation of society and healthcare in particular, there are also concerns about the way in which AI is designed, developed, and deployed. These concerns range from issues concerning data quality and privacy to explainability and transparency of the algorithms, and issues of social and distributive justice (Fjeld et al., 2020). An analysis of current ethical AI guidelines found that while there was convergence of the normative themes (or principles) of transparency, justice, fairness, non-maleficence, and responsibility across many ethical frameworks, principles such as privacy, solidarity, human dignity, and sustainability were underrepresented (Jobin et al., 2019; World Health Organisation 1A, 2021). Notwithstanding the prevalence of ethical instruments–many of which find application in Africa–a critical question is whether AI in Africa is regulated. By regulation, we mean any form of 'hard' law–that is, policies that one can enforce in a court of law. While there is much recent development and debate about the regulation of AI in the Global North, far less attention has been directed toward, and indeed little is known about, the AI regulatory position in the Global South and in Africa, in particular (De Almeida et al., 2021; Schmitt, 2022).

Related work

Research studies have been conducted on the mapping of global AI ethics guidelines, on the ethical challenges presented by AI-driven technologies in healthcare, and on emergent ethical and rights-based approaches to values and principles for AI adoption and global AI governance. However, limited, if any, research has been done on ascertaining the current AI regulatory landscape in the Global South, and in Africa, in particular (Wang and Siau, 2018; Jobin et al., 2019; Fjeld et al., 2020; Gerke et al., 2020). Radu, for example, has conducted a qualitative comparison of the national strategies of 12 countries: Canada, China, France, Finland, Germany, Japan, Singapore, South Korea, Sweden, the United Arab Emirates, the United Kingdom, and the United States (Radu, 2021), Butcher and Beridze provided a synopsis of current AI governance activities globally (Butcher and Beridze, 2019), Larsson analysed the use of ethics guidelines as a governance tool in the development and use of AI with a focus on the Ethics Guidelines for Trustworthy AI published by the EU Commission's High-Level Expert Group on Artificial Intelligence (Larsson, 2020), and Cheng and Zeng reported on the global AI governance initiatives and China's ambition to play a leadership role in nascent global AI governance regimes (Cheng and Zeng, 2022). Concerning Africa, comparatively little research has been done. Brand has reviewed recent international developments and has recounted the insufficiency of only implementing a South African national legal framework–arguing in favour of the introduction of practical instruments of governance, such as an algorithmic impact assessment for measuring and mitigating risk and harm (Brand, 2022). In a not dissimilar vein, Abe and Eurallyah have explored the implications for human rights infringements, and Gwagwa et al. explored the core benefits and challenges of AI adoption in Africa (Gwagwa et al., 2020; Abe and Eurallyah, 2021).

Aim and scope

This article aims to complement current literature by mapping the regulatory landscape of AI in the health context in Africa, with reference to 12 selected countries. By mapping the regulatory landscape, we mean conducting a scoping review of the most relevant regulatory instruments–that is, those we identify and provide references to as regulatory instruments. We consider *AI in healthcare* broadly, including the development of AI-enhanced technology for use in healthcare practice and in health research. As Africa is the second largest and second most populous continent in the world and as it comprises six per cent of the earth's total surface area, we caution against viewing such an extensive and diverse region with such heterogeneous populations as one amassed, singularly constructed entity. Africa is a vast continent consisting of 54 sovereign states recognised by the United Nations. In this contribution, when the words "Africa," "African" or "African continent" are used they are intended to describe in particular the 12 African countries under investigation in our research (and not necessarily all 54 African states or countries)—unless the context dictates otherwise. We have thus restricted the scope of our article, for purposes of practicality, to a selection of 12 African countries: Botswana, Cameroon, The Gambia, Ghana, Kenya, Malawi, Nigeria, Rwanda, South Africa, Tanzania, Uganda, and Zimbabwe. These countries are 1) those English-speaking African countries which 2) hosted research projects as part of the H3Africa programme. The 12 selected countries are not representative of Africa generally, but rather represent a selection of African countries that 1) have regulatory instruments in a language that is understandable to us, the investigators, and 2) which can host health research activity that is relevant from a broader international perspective. These English-speaking African countries were selected for previously hosting Human Heredity and Health in Africa (H3Africa) consortium projects and have been included in the Data Science for Health Discovery and Innovation in Africa (DS-I Africa) Law project, which is funded by the National Institutes of Health (NIH).

Methodology

First, we investigate whether the selected countries have *sui generis* AI regulatory instruments. Next, we identify regulatory instruments in areas of the law that we suggest are most

proximate and relevant to AI in health: digital health law, data protection law, consumer protection law, and intellectual property law. Lastly, we investigate the regulatory authorities in these areas of law, as they can often create regulatory instruments in a dynamic fashion in anticipation of (or in response to) technological developments.

Our investigation follows the style of scoping reviews described in Munn et al. (2018) and Peters et al. (2015). This approach was chosen because of the lack of synthesised comprehensive databases and systematic reviews on the topic (Sucharew and Macaluso, 2019). A scoping review is therefore particularly appropriate to achieve this study's objective of mapping a wide body of regulatory instruments that may affect the emergent legal regulation of AI in healthcare (Munn et al., 2018).

To ensure a comprehensive and systematic search process, searches of various websites as described in Supplementary Annex II, including the Afriwise (Afriwise, 2022) portal and official government websites, were conducted. Where a keyword search was allowed on the website, we used an array of relevant search terms. If not, the sites were manually searched. Supplementary Annex II at Supplementary Tables S1, S2 set out a comprehensive list of the websites searched and the search terms used. The search protocol was developed and pilot-tested initially in South Africa and Kenya, and then applied to other jurisdictions. Each thematic area was surveyed individually, with 97 country-specific databases examined for AI and data protection, 54 for digital health, 11 for consumer protection and ICT law and 21 for intellectual property. In addition, 22 databases were reviewed for information relating to all countries; these include regional/sub-regional organisations and general legal research websites.

From the search results, the researchers downloaded digital copies of all documents relevant to one or more of the study's five themes. The criteria for inclusion in the scoping review were that the document be one of the following types: 1) national statute currently in force; 2) gazetted regulations; 3) draft Bill, 4) published government policy or strategy document; 5) ethics code/guideline/policy by health sector regulatory body or international or regional legislative, regulatory or policy instrument; and 6) applicable in one or more of the 12 jurisdictions. The study excluded private sector documents, documents not publicly available on the internet, and draft documents under discussion. The extracted documents were saved in a shared Google drive folder, and were classified in sub-folders by country and thematic area. Duplicates and documents replaced/repealed by a more recent document were then manually removed. A total of 118 documents (listed in Supplementary Annex I Supplementary Tables S1–S5) were then legally analysed by the researchers.

Limitations

While a full and comprehensive account of the regulatory position in the relevant areas of the law is offered, we do not claim to have captured every provision that may find relevance to AI in health or in health research. The consequences of AI adoption are far-reaching, touching many areas of the law. We have therefore narrowed our enquiry to the areas of the law that are most relevant. It was also not our intention to capture the numerous and varied ethics instruments and other non-regulatory governance measures that may find application to AI in healthcare.

Analysis

Sui generis AI regulation

There are no *sui generis* AI regulatory instruments at a regional African level or in any of the 12 African countries under investigation. However, as AI regulation is more than *sui generis* regulation, certain aspects of the development and deployment of AI are informed by either issue-specific legislation (such as healthcare laws) or sector-specific legislation (such as data protection laws). Our analysis categorises the regulations informing AI adoption in our study along four more generalisable themes: digital health law, data protection law, consumer protection law, and intellectual property law.

Digital health law

Regulatory frameworks form part of a tool to assess the maturity of AI within health. Thus, the absence of clear AI regulatory guidelines and policies may, in certain instances, impede the uptake of AI in the healthcare sector (Broadband Commission and Working Group on Digital and AI in Health, 2020). Although no *sui generis* AI regulation exists in the countries under investigation, healthcare is not unregulated. However, integrating AI into the existing healthcare and health research systems can present challenges. If there are regulatory voids, guidance is, to some extent, sought from existing national health statutes, digital health policy documents, professional codes of conduct, and healthcare and health research guidelines.

The World Health Organisation has implemented an integrated African Health Observatory initiative, together with National Health Observatories aimed at providing an informative digital health platform (World Health Organisation, 2022a). The African Union through its auspice, the African Medical Devices Forum (New Partnership for Africa's Development, 2022), has yet to provide regulatory guidance for the use of AI in clinical healthcare and research. At a subregional level, Kenya and Uganda belong to the IGAD (Intergovernmental Authority on Development) group which has developed a policy and implementation plan on health data sharing (Intergovernmental Authority on Development, 2021a; Intergovernmental Authority on Development, 2022b) efforts. These aspire to integrate cross-border health data sharing that in turn facilitates AI development and healthcare in Africa.

An analysis of statutes governing medical devices in the selected African countries shows that no single piece of legislation explicitly mentions AI or algorithms within the definition of a medical device. Furthermore, when compared to the definition of AI provided by the OECD, the definition of software included in current medical device regulations does not specifically and adequately address novel features of AI software. From the scrutiny of the provisions, 'software' is included in the definition of medical devices in three jurisdictions (South Africa, Kenya, and Uganda) (Matovu, 2018). This is broadly construed to include software as a medical device (SaMD) (Townsend, 2020). However, these regulatory frameworks

do not sufficiently provide AI system risk classifications (critical, serious, or non-serious categories); oversight mechanisms informing the total product life cycle of SaMDs (including pre-market development, post-market management, change management, and ongoing monitoring); guidance on the analytical and clinical validation of the SaMDs; direction on the testing, training, and validation of the datasets; and verification of the veracity and accuracy of the machine learning or other algorithms that underpin the SaMDs.

Reported cases demonstrate that AI has been instantiated in clinical practice in certain countries under investigation. For instance, South Africa used an AI-driven chest x-ray diagnosis application during the COVID pandemic (Staff Reporter IOL, 2020; Philips Foundation team, 2021). Digital health services for medical advice, appointment booking and the delivery of prescriptions to patients through mobile applications, and AI-powered triage systems have been launched in both Rwanda and Tanzania (Babyl, 2022; Elsa Health, 2022). However, the regulation of SaMDs by the medical health regulators is largely undeveloped, with most countries lacking frameworks providing guidance to digital healthcare modalities and applications. These policies are set out in Supplementary Annex I, Supplementary Table S1.

Most countries under study have standalone digital health policies in place, except for Rwanda and The Gambia where such policies are embedded in the broader healthcare policy. Kenya is the only country that has a standalone E-health Bill that regulates digital health. The implementation and monitoring processes of digital health are, however, sporadic and partly a result of a lack of infrastructure and resources (Akanbi et al., 2012; World Health Organisation, 2022b; Butcher et al., 2021; Karamagi et al., 2022; Odenkuhle et al., 2017; Owoyemi et al., 2020). Countries under study that have developed digital health policies have not, however, established professional guidelines for health practitioners, except for South Africa, Kenya, and Zimbabwe (EXCOM, 2014; Ministry of Health, 2017; Health Professions Council of South Africa, 2022). All countries under study have regulations that provide for informed consent, guide health personnel, and stipulate that medical professionals should be registered. In addition, Kenya's policy allows informed consent to be obtained electronically in line with the data protection law (Onetrust DataGuidance Regulatory Research Software, 2021; Kenya Government, 2023). Informed consent is of particular importance in health research as it safeguards the autonomy and dignity of research participants. Kenya's and Ghana's policies allow e-dispensing and e-prescriptions (The Pharmacy Council, Ghana, 2022; Kenya Government, 2023). Only four countries have professional guidelines and policies on telemedicine, namely, South Africa, Kenya, Zimbabwe, and Ghana. For the remainder of the countries, the position on electronic consent, e-dispensing, and e-prescriptions is unclear. Most countries under study do not have guidance on telemedicine which *inter alia* affects its development, adoption, and use in health research. With the absence of telemedicine and digital health guidance, and specific e-health strategies, the call for reform is supported, for example, in South Africa by Townsend et al. and in Botswana by Ncube et al. (Townsend et al., 2019; Ncube et al., 2022). The lack of development, use, and adoption of telemedicine could be related to *inter alia* resource constraints, and to ethical and legal barriers. Lack of adequate healthcare regulations or policies has been noted as a barrier to the adoption of telemedicine in Africa, with Dodoo et al. recommending that governments adopt a comprehensive e-policy framework including the establishment of strict protocols to monitor and evaluate telemedicine practices (Dodoo et al., 2021). These barriers will similarly stand to affect AI in health research.

However, more research is needed as telemedicine solutions are increasingly leveraging AI, as well as new modalities of delivering healthcare services in under-resourced areas, such as Chatbots and mobile applications, to assist community health workers. Regulation that is outdated and not context-specific and culturally appropriate can thus also act as a barrier to digital technology adoption and innovation. In South Africa, for example, there has been a low uptake of telemedicine by healthcare practitioners (Dodoo et al., 2021) and Donnelly has criticised the overly restrictive South African telemedicine guidelines as potentially stifling lawful and ethical development of AI in healthcare (Donnelly, 2022).

The development and adoption of AI in healthcare relies heavily on availability and access to high-quality clinical health data gained from digital health and health research (European Commission, 2022). Therefore, regulatory frameworks associated with the management of digital health data and health research are a foundational element for further development of AI technologies in healthcare. Most countries in the cohort study have legislation on digital health, and regulations that direct the professional conduct of health personnel in clinical and health research are set out in Supplementary Annex I, at Supplementary Table S2. These regulations determine the collection, storage, curation, management, and analysis of digital health data in research, which is vital for AI development and adoption.

In sum, the regulation of AI adoption in healthcare in the countries under study is undeveloped. None of the studied countries have adopted a proactive approach to the development of legislation governing AI in healthcare. The immaturity of AI in healthcare regulatory systems is exacerbated by further impediments including the lack of financial resources, diminished computing resources and structural infrastructure, and inadequate technical expertise. Unfortunately, these factors stand to delay the implementation of digital health in low- to middle-income countries–including the countries under study (World Health Organisation 1B, 2022). Professional guidelines, informed consent provisions, and healthcare and health research regulations provide some guidance and inform AI use in health contexts.

Data protection law

There is a close link between data and AI. AI systems rely on vast quantities of accurate, complete, representative, and quality datasets to train, test, and validate the system. Data that is typically personal - and sometimes sensitive or special category data - is typically 'research' data. AI systems also collect, generate, process, and share data–often on a large scale. Good AI regulation is thus intrinsically shaped by good data regulation. The increasing use and processing of such datasets informs many possible privacy challenges, including issues associated with collection, standardisation, anonymity, transparency, data ownership, and the changing conceptions of informed consent.

AI-enhanced technologies pose risks to data privacy in two ways. First, in the unlawful collection, use, and sharing of a person's personal data, and second in not providing persons with access, control, and autonomy over their data and data use. Legal tensions focus on the increasing requirement to access curated quality datasets and the inherent sensitivity of data, in particular personal information and also the implicit vulnerability to its unethical or unlawful source, use, and disclosure. The use and processing of personal data, and in particular sensitive health data and electronic health records, are well described, as are securing and protecting large-scale data sets against unauthorised collection, access, processing, storage, and distribution (Goodman, 2016; Bari and O' Neill, 2019; Xafis et al., 2019; Townsend and Thaldar, 2020).

Regional developments in Africa have primarily been instantiated through the African Union Convention on Cyber Security and Personal Data Protection, which was adopted in June 2014, and which introduced substantive claims to information privacy in Africa (African Union, 2014).

The AU Convention sought to harmonise African cyber legislation and to elevate the rhetoric of 'protection of personal privacy' to an international level. Moreover, it establishes a normative framework consistent with the African legal, cultural, economic, and social environment, and seeks to balance the use of information and communication technologies with the protection of the privacy of individuals, while guaranteeing the free flow of information across borders. The AU Convention enjoins state parties to establish legal and institutional frameworks for data protection and cybersecurity, encompassing three central issues: electronic transactions, personal data protection, and cybercrimes (African Union, 2014). The AU Convention requires 15 ratifications to enter into force. Recently, on 9 May 2023, it indeed reached 15 ratifications, and is therefore now in force (African Union, 2023).

A further development leading to data protection integration, strengthening collaboration in Africa, and facilitating cross-border data transfers occurred in February 2022 with the endorsement of the AU Data Policy Framework (African Union, 2022). This Framework encourages greater collaboration between AU member states and a coordinated, comprehensive, and harmonised approach to data governance.

In addition, subregional frameworks and agreements as created by the Economic Community of West African States (ECOWAS), the East African Community (EAC), the Economic Community of Central African States (ECCAS/CEMAC), the Intergovernmental Authority on Development (IGAD), and the Southern African Development Community (SADC), have contributed to the protection of the right to privacy and to promoting cyber security and fightingcybercrime (East African Community, 2011; Southern African Development Community, 2013a; Southern African Development Community, 2013b; Southern African Development Community, 2013c; East African Community, 2019; Intergovernmental Authority on Development, 2021a; Intergovernmental Authority on Development, 2021b; Economic Community of Central African States, 2021; Economic Community of West African States, 2021; Intergovernmental Authority on Development, 2022a; Intergovernmental Authority on Development, 2022b).

If Africa once lagged in the development of data protection laws, it has recently remedied this position. Until recently, few, if any, data protection policies had been developed in Africa (Van Gyseghem, 2012; Makulilo, 2015). This, however, has changed significantly. In 2021, of the 145 countries globally with data protection laws, 32 were in Africa with Africa the region of fastest data-protection law expansion (Greenleaf, 2021). The most recent African enactments are Tanzania, Egypt, Uganda, Togo, Nigeria, Kenya, Congo-Brazzaville, Botswana, and Zimbabwe (Tanzania; 2022; ILO, 2020; Uganda, 2019; Togo, 2019; Nigeria, 2019; Kenya, 2023; Congo-Brazzaville, 2019; Botswana, 2018; Zimbabwe, 2021; Wilkinson and Ooijevaar, 2020). Of the 12 countries we investigated, nine had specific data protection laws enacted. Botswana has the Data Protection Act No 32 of 2018 which came into force on 15 October 2021 (with the grace period of 1 year for implementation delayed beyond 15 October 2022). It establishes an Information and Data Protection Commission, yet to be set up, which is mandated to do all things necessary to protect the rights of individuals regarding their personal data and to ensure the effective application of the Botswana Data Protection Act. Both Kenya–one of the few countries whose law contains a specific Privacy-by-Design provision–and Ghana have data protection legislation. In Nigeria, data protection is provided by the Nigerian Data Protection Regulation of 2019, which is subsidiary legislation issued pursuant to the National Information Technology Development Agency Act of 2007. Moreover, the Data Protection Bill, 2020 (anticipated to be passed in 2023) seeks to provide an efficient regulatory framework for the protection of personal data and to regulate the processing of information.

Data protection in Rwanda is governed by law No 058/2021 of 2021 relating to the protection of personal data and privacy. Interestingly, Rwandan law contains a provision in Article 19 giving the data subject the right to request a data controller or data processor to stop processing their personal data which 'causes or is likely to cause loss, sadness or anxiety to the data subject' and a provision in Article 25 permitting a data subject to designate an heir to their personal data. In South Africa, data is protected by the Protection of Personal Information Act No 4 of 2013, which came into effect on 1 July 2020, Uganda by the Data Protection and Privacy Act of 2019, and Zimbabwe by the Data Protection Act No 5 of 2021. Tanzania enacted its first Personal Data Protection Law in late 2022, in terms of which provision is made for conducting transfer impact assessments and the stipulation that data collectors submit their privacy policies to the Tanzanian Data Protection Commission for approval.

Although not all countries have specific data protection legislation in place, all countries under investigation have data or privacy protection in some form or another, often embedded in other legislation. Cameroon, for example, has no specific law relating to data protection, although a degree of protection is provided by law No 2010/012 of 21 December 2010 Relating to Cyber security and Cyber criminality in Cameroon, by Law No 2006/018 of 29 December 2006 to Regulate Advertising in Cameroon, and by Law No 2010/013 of 21 December 2010 Regulating Electronic Communications in Cameroon. Moreover, the Constitution of the Republic of Cameroon provides for the privacy of all correspondence and Decree No 2013/0399/PM of 27 February 2013 for modalities of the

consumers' protection in the electronic communication sector states that "consumers in the electronic communication sector have the right to privacy . . . in the consumption of technologies, goods and services in the electronic communication sector." Cameroon has ratified certain instruments that protect privacy, including the sub-regional CEMAC Directive No 07/08-UEAC-133-CM-18.

In The Gambia, certain data protection and privacy rules relating primarily to information and communications service providers are provided for in their Information and Communications Act, 2009 and the 2019 Data Protection and Privacy Policy sets out the legal framework for data protection and privacy. Although Malawi does not have any specific data protection laws, a Data Protection Bill, 2021, has been drafted. It promotes data security and provides for data protection and related matters, while the Electronic Transactions and Cyber Security Act 33 of 2016 contains data protection-related provisions. We have included a comprehensive list of data protection laws in Supplementary Annex I, at Supplementary Table S3.

Consumer protection law

The debate about AI has focused on data protection requirements and soft law ethics instruments. While general AI regulation remains necessary, it is also vital to address the use of and relationship between AI software as goods that can be sold and the patient as a consumer in respect of the AI product or a healthcare service provided using the AI. Traditional fault-based liability regimes are difficult to implement in relation to harm caused by AI technologies as healthcare practitioners are required to foresee an error and take reasonable steps to meet the required standard of care (Donnelly, 2022; Naidoo et al., 2022). In other words, the law regards a doctor as negligent when they fail to act as a reasonable practitioner would have done in that branch of the profession. Considering the inherent opacity of the complex algorithms that power AI, it is highly unlikely that a doctor could reasonably be expected to anticipate errors that may not even be apparent to the AI developers. Imposing strict liability for harm caused by AI technologies has been extensively explored throughout the literature. However, it may be prudent to first investigate present means of imposing liability before we consider the development of new law/regulation. Many suggest that AI applications may necessitate a more sophisticated product liability regime (Chagal-Feferkorn, 2019), in order to address novel user safety risks found in such systems. The targeted jurisdictions have yet to address this matter and product liability for harm caused by AI is likely to be attributed according to the current consumer protection regime.

All 12 countries provide for consumer protection in relation to the sale of goods. Botswana, Cameroon, The Gambia, Kenya, Malawi, South Africa and Zimbabwe have enacted standalone statutes regulating consumer protection. The position is different elsewhere, where it is regulated alongside (Nigeria and Rwanda) or embedded in (Tanzania) fair competition legislation. While both Ghana (Nkansah, 2015) and Uganda (Zeija, 2018) currently have fragmented frameworks for consumer protection, they too have legislation regulating the sale of goods. The consumer protection legislation that does exist in these jurisdictions is set out in Supplementary Annex I at Supplementary Table S4.

Eleven out of the twelve countries provide for strict product liability of harmful or defective goods in their consumer protection regimes. This means that anyone in the supply chain for the AI product (the goods) can be held strictly liable for harm to the patient (the consumer) if the product does not perform safely or as intended. It is not necessary to prove that the harm arose from any negligence (fault) on the part of the developer or the doctor. Cameroon deviates from this general trend, as the imposition of product liability is negligence-based, that is, a determination of fault is necessary to impose liability (Galega, 2018).

Within current legislation, liability may be wholly or partly imposed on a number of different parties in the distribution chain, such as: the supplier, producer, manufacturer, importer, distributor, trader, seller, retailer, or provider of services (The Gambia, Malawi, and Nigeria). In South Africa, for example, the term supplier is wide enough to include the developer of the AI product and the healthcare establishment or practitioner providing a service using the AI product. Where health researchers intend to commercialise an AI product that they have developed, they too would need to be aware of the legal obligations imposed by consumer protection legislation. In addition, Rwanda's legislation contains a unique provision in terms of which strict product liability for unsafe or defective goods supplied by an enterprise is imposed upon the regulatory body that approved the product for sale.

A consideration of what types or aspects of technology may be included in the definition of goods is necessary. This becomes especially relevant to AI, given the recent CJEU finding that where the supply of software by electronic means is accompanied by a grant of perpetual licence, this will constitute the sale of goods (The Software Incubator Ltd, 2021). However, only Uganda, South Africa and Zimbabwe explicitly include software in the definition of goods. In seven other countries, software could be included by implication, as the term goods is either undefined (Cameroon, The Gambia, Kenya), or the nature of the goods covered is unspecified–but arguably wide enough–to include software. For example, Botswana defines *commodity* to include corporeal or incorporeal property; Ghana defines goods as 'movable property of every description'; while in Nigeria and Tanzania, goods are enumerated as–but not limited to–tangible goods. However, in Malawi, software is excluded because the Act applies to tangible goods only.

Definitions of what constitutes a consumer also vary. Seven countries–Botswana, Cameroon, Malawi, Nigeria, Tanzania, Uganda, and Zimbabwe–provide for the explicit exclusion of persons who purchase goods and services for the purpose of reuse in production and manufacture of any other goods or services for sale, and in Rwanda the Act applies only to goods ordinarily acquired for personal and domestic use. This is particularly noteworthy, given that statistically-based machine learning models used in the healthcare context will invariably be acquired for reuse in the production/manufacturing of other goods (e.g., drug discovery) and services (e.g., disease prediction, patient diagnoses, population health monitoring). Thus, those acquiring data-driven AI technologies for the purposes of health research or use in healthcare practice–where the objective is the sale of a good/service–are not themselves defined as consumers and are thus unlikely to find much protection under consumer legislation. In ensuring compliance with legislation, eight countries–Cameroon,

The Gambia, Malawi, Nigeria, Rwanda, South Africa, Tanzania, and Zimbabwe–allow the relevant consumer protection authority to issue a recall on any goods considered a risk to the public or harmful to human or public health. The Gambia and Tanzania differ in that the supplier or relevant party of the distribution chain is responsible to recall harmful or defective goods. Furthermore, both The Gambia and Malawi provide for an additional safeguard against harmful technology, goods, and services. Here producers or suppliers are intended to attach easily noticeable warnings to products considered harmful or hazardous to human health with the aim that use take place under the strongest possible safety conditions.

In addition, electronic communications and transactions and the protection of e-consumers are regulated in a number of jurisdictions in other legislation. These statutes, which do not refer in specific terms to AI, also do not contain any provisions that could clarify the attribution of liability or address many of the other significant consumer protection concerns that arise from the use of AI in healthcare. In addition, some jurisdictions have laws regulating cybercrimes, content control measures and service provider liability. These safeguards also do not directly address the issue of providing civil redress to individual consumers harmed by an AI application in the healthcare setting.

Intellectual property law

Before one can engage with research, one must first understand the regulatory environment. Importantly, this includes the schemes of protection for any fruits of research. This would be intellectual property. In this section we outline the mechanisms and bodies which are relevant in obtaining such protection. Multiple layers of intellectual property ('IP') protection can apply to a single AI product or process. For this research study we focused on only two IP rights: patents and copyright. These IP rights inform data flow, affect AI research and development, and are critical for AI innovation. Patents generally apply to product inventions (such as AI technologies embedded within products, for example, smartwatches). Copyright applies to literary works, which includes the datasets used to test, train, and validate AI systems. Regional IP frameworks were identified, as was national legislation in each of the selected African countries to denote the relevant avenues of protection and the mechanisms of protection which operate at each level.

The current members of the African Regional Intellectual Property Organisation (ARIPO) include Botswana, The Gambia, Ghana, Kenya, Malawi, Rwanda, the United Republic of Tanzania, Uganda, and Zimbabwe (African Regional Intellectual Property Organisation, 2023). South Africa and Nigeria, while not members under ARIPO, have observer status (Harakenzo World Patent and Trademark, 2023). Under the Harare Protocol, ARIPO can grant and register patents, industrial designs and utility models on behalf of contracting countries. The Protocol is currently in force in 18 of the 19 member countries (the exception being Somalia).

All of the countries under study have enacted patent and copyright statutes which are similar in many ways. The legislation is captured in Supplementary Annex I and Supplementary Table S5. All countries offer copyright protection (and share similar provisions) for the protection of computer programs and compilations of data and/or data tables. Any parties seeking protection for their data records and computer programs can obtain them in all 12 African countries.

Patent protection is available in all selected African countries for AI applications such as core inventions relating to novel advances in model architectures or to the techniques themselves. Other patentable innovations include: novel ways of generating a training set or model; trained models (the most common being AI as a tool to solve a particular problem); and smart AI-enhanced products and health monitoring devices.

Relevant authorities

All jurisdictions have yet to establish authorities or oversight mechanisms mandated to regulate AI. However, regulatory bodies and authorities overseeing data protection, ICTs, and medical devices will play a role in the regulation of AI systems and application in healthcare. The establishment of such authorities is set out in Supplementary Annex I at Supplementary Table S6.

Three of the twelve countries have established relevant committees to guide the uptake of emerging technologies, each of which has produced 4IR strategy documents. In 2018, the Kenyan Cabinet Secretary for ICT appointed the *ad hoc* Distributed Ledger and Artificial Intelligence Taskforce to: critically review AI, contextualise how its application could achieve the goals of, *inter alia*, universal healthcare and enhanced government service provision and to 'prepare an implementation strategy with key performance indicators and clear delivery timelines' (Authority of the Republic of Kenya, 2018). Similarly, in 2018, Uganda established the National Expert Taskforce on the 4IR, which was aimed at determining the state of 4IR technologies in the country, reviewing the legal and policy landscape, recommending a 4IR strategy and national institutional framework, and advising on a national framework intended to solidify the country as a 4IR regional hub (Ministry of ICT & National Guidance, 2022). In South Africa, the Presidential Commission on the 4IR (PC4IR) was mandated (South African Government 1A, 2022) to develop an integrated national strategy and advise on the advancement of global competitiveness, research and skills development. The PC4IR is also tasked with making recommendations to clearly articulate the roles of the state, constitutional actors, and citizens (South African Government 1B, 2022).

In addition, Rwanda and South Africa have established Centres for the Fourth Industrial Revolution–multi-stakeholder initiatives intended to focus on data governance, AI and machine learning (World Economic Forum, n.d.; Centre for the Fourth Industrial Revolution South Africa, 2022). These remain the only African countries that have partnered with the World Economic Forum in developing a network 'connecting technology policy experts and stakeholders across 16 advanced and emerging economies' (World Economic Forum, n.d.).

Conclusion

This work demonstrates that in the 12 selected African countries, AI in healthcare, including in health research, is regulated. However, a diverse and fragmented progress indicates that significant work is yet to be done. Certain selected African countries have made limited progress and all of the 12 selected

African countries are at an early stage in their AI regulatory journey. Notwithstanding regulatory developments, where found, development is often either of general application to all technology or adapted from other older digital technology types.

Encouragingly, certain sectors that inform AI development such as data protection have seen increased development in recent years. This is to be welcomed as exchanging and sharing knowledge, data, and efficiencies between African countries is transformative and can help to build common AI capacity across Africa. This is of particular importance in health research. We have identified the AI-relevant regulators and regulations–and instances where regulatory bodies and regulation are either absent or require strengthening. What is now required is a concerted effort by those regulators to engage with each other, and with health sector stakeholders and health researchers, to address gaps and deficiencies through domestic legal reform and policy development.

Importantly, where a regulatory framework exists, its role, we suggest, should be two-fold: to both prevent AI-related harm and to promote AI innovation across Africa. However, whether extant regulation achieves this and is suitable in the selected target countries for the purposes of AI adoption remains unclear. Where digital health policies and professional guidelines are absent or inadequate they need to be adopted or amended to enable responsible development and deployment of AI both in face-to-face patient care and telemedicine solutions, without stifling innovation. On AI innovation, AI generative tools promise to produce value. However, questions arise about whether these products qualify for intellectual property rights given that there is argument over whether they are created by a human or AI. African countries can certainly benefit by providing guidance on this important matter. In addition, there is limited African scholarship on AI ethics and policy, which makes for important and necessary future research in Africa.

Accordingly, Africa stands to gain from the proliferation of international and sector-specific ethical standards, guidelines, and policies, developed in a response to create "trustworthy," "transparent," and "responsible" AI (European Commission, 2019; Jobin et al., 2019; OECD Expert Group on AI, 2019). While certain jurisdictions outside of the African continent have proposed specific AI legislation, most notably the proposed European Union "Regulation laying down harmonization rules on Artificial Intelligence" (the "EU AI Act") (European Commission, 2021) and the US Algorithmic Accountability Act of 2022, other regions have opted for alternative approaches to AI regulation such as those under consideration in the United Kingdom White Paper on AI regulation published in March 2023 (UK, 2023). In the Global South, the Brazilian Artificial Intelligence Bill, enacted in 2021, contains principles, rights, and duties for the use of artificial intelligence in Brazil, Uruguay adopted an AI strategy on responsible AI in public administration in 2019, Peru and Colombia issued National AI Strategies, and Chile, a National AI Policy.

Africa can certainly draw on these perspectives and benefit from more general and broader policy guidelines and regulation on AI, and specifically on AI in healthcare and health research. The African Union too can play a role in directing such initiatives. The post-colonial reach of digitised data and AI create challenges to Africa's quest for digital sovereignty. However, Africa and indeed most of its nation states have been slow to agree on key digital and data governance measures. For example, as the uptake of the African Union Convention on Cyber Security and Personal Data Protection has demonstrated, progress is often both long and slow (African Union, 2014; Gwagwa and Townsend, 2023). What an appropriate and effective approach for AI regulation would look like for Africa and its individual sovereign nation states and how it may be implemented is an area for urgent and much needed future research.

We identify the role of the local community and African society in establishing principles and in participation and engagement in regulatory policy-making. The AI ecosystem is global, necessitating greater international collaboration and agreement of standards, frameworks, and guidance. Thus, the need exists to align the African position with international standards. However, while the Global North can inform African regulatory development and work at a global level to implement effective AI standards for safety, for example, and can bind countries to certain rules (Metzinger, 2022), we caution against a position where the normative principles and values that guide global AI adoption do not integrate as many perspectives as possible, including African viewpoints. Consideration must be had of the many unique historical and current challenges presenting in Africa. As suggested in Goffi and Momcilovic, we endorse an approach that embraces multiculturalism, and which offers due respect for cultural diversity in AI governance. An approach that is respectful of a variety of ethical perspectives and which involves multilateral debates at local and global levels (Goofy and Momcilovic, 2022).

Notwithstanding the emerging global approaches, we recommend that AI regulation in Africa is best served by being pro-innovation while addressing the many AI practices that carry unacceptable or high risk to health, safety, and human rights infringements. A framework for AI regulation in Africa, we suggest, should follow a cautious, yet proactive and balanced regulatory approach–one that is risk-based, rights-preserving, agile, adaptive, and innovation-supportive. In addition, we suggest that an effective African governance approach should include various governance tools–a combination of hard and soft law-including: 1) mechanisms to capture AI due diligence; 2) principles of transparency, explainability, and accountability; 3) be human-centric; and 4) make provision for AI auditing, assessment, and review. We recommend that an African approach be both risk-based and rights-based. This is premised on the understanding that AI systems have certain characteristics (*inter alia*, an opacity, complexity, dependency on data, and a capacity for autonomous behaviour) that can adversely and significantly affect fundamental human rights–rights to data privacy, transparency and disclosure, autonomy and self-determination, and the like.

Regulators in Africa have an increasing responsibility to address the immediate and significant concerns of algorithmic bias and fairness in the adoption of AI in Africa. AI stands, not only to potentially produce biased outcomes, but also to amplify and perpetuate patterns of general systemic and structural social bias, such as race- and gender-discrimination (Susskind, 2018; Kearns and Roth, 2020). Algorithmic injustice arises when patterns of marginalisation, imprinted in the historical data that shape the training and the testing of the system, produce individual predictive anomalies that, if

left unchecked, inform a pernicious feedback loop of further exacerbating future down-stream systemic and structural injustice within larger groupings (Kearns and Roth, 2020; Glickman and Sharot, 2022). Algorithmic injustice is aggravated where data are under-representative or exclude certain categories of persons resulting in the exacerbation of long-standing societal biases that exist in relation to protected features like race and gender, and are magnified by virtue of their reach and scale.

Better or worse futures in the region will be determined, we suggest, in large part by clearly understanding and articulating the perspectives of previously marginalized and silenced voices and allowing them to be part of the AI conversation. Zimmermann et al. argue that "algorithmic injustice is not only a technical problem, but also a moral and political one, and that addressing it requires deliberation by all of us as democratic citizens." Accordingly, accountability for addressing these injustices becomes shared, rather than that only "offloaded and outsourced to tech developers and private corporations" (Zimmermann et al., 2020).

The overarching idea too is that the higher the risk level, the greater the need for obligations to be placed on the AI system (and those developing and deploying it) and for human protection. Due regard should also be given to those activities that should be prohibited or otherwise curtailed, for example, amongst others, those outlined in the EU AI Act, that is, the use of systems that manipulate human behaviour and/or exploit persons' vulnerabilities and social scoring systems. While AI systems pose many immediate risks to Use short dashes in order to be consistent with the rest of the paper also pose broader, longer-term social harms and large-scale, highly consequential risks that are often difficult to predict ex ante (Kolt, 2023). Further research and focus should be placed on these longer-term risks and on those that have broader social impact in a proposed African AI regulatory solution.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, and further inquiries can be directed to the corresponding author.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

We acknowledge the support by the US National Institute of Mental Health and the US National Institutes of Health (award number U01MH127690). The work of BT was funded by the UKRI project EP/V026747/1 "Trustworthy Autonomous Systems Node in Resilience." The content of this article is solely our responsibility and does not necessarily represent the official views of the US National Institute of Mental Health or the US National Institutes of Health.

Acknowledgments

The completion of the tables in the Annex was fact checked by advice received from: Keneilwe P. Mere (Moribame Matthews, Botswana); Eleng Mugabe (Desai Law Group, Botswana); Hyacinthe Fansi (NFM Avocats Associés, Cameroon); Naa Asheley Ashittey (ÁELEX Ghana Unlimited, Ghana); Susan-Barbara Kamapley (Bentsi-Enchill, Letsa and Ankomah, Ghana); Benedict Nzioki (African Law Partners, Kenya); Frances Obiago (ÁELEX Nigeria Unlimited, Nigeria); Sumbo Akintola (Aluko and Oyebode, Nigeria); Zackiah Nandugwa (K-Solutions and Partners, Rwanda); Karl Blom (Webber Wentzel, South Africa); and Ronald Mutasa (Manokore Attorneys, Zimbabwe). We gratefully acknowledge the assistance of the academic collaborators on the DSI-Africa Law project: Dr Paul Ogendi (University of Nairobi, Kenya); Dr Peter Munyi (University of Nairobi, Kenya); Dr Lukman Abdulrauf (University of Ilorin, Nigeria); Dr Aishatu Adaji (University of Ilorin, Nigeria); Ms Elisabeth Anchancho (University of KwaZulu-Natal, South Africa) and Ms Amy Gooden (University of KwaZulu-Natal, South Africa). In addition, we acknowledge the inputs from the project research assistants: Kiara Munsamy (University of KwaZulu-Natal, South Africa); Jodie de Klerk (University of KwaZulu-Natal, South Africa); Roasia Hazarilall (University of KwaZulu-Natal, South Africa); and Naseeba Sadak (University of KwaZulu-Natal, South Africa). All errors and omissions remain the authors' responsibility.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2023.1214422/full#supplementary-material

References

Abe, O., and Eurallyah, A. J. (2021). Regulating artificial intelligence through a human rights-based approach in Africa. *Afr. J. Leg. Stud.* 14 (14), 425–448. Available from: doi:10.1163/17087384-12340084

African Regional Intellectual Property Organisation (2023). Aripo – AP jurisdiction – general information online patent register and online gazette additional information. https://inspire.wipo.int/system/files/aripo.pdf.

African Union (2014). African Union Convention on Cyber Security and Personal Data Protection. https://au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection.

African Union (2022). *AU data policy framework*. https://au.int/sites/default/files/documents/42078-doc-AU-DATA-POLICY-FRAMEWORK-ENG1.pdf.

African Union (2023). List of countries which have signed, ratified/acceded to the African Union Convention on Cyber Security and Personal Data Protection. https://dataprotection.africa/wp-content/uploads/2305121.pdf.

Afriwise (2022). *Afriwise – African law made simple*. Available from: https://afriwise.com/.

Aitken, M., Leslie, D., Ostmann, F., Pratt, J., Margetts, H., and Dorobantu, C. (2022). *Common regulatory capacity for AI*. London, England: The Alan Turing Institute. Available from: doi:10.5281/zenodo.6838946

Akanbi, M. O., Ocheke, A. N., Agaba, P. A., Daniyam, C. A., Agaba, E. I., Okeke, E. N., et al. (2012). Use of electronic health records in sub-Saharan Africa: progress and challenges. *J. Med. Tropics* 14 (1), 1–6.

Authority of the Republic of Kenya (2018). Kenya gazette notice number 2095 of 2018. Available from: https://gazettes.africa/archive/ke/2018/ke-government-gazette-dated-2018-05-11-no-56.pdf.

Babyl (2022). Babyl – Rwanda's digital healthcare provider. Available from: https://www.babyl.rw/.

Bari, L., and O'Neill, D. P. (2019). Rethinking patient data privacy in the era of digital health. Available from: https://www.healthaffairs.org/do/10.1377/forefront.20191210.216658.

Brand, D. (2022). Responsible artificial intelligence in government: development of a legal framework for South Africa. *eJournal eDemocracy Open Gov.* 14 (1), 130–150. doi:10.29379/jedem.v14i1.678

Broadband Commission, Working Group on Digital and AI in Health (2020). *Reimagining global health through artificial intelligence: the roadmap to AI maturity*. https://www.broadbandcommission.org/publication/reimagining-global-health-through-artificial-intelligence/.

Butcher, J., and Beridze, I. (2019). What is the state of artificial intelligence governance globally? *RUSI J.* 164 (5-6), 88–96. doi:10.1080/03071847.2019.1694260

Butcher, N., Wilson-Strydom, M., Baijnath, M., Orlic, D., Smith, M., Neupane, B., et al. (2021). *Artificial intelligence capacity in sub saharan Africa – compendium report*. Artificial Intelligence for Development Africa.

Centre for the Fourth Industrial Revolution South Africa (2022). *Centre for the Fourth industrial revolution*. Available from: https://www.c4ir.co.za/ (Accessed October 17, 2022).

Chagal-Feferkorn, K. A. (2019). Am I an algorithm or a product? When products liability should apply to algorithmic decision-makers. *Stanf. Law Policy Rev.* 30 (61), 62–113.

Cheng, J., and Zeng, J. (2022). Shaping AI's future? China in global AI governance. *J. Contemp. China* 32, 794–810. doi:10.1080/10670564.2022.2107391

Congo-Brazzaville (2019). *Loi 29-2019 portant protection des donnees a caractere personnel*. Geneva, Switzerland: International Labour Organization.

De Almeida, P. G. R., Dos Santos, C. D., and Farias, J. S. (2021). Artificial intelligence regulation: a framework for governance. *Ethics Inf. Technol.* 23, 505–525. doi:10.1007/s10676-021-09593-z

Dodoo, J. E., Al-Samarraie, H., and Alzahrani, A. I. (2021). Telemedicine use in sub-saharan Africa: barriers and policy recommendations for Covid-19 and beyond. *Int. J. Med. Inf.* 151, 104467. doi:10.1016/j.ijmedinf.2021.104467

Donnelly, D. (2022). First do no harm: legal principles regulating the future of artificial intelligence in healthcare in South Africa. *Potchefstroom Electron. Law J.* 25 (1), 1–43. doi:10.17159/1727-3781/2022/v25ia11118

East African Community (2019). *EAC digital reach initiative strategic plan 2019–2028*. Arusha, Tanzania: East African Community.

East African Community (2011). *EAC draft legal framework for cyber laws 2008/2011*. Arusha, Tanzania: East African Community.

Economic Community of Central African States (2021). *Plan Stratégique Indicatif à moyen terme (2021–2025) de la Communauté Économique des États de l'Afrique Centrale (CEEAC)*. Libreville, Gabon: Economic Community of Central African States.

Economic Community of West African States (2021). *ECOWAS regional cybersecurity and cybercrime strategy adopted in 2021*. Libreville, Gabon: Economic Community of Central African States.

Elsa Health (2022). Elsa Health AI for clinical decision support. https://www.elsa.health/.

European Commission (2019). *High-level expert group on artificial intelligence. Ethics guidelines for trustworthy AI*. Available from: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.

European Commission (2021). *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021) 206 final)*. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206.

European Commission (2022). *Directorate-general for communications networks*. Study on eHealth, interoperability of health data and artificial intelligence for health and care in the European Union: final study report. lot 2, artificial intelligence for health and care in the EU. Luxembourg, Europe: Publications Office of the European Union.

EXCOM (2014). *Medical and Dental Practitioners Council of Zimbabwe Policy on International Medicine*. EXCOM.

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., and Srikumar, M. (2020). *Principled artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI*. Cambridge, MA: *Berkman Klein Center for Internet and Society*, Research Publication.

Galega, S. (2018). *Cameroon product liability in perspective: lessons from abroad*. Cambridge, MA: Academic Publishing.

Gerke, S., Minssen, T., and Cohen, G. (2020). "Ethical and legal challenges of artificial intelligence driven healthcare," in *Artificial intelligence in healthcare*. Editors A. Bohr and K. Memarzadeh (Cambridge, MA: Academic Press), 295–336.

Glickman, M., and Sharot, T. (2022). Biased AI systems produce biased humans. *OSF Prepr.* doi:10.31219/osf.io/c4e7r

Goodman, K. W. (2016). *Ethics, medicines, and information technology: intelligent machines and the transformation of health care*. Cambridge, MA: Cambridge University Press.

Goofy, E. R., and Momcilovic, A. (2022). Respecting cultural diversity in ethics applied to AI: a new approach for a multicultural governance. *Rev. Misión Juríd.* 15 (23), 111–122. doi:10.25058/1794600x.2135

Greenleaf, G. (2021). Global data privacy laws 2021: despite COVID delays, 145 laws show GDPR dominance. *Priv. Laws Bus. Int. Rep.* 169 (1), 1–5. doi:10.2139/ssrn.3836348

Gwagwa, A., Kraemer-Mbula, E., Rizk, N., Rutenberg, I., and De Beer, J. (2020). Artificial intelligence (AI) deployments in Africa: benefits, challenges and policy dimensions. *Afr. J. Inf. Commun.* 26, 1–28. doi:10.23962/10539/30361

Gwagwa, A., and Townsend, B. (2023). Re-imagining Africa's sovereignty in a digitally interdependent world. *Glob. Policy*. Available from: https://www.globalpolicyjournal.com/blog/10/05/2023/re-imagining-africas-sovereignty-digitally-interdependent-world.

Harakenzo World Patent and Trademark (2023). Harakenzo World Patent and Trademark. https://www.harakenzo.com/en/index.shtml.

Health Professions Council of South Africa (2022). Ethical guidelines for good practice in the health care professions booklet 10. Available from: https://aestheticdoctors.co.za/wp-content/uploads/2023/02/Booklet-10-Guidelines-for-the-practice-of-Telemedicine.pdf.

ILO (2020). Egypt Personal Data Protection Law 2020, law No. 151 of 2020. Available from: https://www.ilo.org/dyn/natlex/natlex4.detail?p_lang=en&p_isn=111246&p_count=7&p_classification=01#:~:text=The%20Law%20sets%20out%20the,process%20or%20control%20data%20electronically.

Intergovernmental Authority on Development (2021a). Regional health data sharing and protection policy framework. Available from: https://igad.int/igad-regional-health-data-sharing-and-protection-policy-framework/.

Intergovernmental Authority on Development (2022b). Regional health data sharing and protection policy framework implementation guide. Available from https://igad.int/download/igad-regional-health-data-sharing-and-protection-policy-framework-implementation-guide/ (June 13, 2022).

Intergovernmental Authority on Development (2022a). Regional health data sharing and protection policy. Available from https://igad.int/download/regional-health-data-sharing-and-protection-policy/ (September 20, 2022).

Intergovernmental Authority on Development (2021b). *Regional strategy for the development of statistics*, 2021–2025.

Jobin, A., Ienca, M., and Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nat. Mach. Intell.* 1 (1), 389–399. Available from: doi:10.1038/s42256-019-0088-2

Karamagi, H. C., Muneene, D., Droti, B., Jepchumba, V., Okeibunor, J. C., Nabyonga, J., et al. (2022). eHealth or e-Chaos: the use of digital health interventions for health systems strengthening in sub-Saharan Africa over the last 10 years: a scoping review. *J. Glob. Health* 12, 04090–04099. doi:10.7189/jogh.12.04090

Kearns, M., and Roth, A. (2020). *The ethical algorithm*. Oxford, United Kingdom: Oxford University Press.

Kenya Government (2023). *Kenya national eHealth policy 2016–2030*. Nairobi, Kenya: Kenya National eHealth.

Kolt, N. (Forthcoming 2023). Algorithmic black swans (February 25, 2023). *Wash. Univ. Law Rev.* 101.

Larsson, S. (2020). On the governance of artificial intelligence through ethics guidelines. *Asian J. Law Soc.* 7 (7), 437–451. doi:10.1017/als.2020.19

Makulilo, A. B. (2015). Myth and reality of harmonisation of data privacy policies in Africa. *Comput. Law Secur. Rev.* 31 (1), 78–89. doi:10.1016/j.clsr.2014.11.005

Matovu, B. "Regulation and standards of medical devices in Uganda," in Proceedings of the BME Conference, Ho Chi Minh, Vietnam, June 2018.

McCarthy, J., Minsky, M. L., Rochester, N., and Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence. *AI Mag.* 27 (4), 12–14. doi:10.1609/aimag.v27i4.1904

Metzinger, T. (2022). "Towards a global artificial intelligence charter," in *The Cambridge handbook of responsible artificial intelligence*. Editors S. Voeneky, P. Kellmeyer, O. Mueller, and W. Burgard (Cambridge, MA: Cambridge University Press), 167–175.

Ministry of Health (2017). Kenya standards and guidelines for mHealth. Republic of Kenya. Available from: https://www.health.go.ke/wp-content/uploads/2020/02/Revised-Guidelines-For-Mhealth-Systems-May-Version.pdf.

Ministry of ICT and National Guidance (2022). *Uganda's National 4IR Strategy: a continental 4IR hub that enables a smart and connected Ugandan society*. Kampala, Uganda: Ministry of ICT and National Guidance.

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., and Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med. Res. Methodol.* 18 (1), 143. doi:10.1186/s12874-018-0611-x

Naidoo, S., Bottomley, D., Naidoo, M., Donnelly, D., and Thaldar, D. W. (2022). Artificial intelligence in healthcare: proposals for policy development in South Africa. *South Afr. J. Bioeth. Law* 15 (1), 11–16. doi:10.7196/sajbl.2022.v15i1.797

Ncube, B., Mars, M., and Scott, R. E. (2022). Telemedicine options to address identified health needs in Botswana. *Digit. Health* 8, 20552076221102768. doi:10.1177/20552076221102768

New Partnership for Africa's Development (2022). African Medical Devices Forum. Regulation. Available from: https://www.nepad.org/microsite/africa-medical-devices-forum-amdf.

Nkansah, L. A. (2015). Consumer protection in Ghana: an appraisal of the law. *Joseph Ayo Babalola Univ. Law J.* 2 (1), 185–200.

Odekunle, F. F., Odekunle, R. O., and Shankar, S. (2017). Why sub-Saharan Africa lags in electronic health record adoption and possible strategies to increase its adoption in this region. *Int. J. Health Sci.* 11 (4), 59–64.

OECD Expert Group on AI (2019). Princ. Artif. Intell. *OECD. Recomm. Counc. Artif. Intell. (OECD Leg. Instrum. OECD/LEGAL/0449)*. Available from: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL0449.

Onetrust DataGuidance Regulatory Research Software (2021). *ODCP guidance note on consent*. Atlanta, GA: Onetrust DataGuidance Regulatory Research Software.

Owoyemi, A., Owoyemi, J., Osiyemi, A., and Boyd, A. (2020). Artificial intelligence for healthcare in Africa. *Front. Digital Health* 6 (2), 6–5. doi:10.3389/fdgth.2020.00006

Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., and Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *JBI Evid. Implement.* 13 (3), 141–146. doi:10.1097/XEB.0000000000000050

Philips Foundation team (2021). Philips Foundation deploys AI software in South Africa to detect and monitor COVID-19 using chest X-rays. Available from: https://www.philips-foundation.com/a-w/articles/CAD4COVID.html.

Radu, R. (2021). Steering the governance of artificial intelligence: national strategies in perspective. *Policy Soc.* 40 (2), 178–193. doi:10.1080/14494035.2021.1929728

Schmitt, L. (2022). Mapping global AI governance: a nascent regime in a fragmented landscape. *AI Ethics* 2 (2), 303–314. doi:10.1007/s43681-021-00083-y

South African Government 1A (2022). *Presidential Commission on Fourth Industrial Revolution: members and terms of reference*. South Africa: South African Government.

South African Government 1B (2022). *Report of the Presidential Commission on the 4th Industrial Revolution*. South Africa: South African Government.

Southern African Development Community (2013a). *Computer crime and cybercrime: Southern African Development Community (SADC) Model Law*. Gaborone, Botswana: Southern African Development Community.

Southern African Development Community (2013b). *Data protection: Southern African Development Community (SADC) Model Law*. Gaborone, Botswana: Southern African Development Community.

Southern African Development Community (2013c). *Electronic transactions and electronic commerce: Southern African Development Community (SADC) Model Law*. Gaborone, Botswana: Southern African Development Community.

Staff Reporter, IOL (2020). *South Africa harnesses artificial intelligence, machine learning in Covid-19 fight*. Available from: https://www.iol.co.za/business-report/companies/south-africa-harnesses-artificial-intelligence-machine-learning-in-covid-19-fight-9f9b0dba-65d4-41a8-8196-cd2258fae312.

Sucharew, H., and Macaluso, M. (2019). Progress notes: methods for research evidence synthesis: the scoping review approach. *J. Hosp. Med.* 14 (7), 416–418. doi:10.12788/jhm.3248

Susskind, J. (2018). *Future politics*. Oxford, United Kingdom: Oxford University Press.

The Pharmacy Council Ghana (2022). *Electronic pharmacy policy 2021*. Accra, Ghana: The Pharmacy Council Ghana.

The Software Incubator Ltd (2021). *Computer Associates (UK) Ltd (case C-410/19)*. Norwich, UK: The Software Incubator Ltd.

Togo (2019). Loi 2019-014 relative a la protection des donnees a caractere personnel. Available from: https://jo.gouv.tg/sites/default/files/JO/JOS_29_10_2019-64E%20ANNEE-N%C2%B026%20TER.pdf.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nat. Med.* 25 (25), 44–56. doi:10.1038/s41591-018-0300-7

Townsend, B. A. (2020). Software as a medical device: critical rights issues regarding artificial intelligence software-based health technologies in South Africa. *J. South Afr. Law/Tydskrif vir die Suid-Afrikaanse Reg* (4), 747–762.

Townsend, B. A., and Thaldar, D. W. (2020). Navigating uncharted waters: biobanks and informational privacy in South Africa. *South Afr. J. Hum. Rights* 35 (4), 329–350. doi:10.1080/02587203.2020.1717366

Townsend, B. A., Mars, M., and Scott, R. E. (2019). The development of ethical guidelines for telemedicine in South Africa. *South Afr. J. Bioeth. Law* 12 (1), 19–26. doi:10.7196/SAJBL.2019.v12i1.662

United Kingdom (2023). AI regulation: a pro-innovation approach. Department of Science, Innovation and Technology and the Office for Artificial Intelligence. Available from: https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach.

Van Gyseghem, J. M. (2012). *Model Law on Data Protection Support for Harmonization of ICT Policies in sub-Saharan Africa*. Geneva, Switzerland: International Telecommunications Union ITU.

Wang, W., and Siau, K. (2018). Artificial intelligence: a study on governance, policies, and regulations. *MWAIS 2018 Proc.* 40, 1–5.

Wilkinson, D., and Ooijevaar, M. (2020). Egypt's data protection law enters into force in October. *Priv. Laws Bus. Int. Rep.* 167 (1), 6–8.

World Economic Forum (no date). Centre for the Fourth Industrial Revolution Rwanda (C4IR Rwanda). https://c4ir.rw/ (Accessed October 17, 2022).

World Health Organisation 1A (2022). *Ethics and governance of artificial intelligence for health: WHO guidance*. Geneva, Switzerland: World Health Organisation.

World Health Organisation 1B (2022). *Global strategy on digital health 2020–2025*. Geneva, Switzerland: World Health Organisation.

World Health Organisation (2022a). Region. African Health Observatory. Available from: https://aho.afro.who.int/.

World Health Organisation (2022b). *Regional Committee for Africa. eHealth Solutions Afr. Region Curr. Context Perspect. Resolut. AFR/RC56/R8* 7. Available from: https://www.afro.who.int/sites/default/files/sessions/working_documents/afr_rc60_5_version2.pdf (Accessed August 9, 2023).

Xafis, V., Schaefer, G. O., Labude, M. K., Brassington, I., Ballantyne, A., Lim, H. Y., et al. (2019). An ethics framework for big data in health and research. *Asian Bioeth. Rev.* 11, 227–254. doi:10.1007/s41649-019-00099-x

Zeija, F. (2018). Consumer protection in Uganda: the law in theory and practice. *J. Consumer Policy* 41, 455–471. doi:10.1007/s10603-018-9382-3

Zimmermann, A., Di Rosa, E., and Kim, H. (2020). Technology can't fix algorithmic injustice. Available from: http://bostonreview.net/science-nature-politics/annette-zimmermann-elena-di-rosahochan-kim-techn ology-cant-fix-algorithmic.

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 02 October 2023
DOI 10.3389/fgene.2023.1248747

OPEN ACCESS

EDITED BY
Yann Joly,
McGill University, Canada

REVIEWED BY
Garbiñe Saruwatari,
National Institute of Genomic Medicine (INMEGEN), Mexico
Julian Kinderlerer,
University of Cape Town, South Africa
Lijun Shang,
London Metropolitan University, United Kingdom

*CORRESPONDENCE
Donrich Thaldar,
ThaldarD@ukzn.ac.za

RECEIVED 27 June 2023
ACCEPTED 06 September 2023
PUBLISHED 02 October 2023

CITATION
Thaldar D, Gooden A and Steytler M (2023), Open science and human genetic data: recommendations on South Africa's *Draft National Open Science Policy*. *Front. Genet.* 14:1248747. doi: 10.3389/fgene.2023.1248747

Open science and human genetic data: recommendations on South Africa's *Draft National Open Science Policy*

Donrich Thaldar[1,2]*, Amy Gooden[1] and Michaela Steytler[1]

[1]School of Law, University of KwaZulu-Natal, Durban, South Africa, [2]Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics, Harvard Law School, Cambridge, MA, United States

The *Draft National Open Science Policy*, which was shared by the South African government with stakeholders in 2022, is an encouraging step forward as it aims to promote the practice of open science in South Africa through a system of incentives. Since South Africa is constitutionally committed to be an open and democratic society, this approach is preferable to the approach of state control that characterizes the *Draft National Policy on Data and Cloud*—another data-related policy initiative by the South African government. However, there is room for improvement in the *Draft National Open Science Policy*. In particular, it should: (a) rely on the right to freedom of scientific research to strengthen the policy; (b) rectify the omission of ownership from its policy analysis; and (c) retain a clear differentiation between human and non-human genetic data. This will ensure that the final policy is clearly anchored in the South African Constitution, and that the principle of "as open as possible, as closed as necessary" can be applied to human genetic data in a legally well informed and accountable way.

KEYWORDS

***Draft National Open Science Policy*, freedom of scientific research, human genetic data, open science, ownership, South Africa**

1 Introduction

Science is "an indispensable contribution to the human endeavour" (International Science Council, 2020). Science is necessary to advance society, stimulate innovation, enhance education, develop policies, and protect well-being. But science is most successful when knowledge is freely available (International Science Council, 2020). There is growing concern that science has become too secluded to benefit the common good of society (International Science Council, 2020). Therefore, the philosophical concept of *open science* has emerged as an endeavor to close the science–society gap by democratizing scientific knowledge (Britt Holbrook, 2019). Open science aims to empower all to partake in science, aided by the Internet, which allows broad dissemination of knowledge (Bahlai et al., 2019; Heise and Pearce, 2020; Hanwell, 2022). Open science is a "no-barrier approach to scientific research" (Steger and Hantho, 2019) that is based on the principle of the free sharing of scientific knowledge. This entails taking down the barriers, such as article paywalls, that "chronically impede scientific progress" (Crow and Tananbaum, 2020). South Africa's *Draft National Open Science Policy* (Department of Science and Innovation, 2022) fits well within this philosophical framework.

In this article, we analyze the *Draft National Open Science Policy*, with a focus on a particular kind of scientific knowledge: *human genetic data*. Human genetic data are defined

in the *International Declaration on Human Genetic Data* (the *Declaration*) as "information about heritable characteristics of individuals obtained by analysis of nucleic acids or by other scientific analysis" (United Nations Educational Scientific and Cultural Organisation, 2003). As highlighted by the *Declaration*, human genetic data has a special status, because such data (a) can be predictive of genetic predispositions concerning individuals; (b) may have a significant impact on the family, including offspring, extending over generations, and in some instances on the whole group to which the person concerned belongs; (c) may contain information, the significance of which is not necessarily known at the time of collection of the biological samples; and (d) may have cultural significance for persons or groups. Accordingly, the *Declaration* calls for an appropriate level of protection of human genetic data that recognizes the sensitive nature of such data. At the same time, the *Declaration* also recognizes that the use of human genetic data is of "paramount importance" for the progress of life sciences and medicine (United Nations Educational Scientific and Cultural Organisation, 2003).

We identify three aspects of the *Draft National Open Science Policy* that require reconsideration: First, we suggest that the *Draft National Open Science Policy* should have a clear anchoring in the South African Constitution, and that this can best be accomplished by building a conceptual nexus between open science and the right to freedom of scientific research. Second, we suggest that the issue of ownership of human genetic data is important and consequential, and that its omission from the analysis in the *Draft National Open Science Policy* should be rectified. Third, we suggest that a clear differentiation between human and non-human genetic data is justified, and should be retained.

We also highlight a number of positive aspects in the *Draft National Open Science Policy* that are accentuated when it is compared to the *Draft National Policy on Data and Cloud* (Department of Communications and Digital Technologies, 2021)—another draft policy that is relevant to human genetic data, and which was released by the South African government in 2021.

2 Analysis

2.1 Rely on the right to freedom of scientific research to strengthen the policy

The vision espoused in the *Draft National Open Science Policy* is constituted by the following elements: (a) equality of opportunity; (b) environmental sustainability; (c) democratization of knowledge; (d) inclusive socio-economic development; and (e) scientific research. Given that the South African Constitution is the supreme law of South Africa, and the South African state has a duty to respect, protect, promote, and fulfil the rights in the Bill of Rights, it would significantly strengthen the *Draft National Open Science Policy* if its vision were explicitly grounded in the South African Constitution—in particular its Bill of Rights. This can be achieved by linking the constituent elements of its vision to constitutionally enumerated rights. Such links may be apparent in some instances—such as equality of opportunity and the right to equality—given that the right to equality is well known and often referred to in policy discourse. However, such a link may be less apparent in the case of the element *scientific research*. Does scientific research have any link with the South African Constitution?

The answer is yes. Scientific research enjoys an explicit link with the South African Constitution in the form of the *right to freedom of scientific research* (contained in section 16(1)(*d*)). By invoking the right to freedom of scientific research and unpacking its meaning and purposes, the *Draft National Open Science Policy* can, going forward, provide a more solid basis for the relevance and importance of scientific research—and by extension open science—in South Africa's constitutional dispensation. The right to freedom of scientific research serves purposes that are at the core of our constitutional value system: promoting individual autonomy, facilitating the search for truth, and supporting democracy (Thaldar and Steytler, 2021). We briefly elaborate on each of these purposes.

2.1.1 Promoting individual autonomy

Freedom of scientific research enables individual scientists to find self-fulfillment in pursuing their calling (Case v Minister of Safety and Security, 1996; Steytler, 2021; Thaldar and Steytler, 2021). While this is in itself valuable (Member of the Executive Council for Education: KwaZulu-Natal v Pillay, 2008; British American Tobacco South Africa (Pty) Ltd v Minister of Health, 2012; Van Breda v Media 24 Ltd, 2017; Jordaan, 2009), it also has a powerful knock-on effect on society—and on the autonomy of individuals in society (Jordaan, 2007). Freedom of scientific research has historically been a catalyst for scientific progress; scientific progress, in turn, has played an important role in improving the human condition (Jordaan, 2007), and has "freed a significant portion of humanity from ignorance, poverty and disease" (Corbellini, 2007). An improved human condition broadens the horizons for individual actualization across society.

2.1.2 Facilitating the search for truth

Science has been described as "the search for truths about the natural world" (Lederberg, 1972). Freedom of scientific research facilitates this search for truth by enabling free research and experimentation, the dissemination of results, and the subjection of methodologies, datasets, and results to scrutiny by other scientists (Steytler, 2021). South Africa's Constitutional Court—the country's apex court—has expressed itself in favor of a free marketplace of ideas, based on unfettered supply of, and demand for, ideas (Case v Minister of Safety and Security, 1996).

2.1.3 Supporting democracy

The contemporary understanding of the concept "democracy" is more than just the casting of a vote in an election, and includes values such as transparency, accountability, and participation in public life (Thaldar and Steytler, 2021). Furthermore, South Africa's Constitutional Court has held that the need for *informed* decision-making has become integral to the contemporary understanding of democracy (South African Broadcasting Corporation Ltd v National Director of Public Prosecutions, 2007). This speaks directly to the importance of freedom of scientific research, as science—when practiced freely—seeks to generate reliable, evidence-based knowledge about the world. Such reliable, evidence-based knowledge about the world enables *informed* decision-making and therefore supports democracy.

Open science, by promoting the transparency and accessibility of knowledge, supports and bolsters the practice of freedom of scientific research. It should also be recognized that freedom of scientific research is a necessary condition for open science. Censorship of the dissemination of the results of scientific research would not only be an infringement of the right to freedom of scientific research, but would also inevitably undermine open science. Clearly, open science is intertwined with freedom of scientific research. We suggest that clearly linking and placing reliance on the constitutional right to freedom of scientific research would add significant legal gravitas to the policy initiative of promoting open science.

2.2 Include the issue of ownership in the policy analysis

While the *Draft National Open Science Policy* deals at length with intellectual property rights, it completely omits other kinds of property rights. It is important, for example, to consider not only the copyright in *datasets* of human genetic data, but also common law ownership of the human genetic *data* that make up such datasets. Note that the question of ownership of human genetic data is distinct from—and yet interacts with—data subjects' privacy rights in their genetic data and researchers' possible claims to intellectual property rights related to such genetic data. Stated differently, the legal nature of human genetic data is not one-dimensional, but multidimensional (Thaldar et al., 2022). These various legal dimensions interact with each other—one right can limit another in specific, defined ways (Thaldar et al., 2022). It would therefore be a serious mistake to conceptualize human genetic data in only one or two dimensions and ignore the other dimension(s), as this would render an incomplete, and likely incorrect, understanding of the rights applicable to human genetic data. Yet, this is unfortunately what the *Draft National Open Science Policy* does.

We are not alone in calling for policy engagement with the issue of human genetic data ownership. A 2018 report by the Academy of Science of South Africa (ASSAf) entitled *Human Genetics and Genomics in South Africa*: *Ethical, Legal and Social Implications* (ASSAf, 2018) (the ASSAf report) called for this topic to be "carefully and vigorously debated and clarified for the South African context" (ASSAf, 2018). However, the ASSAf report does proffer a substantive position of its own, namely, that the "custodianship" of human genetic data ought to be preferred, and "ownership" avoided (ASSAf, 2018). We suggest that any normative inquiry about the desirability of human genetic data ownership should be informed by *inter alia* existing common law property rights. This is important, not only because respect for existing rights is a well-established norm, but also because the existence of existing rights may pose significant practical legal challenges to policy options that threaten to encroach on existing rights. For example, if, hypothetically, private research company X is the *owner* of the genetic data of thousands of South Africans, a policy that proposes that all genetic data of South Africans ought to be made public property effectively proposes that the state ought to *expropriate* private research company X's property. This may require an excessive amount of state resources to accomplish, which raises the question of whether the policy objectives (such as greater accessibility of the genetic data) cannot be attained through different means (than making all genetic data of South Africans public property) (Kabata and Thaldar, 2023). However, apart from briefly referring to a "traditionally" held legal view regarding human biological *samples*, the ASSAf report does not present legal analysis on whether human genetic *data* satisfy the criteria for ownership in South African law.

Such an analysis has since been embarked on by Thaldar et al. (2022), showing that a human genetic data instance—i.e., the computer file containing the sequence data—is indeed susceptible of private ownership in South African law. This conclusion is important in the context of developing an open science policy, as private ownership rights in human genetic data can be a powerful tool to either facilitate or hinder greater access to such data. The *Draft National Open Science Policy*'s principle of "as open as possible, as closed as necessary" can only be sensibly applied if there is clarity regarding the parameters of legal rights in human genetic data.

2.3 Retain a clear differentiation between human and non-human genetic data

The *Draft National Open Science Policy* refers to the *Nagoya Protocol to the Convention on Biological Diversity* (Secretariat of the Convention on Biological Diversity, 2011) (the *Nagoya Protocol*) and states that: (a) the *Nagoya Protocol* deals with access to, and benefit sharing of, genetic resources; and (b) the *Nagoya Protocol* has "gained interest with the idea of extension to other genomic data" (Secretariat of the Convention on Biological Diversity, 2011). First, we discuss the exact legal scope of the *Nagoya Protocol*, and second, we suggest that the idea of extending the scope of the *Nagoya Protocol* is controversial and can only serve to detract from the positive vision of the *Draft National Open Science Policy*.

The *Convention on Biological Diversity* (Secretariat of the Convention on Biological Diversity, 1992) (the *Convention*) provides that each state has sovereign rights over its genetic resources, meaning that each state can decide how its genetic resources will be governed, including how rights will be vested in its genetic resources. This principle is often referred to as "genetic sovereignty". But does this principle include *human* genetic data? The *Convention* defines "genetic resources" as "genetic material of actual or potential value" (Secretariat of the Convention on Biological Diversity, 1992); "genetic material", in turn, is defined as "any material of plant, animal, microbial or other origin containing functional units of heredity" (Secretariat of the Convention on Biological Diversity, 1992). Since the word "animal" can be interpreted as including humans, the Conference of the Parties to the *Convention* (1995) clarified that the *Convention* does *not* apply to *human* genetic material. When adopting the *Nagoya Protocol*, the Conference of the Parties to the *Convention* (2010) again recorded that the *Nagoya Protocol* does *not* apply to *human* genetic material. Accordingly, the principle of genetic sovereignty is limited to non-human genetic material.

But, would it not be a good idea to lobby for an expansion of the *Convention* and the *Nagoya Protocol* to also apply to *human* genetic data? After all, human individuals that belong to the same ethnic group share certain genetic similarities. As such, if individual

members of a certain ethnic group participate in a genetic research project, such individuals may be providing valuable genetic information, not only about themselves as individuals but about their entire ethnic group. Should such ethnic groups not be entitled to control access to—and benefit from—such genetic information? Moreover, from a national perspective, should the human genetic data of South Africans not be viewed as a natural resource, similar to water or gold, that should be managed by government as public property? This was indeed the position taken by the ASSAf report (ASSAf, 2018). However, we suggest that this position would be difficult to sustain in the South African context for the following legal and policy reasons.

2.3.1 Reason 1

While non-human biological resources, such as indigenous fynbos flowers or butterflies, cannot decide for themselves whether to provide their genetic material for research and on what conditions, humans can. Underlying the rights entrenched in the South African Bill of Rights is "the constitutional celebration of the possibility of morally autonomous human beings independently able to form opinions and act on them" (British American Tobacco South Africa (Pty) Ltd v Minister of Health, 2012). Provided that individual autonomy is protected through informed consent, how will the South African government (or the leadership of a community) justify restricting individuals' autonomy to donate their genetic data to research projects that they themselves deem worthy?

2.3.2 Reason 2

Building further on this theme, it is important to note that non-human genetic data are *impersonal* in nature, while human genetic data are *personal* in nature (Shabani and Borry, 2018; Thaldar et al., 2019; Costello, 2022). This is an additional reason why the comparison of human genetic data with natural resources is misleading. While there are no personality rights in an indigenous fynbos flower or a butterfly—or in public property such as water or gold—persons have *personality rights* in their own genetic data. Personality rights are inseparably bound up with one's personality, cannot exist independently of the human personality, and are incapable of being transferred (Kumalo v Cycle Lab (Pty) Ltd, 2011). Examples are the right to the integrity of a person, to respect a person's name and reputation, the right to informational privacy generally (as codified in Protection of Personal Information Act 4 of 2013), and the right to control the use of one's image. Accordingly, we suggest that the following is a more appropriate comparison: If persons belonging to, for example, ethnic group X, which is indigenous to South Africa, act as models in a commercial advertisement and are paid handsomely, should all persons identifying as belonging to ethnic group X share in benefits from the use of those individuals' images? Moreover, would the South African government (or the leadership of ethnic group X) be justified in exercising control over the images of the individual persons who are members of ethnic group X and who voluntarily decided to participate in making the advertisement? When considering these questions, bear the following in mind: Model Y may look very similar to her biological sister, but this fact does not give Y's sister any rights over the use of Y's image. Similarly, patient Z who suffers from a heritable condition is at liberty to disclose the nature of her illness to whomever she pleases, despite the fact that such information will imply a certain genetic propensity towards the same heritable condition among her family members. There may be moral and cultural considerations applicable to Z's decision, but legally she is perfectly entitled to disclose the nature of her illness to whomever she wishes.

2.3.3 Reason 3

Health law in South Africa is based on the principle of altruism in research participation (Jordaan, 2016; Thaldar et al., 2021). The National Health Act 61 of 2003 provides (in section 60(4)) that research participants who donate tissue or blood samples may only be compensated for reasonable expenses incurred, and (in section 60(5)) that it is a criminal offence to offer such research participants financial or other reward (apart from reasonable expenses incurred) for their donation. Given the current state of genetic technology, genetic sequence data cannot be obtained directly from a research participant, but must be obtained from a human biological material sample. Accordingly, donating a human biological material sample is a *conditio sine qua non* for genetic research. This clearly restricts the kinds of benefit sharing that research participants in genetics research projects may lawfully receive in South Africa, as any type of benefit sharing that constitutes a "financial or other reward" for the research participant would be unlawful (and criminal) (Thaldar and Shozi, 2023). Importantly, the *Nagoya Protocol* is not self-executory, meaning that it only gains effect in South African law if, and to the extent that, it is incorporated into South African statute law. Accordingly, a hypothetical amendment to the *Nagoya Protocol* (the "idea of extension") to include *human* genetic data would, on its own, not affect the legal reality in South Africa. Note that such an amendment is pure conjecture, as there is no indication that any party to the *Nagoya Protocol* intends to propose such an amendment.

2.3.4 Reason 4

The final reason is the most fundamental in the present context. The idea of genetic sovereignty in the human context—where the state or a community exercises sovereign power over human genetic data—is philosophically opposed to open science. Genetic sovereignty, to have any meaning, will entail access barriers in the form of the state or a community deciding who can access human genetic data, and on what conditions. By contrast, open science entails access to research results—including human genetic data—*free of access barriers* (Steger and Hantho, 2019; Crow and Tananbaum, 2020). As recently argued by Kabata and Thaldar (2023), the idea of state sovereignty over human genomic (or genetic) data may seem superficially attractive, but has no actual utility to African states. Instead, the authors suggest a human-rights-based approach to the governance of human genetic data that focuses on everyone's *right to science*, which is aligned with promoting open science (Kabata and Thaldar, 2023).

For these reasons, we suggest that the *Draft National Open Science Policy* should either remove the sentence about "the idea of extension" of the *Nagoya Protocol* to "other genomic data", or add a clear disclaimer that such an idea is contrary to current South African law and contrary to the objective of the *Draft National Open Science Policy* to promote open science.

TABLE 1 Differences between the *Draft National Open Science Policy* and the *Draft National Policy on Data and Cloud*.

Draft National Open Science Policy	*Draft National Policy on Data and Cloud*
Developed by the Department of Science and Innovation	Developed by the Department of Communications and Digital Technologies
Aims to facilitate free access to data through incremental, incentivized moves towards open science	Aims to facilitate free access to data through the nationalization of all data generated in South Africa
Operates within the existing intellectual property framework	Suggests the disruption of the intellectual property legal framework
Respects the right to private property	Suggests the disruption of private ownership of data

2.4 A brief comparison with the *Draft National Policy on Data and Cloud*

We now turn to the positive aspects of the *Draft National Open Science Policy* that we identified in the introduction above. Our analysis takes the form of a comparison of the *Draft National Open Science Policy* with another draft policy published for public comment by the South African government, namely, the *Draft National Policy on Data and Cloud* (Department of Communications and Digital Technologies, 2021) (see Table 1 below). It is relevant to note that these two draft policies were produced by two different government departments: the former by the Department of Science and Innovation; the latter by the Department of Communications and Digital Technologies. Many of the policy objectives of the *Draft National Policy on Data and Cloud* deserve support. These include: (a) encouraging universal access to broadband connectivity; (b) eliminating regulatory barriers and enabling competition in the data and cloud sector; (c) supporting the development of small, medium, and micro enterprises; and (d) promoting research, innovation, and technological developments in relation to the cloud. However, the *Draft National Policy on Data and Cloud* is premised on the ideological position that (drastically) greater state control of data is the best solution. For the reasons stated below, we are critical of this position.

While there is an important overlap between the objectives of the *Draft National Open Science Policy* and the *Draft National Policy on Data and Cloud*, as both draft policies aim to facilitate South Africans' free access to data, the two policies propose to accomplish this objective through radically different means: the *Draft National Open Science Policy* through incremental, incentivized moves towards open science, within the existing intellectual property legal framework and the right to private property; the *Draft National Policy on Data and Cloud*, on the other hand, through the nationalization of all data generated in South Africa, the disruption of the intellectual property legal framework, and government control of access to data.

At a level of principle, these two approaches are clearly ideologically incompatible. Which of these two approaches would be better aligned with the values of the open and democratic society that South Africa aspires to be? Without doubt, the *Draft National Open Science Policy*. When examining the concept of an "open society" in the South African Constitution, Justice Ackermann (in Ferreira v Levin, 1996) relied on Karl Popper's *magnum opus*, *The Open Society and Its Enemies* (Popper, 1945). This judgment points to the political philosophical source of the concept of an "open society", and opens the door to learn from this source (Jordaan, 2017). In *The Open Society and Its Enemies,* Popper (1945) famously proposed that policy-making (or social engineering) in an open society should be "piecemeal", rather than "utopian". Piecemeal social engineering denotes small-scale social "experiments" that can be modified or reversed based on results in the social "laboratory". Utopian social engineering, on the other hand, denotes large-scale policy interventions that seek to modify human behavior to conform to policy ideas at any cost—i.e., without error-correcting mechanisms along the way (Popper, 1945). The *Draft National Policy on Data and Cloud*'s proposal to nationalize all data generated in South Africa is not only far-reaching, but contains no intermediate steps to provide for learning and adapting based on real-world effects. As such, it leans dangerously towards the kind of utopian social engineering that Popper warned against. Fortunately, the *Draft National Open Science Policy* does not fall into this trap.

Also at a practical level, the *Draft National Open Science Policy* offers a more realistic pathway to reaching the objective of facilitating South Africans' free access to data. It respects the existing legal frameworks and individual rights, while envisioning a cultural change towards open science that will be championed by an official body, the Open Science Advisory Council, and incentivized along the way. By contrast, the *Draft National Policy on Data and Cloud* proposes to break down established legal frameworks and rights, and impose government control of data. It is human nature to respond positively to incentives, and to respond negatively to infractions on one's reasonable expectations, such as the expectation that one's individual rights will be respected. Furthermore, obvious questions can be raised regarding the efficiency of government control of data. Accordingly, if South Africans' free access to data is the public policy objective, the *Draft National Open Science Policy* offers a much more attractive policy pathway to accomplish this objective.

3 Conclusion: towards open science in South Africa

The *Draft National Open Science Policy* is a milestone in South Africa's journey towards a workable and effective national policy that promotes open science at all levels of scientific endeavor. However, there is room for improvement in the three areas that we have highlighted. In our view, the final national open science policy should seriously engage with constitutional law (the right to freedom of scientific research), property law (ownership), and international law (the *Nagoya Protocol*) aspects of human genetic data *qua* research

output. This will ensure that the final policy is clearly anchored in the South African Constitution, and that the principle of "as open as possible, as closed as necessary" can be applied to human genetic data in a legally well informed and accountable way.

Note that the *Draft National Open Science Policy* was not made public by the South African Department of Science and Innovation. Instead, it was only disseminated via email to "stakeholders" within the South African academic community, who were given the opportunity to submit comments. It is anticipated that a subsequent version will, at some future stage, be published for public comment. To assist the reader, we include a summary of the *Draft National Open Science Policy* as Supplementary Material S1.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

Work on this article was supported by the U.S. National Institute of Mental Health and the U.S. National Institutes of Health (award number U01MH127690) under the Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) program. The content of this article is solely our responsibility and does not necessarily represent the official views of the U.S. National Institute of Mental Health or the U.S. National Institutes of Health.

Acknowledgments

The first version of the summary of the *Draft National Open Science Policy* (Supplementary Material S1) was generated by ChatGPT-4.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Author disclaimer

The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the U.S. National Institute of Mental Health or the U.S. National Institutes of Health.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1248747/full#supplementary-material

References

Academy of Science of South Africa (ASSAf), (2018). *Human genetics and genomics in South Africa: Ethical, legal and social implications*. doi:10.17159/assaf.2018/0033

Bahlai, C., Bartlett, L. J., Burgio, K. R., Fournier, A. M. V., Keiser, C. N., Poisot, T., et al. (2019). "Open science isn't always open to all scientists," in *American Scientist*. Available at: https://www.americanscientist.org/article/open-science-isnt-always-open-to-all-scientists (Accessed February 27, 2022).

British American Tobacco South Africa (Pty) Ltd v Minister of Health (2012). 3 All SA 593 (SCA). Available at: http://www.saflii.org/za/cases/ZASCA/2012/107.html.

Britt Holbrook, J. (2019). "Philosopher's corner: open science, open access, and the democratization of knowledge," in *Issues in science and technology*. Available at: https://issues.org/philosophers-corner-open-science-open-access-and-the-democratization-of-knowledge/ (Accessed February 26, 2022).

Case v Minister of Safety and Security (1996). (3) SA 617 (CC). Available at: http://www.saflii.org/za/cases/ZACC/1996/7.html.

Conference of the Parties to the Convention on Biological Diversity, (2010). *Decision adopted by the conference of the Parties to the Convention on Biological Diversity at its tenth meeting*. Available at: https://www.cbd.int/decision/cop/?id=12267.

Conference of the Parties to the Convention on Biological Diversity, (1995). *Second ordinary meeting of the conference of the Parties to the Convention on Biological Diversity decision II/11*. Available at: https://www.cbd.int/decision/cop/?id=7084.

Corbellini, G. (2007). Scientists, bioethics and democracy: the italian case and its meanings. *J. Med. Ethics* 33, 349–352. doi:10.1136/jme.2007.020586

Costello, Róisín Á. (2022). Genetic data and the right to privacy: towards a relational theory of privacy? *Hum. Rights L. Rev.* 22, 1–23. doi:10.1093/hrlr/ngab031

Crow, M. M., and Tananbaum, G. (2020). *We must tear down the barriers that impede scientific progress*. New York: Scientific American. Available at: https://www.scientificamerican.com/article/we-must-tear-down-the-barriers-that-impede-scientific-progress/ (Accessed February 27, 2022).

Department of Communications and Digital Technologies, (2021). *Draft national policy on data and cloud*. GN 306 GG 44389. Available at: https://www.gov.za/sites/default/files/gcis_document/202104/44389gon206.pdf.

Department of Science and Innovation, (2022). *Draft national open science policy*.

Ferreira v Levin (1996) (1) SA 984 (CC). Available at: http://www.saflii.org/za/cases/ZACC/1995/13.html.

Hanwell, M. (2022). "What is open science?," in *Open Source*. Available at: https://opensource.com/resources/open-science#:~:text=Open%20science%20arguably%20began%20in,such%20as%20the%20Royal%20Society (Accessed February 26, 2022).

Heise, C., and Pearce, J. M. (2020). *From open access to open science: the path from scientific reality to open scientific communication*. Thousand Oaks, California: SAGE Open, 1–14. doi:10.1177/2158244020915900

International Science Council, (2020). *Open science for the 21st century (draft ISC working paper 3)*. Available at: https://council.science/wp-content/uploads/2020/06/International-Science-Council_Open-Science-for-the-21st-Century_Working-Paper-2020_compressed.pdf.

Jordaan, D. W. (2009). Autonomy as an element of human dignity in South African case law. *J. Philosophy, Sci. Law* 9 (3), 1–15. doi:10.5840/jpsl2009933

Jordaan, D. W. (2007). Science versus anti-science: the law on pre-embryo experimentation. *S. Afr. Law J.* 124 (3), 618–634.

Jordaan, D. W. (2016). Social justice and research using human biological material: a response to Mahomed, Nöthling-Slabbert and Pepper. *S. Afr. Med. J.* 106, 678–680. doi:10.7196/SAMJ.2016.v106i7.10552

Jordaan, D. W. (2017). The open society: what does it really mean? *De. Jure* 50, 396–405. doi:10.17159/2225-7160/2017/v50n2a11

Kabata, F., and Thaldar, D. (2023). Regulating human genomic research in Africa: why a human rights approach is a more promising conceptual framework than genomic sovereignty. *Front. Genet.* 14, 1208606. doi:10.3389/fgene.2023.1208606

Kumalo v Cycle Lab (Pty) Ltd (2011). ZAGPJHC 56. Available at: http://www.saflii.org/za/cases/ZAGPJHC/2011/56.html.

Lederberg, J. (1972). The freedoms and the control of science: notes from the ivory tower. *South. Calif. L. Rev.* 45, 596–615.

Member of the Executive Council for Education: KwaZulu-Natal v Pillay (2008). (1) SA 474 (CC). Available at: https://www.saflii.org/za/cases/ZACC/2007/21.html.

National Health Act 61 of 2003 (2003). [South Africa].

Popper, K. (1945). *The open society and its Enemies*. London: Routledge.

Protection of Personal Information Act 4 of 2013 (2013). [South Africa].

Secretariat of the Convention on Biological Diversity, (1992). *Convention on biological diversity*. Available at: https://www.cbd.int/doc/legal/cbd-en.pdf.

Secretariat of the Convention on Biological Diversity, (2011). *Nagoya Protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilization to the Convention on Biological Diversity*. Available at: https://www.cbd.int/abs/doc/protocol/nagoya-protocol-en.pdf.

Shabani, M., and Borry, P. (2018). Rules for processing genetic data for research purposes in view of the new EU General Data Protection Regulation. *Eur. J. Hum. Genet.* 26, 149–156. doi:10.1038/s41431-017-0045-7

South African Broadcasting Corporation Ltd v National Director of Public Prosecutions (2007). (1) SA 523 (CC). Available at: http://www.saflii.org/za/cases/ZACC/2006/15.html.

Steger, A., and Hantho, E. (2019). "Open science: removing barriers to knowledge," in *Health Law Matters*. Available at: https://healthlawmatters.squarespace.com/home/2019/4/27/open-science-removing-barriers-to-knowledge (Accessed February 27, 2022).

Steytler, M. (2021). *The right to freedom of scientific research in the age of gene editing. [dissertation]. [Durban (KZN)]*. Durban, South Africa: University of KwaZulu-Natal. Available at: https://researchspace.ukzn.ac.za/handle/10413/20219.

Thaldar, D., Kinderlerer, J., and Soni, S. (2019). An optimistic vision for biosciences in South Africa: a response to the ASSAf report on human genetics and genomics. *S. Afr. J. Sci.* 115 (7/8), 1. doi:10.17159/sajs.2019/6146

Thaldar, D. W., and Shozi, B. (2023). Is benefit sharing with research participants lawful in South Africa? An unexplored question in the governance of genomics research. *J. Law Biosci.* 10 (1), lsad018. doi:10.1093/jlb/lsad018

Thaldar, D. W., and Steytler, M. (2021). Time for cinderella to go to the ball: reflections on the right to freedom of scientific research. *S. Afr. Law J.* 138 (2), 260–288. doi:10.47348/SALJ/v138/i2a2

Thaldar, D. W., Townsend, B. A., Botes, M., and Shozi, B. (2021). "Human biological material," in *Law of South Africa*. Editor Joubert (Durban: LexisNexis).

Thaldar, D. W., Townsend, B. A., Donnelly, D-L., Botes, M., Gooden, A., van Harmelen, J., et al. (2022). The multidimensional legal nature of personal genomic sequence data: a South African perspective. *Front. Genet.* 13, 997595–997611. doi:10.3389/fgene.2022.997595

United Nations Educational Scientific and Cultural Organisation (2003). *UNESCO International Declaration on Human Genetic Data*. Paris, France.

Van Breda v Media 24 Ltd (2017); National Director of Public Prosecutions v Media 24 Ltd 2017 435 (2) SACR 491 (SCA). Available at: http://www.saflii.org/za/cases/ZASCA/2017/97.html.

frontiers | Frontiers in Psychology

TYPE Hypothesis and Theory
PUBLISHED 17 October 2023
DOI 10.3389/fpsyg.2023.1273470

OPEN ACCESS

EDITED BY
Simisola Oluwatoyin Akintola,
University of Ibadan, Nigeria

REVIEWED BY
Nathalie Gontier,
University of Lisbon, Portugal
Opeyemi A. Gbadegesin,
University of Ibadan, Nigeria

*CORRESPONDENCE
Meshandren Naidoo
✉ 214549331@stu.ukzn.ac.za

RECEIVED 08 August 2023
ACCEPTED 26 September 2023
PUBLISHED 17 October 2023

CITATION
Naidoo M (2023) What does it mean to be an agent? *Front. Psychol.* 14:1273470. doi: 10.3389/fpsyg.2023.1273470

What does it mean to be an agent?

Meshandren Naidoo*

School of Law, University of KwaZulu-Natal, Durban, South Africa

Artificial intelligence (AI) has posed numerous legal–ethical challenges. These challenges are particularly acute when dealing with AI demonstrating substantial computational prowess, which is then correlated with agency or autonomy. A common response to considering this issue is to inquire whether an AI system is "conscious" or not. If it is, then it could constitute an agent, actor, or person. This framing is, however, unhelpful since there are many unresolved questions about consciousness. Instead, a practical approach is proposed, which could be used to better regulate new AI technologies. The value of the practical approach in this study is that it (1) provides an empirically observable, testable framework that contains predictive value; (2) is derived from a data-science framework that uses semantic information as a marker; (3) relies on a self-referential logic which is fundamental to agency; (4) enables the "grading" or "ranking" of AI systems, which provides an alternative method (as opposed to current risk-tiering approaches) and measure to determine the suitability of an AI system within a specific domain (e.g., such as social domains or emotional domains); (5) presents consistent, coherent, and higher informational content as opposed to other approaches; (6) fits within the conception of what informational content "laws" are to contain and maintain; and (7) presents a viable methodology to obtain "agency", "agent", and "personhood", which is robust to current and future developments in AI technologies and society.

KEYWORDS

agency, artificial intelligence, autonomy, explanations, personhood, semantics, complex system, mechanics and dynamics

1. Introduction and limitations

This paper aimed to establish a robust account of agency which can be applied to many kinds of systems, including AI systems. This raises further sub-questions, such as (1) what does it mean to be an agent; and (2) what markers are there to determine an agent? An account of the agency must provide answers to those questions in a generally determinable manner. To build an explanatory account of agency, this study evaluates and uses the various logic underpinning "explanations" using the ecological framing of biological organisms as agents of their own evolution. In this light, information-centric quantification tools such as statistical mechanics and bioinformatics would be attractive sources for creating such an account. An important question would be "what is an AI system?" This question is beyond the scope of this article but will be examined in future research. An additional limit is that this methodology describes an empirically testable account of agency, but it will not describe in detail its preferability compared to existing approaches. It is assumed that the reader is familiar with existing approaches.

Evolution is an ecological phenomenon arising from the purposive engagement of organisms with their conditions of existence. It is incorrect to separate evolutionary biology into processes of inheritance, development, selection, and mutation. Instead, the component

processes of evolution are jointly caused by the organismal agency and their ecological relations with their affordances. Purposive action is understood to be agents that use features in their environments as affordances that are conducive to their goals. Furthermore, a Kantian approach (see Part B of Supplementary material) is used. It focuses on accounts of agency and personhood as being the intrinsic purposiveness of the agent/person. A Kantian approach is preferable since it is the common framing for many legal constitutions and is a dominant framing mechanism for questions of this kind. Thus, this research moves away from the erroneous "intentional" approach (Sapolsky, 2017).

2. The nature of explanations and understanding

2.1. Theories and mental models

Explanations usually contain more than theories, in that they involve different bodies of knowledge (Keil, 2006). Explanations create trajectories and lead to understanding among people. They also tend to be more robust than theories. Explanations differ from mental models, which rather speak to formal representations of logical patterns to image-like representations of the works of systems. Mental models are often understood in spatial terms and explanations are not the same as mental modeling. Explanations involve interpretations. The value of explanations in growing knowledge lies in their transactional status and their interpretation (Keil, 2006). Related to this is the question "what does it mean to understand?" When people are probed about their beliefs about the world, coherence often evaporates. Often only fragments about the workings of systems are known, and of these known fragments very few are coherent (Keil, 2006). People's beliefs also tend to contradict one another. They are ignored only until the time when they are made explicit or are pointed out by someone else. This may be because of the limits of working memory. Therefore, not all elements can be considered together at the same time, which would help identify inconsistencies.

2.2. Synchronicity and the nature of oscillation

The question that has plagued humans for a long time is how do we come to an agreement on anything? In language, how do we agree on the meaning of words? In behavioral sciences, it is asked how do we know behaviors? In physics, we ask how entangled particles know what the others are doing? Weiner (1948) published *Cybernetics: Or Control and Communication in the Animal and the Machine*. He discussed the problems of communication and control in systems. He used the example of crickets and how they can synchronize their behavior so that their chirps can follow the progression that it does.

The answer is in oscillations or spin; we can observe this in neurons and non-living things such as pendulums synchronizing with each other, which Christiaan Huygens wrote about in 1665 (Redish, 2019). Mathematics then captured the essence of synchronization. There are populations/families of oscillators. Oscillators are things that repeat themselves. A pendulum, for example, is a mechanical oscillator, and a neuron firing in the brain is a cellular oscillator. Birds moving in unison, flying together, are animal oscillators.

2.2.1. Coupling

What is next needed is a coupling mechanism between individuals in a population. Coupling (Stankovski et al., 2015) depends on the population of concern. For neurons, it is the connections between each of them. For animals, it is sight or sound. For particles, it is spin. You can then capture frequency/pulse. There are also weak and strong couplings. Strong couplings mean that there is a stronger statistical tendency for the oscillation relationship/synchronization to take place. For coupling to take place, either strong or weak, there must be a relatively similar innate frequency between individuals, and they must be local (generally). Many different interlinking oscillators apply to humans and other creatures. The Yoshiko Kuramoto mathematical model (Strogatz, 2000) can explain complicated behaviors in complex systems, including perhaps even semantic information. Oscillation and coupling are key components of understanding (perhaps *the key* components). These components explain not only understanding but also relationality and non-verbal/verbal social communication.

2.2.2. The brain

Robert Moore, Victor Eichler, Frederich Stephan, and Irving Zucker discovered the brain regions responsible for governing circadian rhythms. The key structure is the suprachiasmatic nucleus (SCN), which processes information about light and darkness from the retinas. Damaged SCNs impair animal rhythms. Oscillators are the tools used to interlink and relate to others like us. They define what constitutes an "us". Examples of coupling mechanisms include things such as heat, shape, direction, and vision (eyes, in particular, are a gateway for bonding) (Cornell University, 2022). Previously, the postulation was that mirror neurons enabled us to mimic the behaviors of others in our social group and thus coordinate social or group learning; however, this has not been confirmed (Dickerson et al., 2017). Oscillators and coupling are the modalities of world-building and social organization or communication.

More generally, there are other instances of "understanding" or knowing. These instances involve embodied ontogenetic knowledge: of time, place, circumstance, culture, bodily knowledge (such as sensory information), and the like. For John Vervaeke, this is the four modalities of knowing: (1) participatory knowing; (2) perspectival knowing; (3) procedural knowing; and (4) propositional knowing (Raninen, 2023). Therefore, notions such as "understanding" or "knowing" are not related to thought or mental representations but rather to natural and mechanical processes of relation. This enables a reframing of these concepts such that they do not need to be intimately linked to purely human mental representations.

2.3. Patterns, stances, domains, and social/emotion

We can distinguish different explanations by the causal patterns they employ, the stances they invoke, the domains of phenomena they explain, or whether they are value- or emotion-laden (Keil, 2006). Each of these has different trajectories and properties.

2.3.1. Causal patterns

The most common causal relations to which explanations refer are (1) common cause, (2) common effect, (3) linear causal chains, and (4) causal homeostasis (Keil, 2006). Common cause explanations cite a single cause as having a branching set of consequences. These are usually diagnosis-type explanations (such as a bacterial infection that causes many symptoms or a computer virus). Common effect refers to instances where causes converge to create an event. These are common in historical narratives where several causes are attributed to converge and create an event. Linear chains, on the other hand, are degenerate cases of common cause and effect. With these, there is a unique serial chain from a single initial cause to a series of steps through a single effect (Keil, 2006). Causal homeostatic explanations are fundamental to natural kinds of explanations. These explain why sets of things endure as stable sets of properties. This type does not explain how a cause progresses over time to create effect(s), but rather how an interlocking set of causes and effects results in a set of properties that endure in combination over time as a stable set. This stable set is then of a natural kind. Some explanations are easier to follow, while others are more difficult and hence "unnatural". Furthermore, some explanations are often understood to be domain-based, although this is not necessarily the case (Keil, 2006).

2.3.2. Stances

One can frame explanations in terms of the stance that they take. Dan Dennett is known for drawing this distinction. Each stance speaks to a framing device for explanations. Each stance is general and non-predictive but does speak to certain relations, properties, and arguments that are fundamental to each (Keil, 2006). Dennett highlighted three different kinds of stances: (1) mechanical, (2) design, and (3) intentional. Mechanical stances consider only simple physical objects and their interactions. The design stance considers entities as having purposes and functions that occur beyond mechanical interactions. Some argue that teleology/functional explanation is part of this stance. There are also questions about whether an intentional designer is necessary for teleological explanations. The intentional stance sees entities as having beliefs, desires, and other mental contents/representations that govern their behaviors (Keil, 2006). These mental states then have causal consequences in terms of behavior. This has, however, often been criticized for being based on folk psychology (Woolman, 2013). Each stance describes different insights and distortions and explains different things. They need not exclude each other and can be complementary (see part G of the Supplementary material for more information on intentionality).

2.4. Causation

Causal explanations have been the most dominant explanation, especially in the sciences. However, these are not the only forms of explanations; there are also non-casual explanations, which are called constitutive explanations (Salmon, 1984).

2.4.1. Causal capacities as explanada (etiological)

The object of constitutive explanation is the causal capacity of a system. This capacity describes what a system would do under specified circumstances/conditions (under a certain trigger). Causal capacities speak to what would, or could, or will happen under certain conditions and it includes notions such as ability, power, propensity, and tendency. Causal capacities speak to processes and events: when process (X) happens, event (Y) happens. These explain the changes in properties of a system—that is what an event is (Ylikoski, 2013). They focus on the origin, persistence, and changes in properties of (or in) a system.

2.4.2. Counterfactuals and the Millian method of difference

This is the "Millian Method of Difference" (Encyclopedia Britannica, 2023) or counterfactual approach. Counterfactual explanations (Mertes et al., 2022) are the "knockout" kinds (the gene as the unit of inheritance was established through this approach). Here, if you want to determine whether something (C) as a cause has an effect (E), you perform an experiment whereby you remove (C) and then observe the effects. This can be a literal removal or a conceptual removal. This is often used to explain why something happened, such as a decision, event, or outcome by reference to a particular thing or sequence.

You can also change the values of (C) by making it stronger or weaker, and then observe what happens to (E). We use this to make inferences from the difference observed in effects where (C) is absent or different. Thus, we infer the causal role of (C) based on its presence versus its absence or its changes. This is effective for identifying discrete explanatory privileged causes (Walsh, 2015) (see Part A of Supplementary material for an example and information on its undesirability).

2.4.3. Causation in complex systems

Complex adaptive systems can maintain stable configurations despite perturbations because they can alter the causal relations that happen between their parts. Each part affects, and is affected by, others, and the overall effect is attributable, jointly and severally, to all the parts. The system is thus affected by itself, and these causes are non-separable. Causes are only separable when the effect of a change in one is independent of the effects of changes in others. *If we remove or interfere with one we would also be interfering with others.* Therefore, causal composition/decomposition fails on non-separability—the influence/control factor of each part is non-determinable (thus non-quantifiable), and we cannot attribute differences in effect to specific differences in the causal contributions of the parts. One cannot assume when reviewing a result that the other factors are functioning as they were before

the removal of a factor—they can be operating differently. Thus, we cannot decompose causes and differences in effect by reference to external versus internal influences. Changes in the dynamics of complex adaptive systems can be initiated endogenously through internal perturbations or exogenously through changes in the environment. The system mounts a response to both, and the result of that response is attributable to both internal and external influences as a single cause. Feedback is where the internal dynamics and environment both cause a change in the behavior of a system with signals. Thus, the environment *is part* of the system's dynamic structure. This is why it is difficult or impossible to attribute liability (either for an action or for a composition of a product or artwork) to either an AI system or a human, whereby there is a "commingling" between both. Even distinguishing between "principle causes" and "initiating causes" does not offer an adequate solution.

Complex adaptive systems tend to distinguish between "principal causes" and "initiating causes". Principal causes are those to which we can attribute a large portion of the observable effect. Initiating causes starts the causal process, which ends with an effect. If two identical systems diverge in their outcomes, it is reasonable to afford principal causal responsibility for differences in effect to a factor that initiates the different trajectories (Walsh, 2015) (assuming that all other components contribute as before). In such a case, the principal causes would be initiating causes. However, the inference cannot hold for complex systems. There is logical discord between (1) the proposition that a change in the dynamics of complex systems is *initiated* by changes in exogenous conditions; and (2) the conclusion that *the principle cause* of the overall effect is *that change* in the exogenous condition. All this means is that the usual modes of inferences (cause and effect) do not work in complex dynamical systems.

2.5. Constitution

The constitution explains how things have the causal capacities that they do by relying on their parts and organizations (Ylikoski, 2013). Constitutive explanations ask: "what was it about (X) that resulted in it having disposition (Y)? What is it about (X) that enables a causation event to happen?" They provide different information compared to causative explanations. Fundamentally, these explanations provide modal information for causal possibilities.

To explicate constitutive explanations (Cummins, 1975, 1983, 2000; Craver, 2007a,b; Craver and Bechtel, 2007), their *explanada* must first be described. Constitutive explanations are not related to behavior, reactions, or activities of a system. These explain the properties of a system themselves. The *relata* of causative and constitutive explanations thus differ; causal explanations deal with events and constitutional explanations deal with properties. A constitutive explanation would say, for example, that system (S) has a causal capacity (C) in circumstances (E) because of its components (S1) and (S2) and their organization (O) (Ylikoski, 2013). Therefore, there is an ontological difference between causation and constitution. Both are relations of dependence (Rosen, 2010), but they are metaphysically different. Both, however, must account for explanatory relevance.

Metaphysics posits that the parts, their causal capacities, and their organization constitute the causal capacities of a system/whole. Constitution is synchronous and thus they are atemporal (meaning that it is not based on time and can be instantaneous). This means that if there are changes in the basis, there is an instant change in the causal capacities of a system (hence constitution is process and time-independent).

Importantly, the constitutive relata are not independent existences. In causation, one can insist that the relata of cause and effect are distinct from each other but one cannot insist on the same within constitution relata. Specific causal capacities are direct functions of certain constitutions. Constitutions then do not have independent identities.

Constitutive explanations distinguish themselves from identity, in that identity is a reflexive relation and is symmetric. First, one must distinguish between the constitution of all causal capacities of a system and the constitution of an individual capacity (Ylikoski, 2013). The former is the complete set of causal capacities of a particular system (at a time). We can identify the causal capacities and their causal basis (the organization). To have specific causal capacities, a specific causal basis (organization) is first necessary. Symmetries can be exact and help with allowing for simplicity in explanations, but it does not correlate to being the *identity* of a thing.

We cannot identify individual causal capacities with or as their composite bases (alternative constitution). This is because different objects can have the same causal capacities despite having different compositions (Ylikoski, 2013). This is known as multiple realization (MR). MR implies that we cannot equate a specific property of an object (like fragility) with the specific structural element of an object (molecular structure), but we can attribute a specific property of the object *because* of a specific structure that it has. At the heart of the scientific inquiry are questions about what makes causal powers possible and how changes in the organization of parts affect the total causative capacities of the system. Science largely involves studies of constitution (the study of the relation of dependences). Therefore, the constitution is at the heart of causal inquiries. There is justification then for an approach of the constitution to explain agent status or agency. This explanation offers a method for granting an AI system agent or person status through relations of internal dynamics and dependencies. This understanding of the constitution is what Kant was alluding to in his oft-quoted notion that things are to be understood as ends in themselves and never as means to an end.

The necessary asymmetries are present; the constitution explains causation, and the constitution is composed of parts and the organization of those parts. Systems then are made of causative parts and their organizations. The other asymmetry is existence. This asymmetry means that parts can exist independently of systems, while systems cannot exist without their parts (they can exist without some parts, but not all). The organization of parts is also fundamental for maintaining the status of a system (since systems are not reducible to their parts, they are greater than the sum of their parts). Organization therefore has explanatory relevance. Systems' causal capacities are not just the sum of their

parts; they are also the organization of those parts. Organizations' explanatory relevance stems from their contribution to the causal capacities of the system as a whole (change organizations and the causal capacities of the systems change). Organization is also called contextual causation and is empirically observable. Contextual is similar to downward causation (below), except that it displaces the notion of "downward" and instead posits that parts can influence each other regardless of a relative placement in relation to each other (Ylikoski, 2013). Parts can be of different sizes, different levels of abstraction, and situated at different levels. Causation is not limited to agency nor human agency, but it can also include instances of manipulation/intervention.

Constitution and causation are both explained in terms of their dependencies, which are a particular set of "objective" relations of dependent facts. These facts give explanations a direction and they are the basis for explanatory preferences (explanations must explain the systems' causal capacities in terms of their basis and not vice versa) (Ylikoski, 2013). Constitutive relations involve causal manipulation.

2.6. Downward causation

Downward causation provides an explanation for "emergence" which will also be necessary for an explanation of AI agency. However, downward causation has been criticized. For example, Kim (2006) argues:

> "[d]ownward causation is the raison d'etre of emergence, but it may well turn out to be what in the end undermines it".

However, this argument assumes the causal inheritance principle, which stipulates that the causal powers of complex systems are inherited exclusively from the causal powers of their parts. This has two salient points: (a) If parts do not have causal capacities, then the system as a whole would not (the capacities of the whole counterfactually depend on the capacities of the parts); and (b) in complex entities, nothing other than their parts are relevant to the determination of their causal properties. This then requires the causal powers of an entity to be internal to it.

Internal properties are context-insensitive, and an entity/system has all its internal properties (until there is an internal change) regardless of the context. If causal powers are internal, it is only the internal constitution of a system that confers those causal powers. This, and the assumption of internal causal properties, results in an ontological primacy afforded to the capacities of the parts, as opposed to the capacities of the totality/aggregates. The idea is that complex entities inherit their causal powers from their parts, but that the converse is not true. Complex entities cannot confer on their parts' causal powers which the parts did not have by their internal natures/capacities. Therefore, the properties of complex entities cannot explain why their parts have their causal powers (Walsh, 2015). This is Kim's argument against reflexive downward causation.

Kim's argument against emergence is the assumption of internal causal properties. This kind of thinking may have arisen from the notions of how mass (as a fundamental causal power/property) is context-insensitive. Masses of macroscopic objects are not altered by the masses of other bodies; mass behaves in a context-insensitive manner with regard to forces. An object's mass allows the prediction of its behavior across different contexts where forces act on it. It allows for the assumption that their effects are mutually independent, but do not affect masses.

Context insensitivity of causal powers is present in the analytic method (Cartwright, 2007). The assumption is that as contexts are altered, entities' causal powers remain unchanged because of the internal nature of the causal powers. However, context insensitivity does not equate to internality. The mass itself shows this. It is possible for a mass to be invariant across many contexts, but it is not an internal property of a body. For example, recently, it was discovered that the mass of protons comes from a combination of the masses of their constitutive three quarks, their movements, the strong force that ties them together (the gluon), and the interactions of quarks and gluons (Thomas Jefferson National Accelerator Facility, 2023). Hence, the mass of the proton is emergent. Mass is conferred onto a proton by its relations to something else. Causal powers may therefore be invariant in different contexts; they can be relational properties of things.

If causal powers are non-internal properties conferred on things by contexts, then one can argue that parts of complex systems get their causal powers from the system as a whole (*connubiality*). The parts would not have those capacities if they were not parts of that complex entity. The whole system, in this way, is the context that confers causal powers on its parts. This holds true even if the causal powers of a whole system are completely inherited from its parts.

Therefore, the property of the whole depends on the properties of the parts, and the converse is also true. If properties are understood to be relational (not internal in the strictest sense) and context-sensitive, it becomes easier to understand. Reflexive downward causation can be explained as follows: If they are relational properties, it means that complex systems have the causal powers that they do because of the causal powers of their parts (as in causal inheritance). It is also possible that parts have their causal powers because of the complex system they are part of.

In causally cyclical systems, one can assume that the causal powers of the parts are context-dependent and are conferred by the system in which they are parts. Hence, *emergence* is a fact of complex systems which can transform their parts (Ganeri, 2011). By transform, I mean that they confer on their parts capacities that they counterfactually would not have. These capacities reciprocally fix the properties of the system. Therefore, emergence can arise based on the context. Systems can give their parts causal powers and causal powers of the parts can be explained through reference to the system as a whole and its properties. They are hence relational, and the more suitable framing of this would be "intrinsic" as opposed to "internal". This is developed further at the end of the article.

2.7. Fundamental and emergence

"Fundamental" speaks to things that cannot be decomposed further into smaller resolutions, meaning that we cannot get

a coherent theory if we do so. What is fundamental is thus contingent on knowledge and the era that you find yourself in. Previously, atoms were thought to be fundamental until particle theory was established. However, emergence is different from fundamental, and unlike fundamental, emergence is something that is not conceptually contingent in the same way that fundamental is. "Emergence" can explain many issues in physics, such as how Schrödinger's (1944) order-from-disorder answer in his book *What is Life?* gives us a hint of a theory that incorporates emergence into complex systems. Out-of-equilibrium systems, for example, spontaneously build structures that dissipate energy, and, as they do this, they become increasingly stable and more complex. They have their own intrinsic dynamics. The dynamics of these systems can yield predictions and explanations, not just about the activities of the whole system but also about the activities of the parts. The movement of information toward order is also an emergent property. Vopson and Lepadatu (2022) demonstrated that while the thermodynamic entropy of systems increases (in terms of the studied virus), the *overall* informational entropy decreases (or stays constant). This is "the second law of infodynamics". The law itself works in opposition to thermodynamic entropy; it describes this movement as an *emergent entropic force*. Thus, we can now account for (1) informational emergence in complex systems, which (2) are not considered to be "alive".

2.8. Variance

2.8.1. Multiple possible variance and rarity

The microstate of a system is the configuration of the system (Hidalgo, 2015). Entropy is the logarithm of the fraction of all equivalent states. Entropy is lowest where the states have the least possible variance (order), and it is at its highest where there is the most possible variance (MPV). Rarity (Hidalgo, 2015) is the measure of the possibility of a particular arrangement occurring at random or without intervention. If it is rare, the probability of it occurring *without intervention* is unlikely. Functionality and working conditions are indicators of rarity. The natural state of things is to be in disorder, as opposed to order. States of the disorder have less information, and thus, the destruction of a physical order is also the "destruction" of information (informational content). Creating physical order is creating information (which is embedded in that order). The rareness of a state of order is measured against the number of possible states. One manner to do this is by correlating the connections between states. There is a correlation if one can get from state A to state B with a simple transformation. Information-rich states that involve correlations give the word "information" its colloquial meaning. Most things are made up of information. "Order" is a statistical probability measure of occurrence. Sometimes, the states of systems do not allow changes from A to B, or they impose limitations on the modes of transformation. The modes of achieving disorder outnumber the modes of achieving order.

2.8.2. Covariance, correlation, and mutual information

In statistics, correlation describes *the degree* of *linear dependence*, *association, distance, or relation* between two random variables in data. Correlations and standard deviations apply only in mediocristan non-scalable environments (Taleb, 2007) (Gaussian or bell-curves) wherein magnitude does not matter. In other words, both only have predictive or informational power in that context (see part D of Supplementary material). They can only be used to draw *qualitative inferences*.

Importantly, while correlation applies only to *linear relationships* between variables, this linear relationship (the signal) between variables *does not scale linearly*. Correlation is not additive (Taleb et al., 2023) because the correlation coefficients are non-linear functions of the magnitude of the relations between variables. They cannot be averaged for this reason. In turn, this means that an average of the correlation coefficients does not equal an average correlation itself. For example, a correlation coefficient signal of 0.7 conveys much less information than a coefficient of 0.9, while a signal of 0.3 conveys almost the same relationship as that of 0.5 (Salazar, 2022). Correlation cannot be used in non-linear relationships between variables, which is what characterizes reality (or Extremistan, scalable environments). Using it here will result in an incorrect explanation of the relation between random variables. In short, correlation does not accurately reflect the *informational distance* between random variables (Taleb et al., 2023).

Covariance speaks to the linear measure of the strength of the correlation between two or more sets of random variables. The covariance for two random variables (X) and (Y) each with a sample size of (S) is defined by an expectation value (Weisstein, 2023). Where there is a correlation between the values, the covariance will be non-zero. Where they are not correlated, it will be zero. The covariance can be directly proportional or inversely proportional. Covariance can be infinite, while correlation is always finite (Taleb, 2020). Covariance provides a method for construing features of contexts as "affordances" since this would be a qualitative finding and one that is non-scalable as described below.

The appropriate measurement function is *mutual information* (MI), which is not dissimilar to the Kelly criterion in finance and risk (see Supplementary material). In machine learning, this is known as *relative entropy* and is based on the expectation of the Kullback–Leibler divergence (a measure of similarity between distributions) (Taleb et al., 2023). Machine learning loss functions rely on entropy methods. Mutual information can be understood as a non-linear function of correlation; if mutual information increases, correlation itself increases, *but non-linearly*. Mutual information compares the probability of observing two random variables together with the probability of observing those same two variables independently (Prior and Geffet, 2003). In other words, an MI approach captures non-linear relationships and, importantly, it also *scales to noise*. The MI approach describes the amount of mutual dependence between two random variables; one gains information about a random variable by observing the value of another random variable. It measures this amount of dependence in information (in bits) and is used in instances of non-linear dependencies and discrete random variables. This is an entropy

measure, *and it is additive* (Taleb et al., 2023). This understanding of how "seemingly" random variables are related in terms of how the values or changes in one variable affect the understanding of the values or changes in another is an important tool.

Mutual information maps to the mutual dependence of random variables (how much can I rely on X if I know Y). Therefore, an MI approach would be most applicable to genetic distances (Taleb et al., 2023). Furthermore, an information metric is preferable and suitable for an account of agency or personhood, since DNA is understood as the basis of "life". Mutual information then provides the proper tool for creating a methodology with proper scaling, proper explanatory value, minimal informational loss (Taleb et al., 2023), and avoidance of using linear approaches (such as Cartesian methods or internal–external measures). Conditional mutual information (also known as transfer entropy) provides a suitable manner for causality detection since non-linear relationships (Mukherjee et al., 2019) in data associated with genetics and biological systems make *generalized data impossible*. Transfer entropy provides a consistent method across different conditions.

2.8.3. Intervention

Interventions usually involve notions of manipulations carried out on a variable (X) to determine whether changes in (X) are causally related to a variable (Y). However, any process qualifies as an intervention if it has the right causal characteristics, and not just human activities (Woodward, 2000). Consider this example: First, there is an intervention (I) on variable (X) which is a causal process that changes (X) in an exogenous way. If a change in (Y) happens after this, this change occurs *only because* of the change in (X) and not because of another set of causal factors (Woodward, 2000). One must also define what intervention means; interventions involve exogenous changes that break or disrupt previously existing endogenous causal relationships between variables and system states. This understanding of intervention allows for an extrinsic manner of specifying intrinsic features. It allows us to distinguish between types of correlations and dependencies that reflect causal and explanatory relations and those that do not. Viewing intervention in this way also transparently allows for the epistemological designation of experimentation as the establisher of causal and explanatory relationships. This allows us to make claims about the role behavior plays in causality through the use of interventions (Woodward, 2000). This is a much clearer account of causation and explanation as opposed to the traditional doxa.

2.8.4. Invariance, generalizations, and laws

According to Woodward, generalizations can be used in explanations and depend on invariance rather than lawfulness (Woodward, 2000). A generalization describing a relationship between two or more variables is invariant if it is stable or robust after the occurrence of an intervention or change in various other conditions at an appropriate level of approximation (Woodward, 2000; Maher, 2006). Invariance comes in degrees, and it has other features that capture the characteristics of explanatory generalizations in the social sciences, in particular (Woodward, 2000). In other words, invariance does not appeal to laws for its usefulness in explanations. The set or range of changes over which a relationship of generalization is invariant is known as its *domain of invariance*.

There are two types of changes, and both are fundamental to explanatory powers. The first is changes in background conditions (changes that affect other variables other than those variables which are part of the generalization) (Woodward, 2000). The second is changes in variables that are present solely within the generalization itself [within the Newtonian equation of $F=ma$, the change can occur to mass as (m) or acceleration as (a)].

For a methodology to constitute a law on personhood or agency, it must meet the conditions of laws (see part A of Supplementary material). This includes being a generalization with a higher invariance or wide applicability and being confirmable, predictable, and integrable (not only including being integratable with other laws, but also with philosophical or jurisprudential axioms which may ground legal laws, such as Kantian philosophy). Laws can also replace other older laws where they demonstrate that the older laws were unsuitable or provide less information.

2.8.5. Explanations and invariance

Good explanations require the use of invariant generalizations, which enable the specification of systemic patterns (of counterfactual dependence). This converts information into explanations since it can be used to answer a range of counterfactual circumstances about the explanandum. This allows for better predictive models. There are various kinds of counterfactual dependences, including active and passive ones; active is the type that is necessary for good explanations (Woodward, 2000). Invariance is thus necessary for reliance on counterfactuals and prediction (and to some degree also causal links). Invariance comes in degrees. There is also a connection between the range of invariance and explanatory depths; generalizations with more invariances constitute better explanations, especially for science. Generalizations that are not invariant under any conditions have no explanatory powers. Invariance is also important for building a purposive teleological account and countering the notion of "chance".

2.9. Theories of explanation: teleology and mechanism

2.9.1. What is teleology?

Teleology explains the existence of a feature based on its purpose (Walsh, 2015). The understanding that biological organisms are self-building, self-organizing, or adaptive suggests that they are greater than the sum of their parts. Thus, we can argue that organisms are purposive things. Refer to Sommerhoff (1950) in part B of the Supplementary material for information on how capacities can serve as a criterion of purposiveness.

2.9.2. Mechanism vs. teleology

Mechanists argue that natural selection explains the fit and diversity of organic forms, thus making teleology or purpose explanations unnecessary. The mechanical view is that every event

has a cause, with causes being able to fully explain events. But there are three main arguments against this approach: (1) non-actuality, (2) intentionality, and (3) normativity (Walsh, 2015).

The non-actuality argument states that means come before ends (goals). However, in terms of teleology, ends *explain* their means. Therefore, teleology in this light is inferential: it is the process of positing one's own presuppositions to establish an end. When the means occur, the goal or ends are not yet realized (they are non-actual). How can a non-actual state affect or cause a means?

The intentionality argument states that non-actual states of affairs cannot cause anything but mental representations of them can. One way to solve the teleology non-actual dilemma is to propose mental states as representations of these goals (or ends). Thus, occurrences of actions or events are explained by intentions as mental states of agents. The intentional and mental state argument is the most common justification of teleology (Kant and Bernard, 1790). The issue is that organisms typically do not have intentional states. However, this intentional and mental state justification is most commonly used in teleology. The earliest form of teleology can be found in Plato's *Timaeaus* and in the works of Thomas Aquinas; after all, any perceived forms of an order must presuppose a purpose or an intention. Aquinas argues that whatever lacks intelligence cannot move toward the end unless it is directed by knowledge, "as the arrow is shot to its mark by the archer." Intentionality is the obvious paradigm for teleological framing. Kant (2000) notes that intentionality is our only model for understanding purpose.

The normativity argument suggests that teleology has a normative value. Explaining an action as a consequence of intention is to argue that an agent was rationally required or permitted to act in a particular way to achieve certain goals. Rational actions are those which are required to attain a goal (or end). Thus, a teleological approach must account for an action being rational (Walsh, 2015).

Bedau (1991, 1998) argues that because of the normativity of teleological explanations, goals can have their explanatory roles only if they have intrinsic normative properties. Namely, (*c*) construed as a means toward attaining a goal (*g*) could only be something that a system *ought* to produce, if (*e*) is a state that the system *ought* to attain, but (*e*) could not be an "ought to attain" state unless (*e*) was intrinsically good. The issue is that natural facts are not intrinsically evaluable (Walsh, 2015). A proper account of teleology must account for all these arguments in making space for purpose. Furthermore, a proper teleological account must not be purely metaphysical, but must also operate within a scientific framework. Emergence is an important aspect of the account of agency. The dynamics of agents must be explained by their purposes and affordances. These would be emergent properties that emerge from the relation between agents and their contexts. They are not properties of the systems' parts themselves. Mechanistic explanations tend to exclude emergence since they appeal to the dynamics of complex systems as being entirely explainable through the properties of their parts (Walsh, 2015). Parts are not emergent. However, before solving the emergence issue, I need to account for "purpose".

2.9.3. Teleology and purpose

Teleology explains the existence of a feature based on its purpose (Walsh, 2015; Kampourakis, 2020). We can argue that organisms are purposive things because organisms or agents are self-building, self-organizing, and adaptive, which suggests that they are more than the sum of their parts.

2.9.4. Chance and purpose

In biology, Jacques Monod considered the consequences of a non-purposive nature/biology. He identified a contradiction at the heart of evolutionary biology. This is the "paradox of invariance" (Monod, 1971). The paradox is that living creatures show two contradictory properties: invariance and purpose. Invariance is the ability to reproduce and transmit information, including *ne variateur* information. *Ne variateur* information relates to its own structures and is transmitted from one generation to the next. The purposiveness of organisms is evident in the maintenance of their viability by responding to environments and adaptation. However, many would argue that science does not recognize this kind of purpose because it seems to be a contingent truth instead of an objective one. To explain this, Monod suggested that purposiveness can be explained by the mechanism of molecular invariance (Walsh, 2015).

However, the invariance principle raises complications as evolution is fundamentally about change. Adaptive evolution is a form of environmentally charged biased change. Thus, there should be a source of new variants and a process that is biased toward change. If we argue that new variants are biased in favor of goals and purposes, we may also be undermining science. For Monod (1971), the source of evolutionary novelties must come from unbiased chance. Monod argues that chance must have a requisite role in evolution, and this role is methodological and not metaphysical. This is akin to Democritus, who argues that everything is a result of chance and necessity. With chance and necessity, there is no need for purpose (Walsh, 2015). However, the chance is unsuitable for an account of purposiveness that I want to build.

Aristotle took issue with Democritus's explanation, since chance is, by its nature, not measurable. In *Physics Book II*, Aristotle discussed what an explanation should include. His arguments were developed to counter the atomists' arguments at the time, which are similar to the mechanists' arguments of cause and effect. He did not like explanations that did not account for something—and chance was unaccounted for. He illustrates this (*Physics II.5*) (Barnes, 1991) with the story of a man who is collecting money. The man meets a debtor at the market and collects money owed to him. This, for Aristotle, is a chance encounter since the collector went to the market for a different purpose; he coincidentally also collected his money. This is a mechanistic explanation, and these explanations do not distinguish between occurrences that are regular/purposive or chance. They therefore give incomplete information since they do not distinguish between both. Mechanistic explanations are necessary since every occurrence must have a mechanical cause, regardless of whether it occurred for a purpose or because of chance (Walsh, 2015).

Purposive events are, however, robust (invariant) across a range of alternate initial conditions and mechanisms, whereas chance events are not (they have differing modal profiles). Good explanations must be able to distinguish these. Purposive encounters are those which are insensitive to initial conditions, including locations. Thus, in purposive occurrences, the means counterfactually depend on the ends. Chance occurrences are sensitive to initial conditions and, if the initial conditions are different, the event or ends would not have happened. Unlike chance occurrences, purposive occurrences are sensitive to goals. If an agent's goals were different, the event would now have occurred. If the collector had been elsewhere in the market, then the encounter may have happened elsewhere, at a different time, and by different mechanisms.

Given the counterfactual dependence of mechanisms and ends, events that happen because they serve a purpose can be explained in two ways: (1) the occurrence results from mechanical interactions and (2) the occurrence is conducive to the fulfillment of a goal. However, one thing is certain; one cannot simply disregard purposes. If purposes are ignored, it induces a "selective blindness" to a class of explainable occurrences, namely, those that are structured according to the counterfactual dependence of means on goals. This is not just an error of omission; it also risks misconstruing purposive occurrences as blind chance. To properly account for events, both teleology and mechanistic explanations are needed. I have now explained purposiveness as goals; these purposes can also explain their own means (Walsh, 2015).

2.9.5. Goals

Goal-directed processes are those that are conducive to stable end states and their maintenance. The end state itself is the goal. Thus, a goal is a state that the goal-directed process is aimed toward. Central to studies on natural goal-directed processes is an adaptive and autonomous system, which can achieve and maintain persistent and robust states through the implementation of compensatory changes (Di Paolo, 2005; Barandiaran et al., 2009). These systems can pursue goal states and sustain them in the presence of perturbations. They can effectively implement changes to component processes in ways that correct the effects of perturbations, which could otherwise result in the system not achieving its goal (Walsh, 2015). This will be necessary for an account of purpose and agency.

The architecture of the system underpins the goal-directed capacities and states of the goal itself. These systems are usually comprised of modules. These modules are clusters of causally integrated processes decoupled from other modules. They also demonstrate the capacity to produce and maintain integrated activities across a range of perturbations of influences (robustness). Each model has regulatory influence, using positive and negative feedback, over a small number of other modules. Each part effectively influences other parts in some way. This allows for robustness and plasticity by maintaining stability in the presence of perturbations by enacting new adaptive changes. Robustness describes a property of something which can produce novelty, in response to novel circumstances. Biological organisms display this. What allows organisms or systems to do this is the modularity of their development (Schlosser, 2002).

Thus, goal-directed behavior is a causal consequence of the architecture of adaptive systems. Furthermore, it is an observable feature of systems dynamics. It is the capacity of systems as a whole to utilize the causal capacities of their parts, and the ability to direct them toward attaining a robust and stable end state. That end state or goal is not a mysterious something; it is a complex and relational property—the property of being in a state that a goal-directed process can achieve and maintain. Therefore, goals are natural and observable (Walsh, 2015). Goals are thus not "mental states" and instead are naturally derived from a system's intrinsic dynamics.

But what about the content of teleological explanations? We can determine the conditions under which they apply as explanations, but we must also account for the content of the explanation. There is a fundamental difference. Conditions for teleology can be understood as causal occurrences; however, content cannot be described in causal terms. Teleology is not about explaining causes, it is about explaining goals to which events are conducive (Walsh, 2015). Thus, for agency, we no longer need to rule out an entity based on being "created" or "developed" by something or someone else. The focus is on the entity itself.

2.9.6. Teleological explanations and invariance

To describe a non-mechanistic account of goals, two questions must be answered: (1) How can an event be explained by citing the ends to which it is simply a means; and (2) Why does this explanation not need to be explained through mechanisms of cause and effect?

To address the first question, goals can explain their means of achieving those goals in a way that is similar to how mechanisms explain their effects by using counterfactual invariance relations. Invariance here does not mean the transmutation of stability of form across generations or lineages. Here, it is *Woodwardian invariance*. We can answer the second question by simply demonstrating that they appeal to different invariance relations more than mechanistic explanations do (Walsh, 2015).

Mechanistic explanations demonstrate how activities and characteristics of (X) produce (Y) as the effect including the specific properties related to that effect. Activities *produce* effects, which are related through the notion of counterfactual dependence—effects counterfactually depend on their mechanisms. These activities can be expressed in terms such as "binding", "opening", and "bending". Woodward (2003) called this "relation invariance":

> "[T]he sorts of counterfactuals that matter for purposes of causation and explanation are just such counterfactuals that describe how the value of one variable would change under interventions that change the value of another. Thus, as a rough approximation, a necessary and sufficient condition for *X* to cause *Y* or to figure in a causal explanation of *Y* is that the value of *X* would change under some intervention on *X* in some background circumstances".

Thus, we can use this to explain how events as means are related to their goals. If there is goal (X), which then produces event (A) which is conducive to (X) under conditions (Q), then under

different conditions (V), it would produce event (B), as (B) would be more conducive toward (A). If the system had another goal (Z), it would produce event (C), should (C) be more conducive toward attaining (Z). This is an invariance relation. It is the obverse of the relation of cause and effect. In other words, we explain that causes themselves explain their effects, because when the cause occurs, then so too would the effect. If the cause does not occur, neither does the effect. We can also reason that a goal explains its means because if a system has a goal then the means too would arise, and if there was no goal then the means would not arise.

This explains how events, as means, are related to their goals. Causes explain their effects because when the cause occurs, so does the effect. If it does not, neither does the effect. We can also reason that a goal explains its means. If a system has a goal, the means arise; without a goal, the means do not arise. But, on its own, invariance is insufficient. Explanations are description-dependent, and good explanations enhance understanding. Mechanistic explanations do not simply speak to cause and effect (relations), and they also speak to the appropriateness or accuracy of that relation. The relation itself only exists if it is appropriate. We use concept descriptions such as "push", "pull", and "attract" to describe productive relations. These speak to the nature of the relation, and sometimes also explain the effect.

For teleology, we use the concept descriptor of "conduce/ive". So, the modal relations are (1) causes produce effects; and (2) means are conducive to their ends. *Conducing is not causation.* A means is only considered conducive to its ends if it robustly and reliably brings about the end *ceteris paribus* across a range of counterfactual circumstances. Hence, if the goal is (A) and event (X) causes (A), this does not mean that (X) conduces to (A) (Davidson, 1980). Thus, producing and conducing are descriptions of events, and they have different informational content. Producing specifies an earlier event (time is important here), which is the mechanism for the later event. This describes *how* the later event arose. Conducing specifies the *why* of an event—that it is conducive to realizing or maintaining a goal.

A singular event can be explained in terms of mechanistic (causal) and teleological (conducive) relations. The former explains how things happen, while the latter explains why they happen, and thus they co-exist. They are complementary and non-competing. They are also complete—they do not need each other to explain their own coherence—the how's explain the how's and the why's explain the why's, and we do not need the how's to explain the why's. They both explain different information about events. However, for the completeness or coherence of an explanation as a whole, one needs both types of sub-explanations. Without both, there is an explanatory loss. Thus, both mechanism and purpose are important for explanations but not for independent systems themselves. The non-actual claim, for example, is a conflation between causes and explanations. In terms of the intentionality counter, intentions can be understood as goal-directed activity instead of mental representations. Intentional states are mental representations and are unnecessary for teleology (Walsh, 2015).

In terms of the normativity counter, the goal need not be described as "good" to explain why systems *ought* to act in certain ways, which result in conducing to that goal. Systems will do what it takes to achieve the goal; there is no specific modality to be followed. The modality need not be prescribed, singular, or of a specific nature (such as good or valuable). What matters is *appropriateness*. There is thus no need for an evaluative state of affairs. Aristotelian teleology is not intentional, transcendent, or causation-based. It comes about because of the activities of goal-directed entities which are observable and occur in the natural world. This can be used for both predictive power and explanatory power in the same way that we use other robust regularities (Walsh, 2015).

2.10. Theories of explanation: agents and objects

2.10.1. Natural agents

Natural agents are obtained from the natural purpose explanation. Agency, such as purposiveness, is an observable property of a system's gross behavior. The system can pursue goals and respond to conditions of its environment and its internal constitution in ways that promote the attainment and maintenance of its goal states. The agency is observable in the sense that we see agents negotiating with situations using its dynamics. We can see a range of robust and regular responses to conditions. If we understand its goal, we can understand its behavior. The agency is ecological as a system that can cope with its context and achieve its goals by responding to *affordances as affordances*. An ecological definition of the agency includes three inter-definable factors: (1) goals, (2) affordances, and (3) repertoire (Walsh, 2015). Affordances are opportunities for, or impediments to, a goal; only goal-directed systems can experience its conditions as affordances. Systems can experience affordances only if they have repertoires, which are sets of possible responses that systems can enlist in pursuit of goals (in response to the system's experienced conditions). For repertoires to constitute a response to affordances, repertoires must be biased. Systems must be able to exploit behavioral repertoires in response to conditions in ways that are conducive to the attainment or the maintenance of their goal. The goal of the system is the state that it moves toward attaining/maintaining by directing behavioral repertoires in response to affordances conducing that state. Repertoires come in degrees, and some agents have richer repertoires than others. Systems with wide ranges of repertoires can respond to more affordances and can pursue a wider range of goals. Ecological agency is not all-or-nothing: It comes in degrees. There is a continuum from the most basic agents capable of pursuing a narrow range of goals to those possessing greater repertoires of responses. Cognitive systems tend to have large repertoires, with thinking forming part of their repertoire (Walsh, 2015). This is a model in which we can "grade" or rank the agent status of a system. A system will have a greater agent status grading if it demonstrates a greater repertoire (as variable responses to affordances) for maintaining or improving the conduciveness toward a goal (see Parts E and F of the Supplementary material).

2.10.2. Object and agent theories

There is a difference between object and agent theories. Object theories that we use today aim to describe and explain the dynamics

of objects (Walsh, 2015). To construct these theories, we create a space of possible alternatives for those objects. This is known as a "state space". We then look for principles that may account for various possible trajectories through this state space. The objects in these domains are subject to forces, laws, and initial conditions. Lee Smolin dubs this the "Newtonian paradigm" (Smolin, 2013). This describes system dynamics by the answers to two questions: (1) What potential configurations does the system have; and (2) In each configuration, what forces is the system subject to (Smolin, 2013)? In this paradigm, the laws, forces, and initial conditions are irrelevant to and exist separately from the objects. Object theories are transcendental, and they have an explanatory asymmetry. Transcendental means that the principles that govern the dynamics of the objects in the theory's domain are not part of the domain itself. They do not evolve as the system does, and the laws of nature and the space of possibilities through which the objects move remain constant as the objects change (Walsh, 2015). This allows for the explanation of the changing state of a studied system by appealing to unchanging laws.

2.10.3. Action theories

The Cartesian view holds that agents' thoughts, beliefs, and desires explain their actions only if they cause said actions (Davidson, 1963). This means that contemporary action theory is interpreted as implying that thoughts are mental entities realized as internal physiological mechanisms and that these mechanisms combine with other internal mechanisms to effect actions. They do so by their intrinsic causal properties (Fodor, 1987). Actions are outputs, such as an internal process of computation, and they result from the mechanical interactions of the internal states of the agent. The purposes of agents and their dynamics do not appear in the explanations of actions. The Cartesian model (thought and action) posits that agents are akin to "middlemen" (Walsh, 2015) since they are the connection between the causal activities of their psychological states and the environmental demands that they experience (Walsh, 2015). Haugeland (1998) described the notions of *intimacy* and *commingling*. His conception was in opposition to the Cartesian mind which posited that the mind is entirely internal to the agent, and the position that the environment is entirely external to the mind. In Cartesian dualism, both communicate through perception (which is environment to the mind) and action (which is mind to environment). Haugeland (1998) thus argued that the mind played an active role in constituting the environmental conditions to which it responds. Intimacy in this explanation described the mind as embodied and embedded in the world. This is not just an interdependence but also a "commingling" or integralness of the mind, body, and world, which undermines any separation between them.

2.10.4. The disappearing agent

The standard action theory approach created the issue of the missing agent, which is a consequence of its underlying methodological commitments (Velleman, 1992; Hornsby, 1997). These have arisen from precepts of the Cartesian mechanisms already described. It ignores the fact that actions do not happen to agents: *they are performed by them*. Cartesian mechanisms of action miss this point—an action is something produced by an agent for a reason. A proper account of action involves explaining the doing of agents by highlighting them to be reasonable or rationally justified considering the agent's purposes. The agent's goals will explain the appropriateness/conduciveness of the actions undertaken. Viewing actions as just causal consequences of internal states erroneously misses the fact that actions are purposive activities in lieu of goals. The Cartesian object theory views agents as objects, wherein the actions of agents are explained/caused by extraneous forces that act on said agents. It does not explain actions as products of agency, but rather as effects of extrinsic causes: external environments and internal computation and representation. Thus, it is an exclusion of agency which is both real and natural. This is also present in the understanding of "rational action". Action theory is divided between two conceptions of humans: (1) as objects in the natural world, subject to external causal influences; and (2) as agents able to initiate actions that are guided by reasons (Walsh, 2015).

Merleau-Ponty explains behavior as commencing with an active organismal agent that is problem-solving and goal-pursuing (Matthews, 2002). The agent responds to conditions as meaningful, either obstacles or opportunities. The goals and capacities of the agent give importance to the conditions. Thus, actions are responses initiated by agents to sets of affordances, and these affordances are largely of the agent's making. Agents also co-evolve with these affordances in line with their actions and goals. Agent theories of actions view actions as events that are generated by agents because of agents' pursuit of the goals. These purposes explain and justify the actions and not the other way around. Adaptive evolution is thus a phenomenon of agency. Thus, using an agent theory of this sort enables proper conceptual underpinning for agent status and agency in combination with natural purpose and goals (see Part F of the Supplementary material).

2.10.5. Autonomy

Agents create degrees of freedom for themselves by constituting their affordances through self-maintaining and self-regulating activities. They determine which environmental conditions are important. They also enable the exploitation of opportunities that the environment presents. This is a stronger account of autonomy. The integral processes in autonomous systems are (1) continually dependent on one another in their formation and realization as a network; (2) make up a unity (converge) in their domain of existence; and (3) govern areas of exchanges with the environment (Thompson, 2007). Autonomous agents can "make sense" of circumstances. Making sense means to detect and use the features of one's context, which in turn also constitutes the features/context. This is then the capacity of the agent to mobilize its resources in a way that supports the pursuit of its goals, and by exploiting opportunities or reducing impediments. Agents make features significant in the way they are detected and responded to in pursuit of their goals. In this way, autonomous agents construct and constitute the conditions that they respond to. There is a reciprocity of form and affordance—as form evolves so do affordances (Walsh, 2015). As mentioned above, this is related to the repertoire of capabilities. Thus, systems as agents that demonstrate a greater ability to identify, interpret, utilize,

and implement features as affordances in pursuit of their goals would be graded higher [see Part C of Supplementary material for a supportive moral perspective on AI and agency and the supportive novel Technological Approach to Mind Everywhere (TAME) framing].

3. Constructing the AI agent

3.1. Write-re-write systems: semantic closure

Semantic closure is a concept that refers to the fact that a system can enclose meaning within itself. In biology, for example, a string of DNA and messenger RNA (mRNA), the encoding mechanism between both, has evolved, altering the meaning of DNA by rewriting the genetic code (Clark et al., 2017). In biology, the most important factors related to this concept are the ribosome, transfer RNA (tRNA), DNA, and mRNA (Clark et al., 2017). The tRNA is involved in expression which defines the meaning of DNA by mapping the three bases of DNA to one amino acid. Changing the mapping also means rewriting the genetic code. Hence, the meaning of the genome can itself be altered (Clark et al., 2017). Rewriting in biology is the process of moving from one semantically closed state to another.

It is important then to understand how meaning originated for translating proteins and how it has been altered through evolution. This is an ontogenetic or bottom-up approach (Clark et al., 2017). For this process of moving from one semantically closed state to another, there must be a necessary structure. Von Neumann was the first to describe what an artificial architecture that enables semantic closure would look like. His constructor theory birthed the modern form of universal constructor architecture (Clark et al., 2017). Some of these models have highlighted the necessity of redundancy in maintaining stability in the presence of mutations. In the proposed theorem of chemical construction theory, the authors also highlight the self-referential nature of the genome (it contains descriptions of all other machines in the system, and hence it is its own description) (Clark et al., 2017). In their experiments, the authors demonstrated how alterations in the expressors can lead to novel interpretations of the genome which, in turn, gives rise to pleiotropic effects. Thus, the meaning of the genome has been changed, and this new interpretation of it extends to other molecules, not just the expresser. They also demonstrated that it is not genetic material that evolves but also the mechanisms of copying. Each string can play different functions in many different relations or reactions. Control in this way is distributed throughout the system (there is no explicit or centralized control mechanism). The authors also postulated that the ribosomes may be the biological equivalent of any string that imposes meaning into the system (Clark et al., 2017).

Finally, the authors proposed something interesting: there were *emergent* or transient changes that were expressed and *which did not* appear in genetic records. *These arise through inaccurate expressions*. Their results demonstrate that these "errors", while not reflected in the genome, are reflected in heritable changes in expression (they are covert). Errors in expression in biology are deleterious or non-heritable, since only genomic information is thought to be heritable (Clark et al., 2017). They also provide evidence for misreading errors of this nature, including the streptomycin-dependent phenotypes of *E. coli*. Errors in ribosomic interpretation of DNA have been demonstrated previously (Clark et al., 2017). In this way, they can change meaning. The authors stated that expressors can make a consistent interpretation of a genome (meaning it leads to its own expression). By interpreting its own genetic material, *expressors obtain meaning through self-reference*. From this, we can use semantic information as the central measure for an account of personhood or agency. Importantly, it is not tied to a biological brain, and systems can themselves enclose and change their own semantics. Self-reference in this light provides another framing for personhood and agency. This study also provides backing for "emergence".

3.2. A semantics model for personhood, agent, and agency

3.2.1. Semantics

Historically, semantic information was contrasted with syntactic information. Syntactic information quantifies the kinds of statistical correlations between two systems without giving meaning to those correlations (Kolchinsky and Wolpert, 2018). This is used predominantly with Shannon's information theory, which is a measure of the reduction of statistical uncertainty between two system states which can differ in time.

Some studies (going forward known as the study or this study) have distinguished between syntactic and semantic information in systems (Kolchinsky and Wolpert, 2018). This study attempted to create a formal definition of semantic information that is applied to both "living" and "non-living" beings (any physical system like a rock or cell, for example). Herein, semantic information was defined as *information that a physical system has about its environment which is causally necessary to maintain its existence over time*. The qualitative aspect of semantic information is related to the intrinsic dynamics of systems and their "environments". The *quantitative tools* used to calculate semantics are information theory and non-equilibrium statistical mechanics.

Importantly, the study is distinguished between "meaningful bits" and "meaningless bits". This also allowed for a differentiation between sub-concepts of semantic information such as "value of information", "semantic content", and "agency." Semantic information then is defined as information that enables systems to achieve their goals (maintaining a low entropy state). However, this is not an exogenous (goal derived from or measured from "external" sources) approach. Any "meaning" obtained from exogenous studies is meaningful (in terms of goals) from the *perspective of the observer* or scientist, and *not the system itself*. The difference in this study as compared to others is that the others offer standard teleo-semantic approaches where goals are understood in terms of evolutionary successes such as fitness. These standard approaches are suited more for systems that change according to selection; they do not describe systems that are "non-living" or synthetic (Kolchinsky and Wolpert, 2018). They also tend to be etiological, in that they are based on past histories of the system. The approach presented in this study instead creates an account of semantic information based solely on the intrinsic dynamics of

a system in an environment without regard to its past or origin. Therefore, it presents an attractive model for an account of agency which includes AI systems. This is an autonomous agent model which requires that a not-in-equilibrium agent maintain its own self-existence/maintenance in an environment. This is active self-maintenance where agents use information about the environments to achieve their goals, and hence this information is intrinsically meaningful for them (Kolchinsky and Wolpert, 2018). This kind of perspective also *applies to robots and "non-living" systems*. This intrinsic goal is neither obtained from an exogenous source nor is it based on past histories or origins. Importantly, semantic information is derived from the *mutual information* between the system and its environment (within the initial distribution, which is defined as stored semantic information).

3.2.2. Viability and value

The study coins the term "viability function". Viability functions are used to statistically quantify the system's degrees of existence at any given time (hence, one can say that viability functions describe real-value aspects of systems). For this, a negative Shannon entropy is used (it provides an upper bound on the probability that the system occupies any small set of viable states). Semantic information now means the information exchanged between the system and its environment which causally contributes to the system's existence. It is measured by the maintenance of the *value of the viability function*. To quantify causal contributions, the study used a counterfactual intervened distribution in which there was a *scrambling of syntactic information* between the system and its environment. The value of information was defined as the difference between the system's viability in time after the intervention. A positive difference would mean that there was some syntactic information between the system and its environment which plays a causal role in maintaining its existence. A negative difference would mean that the syntactic information would decrease the system's ability to exist.

To describe the value of information, the study gives the example of a rock. A rock has a very low dynamic and thus it can remain in a low entropy state for longer periods. If information is then scrambled by swopping rocks from their current environment into different ones, this intervention would not make much difference to the rock. However, by doing the same thing with a hurricane (my modified explanation) that requires specific conditions for its maintenance, the result is that the hurricane has a greater set of parameters for its maintenance. If those parameters are not met, it will dissipate (viability decreased)—and thus it has some important semantic information. Therefore, the semantic information is important for hurricanes, and this would likely be greater for hurricanes than rocks. If you put an organism in a new environment it may not be able to find its own food, hence organisms place a higher value on information.

3.2.3. Viability, syntactic, and semantic information

Non-equilibrium systems are those in which the non-equilibrium status is maintained by the ongoing exchange of information by sub-systems. An example of this is the "feedback-control" process in which one subsystem acquires information about another subsystem and then uses this information to apply controls to keep itself or the other system out of equilibrium (like Maxwell's demon). Information-powered non-equilibrium states differ from the traditional non-equilibrium systems considered in statistical physics which are driven by work reservoirs with control protocols, or which are coupled to thermodynamic reservoirs (Kolchinsky and Wolpert, 2018). The reduction of entropy thus carries costs in the expenditure of energy as heat. Within the thermodynamics of information, Launderer's principle states that any process that reduces a system's entropy (by x number of bits) must release energy in the form of heat. Heat generation is also necessary for the acquisition of syntactic information. *Viability* is connected to this reduction of entropy through semantic information acquisition. Semantic efficiency in the study speaks to a quantification value of how much the system is "tuned" to possess only syntactic information which is relevant for maintaining its own existence. The semantic efficiency is related to the thermodynamic multiplier which is the measure of the "bang-for-buck" of information (below). This simply asks, "what types of information would carry more benefit than other types?" Systems with positive values of information and higher semantic efficiency tend to have a larger thermodynamic multiplier (Kolchinsky and Wolpert, 2018). Stored semantic information is not that which is acquired during dynamic exchanges with environments. Rather, it is the *mutual information* between systems and environments that is also *causally responsible for maintaining viability*. It is important to note that systems with low entropy are not the same as remaining within a specific viability set. This means that systems do not need to maintain the same "identity" over time to maintain a low entropy state. Identities can change while still maintaining low entropy states. Hence, a specific identity profile (like a human) is unnecessary for an account of agency.

Observed semantic information in the study speaks to that which is affected by the dynamic interventions that scramble the transfer of entropy from the environment to the agent. This kind of information identifies semantic information which is acquired by dynamic interactions between systems and environments (not mutual or stored information). The syntactic information in the study is scrambled to obtain semantic information. This is how *meaningless* and *meaningful* are obtained (optimal intervention determines this). Any information that can be scrambled without affecting viability is meaningless and that which must be preserved to preserve viability would be meaningful. Both observed and stored information are necessary for viability preservation; however, observed speaks to dynamic interactions between systems and their environments. The semantic efficiency ratio is the ratio of the stored semantic information to the overall syntactic information (Kolchinsky and Wolpert, 2018).

Systems can have a non-unique optimal intervention, namely, *multiple variable and redundant sources of semantic information* which are used to maintain viability (like relating to different food sources, see Kolchinsky and Wolpert, 2018). This is important when considering the different dimensions of society in which systems are integrated. Relevant reservoirs depending on the system, its context, and its function can include sexual reservoirs, ethical/behavioral reservoirs, different knowledge

domain reservoirs, socio-political reservoirs, and socio-emotional reservoirs. This presents a paradigm and mechanism to determine the status and inclusion of certain systems in certain contexts by assessing their suitability to participate adequately in that context. The thermodynamic multiplier provides a means to determine *suitability*.

3.2.4. The thermodynamic multiplier

The thermodynamic multiplier is the stored semantic information (the benefit–cost ratio of mutual information) that provides a manner of comparison for the ability of different systems to use the information to maintain their viability (Kolchinsky and Wolpert, 2018). This would mean that the stored semantic information gains its status *based on its benefit outweighing its cost*. If the information value is positive, then having a low semantic efficiency means that there would also be a low thermodynamic multiplier. Therefore, "paying attention to the right information" in terms of semantic efficiency is also correlated with thermodynamical efficiency. It is a measure of the thermodynamic costs of obtaining new mutual information compared to the viability benefit obtained from that acquisition.

3.2.5. Transfer entropy and semantic efficiency

Observed semantic information can be acquired in dynamic interactions through the use of transfer entropy. This is a measure of information flow and is widely used and understood. The transfer entropy movement from the environment to the system is not necessarily the same as the flow from the system to the environment. Observed semantic information describes dynamic actions and decisions where any information scrambling that comes from the environment to the organism *would result in an impact on viability*. For example, Jack and Jill went up the hill with Jack leaving behind a trail of breadcrumbs to lead them back home. If at some point during their adventure, a wind were to blow away those breadcrumbs then they would not know their way home, affecting their ability to survive or feed themselves. The transfer entropy would speak to the breadcrumbs which would have observed semantic information because the breadcrumbs as an object contain an informational interaction between a system (as agent) and the environment. Thus, the *value of the transfer entropy* is the viability value at a specific time before scrambling versus the viability value after scrambling. This value is then known as *semantic efficiency* (Kolchinsky and Wolpert, 2018).

3.2.6. The agent

An autonomous agent (and autonomous agency) in this system would be a physical system that has a large measure of semantic information (Kolchinsky and Wolpert, 2018). One can identify autonomous agents by finding timescales and system/environment decompositions which maximize measures of semantic information. This would in turn depend on the thermodynamic multipliers, the transfer entropy, and the amounts of semantic information. It is, however, important to remember that semantic information can have a negative viability value. This means that it can be mistaken/misrepresented information that is used in a way that harms the agent's viability value. The study also highlights that semantic information requires an *asymmetrical measure* (unlike syntactic mutual information). This is because this information concerns the *viability of the system* and not the environment. This system also does not require the decomposition into separate degrees of freedom (such as sensors, effectors, membranes, interior, exterior, brain, or body). Thus, it is not about internal representations *but rather about the intrinsic dynamics* of the system and its environment. This can also be used to create an account for "life".

4. Conclusion

My methodology, using successful observable, predictable experiments that provide more information, is more accurate and enables a method of grading or ranking systems as agents according to domain suitability. This relies on the use of semantic information and its relationship with viability. To summarize, viability (reducing or maintaining a low entropy state) is the ability of a system to continue to exist, and it is measured in terms of the viability function. Changes in this viability function are determined by counterfactual dependences obtained through the scrambling of syntactic information. This enables the ascertainment of the more "valuable" semantic information as causally contributing to the system's viability function. There are two kinds of semantic information, both of which affect the viability function: (1) stored and (2) observed. Stored semantic information is the *mutual information* between systems and environments, while observed information is that which is acquired by dynamic exchanges between systems and environments. One can obtain observed semantic information by scrambling the transfer entropy. The observed semantic information is necessary to determine actions and agency since it describes dynamic "active" interactions. Furthermore, survival in this instance is de-linked from "biological" systems and is measured according to maintaining a system's viability based on its own intrinsic dynamics. This presents an attractive way to create a general and invariant account of personhood and agency. I also presented an account of what constitutes rarity. This provides a further attractive way to grade "emergent" information content or properties.

This account, routed in Kantianism, recognizes the explanation and information issues in alternative accounts and provides a more accurate framework. Legal systems and ethics discourse should take note of this account as the usual ways in which these conversations are entertained and are doomed since they tend to rely on poorly understood, ephemeral notions such as "consciousness". Instead, systems should be evaluated according to their own intrinsic properties which enable a better approach to determining suitability (agency and personhood) because it considers agents as agents within their own informational paradigm and not relative to another agent's informational paradigm. In this way, intrinsic bias is made to be a strength when it is considered from the perspective of the system itself.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://osf.io/evna6.

Author contributions

MN: Conceptualization, Investigation, Methodology, Project administration, Resources, Visualization, Writing—original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This article was supported by the U.S. National Institute of Mental Health and the U.S. National Institutes of Health (award number U01MH127690) under the Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) program. In addition, support was also obtained from the National Research Foundation (NRF) under the Doctoral Innovation Scholarship (award MND190619448844).

Acknowledgments

The author would like to acknowledge Amy Gooden for her technical editing of this article and Donrich Thaldar for acting as a critical reader and for his helpful comments about earlier drafts of this article.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Author disclaimer

The content of this article is solely my responsibility and does not necessarily represent the official views of the U.S. National Institute of Mental Health or the U.S. National Institutes of Health.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1273470/full#supplementary-material

References

Barandiaran, X. E., Di Paolo, E., and Rohde, M. (2009). Defining agency: individuality, normativity, asymmetry and spatio-temporality in action. *J. Adapt. Behav.* 17, 1–13. doi: 10.1177/1059712309343819

Barnes, J. (1991). *The Complete Works of Aristotle*. Princeton: Princeton University Press.

Bedau, M. (1991). Can biological teleology be naturalized? *J. Philos.*, 88, 647–655. doi: 10.5840/jphil1991881111

Bedau, M. (1998). "Where's the good in teleology?" in *Reprinted in Nature's Purposes: Analyses of Function and Design in Biology*, eds. C. Allen, M. Bekoff, and G. Lauder (Cambridge: MIT Press), 261–291.

Cartwright, N. (2007). *Causal Powers: What Are They? Why Do We Need Them? What Can and Cannot be Done With Them?* London: Contingency and Dissent in Science Project.

Clark, E. B., Hickinbotham, S. J., and Stepney, S. (2017). Semantic closure demonstrated by the evolution of a universal constructor architecture in an artificial chemistry. *J. R. Soc. Interface* 14, 1–12. doi: 10.1098/rsif.2016.1033

Cornell University (2022). "Eye expressions offer a glimpse into the evolution of emotion," in *ScienceDaily*. Available online at: https://www.sciencedaily.com/releases/2017/04/170417182822.htm (accessed March 6, 2023).

Craver, C. (2007a). *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Oxford University Press.

Craver, C. (2007b). Constitutive explanatory relevance. *J. Philos. Research* 32, 3–20. doi: 10.5840/jpr20073241

Craver, C., and Bechtel, W. (2007). Top-down causation without top-down causes. *Biol. Philos.* 22, 547–563. doi: 10.1007/s10539-006-9028-8

Cummins, R. (1975). Functional analysis. *J. Philos.*, 72, 741–765. doi: 10.2307/2024640

Cummins, R. (1983). *The Nature of Psychological Explanation*. Cambridge: MIT Press.

Cummins, R. (2000). "'How does it work?' versus 'what are the laws?': Two conceptions of psychological explanation," in *Explanation and Cognition*, eds. F. Keil, and R. Wilson (Cambridge: MIT Press), 117–144.

Davidson, D. (1963). Actions, reasons and causes. *J. Philos.*, 60, 685–700. doi: 10.2307/2023177

Davidson, D. (1980). "Agency," in *Essays on Actions and Events*, ed. D. Davidson (Oxford: Clarendon Press), 43–62.

Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenom. Cogn. Sci.* 4, 429–452. doi: 10.1007/s11097-005-9002-y

Dickerson, K., Gerhardstein, P., and Moser, A. (2017). The role of the human mirror neuron system in supporting communication in a digital world. *Front. Psychol.*, 8, 1–6. doi: 10.3389/fpsyg.2017.00698

Encyclopedia Britannica (2023). *Mill's Methods*. Available online at: https://www.britannica.com/topic/Mills-methods (accessed January 6, 2023).

Fodor, J. (1987). *Psychosemantics: The Problem of Meaning in Philosophy of Mind*. Cambridge: MIT Press.

Ganeri, J. (2011). Emergentisms, ancient and modern. *Mind* 120, 671–703. doi: 10.1093/mind/fzr038

Haugeland, J. (1998). *Having Thought: Essays in the Metaphysics of Mind*. Cambridge: Harvard University Press.

Hidalgo, C. (2015). *Why Information Grows: The Evolution of Order, From Atoms to Economies*. New York: Basic Books.

Hornsby, J. (1997). *Simple Mindedness: A Defense of Naive Naturalism in the Philosophy of Mind*. Cambridge: Harvard University Press.

Kampourakis, K. (2020). Students' "teleological misconceptions" in evolution education: why the underlying design stance, not teleology per se, is the problem. *Evo. Edu. Outreach* 13, 1–12. doi: 10.1186/s12052-019-0116-z

Kant, I. (2000). *Critique of the Power of Judgment, translated by Guyer, P., and Matthew E.* Cambridge: Cambridge University Press.

Kant, I., and Bernard, J. H. (ed). (1790). *Critique of Judgment.* New York, NY: Barnes & Noble.

Keil, F. C. (2006). Explanation and understanding. *Annu. Rev. Psychol.*, 57, 227–254. doi: 10.1146/annurev.psych.57.102904.190100

Kim, J. (2006). Emergence: core ideas and issues. *Synthese* 151, 547–559. doi: 10.1007/s11229-006-9025-0

Kolchinsky, A., and Wolpert, D. H. (2018). Semantic information, autonomous agency, and nonequilibrium statistical physics. *Interface Focus* 8, 20180041. doi: 10.1098/rsfs.2018.0041

Maher, P. (2006). *Invariance and Laws*. Available online at: http://patrick.maher1.net/471/lectures/wood9.pdf (accessed April 4, 2023).

Matthews, E. (2002). *The Philosophy of Merleau-Ponty*. McLean, VA: Acumen.

Mertes, S., Huber, T., Weitz, K., Heimerl, A., and André, E. (2022). GANterfactual–Counterfactual explanations for medical non-experts using generative adversarial learning. *Front. Artif. Intell.* 5, 1–19. doi: 10.3389/frai.2022.825565

Monod, J. (1971). *Chance and Necessity: An Essay on the Metaphysics of Life*. New York: Vintage Books.

Mukherjee, S., Asnani, H., and Kannan, S. (2019). CCMI: Classifier based Q15 conditional mutual information estimation. *arXiv [Preprint]*. Available online at: https://arxiv.org/abs/1906.01824 (accessed October 6, 2023).

Prior, A., and Geffet, M. (2003). "Mutual information and semantic similarity as predictors of word association strength: Modulation by association type and semantic relation," in *Proceedings of Eurocogsci 03*, eds. F. Schmalhofer, R. M. Young, G. Katz, K. Graham (New York: Routledge).

Raninen, M. (2023). "Four ways of knowing – a semiotic interpretation," in *PhiloSign*. Available online at: https://philosign.substack.com/p/four-ways-of-knowing-a-semiotic-interpretation (accessed April 30, 2023).

Redish, J. (2019). "Huygens' principle," in *Nexus Physics*. Available online at: https://www.compadre.org/nexusph/course/Huygens'_principle (accessed April 5, 2023).

Rosen, G. (2010). "Metaphysical dependence: Grounding and reduction," in *Modality: Metaphysics, Logic, and Epistemology*, eds. B. Hale and A. Hoffmann (Oxford: Oxford University Press), 109–135.

Salazar, D. (2022). *Correlation is not Correlation*. Available online at: https://david-salazar.github.io/posts/fat-vs-thin-tails/2020-05-22-correlation-is-not-correlation.html (accessed April 28, 2023).

Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. New Jersey: Princeton University Press.

Sapolsky, R. M. (2017). *Behave: The Biology of Humans at Our Best and Worst*. New York: Penguin Books.

Schlosser, G. (2002). Modularity and the units of evolution. *Theory Biosci.* 121, 1–80. doi: 10.1078/1431-7613-00049

Schrödinger, E. (1944). *What is Life?*. Cambridge: Cambridge University Press.

Smolin, L. (2013). *Time Reborn: From the Crisis in Physics to the Future of the Universe*. Boston: Houghton Mifflin Harcourt.

Sommerhoff, G. (1950). *Systems Biology*. Oxford: Oxford University Press.

Stankovski, T., Ticcinelli, V., McClintock, P. V. E., and Stefanovska, A. (2015). Coupling functions in networks of oscillators. *New J. Phys.* 17, 1–12. doi: 10.1088/1367-2630/17/3/035002

Strogatz, S. H. (2000). From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators. *Phys. D: Nonlinear Phenom.* 143, 1–20. doi: 10.1016/S0167-2789(00)00094-4

Taleb, N. N. (2007). *The Black Swan: The Impact of the Highly Improbable*. New York: Penguin Books.

Taleb, N. N. (2020). *Statistical Consequences of Fat Tails: Real World Preasymptotics, Epistomology, and Applications (The Technical Incerto Collection)*. Pittsburgh: STEM Academic Press.

Taleb, N. N., Zalloua, P., Elbassioni, K., Henschel, A., and Platt, D. (2023). Informational rescaling of PCA maps with application to genetic distance. *arXiv [Preprint]*. Available online at: https://arxiv.org/abs/2303.12654 (accessed October 6, 2023).

Thomas Jefferson National Accelerator Facility (2023). "Charming experiment finds gluon mass in the proton: Experimental determination of the proton's gluonic gravitational form factors may have revealed part of proton's hidden mass," in *ScienceDaily*. Available online at: www.sciencedaily.com/releases/2023/03/230330102332.htm (accessed April 5, 2023).

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology and the Sciences of Mind*. Cambridge: Harvard University Press.

Velleman, D. (1992). What happens when someone acts? *Mind*101, 461–481. doi: 10.1093/mind/101.403.461

Vopson, M., and Lepadatu, S. (2022). Second law of information dynamics. *AIP Adv.* 12, 1–7. doi: 10.1063/5.0100358

Walsh, D. M. (2015). *Organisms, Agency, and Evolution*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781316402719

Weiner, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. Hoboken: Wiley.

Weisstein, E. W. (2023). "Covariance," in *MathWorld*. Available online at: https://mathworld.wolfram.com/Covariance.html (accessed April 21, 2023).

Woodward, J. (2000). Explanation and invariance in the special sciences. *Brit. J. Phil. Sci.*, 51, 197–254. doi: 10.1093/bjps/51.2.197

Woodward, J. (2003). *Making Things Happen*. Oxford: Oxford University Press. doi: 10.1093/0195155270.001.0001

Woolman, S. (2013). *The Selfless Constitution: Experimentalism and Flourishing as Foundations of South Africa's Basic Law*. Cape Town: Juta.

Ylikoski, P. (2013). Causal and constitutive explanation compared. *Erkenntnis* 78, 1–28. doi: 10.1007/s10670-013-9513-9

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 07 November 2023
DOI 10.3389/fgene.2023.1282515

OPEN ACCESS

EDITED BY
Go Yoshizawa,
Kwansei Gakuin University, Japan

REVIEWED BY
Bartha Maria Knoppers,
McGill University, Canada
Rispah Torrorey Sawe,
Moi University, Kenya

*CORRESPONDENCE
Faith Kabata,
faithkabata@gmail.com

RECEIVED 24 August 2023
ACCEPTED 16 October 2023
PUBLISHED 07 November 2023

CITATION
Kabata F and Thaldar D (2023), The human genome as the common heritage of humanity.
Front. Genet. 14:1282515.
doi: 10.3389/fgene.2023.1282515

The human genome as the common heritage of humanity

Faith Kabata[1]* and Donrich Thaldar[1,2]

[1]School of Law, University of KwaZulu-Natal, Durban, South Africa, [2]Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics, Harvard Law School, Cambridge, MA, United States

While debate on the international regulation of human genomic research remains unsettled, the Universal Declaration on the Human Genome and Human Rights, 1997 qualifies the human genome as "heritage of humankind" in a symbolic sense. Using document analysis this article assesses whether, how and to what extent the common heritage framework is relevant in regulation of human genomic research. The article traces the history of the Human Genome Project to reveal the international community's race against privatization of the human genome and its resulting qualification as the common heritage of humanity. Further, it reviews the archival records of UNESCO's International Bioethics Committee to discover the rationale for qualifying the human genome as common heritage of humankind. The article finds that the common heritage of mankind framework remains relevant to the application of the human genome at the collective level. However, the framework is at odds with the individual dimension of the human genome based on individual personality rights. The article thus argues that the right to benefit from scientific progress and its applications offers an alternative international regulatory framework for human genomic research.

KEYWORDS

human genome, common heritage of mankind, human genome project, human genomic research, human pangenome project, right to science

1 Introduction

Traditionally, jurists, politicians and scholars have invoked distributive and preservationist aims to decide that certain natural and cultural assets outside national territorial limits should be regulated under the common heritage of mankind framework (Wolfrum, 1983). The UNESCO Universal Declaration on the Human Genome and Human Rights (UNESCO, 1997) affirms the human genome as the heritage of humanity. This places the human genome in the category of outer space, the moon and other celestial bodies, as well as the deep seabed. These are all common resources regulated under the common heritage of mankind framework.

The qualification of the human genome as the common heritage of humanity is, however, so much more. It is the outcome of a titanic battle by the international scientific community between open scientific inquiry versus proprietary science; open and freely accessible data versus proprietary databases; and common resources versus private property. The archival records of the drafting of the Universal Declaration of the Human Genome and Human Rights (Universal Declaration on the Human Genome and Human Rights, 1948) confirm that the qualification of the human genome as the heritage of humanity was coined to emphasise the "need for equitable pooling" of scientific knowledge of the human genome to benefit all of humankind (Committee of governmental experts for the finalization of the declaration on the human genome, 1997).

In this article, using documentary analysis, we demonstrate the relevance of the common heritage of mankind framework to the human genome, and review its adequacy for regulating human genomic research after the Human Genome Project (HGP). Several scholars have noted the common heritage of mankind framework (Gorove, 1972; Wolfrum, 1983; Joyner, 1986). Therefore, there is no need to rehash their detailed analysis, and we rather assess whether, how and to what extent the common heritage of mankind framework is relevant to regulation of research on the human genome using a two-pronged approach. First, we trace and demonstrate its enduring relevance to human genomics research. Second, reflecting on the adequacy of the framework for regulating human genomics research after the HGP, we explore the tension between the framework and the individual dimension of the human genome. In conclusion, we suggest the right to enjoy the benefits of scientific progress and its applications as an alternative international regulatory framework.

The article flows as follows. In part 2 we focus on the meaning of the human genome—based on the Universal Declaration on the Human Genome. In part 3, building on the discussion in part 2, we demonstrate the relevance of the qualification of the human genome as a common heritage of mankind through the history of the HGP and analysis of the drafting records of the Universal Declaration on the Human Genome. After the HGP, we demonstrate the relevance of the framework in relation to the pangenome, and highlight its inadequacies in the context of the individual human genome. In the concluding part 4, we discuss the right to enjoy the benefits of scientific progress and its applications as an alternative regulatory framework.

2 What is the human genome?

Despite the prominence and frequent appearance of the term 'human genome' in human genomic research, there is little conceptual clarity on the meaning of the human genome. Acknowledging the different usages of the term, we confine ourselves to the meaning assigned by international instruments. The Universal Declaration on the Human Genome and Human Rights (Universal Declaration on the Human Genome) (UNESCO, 1997) does not in the body of the Declaration define the human genome. Nonetheless, the Explanatory Notes attached to the Declaration (UNESCO General Conference, 1997) define the human genome as "both to the full set of genes of each individual—in the twin senses of genetic material (DNA molecules) and genetic information—and to the entire range of genes which constitute the human race". Accordingly, the human genome is broadly the individual genome and the collective genome of the human species, and both genetic material and genetic data.

The broadness of this definition renders it meaningless. For instance, for the individual, his/her DNA and genetic information derived from the DNA constitute the human genome, while at the same time all the DNA and genetic information derived therefrom of the entire human species is the human genome. The human genome is thus the individual genome of each individual and also the collective genome of the entire human race. The UNESCO International Declaration on Human Genetic Data does not expressly mention the human genome, but elaborates on both genetic data and genetic material. These are referred to as biological samples, which are addressed as related concepts. This allusion to genetic data and biological samples embraces the individual dimension of the human genome.

Existing scholarship has offered some clarity. First, there is the human genome reference sequence which refers to a baseline map and a compound genome sequence of the human genome derived from the genetic data of several individuals that was generated by the HGP (National Human Genome Research Institute, 2022). The first draft of the human genome sequence was published in 2001, and later refined and updated in 2003 and 2010 (National Human Genome Research Institute, 2022). A complete sequence of the human reference genome, that which closed all the gaps, was released in 2022 (National Human Genome Research Institute, 2022). The definition of the human genome in the Universal Declaration on the Human Genome, the human genome reference sequence, represents the human genome at the collective level. This is because, as stated, it is a compound sequence generated from the genetic data of several individuals.

Second, on the individual dimension of the human genome, Thaldar et al in their analysis of the multidimensional legal nature of personal genomic sequence data offered some conceptual clarity (Thaldar, et al., 2022). The authors noted that personal genomic sequence data refers to individual genomic information that has been sequenced from DNA (Thaldar, et al., 2022).

In the next section we trace the application of the common heritage framework to the human genome and demonstrate its relevance in the regulation of research on the human genome.

3 The common heritage doctrine and regulation of the human genome

3.1 The human genome as "heritage of humankind"

The Universal Declaration on the Human Genome refers symbolically to the human genome as the heritage of humankind (UNESCO, 1997). While the term "heritage of humankind" had relatively little usage in international law, archival records indicate earlier drafts of the Universal Declaration on the Human Genome referred to the human genome as the common heritage of mankind. The International Bioethics Committee (IBC), in initial drafts, referred to the human genome as the common heritage of mankind, but the Committee of Governmental Experts dropped the term in favour of the "heritage of humankind in a symbolic sense" (International Bioethics Committee, 1996). Knoppers attributed these changes to differences among governmental representatives on the implication of the common heritage of mankind framework. Developing countries viewed the framework as allowing appropriation of the human genome by international companies, while the developed countries took a counter-position. They did not favour the communitarian aspect envisaged in the framework and were wary of state sovereignty, and thus preferred to protect the human genome at the individual level (Knoppers, 1999). According to the IBC, "heritage of mankind" was used to disabuse the notion that the human genome could be subjected to commercial appropriation (UNESCO General Conference, 1997).

Notwithstanding use of the term "heritage of humankind", in terms of international regulatory mechanisms, the UNESCO Declaration on the Human Genome qualifies the human genome as the common heritage of mankind.

3.2 The human genome as the common heritage of humankind: the relevance

3.2.1 Keeping the human genome in the public domain: race against privatisation

The qualification of the human genome as the common heritage of mankind traces to the HGP. The HGP is itself a story of many contrasts: international collaborative science versus national human genomic research initiatives; open access scientific inquiry versus private proprietary science; open freely accessible data versus proprietary databases; and the resulting human genome reference sequence as personal yet universal. These contrasts mirror the principles that embody the common heritage framework: national sovereignty versus international governance; sharing of benefits versus commercial principles; and common resource versus proprietary resource. These HGP contrasts sowed the seeds for the formulation of the 1996 Bermuda principles for free public access to the human sequence data and the qualification of the human genome as the common heritage of humanity.

The HGP was launched in 1990 as a 15-year international collaborative initiative involving a group of scientists from the United States, United Kingdom, France, Germany, Japan, Canada and China known as the International Human Genome Sequencing Consortium (Consortium, International Human Genome Sequencing, 2001). The primary objective of the HGP was to map, locate and sequence the human genome, with smaller affiliated projects involving the sequencing of the model genomes of the worm, fruit fly, yeast and mouse (Consortium, International Human Genome Sequencing, 2001). The primary goal was to generate a reference sequence of the human genome. By 1995, the HGP completed the first phase: the construction of the genetic and physical maps of the human genome (Consortium, International Human Genome Sequencing, 2001). The second phase was completed in 2001, and was marked by the release of the first draft of the human reference sequence of the human genome in February 2001 (Consortium, International Human Genome Sequencing, 2001). Updated drafts of the human reference sequence were released in 2003 and 2010, and the final complete version in 2022 (National Human Genome Research Institute, 2022). The results of the human sequence indicate that human beings are 99.9% similar, with the 0.1% accounting for genetic variance among individuals (Consortium, International Human Genome Sequencing, 2001).

However, the above account is the less debated part of the HGP story and does not account for its main legacies—the Bermuda principles for data sharing and qualification of the human genome as the common heritage of humanity. First, it should be noted that national and private enterprise endeavours to sequence the human genome predate the 1990 launch of the HGP. Prior national initiatives included: the USA's human genome research under the Office of the Human Genome Research; the 1981 Japan's Science and Technology Project which aimed to convert genome sequencing into a large-scale project; the French Centre d'Etude du Polymorphisme Humain, established in 1994 with the goal of creating genetic maps of all chromosomes in the human genome; and the UK's Medical Research Council set up in 1988 to coordinate mapping and sequencing of the human genome (Raggio, 2002). Private enterprise endeavours included the Genome Corporation which in 1987 announced plans to sequence the human genome and to commercialise the data (Raggio, 2002). While the national human genomic research interests were mediated in the HGP, threats to commercialise the human sequence data by private enterprise endeavours overshadowed the HGP throughout its life.

Controversy over the public or private nature of the human genome first arose at the beginning of the HGP in 1991 when the National Institutes of Health (NIH) filed 337 patent applications for gene fragments sequenced by Venter, who was then a scientist at the NIH (Eisenberg and Nelson, 2002). The international collaborators of the HGP considered these patent applications by the NIH to be in contradistinction with the primary objectives of the HGP—to sequence the heritage of humanity (Eisenberg and Nelson, 2002). For instance, the French National Consultative Committee on Ethics condemned the patent applications and indicated that the information contained in the human genome was part of common heritage of humanity, and hence could not be monopolised (Dworkin, 1997). In 1992, Venter left the NIH to set up the non-profit Institute for Genomic Research, which was affiliated to the private firm Human Genome Sciences (Mukherjee, 2016; Cook-Deegan, et al., 2017). Human Genome Sciences backed Venter's earlier work on expressed sequence tags and established proprietary databases on these gene sequences which locked out access to researchers in academic institutions (Eisenberg and Nelson, 2002; Cook-Deegan, et al., 2017).

These concerns over the "gold rush" to privatise the human genome were part of the agenda of the 1996 Bermuda meeting captured as patenting of the human genome and data sharing (Cook-Deegan, et al., 2017). The session on data sharing noted: "The fact is that we'd come to realize that the genomic sequence we are producing and dealing with is more than a commodity. It is the essence of biological heritage, the instruction book of living things. The only reasonable way of dealing with the human genome sequence is to say that it belongs to us all—it is the common heritage of humankind" (Bradley, 2005). The final statement from the session read: "It was agreed that all human genomic sequence information generated by centers for large-scale human sequencing, should be freely available and in the public domain, in order to encourage further research and development, and to maximise its benefit to society" (Cook-Deegan, et al., 2017). This data sharing agreement on daily online release of human sequences under the HGP is referred to as the Bermuda principles for data sharing (Cook-Deegan, et al., 2017).

Even after the Bermuda principles, the race to keep the human genome from private enterprise was far from over. In 1998, as the HGP was embarking on sequencing the human genome, Venter broke away from the HGP and established Celera Genomics, a new private company. Under Celera Genomics, Venter announced plans to sequence the human genome using a faster methodology and more cheaply, and aimed to complete the sequencing within three years—four years ahead of the HGP—and establish commercial proprietary databases (García-Sancho, et al., 2022). In addition,

contrary to the Bermuda principles on daily data release, Venter indicated that Celera Genomics would release data every 3 months (Eisenberg and Nelson, 2002; Jasny, 2013). Challenged by the HGP that quarterly release was contrary to the Bermuda principles, Venter retorted: "[w]e are a company . . . We do not have to release the data at all. But if you think about it, quarterly is a lot closer to nightly than it is to never" (Jasny, 2013).

This announcement by Celera Genomics began the most polarising race between the HGP and private enterprise. The implication of the announcement was that it put the utility of the HGP into question, as it suggested that sequencing could be achieved faster and more cheaply and threatened to forever put the human genome in the hands of private enterprise. In response, at the technical level, the HGP revised its strategy: it requested more funding to speed up the sequencing and importantly shifted the priority to producing a 'rough draft' of the human genome by 2000 rather than a complete sequence in 2005 (Eisenberg and Nelson, 2002; Raggio, 2002). At the political level, the United States government pre-empted the race by issuing a joint statement by the US President and United Kingdom Prime Minister in March 2000, which declared that the human sequence DNA should be made freely available to all scientists across the globe (Raggio, 2002). In June 2000, the US President and United Kingdom Prime Minister also presided over a joint release by the HGP and Celera Genomics of the 'rough draft' of the human genome reference sequence (Consortium, International Human Genome Sequencing, 2001), which effectively ended 'the race'. In line with the Bermuda principles, the HGP released its data on the human reference sequence in *Nature* in February 2001, while Celera Genomics published its sequence in *Science* a day later, although with some restrictions to full access. However, contentions abound on the quality of the human genome reference sequence released by Celera Genomics based on the methodology of sequencing and claims that it benefited from the HGP data to generate its own human reference sequence (Waterston, et al., 2002).

The enduring legacy of the HGP was keeping the human genome in the public domain through the daily data release policies and the common heritage of humanity qualification. A key observation from the foregoing is that for the international sequencing community in the HGP, the essence of the human genome as the common heritage of humankind was to protect and promote freedom of research in the scientific community. The race was thus between open and free scientific inquiry versus private proprietary science, and between open and freely accessible data versus proprietary databases.

UNESCO waged an equivalent race to keep the human genome in the public domain. In 1997, the IBC, comprising scientists and legal scholars, affirmed the human genome as the common heritage of humanity in the Universal Declaration on the Human Genome. In at least two ways UNESCO's qualification of the human genome as the common heritage of humanity coincided with events at the HGP. First, Knoppers alluded to the fact that there were already proposals as far back as 1991 to declare the human genome, at the collective level, as the common heritage of humanity (Knoppers, 1999). As noted earlier, the initial attempt at national privatisation of the human genome was in 1991 when the NIH filed for patents for gene fragments of brain cells. The concern then was both for scientists and for other states that had foregone their national genomic research initiatives for the collaborative HGP. Second, the Universal Declaration on the Human Genome drew from the spirit of the 1996 Bermuda principles. As alluded to earlier, the Bermuda meeting stated that "the human genome belongs to us all", and hence it is plausible to link the UNESCO affirmation with the position taken by the international sequencing community.

In addition, archival records of the IBC's discussions reveal that UNESCO was concerned with national appropriation of the human genome by developed countries. This concern was well founded. Besides the USA's 1991 attempt to patent gene fragments, the composition of countries that participated in the HGP validated this concern. According to the HGP architects, the HGP was founded on the principle of inclusivity as the genome was the common heritage of all humankind, and thus any nation could participate by opening mapping and sequencing centres (Waterston, et al., 2002). The HGP participating countries were the United States, United Kingdom, Canada, France, Germany, Japan, China and the European Community—the most developed countries with the technological infrastructure and human capacity. Illustratively, the then Chair of IBC, Mohammed Bedjaoui, stated that, "any advance in knowledge on the human genome must benefit mankind as a whole . . . a common heritage regime is mindful of the inequalities in the development of various regions in the world" (Kuppuswamy, 2009). Therefore, UNESCO was also waging a war against bio-colonialism: the appropriation of the human genome by developed countries without any benefits accruing to developing countries. The common heritage of humankind framework was deployed for regulation of the human genome as a resource belonging to all of humanity. Significantly, while the HGP's main motivation for qualifying the human genome as the common heritage of humanity was freedom of research, UNESCO privileged equity, social justice, and benefit sharing from the Global North-South perspectives. These ideals find expression in the provisions of the Universal Declaration on the Human Genome (UNESCO, 1997) on the regulation of the human genome.

3.2.2 The human genome as the common heritage of mankind

The moon, outer space and the deep sea bed and ocean floor are all considered common resources, which are outside territorial limits of national jurisdiction and from which no person should be excluded and for which there should be no individual or government appropriation. However, they should rather be publicly regulated to distribute the benefits and preserve them for future generations. Consequently, these common resources are internationally regulated under the common heritage of mankind framework. Significantly, unlike the common property doctrine, the common heritage of mankind framework requires that all manage the resources and share in the benefits, including those who do not participate in the exploitation of the resources (Noyes, 2011). What then does it mean for the human genome as the common heritage of humanity?

As discussed in the foregoing, HGP's two principles on human genome sequencing embedded the common heritage of humanity doctrine. First, the collaborative nature of the HGP was informed by the universal nature of the human genome sequence as the common

heritage of humanity transcending national territorial limits (Waterston, et al., 2002). Second, the principle of unrestricted data release as articulated in the 1996 Bermuda principles was founded on the idea that the human genome sequence 'belongs to us all', and that it is a common resource (Waterston, et al., 2002). Furthermore, the HGP was conceived as a universal project to generate the human reference sequence, and therefore it sequenced the individual DNA of a diverse group of anonymous individuals who maintained no further association with the data (Contreras and Knoppers, 2018a). Taken together, the conceptualisation and nature of the HGP pointed to humanity's collective ownership of the human reference sequence. Knoppers and Beauvais have noted that given the nature of the HGP, the human reference sequence was a common resource for humanity that could not be controlled by an individual or private enterprises or government (Knoppers and Beauvais, 2021). The Universal Declaration on the Human Genome defines the human genome as both the individual and collective genome. Records of the drafting sessions of the IBC reveal that in applying the common heritage doctrine in the field of genetics, the IBC indicated that the aim was "safeguarding the integrity of the human species" (International Bioethics Committee, 1996).

A consideration of the above context reveals that the common heritage of mankind doctrine applies to the human genome at the collective level: the human species level. The individual genome thus does not qualify as the common heritage of mankind. Knoppers has alluded to the application of the common heritage of mankind framework to the human genome at the human species level (Knoppers, 2005; Knoppers and Joly, 2007). Applying the common heritage of humankind doctrine to the human genome, two questions arise. What common resources does the doctrine apply to? And what are its elements? On the first question, the doctrine applies to areas outside the territorial limits of states and to the natural resources in those areas (Noyes, 2011). The human genome at the human species level refers to the human reference sequence of humanity, which embodies the universality of the human species unbounded by state territorial limits. It thus qualifies as common resources beyond state territorial limits. On the second question on the elements of the doctrine, although unsettled, consensus exists on the following: (i) a ban on the acquisition of or exercise of sovereignty over the resources; (ii) rights over the resources vest in humankind; (iii) equitable sharing of benefits derived from exploitation of the resource, with particular consideration of the needs of developing states; (iv) common management of the resources; (v) use of the resources for peaceful purposes; and (vi) protection of the environment (Wolfrum, 1983; Noyes, 2011).

Reflecting on the human genome at the collective level, the first four elements noted above are important. On the ban on acquisition of or exercise of sovereignty, the HGP by its very nature, conceptualisation and coordination, as discussed earlier, ensured that the human reference sequence was not appropriated by national states, individuals or corporations. The open and free release of the human reference sequence put the resource in the hands of humanity, from which no entity could be excluded and no entity could claim exclusive control. It then follows that the human reference sequence belongs to all of humanity and humanity has rights over its use and disposal. On the element of sharing of the benefits derived from exploitation of the resources, Wolfrum and Noyes have pointed out its controversial nature (Wolfrum, 1983; Noyes, 2011). The controversy on sharing of the benefits mainly arises from the assertion that this includes preferential treatment for developing states (Noyes, 2011). Discussing the application of the doctrine to the seabed and ocean floor, Wolfrum argued that since all states participate equally, directly or indirectly, in the exploitation of the seabed minerals, the idea of preferential treatment was discarded (Wolfrum, 1983). In addition, a question may be asked about the scope of the shared benefits, and whether it includes the results of scientific research. Viewed from the actual implementation of the common heritage of mankind framework in the law of the sea regime, scientific research results fall within the scope of shared benefits, while preferential treatment in the distribution of benefits for developing states was subjected to market principles (United Nations, 1994). Finally, common management of the resource is anchored on the idea that humankind is vested with rights over control of the resources, where an international entity or forms of cooperative arrangements would be required to act at the instance of humankind (Noyes, 2011). In relation to the human genome, the scientific results of the human reference sequence are available for all and to that extent the element of benefit sharing seems to hold. However, the human sequence as generated by the HGP is a reference map resource, and any health benefits that accrue to humankind would be the result of further scientific research.

3.2.3 After the HGP: the common heritage of humankind framework and the human pangenome reference project

Beyond the HGP, does the common heritage framework have any relevance? As demonstrated above, the main legacies of the HGP were the human reference sequence and the Bermuda principles for data sharing. After the HGP, scholarship has identified concerns in human genomic research. An editorial in *Nature* in February 2021 identified the enduring concerns as: ethical and legal issues such as privacy and consent; representation of both data contributors and users; and challenges in implementation of access to genome data (Nature Editorials, 2021). The editorial provided further elaboration of the concerns as: data collection from the participants; data deposits in publicly accessible and approved databases; and data access (Nature Editorials, 2021). Similarly, Knoppers, Contreras and Cook-Deegan et al. note that after the HGP, ethical, legal and technical issues such as the protection of individual data and researchers' publication priority have chipped away the expanse of data sharing envisioned in the Bermuda principles (Contreras, 2011; Cook-Deegan, et al., 2017; Contreras and Knoppers, 2018b).

In relation to the diversity deficit in the human reference sequence, the human pangenome reference project was initiated in 2019 under the Human Pangenome Reference Consortium and is expected to sequence, assemble and freely share the human pangenome reference which will correctly reflect the diversity of the human species (Miga and Wang, 2021; Liao, 2023). The human pangenome reference project is similar to the HGP in that it is an international collaborative science initiative and involves sequencing DNA from 350 individuals of diverse ethnic backgrounds to create a baseline reference sequence (Miga and Wang, 2021). It is therefore a community resource project aimed at

generating reference data for human genomic research. The first draft of the human pangenome reference sequence was released in May 2023 and consists of 47 sequenced and assembled diverse individual human genomes which feature the diversity within the human species (Liao, 2023). The Human Pangenome Reference Consortium will increase the number of individual human genomes sequenced and assembled to 350 individuals by 2024 (Liao, 2023). Even then, the pangenome project is not without criticism as to its diversity and inclusiveness. It has been pointed out that the pangenome project appears focused on numerics without proper consideration of the communities and nations to collaborate with in order to address the diversity deficit (Cho, et al., 2023). Unlike the human reference sequence under the HGP, with the pangenome sequence the researchers indicated that consent was obtained from 47 individuals for the release of the draft pangenome human sequence (Liao, 2023). The implicit question is whether the common heritage of mankind framework is relevant to the human pangenome sequence.

While there has been a narrowing of the original scope of data sharing under the Bermuda principles, Knoppers and Contreras have noted that the Bermuda principles apply to community resource projects, that is research aimed at generating data for use by the scientific research community (Contreras and Knoppers, 2018a). This position is also affirmed by Cook-Deegan et al., who noted that the data sharing obligations of research projects aimed at generating community resources remained governed by the Bermuda principles, despite a watering down of obligations for hypothesis-focused research (Cook-Deegan, et al., 2017). Therefore, given that the human pangenome sequence project aims to generate data for the scientific community, drawing from the HGP approach, the common heritage framework can apply to the human pangenome reference at the collective level—the human species level.

On the future of the application of the common heritage framework in human genomic research, it is notable that up to now under international law the framework has been implemented only in the law of the sea regime. And, as alluded to earlier, what was operationalised and implemented is a diluted version of the framework, in particular with regard to the sharing of benefits. Noyes, while discussing the application of the common heritage of mankind framework to other common resources besides the seabed minerals, noted that sharing of benefits and common resource management are the most contested elements, because of the finite nature of the resources. Furthermore, the likelihood of extending the framework to other common resources would require redefining the elements of the framework (Noyes, 2011).

Even beyond these general contestations, in human genomics research the conceptual underpinnings of the common heritage of mankind framework present important considerations. These include: it is associated with natural resources, leading to the question of whether the human genome, in particular in its individual dimension, can be considered a natural resource; human genomic data is infinite, and hence the problem of a depletion of resources fear that characterizes the common heritage of mankind framework does not apply; and the preservation ethic aimed at conserving the resources, particularly given that the human genome evolves. These considerations resonate with the ethical, legal and social concerns identified above in human genomic research after HGP.

We now explore these inadequacies of the common heritage of mankind framework from the individual dimension of the human genome.

3.3 Common heritage of mankind framework, individual rights and species preservation

The underpinnings of the common heritage of mankind framework appear to be at odds with the individual dimension of the human genome. First, the notion of a common resource under common heritage raises the following questions: Can individuals, genes and genetic information in the individual genome be considered a common resource; and can the common heritage doctrine be reconciled with individual personality and property rights inherent in genomic resources? Second, the preservationist bias that underpins the common heritage doctrine also raises questions about individual rights such as the right to health, life and to enjoy the benefits of scientific progress and its applications.

3.3.1 Individual personality rights and the "heritage of species"

Thaldar et al., in their analysis of the multidimensional legal nature of personal genomic sequence data, identified individual personality rights in the data as: personal integrity, respect of a person's identity and informational privacy (Thaldar, et al., 2022). The authors also noted that personality rights attach to the individual and cannot be lost. Furthermore, individual personality rights take precedence over any property rights or claims that may be made in relation to the data (Thaldar, et al., 2022). The right to informational privacy entails control over use, access and processing of personal genomic sequence data (Thaldar, et al., 2022). Tied to this is the notion of informational self-determination which gives the individual sovereignty and control over their data (Hummel, et al., 2019). The right to personal identity entails the right of an individual to construct a life narrative of themselves based on what they consider important (De Andrade, 2010). Implicit in this right is the right to individual data sovereignty by controlling its use and processing.

As discussed, the common heritage of mankind framework regulates common resources and is relevant for the human genome in its collective dimension. However, in relation to the individual human genome, as Thaldar et al have noted, individual personality rights take precedence (Thaldar, et al., 2022). The common heritage of mankind framework based on its patrimonial foundations cannot be reconciled with individual personality rights that arise in relation to the individual dimension of the human genome (De Andrade, 2010). The UNESCO Declaration on Human Genetic Data embraces the individual personality rights as it refers consent for collection and use of genetic data and biological samples to the individual (UNESCO, 2003). The Declaration on Human Genetic Data is proclamatory, and thus has no legally binding obligations on states. Rather, it defers the protection of individual personality rights to states.

Flowing from the Declaration on Human Genetic Data, states have put in place mechanisms for the protection of individual personality, including privacy, informational self-determination and respect for personal identity in the context of human genomic research. Equally, states have an obligation under the Covenant on Economic, Social and Cultural Rights to guarantee the right to enjoyment of the benefits of scientific progress and its applications. This takes a cue from the indivisibility of human rights, individual personality rights and the right to the benefit of scientific progress and its applications which are interdependent and interconnected, and no right should take precedence over another. Knoppers and Beauvais noted that enjoyment of the right to the benefits of scientific progress and its applications is premised on data sharing, which invokes individual personality rights as individuals exercise informational self-determination by deciding which data to share or control (Knoppers and Beauvais, 2021). Therefore, states in their obligations to guarantee the right to enjoyment of the benefits of scientific progress and its applications must put in place a regulatory framework that ensures respect for privacy and individual genetic data control, including the right to one's personal identity based on their genetic data.

3.3.2 Individual property rights

As noted earlier, Thaldar et al. noted that personal genomic sequence data can be owned privately, can be public property and also can be common resources under the common heritage of mankind framework (Thaldar, et al., 2022). In addition, the authors posited that since personal genomic sequence data is generated from DNA sequencing, a number of entities, including the research institutions and funders, may lay a claim of ownership (Thaldar, et al., 2022). Furthermore, they suggested that an entity can acquire ownership of personal genomic sequence data through appropriation if it has effective control of the data as a digital object (Thaldar, et al., 2022). However, ownership rights in personal genomic sequence data are subjected to the individual personality rights of the data subject (Thaldar, et al., 2022). In essence, in relation to the individual human genome, the individual has certain entitlements in their personal genomic sequence data, based on personality rights that trump ownership rights.

Therefore, in relation to the individual human genome, while the person or entity in control of the personal genomic sequence data may claim ownership, the personality rights of the research participant limit such ownership. In the context of exercise of individual personality rights in genomic research, the right to informational self-determination would entitle the research participant to control use of and access to their data.

3.3.3 Preservation of the human genome: safeguarding species integrity and the natural evolution bias

In qualifying the human genome as the common heritage of mankind framework, the IBC was motivated to "safeguarding the integrity of the human species" (International Bioethics Committee, 1996). In addition, the IBC took note of the natural evolution of the human genome ascribing to the idea that natural evolution is responsible for the human genome (UNESCO, 1997). The HGP sequence of the human genome revealed that 50% of the human genes are similar to the genes of other model organisms sequenced such as the worm, fruit fly and mouse, thus displacing any special expectations on the human genome (Goes, 2016). This questioned the claim of specialty of the human species and the idea of the species barrier. Harris has argued that the claim of integrity of the human species does not hold. First, he posited that claims of maintaining a species barrier between the human person and non-humans overlook the fact that through diet, drugs, vaccines and xenotransplantation, exchange of biological material often occurs between the human person and animals (Harris, 2011). Harris observed that these instances which involve mixing of the biological matter from animals to the human person are not frowned upon as an interference with the purity of human species (Harris, 2011). Based on this, he questioned barring scientific interventions in the human genome to safeguard the integrity of the human species, while the above practices that involve mixing of human and non-human genes are acceptable. Second, Harris noted that based on evolution theory, the genetic makeup of the human person includes genes from all other creatures that the person has over time evolved from (Harris, 2011). Based on the evolution process, the argument on purity of the human species is flawed and, therefore, there is no basis for safeguarding the integrity of the human species.

Similarly, Knoppers and Joly noted that the idea of safeguarding the integrity of the human species is based on the preference for maintenance of the natural order over scientific interventions, which are viewed as interfering with the purity of the human species (Knoppers and Joly, 2007). They argued that appeals to purity of the human species should not be a justification for human persons not to benefit from scientific progress (Knoppers and Joly, 2007). On the natural order, Knoppers and Joly have called for a reconceptualisation of what is considered natural and a shift away from viewing the human person with the naturalism lens (Knoppers and Joly, 2007). According to Harris, the preservationist ethic embedded in the common heritage of mankind framework is flawed as it ignores that natural human reproduction already changes the human genome, and therefore the human genome cannot be considered as frozen in time (Harris, 2015).

Harris also argued against UNESCO's bias towards the natural order, noting that natural evolution is slow and does not guarantee improvement of the human species, while scientific progress would guarantee improvements in the quality of health and life of humankind (Harris, 2015). He pointed out that the bias towards natural evolution is premised on the wrong assumptions that the natural order is good and not capable of improvement and that natural evolution enhances the human genome for the better (Harris and Søren, 2002). Ultimately, the bias to the natural order impedes the enjoyment of various individual rights as individuals cannot benefit from scientific progress and also exercise informational self-determination.

A common theme in the arguments of Harris and Knoppers and Joly is the effect of the absolute construction of individual personality rights such as privacy, autonomy and human dignity on the right to enjoy the benefits of scientific progress and its applications. Knoppers and Joly have questioned the invocation of human dignity concerns as a bar to application of scientific inventions to human beings (Knoppers and Joly, 2007). The authors have noted that individuals enjoy human dignity by virtue of personhood, and that personhood is not diminished by

application of scientific interventions on the human body (Knoppers and Joly, 2007). Supporting this proposition, Harris drew from scientific experiments in the 1990s which involved interventions from animals to humans—noting that mixing of genes does not change the characteristics of a species (Harris, 2011). Thus, personhood and the human dignity that attaches to personhood is not lost through scientific interventions. In the same line of argument on human dignity, Jordaan, writing on stem cell research in the Brüstle case before the European Court of Justice, observed that human dignity attaches to human beings, but is often deployed as a mask for abstract claims anchored in morality (Jordaan, 2017). Knoppers and Joly and Jordaan raised related concerns on how human dignity should be deployed in relation to genomic research, should it promote a conception of dignity that attaches to abstract humanity or to real personhood. For Knoppers and Joly, the question is whether it should be invoked to promote species purity rather than to advance the right to health and life. Jordaan criticised the invocation of human dignity to abstract embryos (which in many jurisdictions are not considered as human beings), rather than invoking human dignity to advance the right to health. Generally, on human rights, Harris has taken a more blunt view and posited that the concept of humanness should give primacy to the powers and capacities that improve the quality of existence of the human person, rather than deploying human rights as an obstacle to scientific interventions on the human genome (Harris, 2011).

In sum, as noted earlier, human rights are indivisible, and enjoyment of one set of rights does not curtail the enjoyment of other rights. The individual personality rights and the right to enjoy the benefits of scientific progress and its applications are not absolute. There is therefore room for proportionality in the exercise of each set of rights to ensure that there is no hierarchy in rights but instead an interdependence of rights. In addition, on human dignity in human genomic research, the notion of human dignity as a mask for morality and the natural order should be discarded.

4 Conclusion

The article demonstrates the relevance of the common heritage of mankind framework to the human reference sequence and also to the pangenome reference sequence at the collective level. The qualification of the human reference sequence as a common resource open and freely accessible to all humanity, provided a framework for collaboration in sharing of genomic data within the scientific community, hence facilitating the realisation of the right to freedom of research. As noted earlier, the enduring concerns in human genomic research after the HGP are: data collection from participants; depositing data in publicly accessible and approved databases; and data access (Nature Editorials, 2021). And as demonstrated by the article, the common heritage framework is at odds with the individual rights, putting into question the extent to which the framework is able to protect the rights and interests of the diverse stakeholders in human genomic research. Thorogood et al. identified the rights and interests of the different stakeholders as: recognition of data generators; interests of data users in accessing data; rights of participants to benefit from the research and to protection of their data rights (Thorogood, et al., 2015). The authors proposed that the international human rights law framework is best suited for bringing together the multiple interests and rights involved as it is universal and transcends state borders, it has legal and political binding force and it imposes obligations beyond the scientific community to states, private actors and protects the rights of individuals (Thorogood, et al., 2015). Specifically, the authors discussed the right to enjoy the benefits of scientific progress and its applications (right to science) as a possible framework for human genomic research data sharing (Thorogood, et al., 2015).

Drawing from these insights, the article highlights the relevance of the right to science as an alternative framework for sharing of genomic data. The UN Committee on Economic, Social and Cultural Rights elaborated on the right to science in General Comment No. 25 (CESCR, 2020). In the context of human genomic research data sharing, the right to science imposes obligations on states and the international community to safeguard the rights of data generators, data users and research participants. In relation to recognition of data generators, the right requires states to ensure that contractual arrangements provide appropriate crediting and acknowledgment of the contributions of scientific researchers to research outcomes as a consequence of the right to freedom of research. For data users, states have a duty to facilitate international cooperation that enables researchers to freely share data and collaborate internationally. For research participants, the right imposes an obligation on states to ensure access of their population to health benefits that accrue from human genomic research, including fostering a positive balance with intellectual property, as well as adopt normative standards for the protection of privacy and data rights and human dignity (CESCR, 2020).

For actualization of the right, beyond states' implementation of the above discussed obligations, there is a need to reframe the perception of the right. On implementation of state obligations, states should put in place regulatory frameworks that ensure the respect and protection of the rights of individual genomic data in the context of human genomics research. An additional obligation is to conduct public education to promote participation of individuals in the advancement of science, in particular through data sharing. On reframing the perception of the right to enjoy the benefits of scientific progress and its applications, the idea is to relook at implementation of the right in relation to individual personality rights that seek to protect the human person as an autonomous individual. Practices in human genomics research have mainly focused on protecting individuals from harm that would be associated with research, and hence the dominance of rights protecting privacy, human dignity, informational self-determination and identity. However, the aspect of the benefits that accrue from science and its applications appears to be neglected. A reframing of the right to emphasise the benefits dimension will also result in a shift from the absolute construction of individual personality rights in human genomic research to a construction that allows for their interdependence with the right to enjoy the benefits of scientific progress and its applications. In part, there is also a need for appreciation of the right to science as a collective endeavour, in that in genomics research the benefits of science result from human solidarity rather than absolute notions of individual autonomy. Finally, in relation

to human dignity, while it is a fluid concept, given its obsessive repetition in the primary instruments regulating the human genome (UNESCO, 1997; UNESCO, 2005), there is room to define its scope and contours in relation to human genomic research. Currently, as noted above, human dignity has been invoked to mask claims of morality, instead of invoking it to promote enhancement of the human genome that promotes human dignity of individuals and humanity as a whole.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, and further inquiries can be directed to the corresponding author.

Author contributions

FK: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing–original draft, Writing–review and editing. DT: Conceptualization, Formal Analysis, Funding acquisition, Investigation, Methodology, Resources, Supervision, Validation, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Work on this article was supported by the U.S. National Institute of Mental Health and the U.S. National Institutes of Health (award number U01MH127690) under the Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) program. The content of this article is solely our responsibility and does not necessarily represent the official views of the U.S. National Institute of Mental Health or the U.S. National Institutes of Health.

Acknowledgments

We are grateful to Bonginkosi Shozi, Freddy Mnyongani, Peter Munyi, Claude Kamau, and Adrian Bellengère for their insightful comments in developing the article. Any errors are our own.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Bradley, P. (2005). DASH. Available at: https://dash.harvard.edu/handle/1/10015275 (Accessed October 7, 2023).

CESCR (2020). United Nations Human Rights general comment no. 25 (2020) on article 15: science and economic, social and cultural rights. Available at: https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2020-article-15-science-and (Accessed February 27, 2023).

Cho, M. K., Fullerton, S. M., Hammonds, E. M., Lee, S. S., Panofsky, A., and Reardon, J. (2023). Pangenomics: prioritize diversity in collaborations. *Nature* 619, 698. doi:10.1038/d41586-023-02248-7

Committee, International Bioethics (1995). *Report of the second session proceedings*. Paris: UNESCO.

Committee of governmental experts for the finalization of the declaration on the Human Genome (1997). *Universal Declaration on the Human Genome and Human Rights (20 December 1996)*. Paris: UNESCO.

Consortium, International Human Genome Sequencing (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921. doi:10.1038/35057062

Contreras, J. L. (2011). Bermuda's Legacy: policy, patents, and the design of the genome commons. *Minn. J. Law, Sci. &Technology* 12 (1), 61–125.

Contreras, J. L., and Knoppers, B. M. (2018a). The genomic commons. *Annu. Rev.* 19 (1), 429–453. doi:10.1146/annurev-genom-083117-021552

Contreras, J. L., and Knoppers, B. M. (2018b). The genomic commons. *Annu. Rev. Genom. Hum. Genet.* 19, 429–453. doi:10.1146/annurev-genom-083117-021552

Cook-Deegan, R., Rachel, A. A., and Jones, K. M. (2017). Sharing data to build a medical information commons: from Bermuda to the global alliance. *Annu. Rev. Genom. Hum. Genet.* 18, 389–415. doi:10.1146/annurev-genom-083115-022515

De Andrade, G. N. N. (2010). Human genetic manipulation and the right to identity: the contradictions of human rights law in regulating the human genome. *SCRIPTed* 7 (3), 429–452. doi:10.2966/scrip.070310.429

Dworkin, G. (1997). Should there be property rights in genes? *Phil. Trans. R. Soc. Lond. B* 352 (1357), 1077–1086. doi:10.1098/rstb.1997.0088

Editorials, N. (2021). Twenty years of the human genome. *Nature* 590, 183–184.

Eisenberg, R. S., and Nelson, R. R. (2002). Public vs. Proprietary science: a fruitful tension? *Daedalus* 131 (2), 1392–1399. doi:10.1097/00001888-200212001-00011

García-Sancho, M., Rhodri, L., Gil, V., Wong, M., Vermeulen, N., and Lowe, J. (2022). The human genome project as a singular episode in the history of genomics. *Hist. Stud. Nat. Sci.* 52 (3), 320–360. doi:10.1525/hsns.2022.52.3.320

Goes, F. M. A., and Góes, A. (2016). A decade of human genome project conclusion: scientific diffusion about our genome knowledge. *Biochem. Mol. Biol. Educ.* 44 (3), 215–223. doi:10.1002/bmb.20952

Gorove, S. (1972). The Concept of "common heritage of mankind doctrine": a Political, Moral or Legal Innovation. *San. Diego Law Rev.* 9 (3), 390–403.

Harris, J. (2011). Taking the "human" out of human rights. *Camb. Q. Healthc. Ethics* 20, 9–20. doi:10.1017/S0963180109990570

Harris, J. (2015). Germline manipulation and our future worlds. *Am. J. Bioeth.* 15 (12), 30–34. doi:10.1080/15265161.2015.1104163

Harris, J. (2016). Germline modification and the burden of human existence. *Camb. Q. Health Care Ethics* 25, 6–18. doi:10.1017/S0963180115000237

Harris, J., and Holm, S. (2002). Extending human lifespan and the precautionary paradox. *J. Med. Philosophy* 27, 355–368. doi:10.1076/jmep.27.3.355.2983

Hummel, P., Braun, M., and Dabrock, P. (2019). "Data donations as exercises of sovereignty," in *The ethics of medical data donation*. Editors K. Jenny and F. Luciano (Cham: Springer Open), 23–54.

International Bioethics Committee (1996). *Seventh meeting of the Legal Commission of the International Bioethics Committee*. Paris: UNESCO.

Jasny, B. R. (2013). Realities of data sharing using the genome wars as case study: an historical perspective and commentary. *EPJ Data Sci.* 2 (1), 1–15. doi:10.1140/epjds13

Jordaan, D. (2017). Stem Cell Research, Morality and Law: an analysis of *Brüstle v Greenpeace* from a South African Perspective. *South Afr. J. Hum. Rights* 33, 429–451. doi:10.1080/02587203.2017.1392839

Joyner, C. C. (1986). Legal implications of the concept of the common heritage of mankind. *Int. Comp. Law Q.* 35 (1), 190–199. doi:10.1093/iclqaj/35.1.190

Knoppers, B. M. (1999). "Biotechnology," in *Commercialisation of genetic research: ethical, legal and policy issues*. Editors A. C. Timothy and W.-J. Byrn (New York: Kluwer Academic), 1-11.

Knoppers, B. M. (2005). Of Genomics and Public Health: building public "goods"? *CMAJ* 173 (10), 1185–1186. doi:10.1503/cmaj.050325

Knoppers, B. M., and Chadwick, R. (2005). Human Genetic Research: emerging trends in ethics. *Nat. Rev. Genet.* 6, 75–79. doi:10.1038/nrg1505

Knoppers, B. M., Harris, J., Budin-Ljosne, I., and Dove, E. S. (2014). A human rights approach to an international code of conduct for genomic and clinical data sharing. *Hum. Genet.* 133 (7), 895–903. doi:10.1007/s00439-014-1432-6

Knoppers, B. M., and Joly, Y. (2007). Our social genome? *TRENDS Biotechnol.* 25 (7), 284–288. doi:10.1016/j.tibtech.2007.04.006

Knoppers, B. M., and Michael, J. S. B. (2021). Three decades of genetic privacy: a metaphoric journey. *Hum. Mol. Genet.* 30 (20), 156–160. doi:10.1093/hmg/ddab164

Kuppuswamy, C. (2009). *The international legal governance of the human genome*. 1. London: Routledge.

Liao, W.-W., Asri, M., Ebler, J., Doerr, D., Haukness, M., Hickey, G., et al. (2023). A draft human pangenome reference. *Nature* 617 (2023), 312–324. doi:10.1038/s41586-023-05896-x

Miga, K. H., and Wang, T. (2021). The need for a human pangenome reference sequence. *Annu. Rev. Genomics Hum. Genet.* 22, 81–102. doi:10.1146/annurev-genom-120120-081921

Mukherjee, S. (2016). *The gene: an intimate history*. 1. New York: Scribner.

National Human Genome Research Institute (2022). *Human genome project fact sheet*. Available at: https://www.genome.gov/about-genomics/educational-resources/fact-sheets/human-genome-project (Accessed June 13, 2023).

Noyes, J. E. (2011). The Common Heritage of Mankind: past, present and future. *Denv. J. Int'l L. Pol'y* 40 (1), 447–471.

Office of Science and Technology Policy (2000). *J Craig Venter Statement Sub-Committee on Energy and Environment*. Available at: https://clintonwhitehouse4.archives.gov/WH/EOP/OSTP/html/00626_4.html (Accessed July 14, 2023).

Raggio, N. (2002). Cracking the human genome: a look into how competitive forces create and reshape collaborations. SWITCH *(New York Times)* 17 (1), 1–7.

Thaldar, D. W., Townsend, B. A., Donnelly, D. L., Botes, M., Gooden, A., van Harmelen, J., et al. (2022). The multidimensional legal nature of personal genomic sequence data: a South African perspective. *Front. Genet.* 13, 1–11. doi:10.3389/fgene.2022.997595

Thorogood, A., Joly, Y., Dove, E. S., and Knoppers, B. M. (2015). "Reinforcing international data sharing in genomics with human rights," in *Council of Europe's proceedings on "emerging technologies and human rights"* (Strasbough: Council of Europe), 77–84.

UNESCO (1997). *Universal Declaration on the Human Genome and Human Rights*. Paris: UNESCO.

UNESCO (2003). *International Declaration on Human Genetic Data*. Paris: UNESCO.

UNESCO (2005). *Declaration on Bioethics and Human Rights October 19*. Available at: https://www.unesco.org/en/ethics-science-technology/bioethics-and-human-rights (Accessed August 1, 2023).

UNESCO General Conference (1997). Drawing up of a declaration on the human genome: report by the director general. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000109141 (Accessed August 1, 2023).

United Nations (1966). International Covenant on Economic, Social and Cultural Rights. https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights (Accessed May 30, 2023).

United Nations (1994). UNCLOS and agreement on Part XI. Available at: https://www.un.org/depts/los/convention_agreements/texts/unclos/closindxAgree.htm (Accessed July 8, 2023).

Universal Declaration on Human Rights (1948). *Universal Declaration on Human Rights*. Available at: https://www.ohchr.org/sites/default/files/UDHR/Documents/UDHR_Translations/eng.pdf (Accessed February 27, 2023).

Waterston, R. H., Lander, E. S., and Sulston, J. E. (2002). On the sequencing of the human genome. *PNAS* 99 (6), 3712–3716. doi:10.1073/pnas.042692499

Wolfrum, R. (1983). The principle of the common heritage of mankind. *Public Int. Law*, 312–337.

frontiers | Frontiers in Pharmacology

TYPE Perspective
PUBLISHED 23 November 2023
DOI 10.3389/fphar.2023.1238749

OPEN ACCESS

EDITED BY
Go Yoshizawa,
Kwansei Gakuin University, Japan

REVIEWED BY
Julian Kinderlerer,
University of Cape Town, South Africa
Victoria Bronstein,
University of the Witwatersrand, South Africa

*CORRESPONDENCE
Donrich Thaldar,
thaldard@ukzn.ac.za

RECEIVED 12 June 2023
ACCEPTED 31 August 2023
PUBLISHED 23 November 2023

CITATION
Thaldar D (2023), Does data protection law in South Africa apply to pseudonymised data?
Front. Pharmacol. 14:1238749.
doi: 10.3389/fphar.2023.1238749

Does data protection law in South Africa apply to pseudonymised data?

Donrich Thaldar[1,2]*

[1]School of Law, University of KwaZulu-Natal, Durban, South Africa, [2]Petrie-Flom Center for Health Law Policy, Biotechnology and Bioethics, Harvard Law School, Cambridge, MA, United States

The use of pseudonymised datasets is increasingly commonplace as research institutions seek to balance data utility with data security. Yet, a crucial question arises: How does South Africa's Protection of Personal Information Act (POPIA) govern these datasets, especially given their ambiguous state between de-identification and possible re-identification? A thorough examination of POPIA suggests that the determination of whether a pseudonymised dataset is personal information—and thus whether processing the dataset falls within POPIA's purview—must be informed by the *specific context* of the responsible party in possession of the pseudonymised dataset. When a research institution retains both the pseudonymised dataset and its linking dataset, the pseudonymised dataset remains identifiable and is thus personal information that falls within POPIA's purview. However, when only the pseudonymised dataset—without the linking dataset—is transferred to another entity, it is non-personal information in the hands of such a recipient, thus freeing the recipient from POPIA compliance. Such a delineation offers research institutions greater flexibility in sharing and using pseudonymised datasets. Importantly, because the original provider of the pseudonymised dataset (who has the means to re-identify the dataset) remains governed by POPIA, the privacy rights of data subjects are not undermined.

KEYWORDS

code of conduct, data protection, POPIA, pseudonymisation, research, South Africa, transfer

1 Introduction

When sharing health research data, it is a legal and ethical imperative to secure any information that can identify *research participants*—or *data subjects* in privacy law terminology. A common technique used to accomplish this is to replace data subjects' identifying information in the dataset to be used for research—and for sharing with collaborators—with unique codes. This is done while keeping another dataset that links these data subjects' identifying information with their allocated codes. This technique—commonly referred to as *pseudonymisation*—attains non-identifiability of data subjects in certain *specific contexts*, viz where researchers have access to only the pseudonymised dataset and not to the linking dataset. However, pseudonymisation does not attain non-identifiability of data subjects in any or all contexts, as the linking dataset still exists and can be used by someone—perhaps now or in the future—to identify the data subjects in the pseudonymised dataset.

Since the identifiability of data subjects in a dataset is the fulcrum of determining whether statutory data protection law rules apply to such a dataset, it is important to know

whether *context* is legally relevant when working with pseudonymised datasets. However, this has been controversial—so much so that this has already led to litigation in the European Union. In the recent case of *Single Resolution Board v European Data Protection Supervisor* (2023), the European Data Protection Supervisor adopted a context-agnostic stance that focused on the fact that when a dataset is pseudonymised the data subjects remain identifiable because someone, somewhere, still has the linking dataset that can be used to identify the data subjects. However, aligned with previous case law (*Breyer v Federal Republic of Germany*, 2016), the General Court of the EU decided against the European Data Protection Supervisor and held that the identifiability of data subjects must be determined based on the *specific context* of the relevant party before the court. The European Data Protection Supervisor filed an appeal against this judgment (European Data Protection Supervisor, 2023). The appeal will be heard by the Court of Justice of the EU.

This ongoing litigation in the EU raises the pertinent question: What would be the position in South Africa? The South African Protection of Personal Information Act (POPIA) (Protection of Personal Information Act 4, 2013) does not explicitly deal with pseudonymisation. Also, there is no South African case law on the topic, nor any guidance by the country's Information Regulator. In this article, I analyse POPIA and propose that the South African position is that the identifiability of data subjects must be determined based on *specific context*.

2 Interpreting POPIA

I plot two complementary interpretative avenues through POPIA: The first focuses on the definitions of the terms "de-identify" and "re-identify" used in the exclusions clause—in particular, the phrase "reasonably foreseeable method" contained in these definitions—and interpret this reasonability standard by borrowing established legal principles from other branches of the law. The second interpretative avenue uses POPIA's application clause as a point of departure and then analyses POPIA's research exception. As will become evident, both of these interpretative avenues reach the same destination.

2.1 The exclusions clause and the definitions

POPIA's exclusions clause, Section 6, provides that POPIA does not apply to the processing of personal information that has been de-identified to the extent that it cannot be re-identified again. However, it is not clear from this clause *who* must not be able to re-identify the de-identified information: Nobody in the entire world, or the specific responsible party in possession of the information?

To help find an answer to this question, the definitions of the terms "de-identify" and "re-identify" (in Section 1 of POPIA) should be considered. Both definitions use much of the same language and are mirror images of each other. They both relate to information that (a) identifies the data subject; (b) can be used or manipulated by a *reasonably foreseeable method* to identify the data subject; or (c) can be linked by a *reasonably foreseeable method* to other information that identifies the data subject. The difference is that de-identification is the *deletion* of such information [meaning information of type (a), (b), or (c)], while re-identification is the *resurrection* of such information that has been deleted. Note POPIA's use of the phrase a "reasonably foreseeable method". The concept of reasonableness is not unique to informational privacy law (as codified in POPIA) and is regularly used in other branches of South African law—especially the law of delict and administrative law. In these branches of the law, reasonableness is understood to entail an *objective* inquiry (*Cape Town Municipality v Bakkerud*, 1997; *Medirite Ltd v South African Pharmacy Council*, 2013). There is no reason to believe that the same would not also apply to POPIA. An objective inquiry means that when considering whether a pseudonymised dataset is in fact de-identified, and also when considering whether such a pseudonymised dataset can be re-identified, the test is not whether the responsible party *subjectively* foresaw a method that would re-identify the information, but rather whether a reasonable person—an abstraction—would have foreseen a method that would re-identify the information.

Although it is important that the definitions of "de-identify" and "re-identify" require an objective inquiry, this still does not solve the problem of the legal relevance of *context*. Must the reasonable person be conceived of in a context-agnostic way, or conceived of in a specific context? The well-established position in South African law of delict is to conceive of the reasonable person *in the position* of the person whose conduct is considered (*Mukheiber v Raath*, 1999). In other words, the objective inquiry is not context-agnostic, but firmly anchored in a specific context. Therefore, if applied to POPIA, the objective test is not whether a reasonable person *anywhere in the world* would have foreseen a method that would re-identify the information, but rather whether a reasonable person *in the position of the responsible party* would have foreseen a method that would re-identify the information.

Accordingly, the first interpretative avenue leads to the conclusion that the identifiability of data subjects must be determined in an objective, *context-specific* way.

2.2 The application clause and the research exception

The second interpretative avenue follows a different pathway through POPIA but reaches the same conclusion. POPIA's application clause (Section 3) provides that POPIA applies only to *personal* information, which is information relating to an identifiable living natural person and, where it is applicable, an identifiable existing juristic person. Accordingly, in the health research context, the first question is whether the information *relates* to actual human research subjects (in other words, the information is not synthetic). And if the first question is answered in the affirmative, the second question is whether these human research subjects (i.e., the data subjects) are *identifiable* from the information. If the second question is also answered in the affirmative, POPIA applies to such information.

However, analogous to the position with the exclusions clause discussed above, it is not clear from the application clause *who* must not be able to *identify* the data subjects from the information: Nobody in the entire world, or the specific responsible party in

possession of the information? Although this question is not explicitly answered in POPIA, the way in which the word "identifiable" is used elsewhere in POPIA, namely, in the research exception [Section 15 (3) (e)], does suggest the answer.

POPIA's research exception allows for secondary research on personal information already collected without the need to re-consent the data subjects, but on condition that the responsible party ensures that the personal information used in such secondary research "will not be published in an identifiable form". Thus, the data subjects must not be identifiable from the information *that is shared with the public*. Yet, there is no requirement in this section that the research institution must de-identify the personal information that is in its own possession—i.e., not shared with the public. This means that the research institution *itself* can retain its ability to identify the data subjects. Accordingly, POPIA contemplates identifiability to be determined from the perspective of the person or institution that is interacting with the relevant information. In other words, POPIA contemplates identifiability to be *context specific*.

Let me explain this from another angle: In the context of health research, POPIA's research exception envisions the possibility of multiple versions of the same dataset. The dataset that contains personal information (call it "Dataset A") can be used repeatedly for research purposes without the need to re-consent the data subjects. This can continue *ad infinitum*. Whenever an article is published based on the research, and whenever the underlying data must be provided as supplementary material to the journal, Dataset A or the relevant part of it that the article relies upon, must be de-identified (call it "Dataset A1") before submitting it to the journal. This de-identification can be accomplished by either *deleting* all identifiable information in the derivative dataset or by *replacing* such information in the derivative dataset with a *pseudonym* that the public does not have access to. Whichever method is employed, the research institution complies with POPIA's research exception, as the public (excluding the research institution) cannot identify the data subjects. Note that whether the derivative datasets are created by deleting information or by pseudonymisation makes no difference to the fact that the research institution remains in possession of Dataset A itself—the original dataset that contains all the personal information. Datasets A and A1 exist at the same time—one version of the dataset in "identifiable form," another version *not* in "identifiable form". This vision of what is entailed by POPIA's research exception is clearly incompatible with *identifiable* meaning *identifiable by anyone in the world*, as the data subjects will indeed be identifiable by those with access to Dataset A.

In statutory interpretation, according to the principle of internal consistency, it is presumed that the meaning of a term used in a statute remains consistent throughout the statute (*Minister of the Interior v Machadodorp Investments Ltd*, 1957). Accordingly, identifiability should consistently be interpreted in a *context-specific* way.

2.3 Conclusion on interpretation

While POPIA does not overtly elaborate on pseudonymisation, POPIA's provisions, when interpreted contextually and in light of established South African legal principles, lean towards an objective, context-specific understanding of data subject identifiability. Notably, the concept of "reasonably foreseeable method" intertwined with established legal precedents, and the contextual interpretation of the term "identifiability" in POPIA's research exception, both converge on a perspective that grounds data subject identifiability in specific contexts. It is noteworthy that this interpretation aligns with the European position expressed in the *Single Resolution Board*.

3 POPIA's application to pseudonymised datasets

In this section, I consider how the context-specific interpretation of identifiability in POPIA applies to pseudonymised datasets. First, I focus on the practical issue of determining whether a dataset is pseudonymised. I then consider the legal position under POPIA of each of the parties to a data transfer agreement, namely, the provider and the recipient, where the dataset that is transferred is pseudonymised.

3.1 Pseudonymisation in health research practice

When exactly is a dataset pseudonymised? In health research this question might not always have an obvious answer. Consider, for example, a research institution that conducts genomic research. It collects the data subjects' names, phone numbers, gender, age group, race, and takes blood samples that are used to generate genomic data. All of these data are combined in a dataset. If the research institution replaces the data subjects' names and phone numbers with unique codes, is the dataset pseudonymised? The answer depends on an assessment of whether the genomic data can identify a data subject. Say, for example, the research institution conducted genotyping (investigating the differences in individuals' genotypes) and used a targeted approach of focusing only on specific portions of DNA instead of the entire genome. This targeted approach does not mean that the resulting data are not identifiable. In fact, genotyping data may contain unique genetic markers specific to an individual. This means that under the right circumstances or when combined with other datasets, an individual could potentially be identified. Although human whole-genome sequencing is relatively rare in South Africa, the same would obviously apply. On the other side of the spectrum, information on a single allele—even if rare—within a sufficiently large cohort would not be sufficient to identify a person.

If it is determined that the dataset still contains data that can identify a data subject, even after the data subjects' traditional identifiers, such as their names and contact numbers have been replaced with codes, it means that the dataset has only been *partially* pseudonymised. Although this is a good data *security* measure (as it limits the risk of data subjects being identified), from a *legal* perspective it does not change the dataset's status, as it remains inherently identifiable. In other words, for purposes of legal analysis, partial pseudonymisation is not pseudonymisation.

To build on the example above, an important question is whether a dataset that contains identifying genomic data can be pseudonymised? Similar to a dataset that contains high-resolution geolocation data that can be pseudonymised by—over and above replacing names with codes—lowering the resolution of the geolocation data to such an extent that such data can no longer identify any data subject, a dataset that contains genomic data can also be pseudonymised by lowering the dataset's resolution in the sense that only broader, less granular data is retained. For example, exact genetic sequences can be replaced with information about whether a certain genetic marker is present or not. For certain datasets, it might be possible to generalise data by grouping them. However, depending on the kind of analysis that researchers intend to perform on the dataset, these techniques may entail sacrificing useful and valuable data, and their use is therefore not always appropriate or desirable.

In sum, therefore, a dataset is pseudonymised by taking the following steps: Allocating a unique code for each data subject; deleting all the traditional identifiers, such as name and phone number; where applicable, deleting any other identifying information, such as unique genetic markers specific to an individual, or changing such information to the extent that it can no longer identify any data subjects; and creating a dataset that links the unique codes of the data subjects with their identities, and keeping such linking dataset separate, confidential, and secure.

3.2 Transferring a pseudonymised dataset

Consider the following scenario: University X collects health information from research participants (data subjects). From the outset, University X employs a pseudonymisation system to ensure that the health information dataset that it is developing does not contain any identifying information of the data subjects. University X keeps the linking dataset separate, confidential and secure. The following legal questions are pertinent: First, does POPIA apply when University X processes its pseudonymised dataset? Second, if University X shares a copy of its pseudonymised dataset with University Y—but not the linking dataset—does POPIA apply when University Y processes the pseudonymised dataset?

3.3 The pseudonymised dataset in the hands of the provider

Although University X keeps the linking dataset secure, it possesses both the pseudonymised dataset and the linking dataset, and therefore has a reasonably foreseeable method at its disposal to re-identify the pseudonymised dataset. An important data safety measure for University X is having internal policies in place to ensure that the linking dataset is secured and that the researchers who are using the pseudonymised dataset do not have access to the linking dataset. But it does not change the fact that University X *qua* juristic person can re-identify the pseudonymised dataset.

Accordingly, in the hands of University X, the pseudonymised dataset constitutes *personal* information (information relating to *identifiable* living natural persons) and POPIA applies to such a pseudonymised dataset. This means that any processing of the information contained in the pseudonymised dataset by University X must be done in compliance with the relevant conditions for processing, as provided in POPIA. However, does such processing include transfer of the pseudonymised dataset to University Y? I return to this question after discussing University Y.

3.4 The pseudonymised dataset in the hands of the recipient

University Y possesses only the pseudonymised dataset and not the linking dataset, and therefore does not have a reasonably foreseeable method to re-identify the pseudonymised dataset. Accordingly, in the hands of University Y, the pseudonymised dataset does *not* constitute personal information and POPIA does *not* apply. It follows then that when University Y processes the information in the pseudonymised dataset, it is under no legal obligation to comply with any of POPIA's conditions for processing.

3.5 Redux: transfer of the pseudonymised dataset by the provider

At the moment that University X transfers the pseudonymised dataset, the dataset is still personal information in its hands. This seems to suggest that University X must comply with POPIA's rules regarding the transfer of the pseudonymised dataset to University Y. On the other hand, the act of transfer implies that the information will be placed in possession of the recipient. Common sense dictates that the act of transfer of information necessitates an orientation towards the *recipient*, instead of the provider.

This common-sense position can be strengthened by the following legal argument: South African law adheres to the doctrine of purposive interpretation (*Bertie Van Zyl Ltd v Minister for Safety and Security,* 2009). Thus, one should ask: What is the purpose of applying the rules of POPIA to the *transfer* of information? The purpose, I suggest, is to ensure that data subjects' privacy rights are protected when the recipient receives the transferred information. This is why, for example, where the recipient is in a foreign country (see POPIA Section 72), it is legally relevant whether the *recipient* is subject to law, binding corporate rules or a binding agreement which provide an adequate level of protection for the processing of personal information. Would applying the rules of POPIA to a transfer where the transferred information will be non-personal information in the hands of the recipient ensure that data subjects' privacy rights are protected when the transferred information is received by the recipient? The answer is clearly "no". In the hands of the recipient the information is non-personal information. In other words, the recipient has no reasonably foreseeable method of identifying the data subjects and therefore their privacy rights are, from the outset, not at risk. It follows that when University X transfers the information in the

pseudonymised dataset, it is under no legal obligation to comply with any of POPIA's conditions for processing.

4 Conclusion

If my analysis is correct, namely, that identifiability in POPIA ought to be interpreted in a context-specific way, the transfer of pseudonymised datasets by providers and the subsequent processing of such datasets by recipients fall beyond POPIA's scope of application. This result provides significantly more leeway for both providers and recipients of pseudonymised datasets. Does this leeway come at a cost for the privacy rights of data subjects? I suggest not. Nobody but the providers of the pseudonymised datasets—those who hold the key to re-identification of such datasets—have a reasonably foreseeable method of identifying the data subjects. And these providers remain bound by POPIA's rules when they process pseudonymised datasets *within* their organisations, for example, when their own staff analyse the pseudonymised datasets for research purposes.

However, a note of caution is warranted. My argument hinges on the premise that a recipient does not possess reasonably foreseeable means to re-identify a (properly) pseudonymised dataset. Yet, scenarios can be imagined where this premise is challenged. For instance, if University X collected its research data in partnership with University Z, the latter might have the means to re-identify a pseudonymised dataset based on their joint research. However, many years later, staff members from University X might be oblivious to this past collaboration. Therefore, for practical reasons, I propose that the provider of a pseudonymised dataset should (a) internally examine the organisational history of the pseudonymised dataset and (b) query the recipient about any accessible information that could serve as a key to re-identify the dataset. Both (a) and (b) ought to be documented, with the results of (b) ideally being included in the parties' data transfer agreement.

At the end of 2020, the Academy of Science of South Africa (ASSAf) embarked on a project to develop a Code of Conduct for Research (Code) in terms of POPIA. This project offers the opportunity to clarify when and how POPIA applies to pseudonymisation, and how pseudonymisation should be used in research. Academy of Science of South Africa (2023) recently submitted its proposed version of the Code to the South African Information Regulator for its consideration and eventual approval. The Information Regulator (2023) then published the proposed Code for public comment. The proposed Code defines pseudonymisation and embraces it as the default in all high-risk research. However, the proposed CCR does not address the essential issue of the relevance of context in the interpretation of identifiability. Given the widespread use and sharing of pseudonymised datasets in health research in South Africa, I suggest that the final Code should provide clarity on this highly consequential issue and illustrate its application to everyday research activities with practical examples. Moreover, since the use of pseudonymised datasets transcends the research milieu, the Information Regulator should publish a general guidance note to clarify this issue.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

The author acknowledges the support by the US National Institute of Mental Health and the US National Institutes of Health (award number U01MH127690). The content of this article is solely the author's responsibility and does not necessarily represent the official views of the US National Institute of Mental Health or the US National Institutes of Health.

Acknowledgments

The author would like to express his gratitude to Paul Esselaar and to the peer reviewers for their insightful comments on previous drafts of this article. Any remaining errors are the author's alone.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Academy of Science of South Africa (ASSAf) (2023). POPIA Code of Conduct for Research. Available at: https://www.assaf.org.za/wp-content/uploads/2023/04/ASSAf-POPIA-Code-of-Conduct-for-Research.pdf.

Bertie van Zyl (Pty) Ltd v Minister for Safety and Security (2009). ZACC 11 [South Africa]. Available at: http://www.saflii.org/za/cases/ZACC/2009/11.html.

Breyer v Federal Republic of Germany (2016). C-582/14, EU:C:2016:779 [European union]. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:62014CJ0582.

Cape Town Municipality v Bakkerud (1997). 4 SA 356 (C) [South Africa].

European Data Protection Supervisor (2023). Notice of appeal against the judgment of the General Court in case T-557/20. *Single Resolution Board v European Data Protection Supervisor* (Case C-413/23 P) (2023/C 296/26) [European Union]. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:62023CN0413.

Information Regulator [South Africa] (2023). *Notice in terms of section 61(2) of the Protection of Personal Information Act no 4 of 2013 (POPIA) Code of Conduct: the Academy of Science of South Africa (ASSAf)*. Government Gazette no 48589 of 12 May 2023. Available at: https://inforegulator.org.za/wp-content/uploads/2020/07/Government-Gazzette-dated-12-May-.pdf.

Medirite (Pty) Ltd v South African Pharmacy Council (2013). ZAGPPHC 493 [South Africa]. Available at: http://www.saflii.org/za/cases/ZAGPPHC/2013/493.html.

Minister of the Interior v Machadodorp Investments (Pty) Ltd (1957). 2 SA 395 (A) [South Africa].

Mukheiber v Raath (1999). 3 SA 1065 (SCA). [South Africa].

Protection of Personal Information Act 4 of 2013 (2013). [South Africa]. Available at: https://www.gov.za/documents/protectionpersonal-information-act.

Single Resolution Board v European Data Protection Supervisor (2023). T-557/20, ECLI:EU:T:2023:219 [European Union]. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62020TJ0557.

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 06 December 2023
DOI 10.3389/fgene.2023.1273975

OPEN ACCESS

EDITED BY
Donrich Thaldar,
University of KwaZulu-Natal, South Africa

REVIEWED BY
Marco Brandizi,
Rothamsted Research, United Kingdom
Marietjie Botes,
University of Luxembourg, Luxembourg
Scott Hazelhurst,
University of the Witwatersrand, South Africa

*CORRESPONDENCE
Michael S. Pepper,
michael.pepper@up.ac.za

RECEIVED 07 August 2023
ACCEPTED 10 November 2023
PUBLISHED 06 December 2023

CITATION
Strydom A, Van Rensburg J and Pepper MS (2023), A data management plan for the NESHIE observational study. *Front. Genet.* 14:1273975. doi: 10.3389/fgene.2023.1273975

A data management plan for the NESHIE observational study

Adéle Strydom, Jeanne Van Rensburg and Michael S. Pepper*

Institute for Cellular and Molecular Medicine, Department of Immunology, and SAMRC Extramural Unit for Stem Cell Research and Therapy, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

With regard to the use and transfer of research participants' personal information, samples and other data nationally and internationally, it is necessary to construct a data management plan. One of the key objectives of a data management plan is to explain the governance of clinical, biochemical, laboratory, molecular and other sources of data according to the regulations and policies of all relevant stakeholders. It also seeks to describe the processes involved in protecting the personal information of research participants, especially those from vulnerable populations. In most data management plans, the framework therefore consists of describing the collection, organization, use, storage, contextualization, preservation, sharing and access of/to research data and/or samples. It may also include a description of data management resources, including those associated with analyzed samples, and identifies responsible parties for the establishment, implementation and overall management of the data management strategy. Importantly, the data management plan serves to highlight potential problems with the collection, sharing, and preservation of research data. However, there are different forms of data management plans and requirements may vary due to funder guidelines and the nature of the study under consideration. This paper leverages the detailed data management plans constructed for the 'NESHIE study' and is a first attempt at providing a comprehensive template applicable to research focused on vulnerable populations, particularly those within LMICs, that includes a multi-omics approach to achieve the study aims. More particularly, this template, available for download as a supplementary document, provides a modifiable outline for future projects that involve similar sensitivities, whether in clinical research or clinical trials. It includes a description of the management not only of the data generated through standard clinical practice, but also that which is generated through the analysis of a variety of samples being collected from research participants and analyzed using multi-omics approaches.

KEYWORDS

sample, data, management, legislation, NESHIE

1 Introduction

Data management is important in any biomedical research project and facilitates the generation of high-quality and reliable data (Nourani et al., 2022). Broadly, a data management plan (DMP) may have the following benefits (Fadlelmola et al., 2021): 1) it protects the research participants and the project team; 2) it allows compliance with local data protection policies and legislation; 3) it maintains FAIR content; 4) it enables research that is transparent; and 5) it allows compliance with funder requirements. However, when considering a DMP, two key questions usually emerge. The first asks "what is a data management plan?", while the second asks "how do you write one?"

In answering the first question, Stanford University defines a DMP as follows: "a written document that describes the data you expect to acquire or generate during the course of a research project, how you will manage, describe, analyze, and store those data, and what mechanisms you will use at the end of your project to share and preserve your data."[1] While some or all of these issues may have been considered when starting a research project, their documentation validates the DMP construction process. In so-doing, weaknesses in the plan are identified and a record is kept of what is proposed or completed. While potentially labor-intensive, the construction of a DMP is nevertheless viewed as a worthwhile exercise that addresses data management prior to the onset of a research project, rather than in a reactionary or improvised fashion during or towards the end of a project. The aim of any DMP should therefore be to focus attention on available resources and research infrastructure, and identify parties responsible for the inception, implementation and management of the DMP. The DMP should also highlight potential problems regarding long-term preservation and sharing of data and samples. When noting potential problems, some form of recourse or plan of action should accompany the DMP. This is because good data management can assist in preventing 'bad' research. While 'bad' research may result in the retraction of published papers, 'good' research provides data that is documented, stored, and includes reasonable routes for access.

Regarding the question on how to write a DMP, several online tools and questionnaires are available for this purpose (Fadlelmola et al., 2021). While questionnaires provide a guide to the nature of the data management issues that should be considered when writing a DMP, online tools include templates with information and guidance for ready-to-use DMPs. While these tools may be specific to a research project or funder, they typically include text that can be copied and pasted into a customized DMP. They also provide different export formats to support the requirements of funding applications. Examples of such online tools include DMPonline[2] and DMPTool[3]. However, because research is discovery-oriented, the research process sometimes requires a change in direction and a revision of the intended data management path. As such, while the DMP should be constructed prior to the onset of the project, the DMP should also be viewed as a dynamic document that may be altered during the course of the research study. Every time the research plan changes, the DMP should be reviewed to make sure that it still meets the various regulatory and statutory requirements. This includes considering the funder-supplied set of policies and guidelines for data management and sharing (Fadlelmola et al., 2021). Funding bodies increasingly require that DMPs accompany study proposals when submitting funding applications. Importantly, such DMPs are typically required to consider open data sharing models (Wilkinson et al., 2016).

An increase in the need for open data sharing has resulted in the construction of a recommended set of standards for data management known as the "FAIR" principles. These principles have been defined as[4].

- "*Findable*—Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services."
- "*Accessible*—Once the user finds the required data, he/she needs to know how it can be accessed, possibly including authentication and authorization."
- "*Interoperable*—The data usually needs to be integrated with other data. In addition, the data needs to interoperate with applications or workflows for analysis, storage, and processing."
- "*Reusable*—The ultimate goal of FAIR is to optimize the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings."

While data management planning may have technical challenges that include not clearly knowing the benefits and best practices for a research project at inception (Lefebvre et al., 2020), the practical implementation of FAIR principles in low- and middle-income countries (LMICs) may be hindered by a number of additional factors (Fadlelmola et al., 2021). These may include a lack of research funding, inadequate human resources, limited research data management guidelines and policies, a lack of training in research data management, inadequately secure and/or reliable technology, inefficient or inadequate archiving of data, and inefficient support from academic institutions regarding data management. Because of historical, cultural, and ethical concerns, special consideration should also be made in the construction of DMPs when planning to share African-centric data.

Nevertheless, the data management process not only consists of creating study-associated documents such as data sheets or case report forms (CRFs) and consent forms, but also involves training of the research team, creating databases, capturing and validating data, managing data discrepancies, resolving data disagreements, describing the processes of data coding and extraction, access control, recording the data management process, and providing security throughout the duration of a research project (Nourani et al., 2022). When working with databases, electronic data management systems require sufficient hardware, software, communication technologies, policies/guidelines for data collection, quality control of data, and security in order to be operational (Nourani et al., 2022).

Regardless of the form the DMP assumes, the basic principles that govern its construction include the preservation of and (continued) access to the research data (Dunie, 2017). This not only ensures the reproducibility, traceability, and reliability of the research data, but also assists in reducing the costs of performing additional research investigations (Williams et al., 2017).

1 Stanford University. (2023). Data management plans. https://doresearch.stanford.edu/topics/manage-research-data [Accessed 28 August 2023]

2 DMPonline. (2023). Plan to make data work for you. https://dmponline.dcc.ac.uk/[Accessed 5 September 2022]

3 DMPTool. (2023). Create Data Management Plans that meet requirements and promote your research. https://dmptool.org/[Accessed 5 September 2022]

4 GO FAIR. (2023). FAIR Principles. https://www.go-fair.org/fair-principles/ [Accessed 6 September 2022]

Reproducibility and traceability are at risk when robust policies and documentation regarding data management are absent. As such, it is important that issues regarding data sharing and secondary use of data are covered in the DMP, especially in relation to cross-border sharing of samples/data. In most instances, a material transfer agreement (MTA) or data transfer agreement (DTA) between collaborators will resolve data and/or sample transfer issues. Such agreements may be listed in the DMP. Globally, funders and research institutions are promoting open science policies and practices to manage research data (Lefebvre et al., 2020).

Since nearly 50% of medico-legalcases brought against the South African National Department of Health between 2019 and 2020, which totaled ZARR53-billion, were linked to birth asphyxia, neonatal encephalopathy and cerebral palsy, it is important that the management of data linked to any or all of these conditions are carefully considered and well documented.[5,6,7] This paper will consequently describe the general principles used to design and structure the data management plan for the national multi-institutional NESHIE (Neonatal Encephalopathy with Suspected Hypoxic Ischemic Encephalophathy) project being overseen by the Institute for Cellular and Molecular Medicine (ICMM) in the Faculty of Health Sciences at the University of Pretoria (UP), South Africa, in collaboration with the Universities of Cape Town (Cape Town, South Africa), Stellenbosch (Cape Town, South Africa), and the Witwatersrand (Johannesburg, South Africa).

Within this ongoing study, a 'multi-omics' approach is being employed to identifying proximal biomarkers and increase understanding of the pathogenesis of NESHIE in a highly vulnerable population. This study is unique in that genomic, epigenomic, transcriptomic, proteomic and metabolomic ('multi-omic') analyses are being performed on the same individuals on whom large-scale clinical data (up to 1,500 variables per neonate-maternal pair) is being collected. This data includes imaging information (cranial ultrasound and limited magnetic resonance imaging data), placental pathology data, and an additional molecular component investigating the potential pathomicrobiome associated with placental tissue samples. While a full contingent of samples is not necessarily collected for every participant, a DMP that comprehensively describes the management of the clinical and multiple molecular data outputs was, and remains, necessary for the NESHIE study. However, at the onset of sample and data collection for the NESHIE study in 2019, existing DMP templates did not fully capture the data management needs of the study. A detailed DMP was subsequently developed and is presented here for use in research projects or clinical trials involving NESHIE or associated conditions, particularly in LMICs, for clinical research investigations involving multi-omic data outputs, and/or investigations involving vulnerable populations. Importantly, this DMP was constructed in the context of South Africa and was therefore guided by the requirements of the National Health Research Ethics Council (NHREC), the Protection of Personal Information Act 4 of 2013 (POPIA), and the National Health Act 61 of 2003 (NHA) and the Declaration of Helsinki. It was additionally guided by Good Clinical Practice (GCP) principals. It describes the safe, secure and ethical manner in which clinical and multi-omic sample-associated data collected from vulnerable populations may be collected and shared in a research team or amongst collaborators.

The NESHIE study DMP template, which is available as Supplementary Material S1, is user friendly, easy to access and can be adapted to most research projects or clinical trial. To our knowledge, this is the first data management plan of this nature to be published.

2 Framework of the NESHIE data management plan

The NESHIE Data Management Plan consists of a combination of DMP templates with sections relevant to a wide range of clinical data and associated molecular analyses outputs generated as part of the study.[8,9,10] Sections have been modified to the requirements of the study, thus creating a living document that is easy to maintain. The NESHIE study DMP complies with the policies and guidelines of all stakeholders (academic institution and funders) involved in the project and has been approved by the University of Pretoria Research Ethics Committee (REC; reference number 481/2017).

The first page of the DMP template consists of a cover page (Figure 1) followed by the Table of Contents on the next page. In the case of the NESHIE project, a project manager and a DMP coordinator review the document regularly. This process considers amendments made to the study protocol, SOPs, and local and international policies/regulations. These amendments are listed on the cover page of the DMP under 'Revision History' (Figure 1) and indicates the dates and details of the changes and by whom they are made.

The sections that follow the Title page in the NESHIE DMP include.

- A *List of abbreviations* that is specific to the NESHIE project;
- *Definitions* of terms used in the DMP that may otherwise be misinterpreted or misunderstood;
- The *Scope* of the DMP that briefly describes the policies and regulations to which the NESHIE project complies, and explains the application of the DMP; and

5 News24. (2023). State hospitals pay huge legal claims for cerebral palsy, but new study makes surprising claims. https://www.news24.com/fin24/companies/state-hospitals-pay-huge-legal-claims-for-cerebral-palsy-but-new-study-makes-surprising-claims-20210622 [Accessed 5 September 2023]

6 Daily Maverick. (2023). Cerebral palsy in South Africa—medical negligence is just one of many causes. https://www.dailymaverick.co.za/article/2022-07-04-cerebral-palsy-in-south-africa- medical-negligence-is-just-one-of-many-causes/[Accessed 15 March 2023]

7 News24. (2023). Landmark medical negligence ruling orders SA hospitals to treat—not pay—victim. https://www.news24.com/fin24/companies/landmark-medical-negligence-ruling-orders-sa-hospitals-to-treat- not-pay-victim-20230213 [Accessed 5 September 2023]

8 Society for Clinical Data Management. (2023). Data Management Plan.https://scdm.org/wp-content/uploads/2019/12/GCDMP- Data-Management-Plan-2019-Edition.pdf [Accessed 15 March 2023]

9 University of Toronto. (2023). Data Management Plans.https://onesearch.library.utoronto.ca/researchdata/data-management-plans [Accessed 3 August 2023]

10 DCC. (2023). Data Management Plans. https://www.dcc.ac.uk/resources/data-management-plans [Accessed 3 August 2023]

Title of project: NESHIE South Africa – Biomarker Discovery and Neuroprotection

Document name: Sample and Data Management Plan

Document Version Number: Version 3.0

Name of organization where research is being done (list if multiple)	E.g. University of Pretoria
Date current DMP version came into effect	E.g. 7 March 2005
Other administrative details (E.g. IRB/REC approval numbers)	E.g. Protocol nr. 123/2023

DMP Prepared by*	*Dr Adéle Strydom (AS)*
DMP Reviewed by*	*Miss Jeanne Van Rensburg (JVR)* *Prof Michael S. Pepper (MSP)*
DMP Approved by*	*Prof Michael S. Pepper*

* List all individuals if more than one person is involved in any part of this

Revision History:

Version Number	Date	Author(s)	Authorized by	Description of amendments made
E.g. 1	7 March 2022	AS	MSP	E.g. Section # - IP agreements (updated) Annexure: Agreements (updated)
E.g. 1	4 October 2022	AS	MSP	E.g. Annexures updated

FIGURE 1
The cover page for the NESHIE Data Management Plan. Image created by AS.

- A short *Protocol summary* or *Introduction* that provides a broad overview of the NESHIE study protocol and that may include the variables used to analyze critical data.

The points above form the foundation of a DMP and should provide the reader with sufficient background information to facilitate an understanding of the content of the remainder of the DMP.[8,9,10] As summarized in Figure 2, the core elements of a DMP are then described in more detail in the subsequent sections. While the NESHIE study information has been used for this purpose, the details should be amended to suit the needs of each study being performed.

2.1 Project and sample/data collection

2.1.1 Research team and training

A comprehensive DMP includes all role players and organizations, and describes their roles and responsibilities (Michener, 2015). These responsibilities may include collection and entry of data, quality control, creation and management of metadata, submission of data to an archive, and the administration of databases. Furthermore, the research team and training section of a DMP (1) specifies the research team responsible for collecting data and samples at the participating study sites, and (2) describes the project-specific training requirements for the research team.

Table 1 summarizes the roles and responsibilities of the NESHIE research team, and Table 2 describes the NESHIE training requirements for the study team members. Importantly, while this may not be a formal requirement for all studies, given the long-term aspirations of the research associated with the NESHIE study to transition into a clinical trial, Good Clinical Practice (GCP) training is required by all medical officers/staff and research assistants/associates involved in sample and data collection. Renewal of GCP certification is also required. Each study will however have its own requirements regarding GCP training, including identification of those team members who require

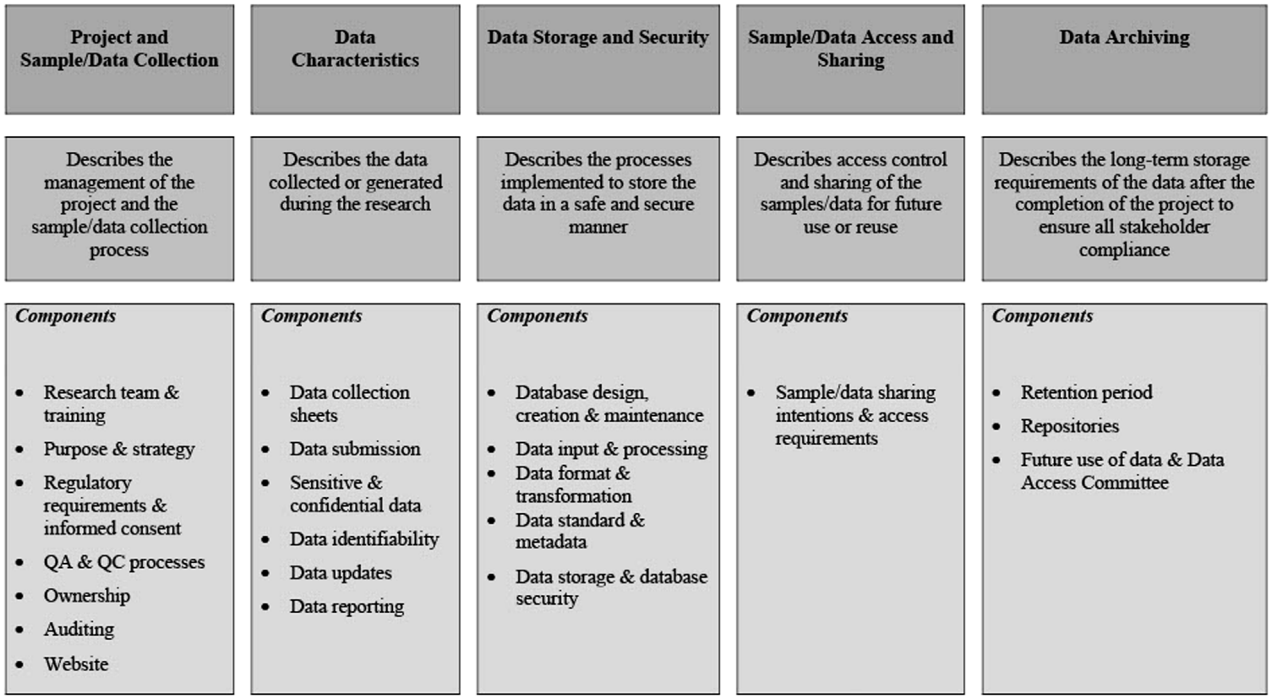

FIGURE 2
Core elements as described in the NESHIE data management plan. Image created by AS.

certification. In relation to the NESHIE study, further training is provided prior to study onset for the following:

- Study protocol;
- Study annexure documents;
- Consent documents and associated process;
- Sample collection; and
- Data collection and capturing to electronic platforms.

Training typically starts at the initiation of the study at each site but may be ongoing. Additional training may be arranged as required or requested by study associates. Training is recorded during the formal training sessions. As best practice, these records should be retained in a study-associated logbook or file for the full duration of the project, as is the case for the NESHIE study.

2.1.2 Purpose and strategy of sample/data collection

A DMP will include information that explains what samples and data are to be collected (Michener, 2015). Typically, a list of various types of samples/data that are expected to be collected or created, which could include biological samples, patient records, or images, will be provided within a DMP. The source of the samples/data is usually also provided. In this section, the NESHIE study DMP outlines what samples and data should be collected/generated and what the collection and future analysis strategy regarding clinical and molecular data is. This information is summarized in Figure 3. Importantly, while Figure 3 provides an overview of the samples and data collection purpose, collection strategies are described in detail following the overview. For the NESHIE study, this includes providing details on the collection of time-sensitive samples and data (e.g., umbilical cord blood, dry blood spot samples, and follow-up data at 9–12 months of age), as well as samples and data unaffected by collection time (e.g., peripheral blood and baseline laboratory values). This is important to provide for context given the need to account for multiple sets of data generated for the same patient at different points in time. For example, umbilical blood will result in the generation of both genomic and transcriptomic data, while dry blood spot samples allow both metabolomic and proteomic datasets to be generated at two different timepoints within 3 days of life.

Lastly, within this section of the DMP, it is also important to state the aims and objectives of the sample and data collection process for context. The aims and objectives were summarized in the NESHIE DMP as follows:

1) Providing a detailed description of grade 2–3 NESHIE in South African tertiary-level hospitals;
2) Biomarker identification; and
3) Determining whether genetic factors are associated with and potentially contribute toward the presentation of NESHIE.

TABLE 1 Examples of roles and responsibilities of the NESHIE research team.

	Project roles	Project responsibilities
Project leads	Principal Investigator Project Manager Lead Neonatologist Lead Placental Pathologist Lead Obstetrician	• Project management team at academic institutions and study sites
Project site support	Site Neonatologists Site Placental Pathologists Site Obstetrician	• Management of sample and data collection from research participants at participating study sites • Management/Co-ordination of neurodevelopmental follow-up with research participants at study sites
Biomarkers and clinical trial components	Principal Investigator Project Manager Radiology/Imaging team Molecular analysis team Clinical trials team	• Imaging biomarkers of research participants (E.g. MRI scans) • Analyses of biological samples (E.g. genomic/transcriptomic/metabolomic/proteomic and placental microbiomic analysis) • Clinical trial protocol development
Electronic data capture platform	Bioinformaticians IT support Data Manage	• Administration and maintenance of electronic data capture platforms (E.g. REDCap)
Study appointees	Medical officers and clinical support staff Scientists Bioethicists Statisticians Administrative staff	• Collection, storage and transport of samples • Data collection and capture • Quality control of data • Molecular data analysis • Radiography and imaging • Sample and data management • Clinical trial management

TABLE 2 Examples of study training requirements.

Training document	Who receives the training	When is training received
Protocol documents	All study members	Prior to study onset
Informed Consent Documents & Associated Process	All study members	Prior to study onset
SOP documents	All relevant study members. Not all study members may require training on all study-associated SOPs. If necessary, SOP training should be segmented according to the respective study roles and responsibilities	Prior to study onset
Good Clinical Practice (GCP)	All clinical and scientific study members. All relevant administrative study members	Prior to study onset; renewable every 2–3 years
Good Laboratory Practice (GLP)	All laboratory-based study members	Prior to study onset; renewable every 2–3 years
Case Report Form Completion & Data capturing	All data-associated study members and project site support team	Prior to study onset

The following additional long-term aims are included with biomarker identification:

1) Predicting susceptibility to NESHIE;
2) Identifying factors other than hypoxia/ischemia that may cause/contribute to the clinical presentation;
3) Quantifying the extent and duration of hypoxia if present; and

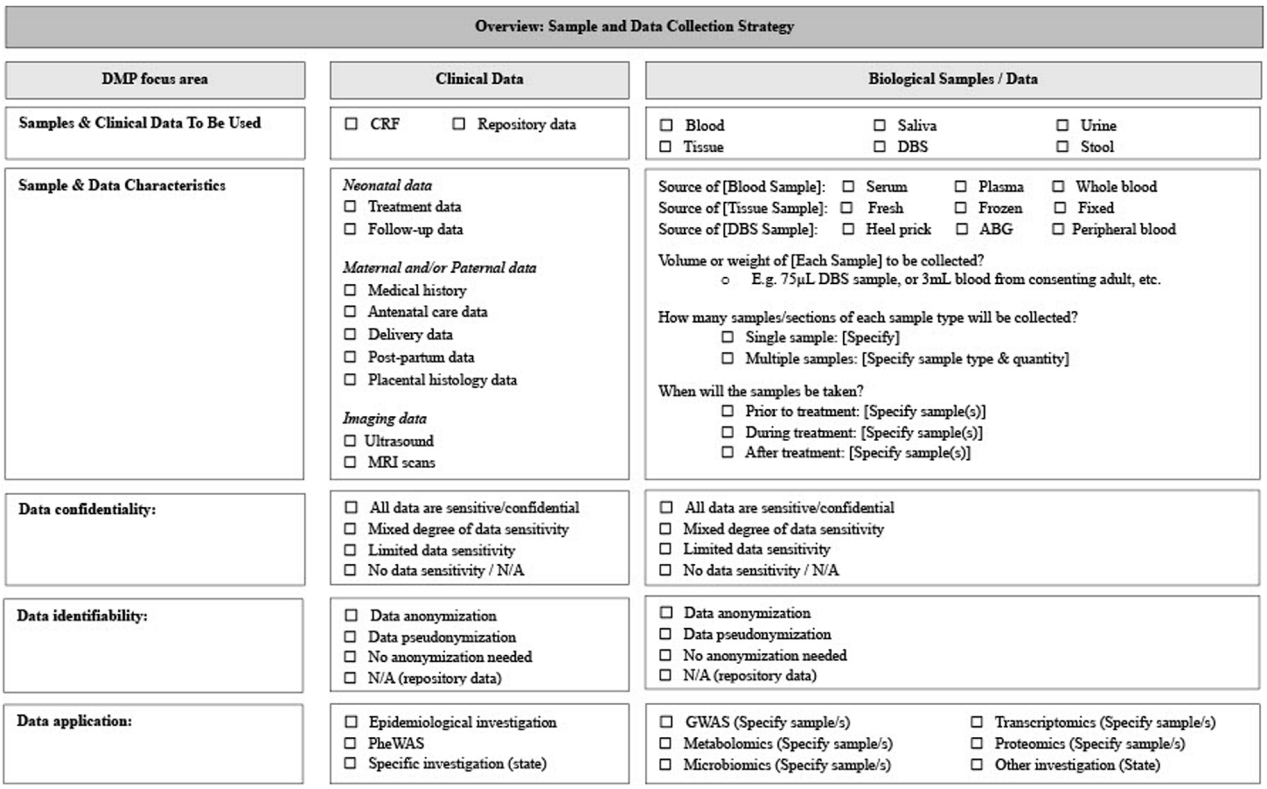

FIGURE 3
Sample and data collection summary template. Image created by JVR.

TABLE 3 Example of a document tracking log.

Document amendment no.	Document name	Document version in use under current amendment	Date of IRB/REC amendment submission	Date of IRB/REC amendment approval
Amendment 1	Study Protocol	Version 1 (Original version)	No changes to submit	N/A
Amendment 1	Study SOP: Sample collection	Version 2	27 March 2023	07 May 2023

4) Defining prognostic factors used to predict the medium-to-long term consequences in affected neonates; thus minimizing risk by increasing awareness and altering management during pregnancy and the peripartum period.

2.1.3 Regulatory requirements and informed consent

Many funders require that researchers receive prior approval from their institutional ethics committees before the submission of a grant proposal and before the start of the actual research (Michener, 2015). The NESHIE study is no exception. Ethics approval was obtained from the UP REC (primary ethics committee) and all other participating institutions for the NESHIE project. Permission was also obtained for the research to be conducted at the respective institutions included in the NESHIE study. This information needs to be recorded in the DMP. While the NESHIE study described this information using bullet points, this information can also be tabulated. The regulatory body's name, study approval number and approval date should be noted. If additional approvals are required for the purposes of a study, for example, hospital or internal review committee approvals, the DMP should capture this information as well.

However, since study documents may require amendment during the course of a study, it is also important that a record of changes to study records be kept. While this data may not necessarily be reflected within a DMP, reference to its storage location should be made, as is the case for the NESHIE study DMP. An example of a document tracking log is reflected in Table 3.

Depending on funder or regulatory body requirements, it may be necessary to explicitly state what the inclusion and exclusion criteria for participation in a study are. While this is the case for the NESHIE study, the DMP may also simplify this by providing a reference to the relevant study document. These details may be provided in text or table format and are at the discretion of the author of the DMP. This information is often provided in order to contextualize the informed consent process since ethics approval

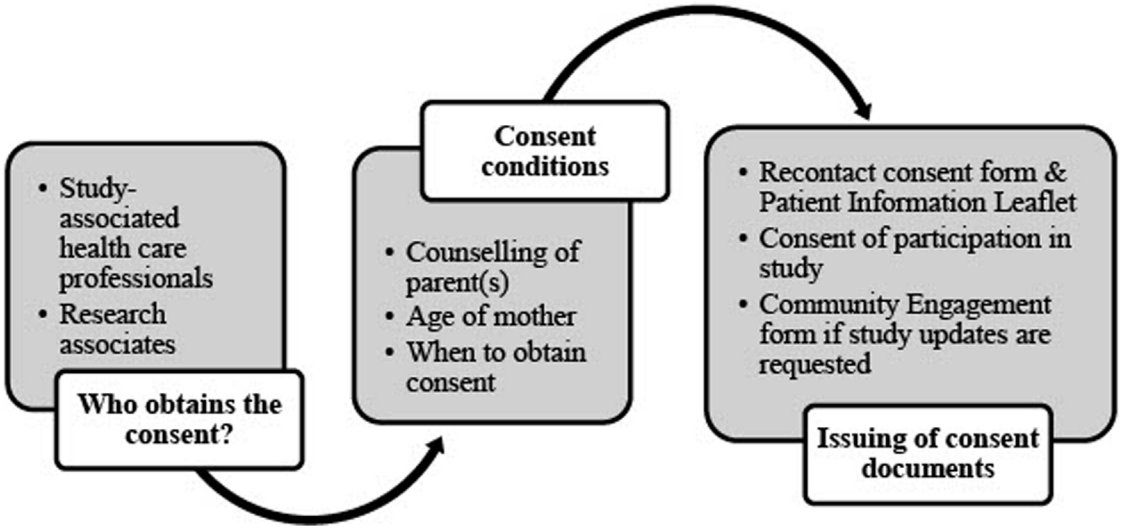

FIGURE 4
Informed consent process in the NESHIE study. Image created by JVR and AS.

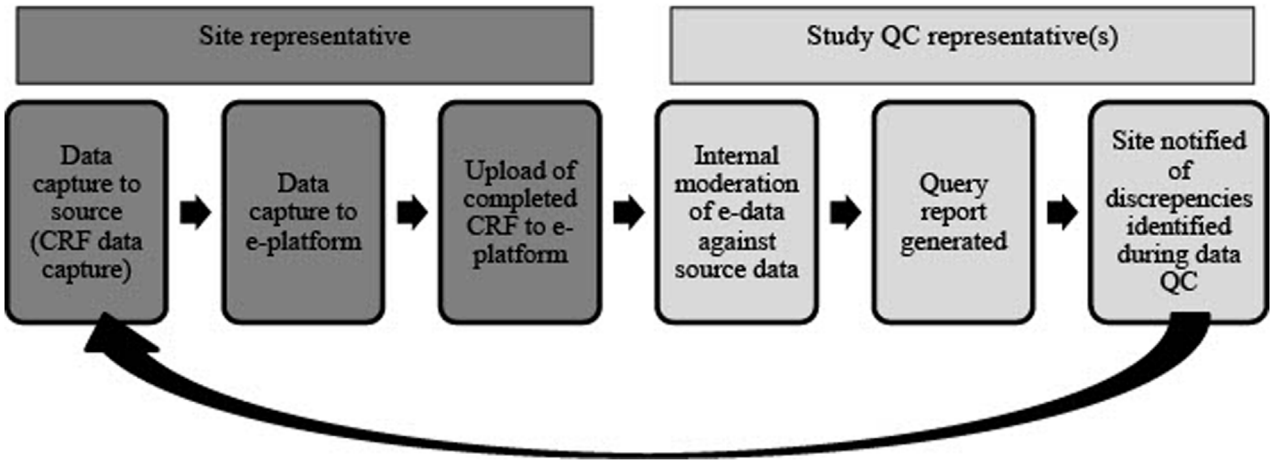

FIGURE 5
An overview of the QC process for study-associated NESHIE data. This review cycle is repeated until all data queries are cleared and data are validated as accurate. Image created by JVR.

requires that informed consent be granted by research participants, that data are de-identified, and that data access and use is restricted.

An informed consent form outlines the terms of research participation and may include or exclude future usage of data (Hardy et al., 2016). Researchers may not re-analyze research data in any form when the informed consent form excludes the future and unrelated usage of that data. In such cases, new applications must be submitted instead of amending current ethics approvals. The DMP needs to refer to the consent documents and where they may be found. Similarly, it describes the consent process and associated vetting thereof. This includes the quality control process in ensuring the validity of each participant's signed consent form(s). The informed consent process is explained in the NESHIE DMP and summarized in Figure 4.

2.1.4 Quality assurance and quality control processes

Quality assurance (QA) and quality control (QC) processes measure, assess, and improve the quality of data, software and other study-related factors. According to the nature of the study and degree of research funding, specific QA/QC guidelines may have to be followed (Michener, 2015). It is however good practice to provide a description of the QA/QC measures employed in a research project, which may involve training, the calibration and verification of instruments, and double-blind data entry. It is critical to state who will be responsible for performing the various QA/QC tasks, when they will be performed, how frequently they will be performed, what potential problems are expected and what contingency plans are in place for this. If templates are used for

the purpose of QA/QC and not already included in the protocol, these should be reflected in the DMP and adjusted according to the needs of the study (as is the case for the NESHIE study).

Quality control (QC) reports for the NESHIE study are done manually or by using REDCap's Data Resolution Workflow which is an inbuilt data query function in REDCap (Figure 5). QC for placental histology and de-identified scans/images are conducted using standardized report forms from blinded individuals. This is done to account for any inter-observer bias. Unblinded QC is performed for informed consent documents and other core anonymized clinical data (e.g., severity grading) using standardized report forms.

2.1.5 Ownership

The DMP must indicate ownership of the data (Fadlelmola et al., 2021). According to Thaldar et al. (2022), ownership in law implies a relationship between the owner and the object (or thing) in respect of which the owner acquires certain legal rights and entitlements. Neither a hospital site nor a research project can acquire, exercise, or enforce any of the rights contemplated in terms of legal ownership. As a consequence, NESHIE research data would belong to the Principal Investigator (PI), Prof Michael S. Pepper, as prescribed by Thaldar et al. (Thaldar et al., 2022).

Intellectual property including Material Transfer Agreements between collaborators and the NESHIE study are also mentioned in the DMP and are listed in the Annexure but retained separately to the DMP owing to confidentiality. Commercialization has been considered, but as the study has not yet reached a stage where this is relevant, a specific framework for commercialization has not yet been developed. Funding agencies for the NESHIE project are clearly indicated. Publication outputs are reflected in the NESHIE study DMP on a regular basis.

Because the NESHIE study is a collaborative project, all collaborators and their roles in the study are provided in the DMP. This list can be updated as additional collaborators join the study and includes the following:

- all hospital sites where participants will be recruited;
- participating academic institutions and their associated clinical facilities and representatives;
- collaborators assisting in establishing systems for sample processing and sample quality control protocols;
- national and international centers where genomic, transcriptomic, proteomic and metabolomic data will be processed and analyzed; and
- collaborators assisting with MR imaging.

2.1.6 Auditing

The DMP should include the details of an audit plan for a clinical trial, or in the case of a research project, refer to the relevant SOPs. Since it is an observational study, the NESHIE study does not have a formal audit plan as stringent as what would be found in a clinical trial. More specifically, auditing for the NESHIE study is performed internally, rather than by auditors/monitors contracted externally to the study. Nevertheless, auditing of the data is a critical component of the study to ensure that data of the highest quality is reported in the public domain. As such, auditing of the data ties closely to data QC processes and can be viewed as a two-step process. Firstly, a data manager will ensure that data captured to case report forms is concordant to what is captured to the study's electronic data capture platform. Where discordance or data missingness is noted, data queries are raised for the study site to attend to until data concordance is observed. Once data is concordant across sources for each section of data, the data section ('instrument') is locked. This first step of data validation/auditing is performed on a continual basis. The second phase of data auditing involves senior members of the research team (project manage and molecular data manager) who review data from all locked instruments. This involves several point and cross-sectional data checks. If no further data errors are noted, the participant record is locked in entirety. The second phase of data auditing is largely dependent on the complete review and locking of data instruments and therefore occurs as needed at variable intervals. In addition, audit trails are automatically generated through the various electronic data collection platforms used as part of the NESHIE study, are downloaded on a weekly basis by the data manager, and retained on an independent local server where periodic checks for completeness are conducted by the project manager. Spot checks of study records are also conducted by the project manager to ensure that enrolled participant's data records are complete. Each study will need to tailor its data auditing needs according to its aim(s), objectives and long-term aspirations.

2.1.7 Website

The DMP should provide details of a website if one has been created for a research project or clinical trial. A website for the NESHIE project is under development and is aimed at providing general information on HIE, for example, as well as current research and publications.

2.2 Data characteristics

2.2.1 Data collection sheets

A brief description of the data/sample collection sheets (case report forms or CRFs) should be provided in a DMP. This may include the development of the collection sheets, general guidelines for the completion of these collection sheets, amendments and recordkeeping practices. The NESHIE data collection sheets have been designed by experts. For example, neonatologists have developed and vetted the neonatal data collection sheets, while obstetricians have developed and vetted the maternal and obstetrics data collection sheets. The design was based on local and international best practices and within the South African context where applicable. Guidelines for the completion of the data collection sheets are provided in the NESHIE study documents, during training of the research team, and/or within the REDCap database. A list of all amendments to the design of the data collection sheets should be submitted to an IRB/REC prior to enforcement, with redline copies kept for the study's recordkeeping purposes. This record is maintained by the project manager of the NESHIE study. An electronic record is also maintained of all the CRF versions.

TABLE 4 Tools used to generate and submit data in the NESHIE study.

Generated Document/Data	Submission
Informed Consent Forms (ICFs)	☐ Electronic: Site and Study records
	☐ Hard-copy: Site records
Community Engagement Form(s)	☐ Electronic: Site and Study records
	☐ Hard-copy: Site record
	☐ Hard-copy: Study Master File
Clinical data storage	☐ Electronic: Site and Study records
	☐ Hard-copy: Site records
	☐ Hard-copy: Study Master File
Basic data for sample storage	☐ Electronic: Sample storage facility (E.g. biorepository)
	☐ Electronic: Site and Study records
	☐ Hard-copy: Site records
	☐ Hard-copy: Study Master File
Molecular data storage	☐ Electronic: Sample analysis facility
	☐ Electronic: Site and Study records
	☐ Hard-copy: Site records
	☐ Hard-copy: Study Master File

2.2.2 Data submission

A DMP for a research project or clinical trial should describe how the data is submitted, whether it is done manually or electronically, or even both. This should also include the tools being used to generate the data, for example, standardized case report forms and electronic data capture platforms (eDCPs).[8,9,10] The mode through which data will be submitted should be compliant with the various guidelines and policies applicable to the research being conducted. For example, if performing a clinical trial and enforcing GCP guidelines, it is necessary to have access to the source documents.[11] In clinical settings, this often relates to accessing hospital records. However, if obtaining hospital records (source documentation) is challenging, the DMP should provide a clear set of instructions on when and how photocopies of the hospital records should be made for the research site's reference and for auditing purposes.

In the NESHIE study:

- all clinical data are captured manually on data capture sheets or CRFs and these serve as the source documents;
- de-identified electronic copies of the source documents are then uploaded onto REDCap; and
- documents with identifying information necessary for QC purposes are uploaded to the study-associated server using LogicalDOC.

Data verification is done at the hospital site where information is captured onto the CRFs and when data are captured onto the eDCP(s). Each NESHIE participant's record is linked to an electronic copy of the CRF. This includes capturing the barcode links affiliated with the specific samples that were collected from each consenting participant and sent for long-term storage in an accredited biorepository. Long-term storage of NESHIE samples is undertaken by Clinical Laboratory Services (Braamfontein, Johannesburg, South Africa) which is a highly reputable biobank in South Africa. This process, which includes the means through which Clinical Laboratory Services captures barcoded sample data to its laboratory information management system and disseminates it, is detailed across the NESHIE DMP and protocol. A simple template that can be modified and used for the purposes of capturing this information is reflected in Table 4.

2.2.3 Sensitive and confidential data

Data sensitivity should be covered in the DMP and an explanation should be provided about how it will be treated (Hardy et al., 2016; Fadlelmola et al., 2021), for example, compliance with research institutional policies or national legislation. It is therefore mandatory that researchers who collect and analyze data should have training on ethical practices which include confidentiality, informed consent and data protection (Hardy et al., 2016). Sensitive data from vulnerable communities require more stringent guidelines. Thus, when a research collaboration becomes large and complicated, the project leads or principal investigators may require multiple approvals for just one project (Hardy et al., 2016). A DMP should include clear instructions on how sensitive or confidential information will be collected, protected and used.

11 SAHPRA. (2023). South African Good Clinical Practice: Clinical Trial Guidelines. https://www.sahpra.org.za/wp-content/uploads/2021/06/SA-GCP-2020_Final.pdf [Accessed 24 July 2023]

Within the NESHIE study DMP, all clinical and molecular data are considered to be confidential due to the vulnerability of the research participants (mothers and infants). Where access to sensitive information is required for the purposes of QA/QC, the DMP clearly indicates the responsible parties for these processes. These processes describe how, when, and the frequency with which this information will be accessed. Similarly, to prevent the identification of individuals, aggregated research findings are reported.

2.2.4 Data identifiability

Participant anonymization is a requirement for all research studies and clinical trials involving human participants when releasing data into the public domain. The protection of research participants' data necessitates constant supervision. Therefore, when a coding system is used to ensure participant anonymity for this purpose, this must be explained in the DMP. Several methods for participant anonymization are available; Rodriguez et al. (2022) provide an excellent review of the many methods used for this purpose.

All research participants in the NESHIE study are assigned a random alphanumeric code at the time of participant screening to protect their privacy. This unique identifier remains unchanged throughout the entirety of the study and is applied equally to clinical and sample-associated data. Individual clinical data are not publicly released, however aggregated data (a form of *k*-anonymization) is consensually made publicly available through journal publications to further protect participants identities. Additional sample anonymization is applied once samples are deposited in a biobank. Regarding sample-associated data, while identification of research participants is very low with proteomic, metabolomic, and transcriptomic data, it is not possible to guarantee the absence of re-identification with genomic data. Nevertheless, metadata for sample analysis in the NESHIE study is typically limited to variables such as sample source (e.g., heel prick), time of collection relative to the time of birth, disease severity and the receptacle(s) used for sample collection (e.g., Whatman 903 protein-saver cards).

However, it is important to note that during the data/sample collection process, some members of the research team will have full access to personal/identifying information. Such detail must be clearly explained within the DMP. In the NESHIE study, for example, this will include those individuals responsible for collecting data from hospital records, as well as those responsible for ensuring that placental histology slides are obtained as part of the study's QC procedures. As such, each site is expected to maintain a site-specific master list which links patient information to the Study IDs. These lists are not stored at the central database level for the purposes of the NESHIE study but must be retained at site-level and maintained at all times. Nevertheless, most members of the team will have access to pseudonymized information in order to fulfil their various roles and responsibilities.

2.2.5 Data updates

The DMP should explain how data will be updated or become redundant when revisions are made and subsequent CRF versions are produced. This also applies to the NESHIE study; the DMP is revised and updated when necessary. Although some data may become redundant when revisions are made, information is retained from any system that is updated for recordkeeping and backup purposes. All changes made on the NESHIE study documents or REDCap database are reflected in the amendments submitted to the REC. A record is kept of all changes made on the NESHIE study documents while data dictionaries are kept that reflect all changes made on the different REDCap versions. This process is explained in the DMP.

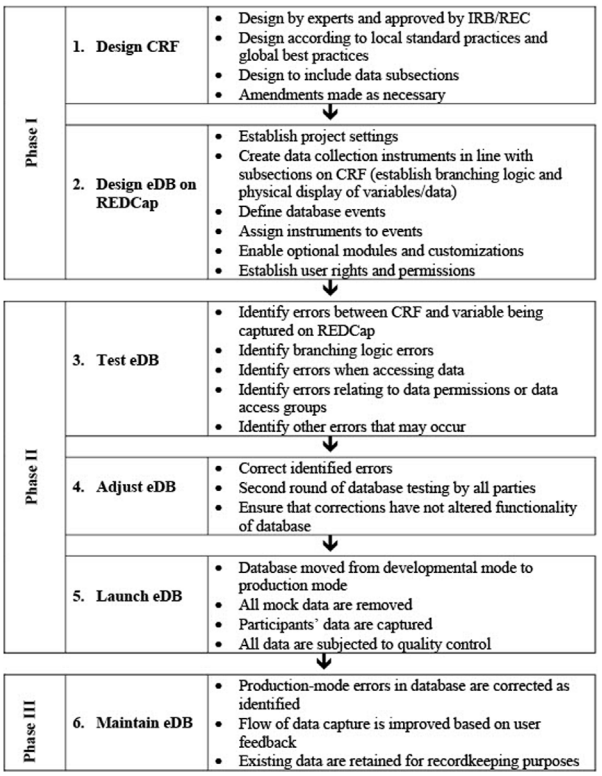

FIGURE 6
The overall flow of database creation and maintenance in the NESHIE study. Image created by JVR and AS.

2.2.6 Data reporting

The requirements and frequency of data reporting should be described in the DMP, and where feasible, should be reflected in the study-associated timeline. In the NESHIE study, the data reporting requirements start when a potential participant is screened for inclusion. Data is reported after validation through the entire QC process and only fully validated data is published. Internal reporting regarding screening, enrolment and sample collection is typically done monthly to the PI but may be adjusted according to study needs. Data obtained from sample analyses is reported on an ongoing basis when sufficiently validated.

2.3 Samples/data storage and security

2.3.1 Database design, creation and maintenance

The process involved in establishing an electronic database platform should be described in a DMP, starting with the requirement analysis, conceptual, logical and physical design,

TABLE 5 Examples of clinical, metabolomic, genomic and imaging data inputs and formats for the NESHIE study.

DATA	Software used	FILE formats
Clinical	• SPSS Statistical Software	.csv and.xlsx files
	• Stata Statistical Software	
	• SAS Statistical Software	
	• R Statistical Software	
Metabolomic	• Liquid Chromatography-Mass Spectrometry data files	.csv and.xlsx files
Genomic	• Illumina short read sequencing	FASTQ
	• BGI-based sequencing platforms	
Imaging	• MRI: Hyperfine Cloud Picture Archiving and Communication System (PACS) or local hospital PACS	• MRI: DICOM files
	• CUS imaging software	• CUS: .jpeg files

until the launch and maintenance of the database(s). Figure 6 summarizes the process for the establishment of the REDCap database for the NESHIE project.

2.3.2 Data input and processing

Guidelines for data entry and data processing should be explained in the DMP. In the NESHIE DMP, data entry and processing are described with regard to the CRFs and the REDCap database. Data entry on the CRFs is guided through instructions on the forms and through regular training of research associates. The REDCap database provides prompts that guide the data input. Data anomalies or missing data are managed through the QC process. A single data entry method is used with on-site and off-site verification of data. Hospital site clinical appointees make corrections on the paper documents and the data manager makes changes on the electronic data. The data manager also keeps records of the data logs/audit trails on REDCap and changes made to the CRFs. Data generated through the NESHIE study is not linked to external/third-party databases except where expressed consent is provided for this (e.g., submission of molecular findings to a data repository), or where it is necessary as part of the sample/data analysis process (e.g., analysis of MR images).

2.3.3 Data format and transformation

As technology changes, some current data and file formats may become obsolete (Michener, 2015). Therefore, a good choice of data and file formats includes those that are non-proprietary and commonly used within the associated research field, for example, comma separated values (CSV) as a replacement for Microsoft Excel® formats (.xlsx). Data formats should be consistently applied throughout the duration of the study and must also be considered in relation to archiving conditions. The study-related data formats must be disclosed in the DMP. An example of the data formats used in the NESHIE study, as described in the DMP, is found in Table 5.

In the NESHIE study, while some molecular data may undergo file format transformations, most of the file formats are accessible in common analysis platforms and are accessible for prolonged periods of time. As such it is important to note that data export options must be reflected within the DMP. Data export should be considered relative to the export for internal review of data, as well as external data review/sharing (as is the case for repositories). Personal health information should be removed from all datasets where authorization to view such data has not been granted and would constitute unethical behavior if shared. Data export will be covered in more detail in a subsequent section.

2.3.4 Data standard and metadata

According to Michener (2015), "metadata are the details about what, where, when, why, and how the data were collected, processed, and interpreted". Metadata allows for the discovery, use, and the accurate citation of data and files. It explains the names, structure, and storage of data and files and details the research environment, experiments, and analyses. A good documentation strategy includes the following three steps (Michener, 2015; Fadlelmola et al., 2021): First, identify the types of information that must be collected which will allow the discovering, accessing, interpretation, usage, and the citation of data. Second, determine whether a community-based metadata standard should be implemented, for example, ICD-11 (International Classification of Diseases, 11th ed). This may be required by a data repository, archive, or domain professional organization. Third, identify software that can generate and manage metadata content. As such, a DMP should adequately describe how the data will be formatted and standardized for present and future use, whether or not a data dictionary is available (if REDCap or similar electronic data capture platform is being used), whether file naming conventions are being applied (e.g., HGNC ID for naming of genes)[12], whether existing metadata is sufficient for data interpretation, how CRF versions and associated data will be tracked across data records, and whether data will adhere to FAIR principles.

In the NESHIE study, formatting of variables and their associated inputs/outputs is reflected in the NESHIE code book. This code book was formulated with the construction of the

12 NCBI. (2023). Guideline for Human Gene Nomenclature. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7494048/#:~:text=Each%20gene%20is%20assigned%20a,or%20%E2%80%9CG%E2%80%9D%20for%20gene [Accessed 24 July 2023]

TABLE 6 Technical aspects of electronic platforms to report in a study DMP.

	Electronic platform name
Installed version	Version # (Date: XXX)
Updated version	Version # (Date: XXX)
Overview description	E.g., Web-based interface for user/server interaction; Document repository? ('Yes', 'No', 'Limited'?); Clinical research platform for data captureetc.
Technical description	E.g., TomCat-based document storage and management; PHP-Apache-MySQL-based clinical research platformetc.
Security	E.g., Secure encrypted website using LetsEncrypt certificates; Dependent on IT infrastructure and environment of host server; Multi-level securityetc.
Servers	Example: Number of servers: web data and database Number of servers: back-up All servers hosted: [insert location name(s) of each server]
Web server requirements	E.g., TomCat 8.5 or higher; PHP 5.3.0 or higher
Database server requirements	E.g., MySQL 5.0+; MariaDB 5.1+
SMTP email server	E.g., Configuration with PHP required on web server
File server	E.g., Files stored on file system of database server; may be separate from database server; files stored behind firewall location for study; WebDAV application available when firewall storage is not used
SSL certificate required	Yes/No
File or Data storage method	Ext4 file system; MySQL back-end with PHP front-end
3rd party server access or use	Not applicable/Applicable [list 3rd party server access/use]
User privileges	Multi-level user privileges at system level (broad or constricted/limited); System administrator assigns initial user privileges at onset
Authentication	Validation of end-users required (State specific authentication methods if applicable)
Auto-logout function	Yes/No
Logging and audit trail	Yes/No
File or Data import function	Yes/No (Describe if 'Yes')
File or Data export function	Yes/No (Describe if 'Yes')
File or Data interoperability	Import files or folders stored at remote locations using FTP, SFTP, *etc.*,; API via API tokens; data import and export; Dynamic Data Pull (DDP) via web service; data import only

REDCap database and allows the interpretation of the data. Since REDCap is used, a data dictionary is available for each iteration of the database, as are CDISC ODM files. File naming conventions for clinical data is not used but is applied to molecular data in accordance with external collaborators associated with sample analysis, funder guidelines and other standard practices. Sufficient metadata exists for data interpretation, and CRF version details are captured to the REDCap system for each participant record. FAIR principals are applied as reasonably as possible due to the vulnerability of the participants/study population.

With regard to data sharing and associated standards in the NESHIE study, while these typically do not apply to the clinical data being collected since a single condition is under investigation, such standards may be relevant for other studies and are important for consideration (e.g., SNOMED/CT or ICD-10/11). Additionally, there are currently no formal reporting standards to describe the metadata at a dataset level. In relation to sample data that would be shared as part of the NESHIE study, standard gene identifiers such as HGNC IDs and rsIDs are used for the genetic and transcriptomic data, while gene annotations (e.g., Gene Ontology) are used for proteomic and metabolomic data. All magnetic resonance image files are generated in a DICOM format.

2.3.5 Data storage and database security

Research projects consisting of multi-sectoral collaborators (including the community or institutional ethics committees) may require significant data protection to prevent data access breaches that may violate participants' rights to privacy (Hardy et al., 2016). Besides noting data security structures in place for a study, the DMP should also explain the storage and protection of the data during the lifecycle of the project (Michener, 2015). Protection of data must include a guide as to how many copies and what format copies of data should take. Copies of data can be in hard- and/or soft-copy format. The storage location of copies should be made known to relevant study representatives; access to data copies should be restricted according to the role and responsibility each research associate holds within the study. A regular backup of the data should also be scheduled.

With regard to the NESHIE study, a three-layered approach to backing up server-hosted data has been taken. First, there is a daily back-up of captured data to a local server. Second, there is a weekly back-up of the entire data ecosystem (which includes uploaded

documents), also to a local server. Thirdly, there is a weekly back-up of the entire data ecosystem to a local server that is independent of the server(s) used for points one and two. This backup schedule is consistently tested to ensure the retrieval of the stored data files. Log files are received and viewed by IT specialists to ensure that these systems operate as expected. Such information may be presented in different ways in the DMP but should be appropriate and relevant for each individual study. The flow of storage and sharing of samples and data for analysis purposes in the NESHIE study is detailed in the DMP using a figure, while security considerations regarding the NESHIE electronic database platforms is described in a table format. A template of the key security considerations used for this latter purpose is summarized in Table 6. Internal review of security settings are usually checked and confirmed at the point of user creation and assignment to projects. In REDCap, this is tested by utilizing the 'view project as' function, while LogicalDOC facilitates a visualized output of user security settings. In both instances, these functions are reserved only to those individuals within the study assigned with admin rights.

Due to the sensitivity of the data collected for this study, cloud storage of our clinical data is not being used. However, data repositories such as the European Genome Phenome Archive (EGA) are being considered for sample-associated data storage under contract and accessibility by way of a data access committee (DAC). Data transfer agreements would stipulate data transfer conditions, including those pertinent to POPIA regulations.

2.4 Sample/data access and sharing

2.4.1 Sample/data sharing intentions and access requirements

Even though good data protection and sharing policies might be in place, this does not mean that data will not be shared (Michener, 2015). Data can be disseminated in a passive or active way. Passive sharing includes posting data on a website or emailing it. Active dissemination, which is preferred, includes submitting the data to an open repository or archive, or publishing the data as articles/Supplementary Material in peer-reviewed research or data journals. Data can be correctly cited by using the guidelines and mechanisms provided by journals and repositories, for example, DOIs. This ensures that researchers are accredited for their data products. Furthermore, data will be more user-friendly and interpretable if it is distributed through standard, non-proprietary approaches. The data should then also include metadata and code which will enable data processing. Licensing and copyright may involve allocating an identifier that is unique to a dataset (Fadlelmola et al., 2021). This will facilitate data discovery and any legal issues when reusing the data.

When writing about the sharing of data in the DMP, it is important to consider data responsibility, accountability and authority (Fadlelmola et al., 2021). These considerations are usually stated in data protection policies such as POPIA or GDPR. Protecting the rights of research participants, especially those who are from vulnerable populations, has become crucial. This has resulted in the adoption of data protection policies globally to provide this framework for both the research participants and data users. This is especially relevant in genomic research as individual genomes are considered personally identifiable information even after participant anonymization (Fadlelmola et al., 2021).

As described by Michener (2015), DMPs should include policy statements regarding the management and sharing of data. These should bear, at the very least, reference to:

- Licensing or sharing arrangements about the use of pre-existing materials;
- Arrangements for retaining, licensing, sharing, and embargoing data, code, and other materials; and
- Legal and ethical restrictions on access and use of sensitive data from research participants.

The level of data access should be determined by a study's management team with sample/data access and sharing plans approved by an institutional ethics committee/board and/or study-associated data access committee (Fadlelmola et al., 2021). The DMP explains whether access is limited or open, who has access to the study-related data, and whether access to the data has to be approved by a DAC. This applies to the NESHIE study where only individuals associated with the study or duly appointed study-associated representatives have access to the data and samples; Table 7 provides a template of how this can be presented within a DMP. Metadata export for the NESHIE study can currently only be done by a limited number of people when necessary for sample/image analysis.

Importantly, local legislation may require that additional sample and data access-related documents be noted within a DMP. In South Africa, sample analyses not being performed in the country require an accompanying export permit. A material transfer agreement (MTA) between the study's host institution and the representative organization of another country is required as part of the application for an export permit. Memorandums of understanding are established as supporting documents to MTAs and may describe sample and/or data processing requirements in more detail. This principal similarly applies to the transfer of data and the need for signed data transfer agreements (DTAs). DS-I Africa Law Research Group from the University of KwaZulu-Natal recently presented a DTA template that is an excellent resource for the provision of DTAs in the South African research context (Swales et al., 2023).

2.5 Data archiving

2.5.1 Retention period

The archive period of data for research purposes should be mentioned in a DMP. Sample and data retention periods are usually stipulated in the institutional regulations/policies and in funder guidelines. The NESHIE data will be archived for at least 15 years from completion of the study according to UP regulations. When access to data is required for analytical purposes, it should be able to be uncompressed, unencrypted, and decoded from standard character encodings such as 16-bit Unicode Transformation Format or UTF-16 (Michener, 2015).

2.5.2 Repositories

Different digital data repositories are available that provide secure and remote access to their web-based platforms (Antonio

TABLE 7 An example of data and sample access specifications.

Clinical data	
Where is the data stored?	
Who does the data belong to?	
Who/what institution has access to datasets for analysis purposes?	
Who/what institution is permitted to perform the analysis of data?	
What platform(s) will be used to perform the data analysis?	
REC reference institution/approval numbers	
Study-associated samples: Specific (e.g. Blood)	
Where are samples stored: *Once collected?* *After analysis?*	
Who is permitted to perform the sample nucleic acid isolations?	
Who/what institution is permitted to perform the nucleic sequencing?	
Who/what institution is permitted to perform the analysis of the nucleic acid sequences?	

et al., 2020). These data repositories store large data sets and are supported by funders and government agencies. This facilitates data sharing across research teams or sharing datasets for a single study where a formal data access application process has been approved. Academic institutions may also provide data repositories that support private and selectively restrictive institutional access to multiple research studies. When selecting a data repository, there are three considerations: physical features (servers and hardware), technical features (software), and administrative features (personnel requirements and support, policies, security and data access). The key function of a data repository is to support secure data management for geographically dispersed research institutions and to provide secure data access, storage, and sharing (Antonio et al., 2020).

If data is archived in an unsecure location for a long-term period, both the researchers and others may not be able to use it as it becomes inaccessible (Michener, 2015). The description of the storage and preservation of data are therefore essential to any good DMP. In other words, three questions have to be considered (Michener, 2015):

1) "How long will the data be accessible?";
2) "How will data be stored and protected over the duration of the project?"; and
3) "How will data be preserved and made available for future use?".

Several factors are involved when answering the first question. First, research funders or institutions may have specific requirements. Second, the core value of the data should be considered in relation to the ease with which it can be generated independently of the initial study. To answer question 3, a robust solution may be necessary to access data 20 years after the finalization of a project.

Funders and research institutions may have identified appropriate data repositories for specific research areas. Certain disciplines maintain specific repositories, for example, GenBank is a repository for nucleotide sequence data. Universities may host institutional repositories or general science data repositories, for example, Figshare[13]. Alternatively, there are the Registry of Research Data Repositories[14] and BioSharing[15] which are discipline-specific and general repositories via online catalogues. The DMP should note the policies of the selected repository, specifically for data privacy and security (Fadlelmola et al., 2021).

Funders may provide a list of approved repositories for data archiving, but if these repositories do not have the required functionalities or compliance to participant consent conditions, researchers may request the use of other repositories. In terms of repositories for the NESHIE study, several repositories recommended by one of the funding bodies were considered and included: MassIVE, Panorama, Pride (Proteomics Identifications Database) and Metabolomics Workbench. However, the UP REC requires that all NESHIE sample-generated data be archived in a repository that is safeguarded by a DAC. Therefore, with approval from the funding body, it was decided that the EGA will be predominantly used for the purposes of the majority of the NESHIE study 'omics' data. NESHIE clinical data is not available for public access.

2.5.3 Future use of data and the data access committee (DAC)

The DMP should provide details regarding how data may be used in the future, for instance, in publications, industry involvement or commercialization. Consideration should also be given to potential funding and/or publication requirements regarding data access. The participant informed consent form should clearly indicate any

13 Figshare. (2023). Figshare as a data repository. https://figshare.com/ [Accessed 9 November 2022]

14 Re3data. (2023). Registry of Research Data Repositories. https://www.re3data.org/ [Accessed 9 November 2022]

15 BioSharing Network. (2023). Blood donor repository. https://bio-sharing.org/ [Accessed 9 November 2022]

envisaged future use of data and must serve as the foundation for future access to and use of study-related data.

If applicable, the DMP should provide a description of a study-related DAC. This description explains the role of the DAC and the documents that govern the terms and conditions on which access is granted to the data by the DAC. Access management through a DAC endorses the benefits of data sharing while diminishing the risk of uncontrolled access to study data that is generated from vulnerable/at risk study populations for uses that may fall outside of the purview and restrictions established by the study and its associated consent conditions. Requests are approved or rejected by the DAC rather than having open access without restrictions (Cheah and Piasecki, 2020). In other words, the DAC regulates access to all data generated from a funded research project. The DAC's purpose includes both the promotion of data sharing and the protection of research participants and their communities, researchers, and research institutions. The establishment of the DAC should adhere to institutional and legal policies together with clear distinctions of responsibility, terms of reference and membership (Cheah and Piasecki, 2020). To accomplish its role, a DAC's members should represent relevant areas of expertise. Some members may be independent as this will address the issues where conflicts of interest may be involved. The application procedure for data access should be transparent, consistent and simple. The data sharing policies of institutions should provide guidelines for the review process while independent DACs should be guided by pre-agreed terms. Elements of the review should include the applicants, what the objectives of data reuse are, which data are requested, and the potential benefits and risks involved. In order to monitor the flow of data requests and the decisions made thereto, one could either design a custom database or use an equivalent platform in a public domain, such as Resource Entitlement Management System (REMS)[16].

Sustainability of a fully-functional DAC is a challenge faced in many studies. This should be specifically addressed and may need to be institutional in order to be sustainable. At the very least, it should not be constituted by study members that form part of a mobile community. Nevertheless, a lack of resources/support at an institutional level for the sustainability of a DAC should not prohibit a study from pursuing such endeavors. This does however create additional expectations for the study to appropriately plan and budget for the establishment and maintenance of a DAC. It is a recognized imperfect system but is often the only means through which 'omics' data generated on vulnerable populations may be disseminated. It is therefore in the best interest of studies involving such populations to carefully consider the long-term requirements of a DAC during the development phase of the study to ensure that an appropriate strategy may be put in place for its ultimate success. Within the NESHIE study, the onus falls on the study to provide the resources and support for the establishment and maintenance of the DAC.

16 GitHub. (2023). REMS. https://github.com/CSCfi/rems [Accessed 5 September 2023]

3 Conclusion

A data management plan should provide a user-friendly road map that guides and explains the governance of data throughout the duration of a research project and also after the conclusion thereof. The DMP template presented here, as drafted from the NESHIE study DMP, provides a thorough design or framework that will require approval by reviewers and funders, and that can be applied to a research study or a clinical trial being conducted in vulnerable populations and/or employs muti-omics analysis methods to achieve the study aims. While there should be limited duplication between the protocol and DMP, the DMP does not need to be a self-standing document if the inclusion of protocol information is able to reasonably add value and context to the content of the DMP. The template can nevertheless be adjusted to reference the relevant protocol section(s) where necessary, should a self-standing DMP be required.

The NESHIE DMP remains a dynamic, living document, that will continue to change as legal, ethical, funding, publishing, resource-related and other factors evolve. As an example, the study's approach to informed consent and the associated documentation and data storage was adjusted with the implementation of POPIA. As a second example, data deposition to a repository was initially not approved as part of the initial ethics approval process. Data deposition to a repository has subsequently been approved, but with important considerations regarding data security given the vulnerability of participants enrolled into this study. As a final example, the NESHIE DMP did not need to make consideration for magnetic resonance image storage at the study's onset. This changed with the introduction of a portable ultra low-field magnetic resonance which required careful consideration and adjustment to the DMP regarding image acquisition and data transfer. It is therefore unsurprising that this DMP will continue to be adjusted as the project continues to evolve.

Nevertheless, it is clear that DMPs are adaptable in part or in whole and can be applied to successive research studies or clinical trials. An accessible DMP may support researchers and funders in data discovery and future collaborators, provide education on data management, and may monitor compliance with policies and regulations. In considering the nature of DMPs, future work will therefore include describing the lessons learnt throughout the NESHIE study in relation to adjustments to the DMP, particularly regarding the areas prone to change. Future work will also focus on creating a machine-readable DMP version.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

AS: Writing–original draft. JVR: Writing–review and editing. MP: Supervision, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the South African Medical Research Council (GIPD 96783) and the Bill and Melinda Gates Foundation (INV-022216 and INV-041062).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1273975/full#supplementary-material

References

Antonio, M. G., Schick-Makaroff, K., Doiron, J. M., Sheilds, L., White, L., and Molzahn, A. (2020). Qualitative data management and analysis within a data repository. *West. J. Nurs. Res.* 42 (8), 640–648. doi:10.1177/0193945919881706

Cheah, P. Y., and Piasecki, J. (2020). Data access committees. *BMC Med. Ethics* 21 (12), 12–18. doi:10.1186/s12910-020-0453-z

Dunie, M. (2017). The importance of research data management: the value of electronic laboratory notebooks in the management of data integrity and data availability. *Inf. Serv. Use* 37 (3), 355–359. doi:10.3233/ISU-170843

Fadlelmola, F. M., Zass, L., Chaouch, M., Samtal, C., Ras, V., Kumuthini, J., et al. (2021). Data management plans in the genomics research revolution of Africa: challenges and recommendations. *J. Biomed. Inf.* 122, 103900. doi:10.1016/j.jbi.2021.103900

Hardy, L. J., Hughes, A., Hulen, E., and Schwartz, A. L. (2016). Implementing qualitative data management plans to ensure ethical standards in multi-partner centers. *J. Empir. Res. Hum. Res. Ethics* 11 (2), 191–198. doi:10.1177/1556264616636233

Lefebvre, A., Bakhtiari, B., and Spruit, M. (2020). Exploring research data management planning challenges in practice. *it - Inf. Technol.* 62, 6229–6237. doi:10.1515/itit-2019-0029

Michener, W. K. (2015). Ten simple rules for creating a good data management plan. *PLoS Comput. Biol.* 11 (10), e1004525. doi:10.1371/journal.pcbi.1004525

Nourani, A., Ayatollahi, H., and Dodaran, M. S. (2022). Data management in diabetes clinical trials: a qualitative study. *Trials* 23 (187), 187. doi:10.1186/s13063-022-06110-5

Rodriguez, A., Tuck, C., Dozier, M. F., Lewis, S. C., Eldridge, S., Jackson, T., et al. (2022). Current recommendations/practices for anonymising data from clinical trials in order to make it available for sharing: a scoping review. *Clin. Trials* 19 (4), 452–463. doi:10.1177/17407745221087469

Swales, L., Botes, M., Donnelly, D., and Thaldar, D. (2023). Towards a data transfer agreement for the South African research community: the empowerment approach. *South Afr. J. Bioeth. Law* 16 (1), 13–18. doi:10.7196/SAJBL.2023.v16i1.827

Thaldar, D. W., Townsend, B. A., Donnelly, D.-L., Botes, M., Gooden, A., Van Harmelen, J., et al. (2022). The multidimensional legal nature of personal genomic sequence data: a South African perspective. *A South Afr. perspect.* 13, 997595. doi:10.3389/fgene.2022.997595

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci. data*, 3160018. doi:10.1038/sdata.2016.18

Williams, M., Bagwell, J., and Nahm Zozus, M. (2017). Data management plans: the missing perspective. *J. Biomed. Inf.* 71, 71130–71142. doi:10.1016/j.jbi.2017.05.004

frontiers | Frontiers in Pharmacology

TYPE Review
PUBLISHED 14 December 2023
DOI 10.3389/fphar.2023.1297353

OPEN ACCESS

EDITED BY
Athanasios Alexiou,
Novel Global Community Educational Foundation (NGCEF), Hebersham, Australia

REVIEWED BY
Matjaž Perc,
University of Maribor, Slovenia
Eike Buhr,
University of Oldenburg, Germany

*CORRESPONDENCE
Donrich Thaldar,
ThaldarD@ukzn.ac.za

RECEIVED 19 September 2023
ACCEPTED 27 November 2023
PUBLISHED 14 December 2023

CITATION
Bottomley D and Thaldar D (2023), Liability for harm caused by AI in healthcare: an overview of the core legal concepts.
Front. Pharmacol. 14:1297353.
doi: 10.3389/fphar.2023.1297353

Liability for harm caused by AI in healthcare: an overview of the core legal concepts

Dane Bottomley and Donrich Thaldar*

School of Law, University of KwaZulu-Natal, Durban, South Africa

The integration of artificial intelligence (AI) into healthcare in Africa presents transformative opportunities but also raises profound legal challenges, especially concerning liability. As AI becomes more autonomous, determining who or what is responsible when things go wrong becomes ambiguous. This article aims to review the legal concepts relevant to the issue of liability for harm caused by AI in healthcare. While some suggest attributing legal personhood to AI as a potential solution, the feasibility of this remains controversial. The principal–agent relationship, where the physician is held responsible for AI decisions, risks reducing the adoption of AI tools due to potential liabilities. Similarly, using product law to establish liability is problematic because of the dynamic learning nature of AI, which deviates from static products. This fluidity complicates traditional definitions of product defects and, by extension, where responsibility lies. Exploring alternatives, risk-based determinations of liability, which focus on potential hazards rather than on specific fault assignments, emerges as a potential pathway. However, these, too, present challenges in assigning accountability. Strict liability has been proposed as another avenue. It can simplify the compensation process for victims by focusing on the harm rather than on the fault. Yet, concerns arise over the economic impact on stakeholders, the potential for unjust reputational damage, and the feasibility of a global application. Instead of approaches based on liability, reconciliation holds much promise to facilitate regulatory sandboxes. In conclusion, while the integration of AI systems into healthcare holds vast potential, it necessitates a re-evaluation of our legal frameworks. The central challenge is how to adapt traditional concepts of liability to the novel and unpredictable nature of AI—or to move away from liability towards reconciliation. Future discussions and research must navigate these complex waters and seek solutions that ensure both progress and protection.

KEYWORDS

artificial intelligence, liability, Africa, healthcare, harm

1 Introduction

Modern artificial intelligence (AI) is the cornerstone of the fourth industrial revolution. Successes in data availability, algorithm design, and processing power (Craglia et al., 2018) have empowered AI systems to make dramatic impacts in disparate sectors including transportation, education, agriculture, public services, finance, and healthcare (Artificial Intelligence for Africa: An Opportunity for Growth, Development, and Democratisation, 2018).

The varying degrees of autonomy with which AI systems can operate distinguish it from other emerging technologies. The advantage of AI lies in its ability to process massive amounts of varied information, and thereby perform valuable functions or draw useful

conclusions inspired by its interpretation of the information. However, the essence of its usefulness is also its most challenging feature. For example, machine learning is a common approach to AI system design in medicine. Instead of programming the system for all possible scenarios with specific instructions, when using machine learning, developers set a broad goal which the system uses to form its own instructions to achieve the goal through repeated experiments and self-research (Rachum-Twaig, 2020). As it processes information, the AI system adjusts the parameters by which it judges inputs to produce more accurate outputs, effectively programming itself (Townsend, 2020). These approaches usually produce more accurate systems and they also require less human control (Grimm et al., 2021). Alarmingly, this and similar approaches to AI system design remove the human element at key stages of development in a way which may complicate inquiries into the attribution of responsibility and liability. This becomes especially pronounced where the AI system is so complex that its operations become inscrutable to humans. These so-called "black-box" algorithms lack the transparency to fully audit how they came to the conclusions they did. In response to this issue, some developers have endeavoured to design and create 'explainable' AI systems and ways of ensuring transparency which would foster an environment of accountability and responsibility and create better evidence when determining liability (Ali et al., 2023).

Determining responsibility will be important in dealing with the social challenges of AI integration. Perc et al. (2019) investigates how AI systems will likely have to choose between acting in favour of one party's interest over another in certain contexts and how this may influence how the technology evolves. Developers may be incentivised to produce systems which favour owners' interests above users' in order to drive sales. The solution may be to require that AI systems act in the interests of the broader community; however, this policy may create its own issues in that it will potentially disincentivise people from buying AI systems which will not protect their interests outright and therefore lead to a lower adoption and investment in AI systems overall. This approach may then fail to fully realise the safety gains which can be had by increased AI usage. Of course, as Perc et al. (2019) consider, another approach may be to leave such decisions for the AI system to decide itself, or simply leave it to chance. This approach, however, suffers from a lack of clear answers to questions of responsibility and liability for the outcomes of decisions. Robust regulation and thoughtful juristic approaches to AI challenges will be necessary to provide adequate responses to responsibility for actions in these cases. This will be vital to supporting the benefits of AI integration whilst properly addressing the risks of the technology. Specifically, in healthcare, AI systems show impressive potential to increase the overall efficiency of healthcare systems and to manage disease outbreaks (Owoyemi et al., 2020). Furthermore, these systems can increase the reach of initiatives, while supplementing an already overburdened sector (Pepper and Slabbert, 2011). However, healthcare institutions deal with patients who are at their most vulnerable, where an incorrect decision could prove fatal. In addition, healthcare practitioners are required to abide by particularly high ethical and legal standards which AI systems may not easily conform to. In particular, the black box nature of some algorithms may prevent physicians from providing enough information to their patients about their treatments to satisfy requirements of informed consent, the emergent abilities of AI systems also raise questions as to how they will be considered in relation to the usual standard of care expected of physicians, and medical liability may need to be redefined for AI use.

Many jurisdictions already have laws and regulations which would encompass AI technologies; however, the specific challenges of AI may mean that these regulations do not provide desirable results when they are relied upon. As a response to this, many jurisdictions outside Africa have begun drafting specific AI law and regulations (Sallstrom et al., 2019). Providing a proper response to the issues posed by AI use in healthcare is essential to providing legal certainty to all stakeholders. This will allow them to order their interactions with AI systems and create an environment of trust in relation to AI use (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019). This trust will be important for the future of AI as a lack of trust could permanently harm the reputation of AI in healthcare, or lead to additional costs through inefficient regulation or repeated amendment (Floridi et al., 2018).

The aim of this article is to set the stage for legal development and policy initiatives in Africa by exploring the legal concepts relevant to the attribution of liability for AI harm. First, we begin by describing current developments and the use of AI in healthcare in Africa in Section 2. Then we discuss the concept of liability broadly in Section 3. In Section 4, we describe how AI presents novel challenges to liability determination, particularly the concept of personal liability. In Section 5, we review the different approaches to determining liability. We provide our concluding thoughts in Section 6.

2 Artificial intelligence in healthcare in Africa

AI systems in healthcare can perform tasks normally requiring human physicians (Joshi. and Morley, 2019). Most current uses are in diagnosis and screening; however, future systems could scan images, discover new drugs, optimise care pathways, predict positive treatment outcomes, and provide preventative advice (Joshi and Morley, 2019). Increased use of AI allows physicians to focus on tasks where, given the current state of technology, they cannot be replaced. Furthermore, AI could further broaden public health initiatives by increasing access and tracking disease outbreaks, while lowering the cost of care (Joshi and Morley, 2019).

For example, DeepMind's AlphaFold is an AI system which accurately predicted the protein structures of the COVID-19 virus, being an important aspect of creating a vaccine (Jumper et al., 2020). This use could greatly reduce vaccine response times in the future. IBM's Watson for oncology is another system which has been able to analyse genomic data of patients in light of medical data from vastly more journals than a person could process, so providing more personalised treatments with high accuracy rates (Chung and Zink, 2018).

While other jurisdictions are considering policy-level AI implementation in healthcare systems (Joshi,and Morley, 2019), Africa has had relatively little meaningful interaction with AI in healthcare both academically (Tran et al., 2019) and clinically (Owoyemi et al., 2020) and, currently, African countries are at a

nascent stage in their AI regulatory policies (Townsend et al., 2023). This is despite AI's utility in developing countries where AI systems could lead to better utilisation of resources and enable new, effective treatments and treatment management systems (Sallstrom et al., 2019). Furthermore, AI systems can provide overarching and effective treatment options that improve standards of living, improve direct patient care, maximise supply-chain efficiencies, reduce administrative tasks, and streamline and improve compliance measures (Sallstrom et al., 2019).

Even though relatively limited, there has been some AI system use in Africa. In South Africa, Vantage, a machine learning-based system developed by BroadReach Healthcare, was used to assess clinics' performance and provide staffing and operational recommendations in HIV clinics in KwaZulu-Natal (Singh, 2020). Further, DrConnect, an application by Discovery Health, provides personal assessments of medical symptoms and advice and remote support using AI technology, by using information from wearable devices such as smartwatches, to give medical and lifestyle advice (Singh, 2020). In Ghana, MinoHealth AI labs have used AI systems for automated diagnostics, forecasts, and prognostics. Also, BareApp is using specialised AI technology to diagnose skin disease and suggest treatments (Eke et al., 2023). In Uganda, AI is being merged with other technologies to develop a specialised system in the management of female chronic diseases (Eke et al., 2023). In Nigeria, Ubenwa is using AI to improve the diagnosis of birth asphyxia in low-resource settings (Owoyemi et al., 2020). Also in Nigeria, AI is proving effective in the identification of fake drugs (Owoyemi et al., 2020).

These examples illustrate the growing use and development of AI systems in Africa. However, as this use grows, it will be vital that African countries position themselves to take full advantage of AI's benefits. Legal regulation will be especially important in directing AI system use and development by providing legal certainty by the formation of proper policies and regulations. A main concern though will be the determination of liability for AI harm.

3 Understanding liability

The nature of emerging technologies is that we need time to understand them and develop policies and regulations which will encourage equitable use (Calo, 2015). AI in healthcare is no different. While AI has the potential to positively influence healthcare, its implementation must necessarily be coupled with appropriate safeguards to minimise risks of harm (European Commission, Directorate-General for Justice and Consumers, 2019). Specific to AI, unforeseeable risks may still arise in apparently well-trained systems where performance is being improved (World Health Organisation, 2021). As it currently stands, when risks arise, our existing policies and regulations will be the basis of determining who is responsible and liable for the harm caused. Assessing whether these policies and regulations are sufficient to properly determine responsibility will be important, as the determination of responsibility plays an important role in determining the basis of legal liability for AI conduct and garnering trust in AI usage more broadly. Currently, this will largely depend on civil liability rules.

Generally, civil liability provides the dual purpose of providing a means for victims of harm to be compensated, while also providing an economic incentive for those held liable to avoid continuing harmful conduct (Buiten et al., 2021). Accordingly, these rules are an important means of protecting patients and providing clarity to businesses on how they may innovate and operate their products (Buiten et al., 2021). However, the varying complexity of AI systems, system updates, algorithms which change from environmental input, and cyber-security concerns may make it difficult to justify claims for compensation and to provide clear pathways for victims to bring claims (European Commission, Directorate-General for Justice and Consumers, 2019). It is also unclear whether the rationale behind current liability regimes will be effective in dealing with AI harm. For example, where AI systems make decisions, it may be difficult for a plaintiff to find a suitable defendant or for a court to determine the standard of care to be expected from an AI system. Therefore, it is currently unclear how current liability regimes will consider AI harm in healthcare.

Proper liability policy formation will consider the outcomes of current liability rules but, in addition, it must necessarily consider the impact which the policy will have on the development and use of AI in the future. This means tailoring policy towards managing AI-specific risks while encouraging positive uses. For example, a lack of legal certainty and fear of unreasonable legal penalties for relying on AI recommendations may discourage healthcare practitioners from using AI systems as active participants in treatment, relegating AI systems' role to the mere confirmation of decisions made by healthcare practitioners (World Health Organisation, 2021). On the contrary, removing penalties may encourage AI systems use; however, this position may be tenable only where existing issues of accountability and responsibility are properly considered.

Of particular concern in healthcare should be determining how an AI system will form part of the standard of care. Such a determination will be essential for providing sufficient information for physicians and patients to make decisions about relying on the technology (World Health Organisation, 2021). The decision of the physician is important as he/she will also likely be responsible for the proper operation, monitoring and maintenance of the technology (Bertolini and Episcopo, 2021), and their decision could be consequential for their employer through vicarious liability (World Health Organisation, 2021).

A concern specific to Africa is that many policy frameworks which would guide the development of AI systems are created in environments outside of Africa. Moreover, a lack of access to high quality data sets and limitations in infrastructure could lead to the use of algorithms which are predominantly developed outside of Africa. These could be potentially prejudicial as they may not be properly designed to work in low-resource environments (World Health Organisation, 2021). Therefore, liability policies will need to consider that developers are situated outside of Africa, and that algorithms are adapted for, rather than designed for, the African context.

The role of an AI policy framework should be to prevent AI harm and to promote AI innovation, following a risk-based, rights-preserving, agile, adaptive, and innovation-supporting regulatory approach (Townsend et al., 2023). Robust and effective regulation will provide important guiding principles for the development and implementation of AI systems in healthcare in Africa (World Health Organisation, 2021). Legal certainty will provide routes for compensation for patients

and ensure accountability and responsibility through integration and innovation in the healthcare system.

4 Challenging our understanding of liability: AI and personhood

AI systems' successful imitation of qualities normally associated with humans has bolstered the inquiry into AI personhood (Abbott and Sarch, 2019). A crucial development in support of AI personhood has been the ability to program generalised goals into AI systems. This approach is markedly different to traditional software as the AI system is programmed to decide what steps it would take to achieve its goal, instead of being programmed with specific, step-by-step instructions (Bostrom and Yudkowsky, 2014). This goal-directed behaviour is what powered IBM's Deep Blue chess robot. Programmers surpassed their own chess skills by encoding the rules of the game into Deep Blue and relying on its superior processing power to find ways of "winning" which the developer would not be able to do (Bostrom and Yudkowsky, 2014). Should this be enough to draw the necessary philosophical conclusions on AI personhood, it is clear that the legal implications would be substantial (Solum, 1992). Where AI systems are considered persons, even in limited form, they may be held responsible for their actions in their own capacity.

However, the utility of recognising AI personhood should not replace thoughtful policy formation. An AI system fulfilling roles normally delegated to humans does not mean that personhood necessarily follows (Thaldar and Naidoo, 2021). This may be illustrated by the recent granting of a patent in South Africa where the sole inventor was AI. Although some would consider "inventing" to be a human characteristic, without the ability to fully contain human emotion and capacity to engage in relationships, it is difficult to see such an AI system as more than a "special species of legal object that has the ability to invent" (Thaldar and Naidoo, 2021). As AI becomes more autonomous, legal rules can be developed to allow for special treatment of AI systems, which would be comparable to the legal rules that provide for the special treatment of animals (Thaldar and Naidoo, 2021).

While it is generally agreed that current AI systems are not capable of being considered legal persons, more sophisticated, generalised, and autonomous systems may change this assumption (Solum, 1992). Current systems can be changed, created, or completely deleted like any other software, but where AI systems enjoy a degree of personhood, our relationship with them may become far more complicated. Legally, the granting of AI personhood would aid plaintiffs of AI harm in that they could gather evidence from the AI system through its examination as a witness (Chung and Zink, 2018). However, this benefit may be somewhat limited in systems that lack transparent reasoning.

More definitively, some scholars insist that a separate legal personality for AI systems will never be necessary (European Commission, Directorate-General for Justice and Consumers, 2019). They contend that even fully autonomous systems' actions are better attributed to individuals or other legal persons than to the system itself (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019).

An important consideration is that AI systems' lack of abstract thought limits their comparison to human personhood and decision-making, particularly in healthcare. Whereas human decision-making in healthcare is largely justified by morality, AI systems lack moral input in decision-making (Chung and Zink, 2018). Moral considerations become vitally important in healthcare and resource-scarce environments where circumstances require difficult decisions to be justified, usually with reference to moral ideals. Therefore, we suggest that in lacking moral capacity, AI systems could be limited in how they could be held accountable if they were considered persons or could lack prerequisites to make decisions in moral contexts.

For scholars who consider AI more than a tool, the lack of moral input is an issue they contend with (Bashayreh et al., 2021). Dignum (2017) suggests that even AI systems acting as assistants may inherit a moral framework for decision-making through incorporating the values of their engineers. However, a mere copy of an engineer's morals may not necessarily lead to satisfactory results as AI systems may not apply moral lessons to their environments in the same way as humans (Bostrom and Yudkowsky, 2014). Dignum contends that identifying and analysing these imbued values will nevertheless improve system performance (Dignum, 2017). This would also ensure that incorporated morals are interpreted in an acceptable way, meaning that, as these systems become more autonomous and powerful, moral assessment may become an essential component of their decision-making, especially in a field such as healthcare (Dignum, 2017).

Accordingly, there is some possibility of future AI systems bearing some form of personhood (Solum, 1992). However, conferring even a limited form of personhood on AI systems presents further practical difficulties. For example, as is commonly suggested, a limited form of personhood may be imbued on AI systems through the extension of the principal-agent relationship. In determining responsibility, however, the standards which apply when adjudicating AI system conduct and under what circumstances AI systems would be considered liable for their conduct would remain unclear. This will be discussed further in the section on the principal-agent relationship below.

Further, a final practical issue of attributing liability directly to AI systems is that it leaves no clear pathways for compensation of victims (Bashayreh et al., 2021). As AI systems are currently incapable of ownership, there are no assets that a victim could claim. To remedy this situation, some scholars have suggested the introduction of an insurance scheme funded by developers from which victims may claim (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019). However, such a scheme may not adequately replace clear and fair liability rules and could lead to high administrative costs, so defeating the cost-saving benefits of a clear claim process (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019). Furthermore, there is a lack of guidance on the value of AI insurance policies as there are no standards against which to assess risk or begin a cost analysis (Bertolini and Episcopo, 2021).

5 Approaches to attributing liability

The subsections below discuss the main approaches to the attribution of liability for harm caused by AI systems in healthcare. Section 5.1 broadly considers the extension of the

principal–agent relationship to include AI systems and the consequences of such an extension. Section 5.2 deals with AI as a product and how consumer protection law standards may be applied to AI system harm. We then comment on current fault-based liability regimes as they apply to AI systems in Section 5.3. This leads to a discussion of efforts to use strict liability to attribute liability for AI harm in Section 5.4. In Section 5.5, we consider an approach to AI harm focusing on improving AI system use in healthcare through reconciliatory forums.

5.1 Principal–agent relationship

Most current AI systems in healthcare act as assistants to healthcare practitioners (Joshi and Morley, 2019). Accordingly, some scholars have suggested extending principal–agent rules to govern liability (Rachum-Twaig, 2020). This approach is mostly modelled after the doctor–medical student relationship whereby a medical student performs tasks under the authority and supervision of a doctor; however, the doctor attracts liability for harm which occurs during the student's duties (Chung and Zink, 2018). IBM's Watson operated under a similar regime, whereby the system would assist physicians in making decisions and providing recommendations; however, the physician carried responsibility for the final decision (Chung and Zink, 2018). This approach would ensure that there is always an identifiable human part of the decision-making process and would be in line with an AI design philosophy called "human-in-the-loop" systems (HITL) (Dignum, 2017). HITL ensures proper oversight of system decisions, while creating a clear party to hold accountable by making a human ultimately responsible for decisions (Dignum, 2017).

Although this approach provides a justification for attributing liability to a specific person, it may disincentivise practitioners from following system recommendations as they would bear the risk of harm. The tension arises where the physician may not be able to understand how the system came to its decisions and therefore be unable to assess the risk of harm himself or herself. He or she will likely, however, justify considering AI recommendations based on AI's profound ability to consider vastly more information than he or she could. This could potentially lead to increased costs of medical care and slower treatments as practitioners seek alternative means of validating their decision to follow or reject AI system recommendations. This may be so until there is guidance as to AI systems' position in the standard of care. Should AI systems form part of the standard of care, there may be an expectation for physicians to follow AI recommendations, potentially only until they have a clear professional duty to act otherwise.

Furthermore, similar to criticism of AI personhood, critics of HITL argue that there is a difficulty in determining the correct standard against which to compare the conduct of the AI system (Kingston, 2016). Initial systems may be comparable to humans; however, as systems begin to outperform humans, another standard may need to be considered (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019). In addition, as systems become more sophisticated, there remains uncertainty as to how disagreements between AI system recommendations and human practitioner recommendations should be resolved. Current norms suggest that claims for damages will favour standard care pathways, even where AI systems recommend non-standard treatments (Tobia et al., 2021). This seems to be true, regardless of the outcome of treatment and healthcare practitioners are more likely to attract liability where they do not follow these standards (Price et al., 2019). This initial bias against non-standard care could limit growth of AI technology use in healthcare, which could limit future AI development as there will be a lack of testing in a medical environment and a lack of opportunity to build trust (World Health Organisation, 2021).

Importantly, healthcare practitioners could be less willing to implement recommendations for AI systems which deviate from standard care procedures where they face liability for acting on AI recommendations. However, as AI systems become commoner in healthcare, the bias against their inclusion could shift, especially where AI systems become part of the standard of care (World Health Organisation, 2021). The attribution of liability to the developer of the system may follow if they are in the best position to prevent harmful outcomes as the creator of the system (Lövtrup, 2020).

5.2 Product liability

Townsend et al. (2023) found that eleven out of twelve African countries surveyed provide for strict liability of harmful or defective goods in their consumer protection laws. Therefore, anyone in the supply chain could in principle be held strictly liable for AI harm to the patient. However, are these consumer protection laws sufficiently equipped to deal with AI-specific risks? Core to consumer protection law is the concept of a product *defect*. To attenuate strict liability, it must be proven that a product had a defect. However, the inherent unpredictability of AI systems makes it difficult to define what constitutes a defect in the context of AI (Bashayreh et al., 2021). The South African Supreme Court of Appeal held that a consumer who is claiming in terms of South Africa's Consumer Protection Act (South African Government, 2009) must prove not only the existence of a defect, but also that the defect is material (Motus Corporation, 2021). Furthermore, it is difficult to prove that a defect caused harm (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019), or that the developer was responsible for the defect (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019). When using multiple systems together, as is common in healthcare, attributing fault may be impossible (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019). Modern regulations were drafted before the AI boom, and therefore are unlikely to have properly considered AI-specific issues (Lövtrup, 2020). Accordingly, patients who have suffered harm caused by AI are likely to face a considerable evidentiary burden when seeking resolution through product liability law.

In the United States, software has generally been considered a tool and courts have been hesitant to extend product liability to healthcare software developers (Gerke et al., 2020). In Europe, the "developmental risk defence" allows a producer to avoid liability on the basis that scientific knowledge at the time of production was unable to detect the existence of a defect in the product (Holm et al.,

2021). Sihlahla et al. (2023) note that in South Africa, a healthcare practitioner or a healthcare establishment sued in terms of the Consumer Protection Act (South African Government, 2009) for harm caused by AI would have a complete defence if they can show that they could not reasonably have been expected to have discovered the defect.

5.3 Fault-based remedies

Generally, fault-based liability is based on a person's intentional or negligent conduct which causes harm wrongfully and culpably (Mukheibir et al., 2010). Liability is attributed based on a determination of who should justly compensate for the damages of the plaintiff (Marchisio, 2021). Currently, there is no case law to guide the application of fault-based liability principles, particularly in cases where the AI suffers from an unknown flaw which was not reasonably foreseeable (Donnelly, 2022).

Accordingly, key elements of such remedies are difficult to prove in AI system cases, specifically causation and fault. Causation is difficult to prove as it may be difficult to show a flawed algorithm was the cause of harm (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019). Similar to product law, it may be difficult to determine what a flaw is, or at what point the flaw was created if the system was developed by multiple parties (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019). Even where a flaw is identified, demonstrating foreseeability for negligence-based claims is still difficult (Holm et al., 2021). Furthermore, establishing vicarious liability would be complicated as, currently, there is no means of determining whether the AI system "acted negligently" or what degree of control a medical practitioner should exert over an AI system (Donnelly, 2022). Accordingly, where there is no causation on the part of the physician, a patient may be left with no recourse (Donnelly, 2022).

Fault-based liability is an important means of deterrence (Buiten et al., 2021). Defendants who are penalised are incentivised to prevent harm in the future (Marchisio, 2021). This is justified as the defendant should be the one best oriented to assess and avoid risk (Marchisio, 2021). However, AI systems' necessary unpredictability may make it impossible for a particular party to act to prevent harm as it would be unforeseeable.

Therefore, it has been suggested that liability, by rule, be shared among the technical and medical stakeholders as part of their joint contribution to the risk of harm in the use of the system (Smith and Fotheringham, 2020). This could be in the form of joint and several liability or proportional liability (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019) using the person's choice to develop or implement the system as the justification to establish causation (Bashayreh et al., 2021).

An extension of this idea is a risk-sharing approach (Bashayreh et al., 2021). Owners and developers would bear liability proportionate to the risk each has accepted in their role in the AI lifecycle, operating to the exclusion of cases of wilful misconduct or gross negligence (Bashayreh et al., 2021). Importantly, developers would need to disclose all risks and potential deficiencies of the system, including the degree to which the system's decision can be explained and all the built-in values of the system (Bashayreh et al., 2021). In addition, owners would disclose their intended use of the product and the environment it will be deployed in (Bashayreh et al., 2021). In the event of harm, liability could be portioned by a court adjudicating on the facts with relevant disclosures.

The creation of responsibilities at different stages of the AI system's lifecycle remains a common approach to justifying liability in fault-based approaches in literature. Current fault-based standards already attach responsibilities to people based on special relationships they may have with an object, such as where the person is in control of a potentially dangerous animal or thing (Marchisio, 2021). Where the animal acts unpredictably, the person controlling it could be held liable (Bashayreh et al., 2021). Failure to fulfil responsibilities to protect others from harm in this type of relationship will justify the attribution of liability. This approach may be useful in AI through the prescription of minimum rules to establish wrongfulness and fault (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019). Where these standards are not upheld, the burden of proof may shift in favour of the victim. Therefore, Rachum-Twaig (2020) suggests the creation of "safe harbours." Safe harbours act as points in the AI lifecycle where a party is responsible for ensuring certain minimum standards. Where the party fails to uphold these standards, they are more likely to incur liability and current fault-based remedies can be employed. Approaches like this form part of a movement towards risk-based liability replacing the foreseeability element of many fault-based regimes (Calo, 2015).

5.4 Strict liability

The clear issues that arise in justifying attribution of liability to certain stakeholders has encouraged some scholars to suggest no-fault or "strict" liability systems as better means of attributing liability (Holm et al., 2021). No-fault liability makes it significantly easier for victims to claim compensation by providing clear pathways to settle claims and removing the necessity of proving fault (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019). This eases the burden on claimants who are already the victim of harm when reporting errors and provides better hope of reconciliation (Holm et al., 2021). No-fault systems also separate the compensation and liability claims (Holm et al., 2021). They remove the necessity of victims to access information to prove fault, which is a particular concern with inscrutable AI systems. The occurrence of harm is made the centre of the claim instead of proving fault.

Concerns raised about this approach have focused on the future development of AI systems (European Commission, Directorate-General of Communications Networks, Content and Technology, 2019). First, strict liability would subject stakeholders to material burdens with no fair opportunity to avoid them (Abbott and Sarch, 2019). Normally, strict liability applies for unexpected harms, but where AI systems are implemented, it is difficult to determine how unexpected harms would be defined, as the systems are necessarily programmed to be unpredictable (European Commission, Directorate-General for Justice and Consumers, 2019). Second,

stakeholders would be at risk of reputational damage resulting from the occurrence of harm which is otherwise not foreseeable (Abbott and Sarch, 2019). Therefore, stakeholders would be subject to significant burdens without an opportunity to take effective measures against the realisation of these harms.

To ease the potential economic impact stakeholders may experience under strict liability, it has been suggested that a stakeholder-funded scheme be created to compensate victims of AI harm (European Commission, Directorate-General for Justice and Consumers, 2019). This may further simplify the pathways for victims to claim; however, a mixed fund would lead to innocent parties effectively being held liable for harm they did not cause (European Commission, Directorate-General for Justice and Consumers, 2019). Furthermore, the burden on blameworthy parties would be eased as they would pay only a portion of any damages claims for harm caused by their systems. This reduction would add to the already perceived loss of the deterrent effect as litigation is no longer available to claimants (European Commission, Directorate-General for Justice and Consumers, 2019). One suggested solution is to model the New Zealand approach to medical matters whereby no-fault systems have been implemented in certain medical matters, but claims are limited to unusual injuries (European Commission, Directorate-General for Justice and Consumers, 2019).

Practically, strict liability could potentially be more expensive than litigation when administrative costs are coupled with more patients being eligible to claim (Holm et al., 2021). Also, a strict liability system may not be capable of being applied cross-jurisdictionally or globally (Rachum-Twaig, 2020). This has led some scholars to suggest that a mixture of fault and no-fault rules could provide equitable AI regulation (Marchisio, 2021).

5.5 Reconciliation

The adversarial nature of the approaches to liability outlined above may be counter-productive to the proper regulation of AI technology—at least during its nascent stage. Naidoo et al. (2022) argue that instead of prioritising questions such as "Who acted?" and "Was the act wrongful?," which causes persons involved to be antagonistic and defensive, the focus should shift to (a) learning how to better use AI in healthcare, and to (b) actively developing guidelines for AI developers and healthcare professionals who are using AI systems. The authors suggest that (a) and (b) can best be attained by establishing a *sui generis* dispute resolution institution for harm caused by AI in healthcare. This institution would replace litigation in the courts, hold broad investigative powers to access all relevant information, resolve disputes through reconciliation, award financial redress to victims of AI-driven harm in healthcare, and—importantly—learn and develop guidelines. In essence the authors argue for reconciliation to replace litigation as they view reconciliation as more conducive to the learning element of a regulatory sandbox.

This approach could draw inspiration from current alternative dispute resolution structures, principally, the South African Commission for Conciliation for Conciliation, Mediation and Arbitration (CCMA). The compensation structure could draw lessons and inspiration from the operation of the South African Road Accident Fund which compensates victims of accidents on public roads for bodily harms. The basis of this system could encompass a more inquisitive approach to litigation, whereby all parties are enabled to share information with the institution taking a more active role in discovery through its investigative powers. A thoughtful use of the institution's powers to adjudicate the matter can help to ensure that power disparities between the parties could be mitigated whilst providing for a just outcome.

The guidelines developed by the *sui generis* dispute resolution institution can over time either become customary law in the field, or be solidified in legislation—depending on the preferences and traditions of the relevant jurisdiction. This would signal that AI technology and the regulation thereof has reached a stage of maturity, at which stage the *sui generis* dispute resolution institution would have served its purpose, and a return to a liability-based approach can be considered.

6 Conclusion

The assimilation of AI technologies in the African healthcare sector is an unprecedented juncture in the continent's journey towards equitable and advanced medical care. As AI solutions make inroads into African medical establishments, they bring along a multitude of autonomy and opacity issues, challenging the longstanding ethical pillars and legal norms ingrained in the diverse cultures of the continent. The quintessential medico-legal principle of informed consent is now juxtaposed against the intricate algorithms of AI, challenging the very essence of transparency and patient understanding. Similarly, the increasing autonomy of AI systems amplifies the intricacies of liability, pushing the boundaries of traditional legal frameworks.

In this article, we tried to provide the reader with an overview of the legal concepts relevant to the issue of AI and liability in healthcare. We started with the contemplation of AI personhood. While captivating, we suggest that it poses substantial challenges in an African context, particularly when addressing tangible redress mechanisms for AI-induced mishaps. Next, the principal-agent framework, although providing a modicum of accountability, could inadvertently stifle the AI adoption rate by placing considerable responsibilities upon local medical practitioners. While product liability law offers another plausible approach, it struggles to categorise the continually evolving nature of AI in the static confines of conventional product definitions. Alternative strategies, such as risk-based liability may offer clearer paths in contexts where fault determination proves onerous. Yet, they too grapple with ensuring specificity and justice. Strict liability, offering more transparent compensation mechanisms, raises concerns about economic implications, reputational risks and, most critically, the challenge of harmonising such policies across Africa's diverse legal landscapes.

An approach based on reconciliation rather than liability potentially provides the best environment for a regulatory sandbox; however, reconciliation in the context of AI-driven harm in the healthcare context lacks the same level of scholarship as the approaches based on liability. We suggest that reconciliation offers much potential and deserves more academic attention.

In distilling these insights, it is evident that Africa's AI journey in healthcare is not solely a scientific or medical transition. It also requires profound legal reflection and evolution.

Author contributions

DB: Conceptualization, Writing–original draft, Writing-review and editing. DT: Conceptualization, Funding acquisition, Supervision, Writing–review and editing.

Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. The first author wishes to acknowledge the support of the National Research Foundation of South Africa (Grant Numbers: 131307). The opinions, findings and conclusions or recommendations expressed in the publication are the author(s) alone, and the NRF accepts no liability whatsoever in this regard. The second author acknowledges the support by the US National Institute of Mental Health and the US National Institutes of Health (award number U01MH127690). The content of this article is solely the authors responsibility and does not necessarily represent the official views of the US National Institute of Mental Health or the US National Institutes of Health.

Acknowledgments

The authors acknowledge the use of ChatGPT4 from Open AI to improve the language and readability of the abstract, introduction and conclusion sections of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Abbott, R., and Sarch, A. (2019). Punishing artificial intelligence: legal fiction or science fiction. *UC Davis Law Rev.* 53, 323–384. doi:10.2139/SSRN.3327485

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., et al. (2023). Explainable artificial intelligence (XAI): what we know and what is left to attain trustworthy artificial intelligence. *Inf. Fusion* 99, 101805. doi:10.1016/j.inffus.2023.101805

Artificial Intelligence for Africa: An Opportunity for Growth, Development, and Democratisation (2018). *Access partnership*. Hatfield Campus, Pretoria: University of Pretoria. Available at: https://www.up.ac.za/media/shared/7/ZP_Files/ai-for-africa.zp165664.pdf.

Bashayreh, M., Sibai, F. N., and Tabbara, A. (2021). Artificial intelligence and legal liability: towards an international approach of proportional liability based on risk sharing. *Inf. Commun. Technol. Law* 30 (2), 169–192. doi:10.1080/13600834.2020.1856025

Bertolini, A., and Episcopo, F. (2021). The expert group's report on liability for artificial intelligence and other emerging digital technologies: a critical assessment. *Eur. J. Risk Regul.* 12, 644–659. doi:10.1017/err.2021.30

Bostrom, N., and Yudkowsky, E. (2014). "The ethics of artificial intelligence," in *Cambridge handbook of artificial intelligence*. Editors W. Ramsey and K. Frankish (Cambridge, UK: Cambridge University Press), 316–334. doi:10.1017/CBO9781139046855.020

Buiten, M., de Streel, A., and Peitz, M. (2021). EU liability rules for the age of artificial intelligence. *SSRN Electron. J.* doi:10.2139/ssrn.3817520

Calo, R. (2015). Robotics and the new cyberlaw. *Calif. L. Rev.* 103, 513–563. doi:10.2139/ssrn.2402972

Chung, J., and Zink, A. (2018). Hey Watson – can I sue you for malpractice? Examining the liability of artificial intelligence in medicine. *Asia Pac. J. Health L. Ethics* 11, 30.

Craglia, M., Annoni, A., Benczúr, P., Bertoldi, P., Delipetrev, B. T., De Prato, G., et al. (2018). *Artificial intelligence a European perspective*. Luxembourg: Publications Office of the European Union.

Dignum, V. (2017). Responsible autonomy. Available at: http://arxiv.org/abs/1706.02513 (Accessed January 19, 2021).

Donnelly, D.-L. (2022). First do No harm: legal principles regulating the future of artificial intelligence in health care in South Africa. *Potchefstroom Electron. Law J.* 25 (1), 1–43. doi:10.17159/1727-3781/2022/v25i0a11118

Eke, D. O., Chintu, S. S., and Wakunuma, K. (2023). "Towards shaping the future of responsible AI in africa," in *Responsible AI in Africa social and cultural studies of robots and AI*. Editors D. O. Eke, K. Wakunuma, and S. Akintoye (Cham: Springer International Publishing), 169–193. doi:10.1007/978-3-031-08215-3_8

European Commission, Directorate-General for Justice and Consumers (2019). *Liability for Artificial Intelligence and other emerging digital technologies*. Luxembourg: Publications Office. Available at:. doi:10.2838/573689

European Commission, Directorate-General of Communications Networks, Content and Technology (2019). *Ethics guidelines for trustworthy AI*. Luxembourg: Publications Office. Available at: https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., et al. (2018). AI4People – an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. *Minds Mach. (Dordr)* 28, 689–707. doi:10.1007/s11023-018-9482-5

Gerke, S., Minssen, T., and Cohen, G. (2020). "Ethical and legal challenges of artificial intelligence-driven healthcare," in Artificial Intelligence in healthcare *(elsevier)*. Editors A. Bohr and K. Memarzadeh (Cambridge, MA: Academic Press), 295–336. doi:10.1016/B978-0-12-818438-7.00012-5

Grimm, P. W., Grossman, M. R., and Cormack, G. V. (2021). Artificial intelligence as evidence. *Northwest. J. Technol. Intellect. Prop.* 19.

Holm, S., Stanton, C., and Bartlett, B. (2021). A new argument for No-fault compensation in health care: the introduction of artificial intelligence systems. *Health Care Anal.* 29, 171–188. doi:10.1007/s10728-021-00430-4

Joshi, I., and Morley, J. (2019). *Artificial Intelligence: how to get it right. Putting policy into practice for safe data-driven innovation in health and care*. London, United Kingdom: NHSX.

Jumper, J., Tunyasuvunakool, K., Kohli, P., Hassabis, D., and the AlphaFold Team (2020). Computational predictions of protein structures associated with COVID-19. *DeepMind*. Available at: https://www.deepmind.com/open-source/computational-predictions-of-protein-structures-associated-with-covid-19 (Accessed September 29, 2021).

Kingston, J. K. C. (2016). "Artificial intelligence and legal liability," in *Research and development in intelligent systems XXXIII*. Editors M. Bramer and M. Petridis (Cham: Springer International Publishing), 269–279. doi:10.1007/978-3-319-47175-4_20

Lövtrup, M. (2020). In brief: artificial Intelligence in healthcare. *Swed. Counc. Med. Ethics* 2. Available at: https://smer.se/wp-content/uploads/2020/06/smer-2020-2-in-brief-artificial-intelligence-in-healthcare.pdf.

Marchisio, E. (2021). In support of "no-fault" civil liability rules for artificial intelligence. *SN Soc. Sci.* 1, 54. doi:10.1007/s43545-020-00043-z

Motus Corporation (2021). *Ltd and another v wentzel (1272/2019) [2021] ZASCA 40 (13 april 2021)*.

Mukheibir, A., Niesing, L., and Perumal, D. (2010). in *The law of delict in South Africa*. Editors M. M. Loubser and R. Midgley (Cape Town, South Africa: Oxford University Press Southern Africa).

Naidoo, S., Bottomley, D., Naidoo, M., Donnelly, D., and Thaldar, D. W. (2022). Artificial intelligence in healthcare: proposals for policy development in South Africa. *South Afr. J. Bioeth. Law* 15 (1), 11–16. doi:10.7196/SAJBL.2022.v15i1.797

Owoyemi, A., Owoyemi, J., Osiyemi, A., and Boyd, A. (2020). Artificial Intelligence for healthcare in africa. *Front. Digit. Health* 6 (2), 6–5. doi:10.3389/fdgth.2020.00006

Pepper, M. S., and Slabbert, M. N. (2011). Is South Africa on the verge of a medical malpractice litigation storm? *S. Afr. J. Bioeth. Law* 4, 29–35.

Perc, M., Ozer, M., and Hojnik, J. (2019). Social and juristic challenges of artificial intelligence. *Palgrave Commun.* 5, 61. doi:10.1057/s41599-019-0278-x

Price, W. N., Gerke, S., and Cohen, I. G. (2019). Potential liability for physicians using artificial intelligence. *JAMA* 322 (18), 1765–1766. doi:10.1001/jama.2019.15064

Rachum-Twaig, O. (2020). Whose robot is it anyway? liability for artificial-intelligence-based robots. *Univ. Ill. Law Rev.* 2020, 1141–1176.

Sallstrom, L., Morris, O., and Mehta, H. (2019). Artificial intelligence in Africa's healthcare: ethical considerations. *Observer Res. Found. Issue Brief* 312, 12.

Sihlahla, I., Donnelly, D.-L., Townsend, B., and Thaldar, D. (2023). Legal and ethical principles governing the use of artificial intelligence in radiology services in South Africa. *Dev. World Bioeth.*, 1–11. doi:10.1111/dewb.12436

Singh, V. (2020). AI and data in South Africa's health sector. *Policy Action Netw.* 6.

Smith, H., and Fotheringham, K. (2020). Artificial intelligence in clinical decision-making: rethinking liability. *Med. Law Int.* 20, 131–154. doi:10.1177/0968533220945766

Solum, L. B. (1992). Legal personhood for artificial intelligence. *North Carol. Law Rev.* 70, 1231.

South African Government (2009). Consumer protection act. Available at: https://www.gov.za/sites/default/files/32186_467.pdf.

Thaldar, D., and Naidoo, M. (2021). AI inventorship: the right decision? *S. Afr. J. Sci.* 117. doi:10.17159/sajs.2021/12509

Tobia, K., Nielsen, A., and Stremitzer, A. (2021). When does physician use of AI increase liability? *J. Nucl. Med.* 62, 17–21. doi:10.2967/jnumed.120.256032

Townsend, B. A. (2020). Software as a medical device: critical rights issues regarding artificial intelligence software-based health technologies in South Africa. *J. South Afr. Law/Tydskrif vir die Suid-Afrikaanse Reg* (4), 747–762.

Townsend, B. A., Sihlahla, I., Naidoo, M., Naidoo, S., Donnelly, D.-L., and Thaldar, D. W. (2023). Mapping the regulatory landscape of AI in healthcare in Africa. *Front. Pharmacol.* 14, 1214422. doi:10.3389/fphar.2023.1214422

Tran, B., Vu, G., Ha, G., Vuong, Q.-H., Ho, M.-T., Vuong, T.-T., et al. (2019). Global evolution of research in artificial intelligence in health and medicine: a bibliometric study. *J. Clin. Med.* 8, 360. doi:10.3390/jcm8030360

World Health Organisation (2021). *Ethics and governance of artificial intelligence for health: WHO guidance*. Geneva, Switzerland: World Health Organisation. Available at: https://www.who.int/publications/i/item/9789240029200.

frontiers | Frontiers in Pharmacology

TYPE Review
PUBLISHED 01 March 2024
DOI 10.3389/fphar.2024.1304950

OPEN ACCESS

EDITED BY
Raphael Zozimus Sangeda,
Muhimbili University of Health and Allied Sciences, Tanzania

REVIEWED BY
Lonni Besançon,
Monash University, Australia
Rogena Sterling,
University of Waikato, New Zealand

*CORRESPONDENCE
Amy Gooden,
215031300@stu.ukzn.ac.za

RECEIVED 17 November 2023
ACCEPTED 12 February 2024
PUBLISHED 01 March 2024

CITATION
Gooden A (2024), A pathway to strengthening open science: comments on the draft *South African Ethics in Health Research Guidelines*. *Front. Pharmacol.* 15:1304950.
doi: 10.3389/fphar.2024.1304950

A pathway to strengthening open science: comments on the draft *South African Ethics in Health Research Guidelines*

Amy Gooden*

School of Law, University of KwaZulu-Natal, Durban, South Africa

The recently released *draft South African Ethics in Health Research Guidelines: Principles, Processes and Structures* (Draft Guidelines) by the National Health Research Ethics Council recognize open data and provide guiding principles for this in the context of health research in South Africa. While its inclusion is a positive development, there is room for improvement. Although the Draft Guidelines leverage the *Draft National Policy on Data and Cloud*, it lacks incorporation of other relevant government policies, notably the *Draft National Open Science Policy*, and fails to sufficiently detail the principles of open science and open access. This limited scope and lack of comprehensive definition and detailed guidance present challenges for researchers in conducting ethical and responsible health research in South Africa. It constrains the Draft Guidelines from fully aligning with national imperatives and from fostering African-centric approaches. To address these issues, it is recommended that the Draft Guidelines integrate broader policies and principles, enhance clarity through comprehensive definitions, provide detailed guidance on open access, and promote African-centric approaches. Implementing these solutions will strengthen the Draft Guidelines, aligning them with national visions of open science, and thereby harnessing the full potential of South Africa's diverse scientific community in advancing health research.

KEYWORDS

data, genomics, health research, open access, open science, policy, South Africa

1 Introduction

In recent years, there has been a proliferation of health research worldwide. Health research contributes to the understanding of disease, the improvement of healthcare systems, the development of new medicines and treatments, and technologies aimed at bettering health and healthcare (DoH, 2015). As such, health research has the potential to benefit the population—especially in South Africa where there is a high disease burden, predominantly from HIV/AIDS and tuberculosis (Abdool Karim et al., 2009; Mayosi et al., 2012; Johnson et al., 2017; Abdool Karim and Baxter, 2022; Kubjane et al., 2022).

With the growth of health research in South Africa came the need to address various ethical concerns in health research, align with international standards, protect research participants, and ensure the proper conduct of health research. In 2015, the Department of Health (DoH) released the second edition of the *Ethics in Health Research: Principles, Processes and Structures* (DoH, 2015) (DoH Ethics Guidelines), to replace the previous 2004 edition. The DoH Ethics Guidelines provide guidance to health researchers in South

Africa and cover certain key aspects of health research, such as informed consent, the need for ethical review, community engagement, benefit sharing, risk assessment, the protection of research participants' rights, and the upholding of ethical principles like autonomy and privacy (DoH, 2015). Importantly, the DoH Ethics Guidelines are not simply "soft law"; they are made legally binding by regulation 2(a) of the Regulations Relating to Research with Human Participants (GN R179 of GG 38000, 2014)—therefore, health researchers in South Africa are legally compelled to comply with the DoH Ethics Guidelines.

Health research has further progressed with the advancement of genome sequencing, which led to genomics research and the use of large datasets. The availability of health research data, which could have huge positive impacts on population health, led to calls for datasets, materials, processes, protocols, findings, results, and software to be made more accessible (Spellman et al., 2017; Ramachandran et al., 2021; Chakravorty et al., 2022). Although the idea of open science has existed for many years and was adopted when science research in universities was thriving (Baca, 2006; Rhoten and Powell, 2007; Scaria and Rangarajan, 2016; Krishna, 2020), in recent years open science has come under pressure due to intellectual property law and policy developments, which has caused research to become commercial and proprietary instead of open (Baca, 2006). However, health research (inclusive of genomics research) has driven calls for the promotion of open science given the vast amounts of data generated by genomics research and the need for collaborative efforts in order to analyze it (International Human Genome Sequencing Consortium, 2001), the need for reproducibility and transparency (Begley and Ellis, 2012), the promise of precision medicine (Collins and Varmus, 2015), and the potential for increased discoveries to be made with access to more data (Venter et al., 2001). The data-intensive, collaborative, and translational nature of health and genomics research has led to it being a driving force in advocating for open science (Hetu et al., 2019; Staunton et al., 2021). Not only does open science accelerate research, but it also lessens the wastage of research resources (Buxton et al., 2021), allows the inspection of research outputs (Besançon et al., 2021), enhances transparency, research integrity, and the responsible use of genomic data (Grant et al., 2022; Haven et al., 2022).

The newly released *draft South African Ethics in Health Research Guidelines: Principles, Processes and Structures* by the National Health Research Ethics Council (NHREC) (NHREC, 2023) (Draft Guidelines), which were circulated amongst stakeholders for comment, are an attempt to revise the second edition 2015 DoH Ethics Guidelines and develop a third edition—giving South Africa with an opportunity to provide guidance for open science in health research, something which was not addressed in the 2015 DoH Ethics Guidelines. This article provides a commentary on the Draft Guidelines, focusing on its handling of open science and open access data. In this article, I highlight several problematic aspects of the Draft Guidelines and suggests potential solutions. I begin by introducing open science broadly, and then examining the concept in South Africa specifically. Thereafter, I analyze the Draft Guidelines' addition of guiding principles for open access, identifying where the Draft Guidelines have succeeded in providing clear guidance, as well as areas in which the Draft Guidelines are lacking. Throughout this article, I provide a way forward for the promotion of open science in South Africa, and emphasize areas where the Draft Guidelines can improve in this regard. Given that there have been recent academic pushes for openness, specifically in genomics research in South Africa (Gooden and Thaldar, 2023a; Thaldar et al., 2023a; Gooden and Thaldar, 2023b), it is imperative that this issue be given due consideration.

2 The imperative for open science

Given that advancements in technology have allowed science to become more "open," open science must be viewed as distinct from the previous *status quo* where, for example, publications were only available to subscribers of journals post publication (Friesike et al., 2015). Various definitions have been utilized to describe open science and what it entails. Broadly, open science aims to make research methods and results freely available in order to promote collaboration and transparency to the benefit of the community (Strydom et al., 2022). However, open science should be distinguished from open access. Open access—as a practice of open science—is a set of principles and procedures allowing research outputs to be freely accessible, without any costs or other access barriers (DSI, 2022). Open access allows for published work to be obtained, while open science provides access to the whole scientific knowledge process (Heise and Pearce, 2020).

A common definition of open science, put forward by Maurer (2003), is that it "tends to connote (a) full, frank, and timely publication of results, (b) absence of intellectual property restrictions, and (c) radically increased pre- and post-publication transparency of data, activities, and deliberations within research groups". Vicente-Saez and Martinez-Fuentes (2018) define open science as "transparent and accessible knowledge that is shared and developed through collaborative networks". Open science is seen to comprise of certain central elements, such as (a) open data, (b) open source, (c) open access, (d) open material, (e) open peer-review, and (f) open educational resources (Levy et al., 2010; Krishna, 2020). In many definitions of open science, there are certain common terms that often feature. These include: (a) open, (b) transparent, (c), accessible, (d) shared, (e) collaborative, (f) available, and (g) replicable (Scaria and Rangarajan, 2016; Vicente-Saez and Martinez-Fuentes, 2018). Open science is vital in advancing research, innovation, and society. It emphasizes accessibility, collaboration, and transparency (Nielsen, 2011; Gewin, 2016). Through open science, the sharing of data, methods, and findings makes research more accessible and reproducible, which enhances scientific discovery, democratizes access to knowledge, grows research impact, and increases public trust in science (Fecher and Friesike, 2014; Nosek et al., 2015; McKiernan et al., 2016; Hardwicke et al., 2018).

Central to the implementation of open science is the FAIR Guiding Principles, which are applicable to scientific data management and stewardship (Wilkinson et al., 2016). These principles aim to minimize barriers to research outputs, thereby allowing others to discover, understand, and re-use such outputs—which may lead to further findings and opportunities, as well as take advantage of existing resources (UCL, 2024). FAIR stands for: (1) findability, which aims to make research more easily

discoverable; (2) accessibility, which entails information on how to access the data; (3) interoperability, which allows the data to be integrated with other data; and (4) reusability, which allows for research outputs to be repurposed (Wilkinson et al., 2016; UCL, 2024). In addition to the FAIR Guiding Principles are the CARE Principles for Indigenous Data Governance. The CARE Principles are people centered, and aim to ensure that research is done in such a way so as to benefit indigenous people, and to highlight the how data can further the innovation and self-determination of indigenous people (GIDA, 2019; DSI, 2022). CARE stands for: (1) collective benefit, where data ecosystems should allow indigenous people to derive benefit from the data (GIDA, 2019); (2) authority to control, which recognizes and allows indigenous people to control their data (GIDA, 2019); (3) responsibility, which requires those working with indigenous data to publicize the ways in which the data is used to promote indigenous people's self-determination and collective benefit (GIDA, 2019); and (4) ethics, which ensures that the rights and wellbeing of indigenous people is central in all research endeavors (GIDA, 2019). In South Africa, open science has been defined as "research and development that is collaborative, transparent and reproducible and whose outputs are publicly available" (DSI, 2022). The Department of Science and Innovation (DSI), previously the Department of Science and Technology (DST), in its *White Paper on Science, Technology and Innovation* (STI White Paper), provides that open science "refers to an approach to research based on greater access to public research data enabled by information and communications technology (ICT) tools and platforms, broader collaboration in science–including the participation of non-scientists–and the use of alternative copyright tools for diffusing research results" (DST, 2019). The African Open Science Platform (AOSP) recognizes that open science tends to refer to open data and open access publishing (AOSP, 2023). However, the AOSP notes that this only provides a limited view of what open science actually is. Open science is not limited to scientists, but should be a more public enterprise that includes the public and private sector, business, policymakers, government, communities, and citizens who engage with scientists to explore solutions to issues facing society (AOSP, 2023).

Open science has not only been promoted by the AOSP in various strategies and reports (ASSAf, 2019; AOSP, 2023), but it is also the subject of the *Draft National Open Science Policy*, which was shared by the DSI with stakeholders in 2022. The Draft National Open Science Policy aims to democratize scientific knowledge and thereby strengthen the research landscape by making research outputs accessible, advancing economic development, and promoting research collaboration (DSI, 2022). The Draft National Open Science Policy is guided by various principles, such as findability, accessibility, reusability, transparency, responsibility, flexibility, and sustainability (DSI, 2022). Open science also features in the STI White Paper, where ideas such as inclusivity, innovation culture, and policy coherence are introduced in order to promote science, technology, and innovation while addressing global challenges like the Fourth Industrial Revolution (DST, 2019). Open science is recognized as a means through which the benefits of collaborative, transdisciplinary approaches to knowledge development, as well as the spread of ideas and research, may be realized (DST, 2019).

Given the importance of open science, one would expect it to appear in most government documents. However, in South Africa, a focus on open science has been lacking, and it has not featured in many recent and relevant publications—such as the *Draft National Policy on Data and Cloud* (Department of Communications and Digital Technologies, 2021), the Protection of Personal Information Act 4 of 2013, 2013 (POPIA) Code of Conduct for Research (ASSAf, 2023), and the *Bio-Economy Strategy* (DST, 2013), to name a few. The Draft Guidelines are no exception—any mention of open science and its promotion in health research in South Africa is absent from the Draft Guidelines. This, I suggest, is a missed opportunity and one that should be addressed by the NHREC.

3 Analysis of the Draft Guidelines

The Draft Guidelines are intended to provide minimum standards for undertaking ethical and responsible research in South Africa (NHREC, 2023). They cover different types of health research, guiding principles for ethical research, processes for ethics review, research ethics committees, health research ethics infrastructure, as well as human biological material and data used in research (NHREC, 2023). Unlike the 2015 DoH Ethics Guidelines, the Draft Guidelines provide principles for open access in health research (NHREC, 2023). This is important because it ensures that valuable knowledge—which may be crucial in bettering population health and developing cures and treatments for disease—is freely available (Smith et al., 2017; Day et al., 2020; Strydom et al., 2022). The inclusion of open access in the Draft Guidelines initially appears as a promising step forward considering South Africa's commitment to open science, which has featured in the STI White Paper, and formed a central part of the Draft National Open Science Policy and the STI White Paper. However, despite having the opportunity to further promote open science and open access databases in South Africa, the Draft Guidelines only refer to the Draft National Policy on Data and Cloud—a policy that, although positive in its vision to facilitate free access to data, has been criticized for the means to achieve it, which entails government control of access to data, nationalizing all data generated in South Africa, and interrupting the intellectual property legal framework (Thaldar et al., 2023b). As such, the Draft Guidelines fail to provide a comprehensive and inclusive pathway for open access databases, and thereby open science, in research in South Africa.

In what follows, I analyze various problematic aspects of the Draft Guidelines, specifically in relation to open science—namely, the failure to consider open science, the definition of open data, the importance of comprehensive definitions, the matter of privacy and consent, and the failure to provide proper guiding principles for open access data—and point towards potential solutions, where relevant.

3.1 The failure to consider open science

Open data, which is explicitly referred to in the Draft Guidelines, is regarded as a "sub-set" of open science (ASSAf, 2019). Concepts such as open data, open access, and open source are all considered within the practice of open science (Strydom et al., 2022). Therefore,

open data—which is mentioned in the context of research—should not be discussed without considering the broader framework of open science. This is something that has been recognized and promoted by the Draft National Open Science Policy, but which the NHREC appears to overlook. However, the Draft Guidelines fail to address open science, and thereby negate a vital aspect of research in South Africa.

In recent years, there has been a push for open science in South Africa and the concept has featured in two government documents: The 2019 STI White Paper and the 2022 Draft National Open Science Policy. However, the Draft Guidelines only focus on one aspect of open science—namely open data—and fail to even mention open science. Therefore, the Draft Guidelines do not promote government policies and strategies intended to further research in South Africa and make it more open and accessible.

However, it should be noted that, without expressly stating so, the Draft Guidelines do appear to point towards open science. The Draft Guidelines recognize that the sharing of data has the potential to *inter alia* enable broad dissemination of research results, increase collaboration, enhance responsiveness to challenges in society, encourage research integrity, and promote greater transparency (NHREC, 2023). In essence, this is open science. Yet, principles that are aligned with open science—such as reproducibility, transparency, and translatability—seem to only apply in the context of animal research and not in terms of research with human participants (NHREC, 2023). Further, international collaboration and the sharing of funding, knowledge, and data—all vital to open science—are only mentioned in the context of public health emergencies, such as the COVID-19 pandemic and not as the norm (NHREC, 2023). It appears as if the Draft Guidelines implicitly recognize open science and its importance, but only in certain contexts such as genomics research, research on animals, and public health emergencies. I suggest that it would be beneficial for the Draft Guidelines to consider explicitly mentioning open science and expanding on its importance in health research, especially given the existence of government policies and strategies that promote it.

3.2 Defining a sub-set of open science: open data

Generally, definitions of open data denote that such data must be freely accessible to be used and re-used by anyone (Scott, 2017; European Commission, 2023; Open Data Charter, 2023; Open Data Handbook, 2023; Open Knowledge Foundation, 2023)—with the only restriction being acknowledgement of the source or share-alike (Open Data Handbook, 2023; Open Knowledge Foundation, 2023).

The Draft Guidelines rely on the definition of "open data" provided in the Draft National Policy on Data and Cloud, which it defines as "data that is made freely available to everyone for use, re-use and republishing as they wish, subject to ensuring protection of privacy, confidentiality and security in line with the Constitution" (Department of Communications and Digital Technologies, 2021). Yet, this is not the only definition of open data available. Although similar, the Draft Guidelines exclude the definition of "open data" provided in the Draft National Open Science Policy, which it defines as "data that anyone can freely access, use and share, subject, at most, to requirements that preserve provenance and openness" (DSI, 2022). Additionally, the *National Integrated ICT Policy White Paper* (ICT Policy White Paper) defines open data as "datasets that can be freely used, re-used and distributed by anyone, only subject to (at the most) the requirement that users attribute the data and that they make their work available to be shared as well" (Department of Telecommunications and Postal Services, 2016). Having regard to these other definitions of open data that exist would provide researchers with a more comprehensive idea of how open data has been defined by various South African government departments. Therefore, I suggest that the Draft Guidelines develop their own definition of "open data"—that aligns with its objectives—but that references those found in the Draft National Policy on Data and Cloud, the Draft National Open Science Policy, and the ICT Policy White Paper.

3.3 The importance of comprehensive definitions

The provision of definitions serves to assist in providing a common understanding of key terms, thereby lessening the chance of ambiguity and misinterpretation, and ensuring consistent implementation (Whitfield, 2012; Podsakoff et al., 2016). In terms of policies and guidelines, a lack of clear and comprehensive definitions leads to a lack of clarity, which may impede the achievement of policy objectives.

The Draft Guidelines lack definitions relevant to open access data, and only contain a definition of "open data" (defined above). However, had the Draft Guidelines placed this within the broader concept of open science, a definition of such would have been beneficial. Notwithstanding this, there are other definitions relevant to open access and data in research that are pertinent to include. For example, the Draft National Open Science Policy defines "open access" as "a set of principles and a range of practices through which research outputs are distributed online, free of cost or other access barriers" (DSI, 2022). This is highly relevant to research in general, and health research specifically. In considering openness, it is not only the data that is relevant, but also the accessibility of such data. Therefore, I suggest that the provision of additional definitions—such as "open access"—in the Draft Guidelines would assist in this regard.

Additionally, the Draft Guidelines seem to make fundamental errors in basic definitions. The terms "open data" and "open access" are not synonymous and should therefore be distinguished. However, the Draft Guidelines refer to "open access," "open data," and "open access data" and appear to conflate these three terms—which causes confusion regarding what is being referred to (NHREC, 2023). "Open data" refers to the data itself that is made freely accessible, while "open access" denotes principles and practices that allow the free sharing of research outputs (which may be inclusive of data). However, the Draft Guidelines only provide a definition of "open data"—which was adopted from the Draft National Policy on Data and Cloud. I suggest that if the Draft Guidelines had regard to other highly relevant policies that deal with open science, open data, and open access—such as the Draft National Open Science Policy—it would be clear that further definitions exist, and which could have been utilized in the Draft Guidelines in order to clarify the different terminology used.

A further point to note is the differences between the two definitions of "open data"—one provided in the Draft National Open Science Policy and the other in the Draft National Policy on Data and Cloud (and utilized in the Draft Guidelines). Both definitions refer to data that is freely available to all and can be used and shared—although the Draft National Policy on Data and Cloud refers to re-use and republishing (Department of Communications and Digital Technologies, 2021), while the Draft National Open Science Policy uses the term "share" (DSI, 2022). However, the second part of both definitions contain a caveat—in the Draft National Policy on Data and Cloud that the rights to privacy, confidentiality, and security as enshrined in the Constitution are protected (Department of Communications and Digital Technologies, 2021), and in the Draft National Open Science Policy that provenance and openness are preserved (DSI, 2022). These parts of the different definitions appear to be at odds: One promotes openness with very little restriction, and the other allows openness, but only insofar as it does not violate rights to privacy, confidentiality, and security. Although the flaws inherent in the definition of "open data" stem from the Draft National Policy on Data and Cloud, its inclusion in the Draft Guidelines means that this antithesis extends to the health research context—where the privacy rights of research participants have come into question given the nature of genomics research where privacy cannot always be guaranteed (Lunshof et al., 2008; Prainsack and Buyx, 2013; Wang et al., 2017).

Given the above, I suggest that the Draft Guidelines consider revising the definitions provided in relation to open access data. The inclusion of additional relevant definitions—such as open science and open access—as well as the provision of a comprehensive and integrated definition of open data will serve to provide greater clarity when interpreting the Draft Guidelines.

3.4 The matter of privacy and consent

Central to health research is the sharing of data and results. Increased access to such data serves to streamline the research process, making it more efficient and participatory by lessening duplication as well as the costs associated with the creation, transfer, and re-use of data (NHREC, 2023). However, on the face of it, such openness seems to be in opposition to privacy. The Draft Guidelines state that there is a "trade-off between protecting privacy and advancing research" (NHREC, 2023). I suggest that positing the interaction between protecting privacy and advancing research as a "trade-off" is a mischaracterization. It is a common myth in the South African context that research is somehow stymied by the new data privacy legislation, POPIA. Respecting privacy rights and advancing research are perfectly compatible, and ought not be conceived of as necessarily in opposition (Thaldar and Townsend, 2020).

The Draft Guidelines also note that although many participants may not want to publicize their health and genetic data, there are some that do and there should be no obstacles to prevent participants, who wish to share their data in an identifiable manner, from doing so—provided that all foreseeable harms resulting from identification are negligible and understood by participants (NHREC, 2023). What is important is that there be an understanding and those that choose to share their data openly do so knowing that their privacy can no longer be guaranteed.

Given the complexities of health and genomics research, as well as the potential risks involved, consent is vital in all health research involving human participants. The Draft Guidelines provide for three types of consent—specific (or narrow) consent, tiered (or differentiated) consent, and broad consent (NHREC, 2023). The Draft Guidelines also mention blanket consent but, where the 2015 DoH Ethics Guidelines stated that blanket consent was "not recommended" (DoH, 2015), the Draft Guidelines do not permit blanket consent as it "cannot sustain fundamental ethical principles, especially that of protection of privacy" (NHREC, 2023). While these modes of consent are relevant, an additional mode of consent that is aligned with the idea of open science is open consent. Open consent was developed by the Harvard Personal Genome Project (PGP) in response to the recognition that, given the nature of genomics research, privacy cannot be guaranteed (Lunshof et al., 2008). It therefore entails individuals openly donating and sharing their data for research without any assurances of anonymity, privacy, or confidentiality (Lunshof et al., 2008). To ensure that consent is informed, individuals are made aware of the benefits and risks of participation (Lunshof et al., 2010), and are additionally required to pass (with full marks) an assessment that tests their understanding of genomics and privacy (Angrist, 2009). By doing away with any expectations of privacy and taking extra steps to ensure that consent is informed, open consent may offer a potential solution to the contention between open access and privacy. Open consent can essentially be viewed as a type of blanket consent to making data open access, as well as an assessment ensuring that the consent is informed (Gooden and Thaldar, 2023a). However, open consent does differ from blanket consent in certain respects. First, while blanket consent may be utilized for data that has been de-identified, open consent makes no such guarantees, and the publishing and sharing of data is unrestricted and identifiable. Second, open consent can be seen to go a step further than blanket consent in requiring potential participants to pass an assessment in order to ensure that consent is informed. Therefore, open consent furthers open science by combining it (and its benefits) with informed consent.

A potential legal and ethical pathway for an open consent model for genomics research and open access databases in South Africa has already been established (Gooden and Thaldar, 2023a; Thaldar et al., 2023a; Gooden and Thaldar, 2023b). Using this as guidance, I suggest that the Draft Guidelines consider the inclusion of such a model as a means to further open science. Furthermore, I suggest that the Draft Guidelines retain the previous provision regarding blanket consent from the 2015 DoH Ethics Guidelines, where blanket consent was not recommended, but was also not prohibited (DoH, 2015). This provides for the possibility of allowing open consent in health research in South Africa.

3.5 Failure to provide proper guiding principles for open access data

The Draft Guidelines deal with, what it refers to as, "guiding principles for open access". The Draft Guidelines provide that because the Draft National Policy on Data and Cloud supports open access to data, there is a need for guiding principles for health

research. Contrary to what is stated in the Draft Guidelines, it is not only the Draft National Policy on Data and Cloud that supports open access to data. Other policies and reports—such as the Draft National Open Science Policy (DSI, 2022), the POPIA Code of Conduct for Research (ASSAf, 2023), the Academy of Science of South Africa (ASSAf) report on *Human Genetics and Genomics in South Africa: Ethical, Legal and Social Implications* (ASSAf, 2018) (ASSAf Report), the STI White Paper (DST, 2019), the *Synthesis Report: South Africa Foresight Exercise for Science, Technology and Innovation* (DSI, 2019) (Synthesis Report), the *Bio-Economy Strategy* (DST, 2013), and the ICT Policy White Paper (Department of Telecommunications and Postal Services, 2016)—also promote *inter alia* open access and open data and some provide pathways for doing so. It is true that there may be a need for principles governing open access data for health research, but it must be questioned why the Draft Guidelines have only used the Draft National Policy on Data and Cloud as its basis for doing so.

Before examining each of the guiding principles for open access in the Draft Guidelines, it should be noted that some of the principles in the Draft Guidelines come from the *Concordat on Open Research Data* (Rylance et al., 2016). This concordat was developed by stakeholders in the United Kingdon (UK) and designed for the UK research community. As such, some of the principles for open access adopted in the Draft Guidelines may not align with South Africa's research space and the principles of open science that are promoted in the country.

3.5.1 Principle (1): data curation is required to preserve data with acknowledged long-term value

Data curation is important in promoting open access (and thereby open science) in research as it maintains the integrity and value of open data. However, the concept of curation is broad and multifaceted, ranging from the selection of data to its management (Lee and Stvilia, 2017). The Draft Guidelines use the term "data curation" in relation to open access, but fail to define it. Further, the Draft Guidelines, in a separate section, require the Principal Investigator to comply with POPIA in terms of *inter alia* data curation (NHREC, 2023). However, there is no mention of data curation in POPIA or in the POPIA Code of Conduct for Research. On the other hand, the Draft National Open Science Policy does refer to data curation. Although not defined, the Draft National Open Science Policy recognizes that those responsible for funding research must also ensure funding for *inter alia* data curation (DSI, 2022). The Draft National Open Science Policy also notes that open science infrastructure is vital in long-term data curation (DSI, 2022). Given the range of meanings that data curation may have, I suggest that it would be beneficial for the Draft Guidelines to provide a definition of their interpretation of "curation" in order to provide clarity to researchers.

The Draft Guidelines mention the preservation of data with "acknowledged long-term value" (NHREC, 2023). But how will this long-term value be determined? Given the nature of health and genomics research that requires vast amounts of data, which can be used and then re-used for different projects, does all data not have some sort of long-term value? Additionally, it cannot be said that data, which is viewed as having little value now, will not be hugely invaluable at some point in the future—especially given the rate at which technology is advancing, and sometimes in unpredictable ways. As such, it does not seem practical or feasible to determine the long-term value of data used in research. Similar to the Draft Guidelines, the Draft National Open Science Policy makes mention of long-term. However, it refers to "long-term data curation" (DSI, 2022), rather than the curation of data with long-term value (NHREC, 2023). The Draft National Open Science Policy also provides a means of ensuring long-term data curation, namely, through data management plans (DSI, 2022).

Although data management plans tend to focus on active research, and long-term data curation deals with the preservation, maintenance, and accessibility of data after the research has been completed (Lee and Stvilia, 2017; NIH, 2023), it is often beneficial to include long-term data curation within a data management plan. This ensures proper planning, visibility and accountability, adequate resource allocation, and provides a consolidated guide that encompasses both current and long-term data management (Coresignal, 2021; UCLA, 2023). Depending on the nature of the research, the type of data collected and its intended use, the research objectives, data sharing, the complexity of the data, and ethical and legal considerations, data curation may need to be more detailed, and may even require a separate document (Lee and Stvilia, 2017; Miller, 2023).

To provide greater clarity to researchers, I suggest that the Draft Guidelines amend this principle to be more in line with the Draft National Open Science Policy. There are two possible ways in which this can be achieved: (1) the Draft Guidelines amend the current principle to "strategies for long-term data curation are required"; or (2) the Draft Guidelines remove the current principle and combine it with principle (4) regarding data management plans, which is discussed below. I suggest that each of the guiding principles for managing open access data provided in the Draft Guidelines contain an explanation in order to expand on the principle and provide proper, and more detailed, guidance to researchers. Therefore, in terms of (1), the Draft Guidelines can explain that detailed long-term data curation may not be required for all research projects, and it depends on the research. In terms of (2), the Draft Guidelines can specify that long-term data curation be included as part of the data management plan—in line with the Draft National Open Science Policy—or, where required and depending on certain factors like the nature of the research and the type of data collected, long-term data curation be detailed separately.

3.5.2 Principle (2): the right of creators of research data to reasonable first use should be recognized

The principle relating to reasonable first use in the Draft Guidelines was adopted from the UK Concordat on Open Research Data (Rylance et al., 2016). Unlike the UK Concordat on Open Research Data, the Draft Guidelines provide no explanation as to what this principle entails. It is evident that a move towards open science requires the sharing of many aspects of research, including original data. According to the UK Concordat on Open Research Data, this may deter researchers from sharing their data openly, given the time and expertise involved, which would create an obstacle in advancing the goals of open science. However, in certain fields, like genomics, swift data sharing is expected (Rylance et al., 2016). The UK Concordat on Open Research Data provides that in order to encourage researchers to develop and share their data, those who create original data

must be granted a reasonable right of exclusive first use for a suitable period, which is to be established through consultation and included in data management plans (Rylance et al., 2016). The right of creators of research data to reasonable first use is not a typical guiding principle for managing open access data. Open access encourages data sharing, but does not specify how data should be used prior to it being shared or the rights of the data creator (Fecher et al., 2015).

I suggest that the Draft Guidelines remove reference to the right to reasonable first use, and instead focus on ownership. In South Africa, the current position is that the data generator can acquire ownership of the data (Thaldar, 2024 forthcoming; Thaldar et al., 2022). Therefore, there is no need to deal with the right to reasonable first use in this context. Recent academic literature has established that in South African law, instances of data are susceptible of private ownership (Thaldar et al., 2022), and further, that research institutions are best positioned to claim ownership of these newly generated data instances (Thaldar, 2024 forthcoming). However, having ownership in research data instances does not mean that research institutions can do as they wish with the data. Research institutions will be subject to: (1) ethics oversight by a health research ethics committee; and (2) the provisions of POPIA (Thaldar, 2024 forthcoming).

It is important that the Draft Guidelines differentiate data ownership from copyright in datasets. While ownership of data is governed by property law—as found in South Africa's common law—copyright in a dataset is governed by intellectual property law—specifically the Copyright Act 98 of 1978, 1978. Although these areas of law overlap, copyright in a dataset provides a layer of legal protection separate from ownership (Thaldar, 2024 forthcoming; Thaldar, et al., 2022; Swales, et al., 2023). In South Africa, the right of first use—or the exclusive right of use—features in copyright law. In terms of section 7(a) of the Copyright Amendment Bill (2018), where public funding is involved in research, the creator of the work may publicize it, even if an exclusive right of use exists. Therefore, it is clear that the focus of this principle lies in copyright and not ownership.

Being the data owner will assist in giving researchers the confidence that they have the right to openly share their data—thereby promoting open access and open science. As such, I suggest that this principle be replaced with the following: "Data generators, as owners of the data, should be encouraged to openly share their data". This revised principle should explain: (1) the position on ownership of data in South African law; (2) the fact that ownership and intellectual property rights should not be confused; and (3) how data generators should promote open access and open science by sharing their data. Additionally, recognition should be given to indigenous people, in line with the CARE Principles. The Draft National Open Science Policy acknowledges that the CARE Principles deal with research that is not unethical or exploitative, and where the design of data ecosystems ensures that indigenous people benefit from such research (DSI, 2022). The Draft Guidelines contain a section on indigenous knowledge, but it does not deal with this in terms of data ownership and related ethical principles (NHREC, 2023). By overlooking data ownership in South Africa, I suggest that the Draft Guidelines are neglecting a vital aspect of open access data, which will only lead to further difficulties.

3.5.3 Principle (3): for sound reasons, openness of research data may be restricted

The Draft Guidelines provide that the openness of research data may be limited if there are "sound reasons" for doing so (NHREC, 2023). However, it is unclear what constitutes a sound reason. This principle in the Draft Guidelines was adopted from the UK Concordat on Open Research Data (Rylance et al., 2016), which provides that, in certain circumstances, open access to research data may be restricted—for example, to protect privacy and confidentiality of participants, to avoid excessive costs, to uphold consent, to manage risks, to safeguard intellectual property rights, and to abide by other legal limitations (Rylance et al., 2016; Besançon et al., 2021). Moreover, the Organization for Economic Co-operation and Development (OECD) *Principles and Guidelines for Access to Research Data from Public Funding* (OECD, 2007) provide that access to, and use of, certain research data may be limited in some instances, such as national security, privacy and confidentiality, intellectual property rights, and legal processes (OECD, 2007). Governance arrangements, based on good practice and grounded in legal, regulatory, and ethical requirements, should be implemented to establish if and how data should be made openly available (Rylance et al., 2016). The UK Concordat on Open Research Data emphasizes that limitations on openness should not constitute a blanket ban, but should be determined on a case-by-case basis (Rylance et al., 2016). In terms of research publications, it has been suggested that, by default, data should be shared—with providing access to raw data as a prerequisite for manuscript submission. Where this is not possible, journal editors should request that the raw data is inspected by a reliable third party to verify the existence of the raw data and confirm the research results (Besançon et al., 2021).

In South Africa, publicly funded research is governed in part by the Intellectual Property Rights from Publicly Financed Research and Development Act 51 of 2008, 2008 (IPR Act). Section 2 of the IPR Act provides that intellectual property arising from publicly financed research must be protected, used, and commercialized in a way that benefits South Africa (Townsend et al., 2023). The Draft Guidelines refer to publicly funded research, stating that it is a public good and should be made openly available without imposing unwarranted or unjustifiable limitations (NHREC, 2023). The Draft National Open Science Policy applies to all publicly funded research, as well as data that is generated or acquired using public funds (DSI, 2022). In following the principle of "as open as possible, as closed as necessary," certain research projects may entail licensing conditions—which will be determined on a case-by-case basis and by balancing open science and intellectual property licensing (DSI, 2022). Although public funders may have conditions regarding accessibility of the research in their contracts, these contracts do not override any statutory obligations to publicize research. Research funded by the private sector is often subject to contractual terms, but the Draft National Open Science Policy is to be applied in the best way possible, while respecting the private sector funding conditions (DSI, 2022). This is an example of an instance in which the openness of research data may be restricted. As such, I suggest that the Draft Guidelines elaborate on situations when the openness of research data may be restricted, what these sound reasons are, and how they will be implemented. Additionally, the guiding principles in the Draft Guidelines should promote open

access wherever possible, rather than restrict it. In line with this, I suggest that the Draft Guidelines rephrase this principle to state that: "openness of research data should be promoted, wherever possible". An explanation can be provided under this guiding principle with a caveat listing instances where openness may be restricted.

3.5.4 Principle (4): a data management plan should be established at the start of the research process

A data management plan is a formal document that details how data will be handled throughout a research project. It addresses the data to be gathered during a research project, its management, analysis, and storage, as well as measures for sharing and preserving data once the research is complete (IBM, 2023; University of Pretoria, 2023). The Draft Guidelines recognize the importance of establishing a data management plan at the beginning of the research process. This is also provided for in the Draft National Open Science Policy, which requires data management plans for all publicly funded research in order to ensure long-term data curation and stewardship of open data (DSI, 2022). A way in which the Draft Guidelines can promote open science in its guiding principles for open access is to require that data management plans, where applicable, describe how data used in research will be made open—such as alignment with government standards and the principles of findability, accessibility, inter-operability, and re-usability (FAIR)—in line with, and as provided for in, the Draft National Open Science Policy (DSI, 2022).

Additionally, the POPIA Code of Conduct for Research contains the relevant information that researchers must include in their research protocol. A research protocol is defined as "documentation that outlines the plan of a research study" and is inclusive of a data management plan (ASSAf, 2023). These research protocols must encompass the data being collected and its purpose, safeguards, and data quality reviews (ASSAf, 2023). Given that the POPIA Code of Conduct for Research deals with health and genomics research, and given that it contains requirements for research protocols, I suggest that the Draft Guidelines make specific reference to the POPIA Code of Conduct for Research when dealing with data management plans. This will ensure that researchers are provided with further, and detailed, guidance that is in line with data protection laws in South Africa.

3.5.5 Principle (5): use of secondary data should be governed by legal, ethical and regulatory frameworks that promote protection of personal information of donor/participants

The Draft Guidelines state that the use of secondary should be governed by legal, ethical, and regulatory frameworks that protect personal information, but fail to expand on what these frameworks are. For example, POPIA—as well as the POPIA Code of Conduct for Research—are specifically designed for this purpose, but are not mentioned in this section of the Draft Guidelines. Without concrete guidance and clarity, the guiding principles for open access data provided in the Draft Guidelines fall short.

Furthermore, it is not just the secondary use of data that is important. The initial processing of data must adhere to data protection laws. Section 13(1) of POPIA requires that personal information be collected for a "specific, explicitly defined and lawful purpose". Section 15(1) of POPIA allows for the further processing of personal information, provided that it is compatible with the purpose for which it was originally collected. Therefore, if data was initially collected for research, any subsequent use of the data for research is allowed in terms of POPIA. Further, where personal information is used for *inter alia* research purposes, section 15(3) (*e*) of POPIA provides that further processing is compatible with the purpose of collection—as long as the information is only processed for research and is not published in an identifiable manner. If the processing involves special personal information—which is inclusive of genomic data—further processing is permitted, provided that it is for research and: (1) the research serves a public interest, which the processing is necessary for, or it would be unfeasible or involve an excessive effort to obtain consent; and (2) the responsible party can assure that the processing does not negatively and disproportionately impact the data subject's privacy (section 27(1) (*d*) of POPIA). POPIA provides the primary protection for the use and secondary use of personal information, but the POPIA Code of Conduct for Research—which was developed to assist in ensuring legal certainty and compliance with the relevant provisions in POPIA (ASSAf, 2023) —offers additional guidance in this regard.

The POPIA Code of Conduct for Research is mentioned in the Draft Guidelines in terms of privacy and confidentiality of participants, and offers a means to ensure that researchers are compliant with POPIA (NHREC, 2023). But the POPIA Code of Conduct for Research is overlooked in terms of the secondary use of data. The POPIA Code of Conduct for Research deals with further processing (or secondary use). This occurs where the purpose for which the personal information is used changes, or the personal information is re-used for a different purpose (ASSAf, 2023). Where personal information collected for previous research is sought to be used for a different purpose, the researcher must provide certain information, including: (1) the circumstances under which the personal information was collected; (2) how assurances will be made that the personal information will only be used for research and will not be published in an identifiable manner; (3) how the notification requirement in section 18 of POPIA will be complied with; and (4) whether permission has been obtained from the responsible party who originally processed the personal information (ASSAf, 2023; Townsend et al., 2023).

The Draft Guidelines, while providing a principle regarding the protection of personal information, only consider secondary use of data (and not initial use) and fail to define the "legal, ethical and regulatory frameworks" that are applicable. This means that there is a lack of guidance regarding this important aspect of research, and which could lead to a contravention of the provisions in POPIA. To amend this, I suggest that the Draft Guidelines revise this guiding principle as follows: "the use and re-use of data should be governed by legal, ethical, and regulatory frameworks that promote the protection of personal information". Additionally, I suggest that the Draft Guidelines: (1) provide for both the initial use, as well as the re-use, of data; and (2) make reference to POPIA and the POPIA Code of Conduct for Research. However, the Draft Guidelines should ensure that they state the law as it exists, rather than attempting to engage in an interpretive exercise.

3.5.6 Principle (6): use of secondary data should include appropriate acknowledgement of the sources of their data and adhere to the terms of access and use

The final guiding principle for managing open access data in the Draft Guidelines provides that use of secondary data should acknowledge its sources and comply with the terms of access and use. This principle in the Draft Guidelines is taken from the UK Concordat on Open Research Data (Rylance et al., 2016). It is important for subsequent users of data to comply with any rules or restrictions placed on the data (Rylance et al., 2016). The UK Concordat on Open Research Data requires that researchers cite all data that they use in order to acknowledge the data source and creator (Rylance et al., 2016). Open access entails the sharing of data, which strengthens the usefulness and impact of data and increases accountability by allowing others to test analyses or utilize different methodologies to replicate findings (Devriendt et al., 2022). However, in order to ensure that open science is promoted, and researchers are incentivized to openly publish their data, original sources and creators should be acknowledged (Devriendt et al., 2022).

While the Draft Guidelines refer to the "use of secondary data," most other policies and strategies in South Africa dealing with open science, open access, and open data refer to re-use. Although POPIA and the POPIA Code of Conduct for Research do not specifically require acknowledgement of the data source, it promotes transparency—a lawful ground for the processing of personal information in POPIA—and it is good practice to acknowledge sources.

The Draft National Open Science Policy, while not specifically referring to "secondary use," does refer to "re-use" and permits data to be used and re-used freely without restriction, and without the need to acknowledge sources (DSI, 2022). On the other hand, both the Draft National Policy on Data and Cloud (Department of Communications and Digital Technologies, 2021) and the ICT Policy White Paper (Department of Telecommunications and Postal Services, 2016) are more restrictive in terms of the re-use of data. The Draft National Policy on Data and Cloud states that "data must be provided under terms that permit re-use and redistribution" (Department of Communications and Digital Technologies, 2021). Part of the definition of "open data" in the ICT Policy White Paper provides that datasets may be used and re-used, but "that users attribute the data and that they make their work available to be shared as well" (Department of Telecommunications and Postal Services, 2016). Moreover, one of the principles of the ICT Policy White Paper is that identified data "should be freely available for redistribution, use and re-use on conditions, including that the source of the data is identified, and that it is redistributed under the same terms and conditions" (Department of Telecommunications and Postal Services, 2016). However, the subsequent principle in the ICT Policy White Paper requires that data be legally open, meaning that it is in the public domain and can be used and re-used without restriction (Department of Telecommunications and Postal Services, 2016). Therefore, in terms of the re-use of data and acknowledgement of the original source, there seem to be conflicting views.

As good practice, I suggest that the Draft Guidelines amend this guiding principle to read as follows: "The re-use of data should include appropriate acknowledgement of the sources and adhere to the terms of access and use". It is important for the Draft Guidelines to clarify what is meant by this guiding principle and what is required of researchers in this regard.

3.5.7 Conclusion on the Draft Guidelines' guiding principles for open access data

In determining guiding principles for open access data, the Draft Guidelines rely solely on the Draft National Policy on Data and Cloud to the exclusion of other relevant policies and documents. However, open data—as the Draft Guidelines define it—cannot be viewed in isolation, and regard must be had to the broader concept of open science. Open science and its related terms—such as open access and open data—feature in several government policies and strategies and offer potential pathways for the open sharing of data. Many of the existing policies and strategies do not provide concrete guidance on open science or open access, but rather call for the establishment of a policy or framework to govern the field (Department of Telecommunications and Postal Services, 2016; DSI, 2019; DST, 2019). However, there are those that are more detailed in offering objectives and principles for open science (including open access and open data). Below, I consider five main government documents—the ICT Policy White Paper, the Draft National Policy on Data and Cloud, the AOSP, the STI White Paper, and the Draft National Open Science Policy. I suggest that the Draft Guidelines be cognizant of these documents and incorporate certain principles, where relevant.

The ICT Policy White Paper aims to utilize Information and Communication Technologies (ICTs) to reduce poverty and inequality in South Africa (Department of Telecommunications and Postal Services, 2016). Part of the ICT Policy White Paper includes a focus on open government and open data. This entails that essential data is freely available, provided that privacy, confidentiality, and security are protected (Department of Telecommunications and Postal Services, 2016). The principles for open data include that: (1) making data open should be the norm, without violating an individual's right to privacy and security; (2) data that is personal and confidential remains protected; (3) identified data should be freely available for redistribution, use, and re-use subject to certain conditions, including identification of the data source and redistribution under the same terms and conditions; (4) data must be available in the public domain without restriction and published in machine readable, non-proprietary formats; and (5) all data must be accessible and discoverable (Department of Telecommunications and Postal Services, 2016).

The Draft National Policy on Data and Cloud aims to promote the socio-economic value of data and create an enabling environment for the data ecosystem to flourish through *inter alia*: (1) the promotion of access to data and cloud services; (2) the establishment of measures for infrastructure protection; (3) the formation of governance mechanisms for data and cloud services; and (4) the provision of research and innovation (Department of Communications and Digital Technologies, 2021). The Draft National Policy on Data and Cloud recognizes that data should be equally available to all for its benefits to be realized, and that open data is vital in the data revolution (Department of Communications and Digital Technologies, 2021). As such, there is a need for an open data strategy in South Africa, informed by 'Data for Good'

principles, to increase the accessibility of data (Department of Communications and Digital Technologies, 2021).

The AOSP recognizes that the shift to open science is necessary (AOSP, 2023). As such, the AOSP suggests the creation of an African Open Science Platform (the Platform) aimed at empowering African scientists with resources and principles for open science. This initiative is designed to foster scientific excellence and promote the practical application of scientific knowledge in various sectors. The AOSP envisions a platform that supports data-driven research focused on solutions, promoting collaboration between scientists and non-scientists within open networks. Through this collaborative approach, the AOSP aims to generate practical knowledge, enhance the credibility and relevance of science, and bolster its socio-political standing in Africa (AOSP, 2023). The AOSP aims to: (1) map the current data and science initiatives in Africa; (2) create a Pan-African open science community; and (3) develop frameworks to guide the Platform (AOSP, 2023). Given that science communities need to be large, diverse, and collaborative in order to succeed, the AOSP believes that the Platform should be Pan-African. Africa is diverse and this strength should be utilized in order to realize its potential. The AOSP suggests that an individual approach to science in Africa, especially where science communities are small and lack funding, would be a missed opportunity (AOSP, 2023).

Among the policy intents of the STI White Paper is ensuring that South Africa's knowledge system is open, diverse, and responsive (DST, 2019). The STI White Paper recognizes the importance of transdisciplinary knowledge and the data-driven nature of research. Open science offers a solution for greater access to existing information and to benefit from collaborative and transdisciplinary approaches to knowledge development (DST, 2019). However, transitioning to open science requires suitable regulatory frameworks and the development of data skills (DST, 2019). The STI White Paper offers several measures that will be taken in adopting open science in South Africa. These include: (1) promoting open science incentives through education and researcher career development programs; (2) evaluating (and removing) barriers to open science and ensuring that legislation and practice support open science principles; (3) reviewing policies and institutions that govern access to research data and publications, and encouraging researchers to upload their data in public repositories and publish in open access journals; (4) identifying a license system for depositing, and using, open data; (5) respecting the data provider by determining who can use the data, and under what conditions; (6) a reconsideration of the IPR Act to ensure that it supports the findable, accessible, interoperable, and reusable (FAIR) guiding principles for the management and storage of data; (7) the development of a model for data storage and the cloud; and (8) the harmonization of data repositories (DST, 2019). Part of the intentions of the STI White Paper in terms of open science are to develop a framework containing guidelines and principles for open science in South Africa (DST, 2019). This resulted in the Draft National Open Science Policy.

Importantly, the Draft National Open Science Policy specifically provides guiding principles for open science in South Africa (DSI, 2022). The guiding principles for open science are based on the following core values: (1) quality and integrity through transparency, critique, and reproducibility; (2) equity, fairness, and collective benefit; and (3) diversity, collaboration, and inclusiveness (DSI, 2022). Additionally, there are guiding principles to assist in implementing open science in South Africa: (1) publicly funded data and results must be findable, accessible, inter-operable, and re-useable (FAIR); (2) cognisance of collective benefit, authority to control, responsibility, and ethics (CARE) principles, which deal with the ethical and non-exploitative framing of research; (3) the principles of transparency, responsibility, user community, and sustainability, and technology (TRUST) be taken into account when evaluating, developing, and maintaining the trustworthiness of data repositories; (4) a flexible approach to open science that is based on its context; (5) the open science model must be financially and operationally sustainable in the long-term; (6) the principles of "as open as possible, as closed as necessary" will be followed, which means that research outputs must be open and align with the objectives of the Draft National Open Science Policy, unless outweighed by other risks (DSI, 2022).

Based on the above, it is essentially only the Draft National Open Science Policy that explicitly provides guidelines for open science in South Africa. Although useful, it is clear that these guidelines are broad and are not tailored to the specific area of health research. Nevertheless, I suggest that the Draft Guidelines place greater reliance on the various government policies and strategies in existence as they are essential in the realization of open science in South Africa. The Draft Guidelines should be cautioned against adopting principles from other jurisdictions, as was done through reliance on the UK Concordat on Open Research Data (Rylance et al., 2016).

The AOSP highlights that, in adapting to open science, Africa should do so in its own way and based on its own priorities, rather than following other jurisdictions (AOSP, 2023). The AOSP recognizes that Africa should create its own open science platform, with the prospect of promoting science, society, and economic development (AOSP, 2023). A failure to do so will result in dependence on, and requiring skills from, other countries which will not serve to advance science and research (AOSP, 2023). As such, by using guiding principles from the UK Concordat on Open Research Data, the Draft Guidelines do little to serve and further the African agenda.

4 Suggestions for improving the Draft Guidelines

The guiding principles for managing open access data provided by the Draft Guidelines lack concrete guidance on a pathway for the use and sharing of open access data in health research in some respects. These guiding principles appear more as values that have little to do with promoting openness and access, and rather focus on the protection and limitation of such data. As such, there is certainly room for improvement, specifically in terms of the guiding principles for managing open access data. Below, I provide consolidated suggestions for improving the Draft Guidelines based on my analysis above.

4.1 Principle (1): strategies for long-term data curation are required

The suggestions for principle (1) are as follows: (1) provide a definition of "curation" in order to provide clarity to researchers; (2)

remove reference to data curation in terms of POPIA as it does not appear in the Act; and (3) clarify how long-term value will be determined, or acknowledge that in the context of health research, it is likely that all data will be valuable in the long-term. The Draft Guidelines can explain that detailed long-term data curation may not be required for all research projects, and it depends on the research. Alternatively, this principle can be combined with principle (4) regarding data management plans below, in which case the Draft Guidelines can specify that long-term data curation be included as part of the data management plan or, where required and depending on certain factors like the nature of the research and the type of data collected, long-term data curation be detailed separately.

4.2 Principle (2): data generators, as owners of the data, should be encouraged to openly share their data

The suggestions for principle (2) are as follows: (1) remove reference to the right to reasonable first use, and instead focus on ownership; (2) explain the position on ownership of data in South African law; (3) differentiate data ownership from copyright in datasets; (4) promote the open sharing of data by data generators; and (5) recognition should be given to indigenous people and their data in terms of the CARE Principles.

4.3 Principle (3): openness of research data should be promoted, wherever possible

The suggestions for principle (3) are as follows: (1) elaborate on situations when the openness of research data may be restricted, what these sound reasons are, and how they will be implemented; and (2) provide an explanation under this guiding principle that contains a caveat listing instances where openness may be restricted.

4.4 Principle (4): a data management plan should be established at the start of the research process

The suggestions for principle (4) are as follows: (1) require that data management plans, where applicable, describe how data used in research will be made open; and (2) make specific reference to the POPIA Code of Conduct for Research, which contains requirements for research protocols.

4.5 Principle (5): the use and re-use of data should be governed by legal, ethical, and regulatory frameworks that promote the protection of personal information

The suggestions for principle (5) are as follows: (1) provide for both the initial use, as well as the re-use, of data; and (2) make reference to POPIA and the POPIA Code of Conduct for Research as the "legal, ethical and regulatory frameworks" that are applicable. The Draft Guidelines should be cautioned against interpreting the law, and should rather state the law as it exists.

4.6 Principle (6): the re-use of data should include appropriate acknowledgement of the sources and adhere to the terms of access and use

The suggestions for principle (6) are as follows: (1) remove reference to "secondary data" and replace it with "re-use"; and (2) clarify what is meant by this guiding principle and what is required of researchers in this regard.

In addition to the guiding principles for open access data, there are additional considerations that I suggest the Draft Guidelines take into account: (1) avoid placing sole reliance on the Draft National Policy on Data and Cloud and adopting principles from the UK Concordat on Open Research Data that may not apply in South Africa in their current form; (2) explicitly mention open science and expand on its importance in health research; (3) develop a comprehensive definition of "open data" that takes into account other definitions provided by the Draft National Open Science Policy and the ICT Policy White Paper; (4) provide other definitions relevant to open access and data in research, such as "open science" and "open access," and differentiate between "open data," "open access," and "open access data"; (5) provide a potential pathway for open consent to further open science; (6) retain the previous provision in the 2015 DoH Ethics Guidelines regarding blanket consent to allow for the possibility of open consent; (7) refer to other South African government documents that deal with open science, open access, and open data to bolster the Draft Guidelines; and (8) include reference to South African legislation, where relevant. I also suggest that each of the guiding principles for managing open access data provided in the Draft Guidelines are accompanied by an explanation in order to expand on the principle and provide proper, and more detailed, guidance to researchers.

5 Conclusion

Health and genomics research in South Africa have a vital role to play in bettering the health of the population through an increased understanding of various diseases and the ability to develop more effective treatments and advance healthcare and technologies. However, its full potential cannot be realized if data and resources are not open and accessible to others. The Draft Guidelines serve to guide researchers in conducting health research in an ethical and responsible manner. Although the Draft Guidelines set the benchmark for health research in South Africa and are invaluable in certain respects, the inclusion of open access databases in the Draft Guidelines requires improvement. By only relying on one draft government policy—namely, the Draft National Policy on Data and Cloud—and overlooking other drafts that are relevant, such as the Draft National Open Science Policy, the Draft Guidelines cannot provide a comprehensive and context-specific pathway for open access data in research. Additionally, and from a policy perspective, the Draft Guidelines have an obligation to consider, and align with, principles of open science.

By failing to expressly do so, the Draft Guidelines fall short in this regard.

While the Draft Guidelines and its inclusion of open access, especially in the context of health research, is a positive step towards open science and the transformation of the research landscape in South Africa, there is room for improvement. Specifically, the Draft Guidelines should: (1) specifically include reference to open science and its importance in South Africa; (2) add additional (and comprehensive) definitions for clarity, such as "open science" and "open access"; (3) consider the pathway for open access databases in South Africa by relying on an open consent model; and (4) have regard to the guiding principles for open access data and ensure that detailed guidance is provided to researchers, with reference being made to other relevant South African legislation and policy. The Draft Guidelines can also place reliance on existing policies and strategies that deal with open science and open access in order to align the Draft Guidelines with national imperatives. The implementation of these suggestions will serve to strengthen the Draft Guidelines and its position on open access databases.

Author contributions

AG: Formal Analysis, Writing–original draft, Writing–review and editing, Investigation.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. Work on this article was supported by the U.S. National Institute of Mental Health and the U.S. National Institutes of Health (Award Number. U01MH127690) under the Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) program.

Acknowledgments

The author would like to acknowledge the Google PhD Fellowship Program. The author would also like to thank Donrich Thaldar for his useful comments on this manuscript.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

The content of this article is solely my responsibility and does not necessarily represent the official views of the U.S. National Institute of Mental Health or the U.S. National Institutes of Health.

References

Abdool Karim, Q., and Baxter, C. (2022). COVID-19: impact on the HIV and Tuberculosis response, service delivery, and research in South Africa. *Curr. HIV/AIDS Rep.* 19, 46–53. doi:10.1007/s11904-021-00588-5

Abdool Karim, S. S., Churchyard, G. J., Abdool Karim, Q., and Lawn, S. D. (2009). HIV infection and tuberculosis in South Africa: an urgent need to escalate the public health response. *Lancet* 374 (9693), 921–933. doi:10.1016/S0140-6736(09)60916-8

Academy of Science of South Africa (ASSAf); Department of Science and Technology (DST). (2018). Human genetics and genomics in South Africa: ethical, legal and social Implications. doi:10.17159/assaf.2018/0033

Academy of Science of South Africa (ASSAf) (2019). African open science platform – Part 1: landscape study. Available at: https://aosp.org.za/wp-content/uploads/2022/05/aosplandscape-12aug2019_compressed.pdf.

Academy of Science of South Africa (ASSAf) (2023). POPIA Code of conduct for research. Available at: https://www.assaf.org.za/wp-content/uploads/2023/04/ASSAf-POPIA-Code-of-Conduct-for-Research.pdf.

African Open Science Platform (AOSP) (2023). The future of science and science for the future. Available at: https://aosp.org.za/wp-content/uploads/2023/05/AOSP-Strategy.pdf.

Angrist, M. (2009). Eyes wide open: the Personal Genome Project, citizen science and veracity in informed consent. *Per. Med.* 6 (6), 691–699. doi:10.2217/pme.09.48

Baca, M. R. (2006). Barriers to innovation: intellectual property transaction costs in scientific collaboration. *Duke L. Tech. Rev.* 5, 1–14.

Begley, C., and Ellis, L. (2012). Drug development: raise standards for preclinical cancer research. *Nature* 483, 531–533. doi:10.1038/483531a

Besançon, L., Peiffer-Smadja, N., Segalas, C., Jiang, H., Masuzzo, P., Smout, C., et al. (2021). Open science saves lives: lessons from the COVID-19 pandemic. *BMC Med. Res. Methodol.* 21, 117–118. doi:10.1186/s12874-021-01304-y

Buxton, R. T., Nyboer, E. A., Pigeon, K. E., Raby, G. D., Rytwinski, T., Gallagher, A. J., et al. (2021). Avoiding wasted research resources in conservation science. *Conservation Sci. Pract.* 3, 1–11. doi:10.1111/csp2.329

Chakravorty, N., Sharma, C. S., Molla, K. A., and Pattanaik, J. K. (2022). Open science: challenges, possible solutions and the way forward. *Proc. Indian Natl. Sci. Acad.* 88, 456–471. doi:10.1007/s43538-022-00104-2

Collins, F. S., and Varmus, H. (2015). A new initiative on precision medicine. *N. Engl. J. Med.* 372 (9), 793–795. doi:10.1056/NEJMp1500523

Copyright Act 98 of 1978, 1978 [South Africa].

Copyright Amendment Bill (2018) [South Africa].

Coresignal (2021). Data curation: benefits, goals, and best practices. Available at: https://coresignal.com/blog/data-curation/ (Accessed September 27, 2023).

Day, S., Rennie, S., Luo, D., and Tucker, J. D. (2020). Open to the public: paywalls and the public rationale for open access medical research publishing. *Res. Involv. Engagem.* 6 (8), 8–7. doi:10.1186/s40900-020-0182-y

Department of Science and Technology (DST) (2013). The bio-economy strategy. South Africa. Available at: https://www.gov.za/sites/default/files/gcis_document/201409/bioeconomy-strategya.pdf.

Department of Communications and Digital Technologies (2021). *Draft national policy on data and cloud. GN 306 GG 44389*. South Africa. Available at: https://www.gov.za/sites/default/files/gcis_document/202104/44389gon206.pdf.

Department of Health (DoH) (2015). *Ethics in health research: principles, processes and Structures*. South Africa. Available at: https://www.health.gov.za/wp-content/uploads/2022/05/NHREC-DoH-2015-Ethics-in-Health-Research-Guidelines-1.pdf.

Department of Science and Innovation (DSI) (2019). *Synthesis report: South Africa Foresight exercise for science, technology and innovation 2030*. South Africa. Available at: https://www.naci.org.za/wp-content/uploads/2020/07/South-African-Foresight-Exercise-For-Science-Technology-and-Innovation-2019.pdf.

Department of Science and Innovation (DSI). (2022). Draft national open science policy [South Africa].

Department of Science and Technology (DST) (2019). *White paper on science, technology and innovation*. South Africa. Available at: https://www.dst.gov.za/images/2019/White_paper_web_copyv1.pdf.

Department of Telecommunications and Postal Services (2016). National integrated ICT policy white paper. South Africa. Available at: https://www.gov.za/sites/default/files/gcis_document/201610/40325gon1212.pdf.

Devriendt, T., Borry, P., and Shabani, M. (2022). Credit and recognition for contributions to data-sharing platforms among cohort holders and platform developers in Europe: interview study. *J. Med. Internet Res.* 24 (1), e25983. doi:10.2196/25983

European Commission (2023). What is open data. Available at: https://data.europa.eu/en/dataeuropa-academy/what-open-data (Accessed September 27, 2023).

Fecher, B., and Friesike, S. (2014). "Open science: one term, five schools of thought," in *Opening science*. Editor S. Bartling (Cham: Springer), 17–47.

Fecher, B., Friesike, S., and Hebing, M. (2015). What drives academic data sharing? *PLoS ONE* 10 (2), e0118053. doi:10.1371/journal.pone.0118053

Friesike, S., Widenmayer, B., Gassmann, O., and Schildhauer, T. (2015). Opening science: towards an agenda of open science in academia and industry. *J. Technol. Transf.* 40, 581–601. doi:10.1007/s10961-014-9375-6

Gewin, V. (2016). Data sharing: an open mind on open data. *Nature* 529, 117–119. doi:10.1038/nj7584-117a

GIDA (2019). CARE principles for indigenous data governance. Available at: https://www.gida-global.org/care (Accessed January 24, 2024).

Gooden, A., and Thaldar, D. (2023a). Toward an open access genomics database of South Africans: ethical considerations. *Front. Genet.* 14, 1166029–1166037. doi:10.3389/fgene.2023.1166029

Gooden, A., and Thaldar, D. (2023b). *The time is ripe for a large-scale South African genome project*. Johannesburg, South Africa: Mail and Guardian. Available at: https://mg.co.za/thoughtleader/opinion/2023-06-13-the-time-is-ripe-for-a-large-scale-south-african-genome-project/ (Accessed September 14, 2023).

Grant, S., Wendt, K. E., Leadbeater, B. J., Supplee, L. H., Mayo-Wilson, E., Gardner, F., et al. (2022). Transparent, open, and reproducible prevention science. *Prev. Sci.* 23, 701–722. doi:10.1007/s11121-022-01336-w

Hardwicke, T. E., Mathur, M. B., MacDonald, K., Nilsonne, G., Banks, G. C., Kidwell, M. C., et al. (2018). Data availability, reusability, and analytic reproducibility: evaluating the impact of a mandatory open data policy at the journal *Cognition*. *R. Soc. Open Sci.* 5 (8), 180448–180518. doi:10.1098/rsos.180448

Haven, T., Gopalakrishna, G., Tijdink, J., van der Schot, D., and Bouter, L. (2022). Promoting trust in research and researchers: how open science and research integrity are intertwined. *BMC Res. Notes* 15, 302–305. doi:10.1186/s13104-022-06169-y

Heise, C., and Pearce, J. M. (2020). *From open access to open science: the path from scientific reality to open scientific communication*. SAGE Open, 1–14. doi:10.1177/2158244020915900

Hetu, M., Koutouki, K., and Joly, Y. (2019). Genomics for all: international open science genomics projects and capacity building in the developing world. *Front. Genet.* 10, 95–99. doi:10.3389/fgene.2019.00095

IBM (2023). What is a data management plan? Available at: https://www.ibm.com/topics/data-management-plan (Accessed September 27, 2023).

Intellectual Property Rights from Publicly Financed Research and Development Act 51 of 2008, 2008 [South Africa].

International Human Genome Sequencing ConsortiumLinton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921. doi:10.1038/35057062

Johnson, L. F., May, M. T., Dorrington, R. E., Cornell, M., Boulle, A., Egger, M., et al. (2017). Estimating the impact of antiretroviral treatment on adult mortality trends in South Africa: a mathematical modelling study. *PLoS Med.* 14 (12), e1002468. doi:10.1371/journal.pmed.1002468

Krishna, V. V. (2020). Open science and its enemies: challenges for a sustainable science–society social contract. *J. Open Innov. Technol. Mark. Complex.* 6 (3), 61–15. doi:10.3390/joitmc6030061

Kubjane, M., Osman, M., Boulle, A., and Johnson, L. F. (2022). The impact of HIV and tuberculosis interventions on South African adult tuberculosis trends, 1990-2019: a mathematical modeling analysis. *Int. J. Infect. Dis.* 122, 811–819. doi:10.1016/j.ijid.2022.07.047

Lee, D. J., and Stvilia, B. (2017). Practices of research data curation in institutional repositories: a qualitative view from repository staff. *PLoS One* 12 (3), e0173987. doi:10.1371/journal.pone.0173987

Levy, E., Marden, E., Warren, B., Hartell, D., and Filaté, I. (2010). Patent pools and genomics: navigating a course to open science? *BU. J. Sci. Tech. L.* 16, 75–101.

Lunshof, J. E., Bobe, J., Aach, J., Angrist, M., Thakuria, J. V., Vorhaus, D. B., et al. (2010). Personal genomes in progress: from the human genome project to the personal genome project. *Dialogues Clin. Neurosci.* 12 (1), 47–60. doi:10.31887/DCNS.2010.12.1/jlunshof

Lunshof, J. E., Chadwick, R., Vorhaus, D. B., and Church, G. M. (2008). From genetic privacy to open consent. *Nat. Rev. Genet.* 9 (5), 406–411. doi:10.1038/nrg2360

Maurer, S. M. (2003). "New institutions for doing science: from databases to open source biology,". 19 November 2003 (University of Maastricht) in *European policy for intellectual property conference on copyright and database protection, patents and research tools, and other challenges to the intellectual property system* (Netherlands). Available at: http://www.merit.unimaas.nl/epip/papers/maurer_paper.pdf.

Mayosi, B. M., Lawn, J. E., van Niekerk, A., Bradshaw, D., Abdool Karim, S. S., Coovadia, H. M., et al. (2012). Health in South Africa: changes and challenges since 2009. *Lancet* 380 (9858), 2029–2043. doi:10.1016/S0140-6736(12)61814-5

McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., et al. (2016). How open science helps researchers succeed. *Elife* 5, e16800–e16819. doi:10.7554/eLife.16800

Miller, O. (2023). 5 steps to creating a data curation plan. *Medium*. Available at: https://medium.com/@oliviamiller048/5-steps-to-creating-a-data-curation-plan-ffe784789c4a (Accessed September 27, 2023).

National Health Research Ethics Council (2023). South African ethics in health research guidelines: principles, processes and Structures. Available at: https://www.samrc.ac.za/research/rio-hrec-guideline-documents.

Nielsen, M. (2011). *Reinventing discovery: the new era of networked science*. New Jersey: Princeton University Press.

NIH (2023). Data curation. Available at: https://www.nnlm.gov/guides/data-glossary/data-curation (Accessed September 27, 2023).

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., et al. (2015). SCIENTIFIC STANDARDS. Promoting an open research culture. *Science* 348, 1422–1425. doi:10.1126/science.aab2374

Open Data Charter (2023). International open data charter. Available at: https://opendatacharter.net/principles/ (Accessed September 27, 2023).

Open Data Handbook (2023). What is open data. Available at: https://opendatahandbook.org/guide/en/what-is-open-data/ (Accessed September 27, 2023).

Open Knowledge Foundation. (2023). Open data. Available at: https://okfn.de/en/themen/open_data/#:~:text=Open%20data%20is%20data%20that,not%20apply%20to%20personal%20data. (Accessed September 27, 2023).

Organization for Economic Co-operation and Development (OECD). (2007). Principles and guidelines for access to research data from public funding.

Podsakoff, P. M., MacKenzie, S. B., and Podsakoff, N. P. (2016). Recommendations for creating better concept definitions in the organizational, behavioral, and social sciences. *Organ. Res. Methods* 19 (2), 159–203. doi:10.1177/1094428115624965

Prainsack, B., and Buyx, A. (2013). A solidarity-based approach to the governance of research biobanks. *Med. Law Rev.* 21 (1), 71–91. doi:10.1093/medlaw/fws040

Protection of Personal Information Act 4 of 2013, 2013 [South Africa].

Ramachandran, R., Bugbee, K., and Murphy, K. (2021). From open data to open science. *Earth Space Sci.* 8, 1–17. doi:10.1029/2020EA001562

Regulations Relating to Research with Human Participants GN R719 of GG 38000, 2014 [South Africa].

Rhoten, D., and Powell, W. W. (2007). The frontiers of intellectual property: expanded protection versus new models of open science. *Annu. Rev. Law Soc. Sci.* 3, 345–373. doi:10.1146/annurev.lawsocsci.3.081806.112900

Rylance, R., Wingham, D., Wright, N., Bruce, R., Hammonds, W., Arrowsmith, J., et al. (2016). Concordat on open research data. Available at: https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf.

Scaria, A. G., and Rangarajan, R. (2016). Fine-tuning the IP approaches for fostering open science: some insights from India. *WIPO J.* 8 (1), 109–122. doi:10.2139/ssrn.2844625

Scott, A. (2017). *What is open data and why should we care*. London, United Kingdom: Open Data Institute. Available at: https://www.theodi.org/article/what-is-open-data-and-why-should-we-care/ (Accessed September 27, 2023).

Smith, E., Haustein, S., Mongeon, P., Shu, F., Ridde, V., and Larivière, V. (2017). Knowledge sharing in global health research: the impact, uptake and cost of open access to scholarly literature. *Health Res. Policy Syst.* 15 (1), 73–10. doi:10.1186/s12961-017-0235-3

Spellman, B., Gilbert, E. A., and Corker, K. S. (2017). Open science: what, why, and how. PsyArXiv, 1–84. doi:10.31234/osf.io/ak6jr

Staunton, C., Barragán, C. A., Canali, S., Ho, C., Leonelli, S., Mayernik, M., et al. (2021). Open science, data sharing and solidarity: who benefits? *Hist. Philos. Life Sci.* 43 (4), 115–118. doi:10.1007/s40656-021-00468-6

Strydom, A., Mellet, J., Van Rensburg, J., Viljoen, I., Athanasiadis, A., and Pepper, M. S. (2022). Open access and its potential impact on public health – a South African perspective. *Front. Res. Metr. Anal.* 7, 975109–975118. doi:10.3389/frma.2022.975109

Swales, L., Botes, M., Donnelly, D., and Thaldar, D. (2023). Towards a data transfer agreement for the South African research community: the empowerment approach. *S. Afr. J. Bioeth. Law* 16 (1), 13–18. doi:10.7196/SAJBL.2023.v16i1.827

Thaldar, D. (forthcoming) (2024). The wisdom of claiming ownership of human genomic data: a cautionary tale for research institutions.

Thaldar, D., Gooden, A., and Donnelly, D. (2023a). Toward an open access genomics database of South Africans: legal considerations. *S. Afr. J. Sci.* 119 (7/8), 1–4. doi:10.17159/sajs.2023/15069

Thaldar, D., Gooden, A., and Steytler, M. (2023b). Open science and human genetic data: recommendations on South Africa's draft national open science policy. *Front. Genet.* 14, 1248747. doi:10.3389/fgene.2023.1248747

Thaldar, D., and Townsend, B. (2020). Genomic research and privacy: a response to Staunton et al. *S. Afr. Med. J.* 110 (3), 172–174. doi:10.7196/SAMJ.2020.v110i3.14431

Thaldar, D. W., Townsend, B. A., Donnelly, D. L., Botes, M., Gooden, A., van Harmelen, J., et al. (2022). The multidimensional legal nature of personal genomic sequence data: a South African perspective. *Front. Genet.* 13, 997595–997611. doi:10.3389/fgene.2022.997595

Townsend, B., Gooden, A., Botes, M., and Thaldar, D. (2023). Repurposing research data for commercial use: POPIA, a foil or a facilitator? *S. Afr. J. Sci.* 119 (7/8), 1–5. doi:10.17159/sajs.2023/15075

UCL (2024). 8 pillars of open science. Available at: https://www.ucl.ac.uk/library/open-science-research-support/open-science/8-pillars-open-science#:~:text=The%208%20pillars%20of%20Open,Rewards%20and%20Initiatives%2C%20and%20EOSC (Accessed January 24, 2024).

UCLA (2023). Digital humanities. Available at: https://guides.library.ucla.edu/c.php?g=180354&p=9650260 (Accessed September 27, 2023).

University of Pretoria (2023). Research data management (RDM): data management plans. Available at: https://library.up.ac.za/c.php?g=356288&p=2412447 (Accessed September 26, 2023).

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., et al. (2001). The sequence of the human genome. *Science* 291 (5507), 1304–1351. doi:10.1126/science.1058040

Vicente-Saez, R., and Martinez-Fuentes, C. (2018). Open Science now: a systematic literature review for an integrated definition. *J. Bus. Res.* 88, 428–436. doi:10.1016/j.jbusres.2017.12.043

Wang, S., Jiang, X., Singh, S., Marmor, R., Bonomi, L., Fox, D., et al. (2017). Genome privacy: challenges, technical approaches to mitigate risk, and ethical considerations in the United States. *Ann. N. Y. Acad. Sci.* 1387 (1), 73–83. doi:10.1111/nyas.13259

Whitfield, G. (2012). The importance of proper definition. Available at: https://piadvice.wordpress.com/2012/06/13/the-importance-of-proper-definition/ (Accessed September 23, 2023).

Wilkinson, M., Dumontier, M., Aalbersberg, I., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3, 160018–160019. doi:10.1038/sdata.2016.18

frontiers | Frontiers in Pharmacology

TYPE Policy and Practice Reviews
PUBLISHED 12 March 2024
DOI 10.3389/fphar.2024.1333672

OPEN ACCESS

EDITED BY
Yann Joly,
McGill University, Canada

REVIEWED BY
Rania Mohamed Hassan Baleela,
University of Khartoum, Sudan
Lijun Shang,
London Metropolitan University,
United Kingdom

*CORRESPONDENCE
Donrich Thaldar,
thaldard@ukzn.ac.za

RECEIVED 05 November 2023
ACCEPTED 28 February 2024
PUBLISHED 12 March 2024

CITATION
Esselaar P, Swales L, Bellengère D, Mhlongo B and Thaldar D (2024), Forcing a square into a circle: why South Africa's draft revised material transfer agreement is not fit for purpose. *Front. Pharmacol.* 15:1333672. doi: 10.3389/fphar.2024.1333672

Forcing a square into a circle: why South Africa's draft revised material transfer agreement is not fit for purpose

Paul Esselaar, Lee Swales, Devarasi Bellengère, Banele Mhlongo and Donrich Thaldar *

School of Law, University of KwaZulu-Natal, Durban, South Africa

The South African National Health Research Ethics Council (NHREC) recently released a final draft revision of the standard material transfer agreement (MTA) that was promulgated into law in 2018. This new draft MTA raises pertinent questions about the NHREC's mandate, the way in which the draft MTA deals with data and with human biological material, and its avoidance of the concept of ownership. After South Africa's data protection legislation, the Protection of Personal Information Act (POPIA), became operational in mid 2021, the legal landscape changed and it is doubtful that the NHREC has a residual mandate to govern personal information in health research. Furthermore, data is dealt with in a superficial, throw-away fashion in the draft MTA. The position with human biological material is not substantially better, as the draft MTA fails to recognise that human biological material can contain pathogens, which has important legal and ethical ramifications that are not sufficiently addressed. A central problem with the draft MTA is its use of the term 'steward', and avoidance of the legal concept of 'ownership'. This is not only misaligned with the South African legal framework, but also fails to consider the ethical case for recognising ownership. Finally, a call to embrace decolonial thinking in health research underscores the importance of recognising ownership in order to foster the growth of the local bio-economy. Key recommendations to reshape the draft MTA include: Making use of the eventual revised MTA optional, and allowing it to evolve with input from scientific and legal communities; regulating the transfer of associated data in a separate data transfer agreement that can be incorporated by reference in the MTA; enhancing guidance on liability and risk management in respect of human biological material that contains pathogens; and, finally, adopting a decolonial approach in health research governance, which requires recognising the ownership rights of South African research institutions.

KEYWORDS

data, decolonial, human biological material, material transfer agreement, ownership, pathogens, POPIA, stewardship

Background

In 2018, the South African Minister of Health published a standard material transfer agreement (MTA) in the Government Gazette and gave notice that research institutions sharing human biological material for health research or clinical trials must use this MTA (SA MTA) (Material Transfer Agreement for Human Biological Materials, 2018). The SA

MTA was controversial from the outset (Thaldar, 2020; Thaldar et al., 2020). The notion of the state in a supposedly open and democratic society, forcing the use of a single template onto everyone is clearly suspect. The situation could, however, have been palatable had the SA MTA been a well-drafted document. However, it was not. Thaldar et al. (2020) highlighted several problems with the SA MTA, ranging from misalignment with extant law to absurdly overbroad clauses.

The only saving grace was that the SA MTA described itself as a "framework," hence leaving latitude for parties that are legally forced to use it to amend the substantive provisions—and hopefully in the process resolve the problematic aspects (Thaldar et al., 2020; Steytler and Thaldar, 2021; Thaldar and Shozi, 2021; Swales et al., 2023a). Using this latitude, a group of South African law academics developed a revised version of the SA MTA in an attempt to rectify the most serious issues while remaining within the bounds of the framework of the original version of the SA MTA (Pope, 2020). The aim of this revised version—called "SA MTA 1.1" and dating from 2020—was to provide the South African research community with a *usable* version of the SA MTA that would still comply with the law.

Next, in 2022, a research group at the University of KwaZulu-Natal in South Africa started with the development of a data transfer agreement (DTA) template for the South African research community. The rationale was that data sharing between researchers requires an expertly drafted agreement that is aligned with South African law—in particular the Protection of Personal Information Act 4 of 2013 (POPIA) that was brought into full operation on 1 July 2021; however, many—if not most—research organisations in South Africa do not have the inhouse legal expertise to have such an agreement drafted (Swales et al., 2023a). Accordingly, the aim was to develop a comprehensive, professionally drafted DTA template and to make it freely available for anyone to use (Swales et al., 2023a). The DTA template was also complemented with an explanatory memorandum to guide users on how to use and amend the template for their own circumstances (Swales et al., 2023b).

The authors of the DTA template explicitly distanced themselves from the authoritarian practice of forcing the use of a document on a country (Swales et al., 2023a; Swales et al., 2023b). Instead, they stated that the South African research community should use the DTA template and explanatory memorandum because these are top quality documents that answer a need, not because they are forced to do so as is the case with the SA MTA (Swales et al., 2023a; Swales et al., 2023b).

However, the controversial SA MTA (the original version) remained in South Africa's lawbooks. Eventually, South Africa's Department of Health decided that the best policy solution was to task a statutory body that functions under its aegis, the National Health Research Ethics Council (NHREC), with revising the SA MTA. In August 2023, the NHREC distributed a final draft version of their revised SA MTA to stakeholders for comment (National Health Research Ethics Council, 2023). An interesting observation is that the NHREC's final draft is largely based on SA MTA 1.1, rather than on the original SA MTA. Thus, the NHREC's final draft benefits from avoiding the well-documented pitfalls of the original SA MTA. However, the NHREC made some consequential changes to SA MTA 1.1. It is also important to note that the South African legal landscape has changed since SA MTA 1.1 was developed. We have already mentioned POPIA's coming into operation. This raises the important question of whether the NHREC's final draft MTA is aligned with POPIA?

In the sections that follow, we delve into a comprehensive examination of the NHREC's draft MTA and its implications for South Africa's research community and their international collaborators. We investigate four questions: First, do the Minister of Health and the NHREC have the mandate to regulate data in the health research context, or are they overstepping their respective mandates? Second, does the draft MTA provide sufficient protection for data? Third, is there sufficient guidance on biological material in the draft MTA? Fourth, why does the draft MTA shy away from the concept of ownership? Flowing from our analyses of these four questions, we propose an alternative approach to the draft MTA, and offer recommendations to address the identified shortcomings and to align the draft MTA with legal standards and with the needs of the scientific community.

Main text

Are the Minister of Health and the NHREC overstepping their respective mandates?

At a fundamental level, the question must be posed: Do the Minister of Health and the NHREC have the mandate to regulate data in the health research context, or are they overstepping their respective mandates? These entities receive their regulatory mandates from the National Health Act 61 of 2003 (NHA). Chapter 8 of the NHA, in particular, together with relevant regulations, governs the use of human biological material and research with human participants. However, there is also a later statute that is relevant in the health research space, namely POPIA, which deals with personal information. Data in the health research space often includes personal data—or to use POPIA terminology, "personal information." Moreover, data in the health research space are often *sensitive personal* data—or to use POPIA terminology, "special personal information." Accordingly, there is an overlap between the scopes of application of the NHA and POPIA. The question then is: In the case of a conflict, which statute prevails? We consider two relevant legal principles.

In the context of health research, the NHA is *general* legislation, while POPIA is *special* legislation, meaning that POPIA governs only a specific part of health research, namely the way in which the personal information of research subjects is dealt with (National Health Act, 2003). Accordingly, the maxim *generalia specialibus non derogant* (general words and rules do not derogate from special ones) applies. This means that the scope of application of the general statute must be constrained by the presence of the specific legislation (Minister of Justice and Constitutional Development v Southern African Litigation Centre, 2016). Applied to health research, this principle means that the governance of personal information is now governed by POPIA first and the NHA second. It follows that the Minister of Health and the NHREC—who get their respective mandates from the NHA—no longer have a mandate to regulate personal information in the health research milieu. This is now done by POPIA and its implementation mechanism, the Information

Regulator. In turn, the Information Regulator can, among others, issue guidance notes, and approve codes of conduct and compliance frameworks. For example, ASSAf developed a draft Code of Conduct for Research (Academy of Science of South Africa, 2023), which is likely to be converted into a compliance framework based on the Information Regulator's feedback.

Furthermore, POPIA itself contains a supremacy clause in section 2(a). In the context of the processing of personal information, POPIA (2013) supersedes any other legislation that is inconsistent with it. There is however an exception to POPIA's supremacy clause in section 2(b). If any other legislation provides for conditions for the lawful processing of personal information that are "more extensive" than those set out in POPIA, the more extensive conditions in the other legislation prevail. Although some have argued that the NHA is more extensive (in the sense that it is certainly more voluminous), this is mistaken (Bronstein and Nyachowe, 2023). In context, "more extensive" clearly refers to *better protection of data subjects*, not to being more voluminous (Thaldar, 2023). This exception may apply in specific instances where other legislation provides better protection of data subjects. However, as we discuss below, this is evidently not the case with the draft MTA. Accordingly, there is no realistic possibility of relying on the exception to POPIA's supremacy clause.

As a result, to the extent that the NHREC's draft MTA contains provisions regarding personal information, it is beyond the Minister of Health's and the NHREC's statutory mandate. The Minister of Health and the NHREC are overstepping into the terrain of the Information Regulator. To the extent that they overstep, their conduct is invalid and can be challenged in a court of law (Sasol Oil Pty Ltd v Metcalfe, 2004). The solution to this problem is obvious: The NHREC should remove all references to "associated data" in its draft MTA.

Next, we analyse the way in which the draft MTA deals with "associated data."

Is there sufficient protection for the associated data in the draft MTA?

Although trite, it bears repetition: POPIA sets out eight conditions for the lawful processing of personal information (De Stadler et al., 2021; Burns and Burger-Smidt, 2023). These conditions are aimed at protecting the rights of data subjects, but POPIA also recognises that a balance must be struck between the right to privacy and the right of access to information and freedom of speech. POPIA therefore establishes conditions that regulate how personal information may be processed. For the avoidance of doubt, POPIA applies to the processing (including transfer) of all personal information, including personal information derived directly and indirectly from health research, such as genetic data generated from human biological material.

In terms of current best practice—in South Africa and internationally—an agreement that facilitates the transfer of data containing personal information should contain *detailed provisions* articulating compliance with applicable data protection legislation. Parties to an MTA must be aware that by transferring data that contains personal information, several legal obligations arise—and these obligations require careful consideration. The parties must determine, *inter alia*, the nature of the personal information being transferred, the identity of the responsible party, and the data privacy obligations on each party. Critically, sections 107 and 109 of POPIA (2013) provides that failure to comply with POPIA can result in a fine of up to R10 million or imprisonment for a period not exceeding 10 years, or to both a fine and such imprisonment—as well as significant reputational harm.

However, the draft MTA fails to live up to best practice. The draft MTA's "Guidance" section notes that the draft MTA is a template that contains "minimum standards." However, as it stands, there are simply no minimum standards in the draft MTA dealing with data protection. The draft MTA refers to data in a superficial, throw-away fashion.

The "Guidance" section further provides that where "data alone" is transferred a data transfer agreement (DTA) is "appropriate." We suggest that in all circumstances where data containing personal information are transferred, in order to ensure full compliance with POPIA, and to abide by international best practice, a DTA is not only appropriate, but necessary. Although it is true that some of the content in a DTA will be similar to a MTA, the similarity relates only to standard legal clauses, and not to the actual substance of the agreement. The primary purpose of the agreements will be entirely different, and both will seek to comply with distinct pieces of legislation. For this reason, the decision to conflate data with human biological material—something inherited from the original SA MTA via SA MTA 1.1—is a mistake.

To illustrate the issues caused by this conflation, consider the following three definitions:

- "Material" is defined as including both human biological material and associated data.
- "Associated data" includes personal information relating to human biological material.
- "Permit" is defined as "authorisation of the National Department of Health to transfer and/or export Material."

However, in relation to personal information (which is part of 'Material' as defined above), the National Department of Health plays no role in its regulation.

Some of the changes that the NHREC's draft MTA introduced to SA MTA 1.1 seem not to have been sufficiently considered. For example, consider the second sentence added to clause 3.5. The clause now reads as follows: "The Provider must inform the HREC [health research ethics committee] and wherever possible the Participant/s if the Provider is informed that the Material has Become Identifiable for any reason whatsoever. This must be clarified as Material remain [sic] coded and hence potentially identifiable." The second sentence is not comprehensible.

Another example is the definition of "Becomes Identifiable." In the draft MTA the word "directly" was added before "personally identified." This is ill-advised, as it makes the draft MTA narrower than POPIA, which can lead to inconsistency and confusion.

The NHREC's draft MTA is inadequate in relation to the transfer of data. We suggest that the conflation between data and biological material be avoided. These concepts should be dealt with distinctly, as they are governed by different disciplines in the law. Preferably, the envisioned MTA should avoid regulating the transfer of data altogether—rather, it should only regulate the transfer of

human biological materials to avoid misalignment with POPIA. In conjunction with such a pure MTA, parties must consider the use of a professionally drafted DTA that takes account of applicable legislation and is designed to lawfully manage the processing of data. Here there is a ready solution, namely the DTA template that was developed for South Africa's research community. It is fully aligned with POPIA and freely available (Swales et al., 2023b).

Next, we move the focus from the incorporeal to the corporeal—from data to biological material.

Is there sufficient guidance on biological material in the draft MTA?

It is interesting that the NHREC's draft MTA—similar to its predecessors—focuses only on human biological material, to the exclusion of other biological material that is important in health research, such as human pathogens. However, 'Human Biological Material' is defined sufficiently broadly in the draft MTA as to include pathogens. The definition reads as follows:

> 'Human Biological Material' means a biological sample or tissue from a person, living or deceased, including Deoxyribonucleic Acid (DNA), Ribonucleic Acid (RNA), blastomeres, polar bodies, cultured cells, embryos, gametes, progenitor stem cells, growth factors and blood specimens, biopsy tissue and any modifications or derivatives thereof

Consider the following scenario: When, during a pandemic, blood samples are drawn from infected persons and sent from one research institution to another, the blood sample would qualify as "Human Biological Material." However, such "Human Biological Material" would also contain a human pathogen, such as a bacteria or a virus. This is a matter of concern, as the draft MTA does not sufficiently cater for such a possibility. For a researcher there is a vast difference between a human biological material sample that contains a pathogen and a sample that does not, and the procedure for dealing with each is quite different.

Ultimately, the person that the revised SA MTA will govern will be the person who will need to organise the transfer of human biological material—which may include pathogens—and so they need to be aware of the legal and physical dangers relating to this. It may not be apparent that human biological material could be a weapon of mass destruction and yet that is exactly what it could be if the human biological material contains certain pathogens, such as the Ebola virus. This is acknowledged in the Non-Proliferation of Weapons of Mass Destruction Act (1993) and yet the draft MTA does not mention this important consideration. International standards, such as the WHO Manual on Laboratory Biosafety (World Health Organisation, 2020), the National Institutes of Health Shipping Policies and Procedures (National Institutes of Health, 2022), the International Air Transport Association's Infectious Substances Shipping Regulations (International Air Transport Association, 2023), and the Centers for Disease Control and Prevention's Guideline for Disinfection and Sterilization in Healthcare Facilities (Centers for Disease Control and Prevention, 2008) are examples of useful links. However, these are also omitted from the draft MTA leaving it up to scientists to source the relevant material on their own. In this respect, the draft MTA misses a vital opportunity to help and educate scientists by alerting them to the requirements that they need to comply with in order to transfer certain kinds of human biological material.

Apart from missing this opportunity to create awareness among scientists, the issue of pathogens being present in human biological material also opens up the issue of legal liability. At present, the draft MTA includes a provision that obliges the recipient to indemnify the provider of material from any liability, except insofar as the provider is required to be liable in law. The recipient is also required to maintain "adequate" insurance cover against liability to third parties. However, the draft MTA provides no assistance as to when the provider will be liable in terms of the law, nor does it require checks and balances to avert the harm that may or may not be covered by the "adequate" insurance. It is important to consider that the agreement may deal with the transfer of a biological weapon of mass destruction and so liability could be huge, possibly even worldwide. It is unlikely that this type of harm could be cured by any insurance cover and therefore greater effort should be invested in the eventual revised SA MTA to ensure that the harm does not occur.

The provider of the biological material should consider the infectious nature, volume and frequency of the transfer (among other factors) when considering the risk posed by the transfer of the human biological material. The identified risks would also influence the safeguards the provider would need to adopt. In this regard, the process to identify and deal with risks as set out in section 19(2) of POPIA could be considered to be a template for this purpose. The provider can, for example, create an appropriate risk matrix to be added as an annexure to the agreement. In addition, a right to audit compliance by either party should be included. The exercise of this right should be based on the risk profile of the other party.

We now proceed to the last research question, which pertains both to human biological material and associated data: the issue of ownership.

Why shy away from ownership?

The legal ownership of human biological material

The NHA is clear that the only way in which a research participant can provide a sample of his or her bodily material for research, such as tissue or blood, is by *donating* it to a research institution (section 63). Donation is a legal technical term for a nominate contract that entails the *transfer of ownership* from the donor to the donee (Mankowitz v Loewenthal, 1982, para. 765A; Thaldar and Shozi, 2021). Accordingly, when a research participant provides a sample of his or her bodily material for research, the *only legal way* in which this can transpire is for the research participant to *transfer ownership* to the research institution (DE v CE, 2020, para. 24; Thaldar and Shozi, 2021). That means that the research institution is the *owner* of the human biological material that it collects for research (Thaldar and Shozi, 2021).

Moreover, the Regulations regarding the General Control of Human Bodies, Tissue, Blood, Blood Products and Gametes (2012) also provides that a person who acquires human biological material in terms of the NHA acquires *exclusive rights* in such human biological material (Regulation 26). This is not only consistent

with the transfer of ownership to the research institution, but it also makes it clear that the transfer of ownership must be absolute and unqualified (National Health Act, 2003, s. 63; Regulations regarding the General Control of Human Bodies, Tissue, Blood, Blood Products and Gametes, 2012, reg. 24). In other words, the donor is not allowed to retain any rights whatsoever in the donated human biological material.

However, despite these clear statutory provisions, the NHREC decided to obfuscate and confuse the issue by introducing the concept of a "steward"—a concept that is not part of any branch of South African law that is relevant to health research (National Health Research Ethics Council, 2023). The NHREC defines "steward" as "a person or entity *entrusted* by the Participant to *safeguard and protect* the Material" (National Health Research Ethics Council, 2023) (Emphasis added). This is misaligned with ownership, for two reasons: First, an owner has the *right to destroy* the owned object—most certainly *not the duty* to "safeguard and protect" the owned object (Pope, et al., 2020). Second, the word "entrusted" points to a trust relationship between the research institution and the research participant with respect to the donated material (National Health Research Ethics Council, 2023). This is in conflict with the Regulations regarding the General Control of Human Bodies, Tissue, Blood, Blood Products and Gametes (2012) which provides that the research institution enjoys *exclusive* rights in the donated material (Regulation 26).

South Africa's NHA was enacted by the democratically elected representatives of the people of South Africa. It embraces ownership of human biological material by research (National Health Act, 2003, s. 63; Regulations regarding the General Control of Human Bodies, Tissue, Blood, Blood Products and Gametes, 2012, reg. 24). However, the NHREC is not respecting the democratic process. The NHREC is promoting ownership-denial. We suggest that the NHREC should take the law of South Africa more seriously.

The legal ownership of data

"Material" as defined in the MTA includes "associated data." The problematic nature of conflating these two very different kinds of object—human biological material and data—into one term was highlighted above. "Associated data" is defined as "the information associated with the Human Biological Material, including personal information, derived directly or indirectly prior and during the conduct of the research Project" (National Health Research Ethics Council, 2023). Accordingly, associated data includes all data—personal and non-personal—that are in any undefined way 'associated' with the human biological material (National Health Research Ethics Council, 2023). Superficially, the notion of a *data* steward seems to make sense, given that POPIA (2013) places various duties on a responsible party in relation to personal data. These statutory rights of the data subject qualify the common law ownership rights that a research institution may have in the personal data (Protection of Personal Information, 2013). However, the problem that lurks below the surface is that associated data as defined in the MTA are not limited to personal data but can also include de-identified data (Thaldar, et al., 2020). Consider that POPIA (2013) applies only to *personal* data, and ceases to apply when that same data is not personal or is de-identified to become non-personal data (section 3(1)). However, the NHREC's final draft would have a data steward safeguard and protect associated data even if it is not personal data. This makes no sense and is counter-productive. Moreover, the creation of a data steward does not consider the role of the Information Officer, who plays a crucial role in POPIA.

Using human genomic sequence data as an example, and applying the well-established requirement for private ownership in South African law, Thaldar et al. (2022) argue that a data instance—i.e., the computer file containing the data—is a digital object that is susceptible of private ownership in South African law. The authors further consider the rules concerning the acquisition of ownership in South African law, and suggest that the research institution that generates genomic sequence data is in the best position to acquire ownership in the data instances that it generates (Thaldar et al., 2022). In line with this conclusion, the DTA template embraces data ownership (Swales et al., 2023a; Swales et al., 2023b). Because data is a new kind of object and data ownership is not yet well established in the law, it is *essential* that data owners—South African research institutions—should clearly and explicitly record their ownership of the data that is being shared in their DTAs (Swales et al., 2023a).

Some may think that since data is incorporeal, ownership of data is an *intellectual* property right. However, this is mistaken. As analysed by Thaldar et al. (2022), common law ownership is not limited to corporeal objects. In fact, at least since the Second Century, when the Roman jurist Gaius wrote his Institutes, property law included incorporeal objects (Gaius, 1946). More recent examples of private ownership of incorporeal objects are, *inter alia*, digital money, digital books, and digital music (Nightingale v Devisme, 1770; Nissan South Africa Pty Ltd v Marnitz, 2006; S v Ndebele, 2012; Competition Commission v British American Tobacco South Africa Pty Ltd, 2009; S v De Vries, 2008; Curemed CC v Van Onselen, 2015). Millions of people buy music (as digital objects) on their smart phones using digital money (which is also a digital object). *Intellectual* property law, by contrast, is a more recent branch of the law, mostly found in statute and not in common law, and only applicable to specifically defined *kinds* of incorporeal objects, such as inventions and artistic creations (Copyright Act, 1978, s. 2; Patents Act, 1978, s. 3). It is however possible for intellectual property rights to overlap with common law property rights (Thaldar et al., 2022). Intellectual property law would typically not apply directly to data, but rather indirectly (Thaldar et al., 2022). This would be the case if, for example, data is used in an invention (patent law) or as part of a database (copyright law) (Copyright Act, 1978; Patents Act, 1978). However, the application of intellectual property law in no way overrides or supplants ownership in a data instance (Thaldar et al., 2022). Various rights can co-exist and qualify one another. For example, if one buys a book, one becomes the owner of the book, but the author still retains copyright in the content (Thaldar et al., 2022). The author's copyright qualifies the owner's rights in the sense that the book owner may not make copies of the book without the author's consent (Thaldar et al., 2022). The ways in which the rights emanating in various branches of South African law interact in the context of data are explored in detail by Thaldar et al. (2022).

The ethical case for owning the data that one generates

Not only is there a solid *legal* case for data ownership in the health research context, but there is also an *ethical* case, provided by John Locke's labour theory of property (Locke, 1963). In brief, this

entails that persons ought to acquire ownership in the fruits of their own labour (Locke, 1963). Applied to the generation of data, it is the research institution that *collects* the pheno-clinical data from research participants, and that *generates* genetic and genomic data by sequencing DNA isolated from samples donated by research participants. In other words, the research institution is the party that invests its *labour* into producing the data, and therefore ought to own such data. In health research, this typically requires significant investment in expensive equipment and highly trained human resources. Accordingly, it is ethically justified for the research institution to actively claim the fruits of its labour. Why does the NHREC shy away from supporting research institutions to claim what they are ethically entitled to?

Decolonial thinking about health research

In the colonial way of thinking about health research, global health research is conceptualised as an eternal cycle where Africa provides raw "genetic resources" to the Global North, while the Global North conducts value-added research on the "genetic resources" of Africa and owns the intellectual property in inventions such as new precision medicines, which are then sold to Africa for profit. Although this colonial way of thinking about health research is based on historical and (sometimes at least) current facts, it can become self-perpetuating when simply assumed and used as the basis for policy-making.

Allow us to explain: If policymakers make it more difficult for commercial research companies to acquire and control human biological samples and derivatives therefrom, such as DNA, cell-lines and data, the policymakers may think—because of the colonial paradigm in which they conceive health research—that they are protecting Africa from possible exploitation. However, what they may also be doing at the same time is to suppress the growth of the nascent biotechnology sector in Africa itself. In this way, the policy measure that are intended to protect Africa have the perverse effect of ossifying the colonial power structure and hence perpetuating the colonial paradigm of conceiving health research.

We therefore call for *decolonial thinking about health research*. Policymakers should reflect on their paradigms and how their resulting policy decisions can self-perpetuate the colonial power structures. Policymakers should actively strive to think anew about health research, and envision a (future) vibrant and sustainable African bio-economy, and then consider what policy choices would best assist the country to achieve that vision. To the extent that the NHREC has decided to become involved in health research policy development—revising the SA MTA is indeed policy development—the NHREC members should ensure that they are intimately familiar with South Africa's *Bio-economy Strategy* (Department of Science and Technology, 2013). If their answer is "but our mandate is ethics," then they should rethink why they have taken up the project of revising the SA MTA.

Firmly acknowledging research institutions' *data ownership* is not only ethical, but also core to developing a bio-economy that can compete globally in the Knowledge Economy. In the decolonised vision that we propose, South African biotech companies will act in lawful and ethically appropriate ways towards research participants, including respecting the research participants' privacy rights in the personal data that relate to them. In this way, the data owner can also have a "custodian" or "steward" function, by ensuring the safety of personal data. However, without clarity on ownership, being a mere "custodian" or "steward" is legally toothless (Thaldar, 2024). Furthermore, in the decolonised vision that we propose, South African biotech companies will build South Africa's bio-economy by generating a wealth of data. These data can be used for research in South Africa, and can be monetised by licencing access to such data in trusted research environments, or, where such data is de-identified, licencing access in less restrictive ways, such as data transfers. However, if a biotech company is merely the "steward" of the data that it generates, with uncertainty about ownership, there is *no legal basis* for any of these commercial actions (Thaldar, 2024). *Ownership* provides this essential legal basis (Thaldar, 2024). Without it, South Africa will be a *knowledge colony*.

Conclusion

At this point, it must be clear that we believe that the NHREC's draft revised version of the SA MTA is misdirected in several respects and the entire paradigm underlying the creation and content of the draft MTA needs to be considered anew. The NHREC needs to return to first principles to determine what they seek to achieve with a revised SA MTA and whether those are appropriate goals. In order to assist with this, we have the following four main recommendations on how the draft MTA could be reimagined:

Recommendation 1: make the use of the SA MTA voluntary, not mandatory

The draft MTA is not a mature document and it will take some time for it to reach a level of maturity where it is appropriate for it to be considered to be mandatory. The scientific and legal community should be encouraged to work together to progressively improve on the content of the draft MTA and to stress-test it against the actual lived experience of scientists who transfer human biological material. In addition, a mandatory document is inherently less flexible—and thus less able to be updated regularly—than a voluntary document.

Recommendation 2: data should be dealt with separately

The transfer of human biological material and the transfer of data are different disciplines, and different legal rules apply. Moreover, the current definition of "Associated Data" in the draft MTA merges the concept of personal and non-personal information in an unfortunate and unhelpful manner which contributes to confusion. If there are any associated data transferred alongside human biological material, it would be more appropriate to indicate that the DTA template developed for the South African research community (Swales et al., 2023b) as amended by the parties would govern such data.

Recommendation 3: enhanced liability and risk management provisions

Given that the draft MTA involves handling human biological materials—which can contain pathogens—the potential for harm is significant. The MTA should require providers of human biological material to assess the transfer risk by considering factors such as the infectious nature, volume, and frequency of the transfer. Recognising these risks will determine the necessary safeguards that can also be built into the eventual revised SA MTA, such as contractual warranties by the provider. Moreover, the eventual revised SA MTA should include a provision allowing either party to audit compliance, with the decision and scope of the audit informed by the other party's risk profile.

Recommendation 4: adopt a decolonial approach in the governance of health research

In the context of South Africa, it is crucial to ensure that health research does not perpetuate colonial legacies. Adopting a decolonial approach entails having a clear vision of a thriving bio-economy in South Africa—built not merely on being a raw material provider, but on adding value to such material—and strategically aiming for policy decisions that achieve this vision. Clarity on ownership of human biological material and associated data (primary and inferential) is crucial in order to have confidence and certainty in transactions entailing the transfer of these (corporeal and incorporeal) objects. This in turn is vital for building a thriving bio-economy in South Africa. Accordingly, policy instruments such as the eventual revised SA MTA should strive to empower local research institutions by clearly recognising their legal ownership of the material that they share.

Note that the draft MTA was not made public by the NHREC. Instead, the NHREC disseminated it to the health research ethics committees, who were given the opportunity to submit comments.

Author contributions

PE: Writing–review and editing, Writing–original draft, Conceptualization. LS: Writing–review and editing, Writing–original draft, Conceptualization. DB: Writing–review and editing, Writing–original draft, Conceptualization. BM: Writing–review and editing, Writing–original draft, Conceptualization. DT: Writing–review and editing, Writing–original draft, Funding acquisition, Conceptualization.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The authors acknowledge the support by the US National Institute of Mental Health and the US National Institutes of Health (award number U01MH127690).

Acknowledgments

The authors would like to thank Roasia Hazarilall for her assistance with the technical editing of this article, and acknowledge the use of ChatGPT4 from Open AI to improve the language and readability of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Author disclaimer

The content of this essay is solely the authors' responsibility and does not necessarily represent the official views of the US National Institute of Mental Health or the US National Institutes of Health.

References

Academy of Science of South Africa (2023). [Proposed] POPIA Code of conduct for research. Available at: https://www.assaf.org.za/wp-content/uploads/2023/04/ASSAf-POPIA-Code-of-Conduct-for-Research.pdf (Accessed October 6, 2023).

Blood Products and Gametes (2012). *Regulations regarding the general control of human bodies, tissue, blood, blood products and gametes. GN R180, GG 35099 of 2 march 2012*. Available at: https://www.gov.za/sites/default/files/gcis_document/201409/35099rg9699gon180.pdf (Accessed October 6, 2023).

Bronstein, V., and Nyachowe, D. T. (2023). Streamlining regulatory processes for health researchers: to what extent does POPIA apply? *S. Afr. Med. J.* 113 (8), 30–33. doi:10.7196/SAMJ.2023.v113i8.781

Burns, Y., and Burger-Smidt, A. (2023). *Protection of personal information: law and practice*. Johannesburg: LexisNexis.

Centers for Disease Control and Prevention (2008). Guideline for disinfection and sterilization in Healthcare Facilities. Available at: https://www.cdc.gov/infectioncontrol/guidelines/disinfection/index.html (Accessed October 6, 2023).

Competition Commission v British American Tobacco South Africa (Pty) Ltd (2009). *Competition commission v British American Tobacco South Africa (pty) Ltd [2009] ZACT 46*. Available at https://www.saflii.org/za/cases/ZACT/2009/46.html (Accessed October 6, 2023).

Copyright Act 98 (1978). *Copyright act 98 of 1978*. Available at: https://www.gov.za/sites/default/files/gcis_document/201504/act-98-1978.pdf (Accessed October 6, 2023).

Curemed CC v Van Onselen (2015). ZAGPPHC 176. Available at https://www.saflii.org/za/cases/ZAGPPHC/2015/176.html (Accessed October 6, 2023).

Department of Science and Technology (2013). The bio-economy Strategy. Available at: https://www.gov.za/sites/default/files/gcis_document/201409/bioeconomy-strategya.pdf (Accessed October 6, 2023).

De Stadler, E., Hattingh, I. L., Esselaar, P., and Boas, J. (2021). *Over-thinking the protection of personal information act*. Cape Town: Juta.

DE v CE (2020). 1 all SA 123 (WCC). Available at https://www.saflii.org/za/cases/ZAWCHC/2019/142.html (Accessed October 6, 2023).

Gaius (1946). Institutes, book II §12. Available at: http://www.law.harvard.edu/faculty/cdonahue/courses/rlaw/mats/c09.pdf (Accessed October 6, 2023).

International Air Transport Association. (2023). Infectious substances shipping regulations (ISSR). Available at: https://www.iata.org/en/publications/store/infectious-substances-shipping-regulations/ (Accessed 6 October 2023).

Locke, J. (1963). *Second treatise of government*. Indianapolis: Hackett Publishing.

Mankowitz v Loewenthal (1982). *Mankowitz v loewenthal (3) SA 758 (A)*.

Material Transfer Agreement for Human Biological Materials (2018). *Material transfer agreement for human biological materials, GN 719, GG 41781 of 20 july 2018*. Available at https://www.gov.za/sites/default/files/41781_gon719.pdf (Accessed October 6, 2023).

Minister of Justice and Constitutional Development v Southern African Litigation Centre (2016). *Minister of Justice and constitutional development v southern african litigation Centre 2016 (3) SA 317 (SCA)*. Available at https://www.saflii.org/za/cases/ZASCA/2016/17.pdf (Accessed October 6, 2023).

National Health Act 61 (2003). *National health act 61 of 2003*. Available at: https://www.gov.za/sites/default/files/gcis_document/201409/a61-03.pdf (Accessed October 6, 2023).

National Health Research Ethics Council (2023). *Revised material transfer agreement template*. Pretoria: Unpublished final draft.

National Institutes of Health (2022). 26101-42-F shipping Policies and procedures. Available at: https://policymanual.nih.gov/26101-42-f (Accessed October 6, 2023).

Nightingale v Devisme (1770). *Nightingale v Devisme (1770) 5 burr 2589 (98 ER 361)*.

Nissan South Africa (Pty) Ltd v Marnitz (2006). *Nissan South Africa (pty) Ltd v Marnitz [2006] 4 all SA 120 (SCA)*. Available at https://www.saflii.org/za/cases/ZASCA/2004/98.pdf (Accessed October 6, 2023).

Non-Proliferation of Weapons of Mass Destruction (1993). *Non-proliferation of weapons of mass destruction act 87 of 1993*. Available at: https://www.gov.za/sites/default/files/gcis_document/201504/act87of1993.pdf (Accessed October 6, 2023).

Patents Act 57 (1978). *Patents act 57 of 1978*. Available at: https://www.gov.za/sites/default/files/gcis_document/201504/act-57-1978.pdf (Accessed October 6, 2023).

Pope, A. (2020). *South African material transfer agreement 1.1*. Available at https://researchspace.ukzn.ac.za/handle/10413/19095 (Accessed October 6, 2023).

Pope, A., Du Plessis, E., and Badenhorst, P. J. (2020). *The principles of the law of property in South Africa: private law*. Cape Town: Oxford University Press.

Protection of Personal Information 2013 *Protection of personal information act 4 of 2013*. Available at: https://www.gov.za/sites/default/files/gcis_document/201409/3706726-11act4of2013popi.pdf (Accessed 6 October 2023).

Sasol Oil (Pty) Ltd v Metcalfe (2004). *Sasol Oil (pty) Ltd v Metcalfe 2004(5) SA 161 (W)*. Available at: https://www.saflii.org/za/cases/ZAGPHC/2004/25.pdf (Accessed October 6, 2023).

Steytler, M., and Thaldar, D. W. (2021). Public health emergency preparedness and response in South Africa: a review of recommendations for legal reform relating to data and biological sample sharing. *S. Afr. J. Bioeth. Law* 14 (3), 101–106. doi:10.7196/SAJBL.2021.v14i3.772

S v De Vries (2008). ZAWCHC 36. Available at: https://www.saflii.org/za/cases/ZAWCHC/2008/36.pdf (Accessed October 6, 2023).

S v Ndebele (2012). *S v Ndebele (3) SA 226 (GSJ)*. Available at: https://www.saflii.org/za/cases/ZAGPJHC/2011/41.pdf (Accessed October 6, 2023).

Swales, L., Botes, M., Donnelly, D.-L., and Thaldar, D. W. (2023a). Towards a data transfer agreement for the South African research community: the empowerment approach. *S. Afr. J. Bioeth. Law* 16 (1), 13–18. doi:10.7196/SAJBL.2023.v16i1.827

Swales, L., Ogendi, P., Botes, M., Townsend, B., Donnelly, D.-L., Abdulrauf, L., et al. (2023b). *A data transfer agreement template for South Africa*. doi:10.5281/zenodo.7537396

Thaldar, D. W. (2020). One material transfer agreement to rule them all? A call for revising South Africa's new standard material transfer agreement. *Humanit. Soc. Sci. Commun.* 7 (105), 105–5. doi:10.1057/s41599-020-00600-0

Thaldar, D. W. (2023). POPIA does indeed apply to health research: a response to Bronstein and Nyachowe. *S. Afr. Med. J.* 113, 4–5. doi:10.7196/samj.2023.v113i11.1345

Thaldar, D. W. (2024). The wisdom of claiming ownership of human genomic data: a cautionary tale for research institutions. *Dev. World Bioeth.* doi:10.1111/dewb.12443

Thaldar, D. W., Botes, M., and Nienaber, A. (2020). South Africa's new standard material transfer agreement: proposals for improvement and pointers for implementation. *BMC Med. Ethics* 21 (1), 85. doi:10.1186/s12910-020-00526-x

Thaldar, D. W., and Shozi, B. (2021). The legal status of human biological material used for research. *S. Afr. Law J.* 138 (4), 881–907. doi:10.47348/SALJ/v138/i4a9

Thaldar, D. W., Townsend, B. A., Donnelly, D.-L., Botes, M., Gooden, A., Van Harmelem, J., et al. (2022). The multidimensional legal nature of personal genomic sequence data: a South African perspective. *Front. Genet.* 13 (1). doi:10.3389/fgene.2022.997595

World Health Organisation (2020). *Laboratory biosafety manual*. 4. Available at: https://www.who.int/publications/i/item/9789240011311 (Accessed October 6, 2023).

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 03 April 2024
DOI 10.3389/fgene.2024.1272924

OPEN ACCESS

EDITED BY
Simisola Oluwatoyin Akintola,
University of Ibadan, Nigeria

REVIEWED BY
Abdulwasiu Yusuff,
Obafemi Awolowo University, Nigeria
Opeyemi.A. Gbadegesin,
University of Ibadan, Nigeria

*CORRESPONDENCE
Larisse Prinsen,
prinsenl@ufs.ac.za

RECEIVED 04 August 2023
ACCEPTED 18 March 2024
PUBLISHED 03 April 2024

CITATION
Prinsen L (2024), Introducing dynamic consent for improved trust and privacy in research involving human biological material and associated data in South Africa.
Front. Genet. 15:1272924.
doi: 10.3389/fgene.2024.1272924

Introducing dynamic consent for improved trust and privacy in research involving human biological material and associated data in South Africa

Larisse Prinsen*

Department of Public Law, University of the Free State, Bloemfontein, South Africa

Biomedical research using human biological material and data is essential for improving human health, but it requires the active participation of many human volunteers in addition to the distribution of data. As a result, it has raised numerous vexing questions related to trust, privacy and consent. Trust is essential in biomedical research as it relates directly to the willingness of participants to continue participating in research. Privacy and the protection of personal information also influence trust. Informed consent has proven to be insufficient as it cannot overcome the informational deficit between primary and unknown future uses of material and data and is therefore not fully informed and invalid. Broad consent is also problematic as it takes full control of samples and data flow from the research participant and inherently requires that a participant must trust that the researcher will use their material or data in a manner that they would find acceptable. This paper attempts to offer some insight into how these related issues can be overcome. It introduces dynamic consent as a consent model in research involving human biological material and its associated data. Dynamic consent is explained, as well as its claims of superiority in instances where future research is possible. It is also shown how dynamic consent contributes to better control of the samples and data by the research participant, and how trust may be improved by using this consent model. Dynamic consent's co-existence with and support of the South African Protection of Personal Information Act of 2013 is also assessed. The limitations of dynamic consent are also discussed.

KEYWORDS

biomedical research, dynamic consent, participant-centric initiative, protection of personal information, privacy, trust

1 Introduction

Biomedical research that makes use of human biological material and data is vital for increasing our understanding of biological and molecular mechanisms underlying illness and disease, testing the efficacy of new medications, medical devices and interventions, and for moving towards models of personalised medicine (Budin-Ljøsne et al., 2017).

Biomedical research promises significant societal benefits. However, in order to deliver on this promise, the research community requires the active participation of many human volunteers in addition to the distribution of data. Also, biomedical research requires the

continued collection of human biological material samples, health and outcome data from these samples, and also follow ups (Budin-Ljøsne et al., 2017).

Research participants are therefore essential partners in research endeavours since, through their voluntary participation, researchers have access to biological samples and data. These samples and data are vital to research and are available only by generous donation from participants (Horn et al., 2011). These participants may choose to participate in research for varied reasons including receiving the benefits of investigational medication, improving future healthcare practices or contributing to scientific knowledge. However, trust is always a key component of their participation (Horn et al., 2011) and a loss of participant or public trust can threaten continued research that uses human biological material and data (Williams et al., 2015). There are great concerns about the potential abuses and misuses of the collected data (Erlich et al., 2014), such as the invasion of privacy which may involve deeply personal issues. New forms of data and participant-led research are also challenging the traditional mechanisms of oversight and are raising questions about the ethics of partnership and collaboration between research participants and researchers (Tauginienė et al., 2021).

As mentioned above, in order to fulfil the promise of biomedical research, participant involvement is needed as well as analysis of large datasets containing the information of these participants. Sharing this information, however, requires protecting participants from potential harm (Erlich et al., 2014), such as exploitation and confidentiality or privacy breaches. Traditionally, regulatory frameworks have protected the rights and welfare of research participants as passive subjects, relying strongly on paternalistic views that research participants may not be able to assess correctly the risks and benefits involved in the research process or study (Tauginienė et al., 2021). However, because of strong human and consumer rights movements, the protection of human research participants has shifted and is now guided by informed consent or by Institutional Review Board or Research Ethics Committee procedures (Tauginienė et al., 2021). That said, the wide scope and nature of biomedical research that uses human biological material and data challenges a one-size-fits-all approach to obtaining consent. Furthermore, review mechanisms and both informed and broad consent have been shown to be insufficient consent models in biomedical research involving human participants (Prinsen, 2023).

Contemporary data protection models rely mainly on de-identification and de-identified data is largely allowed to flow freely. However, the flow of personal information is restricted and usually explicit consent from the participant allowing the dissemination of information or proof that the risk of re-identification has been minimised, is required (Erlich et al., 2014). De-identification and standard data security measures fall short in three important aspects (Erlich et al., 2014):

1. Standard data security controls may adequately protect data from unauthorised access but may be insufficient against abuses by a legitimate recipient of the information.
2. Advances in re-identification attacks have reduced the utility of de-identification techniques.
3. De-identification does not allow an individual control over data, which is a core element of privacy.

Considering these shortcomings and the limitations of de-identification, participants may, at best, be faced with cumbersome and poorly understood informed consent procedures that attempt to predict the future or, alternatively, ethically problematic broader consent processes (Prinsen, 2023). At worst, they may be given empty promises of anonymity. Researchers and the guardians of data, on the other hand, are then faced with manoeuvring between data utility and privacy (Erlich et al., 2014). This Sisyphean trap may, however, be overcome if trust and trust-enabling frameworks between participants and researchers are established. Such frameworks can be established by following the principles that transparency creates trust, that increased control enhances trust, and that reciprocity maintains trust (Erlich et al., 2014). These principles are discussed in more detail below.

Current data-management discussions frame the value of data against the risks to participants as a zero-sum, meaning that whatever is gained by one side is lost by the other. Erlich et al. (2014) have, however, suggested that a trust-based framework would be advantageous as both research participants and researchers can benefit from data sharing. Dynamic consent, which is discussed below, may be able to support a trust-based framework.

Given the exponential speed at which innovations in technology develop, researchers need flexibility in conducting their research in order to react quickly; thus, traditional approaches to the planning and conducting of biomedical research are unsatisfactory (Budin-Ljøsne et al., 2017). Dynamic consent is a strategy to involve participants, support the principle of informed consent and address the stationary aspect of consent by technological constructs such as communication platforms that establish a continuous two-way communication between researchers and participants (Tauginienė et al., 2021). It is seen as consent which is supported by the necessary information for participants to actively consent to their participation and also consent that is dynamic, regularly revisited, and not static or negotiated in a one-off process (Tauginienė et al., 2021). Dynamic consent may enhance participants' understanding of research and increase their scientific literacy and thereby positively affect their willingness to remain in a research project (Budin-Ljøsne et al., 2017).

This article therefore discusses trust, dynamic consent and privacy in order to introduce dynamic consent as a mechanism to benefit and improve trust and privacy in biomedical research involving human biological material and associated data.

2 Trust

2.1 What is trust?

Trust has been described as a mechanism that enables people to deal with situations of risk or uncertainty (Van der Geest et al., 2005). Various forms of trust have been identified, such as general trust and social trust (Van der Geest et al., 2005). However, in the context of biomedical research, two forms of trust are relevant and they depend on the person on whom trust is declared and on the circumstances. The first form of trust is known as "personal trust" and this is trust between two individuals. The second form of trust is trust directed at entities such as institutions, professional bodies or

governments, and is referred to as "institutional trust" or "impersonal trust" (Kerasidou, 2017). Regardless of the different forms trust may take, trust relationships share the common characteristics of an assumption of vulnerability by the trustor, an attitude of goodwill by the trustee, and voluntariness (Kerasidou, 2017).

In the context of biomedical research, the trust relationship may be personal (between the research participant and the researcher) or institutional and thus between the research participant and the research institution or between two or more institutions (Kerasidou, 2017).

A further concept in need of clarification is "trustworthiness". This relates to the person being trusted, the trustee, and the exhibition of characteristics which indicate that this person has goodwill towards the trustor. Trustworthiness is shown by the data controller and trust is given by the participant (Schuler Scott et al., 2019). For people to have trust, institutions must show trustworthiness (Schuler Scott et al., 2019). A person may be seen as trustworthy when they acknowledge the value of the trust vested in them and use that to rationally decide how to act (Kerasidou, 2017). Dynamic consent may be a mechanism of showing goodwill towards a research participant, since researchers are able to openly communicate with participants and show their intentions of acting in the greater good by achieving medical advances, for example.

2.2 Why do we need trust in biomedical research?

Some have argued that trust is not necessary in biomedical research since there are numerous instruments, laws, rules and authorities to regulate and oversee all research activities (Kerasidou, 2017). However, studies over the last few decades have found that trust plays an important role in the willingness of persons to participate in health research (Resnik, 2021) and a lack of trust may be seen as a great threat which can jeopardise consenting to participate in biomedical research (Kerasidou, 2017). As mentioned above, human participation in biomedical research is vital and thus trust cannot be regulated away.

Research participation always entails some level of risk to the participant, be it physical or informational, which means that in research involving human participants the participants make themselves vulnerable. Vulnerability and belief in the trustee's goodwill are the basis of the participant-researcher relationship (Kerasidou, 2017). By becoming a participant, a person surrenders their health and health-related information to the researchers and institutions. Systems of control and regulation do not fully compensate a person for putting themselves in a vulnerable position. This is where trust becomes relevant in that the participants have to trust that the researcher has an attitude of goodwill towards them. In research, this goodwill means that the researcher acknowledges the vulnerability of the participant and takes it into account when considering how to design, conduct and implement their research (Kerasidou, 2017).

Viewing trust as the cornerstone of the participant-researcher relationship suggests a trust-based relationship. As mentioned above, the data utility *versus* privacy challenge may be overcome by establishing a trust-based framework. The first principle in doing so holds that transparency creates trust. This means that transparency between the parties involved is key and that research participants must be informed of not only the intended but also the actual use of data (Erlich et al., 2014). This is a common feature of information privacy statutes and may be seen in the provisions of the Protection of Personal Information Act of 2013, which is discussed in more detail below. It further demonstrates that trust and privacy are interconnected and cannot be separated. The second principle of a trust-based framework holds that increased control enhances trust. Keeping in mind the uncertainties involved in biomedical research, it is virtually impossible to make fully informed decisions about future data uses and risks, but this issue may be overcome when a research participant is given control over the future use of data.

While clear communication about possible risks is critical in ensuring informed consent, current informed consent practices require participants to make one-off decisions about future data sharing with unknown risks. Broad or blanket consent practices are also problematic as the participant has to surrender their control to another person and trust that they will act with goodwill towards the participant. These consent practices further do not accommodate changing privacy preferences over time (Erlich et al., 2014). Dynamic consent is, however, able to do so—as will be explained below. The third principle of a trust-based framework holds that reciprocity maintains trust. This means that mechanisms whereby participants reward researchers who act appropriately by continuing to participate, while punishing those who violate their trust by withdrawing their participation, may provide valuable incentives for win-win behaviour (Erlich et al., 2014) which enables ongoing research. Building on these principles, a participant-centric bilateral consent framework is suggested and it is further suggested that dynamic consent offers such a framework.

A bilateral consent framework, at its core, enables participants to have dynamic control over access to their data. In current consent frameworks the participant delegates control to the researcher who, upon completion of a study, may delegate further use decisions to an Internal Review Board or to Research Ethics Committees. In a bilateral consent framework, the data control remains primarily with the participant. The researcher may then approach the participant with information about secondary uses of the data and the participant may choose to (re-)consent or to revoke and withdraw consent (Erlich et al., 2014). A bilateral consent framework thus engages the participant by making it possible for the researcher to solicit participant data, while at the same time empowering the participant to change their preferences. This framework therefore emphasises reciprocity and agency, envisions data sharing, and sees consent as a shared process which requires iteration and feedback (Erlich et al., 2014). Dynamic consent is a bilateral consent framework.

Trust is important and is influenced by the sense of control and by privacy concerns (Van der Geest et al., 2005). Creating trust gives participants control, and requesting consent is an essential condition for solving privacy issues (Van der Geest et al., 2005).

2.3 Consent and trust

The history of biomedical research is marred by scandals such as the experiments conducted by Nazi doctors, the Tuskegee Syphilis

Study and the Hwang Woo-suk scandal—to name a few. These incidents have undermined public trust in biomedical research and in an attempt to restore this trust, regulatory instruments, laws and rules have been created, such as The Nuremberg Code of 1948, the 1964 Declaration of Helsinki or the Belmont Report of 1979. In addition, Institutional Review Boards and Research Ethics Committees were created to increase accountability and transparency (Kerasidou, 2017) and consent became a recognised and indispensable requirement for conducting any medical or scientific research involving human participants.

The role of informed consent is to allow participants to make decisions and to safeguard trust in research endeavours (Dankar et al., 2020). Valid informed consent for medical research participation has traditionally required mental capacity to make a reasonable decision, voluntariness and the absence of any form of coercion or undue influence, the provision of information necessary to make a decision, understanding of the given information, and the expression of the decision. When these elements are seen as a whole, it suggests that a consenting person is an autonomous, rational agent making an informed and voluntary decision in line with their own values (Resnik, 2021).

Research has, however, shown that decisions are often not autonomous as the full provision of information has not occurred, the consenting person does not understand the information or other factors may be present which interfere with comprehension (Resnik, 2021). Trust may be seen as a crucial element which compensates for the lack of understanding as it may assure the consenting person that they are not being manipulated, exploited or deceived (Resnik, 2021).

The relationship between trust and consent is reciprocal in that trust may develop during the consent process as the parties become acquainted with one another and form a relationship, but a degree of trust must be presupposed before the consent process begins. Research participants are more likely to participate in research if they have trust in the researchers, the research organisation and the research project itself (Resnik, 2021). It has also been argued that consent and the requirement therefore is a method of trust building (Kerasidou, 2017).

As mentioned above, participants make themselves vulnerable when participating in research. At the start of participation, the participant must give consent and consent lays down the conditions of the relationship between the participant and the researcher. Prospective participants are informed of the risks and possible benefits of the research project, what is required of the participant, and what may be expected from the research, researcher and the research institution. Participants are then given the opportunity to voluntarily make a decision about their willingness to participate or not. The provision of information about the risks and benefits does not, however, absolve researchers of their duty to minimise these risks, to ensure fair distribution of the benefits or to protect the welfare of the participants. In order for participants to trust researchers, they need to believe that the researchers will conduct the research with an attitude of goodwill towards them (Kerasidou, 2017). This means that when a participant gives consent to participation, they do not confirm trust, but rather presuppose it, or differently stated, consent consolidates trust.

Human beings are social animals and trust may come naturally to some, but once trust has been broken it can be difficult to restore. Trust is a form of social capital and, therefore, activities which promote honest, open and respectful communication and dialogue may help in building, maintaining and restoring trust in research (Resnik, 2021). This also means that building, maintaining and restoring trust means doing the same in regards to trustworthiness (Kerasidou, 2017).

Trust may be built through numerous mechanisms, such as developing relationships with participants, demonstrating a track record of accountability, or showing concern for the best interests or goodwill of others. Building trust in biomedical research is essential and becomes more challenging the more distant the research becomes from the participant (Horn et al., 2011). This is often the case when taking into account the unknown secondary and future uses of human biological material and data. Although consent processes alone cannot build trust, it may be considered a minimum effort in terms of doing so (Horn et al., 2011). Consent is essential in nurturing trust but it has been described as the fruit of the tree, rather than its roots (Resnik, 2021).

Trust, control and privacy are strongly related and, accordingly, enabling participants to exert control over their own information may increase trust and thus reduce privacy concerns (Van der Geest et al., 2005). Consent is a mechanism that allows participants to exert control. Control and information are considered central to consent and give rise to two effects. First, participants are informed of relevant matters such as risks or harms and benefits which may flow from participation in research. Second, consent enables participants to exert control and retain responsibility over what they feel is sensitive information. Where consent is used as a process involving participants who control their own personal data, a valuable strategy for dealing with trust and privacy concerns may be found (Van der Geest et al., 2005).

Practices such as informed consent and the use of Institutional Review Boards and Research Ethics Committees have been useful in ensuring the protection of future research while fostering public trust in biomedical research (Kerasidou, 2017). Informed consent, however, is insufficient in biomedical research and broad consent, which was developed to overcome informed consent's shortcomings, is also problematic as it removes control from the participant (Prinsen, 2023). Dynamic consent may therefore offer a valuable solution to the issue of consent in biomedical research.

3 Dynamic consent

New ways of conducting research have given rise to new ethical norms, practices and standards. The status of participants and their level of involvement in research has been a particularly prevalent new question, along with concerns about the appropriate format of consent. A clear shift has also taken place towards more participant-centric initiatives which place the participant in a partnership relationship with the researcher in both the decision-making process and the research study (Steinbekk et al., 2013). Research participants want to be involved, valued and engaged in research and many have shown the desire to be more actively involved in medical research and their health. They further desire transparency and openness and do not wish to be surprised about the use of their donated materials and data (Horn et al., 2011). Participants want to trust that researchers are conducting research with their best

interests and goodwill in mind. Various participant-centric research initiatives use emerging technologies to engage participants in new ways and these systems or frameworks enable participants to exercise as much or as little control over their material and information as they prefer. Dynamic consent is one such technology-backed, participant-centric initiative (Horn et al., 2011).

The dynamic consent framework is founded on the work of an expert group that studies legal, social, technical and compliance aspects of consent (Prinsen, 2017) and has the potential to radically alter the nature of consent in research (Kaye et al., 2015) as it supports the flow of new knowledge between the laboratory, clinic, researcher and participant (Mason and O'Neil, 2007). Dynamic consent, with its two underlying concepts of allowing the revocation of consent for data use, and engaging in communication about data use (Schuler Scott et al., 2019), was created to address problems with one-off consent, to develop trust and improve participant recruitment and participation, and is founded on the principles of revocation and engagement (Schuler Scott et al., 2019).

In the past, and sometimes still today, consent was or is obtained in paper format which was then filed away when a person decided to participate in research. Dynamic consent uses electronic systems which enable a participant to keep track of their data, including records of donated human biological material and what this material has been used for. It further allows a participant to monitor and update consent choices over time. For example, a participant may wish to allow the use of their material or data in a new research project or may wish to limit the research which may be conducted using such a sample or data. Accordingly, this model of consent allows control over past and currently donated materials in addition to any future material to be donated (Prinsen, 2017).

3.1 The reason for dynamic consent

The requirement that consent be obtained by researchers prior to initiating a proposed study is a fundamental principle of research ethics and law. It has also repeatedly been shown to underpin respect for persons and their autonomy. Consent has become a method of recording individual involvement, and for determining the scope of what is included under consented to activities. Therefore, it may be seen as the formalisation of an implicit social contract between the public and researchers (Kaye et al., 2015). New forms of biomedical research, however, challenge the meaning of informed consent and question the current processes of engaging with participants. The uncertain scope of consent is especially controversial and, in an attempt to address this, broad consent has been suggested as an understandably practical solution. However, for various reasons, a broad consent approach is insufficient for meeting the requirements of meaningful informed consent (Prinsen, 2023).

Unlike traditional research, biomedical research does not follow a single experimental procedure in which participants are asked to participate. Rather, it is a request to participate in an ongoing, continuous inquiry with multiple investigations and methods that involve unknown risks, and it is suggested that new research trends demand new models of consent (Kaye et al., 2012). Differently stated, the consent procedure must also be an ongoing one.

Ethically, it is necessary to enable a participant who has given consent under a set of circumstances to review this consent as new research possibilities, using the same material samples or data, emerge. Also, the possibility exists that research participants may benefit clinically from updated information about their data and samples (Kaye et al., 2012). Legally, by providing a more comprehensive form of consent more bases are covered and thus the legal liabilities which could arise because of the absence of consent are reduced, since the scope of consent is more clearly defined.

As biomedical research changes, so too must the role of the research participant change and evolve (Kaye et al., 2015). Furthermore, individual autonomy is not static and involves changing choices, opinions and preferences. Consent is a mechanism whereby participants' rights may be protected in research and consent decisions must have the ability to change over time since people are prone to changing their minds (Schuler Scott et al., 2019). Research participants are no longer passive human subjects, but are active and interested participants and consent must now be seen as a process of ongoing interaction between a researcher and a participant.

O'Neil (2006) stated that true consent is reliant on access to extendable information, the concept of rescindable consent, and the right to veto certain activities. Respect for a participant and their autonomy therefore means that participants must be given as many choices and as much control over information, their material and data as possible (Kaye et al., 2015).

Theoretically, dynamic consent benefits both the researcher and the participant, since the participant is given information related to their material, there is transparency regarding their information usage and sharing which enhances trust, and there is the option to revoke consent. This promotes the relationship between research participant and researcher. The researcher benefits from dynamic consent as they gain a business edge by setting best practice, and the relationship with the participant is flexible–meaning that newer and more refined usage may be allowed (Prinsen, 2017).

3.1.1 Dynamic consent as a participant-centric initiative

A participant-centric initiative may be understood as meaning "a tool, program and project that empowers a participant to engage in research processes using IT" (Prinsen, 2017). By making use of an IT interface, it provides an ongoing, continuous interactive means of obtaining consent and maintaining communication between participants and researchers (Kaye et al., 2012). The key characteristics of a participant-centric initiative are that it is based on respect, promotes the empowerment of participants, and is focused on participation. The researcher and the participant are central in decision-making and are equal partners in the research process (Kaye et al., 2012). Participant-centric initiatives therefore greatly emphasise autonomy.

Participant-centric initiatives exhibit four functions. First, they serve a "matchmaking" function which enables the recruitment of research participants. Second, they provide a "direct-to-consumer" service by offering participants genetic testing and analyses and give them the opportunity to participate in research projects (Tamir, 2010). Third, "dynamic control" aids ongoing interaction between the researcher and the participant. Fourth, the initiatives have a "citizen science function" which engages participants in facilitating, designing and executing research projects (Kaye et al., 2012).

Using participant-centric initiatives may greatly benefit research governance by ensuring adherence to basic ethical and legal principles, improving recruitment methods, and maximising participant retention. They may also minimise costs, enhance knowledge and understanding of the research process, and encourage and sustain public trust through greater involvement, accountability and transparency. Participant-centric initiatives are able to achieve these benefits by streamlining the consent process, decreasing the need for de-identified data, facilitating participant recruitment, facilitating participant retention, promoting the delivery of better quality healthcare, improving the quality of research, and sustaining public trust and confidence in research since greater involvement in research has a dual effect. It firstly improves knowledge of the research process and secondly ensures transparency and accountability on the part of the researcher. Research may then be conducted at a higher standard and will be in tune with societal expectations and concerns, which will result in enhanced public confidence and trust (Kaye et al., 2012).

3.2 The meaning of dynamic consent

As suggested by the name, dynamic consent is dynamic in that given consent is changeable and adaptable. This idea, however, has a narrow and a broad meaning:

- In a narrow sense, it is a personalised communication platform which enables greater participant engagement in research by enabling an interactive relationship between researchers and participants (Steinbekk et al., 2013). Researchers must foster a relationship of confidence, understanding and trust to establish true insight into what is at stake in the course of research. Dynamic consent may be defined as a new approach to engaging persons in the use of their information and material.
- In a broader sense, it is also an interactive and personalised platform which enables participants to engage in research as much or as little as they prefer and to amend their consent decisions in real time (Kaye et al., 2015).

Dynamic consent is seen as dynamic in that it enables the giving and revocation of consent to the use of materials or data, it centralises transactions and interactions, allows participants to be approached for different projects or feedback, and allows for consent processes to be modified over time (Schuler Scott et al., 2019). At its core, dynamic consent is a mechanism of enabling communication between participant and researcher, and offers the participant the opportunity to be continually informed and in control of their information and material (Wee et al., 2013).

Dynamic consent as a participant-centric initiative places research participants at the centre of the decision-making process by providing an interactive IT interface. It is a dynamic approach as it allows interaction over time, enables renewal of consent to new projects, enables consent to be amended in real time as circumstances change, and gives participants the confidence that their amendments will have an effect (Kaye et al., 2015). When a person initially agrees to any processing of their personal material or data, they may do so without fully understanding the implications of what they are consenting to. After some time, they may wish to review or revoke the initial consent in order to create a new agreement which is more in line with their preferences. With dynamic consent, a participant may control the use and flow of their data and material and change their consent about what is permitted and what is not.

Dynamic consent has certain characteristic features. The first is that it comprises different consents. It is not locked in time at the onset of a project and, depending on the nature of the research project, participants are able to consent to a wide range of uses of their material and data, or they may choose to be approached on a case-by-case basis or may create varying preferences for different research types. These preferences may be opt-in or opt-out and the participant is therefore able to tailor his profile to receive certain information at certain times (Kaye et al., 2015).

The second feature of dynamic consent pertains to its tailored aspect. Since a dynamic consent interface acts as a personalised communication forum, a source of information and a platform on which consent may be modified, all aspects of the interface may be tailored to the preference of the participant. Persons may choose how and when they are to be contacted and what information they wish to receive (Williams et al., 2015)—which emphasises an improvement in control exerted by participants.

The third feature of dynamic consent involves the customisation of research needs. This consent model clearly incorporates a flexible design able to accommodate researchers and participants. All aspects of the interface may be tailored to the proposed project and in this manner extend the interaction between the parties (Kaye et al., 2015).

Dynamic consent improves trust in how data is used, as control of the data is passed to the participant (Schuler Scott et al., 2019). From the features discussed above and the repetitive emphasis of tailoring to the preferences of the participant, it should become obvious why and how dynamic consent may improve trust in biomedical research. Where the participant can tailor their experience, they are the controllers thereof–meaning that they are able to trust the experience.

3.3 Dynamic Consent's benefits and claims of superiority

In understanding, recognising and supporting biomedical research as a partnership between researchers and participants, dynamic consent enables research while also improving the research experience. Dynamic consent therefore offers participants engagement in the process, better respects their autonomy, and also offers meaningful consent. Researchers derive benefit from engaged participants, streamlined participant recruitment, and improved trust. Legally, dynamic consent is valuable as it offers better protection by eliminating ambiguity and vagueness. Ethically, dynamic consent may also be seen as beneficial since it allows for the true expression of autonomy (Prinsen, 2017).

In addition to these benefits and improved trust and privacy as discussed in this article, dynamic consent is also beneficial in providing for the facilitation of efficient re-contact, conformity to the highest legal standards, fine-grained withdrawal mechanisms,

the enabling of better communication, improved scientific literacy, and transparency and risk management.

3.3.1 Facilitation of efficient re-contact

Re-contact is often impractical. Dynamic consent offers a method of easy re-contact with participants which grants them accessible information and allows the participant to make an informed decision (Kaye et al., 2015). Maintained contact with participants assists researchers in addressing numerous ethical and legal issues which may arise in unforeseen circumstances. Dynamic consent has also been touted as being able to overcome other ethical challenges encountered in biomedical research (Tauginienė et al., 2021). However, a full discussion thereof falls outside the scope of this article.

3.3.2 Conformity to the highest legal standards

Freely given consent is universally regarded as a requirement of biomedical research as seen in legal and regulatory documents across the globe. Dynamic consent provides a flexible and responsive mechanism of addressing changing legal and ethical requirements (Tauginienė et al., 2021). It may even provide better protection of autonomy than current international standards (Kaye et al., 2015). It is in this flexibility that dynamic consent can accommodate the slightest change in circumstances associated with the consent and therefore the fine-grained functioning of dynamic consent is also seen as beneficial. This also benefits trust as legal compliance and sensitivity may lead to participants feeling better protected.

3.3.3 Fine grained withdrawal

Research participants have the right to withdraw their consent, material or data by requesting that it not be made available for certain secondary or future research projects or even that it be destroyed. Dynamic consent enables a more nuanced choice by offering more information and preference-related options to a participant and, in doing so, excludes the zero-sum "all or nothing" mode of withdrawal which is often found in withdrawal circumstances (Kaye et al., 2015). This not only improves retention of participants but also trust.

3.3.4 Enablement of better communication

Traditionally, consent procedures involve an initial engagement session with the participant at the start of the research project but rarely provide for mechanisms of continued communication (Mascalzoni et al., 2009). In addition, research findings are seldom conveyed to the participants. Dynamic consent uses an online personalised consent and communication platform in order to facilitate the consent process and two-way, ongoing communication between researchers and participants (Budin-Ljøsne et al., 2017). Dynamic consent enables the return of findings according to the participant's selected preferences. It also creates a means whereby broader engagement may be nurtured, which extends beyond an information sheet. This adds value to the research study (Kaye et al., 2015).

3.3.5 Improved scientific literacy

By implementing a user-friendly and accessible platform, dynamic consent gives participants additional opportunities to gain knowledge and understanding of the information provided in their own time. Participants are granted time for reflection and consideration and they are thus empowered to control the type and amount of information they receive and when they wish to receive it. This may lead to a more realistic understanding of research as an interactive and long-term process, may improve participant confidence by transparency and accountability which leads to improved trust, and may support the development of appropriate expectations of what research may achieve (Kaye et al., 2015).

3.3.6 Improved transparency and risk management

Transparency and accountability may be improved by dynamic consent as the research process, the use of material or data and consent may be traced throughout all the studies. This therefore provides for operational control over risk (Kaye et al., 2015). Participants may also be contacted about controversial issues, such as the protection of personal information, and in this manner trust is safeguarded.

In addition to these benefits, dynamic consent presents six claims of its superiority over other forms of consent (Steinbekk et al., 2013):

1. Dynamic consent offers greater respect for participant autonomy than other consent models as it is better able to meet the specifications of autonomy embedded in specific or informed consent requirements. Dynamic consent enables participants to exercise their autonomy by providing consent to new types of research, in real time, as opposed to once-off consent (Kanellopoulou et al., 2011). This means that since participant preferences are used as the point of departure when establishing potential uses of their material and data, participants are given the opportunity to consent to primary and secondary uses of their material and data.
2. Participants are kept better informed by dynamic consent. The ability to keep participants informed is seen as essential in all research consent processes, and dynamic consent is better suited to fulfil the ideals of disseminating detailed information (Whitley et al., 2012). Additional information may appeal to participants who wish to have control or who are uncertain about the specifics of what they are participating in.
3. Dynamic consent also claims to be a solution to other biomedical research-related challenges such as participant recruitment and retention (Budin-Ljøsne et al., 2017). The dynamic consent model encourages participation in biomedical research. Since trust is created by transparency and accountability, proponents of dynamic consent argue that it will positively affect not only participant recruitment but also retention—and this ultimately results in sustainable biomedical research (Kaye et al., 2012). Also, dynamic consent addresses any criticism that participants are regarded as a mere supply of biological material as the participant becomes an active partner in research (Saha and Hurlbut, 2011). Furthermore, public insight and knowledge are increased by dynamic consent. However, it may be argued that possible participants may be deterred by being confronted with, or even intimidated by, all the details and complexities of biomedical research and then being asked for consent over

and over again. As a result, dynamic consent may be described as a two-edged sword in the context of participant recruitment. It could increase trust since participants are given different choices and trust is raised by transparency, and the participant's sense of control is also increased. In addition, it seems that reciprocity is amplified since dynamic consent accommodates the return of information. On the other hand, participants may then have exaggerated hopes and expectations of what a research project could yield. When these hopes are not realised, trust may be breached and recruitment may decrease.

4. Dynamic consent transfers control to the participant. Concerns about the lack of participant control over both the research and the results are addressed (Wagstaff, 2011). This may be the strongest argument in favour of dynamic consent and may potentially lead to new participant rights (Whitley, 2009).
5. Ethical responsibility is transferred from Research Ethics Committees to participants. This would result in a move towards an open and democratic scientific process which ensures socially robust knowledge. Since new consent must be provided for new research projects, the need for Ethics Review Boards is eliminated or decreased (Kaye, 2012).
6. Dynamic consent enables the return of results and incidental findings in an easy, user-friendly and tailored manner. Proponents of dynamic consent argue that the return of results and findings is necessary as it respects participant autonomy as well as reciprocity and beneficence (Steinbekk et al., 2013).

3.4 How dynamic consent works

Dynamic consent as a consent platform is achieved by using technical solutions, compliance services and legal accountability (Prinsen, 2017). Dynamic consent therefore entails a new digital system which allows participants to give consent electronically and by offering dynamic consent along with online services participants can monitor the possible uses of their material samples and personal information or data and make decisions about how these may be used in future (Prinsen, 2017).

Dynamic consent works in a reciprocal fashion where the research participant or data subject is approached to participate in a research project and is provided with relevant information about the project by the researcher or data controller. The participant then consents to the project. During the course of and after the completion of a project, the researcher provides feedback on the findings of the project or notifies the participant of new enquiries or uses of their donated materials or data and again provides the relevant information to the participant. The participant is then given the opportunity to change their preferences which may mean re-consenting to the secondary uses, revoking their consent, or withdrawing from the study. Once these changed preferences are received by the researcher, they can adjust their actions accordingly. This process is continual and facilitates keeping the participant up-to-date, which translates into accountability and thus trust, control and even improved privacy (Prinsen, 2023).

Dynamic consent uses web-based technology features to overcome the problem of the lack of specific "real-time" information about individual research projects (Whitley et al., 2012). The platform must, however, be able to provide a flexible mechanism which provides different degrees of control to participants based on their personal preferences (Kaye et al., 2012). Dynamic consent endorses a process which emphasises continual re-contact with participants by providing real-time information on research projects, and allows for easy user-friendly revocation of any previously given consent (Steinbekk et al., 2013).

4 Dynamic consent and trust

Behavioural psychology has found that empowering participants establishes trust and approval, which results in greater participation (Dankar et al., 2020). Viewing participants as partners, as envisioned by dynamic consent, empowers participants and therefore improves trust. Furthermore, trust is fundamental in the successful use of data and dynamic consent may provide a flexible, transparent and user-friendly manner of providing information and maintaining trust (Williams et al., 2015).

Using consent as the basis of sharing data addresses the limitations associated with de-identification and anonymisation, while still respecting the autonomy of participants. Having a dynamic form of consent may allow participants to more readily provide or withdraw consent over time, while simultaneously providing information to the participants about how their material and data are used (Williams et al., 2015). Ongoing communication with researchers has also improved trust (Chen et al., 2020). Dynamic consent, a participant-centric consent approach, uses technology to allow ongoing engagement of research participants' consent preferences as well as continual communication. Participants are also able to track and audit the use of their information, change their privacy settings and choose how and if they wish to be contacted. Dynamic consent thus enhances trust by giving the participants control over data flow (Williams et al., 2015).

Dynamic consent further fosters trust by providing mechanisms of accountability and transparency about the use of information and data as well as the sharing thereof (Prinsen, 2017). The improved scientific literacy offered by dynamic consent may also improve trust as it removes the fear of the unknown, so giving the participant confidence and, in turn, promoting better trust.

Dynamic consent also claims to allow for better adherence to regulatory systems (Kaye et al., 2015) and this sense of improved lawfulness may improve trust. Where participants feel that authorities bigger than themselves have approved an action, they may be more likely to believe that certain checks and balances are in place and that they and their goodwill are protected. Regulatory regimes may be considered as a vetting process, which allows the participant to better trust in the research process.

5 Dynamic consent, privacy and the protection of personal information

Trust, control and privacy are inextricably connected. If participants are able to have control over the protection of their

privacy, they will be more likely to trust in the research endeavour and ultimately consent to participate in research. Attention must, however, be given to privacy and the benefit of using dynamic consent is highlighted.

5.1 Privacy

Privacy concerns may influence trust and, normally, participants wish to know who has access to their information, for what purpose, and who will make decisions in an ongoing manner since consent to participation is obtained prior to the start of participation. Participants need to be able to trust that their data is being used in accordance with their preferences (Horn et al., 2011).

Support for the sharing of data is often founded in privacy protecting safeguards, such as those envisioned by the Protection of Personal Information Act, which is discussed below. Where concerns have been raised in this regard, these were related to who the recipient of the information will be, anonymity and the type of information being shared, with indications of participants being more concerned the more personal the information is (Williams et al., 2015).

Current participant protection frameworks use de-identification of human biological materials and data, but according to Horn, Edwards and Terry (2011), participants are willing to make trade-offs for privacy if it means that they are able to stay connected to a study in which they are participating. De-identification severs the connection between the participant and the research and some participants may wish to stay connected by learning about the results of a study, for example,. The complete de-identification of material or data may also be harmful to participants as it hinders the return of incidental findings where appropriate, prevents researchers from obtaining follow-up information, and limits the participant's ability to continually direct information (Horn et al., 2011). Dynamic consent may be indicated in these instances as it is premised on keeping participants connected and engaged to the research project and may thus offer a fair trade-off. Privacy enables the opportunity to negotiate how others access or use their information and the attitude towards these others is influenced by the level of trust in them (Schuler Scott et al., 2019). Furthermore, dynamic consent enables the return of findings and allows the researcher to obtain follow-up information.

Dynamic consent may also be useful in protecting participants from threats to their privacy as they are empowered to largely control access to their information. Four types of threats to privacy may be identified (Van der Geest et al., 2005): improper acquisition of information; improper use of information; improper storage and control of personal information; and privacy invasion. Dynamic consent and the use of participant preferences may be able to address all of these threats.

Privacy is often associated with ideas of self-identity, which are related to autonomy, and people have been shown to want control over their personal information and the decisions they may make about their data (Schuler Scott et al., 2019). Dynamic consent rests on participant engagement and facilitation of data, participation, and revocation of consent when needed. It protects tangible privacy interests rather than protecting privacy as an abstract concept (Schuler Scott et al., 2019). Claiming that dynamic consent is a privacy control means that it may be used by participants to manage how their information is shared and to what extent it may be shared even further (Schuler Scott et al., 2019).

5.2 Protection of personal information

The Protection of Personal Information Act of 2013 (POPIA) is the most recent addition to the collection of data protection Acts in South Africa and its purpose is to give effect to the constitutional right to privacy, to regulate how personal information may be processed, provide for rights and remedies to protect personal information, and to establish voluntary and compulsory measures to ensure respect, promotion, enforcement and fulfilment of the right to privacy (POPIA, s 2). It is suggested here that dynamic consent and the protection of personal information and privacy as provided for by POPIA are symbiotic and that dynamic consent may help overcome obstacles faced by health research because of the provisions of the Act.

POPIA requires specific consent for the processing of "personal information", which is defined as (POPIA, s 1): information relating to an identifiable, living, natural person, and where it is applicable, an identifiable, existing juristic person, including, but not limited to—

(a) information relating to the race, gender, sex, pregnancy, marital status, national, ethnic or social origin, colour, sexual orientation, age, physical or mental health, wellbeing, disability, religion, conscience, belief, culture, language and birth of the person;
(b) information relating to the education or the medical, financial, criminal or employment history of the person;
(c) any identifying number, symbol, e-mail address, physical address, telephone number, location information, online identifier or other particular assignment to the person;
(d) the biometric information of the person;
(e) the personal opinions, views or preferences of the person;
(f) correspondence sent by the person that is implicitly or explicitly of a private or confidential nature or further correspondence that would reveal the contents of the original correspondence;
(g) the views or opinions of another individual about the person; and
(h) the name of the person if it appears with other personal information relating to the person or if the disclosure of the name itself would reveal information about the person.

Some other definitions also need clarification in order to understand the discussion below. These definitions are (POPIA s 1):

1. Biometrics: a technique of personal identification that is based on physical, physiological or behavioural characterisation including blood typing, fingerprinting, DNA analysis, retinal scanning and voice recognition;
2. Consent: any voluntary, specific and informed expression of will in terms of which permission is given for the processing of personal information. It is suggested that dynamic consent meets all these requirements;

3. Data subject: the person to whom personal information relates. In the context of biomedical research, this is the participant;
4. Processing: any operation or activity or any set of operations, whether or not by automatic means, concerning personal information, including— (a) the collection, receipt, recording, organisation, collation, storage, updating or modification, retrieval, alteration, consultation or use; (b) dissemination by means of transmission, distribution or making available in any other form; or (c) merging, linking, as well as restriction, degradation, erasure or destruction of information;
5. Record: any recorded information— (a) regardless of form or medium, including any of the following: (i) writing on any material; (ii) information produced, recorded or stored by means of any tape-recorder, computer equipment, whether hardware or software or both, or other device, and any material subsequently derived from information so produced, recorded or stored; (iii) label, marking or other writing that identifies or describes anything of which it forms part, or to which it is attached by any means; (iv) book, map, plan, graph or drawing; (v) photograph, film, negative, tape or other device in which one or more visual images are embodied so as to be capable, with or without the aid of some other equipment, of being reproduced; (b) in the possession or under the control of a responsible party; (c) whether or not it was created by a responsible party; and (d) regardless of when it came into existence;
6. Responsible party: a public or private body or any other person which, alone or in conjunction with others, determines the purpose of and means for processing personal information. In the context of biomedical research, this is the researcher or research institution;
7. Special personal information: personal information as referred to in section 26, which includes health or biometric information of a data subject.

These definitions suggest that POPIA applies to numerous health research-related activities which range from the collection of health information, recording DNA analyses, storing health and biometric information and sharing such information (Thaldar and Townsend, 2021). Special personal information includes various types of research data such as genetic information and is subject to additional processing requirements. Considering the broad nature of biomedical research activities, it is suggested that it falls under the range of activities included under POPIA's ambit. This application does not extend to the physical human biological material samples used in research but does include the information related to the sample, such as the participants' particulars, and the data derived from the sample, such as genetic information which is then recorded (Thaldar and Townsend, 2021). Given that POPIA expressly aims to protect personal information, any information which has been de-identified beyond any chance of being re-identified does not fall within the scope of the Act.

In the Act, eight conditions which correspond to certain sections have been set for processing personal information. These conditions are:

1. Accountability: anyone who controls personal information of another person must appoint an Information Officer to ensure compliance with POPIA and its principles (POPIA, s 8);
2. Processing limitation: lawful processing, minimality of collected information, consent, justification and objection as well as collection of personal information directly from the data subject is provided for (POPIA, s 9–14);
3. Purpose specification: personal information must be collected for a specific purpose only and the person from whom the information is collected must be made aware of this purpose (POPIA, s 13 and 14);
4. Further processing limitation: this builds on the previous principle as it requires that where information must be further processed by a third party, processing must still be in accordance to the purpose specified (POPIA, s 15);
5. Information quality: the responsible party, the Information Officer, must take reasonable steps to guarantee that all the collected information is complete, accurate, not misleading and up-to-date. This must also be done in line with the purpose for which the information was collected (POPIA, s 16);
6. Openness: the Information Officer must be open regarding the collection of personal information. The Information Regulator as created by the Act must be notified if personal information is processed and where information is collected, the Information Officer must take reasonably possible steps to ensure that the data subject has been informed that their information will be collected (POPIA, s 17 and 18);
7. Security safeguards: the Information Officer must ensure that the integrity of the information over which they exert control is secured through technical and organisational measures (POPIA, s 19–22); and
8. Data subject participation: data subjects have the right to request that an Information Officer confirm whether they hold information on the data subject and they may further also request a description of this information (POPIA, s 23–25).

For this discussion, only conditions 2 to 4, 6 and 8 are of most pertinence. Conditions 1, 5 and 7 are also important but are administrative and technical and will be discussed briefly below as they relate to dynamic consent.

Condition 2, the processing limitation, requires legal grounds for the processing of information and the legal ground relevant to this discussion is consent by the data subject. As mentioned above, consent is defined as being voluntary, specific and informed in terms of the Act. The data subject must thus provide consent for the collection of information, recording of DNA analysis of a taken material sample, storing health or biometric information, using such information in conducting research, and sharing the information (Thaldar and Townsend, 2021).

Condition 3, the purpose specification, focuses on two types of processing: collection of information for a specific purpose, and retention and restriction of personal information records. For the collection of information, the purpose must be lawful, specific and explicitly defined. In the context of biomedical research and the very real possibility of secondary and future use of information and data, this is obviously problematic (Thaldar and Townsend, 2021). Information may then also not be retained for longer than needed to achieve this specific purpose. This may also be problematic but less so as an exception is provided by the Act in that information can be retained for a longer period with the condition that safeguards against the use of the records for any other purpose are provided for (POPIA, s 14).

Condition 4, the further processing limitation, requires that any further processing of information must be in line with the purpose for which the information was originally collected. Again, given the nature of biomedical research and the potential secondary and future use of data gathered during a research study, this condition is problematic. Consent may, however, be sought from the data subject to further process the information (Thaldar and Townsend, 2021).

Condition 6, openness, provides that a participant must be informed when their personal information is processed or collected.

Condition 8, participation, means that a research participant is entitled to request a researcher or research institution to provide them with the record or a description of the information held by them relating to the participant, which includes information on any third parties who have or have had access to their information. The participant may further request the correction or deletion of this information (Thaldar and Townsend, 2021).

Dynamic consent is in line with condition 2 as it is a form of consent which is "extra informed" as participants are informed about all new developments related to their material or data and it is also specific in that consent becomes fine-tuned and tailored by using participant preferences as was discussed above. Dynamic consent may be helpful in overcoming the requirements as set out by condition 3, as it allows the participant to specify the purpose for which their material or data may be collected. Dynamic consent also allows for preferences to be reset regarding secondary or future purposes for which their material or data may be used, thereby extending the period of time during which material or data may be retained. The same may be said of dynamic consent and condition 4. Dynamic consent is a platform of continual communication between the research participant and the researcher and this includes informing the participant of the use, or the collection or processing, of their material and data. This means that dynamic consent and condition 6 are symbiotic. Condition 8 is also enabled by dynamic consent as it is founded on the participation and engagement of the participant as a participant-centric initiative and by easier modification, withdrawal or revocation of consent.

In addition, dynamic consent may be useful in meeting the administrative and technical requirements set by conditions 1, 5 and 7. The requirement of an Information Officer who is responsible for seeing to POPIA compliance, thus protecting personal information, would be eased as the participant themselves will be involved in protecting their own information. Condition 5 which requires complete, accurate and up-to-date information would also be assisted by dynamic consent as the participant is enabled to change and update their information and preferences on an ongoing, real-time basis. Lastly, dynamic consent as an online, platform and interactive interface may help provide security safeguards.

5.2.1 Exemptions from processing conditions

POPIA also allows for exemptions from the processing conditions and, again, dynamic consent may be able to ease some of the issues relating to this.

An exemption from any of the processing conditions may be granted where it may be shown that public interest in the processing of the information substantially outweighs any interference with the privacy of the participant. Section 37(2)(e) of POPIA includes research activities as a matter falling under public interest. This also includes health research (Thaldar and Townsend, 2021) which would in turn include biomedical research. However, showing that public interest substantially outweighs a constitutionally protected right, namely, privacy (Constitution of the Republic of South Africa, 1996, s 14) is more difficult (Thaldar and Townsend, 2021). This challenge may be overcome with the consent of the data subject but on the condition that specific consent was obtained from the data subject at the time of data collection. Although dynamic consent cannot be used retroactively to overcome this provision, if it were to be implemented now it would enable easier consent in future and would assist researchers with contacting participants to obtain consent. This would mean that using public interest as an exemption justification would be de-emphasised by the consent of the participant who then exerts control over their own privacy while exercising their constitutional right.

Special information may also not be processed unless consent has been obtained from the data subject or an exemption based on processing for historical, statistical or research purposes has been granted. This exemption will only be granted if the processing purpose serves a public interest, or if it is impossible or would involve a disproportionate effort to obtain consent for such processing and the necessary safeguards are in place to ensure that the data subject's privacy is not disproportionately adversely affected. Again, although dynamic consent cannot change the past or do the impossible, if implemented now, it would be able to facilitate the obtaining of consent without causing disproportionate effort on the part of the researcher.

As illustrated, dynamic consent may be seen as able to symbiotically coexist with POPIA, assist in administrative and technical issues and even enable its functioning and application in a simplified manner in future. This means the protection of personal information as well as privacy benefit from the use of dynamic consent.

6 Limitations and implementation challenges

Although dynamic consent holds great promise, it is not without challenges or free of limitations. Implementing dynamic consent will require cultural changes both by participants and researchers and it will necessitate research relationships which are transparent, open and engaging, and which appreciate the role that participants play in research endeavours as the sources of material and information. Personal responsibility is problematic as this may place participants in a situation, real or perceived, that they are responsible for making decisions about complex issues that they do not fully grasp or are in no position to properly assess (Budin-Ljøsne et al., 2017). These systems will also have to accommodate participant responsiveness to the duration of a study in order to avoid withdrawal at a later stage in a study (Erlich et al., 2014). In addition, information fatigue will have to be guarded against (Teare et al., 2021).

A legitimate concern raised by dynamic consent relates to the creation of new ethical questions about co-responsibility and social exclusion. Representative uptake of participants may be an issue as groups of persons with lower socio-economic status may be less

likely to engage in opt-in models of consent such as dynamic consent (Williams et al., 2015). Research Ethics Committees and Institutional Review Boards may not be familiar with dynamic consent which may hold up the approval of research projects and studies, thus negating the "quick reaction to change" advantage of dynamic consent (Budin-Ljøsne et al., 2017). Dynamic consent will also require the development of new policies and standards of practice. The consent mechanism and language will need to accommodate and adhere to existing regulatory schemes (Erlich et al., 2014).

This consent model requires technical capacities allowing research facilities and participants to engage and exchange information. For this reason, it will demand resources, including time, expertise, money and commitment from researchers, institutions and governments (Kaye et al., 2015). On an institutional level, implementation of dynamic consent may be difficult as it requires a certain e-infrastructure that is able to collect consent, to allow data preferences in order to direct the flow of such data, to capture a complete trail of data recipients, and to receive up-to-date lay summaries of research findings to return to the participants. Scalability is thus constrained by the provision and maintenance of such systems and infrastructures (Williams et al., 2015). Cost and maintenance of a dynamic consent platform may be very high as it requires staff with good communication and IT skills and may also require equipment where participants do not have their own devices (Budin-Ljøsne et al., 2017).

Unfortunately, heavy reliance on electronic communication strategies will exclude some individuals from participating in activities (Steinbekk et al., 2013). The implementation of dynamic consent also introduces issues which are not only of a technical nature but also concern the deeper ethics related to the digital divide (Wee et al., 2013). In developing countries such as many African nations, this may perhaps be the greatest impediment to implementation of a system of electronic dynamic consent. Access to technology is still largely exclusive and unequally distributed. Although numerous new methods of online engagement are becoming more commonplace, universal access is still a long way off. In addition to some participants not having access to the internet or devices, they may also not have the ability to use these technologies (Budin-Ljøsne et al., 2017) and so IT literacy may be problematic.

A further limitation also relates to research using samples and data already collected. It is, however, suggested that dynamic consent should not be seen as attempting to retroactively catch up with history—but should be implemented moving forward, for new projects starting off.

Some of these limitations may, perhaps, be overcome by using dynamic consent as complimentary to more traditional informed or broad consent processes, unique circumstances of a proposed research project permitting. Regardless, however, of the challenges in implementing a system of dynamic consent, it holds great potential for fostering and encouraging the rights and interests, trust and privacy of research participants.

7 Conclusion

This article discussed trust, dynamic consent and privacy in order to introduce and illustrate how dynamic consent may benefit and improve trust and privacy in biomedical research. Biomedical research is vital for increasing our understanding of health-related issues and holds great promise in this regard. However, to deliver on this promise, active participation of many human volunteers and distribution of data are needed. These participants have been protected by paternalism and by informed consent or by Institutional Review Board or Research Ethics Committee procedures.

However, the wide scope and nature of biomedical research challenges a one-size-fits-all approach to obtaining consent and to review mechanisms. Both informed consent and broad consent have been shown to be insufficient in biomedical research involving human participants. Researchers also need flexibility in conducting research in order to react quickly and therefore traditional approaches to the planning and conducting of biomedical research are unsatisfactory. Actively engaged research participants are becoming more commonplace in research and, accordingly, researchers have the potential to gain access to richer datasets and continued supporters of their research. When using trust-based frameworks, doing the right thing becomes easy and scientific progress becomes ever more possible.

While arguments have been made that trust is not necessary in biomedical research, studies over the past few decades noted that it plays an important role in the willingness of persons to participate in research and human participation in biomedical research is vital. Therefore, trust cannot be argued away.

This article also stated that contemporary data protection models rely mainly on de-identification but de-identification and standard data security measures are fallible. Trust and trust-enabling frameworks between participants and researchers may, however, be established by following principles whereby transparency creates trust, increased control enhances trust, and reciprocity maintains trust. Dynamic consent inherently entails all these principles.

Research participants are more likely to participate in research if they trust the researchers, the research organisation and the research project itself. The trust-consent relationship is one of reciprocity where consent is essential in nurturing trust–but it is the fruit of the tree, rather than its roots. Consent is also a mechanism which allows participants to exert control.

Since new research trends demand new models of consent, dynamic consent was introduced as such a new form of consent. Ethically, legally and theoretically, dynamic consent may benefit biomedical research and, by the participant empowering foundation of dynamic consent, trust may be established and fostered. Dynamic consent does this by:

- addressing the limitations associated with de-identification and anonymisation, while still respecting participant autonomy;
- enabling ongoing communication between participants and researchers;
- passing the control over data flow to participants;
- providing mechanisms of accountability and transparency for the use of information and data and the sharing thereof;
- improving scientific literacy; and even
- allowing for better adherence to regulatory systems.

Trust, control and privacy are inextricably connected and where participants are able to have control over the protection of their privacy they are likely to trust the research endeavour and, ultimately, participate in the research.

Current participant protection frameworks use de-identification of human biological materials and data, but it has been found that participants are willing to make trade-offs for privacy if they are then able to stay connected to a study in which they are participating. Dynamic consent was suggested in these instances as it is premised on keeping participants connected and engaged with the research project and this may therefore offer a fair trade-off. It also enables the return of findings and allows the researcher to obtain follow-up information.

This article illustrated how dynamic consent is in line with the information processing conditions provided for by POPIA and how it is symbiotic with the protection of personal information and may aid in the administration and technical requirements set by the Act.

Although dynamic consent cannot be used retroactively to overcome certain provisions of POPIA, it may enable easier consent in future and would assist researchers in contacting participants to obtain consent. It may also be of value in overcoming any disproportionate efforts in obtaining consent on the part of the researcher.

Although dynamic consent holds great promise, it is not free of limitations and these have been discussed. Regardless of these challenges, it still holds great potential for fostering and encouraging the rights and interests, trust and privacy of research participants. It should be strongly considered for biomedical research using human biological material and data.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

LP: Writing–original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The University of the Free State Faculty of Law Research Committee and the Open Access Publishing Funding Support Office for assistance in funding the publication of this article. The DS-I Africa Law project, funded by the U.S. National Institutes of Health, under the Data Science for Innovation in Africa—Law grant (award number U01MH127690), for their assistance in funding the language editing of this article.

Acknowledgments

The author wishes to thank David Barraclough for assistance in proofreading, editing and rendering of the manuscript in its present form. All errors are my own.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Author disclaimer

The content of this essay is solely the author's responsibility and does not necessarily represent the official views of the US National Institutes of Health.

References

Budin-Ljøsne, I., Teare, H., Kaye, J., Beck, S., Bentzen, H., Caenazzo, L., et al. (2017). Dynamic Consent: a potential solution to some of the challenges of modern biomedical research. *BMC Med. Eth* 18, 4. doi:10.1186/s12910-016-0162-9

Chen, C., Lee, P., Pain, K., Delgado, D., Cole, C., and Campion, T. (2020). Replacing paper informed consent with electronic informed consent for research in academic medical centers: a scoping review. *AMIA Jt. Summits Transl. Sci. Proc.* 30, 80–88. PMID: 32477626; PMCID: PMC7233043.

Constitution of the Republic of South Africa (1996). *Constitution of the republic of South Africa*.

Dankar, F., Gergely, M., Malin, B., Badji, R., Dankar, S., and Shuaib, K. (2020). Dynamic-informed consent: a potential solution for ethical dilemmas in population sequencing initiatives. *CSBJ* 18, 913–921. doi:10.1016/j.csbj.2020.03.027

Erlich, Y., Williams, J., Glazer, D., Yocum, K., Farahany, N., Olson, M., et al. (2014). Redefining genomic privacy: trust and empowerment. *PLOS Bio.* 12, e1001983. doi:10.1371/journal.pbio.1001983

Horn, E., Edwards, K., and Terry, S. (2011). Engaging research participants and building trust. *Genet. Test. Mol. Biomarkers* 15, 839–840. doi:10.1089/gtmb.2011.1526

Kanellopoulou, N., Kaye, J., Whitely, E., Creese, S., Lund, D., and Hughes, K. (2011). Dynamic consent: a solution to a perennial problem? BMC. 343:d6900.

Kaye, J. (2012). Embedding biobanks as tools for personalised medicine. *Nor. Epidemiol.* 21, 169. doi:10.5324/nje.v21i2.1490

Kaye, J., Curren, L., Anderson, N., Edwards, K., Fullerton, S., Kanellopoulou, N., et al. (2012). From patients to partners: participant-centric initiatives in biomedical research. *Nat. Rev. Genet.* 13, 371–376. doi:10.1038/nrg3218

Kaye, J., Whitley, E., Lund, D., Morrison, M., Teare, H., and Melham, K. (2015). Dynamic consent: a patient interface for twenty-first century research networks. *Eur. J. Hum. Genet.* 23, 141–146. doi:10.1038/ejhg.2014.71

Kerasidou, A. (2017). Trust me, I'm a researcher!: the role of trust in biomedical research. *Med Health Care Philos* 20, 43–50. doi:10.1007/s11019-016-9721-6

Mascalzoni, D., Hicks, A., and Pramstaller, P. (2009). Consenting in population genomics as an open communication process. *SELT* 3, 2. doi:10.2202/1941-6008.1085

Mason, N., and O'Neil, O. (2007). *Rethinking informed consent in bioethics*. New York: Cambridge University Press.

O'Neill, O. (2006). "Transparency and the ethics of communication," in *Transparency: the key to better governance*. Editors C. Hood and D. Heald (Oxford: Oxford University Press), 74–90.

Prinsen, L. (2017). *An analysis of consent with specific regard to stem cell therapy and research*. Pretoria: University of Pretoria. [LLD thesis].

Prinsen, L. (2023). Reconsidering consent for biomedical research using human biological material and associated data. *SAJS* 119 (11/12). doi:10.17159/sajs.2023/15853

Protection of Personal Information Act (2013). *Protection of personal information act 4 of 2013*.

Resnik, D. (2021). Informed consent, understanding and trust. *Am. J. Bioeth.* 21, 61–63. doi:10.1080/15265161.2021.1906987

Saha, K., and Hurlbut, J. (2011). Research ethics: treat donors as partners in biobank research. *Nature* 478, 312–313. doi:10.1038/478312a

Schuler Scott, A., Goldsmith, M., Teare, H., Webb, H., and Creese, S. (2019). "Why we trust dynamic consent to deliver on privacy," in *Trust management XIII. IFIPTM 2019. IFIP advances in information and communication technology*. Editors W. Meng, P. Cofta, C. Jensen, and T. Grandison (New York: Springer), 563, 28–38.

Steinbekk, K., Myskja, B., and Stolberg, B. (2013). Broad consent versus dynamic consent in biobank research: is passive participation an ethical problem? *Eur. J. Hum. Genet.* 21, 897–902. doi:10.1038/ejhg.2012.282

Tamir, S. (2010). Direct-to-consumer genetic testing: ethical-legal perspectives and practical considerations. *Med. L Rev.* 18, 213–238. doi:10.1093/medlaw/fwq011

Tauginienė, L., Hummer, P., Albert, A., Cigarini, A., and Volhand, K. (2021). "Ethical challenges and dynamic consent," in *The science of citizen science*. Editors K. Volhand, A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perello, M. Ponti, et al. (New York: Springer), 397–412.

Teare, H., Prictor, M., and Kaye, J. (2021). Reflections on dynamic consent in biomedical research: the story so far. *Eur. J. Hum. Genet.* 29, 649–656. doi:10.1038/s41431-020-00771-z

Thaldar, D., and Townsend, B. (2021). Exempting health research from the consent provisions of POPIA. *PER/PELJ* 24, n1–n32. doi:10.17159/1727-3781/2021/v24i0a10420

Van der Geest, T., Pieterson, W., and De Vries, P. (2005). Informed consent to address trust, control and privacy concerns in user profiling, User Modeling UM, Edinburgh, 24–29 July.

Wagstaff, A. (2011). International biobanking regulations: the promise and the pitfalls. *Cancer World* 42, 23.

Wee, R., Henaghan, M., and Winship, I. (2013). Ethics: dynamic consent in the digital age of biology: online initiatives and regulatory considerations. *J. Prim. Health Care* 5, 341–347. doi:10.1071/hc13341

Whitley, E. (2009). Informational privacy, consent and the 'control' of personal data. *Inf. Secur Tech. Rep.* 14, 154–159. doi:10.1016/j.istr.2009.10.001

Whitley, E., Kanellopoulou, N., and Kaye, J. (2012). Consent and research governance in biobanks: evidence from focus groups with medical researchers. *Public Health Genom* 15, 232–242. doi:10.1159/000336544

Williams, H., Spencer, K., Sanders, C., Lund, D., Whitley, E., Kaye, J., et al. (2015). Dynamic consent: a possible solution to improve patient confidence and trust in how electronic patient records are used in medical research. *JMIR Med. Inf.* 3, e3. doi:10.2196/medinform.3525

frontiers | Frontiers in Genetics

TYPE Hypothesis and Theory
PUBLISHED 30 May 2024
DOI 10.3389/fgene.2024.1290658

OPEN ACCESS

EDITED BY
Robert R. Downs,
Columbia University, United States

REVIEWED BY
Heru Nugroho,
Telkom University, Indonesia
Christopher Mierow Maylahn,
New York State Department of Health,
United States

*CORRESPONDENCE
Meshandren Naidoo,
214549331@stu.ukzn.ac.za

RECEIVED 07 September 2023
ACCEPTED 02 May 2024
PUBLISHED 30 May 2024

CITATION
Naidoo M (2024), The open ontology and information society.
Front. Genet. 15:1290658.
doi: 10.3389/fgene.2024.1290658

The open ontology and information society

Meshandren Naidoo*

School of Law, University of KwaZulu-Natal, Durban, South Africa

Information, as the most elusive subject, is central to all forms of thought, governance, economic structure, science, and society. Regulation of information, especially within the healthcare field, is proving to be a difficult task globally, given the lack of a qualitative framework and understanding of the concept and properties of information (or data) itself. The presentation of the overall qualitative framework, comprising a qualitative analysis of information, data, and knowledge, will be valuable and of great assistance in delineating regulatory, ethical, and strategic trajectories. In addition, this framework provides insights (and answers) regarding (1) data privacy and protection; (2) delineations between information, data, and knowledge based on the important notion of trust; (3) a structured approach to establishing the necessary conditions for an open society and system, and the maintenance of said openness, based on the work of Karl Popper and Georg Wilhelm Friedrich Hegel; (4) an active agent approach that promotes autonomy and freedom and protects the open society; and (5) a data governance mechanism based on the work of Friedrich Hayek, which structures the current legal–ethical–financial and social society. This is insightful for questions relating to the extent of rights and duties, the extent of biological bodies and freedom, and the structure of relations in distributed networked systems. There is great value offered in this framework; furthermore, it provides critical insights and thoughts about (and uncovers the interplay between) academic culture, politics, science, society, and societal decay. Note that, in line with the ideas expressed in this manuscript, such as incorporation of personal experience (thereby mending the Kantian and Cartesian gap), a first-person perspective will be used, where relevant.

KEYWORDS

information, data, governance, trust, philosophy, privacy, logic, law

1 Introduction

1.1 Interdisciplinary fib

I will begin with a brief critical analysis of some structural axioms, which form the foundation of society and thinking today. These axioms are typically "swept-under-the-rug," hence hiding them from criticism, while also entrenching their status, creating a closed oligopoly-like scenario. This paper aims to bring hidden links out from under the rug so that an open culture of critical reflection and change can start by reflecting upon those very foundations.

It was René Descartes (1641) who separated the (non-physical) mind (as pure thought, knowledge, objective, and universal truths) from the physical body, emotions, experiences, and empiricism. In doing so, Descartes's body of work/thought presented two important structural axioms to the world: (1) Cartesian mind–body dualism and (2) Cartesian co-ordinates. The former is of relevance here.

In keeping with the goal of not separating thought from experience or emotions, I integrate observations and experiences of my own regarding present academic cultures. Cartesian dualism presents itself in academic culture, whereby self-experiential evidence in academic work tends to be a taboo. Another instance is that qualitative analyses, which integrate human psychology or phenomenology, are reserved mainly for philosophical, psychological, or theoretical journals. Typically, the "hard science" journals are reserved for major quantitative results that are supported by empirical evidence. Unfortunately, the idea that qualitative hypotheses or studies belong in specific locations (and not others), because the "tastes" of an audience will not be satisfied, presupposes the audience's desires and their status as static. A preclusion of this sort serves to reduce the freedom of choice of the audience and entrenches a strong division between subject matter. This hinders the evolution of knowledge (since hidden links remain hidden in this culture). Why would literature about human psychology be any different from literature about astronomy? Does *thought* and *logic* not structure both? Previously, Copernicus believed that the Sun revolved around the Earth, thus centralizing the human as being the *universal subject* (Deutsch, 1998). Now, it seems that the human, as the observer, represents another type of universal, being the *universal object*, which is apparently removed from that of which it is a part. Both are incorrect.

Immanuel Kant then opened the doors for "relativity" (Kant demonstrated logically that the concept of the "Universal" must be limited in some way). However, Kant (1890) backtracked from the radicality within his oeuvre by positing instead that the antinomies (contradictions) of pure reason are the limits of reason, not the limits of reality itself. In this way, as accused by Hegel, Kant remained attached to pre-critical metaphysics, which posits a realm of purity, like the allegory of the Platonic cave (the truth is outside of the cave, not within it).

The Kantian Copernican revolution thus created a split between epistemology and ontology. This is a split between epistemology (as philosophy, qualitative analyses, theory, logical operations, interpretive methods, technique, structure, systems, form, objectivity, thoughts about thought itself, and universalism) and ontology (subjective knowledge, practice, content of thought, experience, phenomenology, empiricism like the sciences, and quantitative analyses). I suggest that this accurately reflects the very notion of "interdisciplinary," which only serves to entrench (1) unfounded distinctions and (2) the Kantian axioms pervading academic culture. This is a closed system since it resists any critical analysis of its very own presuppositions/assumptions (resisting change as a result). This is contrary to what the scientific method *was supposed to be* and reflects a stifling of *imagination*, *creativity*, *and critical thought*—all of which are necessary for an open system. We often tend to think of philosophy, or to be more accurate, epistemological assumptions, as being displaced from everyday life, but we do not see how those very assumptions (from thinkers like Kant and Descartes, for example) structure everyday life, all disciplines, and all sources of knowledge.

There are two reasons for the maintenance of this fib. The first is the *raison d'etre* of capitalism, wherein relations between individuals are *treated as transactional relations between objects* (Marx, 1848; Althusser, 2014). An exemplar of the aforementioned is the logical structure of Churchillian dialectics, which involves *reciprocal causation* (Naidoo, 2023d). Churchillian dialectics is a process where subjects function as objects that freely contract with other objects in society to create/consume objects. These object relations are thought to satisfy desire and provide fulfillment (Althusser, 2014), given that these objects seemingly validate a subject's sense of freedom of choice and the ability to contract freely. This is known as the classic liberalist interpretation of freedom—which is freedom of choice *in relation to objects*. Marx (1848) called this *commodity fetishism*.

The commodity fetish presents freedom as being tied to objects, and thus, freedom is increased where there is a greater number of objects to choose from. However, while freedom of choice *between or of objects* may be increasing, *meta-freedom*, which is the *concept of freedom of choice* itself, is *decaying*. In other words, people have more choices related to objects *but less scope to construe freedom as something else or choose among different types of freedoms.* Unfortunately, this involves a degradation of *qualitative freedom*. Increasing degradation relating to the scope of meta-freedom is axiomatic of an *unhealthy and closed society*, as Theodore Adorno (Horkheimer and Adorno, 1989) and Louis Althusser each argued in their various works. For additional information on freedom and unfreedom, choice, and the paradigm of relations between objects that characterize modern society, please refer to part E of Supplementary Material.

The second, as Sapolsky (2017) demonstrated, is that the dopamine system in the human brain does not support the capitalistic understanding of "satiation through objects." The dopamine system uses objects for the pursuit, which is the goal (thus, this is contrary to the typical capitalistic understanding, as expressed previously). This was originally a Freudian insight, where Freud described the "objectless drives" (or the Lacanian *object-cause-of-desire*) (Naidoo, 2023b). Importantly, increases in abundance, in terms of object access, result in less uncertainty, and thus less dopamine release upon acquisition. More predictability leads to less dopamine (and hence that good feeling). Thus, neurobiology/neurochemistry does not support the notion that satisfaction is obtained through choices among material objects. Rather, satisfaction is geared toward the *pursuit of an uncertain trajectory*.

Marxist dialectic materialism is different from Churchillian dialectics, in that the former posits a *reciprocal constitution* instead of reciprocal causation (Naidoo, 2023d). This paradigm has been confirmed by scientific evidence. Sapolsky (2017) pointed out that the prefrontal cortex, as the executive region of the brain, only matures in a human's late twenties and is more susceptible to contextual influences, as opposed to genetic ones. Furthermore, Dawkins (1976), in *The Selfish Gene*, demonstrated that phenotypes develop a kind of freedom from their genetic constituents and are more susceptible to contexts as opposed to said constituents. Dialectical materialism thus highlights the importance of *contexts* within the paradigm of *development*. Contexts, as described below, *enable more meta-degrees of freedom precisely because contexts modulate rapidly in open systems.*

In terms of the second reason for maintaining the fib, the capitalistic system requires that subjects repress, ignore, or are prevented from grasping the hidden links between different subjects, not dissimilar to the hidden variables hypothesis, or the

EPR paradox, formulated by Einstein–Podolsky–Rosen (Einstein et al., 1935). For the capitalist system to maintain itself, these links *must remain hidden*. If links remain hidden, people are less likely to question the *appropriability* of the system itself.

Hegel (2018), who preceded Marx, introduced the concept of "embodied cognition" (as the substance equals the subject). This has since come to the forefront within the sciences and ethico-regulatory conversations (thanks to developments in artificial intelligence) (Juarrero, 2023). However, these discussions do not go far enough; although they demonstrate the false separation between the mind, body, and emotion; for example (Damasio, 2005), they do not seek to mend the Kantian split. Lastly, the maintenance of this fib, and the entrenchment of a strict divide of the disciplines, serves to hamper the cohesion of the overall total distributed network, which is that of knowledge acquisition (described below). To assist the reader, a brief navigational map is included in part G of the Supplementary Material.

2 Knowledge society

2.1 Enlightenment and the knowledge enterprise

The European Enlightenment greatly impacted the foundations of all system-building, including legal, economic, social, political, financial, ethical, religious, and scientific systems. The European Enlightenment took three main forms: (1) a political thesis for better governance; (2) a philosophical thesis for a secular foundation based on rationality and science; and (3) an economic thesis on creating more wealth (Mokyr, 2012).

What allows knowledge, ideas, and thought to flourish is an open society (Popper, 1945; Thaldar, 2017). An open society is one in which the thoughts of mad men, who go against the grain, are not rejected or punished as they were in 17th- and 18th-century Britain. These rebels must be protected and afforded the space to be the mad men that they are to protect the openness of a society; it is the acts of such mad men that *ensure that society reflects upon itself*. As described by Žižek (2003) and Naidoo (2023a), it was Immanuel Kant who put the first crack in the concept of the Universal, the infinite, or the concept of "objective," followed by others, such as Hegel (2010) (a complete ontology *of incompleteness*), Karl Marx (1848) (false consciousness), and in mathematics, Gödel (a conditional mathematics), George Cantor (the Cantor set describes a limited infinity), and Hans Peter Luhn (the computer science concept of hashing, which is analogous to the Cantor set) (Naidoo, 2023b).

The core of *liberalism* stemmed from the above, which is the *marketplace of ideas* (the marketplace is the necessary condition for the social contract). In the marketplace, ideas compete, fight, coalesce, triumph, and hibernate. The marketplace has a structure that reflects varying interests, like economic, social, and political interests (Mokyr, 2012). None of these interests solely determine outcomes within the marketplace, and each interest typically has varying degrees of importance; there is thus another marketplace within the marketplace. The ultimate marketplace, within which all others are nested and structured, is the *theory of reality*. The ultimate boundaries of a marketplace are thus ideological and based on theories of knowledge and modes of structuring or validating said knowledge. Each mode or theory turns on questions like, what makes knowledge possible? What is persuasive? What kinds of evidence and logic are possible? What kinds of experiments are necessary? How do we structure what the content of truth is? What does it mean to be correct? What is a true statement?

To provide answers or to construct new questions and answers, it is important to determine whether the structures of any given marketplace being visited are suitable or unsuitable. Typically, debates on this issue lead to reflections involving the Industrial Revolution and the Enlightenment (Mokyr, 2012). The English Enlightenment concerned the removal of ancient and conservative governance structures and values. However, this new modern form seems to have just translated some of the older values into a different language (Mokyr, 2012). The English Enlightenment was not in opposition to Protestantism; rather, it took the Protestant ethics and created a new, secular society around it. The new capitalist society is Protestant, often without knowing it. Herein, a liberal form of Christianity was used to justify the pursuit of one's self-interests, which perfectly suited the industrial epoch. Hence, science, politics, law, ethics, religion, and the like are shaped by the same schema.

2.2 Scientific method

In the 18th century, leading up to the invention of the steam engine, science was very different from today. The common quip is that the invention of the steam engine did more for science than science did for the invention of the steam engine. This "pre-modern" science shaped important developments within empiricism, which became known as the "scientific method."

There were many important discoveries in the sciences, such as formulae and understandings relating to heat, thermodynamics, and electricity (Rosen, 2010), which catalyzed perspective shifts as to how the world, physics, and chemistry worked. Given these discoveries and the "force model" imposed by Newton, causation and determinism were the dominant ideas of reality. This catalyzed a move away from the teleological purposiveness, inherent to Aristotelian philosophy and methodology (Naidoo, 2023d).

The scientific endeavor involved testing conclusions, ideas, and premises and falsifying non-repeated results since science, at this time, believed that invariance and physics should dominate knowledge validation and status as such (given the huge successes of the cause–effect relations of Newton's force models). What science sought out was evidence in the search for knowledge, and in doing so, it also moved away from the influence and dominance of the Catholic Church. From this cultural shift, the concept of "replication" was born, and inventing was viewed as a social enterprise, instead of a private one. Previously, inventing was seen as the practice of a "lone genius" (Rosen, 2010). The currency for inventors during this time was mostly recognition or fame (Rosen, 2010).

The notion of perfection, ideal, objective, or universal, at least in Western scholarship, traces back to Plato (2002) and his World of Ideas and Pure Forms. Kant and others built off these ideas through

the introduction of *a priori* pure forms which structure the mind, experience, and thought.

Technological advancements in this period were nominal, mostly consisting of simple improvements on instrumentation and navigational tools (Rosen, 2010). However, there were many novel understandings, scientific principles, formulae, and logical deductions that were understood as *a priori* knowledge. Society at that time, did not benefit from the unearthing of *a priori* concepts. In other words, no tangible value in the form of usable forms of technology was acquired from their derivation. Developing tangible or usable forms required a lot of time and other resources, much more than one individual could possibly provide. Hence, inventing became a social endeavor, wherein many could collaborate, thus sharing resources and responsibilities in the production process (enhancing usefulness, accessing, recording, storing, working, and exploring different avenues). This is what a distributed, networked system is. Hence, the scientific method sought *value* in the *form of physical objects*, and the production of the said value required bridges to be built between *a priori* concepts and utility; theory and empiricism; and theorists and artisans (Rosen, 2010; Mokyr, 2012). The scientific method, as the vital cog, transformed the Age of Enlightenment into the age of utility, objects, and the later Industrial Revolution. The purpose of the formation of the 1836 Select Committee on Arts and Manufactures (Sproll, 1994) was for the committee to determine the best ways in which to disseminate knowledge and principles in the fields of knowledge production and that of makers or manufacturers (Sherman and Bently, 2003). The singular value of *utility through the transformation of knowledge into physical objects/products* was siphoned into the jurisprudence of patent law. Inventing, in law, was now solidified as a social enterprise, which included the *acquisition and validation of data and knowledge*. The dominant paradigm prior to this Enlightenment was that of the Lockean private property. In this view, knowledge was understood to be within the personal domain of a person and thus private. Importantly, this underlies the typical conception of personal identity and privacy, which many are unaware of being traced back to Locke. The purpose of the Enlightenment was to bring out, and ensure transparency and access of/to information and knowledge, which was hidden away (hence the name "Enlightenment") under the veils of preceding religious and cultural practices (Dolar, 1991). This was the first move toward an open society, based on information and knowledge being freely accessible.

The scientific method ushered in a new understanding: to enhance conversion and utility, knowledge and data had to be recorded and shared for testing. Knowledge was then understood to be conditional on trust and not absolute (Rosen, 2010). Knowledge and conclusions could be improved and replaced. No longer did logic hold center stage as the singular source of validation. Experimentation became the defining methodology in the social enterprise of knowledge (Rosen, 2010).

Unfortunately, the "practical" or utility justification of innovation pervades many academic explorations, and the necessity and difficulties associated with obtaining funding only serve to entrench four fibs: (1) scientific empirical *experiments* are necessary to validate knowledge; (2) there is a strong separation between "thinking" and "doing." As demonstrated by thinkers such as Kant, Hegel, Freud, and Marx, knowledge can be validated through thought alone (second-order inferences). Sapolsky (2017) also echoes the argument that one does not need to perform empirical experiments to simply confirm something that is observationally clear. Fib (3) is that there is a strong separation or difference between science and religion. The final fib (4) is that there is such a thing as *progress*. Taleb (2007) demonstrates that progress is not a *linear* concept (nor averageable, much like knowledge), nor is there any consensus as to what constitutes progress. Given that neither progress nor knowledge is linear, both are impossible to evaluate in extremist systems.

3 Freedom and surprise

3.1 Semantics and meaning

The breakthrough of communications theory, later dubbed "information theory," occurred when Claude Shannon and Weaver (1964) published their book *The Mathematical Theory of Communication*. Both felt it was a mathematical theory of communication, as opposed to a theory of information. The mathematical formulations therein describe the necessary conditions for communication using concepts like symbols, signals, and carriers. Therein, Weaver described information as the measure of one's degree of *freedom of choice* when selecting a message (Shannon and Weaver, 1964).

Shannon and Weaver (1964) were concerned with *transmitting information,* and not *meaning*. Thus, Shannon created a formula that could transmit messages in the presence of *noise*. This formula for encoding messages with maximum efficacy was the same as the one created by Ludwig Boltzmann half a century earlier (Hidalgo, 2015). Both formulas treated *information as physical*. For Shannon, informational entropy is the *minimum volume of data necessary to specify any type of message* (Shannon and Weaver, 1964).

In this paradigm, information is *meaningless* (and meaning is not information); *it is the receiver/interpreter/perceiver who weaves-in/transmits meaning* into information. It is *not valuable* either. Meaning and information are often confused because of this automatic (unconscious) transposition (Taleb, 2007). This creates the illusion of a hidden depth within the information. This automatic transposition occurs because of a need *to reduce information and conserve energy*. Thus, meaning is not the message, and meaning lies within the receiver, contexts, and prior knowledge (Hidalgo, 2015). Meaning is a tool used to communicate the physical order of things.

Meaning (dubbed semantics, from here, as semantics are associations of meaning) exists in various spaces and times, within everyday life, and history. Depending on the spatial, temporal, or spatiotemporal co-ordinates, the semantic content of each co-ordinate will vary. Different co-ordinates or contexts entail different semantics, and *vice versa*. Hence, semantics and co-ordinates are co-constitutive (Naidoo, 2023d); semantics constitute spatiotemporal coordinates as much as spatiotemporal constitutes semantics. Semantics are also not just the product of, or constituted by, human brains, but rather exist as varying types of constraints, which can be context-dependent or context-independent (Naidoo, 2023d). Nonetheless, human beings do

bias semantics within spatiotemporal co-ordinates, within their favor (as individuals or as a species), to maintain their viability and structure relations within a networked, distributed system. The semantic content of a co-ordinate is constructed through processes of trial and error and jury rigging, testing, validation, correction, updating, and rejection. These processes require assessment criteria, which bias some semantic content over others, known as *selection criteria*. Selection criteria are also the product of context-dependent or context-independent constraints. For example, human society would not designate that a volcano is a semantically suitable location for a school, but in human history, it was a suitable location (the heat and noxious fumes are context-independent constraints) for sacrificing virgins to appease the gods. Hence, the semantics of contexts shape human selections.

What is required to maintain the semantic content of a spatiotemporal location is continuous, repeated observations and assessments, known as *measurements*. It was Sigmund Freud (Naidoo, 2023b) who first suggested that repetition is not something that is defined by humans but something that defines humans (Lacan, 1979). Hence, the *validity* or *continuous appropriability* of semantic content is determined through repeated observational or experiential assessment (measurements) consisting of taking in, keeping up-to-date, and updating factual information in the form of *events*. Importantly, events are not removed from observation; observation and events are reciprocally constitutive.

3.2 Passive and active mind

Understanding of David Chalmers' (1996; 2010) "hard problem of consciousness" requires knowledge of what preceded the hard problem. In 1799, Friedrich Heinrich Jacobi (di Giovanni and Livieri, 2018) wrote a letter to Johann Gottlieb Fichte in which he expressed his unhappiness about the loss of subjectivity, which arose because of Spinozism and the rationalistic physical sciences. Jacobi wanted to "save" humanity from the perils of nihilism, tasking Fichte, as the "true disciple" of Kant, with this duty (di Giovanni and Livieri, 2018). In *The Spinoza Letters* (Spinoza, 1995), Jacobi quite clearly foresaw that humanity would be saved by a return to inner experience and feeling (affect) (di Giovanni and Livieri, 2018). However, thoughts like Fichte's only served to entrench nihilism, which Hegel argued, since there was a reliance on "spurious infinities," such as the linear flow of time or advancement. Neither time nor progression is linear (Deutsch, 1998).

An important development of the human mind (and of identity) was John Locke's (1860) *An Essay Concerning Human Understanding*, wherein Locke proposed an account of the passive mind as a tabula rasa. Locke's mind simply served to reflect what was perceived. This passive mind was challenged by others, like Freud, who proposed, in his topological economic theory of mind, that the mind was active. In Freud's (1915) breakthrough work, *The Interpretation of Dreams*, he proposed that the separation of the conscious and the unconscious was a *defensive threshold*, implemented by the psyche, to protect against high energy levels of states and excitation. The purpose of this defense was to *maintain a dynamic equilibrium/homeostatic stability*. The mind *is active*. Consciousness and unconsciousness were separated because excessive energy is damaging; the unconscious, which contains the highly energy-invested (*bestsum*) forms of thought and ideas, *is limited by consciousness*. In other words, *consciousness exists as nothing but a limit to energy*. These views have been confirmed by modern science (refer to part D of Supplementary Material).

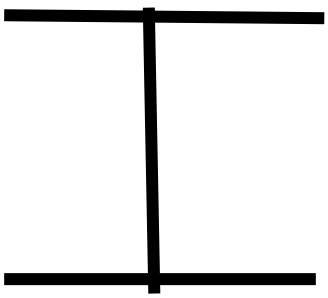

FIGURE 1
Composed by author, 2023.

Chalmers (1996; 2010) also queried why there exists the subjective "I." Naidoo (2023a; 2023c), relying on Lacanian thought, and Benveniste (1966) noted that the "I" *is a stabilizing referent*. The purpose of the "I" is to indicate the spatial location/identity of the speaker. At a deeper level, and importantly, the subjective "I" represents *a resistance to symbolic representation*, meaning a failure to obtain a positive identity through positive knowledge of *what the "I" is*. The resistance to symbolic incorporation (a deadlock or knot) is important for maintaining an open system and society, as argued by Althusser (2014) in his critique of the overdetermining effect of state apparatus. In other words, the "I" serves as an irreducible, indivisible link enabling constitutive couplings, interrelations, and entanglements that serve to create an autocatalytic feedback loop between a system and its previous context [which now becomes a "niche" (Naidoo, 2023d)]. This constitutive coupling enables for the structuring of language, identity, and knowledge. Indivisible knots are used to construct distributed, networked systems and maintain their status as "open," which is a dynamical (non)equilibrium state of stability (homeostasis). The content of the "I" is "unknowledge," meaning that there is no positive knowledge as to *what it is*, but rather knowledge *as to what it is not*. Consensus on the "I" is thus reached through *determinate negation*, instead of *affirmation*. Consensus is thus constituted through *a lack of positive affirmation*. I call this constitutive consensus dis-consensus, which in social matters takes the form of "we agree to disagree." Importantly, this kind of reasoning is analogous to the Kantian *infinite judgment*, the Gödel incompleteness theorem, or the Hegelian logic of *Aufhebung* (or chirality, present in biology and chemistry) (Naidoo, 2023b). To examine the appearance of the "I," we see that it is constituted by two parallel horizontal lines and a single perpendicular line joining the two, as in Figure 1 below.

In other words, the parallel lines are kept apart but also kept together by a vertical line, which separates and joins them. Returning to Hegel (2010), this is precisely the logic of *Aufhebung*, or chirality, which describes the holding apart, while simultaneously holding together something. They are held together by what I (based on Lacan) call a pure difference (Naidoo, 2023b). That pure difference is

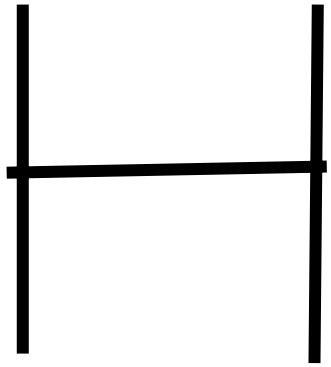

FIGURE 2
Composed by author, 2023.

an indivisible link, as described. When turned on its side, it resembles Figure 2, transitioning from the subjective "I" (also the "I" of information) to the "H" of homeostasis.

The logic of *Aufhebung* (Naidoo, 2023b) is coupled with that of an autocatalytic-feedback-loop, akin to the Belousov–Zhabotinsky chemical reaction, wherein the presence of a fourth step (the observer in the sciences) encloses an open-dynamism intrinsic to a system (Juarrero, 2023). Such a logic provides the necessary enabling conditions for a system to incorporate contextual information into its very constitution, increasing its informational content (Yoshida, 2010). The Hegelian dialectic is topological (relating to circles)—being similar mathematically to Poincare's cobordism, which he explained in *Papers on Topology: Analysis Situs and Its Five Supplements* (Poincaré, 2010). The four categories of cobordism are almost mirrored by Hegel's four-fold infinities (Naidoo, 2023b). Cobordism describes how two circles can be morphed into two other circles, like a pair of pants (Dimitrov, 2015). In Figure 3, the Hegelian fourfold infinities are represented. In Figure 4, the Hegelian ontology, which produces an "internal pair of pants" within the original, is presented. Figure 3 represents what is known as *autopoiesis*, *self-causing logic*, *or Aufhebung*.

Both *Aufhebung* and autocatalytic feedback loops are necessary conditions for open systems, open societies, and the creation of a dynamic equilibrium (a stable state of non-equilibrium, or homeostasis). It is necessary to note that knowledge is not static, but dynamic and evolving, in such systems. On a neurobiological level, the subjective state, as pointed out by Sapolsky (2017) and Naidoo (2023b), operates as a simulated, stochastic risk management strategy aimed at maintaining the viability of the system itself.

3.3 Information gain and surprise

Information is a measure of the *degree of surprise* obtained by an agent who observes or experiences an event (thus the link to

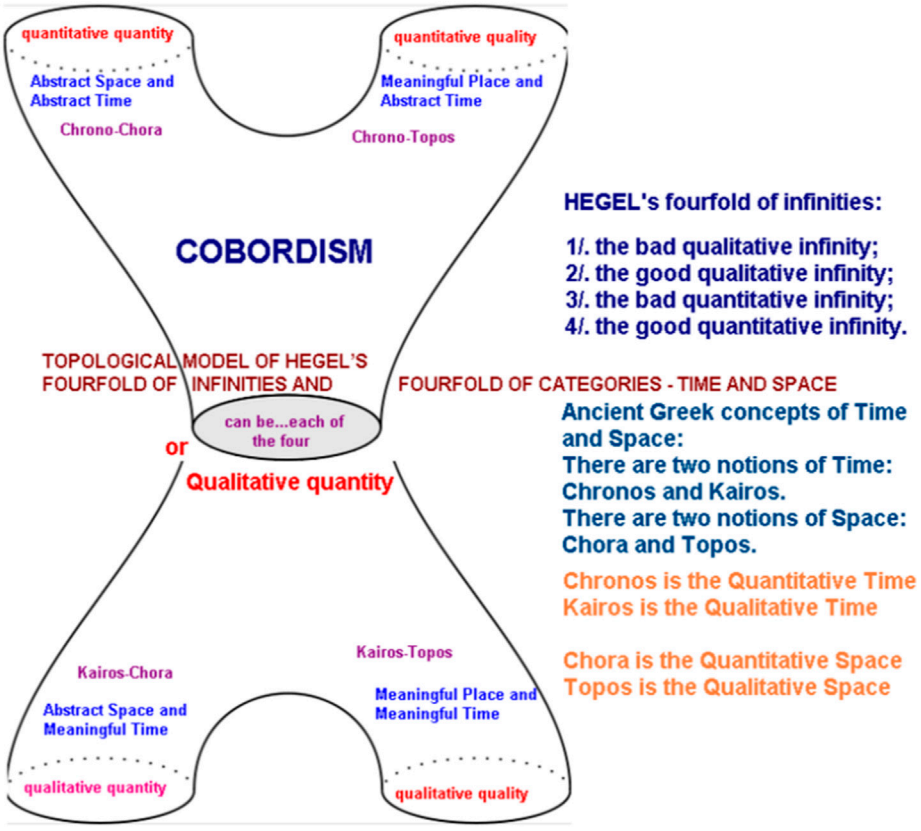

FIGURE 3
Dimitrov (2015).

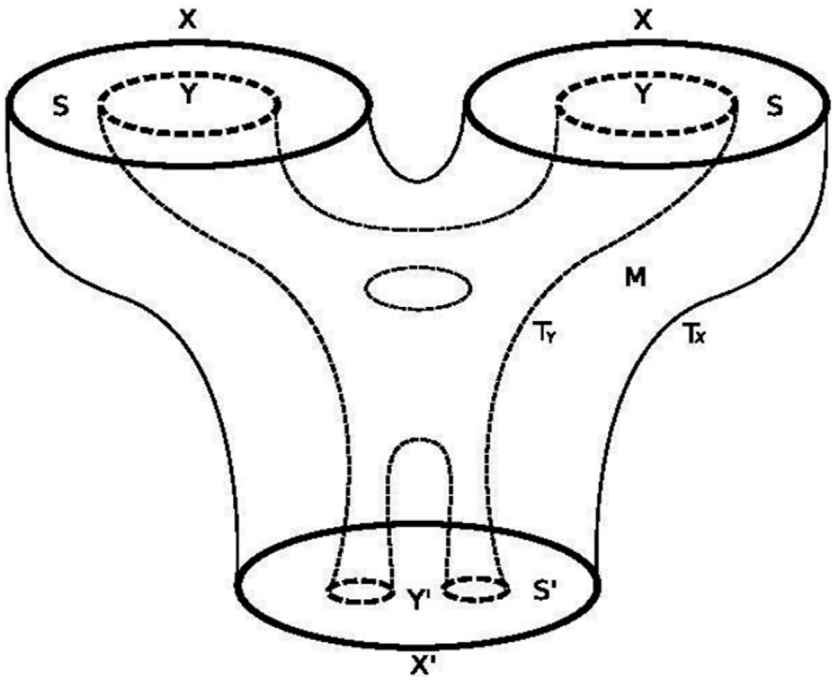

FIGURE 4
Dimitrov (2015).

probability). This degree of surprise of an event is described by *self-information*. Information (I) is *thus a function* called self-information. Self-information is the informational content inherent to any event or occurrence. *Information entropy* extends this idea to discrete random variables (X). The entropy of (X) is the *average of self-information over all possible outcomes of (X).* Conversely, the entropy of a random variable describes *the average degree of surprise obtained by the outcome* of (X). There is an increased information gain after a surprise, as opposed to expected or predicted events. Information gain is tied to *reductions in entropy*. Entropy, to conclude, is about the degree of surprise obtained from outcomes based on prediction (part B of Supplementary Material).

4 Building an ontology

4.1 Negative definition

Any ontology needs to begin with a definition; so, *what is information?* Currently, there is no *qualitative* consensus on an account of information; only a compendium of various and vague axioms exists (Floridi, 2009). We know that information is quantifiable, additive, storable, and transmittable. It is also a golden thread that runs through all disciplines (Deacon, 2007). It is certainly incorporeal and intangible because one cannot physically handle or manipulate information. In the 21st century, the importance of information increased, resulting in heavy commodification (Badiou, 2006; Hidalgo, 2015). Information was thought to be physical because of its *physical embodiment*. It was not construed as a "thing," but instead *a physical arrangement*, taking the form of a thing. Information was that which differed from its surroundings, based on its *identity*, which is its *appearance*, or the *physical order of its arrangement* (Hidalgo, 2015). As time passed, the nature of information morphed into being understood as digital, immaterial, and non-physical.

A founding father of cybernetics, Norbert Wiener (1961), provided a *negative* (exclusionary) definition of information in *cybernetics*. Wiener (1961) said that

> ". . .The mechanical brain does not secrete thought "as the liver does bile," as the earlier materialists claimed, nor does it put it out in the form of energy, as the muscle puts out its activity. Information is information, not matter or energy. No materialism which does not admit this can survive at the present day."

The last part of this quote is insightful; *information is not matter or energy*. This is a negative definition since Wiener does not propose to know what information is, but he suggests *what it is not*. Although information is neither matter nor energy, *it needs matter for embodiment and energy for its communication*.

4.2 Reduction, patterns, and data

Entropy is a limit on efficient communication of the outcome of (X). In other words, it describes how much compression can take place while maintaining the efficiency of the communication. In communication theory, there is a correspondence between the base of the logarithm and the symbol quantity, which is used in a hypothetical scenario involving two agents who communicate

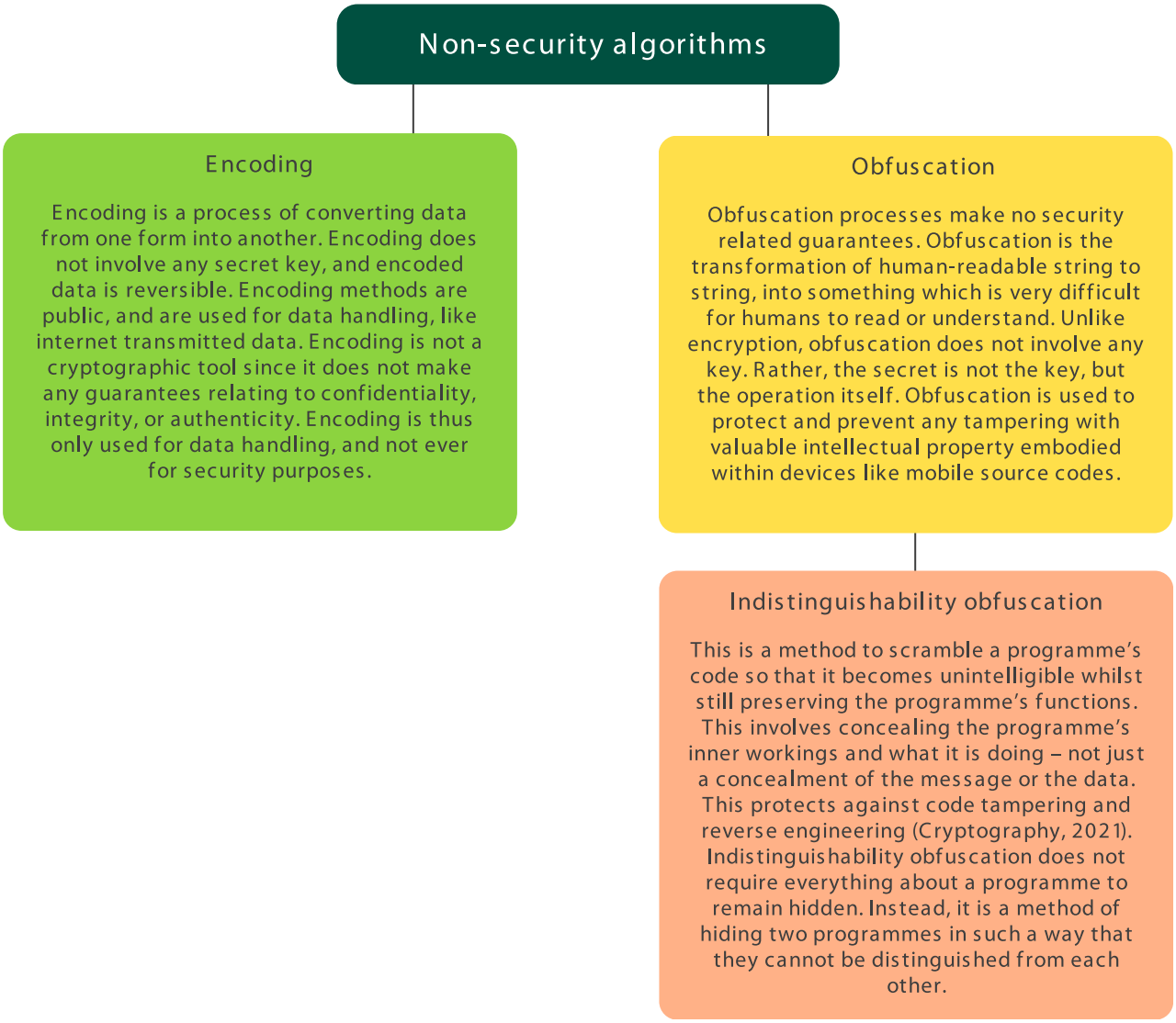

FIGURE 5
Cryptography. Composed by author, 2023.

surprising, random events (Brownlee, 2019). Any change in the base of the logarithm results in a corresponding change in the entropy equation (Bernstein, 2020). This means that informational entropy describes the minimum number of symbols (lower bound) needed to communicate an outcome of (X). The base of the logarithm of the self-information function is also the lower bound on the number of symbols required to communicate, as above.

Information is expensive (energy-wise) (Taleb, 2007) to obtain (the rule-finding, updating, learning, and executive region of the human brain, being the prefrontal cortex, is highly metabolic) (Sapolsky, 2017). It is similarly expensive to store, order, manipulate, and retrieve information. To combat this, systems need energy-efficient ways to handle and store information. The solution to the energy-efficiency issue is to order information, thus making information less random, which requires an association or attachment to narrations, words, or symbols (Taleb, 2007). As Shannon and Weaver (1964) demonstrated, for efficient communication, strings of symbols and signals can be used. Shannon (1940) also demonstrated that some symbols will be used *more frequently* than others (the source of the code will have a higher frequency with regard to certain symbols or knowledge). Based on this, it was possible to assign shorter codes to more frequently occurring symbols (including symbol pairings), reducing the total length of the required code. Natural language processing (in artificial intelligence) uses the same principle, namely, the frequency of a letter depends on what precedes the said letter. This is Shannon's empirical entropic frequency distribution formula (Hartnett, 2022). This is *a predictive functionality*. Importantly, Hawkins and Dawkins (2021) demonstrated that the brain functions according to a single cortical algorithm—*which is prediction*. Humans build their own subjective (and inter-subjective) world-maps according to this system, comprising spatiotemporal semantic relations.

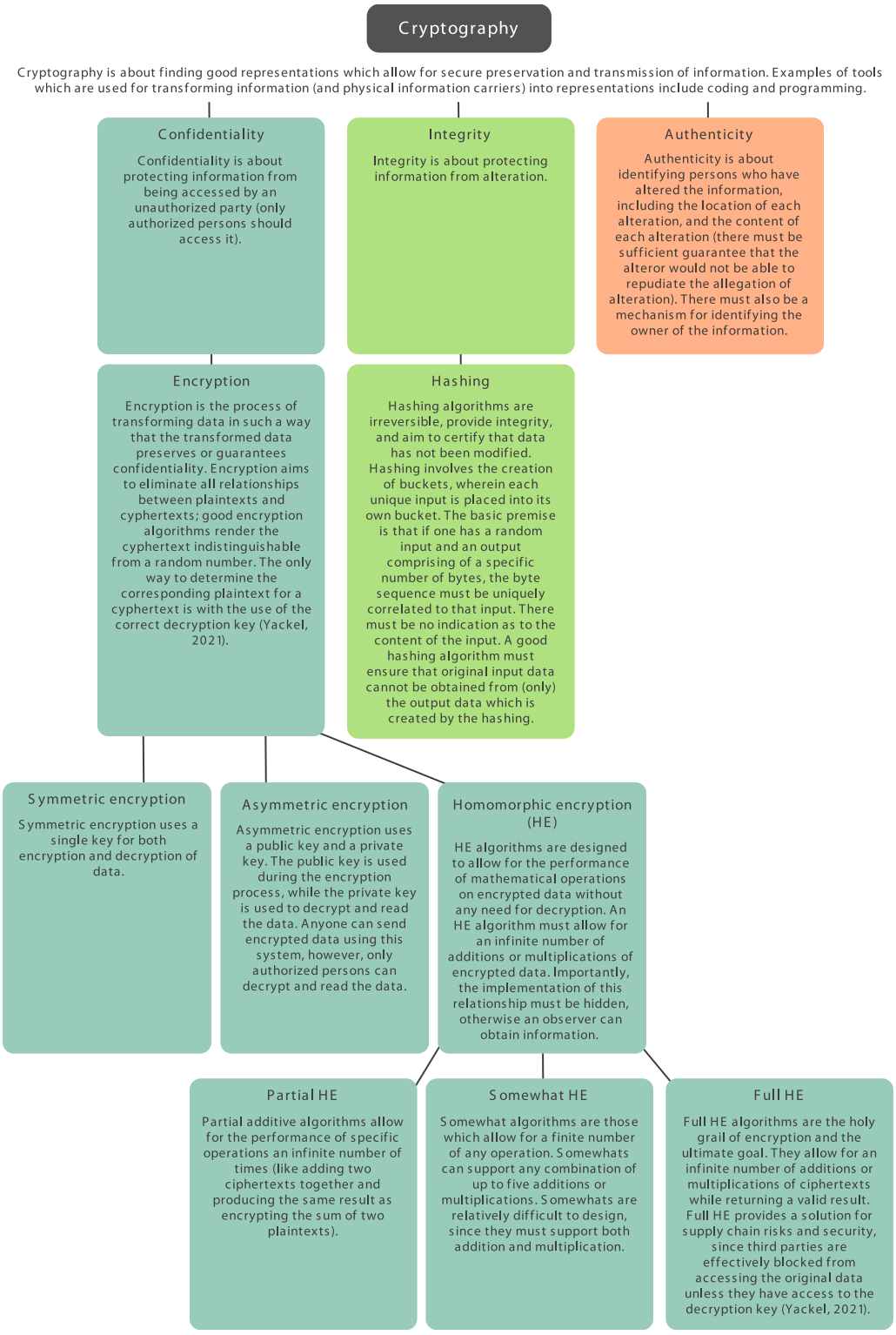

FIGURE 6
Non-security algorithms. Composed by author, 2023.

In other words, *compression* is fundamental for energy efficiency. Compression is achieved through the creation of tunnels, called narratives, wherein the *dimensionality* of information is reduced (Taleb, 2007). A side effect of compression or reduction is that vast chunks of information are typically ignored. Along with reduction, compression requires an

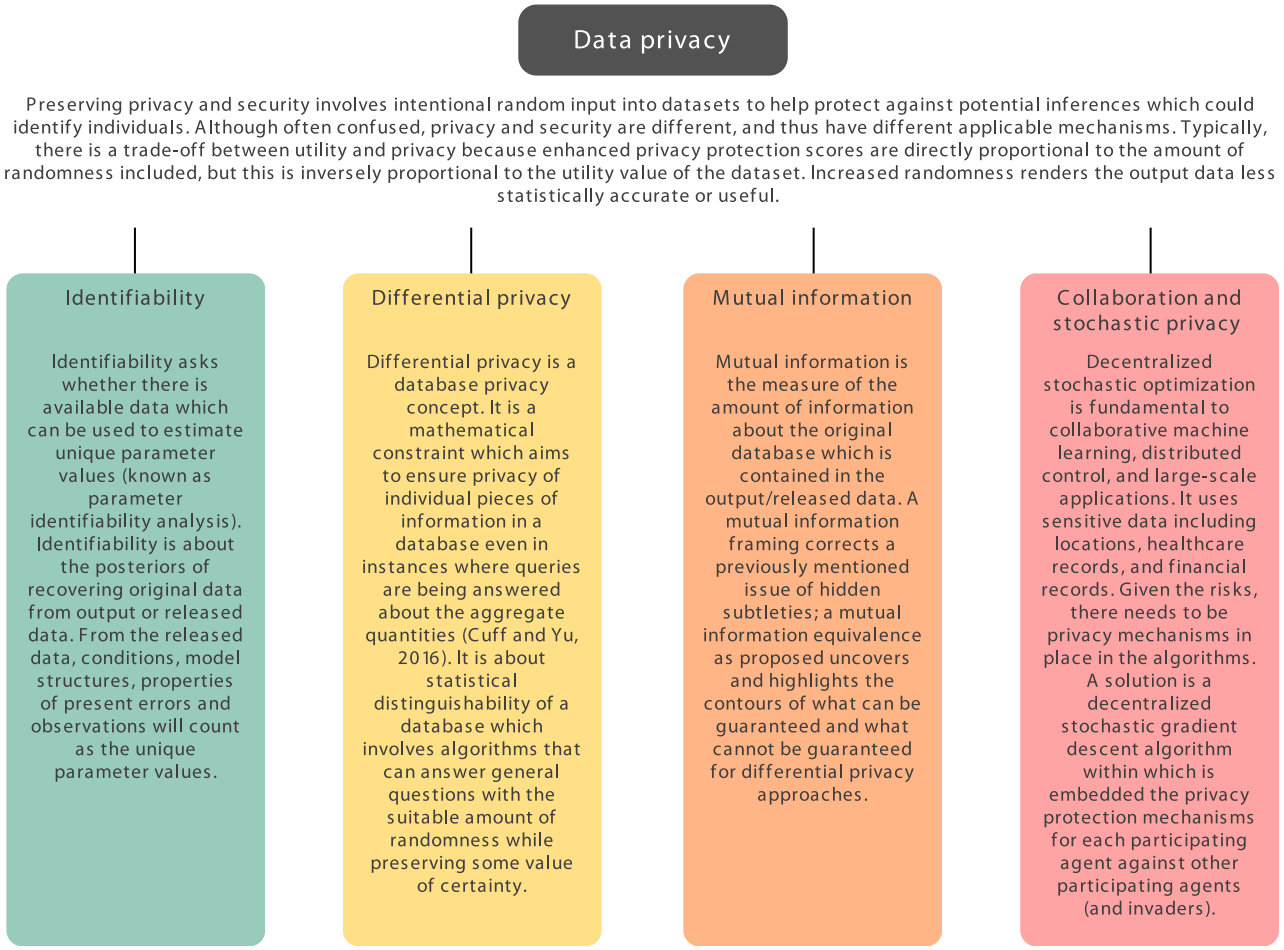

FIGURE 7
Data privacy. Composed by author, 2023.

active function of repeated, continuous assessment and observation, wherein information is abstracted (reduced in dimensionality). The abstracted information is reduced through the combination of both ignoring chunks and the creation of narrations (association links or chains of sequential chunks of observed events). These selective narrations, as sequential observations, are known as *patterns*, which connect events, thus *making prediction possible*.

Patterns are abstractions or *representations of reduced information*. Abstraction (as opposed to concretization) is fundamental to mathematics and physics. Abstraction describes the derivation of non-physical patterns. Concretization refers to the creation of physical objects. Abstraction is similar to Plato's (2002) concept of *ideal*. The transformation of the Platonic ideal into the language of sciences by philosophy is important to understand. It is this concept that structures others, like Newton's universal clock. The Platonic and Aristotelian triadic structure consists of three components: (1) the physical/material world; (2) the mental world; and (3) the world of structures. Abstraction is thus the *mental process* of removing properties from an (X), followed by attaching a name/identity to those properties. These abstractions belong in the *world of mentality*, which can be *individualized or collectivized*.

Patterns enable efficient informational manipulation and storage. It is impossible to use and store all available information, given the several orders of magnitude of energy required. Hence, only useful bits or patterns are used to make generalized knowledge possible. Examples/descriptions of patterns include summaries, compressions, narrations, episodes, sequences, slices, and foliations of information. From patterns, rules or laws can be derived, which are generalized forms of knowledge (Woodward, 2000). Generalized knowledge is that which is judged to be invariant (Woodward, 2000) for a given set of contexts, thus holding the status of governing or ruling constraints (Juarero, 2023), which allows for greater predictability, less uncertainty, and the performance of experiments (which test and maintain the validity of the said rule or law).

Rules and laws, as governing constraints, are compact, and because of their reliability value (Woodward, 2000), they enable functions to occur with less energy expenditure, while also enabling coherence, comprehensibility, and understanding. Media, in the forms of books, magazines, plays, stories, videos, movies, paintings, poetry, and all sciences are based on this principle of compressed bits of information (Taleb, 2007). Patterns also take the form of *ideas or concepts*.

The name of this compressed information, such as patterns, ideas, concepts, episodes, movies, or any of the aforementioned, is *data*. Data describe *identifiable*, embodied, encoded, or nested

patterns within various physical mediums, structures, or forms (like vehicles). In a non-physical medium, such as a mind, patterns or data exist in the form of ideas, concepts, or thoughts, each of which is largely represented by analog encoding through synchronized neuronal spiking and synaptic connections in the brain (Hawkins and Dawkins, 2021). These sequences form through continuous, repetitive observations. As Hawkins and Dawkins (2021) note, world maps are built up of reference frames (mental structures), and thinking is a virtual movement through reference frames.

Observation, I suggest, *is a dual function*. The *passive form* of observation is known as measurement, quantizing, foliation, and hermeneutics. These are repeated processes of *automatic assessments and updates*. These passive observations consist of subjective, yet autonomic behaviors, aimed at updating and maintaining a system's world map (Hawkins and Dawkins, 2021). In doing so, these processes maintain the system's dynamic equilibrium. Observations are thus subjectively sliced into qualitative and quantitative "pieces" of space, time, or a combination of both. Passive observation can occur through perception and experience (if it is possible to even distinguish between them). Observation is thus precisely a method of *biasing information in favor of the observer*.

Embodied knowledge, as passive observation, is difficult to acquire, communicate, store, and copy. The knowledge and knowhow contained in the human body (and mind) is "heavier"; knowledge and knowhow contained in objects is *relativity easier to move*, as objects can be carried and communicated through mediums, such as books and the internet (Hidalgo, 2015). Embodied knowledge is comparatively slower to acquire. This includes technical proficiency like scientific, programming, and legal skills/techniques. It is known as *expertise*. It includes the knowledge and abilities of other team members and knowledge of contextual circumstances. Embodied knowledge is *biased in terms of sociality*; it is accumulated and translated through *social learning and experience*. Beginners learn from more experienced persons; thus, it is not an individual endeavor. There is a social and experiential learning curve, which makes its accumulation time-consuming and *limits* the speed at which individuals can develop it. Embodied knowledge also biases geographic locations, which have greater quantities of some quality. To solve the issue of embodied knowledge distribution, society breaks up knowledge and knowhow among different individuals in a distributed, networked structure.

These individuals then utilize their specific forms of knowledge and knowhow *as a social network of individuals performing as a team* (Hidalgo, 2015). Through networks, the collective body of this knowledge and knowhow *can be increased*, which is *greater than what an individual can produce*. Importantly, these networks must be able to distribute knowledge evenly and ensure there is cohesion and a combination of individual parts to produce the result. It is harder to maintain a cohesion rather than to ensure everyone performs their roles. *The whole is thus greater than the sum of its parts*. For this, there must be timeous and *performative cohesion*, *shared responsibility*, *social practice*, *updates*, *corrections of mistakes*, *assessments, proper communication*, *and trust*.

The Freudian "id" described the *automatic*, unconscious aspect of Freud's topological mind. In modern terms, this would describe the *lizard brain*, which performs *autonomic* functions like maintaining homeostasis in the body (Sapolsky, 2017). As I have described, patterns or *id*eas describe subjective (or relative) perspectives intrinsic to the constitution of any pattern or data. In other words, *id*eas are inherently linked to the concept of *id*entity. The processes of observation and abstraction described above are automatic functions of the brain and body (Hawkins and Dawkins, 2021). Both observation and abstraction aim to maintain autonomic stability or homeostasis (Sapolsky, 2017). These are autonomic/automatic forms of applied *reason*; hence, the separation between observation and reason is misplaced. Reason and observation are reciprocally constitutive; memories, for example, influence reason, and reason influences and alters memories (Taleb, 2007). Reason is thus not something which one does; reason is *something which one is*. In this light, obtaining patterns is then *not* a matter of a strong form of *labor*, *expenditure*, *autonomy*, *or creativity*. As argued below, it is the modulation of contexts, which are in a reciprocally constitutive relation with subjects, which enables patterns to emerge. These are known as *discoveries*.

The *active* form of observation, which I define as *assessment* (or meta-assessment, to be more accurate), entails *active selection* and *direct participation*. Meta-assessment would constitute a strong form of labor, expenditure, autonomy, and creativity in the form of *second-order inferences*, as described below. These are considerations of *meta-suitability* or *meta-reasoning*, which are examples of reasoning about the various modalities (and viabilities) of reason itself. In other words, this process involves *biasing certain forms of reason* (and, by extension, the products of reason) over others.

4.3 Imagination and an open system

Reasoning is the name given to processes used in abstracting and compressing (and ignoring) information to form sequential narrations. Reasoning is thus a *technique* or the various employed methods of constructing sequences, forming patterns, and collecting data. The various types or techniques of reasoning are called *inferences*. Inferences (like induction, deduction, and abduction) are used to validate, falsify, cast in doubt, maintain, or update sequences or patterns. Hence, reasoning involves constructing *various types of coherences* (different techniques can obtain different patterns or informational content from the same information). Some techniques are more contextually suitable than others.

The applications of these various techniques serve to slice subjective perceptions or experiences (Hegel, 2010; 2018). This process converts each piece (through construction) into spatiotemporal sequential patterns, as functionally usable or functionally relevant data. Reasoning enables the construction of *associative semantic relationships* (a semantic network or semantic web) between observational and/or experiential information with a spatiotemporal location within a subject's world map. In other words, reasoning creates an *interrelation of dependencies*. Using reason, subjects can construct their own affordances (affordances are advantages or adaptations, which enable and create agency) based on the value and suitability of said data, in a given context. The data's functional usage has value since it enables the subject to persist (delay the thermodynamic equilibrium of the second law of thermodynamics), maintain, or enhance its viability values and

update its world maps (Naidoo, 2023d). Second-order inferences are processes of *repeated meta-assessment*. They are analogous to the Freudian death-drive (Naidoo (2023b). In psychoanalytic terms, this oscillation is known as *hysteria* (Žižek, 2014).

I conceive second-order inferences as those that simultaneously target current and previously obtained data (including memories). These inferences also target the reasoning techniques used to obtain said data, including any data obtained about the catalog of reasoning techniques, thus determining the appropriability of the data and the applied techniques (in terms of maintaining a dynamic equilibrium by ascertaining the viability value of data relating to reason and data relating to obtained data).

Second-order inferences are thus observations about observations or thoughts about thoughts or reason about reason (hence, *meta-reasoning*). These inferences are typically ignored in favor of observed data (Hossenfelder, 2016), especially in the sciences. These kinds of inferences are often labeled as "mere philosophy" in my experience and ignored. However, they are most important since they ensure *that a system remains open and viable*, or, in Popper's (1945) understanding, a closed system of totalitarianism does not ensure. These inferences are aimed at questioning the natural or accepted order of things and serve to undermine settled positions by demonstrating their inherent contradictions, *a la* Hegel. In other words, transposed into Schrödinger's (1944) terms, second-order inferences enable a system to persist, avoid their (systems) own entropy increases, and thus avoid thermodynamic equilibrium associated with heat death. Second-order inferences are thus those that maintain a stable-non-equilibrium state, known as a dynamical equilibrium, using *negentropy* (Schrödinger, 1944). In psychoanalytic theory, a closed system is one that has *psychotic* foreclosure, wherein "things" are accepted without question. The goal of psychoanalysis is to move a subject from a state of psychotic foreclosure to a state of hysteria.

Three important questions to answer are as follows: (1) when does data become knowledge? (2) what are the conditions for the acquisition of the status of knowledge (and acquisition of knowledge as such)? and (3) what is learning?

The first two questions are strongly linked and can be dealt with concurrently. Data becomes, or is converted into knowledge *only upon gaining a certain grade of trust*. Trust grading is based on various considerations, such as the invariance (Woodward, 2000) and the value of the data, both of which are related to the aim of maintaining a system's dynamic equilibrium. Thus, knowledge is data *trusted* to maintain or enhance a system's dynamic equilibrium. To establish trust, there must be repeated observations and assessments (mainly second-order inferences). The observations and assessments must also target modes of data acquisition/creation, like sampling frequency, error rates, subjective and contextual conditions associated with sampling, the timeframe of the sampling, the sample size, integration with other knowledge, and many others. Trust, like knowledge, is thus *context-sensitive* or relative and needs to be continuously maintained. Both trust and knowledge fluctuate, degrade, or modulate slower in comparison to contextual information and data acquisition. In this way, data and knowledge (as context-dependent constraints) influence and reciprocally constitute one another. Contexts, *as a concept*, I understand as being *the* overarching context-independent constraint. Trust and knowledge exist in a dynamic equilibrium, both serving to maintain an open-dynamic-equilibrium system state. Repeated assessment is necessary, not just repeated observation.

In terms of the truth–knowledge dichotomy, knowledge is *not absolute* (universal or objective) but *relatively-absolute*. Relatively absolute, instead of *absolutely relative,* highlights the distinction between a *postmodern insight* and *insights into postmodernism*. The former would posit that knowledge is completely contextual, which would delight Locke (given his passive mind) and a contextualist like Jacques Derrida. The latter, on the other hand, would suggest that the former replaces one false universal/idol (absolute objectivity or universality) with another false universal/idol (absolute relativity). Thus, the latter attacks the concept of absoluteness/*universality* itself. Hence, absolutely-relative can be explained with reference to art: in societal terms, art cannot be absolute subjectivity or just anything anyone says it is. Art is a singular thing; that singular thing is where *current consensus* lies, particularly in art, and it is situated in an art gallery. In Kantian–Lacanian terms, art, or knowledge, is that which is sublime or imbued with fundamental fantasy (Naidoo, 2023b). The sublime, or the fundamental fantasy, *is trust*. Absolutely relative, as the postmodern, contextualist account of knowledge, typically ignores the underbelly consensus of *unknowledge* or that which is deemed *not trusted*. For example, if I write a paper about the origins of life, which is then disproven by someone else, it may seem like the state of knowledge has not been improved. However, this is not true because invalidation itself serves as a reduction. The state of knowledge *knows what does not describe the origins of life*, which is my invalidated paper.

As pointed out by Popper (1962), the aim of science is not *truth* but rather *knowledge*. Knowledge is conditional; knowledge is only acquired as such through confirmation, repeated assessment, and integration (Woodward, 2000). Science thus aims to build generalized knowledge. Truth is impossible to obtain, argued Popper, because there are infinite paths in history from which knowledge could have originated, making the endeavor of obtaining truth fruitless. Science thus *does not prove; science only confirms through corroboration or refutes*. This means that while knowledge may be relative to space, time (epochs), or spatial–temporal locations, it nonetheless is *relatively absolute* since knowledge is only knowledge as such if it contains *selective trust* (which is an expression of *societal autonomy*). Hegel (2010), anticipating Popper, presented this insight in a different way through his explication of the necessity of contingency. *Facts* are thus forms of knowledge, which have a higher trust value and appear most often in a social context; *but facts are not truth*.

Knowledge and trust are thus based on *consensus*. Consensus is ultimately about building-in, indivisible, irreducible knots into a distributed, networked system (a multi-agent system) (OpenCSF, n.d.). This allows a system to function and maintain a dynamic equilibrium. The issue is that consensus is often interpreted in a closed manner, meaning that it requires the constitution of consensus to be an *agreement* among all participants or agents. However, as voting within politics demonstrates, the presence of a winner does not mean that there is consensus.

Total agreement is unnecessary when the function of consensus, which is irreducibility, is unearthed. In this light, consensus can also be a dis-consensus, meaning a consensus based on a *failure to reach consensus*. This kind of conflict takes the form of "we agree to

disagree." This is an irreducible link, which entangles polemic positions in such a way that it keeps the overall system in an open, dynamic equilibrium (hysteria). Žižek (2008) calls this "oppositional determination," which, in computer science, is known as "a split brain" in distributed, networked systems (the split brain can be traced back to Kant, who demonstrated the split between understanding and reason. This split was confirmed by Sapolsky (2017) as analogous to the neurobiological workings of the amygdala and the prefrontal cortex. Hegel also described the idea of the split as "unhappy consciousness"). I call this a *pure difference*, which is how Lacan described the way *sexual difference* is articulated in society (Naidoo, 2023b). It is not that there are differences between polar positions or contested points of view; instead, a *meta-difference* is introduced, wherein the *difference itself is conceived differently*. If difference itself is construed differently, there can be *no consensus*, and a dis-consensus results. Dis-consensus ensures that there is a *radical enclosure of openness* within a distributed networked system, enabling the persistence of its status as dynamic, as opposed to static. This is also known as an open society in political terms. In this society, opposing sides remain linked while in a state of continuous observation and assessment because of their very (intentional, unbeknownst to them) oppositional determinations. It enables a system of this sort to continuously and dynamically seek out new gradients of energy or information, which are relevant for viability maintenance. Hence, it is not that identities are used for violence; rather, intentional, conceptual violence is performed *to create identities*. The creation of an indivisible knot as such, which structures a split brain, requires the use of second-order inferences.

On question 3, learning is the process of repeated observation, assessment, validation, correction, storage (memorization), updating, degrading, and relating useful or important semantic, spatiotemporal information within a system's world map. It is thus an active process intrinsically related to the constitution (construction) or destruction of data, knowledge, and trust. This is the *dual role of imagination. Imagination is absent in a closed system*.

4.4 Launderer and physicality of information

Ralf Launderer (1961) suggested that information, as a mathematical object, plays a crucial role in physics. His intention was to find the minimum energy required for computation using standard thermodynamics. He used the Launderer reset, which comprises a starting state (say 0) and a binary switch, the latter of which consists of "1" and "0." Each binary state is a possible logical state for this binary switch. This operation is often referred to as "information erasure" since it reduces the amount of information that can be associated with the binary switch. Before operation, there are two possible states; after operation, there is one. According to thermodynamics, a reduction in the number of possible states for a physical device requires a minimum energy expenditure, which is computable thanks to Boltzmann's equation. Launderer (1961) then proceeded to deduce the logically irreversible concept, arguing that it implies physical irreversibility. His ultimate deduction was thus that information *must be physical*. However, this spawned much research into logical reversibility, famously by Charles Henry Bennett (1973) and others. Bennet demonstrated that Launderer erred (refer to part B of Supplementary Material).

The claim that information is physical has been refuted, and the current consensus is that it is not physical. An experiment at the NiPS laboratory demonstrated that the logically irreversible gate can be reversed logically with a small amount of energy expenditure (López-Suárez et al., 2016), concluding that there was no fundamental limit and that reversible logic is not required to operate computers with zero energy expenditure. This means that there is no limit as to how much we can lower energy consumption during computation. In turn, this means that information *cannot be physical*; it only has a *physical representation* (Burgin and Mikkilineni, 2022).

Information and physical carriers/representations are different. Different physical carriers can carry the same kinds of information (different brands of pens still carry the information, that it is a pen, used for writing, despite different appearances). Studies like those done by Vopson and Lepadu (2022) demonstrate valuable insights into the teleological movement of information but also have (in conjunction with later studies) *incorrectly described the nature of information* by confusing physical carriers with information. The properties of informational representations and informational carriers are different (Burgin and Mikkilineni, 2022). Information is never directly interacted with; rather, methods of dealing with informational representations and carriers of information are used, like computation, for example. Humans are prone to *conflating* metaphorical symbols with literals due to the recent evolution and organization of the brain, with the prefrontal cortex being an honorary member of the emotional limbic system (Sapolsky, 2017).

4.5 Does information have mass?

The current consensus is that information is massless (Burgin and Mikkilineni, 2022). The physical representation of information has mass, which means it would comply with physical laws. This is important to keep in mind.

4.6 Place of information

Information forms part of the *world of structures*. In the physical world, entities like genes and neurons process, communicate, and convert information into data (and then knowledge). They communicate information first through a representative analog form, such as biological and neurological structures. Second, communication of such form is achieved with the use of chemical or electrical signals (Burgin and Mikkilineni, 2022). For example, "bits" in the digital world are information, which can have many physically representative forms (like symbols, electrical voltage, or pulses). Information is "carried" by these physical representations in the same way that temperature is "carried" by thermometers (Burgin and Mikkilineni, 2022). The General Theory of Information (GTI) describes and distinguishes the *properties of information* from those of representations and carriers of information.

4.7 General theory of information and the physical world

Material structures in the physical world carry information, which represents the *state and the dynamics* of the analog structures mentioned (Burgin and Mikkilineni, 2022). In the physical world, physical or material things are *governed by the transformation laws of matter and energy*; energy can create or change material structures. All physical (including chemical) structures, which are created or changed by the transformation of matter and energy, are governed by and obey transformation laws. Hence, this is how physics distinguishes *what is physical* from *what is not*.

All physical structures contain information, which characterizes their structures, functions of their components (including the interactions of the components with their surroundings), and their behaviors upon the occurrence of fluctuations. Factually, there is a relationship between the characteristics of physical objects, which allows for the conversion of mass into the energy of physical objects described by these characteristics. Einstein's mass–energy equivalence equation, $E = mc^2$, interlinks the energy and mass of physical objects. However, this formula does not mean that substance (matter) is equal to energy; rather, it describes the maximal amount of energy in a physical object with a given mass.

Thus, the states of physical structures and the regularities of their evolution are described by the laws of physics, which are *mental structures created by humans*. "Living" organisms developed physical structures, which exploit matter and energy transformations, to acquire unique identities and the ability to sense and process information, which is carried by material/physical structures. They can do this by converting it into data or knowledge, which are *mental structures*.

All living organisms have varying degrees of perceptive, processing, magnification, and information-to-data-to-knowledge conversion abilities. Humans can typically represent and manage mental structures using *ideal structures or categories* like named sets or *fundamental triads* (Burgin, 2010). Triads provide the schema, including the necessary operations for creating organized forms of data and knowledge, like entities, relationships, and evolutions, based on events and behaviors (Burgin et al., 2020; Mikkilineni, 2022a; Mikkilineni, 2022b). These are world maps.

Events are caused by (1) fluctuations in the interactions among the components of structures and (2) fluctuations among components and their niches (Naidoo, 2023d). Function, structure, and fluctuations play important roles in a system's microscopic and macroscopic behaviors (Prigogine, 1978). Mental models, created by information processing, are *observer-dependent*, as they are conditional on subjective foliations, previous knowledge of the observer, and various other idiosyncratic variables.

4.8 General theory of information and the ontological principle

According to this principle, information plays the same role in the world of structures as energy plays in the physical, material world. Despite this link, information is *not part* of the physical world. It can only be materialized in a physical form (Burgin and Krzanowski, 2022; Burgin and Mikkilineni, 2022).

For any portion of information (I), there is always a representation (R) for this (I) in a system. This representation is often material, and because of this, information *seems* physical (Burgin and Mikkilineni, 2022). The physical representation, rather, is the materialization or manifestation of this information and is not the same kind of thing as the information itself. This material form enables the possibility of *social exchange*, given that it allows other subjects to read, process, obtain, and transfer information. DNA is an example of an inanimate transformation and transmittance of information from one physical representation to another. It is the physical/material *representation* of information that complies with physical laws, and not the information itself. Mental processes themselves are also not physical; they are tied to something physical, like the brain, but are themselves not (Davis et al., 2012). In other words, *semantics are not physical*.

In terms of this principle, information in any system can precipitate the potential for, or cause, transformations within the system itself (like changing its structural or logical elements) (Burgin and Krzanowski, 2022).

4.9 General theory of information and the representability principle

According to this principle, for any part of information (I), there is always a representation (R) of this part of (I) for a system (S). (R) is a material representation of said information, and it is only (R) only that obeys physical laws.

4.10 General theory of information and the embodiment principle

In terms of this principle, for any information (I), there is always also a carrier (C) in a system (S). As a rule, (C) is typically material; hence, (I) is *present* in the material world. (C) is an instance of *materialization* of the information, which I call *the second level materialization sub-principle* (or SLM for short). Consider this example: a piece of paper, as a carrier, requires the materialization of symbolic information (enaction or inscription, in the form of writing letters forming a language) via an instrument, such as a pen. In this example, information is materialized and hence present in the material world when embodied within a carrier, but the materialization in the form of the inked-in written words is only a physical representation of information. The symbols, being the letters of the language, *are also carriers*. This is supported by Shannon and Weaver, who separated the message from semantics.

Thus, any (C) of (I) is a physical something within which (I) is embodied. A physical (R) is also a physical (C) if it allows for the *direct extraction* of the said information. The key difference between (C) and (R) is that any physical representation is a physical carrier, *but not every physical carrier is a physical representation* (Burgin and Mikkilineni, 2022).

To illustrate, consider the following: an envelope is a physical carrier of information (the envelope contains a paper letter with writing on it). The paper letter is also a physical carrier *of the same information* as the envelope since the information embodied within

both the envelope, and the paper letter is the text written on the paper. Given that direct extraction is only possible through viewing the text (reading it), it is not possible to extract this text from the envelope without opening it and reading the letter. One also cannot directly extract the text from the paper letter *itself* but *only from the visible writing embodied* on the paper letter. For example, if the letter is written in a visible foreign language, being in possession of the letter does not mean that one can extract the information embodied within it. It is the visible symbols themselves, from which direct abstraction is possible and proceeds (not the paper letter). Hence, the envelope and paper letter within it are only (C). Neither, however, is (R) of the information contained within the letter, since the extraction process cannot be performed on either the paper letter or the envelope. Hence, the difference between FLM (representations) and SLM (carriers) is that FLM comprises SLM, but SLM does not necessarily comprise FLM. A (C) of (I), which is not an (R) (like the envelope or paper letter), is called an enveloping carrier of (I).

The mental worlds of living biological organisms are structured by scaffolding (Naidoo, 2023b). Information obtained from the environment through the senses enables mental representation, which is then converted into mental structures in the form of triads. There are two types of mental structures: (1) those derived from external observations and (2) those created by human minds serving to represent ideal structures. Mathematics is used to represent ideal structures and operators; it is also used to model systems from the material world, their states, and their evolution (Burgin and Mikkilineni, 2022). The mental world/reality contains different mental structures that are involved in transforming information and data into knowledge. These processes are physical processors, namely, genes and neurons.

4.11 General theory of information and the rightful placing of information

Information is non-physical (Timpson, 2004; Timpson, 2008; Timpson, 2013), but it is tied to physical and mental structures and processes. Informational (R) and (C) are embodied in other physical and mental structures (Burgin and Mikkilineni, 2022). If the physical (R) is altered, the information changes too. Erasing a representation (like erasing writing)results in (R) losing its status as such since it would no longer embody information. The status as a (C) likewise can be constituted or un-constituted as such.

Symbolic (R) of information is involved in logical or abstract computation (like linguistics), whereas physical computation works with physical (R) and (C) of information (Burgin and Mikkilineni, 2022). GTI locates information not in the world of abstract objects (information exists in things outside of mentality) *but rather within the world of ideal structures. Information appears in mental and physical worlds through materialization and mentalization.* Abstract objects are mental representations of information from the world of ideal structures. They are structures themselves, but they do not belong in the world of ideal structures; they are rather external structures within the general theory of structures.

With regard to living organisms, information can be conceptually (not physically) separated into *ontological information* and *mental information*. The former is that which precipitates formations and transformations of structures within the physical world/physical systems (Burgin and Krzanowski, 2022). Ontological information functions within the physical world; hence, it is used in treating natural phenomena. Mental information (also known as epistemic information), on the other hand, is that which facilitates formations and transformations of structures within the mental world/systems (Burgin and Krzanowski, 2022).

According to GTI, *physical energy* is a type of generalized information situated within the physical world and is that which precipitates the *changing or preserving* of physical systems (Burgin and Krzanowski, 2022). There is a key difference between ontological information and energy as generalized information; the former (as genuine information that can precipitate alterations or preservations of physical systems) acts only on physical systems, which have physical representations, and are embedded in a physical carrier. The latter, energy, *directly acts on physical systems*. Ontological information can also have physical energy as its representation (Burgin and Krzanowski, 2022). This position has also been supported by the demonstration of Maxwell's demon in laboratories (refer to part B of Supplementary Material) (Hossenfelder, 2016). *Information can be converted into work.* This means that it is possible to replace the transfer of energy from a sender to a receiver with a transfer of information, and this information transfer can occur with much less energy than what the receiver gains from the information (Hossenfelder, 2014).

GTI also distinguishes between mental/epistemic information and mental/psychic energy. Mental/psychic information is generalized information in the mental world that can precipitate the change or preservation of mental systems. Mental information, as epistemic information, is genuine information that can precipitate changes or the preservation of mental systems because of the way it behaves. The difference between both is that mental/psychic information directly acts on mental systems, while mental information, as epistemic information, acts only on systems with a mental representation and embedded within a mental carrier. For example, knowledge is embedded within the mentalities/minds of people. Mental information can have mental energy as its representation (Burgin and Krzanowski, 2022).

5 Data governance

Utilizing the framework proposed by Friedrich Hayek (1945) (refer to part C of Supplementary Material), one can answer some of the important questions relating to data governance. The first is ownership of data and the second is the issue of personal data migrations, which bring into play many different, stringent, national and international ethical and regulatory frameworks.

5.1 Attribution, not ownership

5.1.1 Objects and order

Imagination is a process of ideation, as I described above, which is the process of constructing and destructing sequential, semantic relations of association. Objects are "crystallized" forms of the

imagination, in reference to Erwin Schrödinger (1944) and Ilya Prigogine, 1978. Schrödinger (1944), in *What is Life*, explained that the persistence and resistance of information (moving against thermodynamic equilibrium) are abilities gained from their crystal structures, which keep systems in a dynamic equilibrium. The information is embodied within these solid, physical crystals as patterns/data. Corporeality or solids have shielding properties, enhancing the "stubbornness" of the embodied information. The aperiodicity of solids was fundamental for the evolution of life, as pointed out by Schrödinger (1944). In social systems, humans build houses and take photographs for the same reason.

Imagination is the name of a triadic structure composed of what I call *the big three*. The big three are (1) information, (2) knowledge, and (3) knowhow. Bringing objects from imagination to life requires each of the big three. This often requires assistance from other subjects, such as a structured supply-chain (a distributed network). This is a collaborative effort. A distributed network, as a supply-chain, enables a robust and efficient way to obtain each of the big three and to structure the relations and roles between the different actors of the big three. Each of the big three can originate from different sources. For example, the desire to create an object (a concept)—like a new type of flamethrower—can be my own. The desire and idea are attributable to me; however, I have no *knowledge* of the scientific (or legal) laws and mathematics required to design it, use it, or analyze its possibility of existence. I also do not have the *knowhow*, resources, and technical skills required to bring it into physical manifestation/existence. Of the big three, knowledge and knowhow are the *rarest* (and harder to accumulate). However, knowledge and knowhow are different from the notion of *value*, despite often being equated. The value of products and the value of knowledge and knowhow are *qualitatively different*.

The value associated with products (be it notoriety or economic value) is qualitatively different from the value associated with a person who displays/possesses knowledge sets and knowhow as specific skills, both of which translate into the ability to create products. Both types of value are subject to supply and demand; however, *knowledge and knowhow are applicable to different contexts and different creations/products (higher invariance values)*. In other words, knowledge and knowhow display more invariance because they can directly translate to different contexts and different objects (the knowhow of drilling, for example, can be used to create many different objects). The qualitative difference is that *one value is tied to objects, while the other value is tied to its creator*. Hence, it is the *ability to create* that is most important, *not the creation itself*. That is why the former is afforded privilege, through *attribution*, and requires societal nurturing.

Viewing objects as crystals of imagination explains both their social and economic value. In terms of the former, they enable subjects to *feel* socially linked to one another, thus sharing in a social experience or interrelation. This is the fantasy of equivalency, wherein, through objects, one seemingly feels equivalent to the lived experience of another.

In the latter, statistics relating to imports and exports are important. The objects of import and export are (the exchange of) *crystallized/embodied forms of imagination*. Export structures and statistics reveal information about a country's ability (and requisite resources) to bring objects from imagination into physical reality. Hence, exports provide information about a country's knowledge and knowhow. If viewed through the paradigm of crystalized imagination, traditional economic concepts like the balance of trade are ill-suited to their task. An alternative lens is an analysis centered on *balances of imagination*, which involves an imagination exchange (embodied by objects). This also reframes common understandings of exploitation, which are typical in "developing countries" (the idea is that it is exploitative to buy raw materials from a developing country and then sell back to that country an object of higher economic value). The paradigm shift enables linking economic value to the source of imagination, instead of the source of raw materials. Economic prosperity relies on imaginative utility, not on consumption. For example, inventors like Faraday and Tesla developed theoretic frameworks of electromagnetism and methods to make practical applications feasible. These inventors provided imagination for "developing countries" to then see the value of their raw materials. In other words, developing countries are capitalizing on the imagination of others (Hidalgo, 2015).

Hence, economic value is not only understood in terms of the origins of physical order but also includes *the context in which these objects and orders are utilized*. Physical forms of order, or objects, allow for certain functional performances in certain contexts. These functional performances are interrelated with contexts and the arrangements of order embodied as said objects. Arrangement orders precipitate function; the need for different functions precipitates different arrangements.

Medication, for example, is an instance of embodied information, which has greater economic value in some contexts as opposed to others (hence, the economic value is modulated by the contextual conditions). This means that where medicine is produced, *and used*, is important. An instance of medication, being a pill, allows for a deeper description. Intrinsic to the pill's constitution are the practical uses of the creators' knowledge, imagination, and knowhow (this is social value). The creators precipitated a disclosure of the potential biological effects of chemical compounds contained in the pill. What is not present in the pill is information relating to how the creators obtained their understanding of these effects and how to synthesize said effects. The practical uses of the pill exist within the context of its use. The development process would have required many resources in the form of the big three. This is the background context, being knowledge of the "what" and the "how" (what connections, rules, and laws were important) and the knowhow, which is the technical skill required to bring it into physical existence (Hidalgo, 2015). It is the background context that renders the pill economically and socially valuable (utility). Contexts also provide for and modulate the value of creators. It is the variety of knowledge and knowhow embodied by various people and objects, which enables better and more creative information processing.

There is *a qualitative difference between practical uses of knowledge, the knowhow embodied within physical objects, and the knowledge and knowhow embodied in people*. The knowledge and knowhow embodied in people are related to the human experience, body, and reality—not the thing (the object) itself. It is acquired throughout development in life, scaffold (Naidoo, 2023b) thinking, ideas, and abilities. This is known as "tacit" knowledge, which is a term attributable to Hayek (1945). Tacit knowledge, in my ontology, is descriptive of the type of knowledge that arises due to

the *analog form of the human and analog biological structures*, including the brain, neuronal patterns (data), organization of patterns, perception, memories, and experiences. This is the process of *passive observation* (and the content and form of the subjective world map) I detailed above.

5.1.2 Creativity

These patterns, in both subjects and objects, are not subject to *ownership*, but are only *attributable* to persons in terms of said persons being recognized as the rightful creators of said patterns. Here, I briefly discuss the formation and settlement of these ideas, which took place in the early formation of intellectual property (IP) law (refer to part F of Supplementary Material). In the Lockean era, *labor* (as performance) was the source of property and proprietorship. Labor pertains to the *work that went into creating an intangible entity*. The labor concept was flawed, as it did not answer questions related to identifiability; thus, the notion of *creativity* evolved to supplant it.

Creativity describes a specific form of labor, namely, *mental labor*. Mental labor performed by minds enacts processes of creativity, which precipitate in protectable intangibles (in patent law). Although the *common law literary debate* used the *language of identity* more than creativity[1], I would suggest that creativity described *both the rules and the consequences* of mental performance, whereas the identification aspect added a *surplus rule* (the identifiability of the work with its progenitor). Creativity was an *internal performance*, and identity was the linking of said internal performance of the creator with the created external object.

This understanding of creativity became widespread in the mid-19th century, with leaders such as Thomas Webster (1853), who, in his *Treatise on Designs and Patents*, stated that any products

> "of the mind or intellectual labour when embodied in a practical form, whether in books, music, paintings, designs, or inventions in the arts and manufactures' have the peculiar claim derived from the nature of the subject namely, that the subject matter of such property did not exist like land, the air, or wild-animals . . . such property is, in the strictest sense of the term, a creation" (Webster, 1853; Burke Inlow, 1950; Sherman and Bently, 2003).

"Creativity," however, still needed to be described. The understanding was that inventions involved creativity, whereas discoveries were simply *observations of existing natural patterns*, which were not patentable, because this observational process d*id not involve qualitative mental labor, which constituted creation*. The conceptual bridge then to creativity, from discovery, required qualitative mental labor.

Discoveries were understood to be already existing *a priori*, which are context-independent constraints, and always existed independent of human interventions (Webster, 1853; Sherman and Bently, 2003). A genius is one who, as the height of human ability, could ascertain the pool of *a priori*'s, which consisted of scientific laws, ideas, and principles (like gravity or electromagnetism). Like the exclusion of ideas in the literary common law property debate, these *a priori*'s were excluded because of their *universality* (Godson, 1833; Sherman and Bently, 2003). Inventions, however, were objects derived from these *a priori*'s. Inventions were protectable through patent law and *attribution*. Attributions in patent law attributions thus recognize a derived utility from *a priori*'s, arising from creativity. Webster (1853) said

> "discoverer is one thing and an inventor is another. The discoverer is one who discloses something which exists in nature, for instance, coal fields, or a property of matter, or a natural principle: such discovery never was and never ought to be the subject of a patent . . . The Subjects of discovery are indeed sown broadcast; they exist in nature."

Webster (1853) goes on to note that while there may have been great expenditure in making discoveries, discoveries are not inventions (nor subject to ownership). Instead, discoveries are attributable to other mechanisms, such as awards or recognition. Inventions are those that involve a conversion from *a priori* abstract laws, existing in the minds of men, into something physical and useful. This became the *reduction to practice* requirement in patent law (Sherman and Bently, 2003). Hence, empirical embodiment, drawn from *a priori* concepts and *converted into physical reality*, constitutes an invention. In *Boulton and Watt v Bull* (1795), Justice Buller said that patents "were granted for some production from these elements and not for the elements themselves" (*Boulton and Watt v Bull*, 1795). Thus, the logic of creativity was a *human creation*, and the *object of protection* afforded by patent law was *the human element of creation and creativity in the empirical embodiment of a physical object or process* (Sherman and Bently, 2003). This is the foundation for the *conception* element in patent law, which determines the attribution of data or patterns.

5.1.3 Attribution, not ownership

An important issue to address is the common law ownership of ideas or data (part F of Supplementary Material). The common law ownership of ideas, data, or patterns was explicitly rejected for many legal and social reasons, including the maintenance of a social dynamic equilibrium and an open system. The very constitution of IP as a whole served to close common law property in ideas, data, and the like. The very existence of IP thus serves as a *consensus* regarding the limits of common law protections for legal, viability, and social reasons. IP is thus not something different from the common law of property, but it is *its limit*. IP exists as a context-independent constraint and as a designation for that *which is not* susceptible to common law ownership.

It is important to provide clarity on this position, given that recent works, such as that of Thaldar. et al. (2022), have provided an incorrect and irrational account of the data ownership question within the context of South African law. There are several substantive errors made by Thaldar. et al. (2022), which render their conclusions void. To begin, the relevant question to be answered by the authors is whether data are subject to common law ownership under South African property law. The authors note their methodology/protocol for the article as

1 There are finer distinction and details, but they do have much in common as they are both performance-based.

> "However, *the purpose of this article is not to engage in a normative analysis (what should the law be?), but to engage in a positivist analysis (what is the law?)* that draws attention to the multidimensional legal nature of personal genomic sequence data. Accordingly, we do not develop these normative arguments further in this article" (my emphasis) (Thaldar., et al. 2022).

In exploring the relevant law, the authors note that *physical control* is required as a necessary/peremptory for data to be considered property under the South African common law of ownership. The authors note

> "A *potential obstacle to conceiving of data in this way could be the requirement of physical control*, given that personal genomic sequence data are not a corporeal object. It may be recorded on physical devices" (my emphasis) (Thaldar., et al. 2022).

In this light, the authors acknowledge that physical control is a peremptory requirement for data to be considered property under the South African property law. However, in breaking with their methodology, the authors now reject the positivist requirement of physical control and then make a normative argument stating that physical control is outdated. The authors say

> "*We suggest that the requirement of physical control is outdated in today's world where so many valuable assets have a digital rather than a physical existence*, and where these digital assets are effectively controlled via digital device interfaces" (my emphasis) (Thaldar., et al. 2022).

Based on a rejection of a peremptory requirement and thus a positivist analysis as required by the authors' own methodology, a normative argument is presented so that a conclusion can be constructed that data qualifies as an object under the current South African property law, which is not the case. The conclusion reached by the authors on this matter reflects the following:

> "Accordingly, personal genomic sequence data—understood not in the abstract, but as a specific instance of personal genomic sequence data—qualify as property" (Thaldar., et al. 2022).

Had the authors followed their methodology, which is to present the law on the issues of whether data qualify as property under the South African common law of property and thus whether data are susceptible to private ownership, the exact opposite conclusion would have been reached. The rightful and rational conclusion, given the peremptory requirement of physical control, leads to the conclusion that data neither qualify as property nor qualify as being susceptible to private ownership under the South African common law of property.

Additionally, the authors claimed that an instance of personal sequence data is unowned property

> "Personal genomic sequence data are *res nullius*—something that belongs to no one" (Thaldar., et al. 2022).

However, the authors have hypostasized data as being property, without meeting the necessary criteria for something to qualify as property. The authors state the following:

> "For the purposes of this discussion, the term "property" denotes a legal object (or "thing") that is susceptible of ownership. *Property can be corporeal (such as a house) or incorporeal (such as intellectual property)*" (my emphasis) (Thaldar., et al. 2022).

However, the authors have made the mistake of conflating corporeality or incorporeality with physicality and non-physicality. For something to qualify as an object of property law, the said object must be *physical*. The authors have made category errors by confusing physical objects, of which corporeality and incorporeality are sub-categories, with non-physical instances such as knowledge, data, or information. Physical objects are objects of the common law, including intangibles. For example, gas is capable of being an object if it is enclosed; *however, qualitatively, gas is categorized as matter, which is both physical and empirical* (Rowlands, 2007). *Data are not categorized as matter, and that is because data is information or knowledge*, which is the object of intellectual property, and subject to attribution, not ownership.

The objects of both patent law and copyright are data and knowledge (identity in the form of style, in terms of copyright). Both forms exist because of a consensus regarding the societal value of knowledge and the value of the knowhow (the knowledge producer). It is not the physical body of the object that is protected (this is explicitly reserved for the common law); it is the *new knowledge/data* embodied within the object that is protected. This protection is a recognition of the *value of the knowhow producer* through the attribution of legal rights. IP was explicitly created *to govern imagination* (as information, data, knowledge, and knowhow). One of the reasons for this was the perpetuity aspect of the common law, which is harmful to societal equilibrium, and new data/knowledge production. Objects of the common law are physical embodiments of information, not the information, knowledge, or data itself. See part F of the Supplementary Material for more information.

5.1.4 Misunderstanding of physicality and corporeality

Any claim that data can be owned makes the logical error of reification or hypostasis. This is confusing a symbolic pattern with the literal (physical) object. The very first criterion for an object to qualify as such in common law property is *physicality*. I have already described what is considered physical. It is a *category error* to conflate physical objects, of which corporeality and incorporeality are sub-categories, with non-physical instances, such as data, knowledge, or information. Objects can be intangible, incorporeal, corporeal, or fungible and *still be physical*. For example, typically, gas is capable of being an object of the common law only upon its enclosure. Gas, despite its intangible nature, is *qualitatively categorized as matter, which is both physical and empirical* (Rowlands, 2007). This is not to suggest that the premises of the IP law should not be reconsidered.

5.2 Distributed, tiered approach

Following Hayek (1945) once more, we can obtain answers regarding the issue of data migrations. It is possible instead to send algorithms to personal/private data storage instead of the other way around. Data produced by the usage of a migrated algorithm would be attributable to the sending party. This potentially limits the risks associated with legal and ethical cross-border data transfers. In terms of data integrity, a Hayekian approach would suggest leaving the decisions to the man on the ground. A solution is simply allowing data analytics companies (any company that performs analytics on datasets. These can be genomic companies, financial companies, legal companies, and so on) to construct their own tiered quality/integrity approach, wherein they charge different amounts depending on a guarantee the company makes related to the quality of the data. Alternatively, data analytics companies can apply the same tiered approach to in-house analytic tools, which can be applied to their in-house datasets. Higher amounts can be charged for higher-quality data or analytic tools. This way, both parties can decide for themselves and rely on contractual provisions and guarantees to regulate their relations.

6 Data security and data privacy

Figures 5–7 present serve to add to the ontology. Each figure outlines important concepts for regulators, ethicists, and the private industry to be aware of. For example, often, mechanisms aimed at data security are mixed up with those that are aimed at data privacy (and *vice versa*). The information is introductory, for conceptual clarity and exploration in further work (Figures 5–7).

7 Conclusion

Avoiding any lengthy conclusions due to complexity and because I wish any readers of this work to draw their own, I will simply conclude with two things. The first is that I have presented a usable and adaptable framework for an open, dynamically stable, information society, which is applicable to regulatory considerations based on trust and conditionality. Second, I have only provided a negative definition of information. My definition of information is hidden within the first part of this conclusion. Information *is reflection*.

Data availability statement

Publicly available datasets were analyzed in this study. These data can be found at: Naidoo (2023a). An (ontogenetic) transcendental (de) materialist theory of subjectivity. DOI: 10.31219/osf.io/cwugk: https://osf.io/preprints/osf/cwugk. Naidoo (2023b). Contradiction and desire. DOI: 10.31219/osf.io/up5f4: https://osf.io/preprints/osf/up5f4. Naidoo (2023c). The hard problem, qualia, agency, intelligence, and Freud. DOI: 10.31219/osf.io/gqspk: https://osf.io/preprints/osf/gqspk. Naidoo (2023d). What does it mean to be an Agent? DOI: 10.31219/osf.io/evna6: https://osf.io/preprints/osf/evna6 (now published on Frontiers at https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2023.1273470/full).

Author contributions

MN: conceptualization, formal analysis, investigation, methodology, resources, writing–original draft, and writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. Work on this article was supported by the U.S. National Institute of Mental Health and the U.S. National Institutes of Health (award number U01MH127690) under the Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) program. The content of this article is solely my responsibility and does not necessarily represent the official views of the U.S. National Institute of Mental Health or the U.S. National Institutes of Health. In addition, support was also obtained from the National Research Foundation (NRF) under the Doctoral Innovation Scholarship.

Acknowledgments

The author would like to thank Amy Gooden for her assistance with the technical editing of this manuscript, as well as the figures and tables included herein.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2024.1290658/full#supplementary-material

References

Althusser, L. (2014). *On the reproduction of capitalism: ideology and ideological state apparatuses*. London: Verso.

Badiou, A. (2006). Logics of worlds: being and event II. *Continuum*.

Bennett, C. H. (1973). Logical reversibility of computation. *IBM J. Res. Dev.* 17 (6), 525–532. doi:10.1147/rd.176.0525

Benveniste, E. (1966). *Problems in general linguistics*. Chicago: University of Chicago Press.

Bernstein, M. N., and Castellano, G. (2020). Information entropy. *Found. Inf. theory Part 2*, 249–257. doi:10.1201/9780429022937-23

Boulton and Watt v Bull (1795). *126 er 662*.

Brownlee, J. (2019). *A gentle introduction to information entropy*. Machine Learning Mastery. Available at: https://machinelearningmastery.com/what-is-information-entropy/ (Accessed August 2, 2023).

Burgin, M. (2010). *Theory of information: fundamentality, diversity, and unification*. World Scientific Publishing Company.

Burgin, M., and Krzanowski, H. R. (2022). World structuration and ontological information. *Proceedings* 81 (1), 1–4. doi:10.3390/proceedings2022081093

Burgin, M., and Mikkilineni, R. (2022). Is information physical and does it have mass? *Information* 13 (11), 1–10. doi:10.3390/info13110540

Burgin, M., Mikkilineni, R., and Phalke, V. (2020). Autopoietic computing systems and triadic automata: the theory and practice. *Adv. Comput. Commun.* 1, 16–35. doi:10.26855/acc.2020.12.003

Burke Inlow, E. (1950). *The patent grant*. Baltimore: The Johns Hopkins Press.

Chalmers, D. (1996). *The conscious mind: in search of a fundamental theory*. Oxford: Oxford University Press.

Chalmers, D. J. (2010). *The character of consciousness*. Oxford: Oxford University Press.

Cryptography (2021). What is indistinguishability obfuscation? Available at: https://crypto.stackexchange.com/questions/44770/what-is-indistinguishability-obfuscation (Accessed July 11, 2023).

Cuff, P., and Yu, L. (2016). Differential privacy as a mutual information constraint. doi:10.1145/2976749.2978308

Damasio, A. (1991). *Somatic markers and the guidance of behavior*. Oxford: Oxford University Press.

Damasio, A. (2005). *Descartes' error: emotion, reason, and the human brain*. New York: Penguin Books.

Davis, S. F., Palladino, J. J., and Christopherson, K. (2012). *Psychology*. London: Pearson.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Deacon, T. W. (2007). Shannon – Boltzmann — Darwin: redefining information (Part I). *Cogn. Semiot.* 1, 123–148. doi:10.1515/cogsem.2007.1.fall2007.123

Descartes, R. (1641). Meditations on first philosophy *(translated by elizabeth S. Haldane)*.

Deutsch, D. (1998). *The fabric of reality*. London: Penguin Books.

Di Giovanni, G., and Livieri, P. (2018). *Friedrich Heinrich Jacobi*. Available at: https://plato.stanford.edu/entries/friedrich-jacobi/ (Accessed October 14, 2022).

Dimitrov, B. G. (2015). *Topological (in) Hegel: topological notions of qualitative quantity and multiplicity in Hegel's fourfold of infinities*. PhD thesis, University St. Kliment Ohridski).

Dolar, M. (1991). I shall Be with you on your wedding-night: lacan and the uncanny. *Lacan uncanny* 58, 5–23. doi:10.2307/778795

Einstein, A., Podolsky B, B., and Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* 47, 777–780. doi:10.1103/physrev.47.777

Floridi, L. (2009). "Philosophical conceptions of information," in *Formal theories of information: from Shannon to semantic information theory and general concepts of information muenchenwiler seminar (Switzerland)*. doi:10.1007/978-3-642-00659-3_2

Freud, S. (1915). The Interpretation of Dreams *(translated by brill, A. A.)*. The Macmillan Company.

Godson, R. (1833). *Law of patents*. Hansard col, 15.

Hartnett, K. (2022). How Shannon entropy imposes fundamental limits on communication. *Quanta Mag.* Available at: https://www.quantamagazine.org/how-claude-shannons-concept-of-entropy-quantifies-information-20220906/ (Accessed January 18, 2023).

Hawkins, J., and Dawkins, R. (2021). *A thousand brains: a new theory of intelligence*. New York: Basic Books.

Hayek, F. A. (1945). The use of knowledge in society. *he American Economic Review* 35 (4), 519–530. http://www.jstor.org/stable/1809376.

Hegel, G. W. F. (2010). Science of logic *(translated by G. Di Giovanni)*. Cambridge: Cambridge University Press.

Hegel, G. W. F. (2018). Phenomenology of Spirit *(translated by T. Pinkard)*. Oxford: Oxford University Press.

Hidalgo, C. (2015). *Why information grows: the evolution of order, from atoms to economies*. New York: Basic Books.

Horkheimer, M., and Adorno, T. W. (1989). *Dialectic of enlightenment*. New York: The Continuum Publishing Company.

Hossenfelder, S. (2014). The remote Maxwell demon. Back ReAction. *arXiv*. doi:10.48550/arXiv.1408.3797

Hossenfelder, S. (2016). The remote Maxwell demon as energy down-converter. *Found. Phys.* 46, 505–516. doi:10.1007/s10701-015-9981-7

Juarrero, A. (2023). *Context changes everything: how constraints create coherence*. Cambridge: MIT Press.

Lacan, J. (1979). The Seminar of Jacques lacan, book XI: the four fundamental Concepts of psychoanalysis *(translated by A. Sheridan)*. W. W. Norton and Company.

Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM J. Res. Dev.* 5 (3), 183–191. doi:10.1147/rd.53.0183

Lock, J. (1860). *An Essay concerning human understanding*. Available at: http://www.philotextes.info/spip/IMG/pdf/essay_concerning_human_understanding.pdf (Accessed January 18, 2023).

López-Suárez, M., Neri, I., and Gammaitoni, L. (2016). Sub-kBT micro-electromechanical irreversible logic gate. *Nat. Commun.* 12068 (7), 1–6. doi:10.1038/ncomms12068

Marx, K., and Engels, F. (1848). "Manifesto of the communist party," in *Marx/engels selected works*. Editors K. Marx and F. Engels (Moscow: Progress Publishers), 98–137.

Mikkilineni, R. (2022a). A new class of autopoietic and cognitive machines. *Information* 13, 24–14. doi:10.3390/info13010024

Mikkilineni, R. (2022b). Infusing autopoietic and cognitive behaviors into digital automata to improve their sentience, resilience, and intelligence. *Big Data Cogn. Comput.* 6, 7–15. doi:10.3390/bdcc6010007

Mokyr, J. (2012). *The enlightened economy: an economic history of britain 1700-1850*. New Haven: Yale University Press.

Naidoo, M. (2023a). *An (ontogenetic)transcendental (de)materialist theory of subjectivity*. doi:10.31219/osf.io/cwugk

Naidoo, M. (2023b). *Contradiction and desire*. doi:10.31219/osf.io/up5f4

Naidoo, M. (2023c). *The hard problem, qualia, agency, intelligence, and Freud*. doi:10.31219/osf.io/gqspk

Naidoo, M. (2023d). What does it mean to be an Agent? doi:10.31219/osf.io/evna6

OpenCSF (n.d.). *Consensus in distributed systems*. Available at: https://w3.cs.jmu.edu/kirkpams/OpenCSF/Books/csf/html/DistConsensus.html (Accessed May 17, 2023).

Plato (2002). The republic. Available at: https://www.sciencetheearth.com/uploads/2/4/6/5/24658156/plato_-_the_republic.pdf (Accessed May 17, 2023).

Poincaré, H. (2010). Papers on Topology: analysis Situs and its five Supplements *(translated by J. Stillwell)*. American Mathematical Society.

Popper, K. (1945). *The open society and its enemies*. London: Routledge.

Popper, K. R. (1962). On the sources of knowledge and of ignorance. *Philosophy Phenomenological Res.* 23 (2), 292–293. doi:10.2307/2104935

Prigogine, I. (1978). Time, structure, and fluctuations. *Science* 201 (4358), 777–785. doi:10.1126/science.201.4358.777

Rosen, W. (2010). *The most powerful idea in the world: a story of steam, industry, and invention*. London: Jonathan Cape.

Rowlands, P. (2007). *Zero to infinity: the foundations of physics*. World Scientific Publishing Company.

Sapolsky, R. M. (2017). *Behave: the biology of humans at our best and worst*. New York: Penguin Books.

Schrödinger, E. (1944). *What is Life?* Cambridge: Cambridge University Press.

Schrödinger, E. (1967). *What is life? The physical aspect of the living cell with mind and matter & autobiographical sketches*. Cambridge: Cambridge University Press.

Shannon, C. E. (1940). *A symbolic analysis of relay and switching circuits PhD Thesis* (Massachusetts Institute of Technology). Available at: https://dspace.mit.edu/handle/1721.1/11173.

Shannon, C. E., and Weaver, W. (1964). *The mathematical theory of communication*. Illinois: The University of Illinois Press.

Sherman, B., and Bently, L. (2003). *The making of modern intellectual property law: the British experience, 1760–1911*. Cambridge: Cambridge University Press.

Spinoza, B. (1995). Spinoza: the letters *(translated by S. Shirley)*. Indianapolis: Hackett Publishing Company.

Sproll, P. A. C. (1994). Matters of taste and matters of commerce: British government intervention in art education in 1835. *Stud. Art Educ.* 35 (2), 105–113. doi:10.2307/1320824

Taleb, N. N. (2007). *The black swan: the impact of the highly improbable*. New York: Penguin Books.

Thaldar, D. (2017). The open society: what does it really mean? *De. Jure* 50 (2), 396–405. doi:10.17159/2225-7160/2017/v50n2a11

Thaldar, D. W., Townsend, B. A., Donnelly, D. L., Botes, M., Gooden, A., van Harmelen, J., et al. (2022). The multidimensional legal nature of personal genomic sequence data: a South African perspective. *Front. Genet.* 13, 997595. doi:10.3389/fgene.2022.997595

Timpson, C. (2004). *Quantum information theory and the foundations of quantum mechanics*. Oxford: University of Oxford. [PhD thesis].

Timpson, C. (2008). "Philosophical aspects of quantum information theory," in *The ashgate companion to the new philosophy of physics*. Editor D. Rickles (London: Routledge), 197–261.

Timpson, C. (2013). *Quantum information theory and the foundations of quantum mechanics*. Oxford: Oxford University Press.

Vopson, M., and Lepadatu, S. (2022). Second law of information dynamics. *AIP Adv.* 12, 1–7. doi:10.1063/5.0100358

Webster, T. (1853). *On property in designs and inventions in the arts and manufactures*. Chapman & Hall.

Wiener, N. (1961). *Cybernetics: or the control and communication in the animal and the machine*. 2 ed. Cambridge: MIT Press.

Woodward, J. (2000). Explanation and invariance in the special sciences. *Brit. J. Phil. Sci.* 51, 197–254. doi:10.1093/bjps/51.2.197

Yackel, R. (2021). What is homomorphic encryption, and why isn't it mainstream? *Key Factor*. Available at: https://www.keyfactor.com/blog/what-is-homomorphic-encryption/ (Accessed May 4, 2023).

Yoshida, R. (2010). Self-oscillating gels driven by the Belousov-Zhabotinsky reaction as novel smart materials. *Adv. Mater.* 22 (31), 3463–3483. doi:10.1002/adma.200904075

Žižek, S. (2003). *The puppet and the dwarf: the perverse core of christianity*. Cambridge: MIT Press.

Žižek, S. (2008). *The ticklish subject: the absent centre of political ontology*. London: Verso.

Žižek, S. (2014). The most Sublime hysteric: Hegel with lacan *(translated by T. Scott-railton)*. Cambridge: Polity Press.

frontiers | Frontiers in Pharmacology

TYPE Hypothesis and Theory
PUBLISHED 17 June 2024
DOI 10.3389/fphar.2024.1331237

OPEN ACCESS

EDITED BY
Simisola Oluwatoyin Akintola,
University of Ibadan, Nigeria

REVIEWED BY
Marietjie Wilhelmina Maria Botes,
University of KwaZulu-Natal, South Africa
Dmytro Chumachenko,
National Aerospace University—Kharkiv Aviation Institute, Ukraine
Pier Luigi Gentili,
Università degli Studi di Perugia, Italy

*CORRESPONDENCE
Meshandren Naidoo,
214549331@stu.ukzn.ac.za

RECEIVED 31 October 2023
ACCEPTED 20 May 2024
PUBLISHED 17 June 2024

CITATION
Naidoo M (2024), Open optimism as an "embodied-health" ethic for the information era.
Front. Pharmacol. 15:1331237.
doi: 10.3389/fphar.2024.1331237

Open optimism as an "embodied-health" ethic for the information era

Meshandren Naidoo*

School of Law, University of KwaZulu-Natal, Durban, South Africa

This article forms part of a series on "openness," "non-linearity," and "embodied-health" in the post-physical, informational (virtual) era of society. This is vital given that the threats posed by advances in artificial intelligence call for a holistic, embodied approach. Typically, health is separated into different categories, for example, (psycho)mental health, biological/bodily health, genetic health, environmental health, or reproductive health. However, this separation only serves to undermine health; there can be no separation of health into subgroups (psychosomatics, for example). Embodied health contains no false divisions and relies on "optimism" as the key framing value. Optimism is only achieved through the mechanism/enabling condition of *openness*. Openness is vital to secure the embodied health for individuals *and societies*. Optimism demands that persons become active participants within their own lives and are not mere blank slates, painted in the colors of physical determinism (thus a move away from nihilism—which is the *annihilation* of freedom/autonomy/quality). To build an account of embodied health, the following themes/aims are analyzed, built, and validated: (1) a modern re-interpretation and validation of German idealism (the crux of many legal–ethical systems) and Freud; (2) ascertaining the bounded rationality and conceptual semantics of openness (which underlies thermodynamics, psychosocial relations, individual autonomy, ethics, and as being a central constitutional governmental value for many regulatory systems); (3) the link between openness and societal/individual embodied health, freedom, and autonomy; (4) securing the role of individualism/subjectivity in constituting openness; (5) the vital role of nonlinear dynamics in securing optimism and embodied health; (6) validation of arguments using the methodological scientific value of *invariance* (generalization value) by drawing evidence from (i) information and computer sciences, (ii) quantum theory, and (iii) bio-genetic evolutionary evidence; and (7) a validation and promotion of the inalienable role of *theoretic philosophy* in constituting embodied health, and how modern society denigrates embodied health, by misconstruing and undermining theoretics. Thus, this paper provides and defends an up-to-date non-physical account of embodied health by creating a psycho-physical–biological–computational–philosophical construction. Thus, this paper also brings invaluable coherence to legal and ethical debates on points of technicality from the empirical sciences, demonstrating that each field is saying the same thing.

KEYWORDS

information, freedom, autonomy, artificial intelligence, quantum mechanics, bioethics, non-linear dynamics, health

1 Introduction

In 2023, Geoffrey Hinton, better known as "the Godfather of artificial intelligence," given his role in constructing the regressive function, quit his job at Google, citing the immanent dangers of the recent advances in the field. Hinton's basic argument is that humanity will end up being slaves to AI, as AI is soon to supersede human intelligence.

Parsimony describes an action, or mode of acting, with the least amount of resource expenditure, which requires the least number of assumptions being made, prior to any act or explanation. In other words, parsimony describes a chosen trajectory or path that is selected based on efficiency. Efficiency thus involves the choosing of a trajectory that involves the least amount of risk, uncertainty, or resistance (as resistance requires more resource expenditure). Hinton understands the concept of parsimony well, given that movements of efficiency are mirrored by the mechanisms of backward propagation/regressive function (below), which are implemented in gradient descent algorithms. Freud's (now validated) economic theory of mind is based on the notion of energy efficiency. See "Freud" and "Sedimentation and entrenchment" given in Supplementary Material.

To alter Hinton's claim slightly, Hinton seems to claim that humans and societies are slaves to *intelligence*—which is not merely a "human" phenomenon. Intelligence, in terms of Hinton's claim, is reflected by the linear *movement of efficiency*. Efficiency is the process of reducing the degrees of freedom and degrees of meta-freedom (Naidoo, 2023a). This movement *creates linearity*. The way in which nihilism arises and is presented in societies is explored in this paper. As an introduction to the concept of nihilism, typically, nihilism arises (in philosophic presentation) because of linear dynamics, in the form of (1) "something" that linearly arises from the *void* of nothingness or (2) something that proceeds in a linear movement toward nothingness, as the *abyss,* or endpoint. The solution to (1) is simply to understand that *quantity* can only exist by being derived from and attached to some *quality*—(no)thing is not nothing (Naidoo, 2023e). Quality is theory itself or subjective rationalizations, hence the fundamental need to protect subjective freedoms and meta-freedoms. The solution to (2) is to secure meta-freedom, which is the ability to construct/define and reconstruct/redefine subjective qualities, as opposed to promoting quantity, as done by the society. Quality can also be understood as the production of creative or critical thought (abstract thinking), whereas quantity can be understood as the production of said quality (in the form of objects) and consumerism.

2 Stochastics and human biology

Stochastic reasoning/planning is a process used to account for risk and uncertainty when outcomes are unpredictable. One can map risks according to qualitative and quantitative stochastic modeling. Simulation generators (like Monte Carlo generators) are used to model many alternative sample paths/histories. These are modeled paths and not just outcome predictions (Taleb, 2004). The concept of "paths" embodies a wider range of contingencies, in comparison to a direct outcome analysis. Path analyses involve the *qualitative* study of sequential information, within any scenario, for all possible paths, over a certain period. One can uncover what outcomes are (im)possible, what event sequence leads to which outcomes, and possible stopping points upon path progression, including information as to which stops affect an outcome (and in what way). Although humans are poor at learning from history, alternative histories through stochastic modeling improve risk management strategies by rendering them antifragile (Taleb, 2004). These models highlight possibilities and omissions. Risk strategies can be built by comparing ratios and other *qualitative inferences* (Taleb, 2004).

Stochastic modeling occurs naturally in human biology. The prefrontal cortex (PFC) is part of the cortical system and the "emotional" limbic system (Sapolsky, 2017). The PFC is divided into (1) the dorsolateral PFC (PFCDL) and (2) the ventromedial PFC (PFCVP). The PFCDL is the rational, cognitive, utilitarian, and unsentimental decision-maker. It is also the last region of the brain to fully mature. The PFCVP is concerned with the emotional aspects of decision-making. Decisions and thoughts are thus intermingled with the "emotional" limbic system (Sapolsky, 2017). Both regions run real-time simulations of alternative histories. The PFCDL is concerned with utilitarian outcomes, while the PFCVP is concerned with the subjective "how would I feel." This is responsible for that intuition one has about a course of action. The "correct" course of action resolves around the negation (or repression) of "failed" or unsuitable *non-existing*, but possible alternative histories, which are produced through the simulations. These simulations are a mechanism for determining the most *efficient* routes of risk management. This is thus a *repression of nothing,* as Freud noted about the unconscious (it is the repression of the fact that there is nothing to repress, which constitutes the unconscious). Details on "Freud" are given in Supplementary Material. It is a difficult concept to understand and accept, but it is trite that history runs forward and not backward (Taleb, 2007), which is the movement of regression functions, as mentioned above. The author has provided a navigational schema for the reader to follow demonstrating the various levels of interconnectivity and supplemental explanations in Supplementary Material.

3 Frame axioms: epistemology and computation

The frame axiom problem in programming involves how to frame or reason about problems—and also describes the problems in determining what changes, and does not change, as a consequence of certain events or actions. Frame axioms determine methods of making assumptions about the world, thus enabling agents to make predictions regarding actions and possible consequences (Allen and Ferguson, 1994). There are two issues to consider when deconstructing problems and predictions: (1) the epistemological problem and (2) the computational problem.

The epistemological problem involves the kinds of assumptions made about the world. The computational problem concerns the issues involved in determining how to compute and use those epistemological assumptions in a formulism (Allen and Ferguson, 1994). In terms of the epistemological problem, for example, intelligent design programmers must decide whether to make assumptions about any changes in properties or event

occurrences. Computational problems, on the other hand, involve the issues in determining which kinds of techniques should be used in a model to implement those epistemological assumptions. Computational problems typically involve the use of explicit axioms, such as the situation calculus approach or the explanation closure technique.

4 Planning, prediction, and explanation

Sense (coherence) making and reasoning require three concepts: (1) explanations; (2) planning; and (3) predictions. Planning means having an initial description of a world or context and a desired goal. Actual planning describes finding a course of action that will be conducive to a goal. Explanations involve finding the best-fit system or model, which best fits the sets of observational data (Allen and Ferguson, 1994).

Plan recognition involves the prediction of an agent's top-level plans based on observing its actions. It is an *abductive reasoning task*, where plans are inferred that best explain observed actions. First-order logic is often used for plan recognition, but these methods cannot handle uncertainty in data. The other option is probabilistic graphic models, but these cannot handle representations (ScienceDirect, n.d.). It is a problem when one observes other agents' actions (data points) and wishes to obtain the agents' plan in order to construct an explanation. This is known as the *plan recognition problem*.

As described in "The brain" in Supplementary Material, prediction is the process modality for world-building and correction (Hawkins and Dawkins, 2021). Prediction involves "foreseeing" or predicting the effects of new actions and events that will occur and then updating the model as required (Allen and Ferguson, 1994). If temporal logic is used, the world model will likely contain some information relating to the past or future actions and events. Prediction requires the creation of sequences of events or experiences based on the world models created and the information contained therein. Planning requires prediction to determine what actions will accomplish the set goals. Planning can be decomposed into (1) generating a set of candidate actions and (2) evaluating whether these actions will be successful (Allen and Ferguson, 1994). Explanations also require prediction. Explanations can be decomposed into the generation of a possible set of events that might explain observations, followed by a verification of whether the said events would cause the observed effects (Allen and Ferguson, 1994). Both thus would work on a predictive model to generate these steps.

Prediction is a probabilistic graphic model (Allen and Ferguson, 1994). Predictive information describes information within channel input about the channel output. Hence, this is a temporal movement "forward" in time. This information then relates to objects or events that do not exist at the time when the information is processed. Restorative information, on the other hand, is information contained in the channel output about the input (hence a "backward" temporal movement). This information relates to something that, at the time of processing, does not exist (in a specific form) as it did prior to the processing. A non-symmetric channel can have different effects on different symbols (relative), and the outputs can change. In these kinds of systems, there can be large distribution changes (because the physical processes that induce transmission are different for different symbols), including average abstract information quantity gains being made if the channel is used frequently.

Simple models, using a standard backward chaining planning algorithm, function by chaining backwards from the goal state (regression). The starting point is the goal state, which is compared to the initial state. This is followed by using a set of propositions, which differ in truth value between the states. Then, an action is enacted, which results in the obtaining of one of those propositions. The state of the world prior to the enacted action is computed using regression (which inverts the add and delete lists in an action definition). The new state now becomes the goal state, and the process continues until the initial state is derived. After this process, the algorithm now has a sequence of actions that lead from the initial state to the desired goal state. The predictive model here functions according to two broad assumptions: (1) no other events or changes occur in the world except for planned actions and (2) the action definitions completely describe all changes that occur as the result of the action (Allen and Ferguson, 1994). Prediction can be accomplished using these two steps. Regression techniques were designed specifically to exclude an explicit prediction step because of the assumptions (related to the incomplete absolute discussed below). This technique allows for the regression of an operator from state B to state A (as the preceding state) and the guarantee that predicting from state A with the same actions would yield state B. Using this, a plan can be constructed *in a backwards fashion*. Once a plan is found, it will achieve the said goals.

5 Explanation closure

Explanations are deductions, based on axioms, which assist with "sense-making" or reducing uncertainty. There are many difficulties in designing artificial intelligence systems and explanation axioms. When dealing with predictive models (the human brain too operates via prediction [Hawkins and Dawkins, 2021]), designers are faced with the problem of assumptions (epistemology, which is necessary in any design) that do not work. Reasoning processes are far more complex, which simply follow designed-in assumptions.

Situation calculus and temporal logic approaches, on the other hand, do not require those operational assumptions. These theories of logic do not commit to how states resulting from actions relate to states before actions. Instead, the properties of the resulting state must be specified through axioms; the framing problem revolves around how best to specify these properties (Allen and Ferguson, 1994). The first approach is to use explicit frame axioms, where each axiom is stipulated, each of which describes which properties are not to be changed by actions. However, this is an impossible task because there are too many axioms one would need to create (Allen and Ferguson, 1994). To overcome this, the frame axiom approach was relegated in intelligent designs.

The solution to the epistemological problem was to build models that utilize persistence or inertia assumptions. Here, the assumption is that all changes caused by an action are specified, and all properties not asserted to change do not change (Allen and Ferguson, 1994). This is an undesirable approach because if there is uncertainty about whether a property might change, the approach will erroneously assume that the property does not change. Other approaches focus on minimizing property changes or have

constraints imposed on temporal ordering or properties or causal relationships. However, these approaches are problematic because they cannot handle simultaneous actions or external events well. Assumptions that are rather *based on events* lead to a more intuitive characterization of problems, wherein the logic would be related to an intuitive fact about the world (Allen and Ferguson, 1994). This approach handles a wider range and more complex problems.

The better approach is to *specify for each property what events can change it*, instead of trying to specify a host of different actions that could change a property (Allen and Ferguson, 1994). This reduces the problem of reasoning about changes to instead *reasoning about what events may or may not occur*. This includes both "external" events and actions of the agent itself. This is called *explanation closure*. Assumptions are still present in this technique but enable a large reduction in the number of frame axioms required to produce workable sets of axioms for a problem (Allen and Ferguson, 1994). Event-centric approaches to physics have gained ground recently, showing great potential to solve issues within physics. To observe a physical state associated with the probabilities generated from quantum mechanics is impossible because one needs an uncountable set of identically prepared quantum states and measurement apparatus to perform continuous measurements. Every bit of information collected from the environment is a consequence of discrete interactions between material objects—these interactions are called events. Instead of the notion of "particles" or "fields," events offer different axioms for physical systems. Within an event-centric paradigm, the continuous evolution of particles/fields is replaced with the discrete evolution of causal networks of events (Powers and Stojkovic, 2023).

Explanation closure axioms are a method of treating events on a case-by-case basis. This allows for idiosyncratic events and properties to be represented, even though they do not fit the norm. These approaches code solutions into the axioms themselves. Some have dubbed this "cheating" since it provides the agent with some explicitly encoded assumptions that would make the representation work (Allen and Ferguson, 1994). However, this is part of a common-sense logic of the world that agents would require. They are also problem-independent. Furthermore, it is not accurate to suggest that all these axioms need to be "programmed in" by coders—they can in fact be generated automatically during action or "on the fly" (Green, 1969), which means that they are *separable from the formulism*, unlike other approaches (Allen and Ferguson, 1994). This is known as contextual import. Mechanisms of backward propagation loops or regressive functions include the processes of *iteration* or *recursion*. Iteration is a context-independent constraint that *feeds back information from the output* of one trial run *into the initial conditions of the next run*. Hence, iteration acts as a temporal constraint, *which alters the probability of the next output*. The iteration process is repeated. *Recursive iteration* is a process wherein full sequences are fed back on themselves. This form of looping results in processes and sequences becoming *self-referential*. When the last step of a sequence feeds back into itself *to become the first step of the next iteration, a self-referential configuration* with *non-linearities* is created. The latter introduces multiscale and multidimensional interdependencies (Juarrero, 2023). Thus, iteration and recursion are processes that import/incorporate meaningful information from contexts/the world back into the system through the alteration of the weights of the middle-layer connections in the algorithm. The system can thus become more suitable and calibrated to its context. This also allows for *qualitatively novel* results/features to emerge. Contextual import *enables representation to be possible*. Properties here only change upon the occurrence of certain events. Furthermore, importantly, these assumptions do not need to be correct; where wrong or false, they also need to be made explicit in the representations (Allen and Ferguson, 1994).

The solution to the computation problem (or what mechanism can be used to make assumptions) is to either (1) explicitly add axioms that encode all assumptions or (2) use a nonmonotonic model theory that defines a new notion of entailment, including the assumptions (Allen and Ferguson, 1994). There are many ways to do this, with each having positives and negatives (which are mostly reduced to the ease at which a formulism can be achieved). Explanation closure axioms, on the other hand, allow for a flexible system that can handle complex issues in representing actions. The representation that results from this will operate in standard first-order logic, thus making it relatively easy to determine if consequences follow from axioms. Furthermore, the handling of exceptional cases does not require extending language syntax to include special predicates, which complicates the reasoning process and can lead to unintuitive formalisms (Allen and Ferguson, 1994).

6 Hegel's theology: events and virtuality

Hegel's radical re-interpretation of Christian theology involved the birth of Christ as God only being able to recognize his own existence through the Othering of himself (as the form of Christ) (Žižek et al., 2011). The death of Christ on the cross symbolized the death of God (the transcendental God, or the Platonic God) or his belief in himself. Hegel, in forming his account of embodiment, merged the transcendental and materialist positions—the Absolute is to be understood as both substance and subject. Žižek (2003) elucidates this position and calls for the abandonment of the traditional view of Hegelian Spirit:

> "The point this reading misses is the ultimate lesson to be learned from the divine Incarnation: the finite existence of mortal humans is the only site of the Spirit, the site where Spirit achieves its actuality . . . Spirit is a virtual entity in the sense that its status is that of a subjective presupposition: it exists only insofar as subjects act as if it exists. Its status is similar to that of an ideological cause like Communism or Nation: it is the substance of the individuals who recognize themselves in it, the ground of their entire existence, the point of reference which provides the ultimate horizon of meaning to their lives, something for which these individuals are ready to give their lives, yet the only thing that really exists are these individuals and their activity, so this substance is actual only insofar as individuals believe in it and act accordingly. The crucial mistake to be avoided is therefore to grasp the Hegelian Spirit as a kind of meta-Subject, a Mind, much larger than an individual human mind, aware of itself: once we do this, Hegel has to appear as a ridiculous spiritualist obscurantist, claiming that there is a kind of mega-Spirit controlling our history . . . This holds especially for the Holy Spirit: our awareness, the (self)consciousness of finite humans, is its only actual site . . . although God is the

> substance of our (human) entire being, he is impotent without us, he acts only in and through us, he is posited through our activity as its presupposition."

Hegelian Spirit is thus a *virtual entity*, whose existence actualizes (or becomes) upon the recognition within the registers of the subjects—via belief, for example. In this way, the *virtual becomes actual* within subjective beliefs; however, that actualized virtual is not reducible back to the said subjective beliefs. The virtual event is constituted by the subject as such—because *beliefs are a response* to events (they occur after events). *Subjects interpret events as (to be) events.* The critical question here is whether events are only events upon recognition and registration by subjects.

Some have argued that the ability of subjects to register events as such is not a requisite for events to be events in themselves (Žižek et al., 2011). The counter argument is that events themselves are instantiated into things as part of the virtuality of the substance itself, which, when expressed, are then registered by subjects and their beliefs. Considering this, the Spirit would be that which arises in response to events and performs an interpretation after the occurrence of the event. Žižek's position is that of a "subjective presupposition"; however, the counter argument advocates that events can be constituted without a subject. The virtuality of events may be registered by subjects, but it can also be registered in the energies of "things" and, hence, may be "felt" in things themselves. The virtual, in this way, is also substance and subject, which is *ex post* realized as such by subjects, who then name it as such. Thus, in the Hegelian ontology, God takes the *form of events*.

7 Dialectical discrete event-centric physics: events and observers

Events can be construed to be the building blocks of spacetime. In the current relativity theory, events are understood to be discrete units of volume in spacetime. Importantly, an event-centric approach allows for the mathematical structures in general relativity to be constructed using *discrete elements*. This offers solutions to the two most important issues in modern physics and, hence, can be used to construct a new interpretation of physics (Powers and Stojkovic, 2023).

Primitive network elements can form part of a bigger causal network of events. These primitive network elements consist of two events, which can share direct or indirect causal connections. A direct causal connection implies that they are related through a third event. Indirect causal connections, on the other hand, correspond to experiments involving entangled particles, such as the Bell test (Powers and Stojkovic, 2023).

Observers are to be construed as entities capable of assembling information about events (Naidoo, 2023a; see "Topology and spacetime" of Supplementary Material). Observers can use this information to construct physical models, which can be used to infer properties of future events. To obtain information about events, an observer must participate in the said event, which means that events must have a structure. This structure is a one-part system and one-part observer (Powers and Stojkovic, 2023). This means that an observer and an event are *reciprocally constitutive* (Naidoo, 2023f). Thus, as mentioned above, events are only events as such upon registration in the registers of subjects. This accords with the QBism (Naidoo, 2023e) interpretation of quantum mechanics, whereby quantum states (as a navigational tool) are interlinked with subjective beliefs regarding the outcomes of experiments (Caves et al., 2002). This thus involves an inherent interlinking with a (subjective) Bayesian (and surprise-oriented) approach to probability (Naidoo, 2023a; Naidoo, 2023f). Bayesian probabilistic inferential reasoning is a highly accredited theory regarding human reasoning in the cognitive and sensory domains (Pouget et al., 2013).

The notion of observers is thus important and should be part of any model of physics being proposed (Powers and Stojkovic, 2023). The act of measurement ought to be understood as the performance of an experiment and can be described as the revelation of the pre-existing state of the system under the study. "System" is to be understood as two causally connected events (below), which has a wide range of implications. The kind of event-centric view used in the study (Powers and Stojkovic, 2023) is naturally implied by quantum contextuality and is also related to Bell's inequality.

8 Maps and events

Maps are to be used as the basis for physically observable variables (see "The brain" and "Topology and spacetime" in Supplementary Material). For a single event, a map can be used to generate a second event in a way that the important aspects of the causal relationship between those events are encoded within the map (Powers and Stojkovic, 2023). If taken independently, however, neither the initial event nor the map will contain enough information to completely determine the second event. The information is stored in the causal relationship between two events, which are related through the map. To summarize, a map thus connects two events.

Novel properties can emerge from this causal relationship, and these novel properties can be distinguished from those associated with events and maps using the notion of *locality* (Powers and Stojkovic, 2023). Locality describes properties that are associated with either event 1 or event 2 or of the map itself (see "Quantum theory" in Supplementary Material). Locality thus describes *non-emergent properties*. *Non-locality*, on the other hand, describes properties which are emergent from the causal relationship itself. These are *non-local degrees of freedom*. Events can be described as measurements that are performed by observers. In other words, events are only as such, when they are inscribed into the registers of subjects, as observers (Naidoo, 2023a). The nature of this registration thus depends on how, or the type of rationalization/explanatory theory/interpretation, the subject affords to the event.

9 The Absolute

The "Absolute" is a concept originating from German philosophy and is a general term used to describe the metaphysical conception of a fundamental "totality." The Absolute is thus that which is self-sufficient, meaning the "thing" upon which all other things depend, but which itself supposedly

depends on nothing outside of itself (Žižek et al., 2011). The Hegelian conception of the Absolute aimed to overcome dualisms, especially those in Kantian philosophy (such as phenomenal versus noumenal worlds), by providing a metaphysical grounding for these dualisms. The Hegelian conception of the Absolute thus was a unifying metaphysical ground for all knowledge (Moyar, 2017). In legal–political terms, the Absolute corresponds with the governing state, and its corresponding laws, which delineates the bounds of freedom for persons.

The concept of the Absolute as such was ambiguous, in that it undermined nihilism, while also threatening to be nihilistic. In the first sense, it undermined nihilism as non-linearity was implied, thus simultaneously undermining the concepts of "finite" and "finitude." In the second sense, it is nihilistic in itself as it seemingly eliminates "free subjectivity" because it implies absolute predetermination (and not independence) by an Absolute totality (Moyar, 2017). This means that there would be no individuality, free choice, or free will, and the finite would be illusory and be akin to the Kantian "puppet-on-a-string" (Žižek, 2003). Hegel's solution was thus to construct a theory wherein finite subjects, and finite objects, are positive "moments" in the Absolute. Thus, Hegel posits the Absolute as being an *internally differentiated whole* (Moyar, 2017).

10 The incomplete Absolute

During its unfolding, the Hegelian Absolute passes through a process of its own becoming, which involves the Absolute realizing that a part of itself must always remain beyond itself. The Absolute thus sublates its otherness, in its identity. Gabriel (2011) noted how this also means that the Absolute is simultaneously finite and infinite:

> "The absolute idea is only grasped in the context of a theory of self-constitution of logical space, i.e. of the concept in an eminent singular . . . The answer, therefore, to the question: how does the infinite become finite? is this: that there is not an infinite which is first of all infinite and only subsequently has need to become finite, to go forth into [herausgehen] finitude; on the contrary, it is on its own account just as much finite as infinite."

Gabriel (2011) noted that

> "This movement of the negation of negation is precisely what takes place in the chapter on 'the Absolute' in the Logic, the introduction and first subchapter (A) of which proceed in three steps. First the absolute is determined as absolute transcendence, or as absolute identity which outstrips our conceptual capacities. It can only be paradoxically determined by the negation of all predicates. Second this movement, which is a movement of reflection, is made transparent as reflection. In order to steer clear of the problem of absolute transcendence, the finite is determined as an image of the absolute, which has being far more than any finite being due to its pure positivity, a position Hegel ascribes to Spinoza . . . Third this whole movement is presented as a process by which we eventually arrive at the form determination of the absolute form, where form and content of reflection coincide in the 'self-exposition' of the absolute, i.e. in the reflection of reflection."

This means that the Absolute cannot exist as a preconceived/presupposed substance prior to the process of its own becoming; hence, the Absolute retroactively posits/manifests itself through the logical process of self-negation and the negation of determinate concepts. Hence, it is a retroactive position of the Absolute from a logical space (Gabriel, 2011). This Absolute is both the form and the content of logic, and it is not something prior to the manifestation of itself in logical thought. This is the logic of the "return of the repressed," which Freud later expressed, that constituted the unconscious as discussed.

This kind of unfolding is thus a movement of *pure thought*, which Hegel (2010) described in the *Science of Logic*. This movement of pure thought means that the Absolute is just a grounding concept that makes the finite intelligible, but the Absolute itself does not have any content other than being the grounding that separates two relata (and, thus, two entities). Why is this so? Because of reflection. By reflecting on a transcendental Absolute, reflection must think "beyond itself" (Gleason, 2021). Hence, reflection must then negate all predicates (of existence). However, by doing this, reflection just creates the Absolute as a transcendent only insofar as the Absolute cannot be determined by predicates. The Absolute in this way is undermined because it would then be determined as that which cannot determine it. Hence, it would not be a true Absolute since it is dependent in this way. Hegel then moves from external reflection to Absolute reflection, the result of which posits the difference as being internal to the Absolute, and not external to it. It is this very movement that constitutes the Absolute. Žižek (2014) explains this as follows:

> "This, then, is the dialectical process: an inconsistent mess (first phase, the starting point) which is negated, and through negation, the Origin is projected or posited backwards, so that a tension is created between the present and the lost Origin (second phase). In the third phase, the Origin is perceived as inaccessible, relativized – we are in external reflection, that is, our reflection is external to the posited Origin which is experienced as a transcendent presupposition. In the fourth phase of absolute reflection, our external reflexive movement is transposed back into the Origin itself, as its own self-withdrawal or decentring. We thus reach the triad of positing, external reflection, and absolute reflection."

This means that the Absolute is internally incomplete. The transcendent is internal to the immanent as the Absolute is always already beyond itself. In fact, the Absolute is constituted precisely by its own failure to fully grasp itself. The expanse between the transcendent and the immanent is internal to the constitution of the immanent–transcendent, as noted by Gabriel (2009) on this–"The crucial point of Hegel's dialectic of the absolute is that metaphysical reflection must not be external reflection. We cannot determine the absolute as absolute substance ontologically anteceding our conceptualization of it. Therefore, reflection has to become absolute, i.e. self-referential."

The Absolute self-discovers itself as infinite, and in doing so, this is simultaneously how it limits itself and surpasses its own self-imposed limits. Hence, there are parts of the Absolute that are unknown to the Absolute itself. The infinite is then the Absolute's ongoing process of self-constitution, which is not determined by anything external to the process itself. This means that there is a realization of the overall coherence of the overall movement. Simply put, consciousness discovers that it is the Absolute itself that moves between self-transcendence and identity (Gleason, 2021). Absolute reflection then is the difference between the immanent and transcendent as being internal to the Absolute itself. It operates in such a way that it was the Absolute itself that distanced itself from itself. The Absolute is thus the oscillation of thought between indeterminacy and determinacy (and *vice versa*) in time (Gleason, 2021). The immanence of the Absolute is, thus, not exclusionary to the world of experience (see "Retroactivity and logic" in Supplementary Material for a noteworthy summative quote).

11 Epistemic constraints and possibility

The study accounts for the seemingly non-deterministic property of nature through enforcing an *epistemic constraint on observers*. The information, which observers are permitted to have about the causal relationship between two events, is limited to *local counts* (for information on locality, see Naidoo, 2023c). This implies that the information stored in the ordering of symbols (see "Symbols" in Supplementary Material), within events and maps, is *hidden* from observers. This accords with Hegel's assertion about the *content hidden in the form* (Naidoo, 2023b) (see "Content hidden in form" in Supplementary Material). This hidden information is that of qualitative meta-degrees of meta-freedom, which amounts to the freedom to interpret and re-interpret (Naidoo, 2023a; Naidoo, 2023b).

Under this constraint, *one can interpret a particular choice of local counts* as specifying the macroscopic state of the causal network, which is analogous to a quantum state (Powers and Stojkovic, 2023). For any macroscopic state, there will typically be many associated microscopic states, each of which is modeled by a unique sequence. These microstates represent possible ontic states (or real physical states) of the underlying causal network. This epistemic constraint secures an *irreducible* ambiguity, which, in turn, ensures that inferential and interpretative capacity is non-eliminable (Naidoo, 2023a). In other words, shades of freedom lie within *qualitative theorization,* as opposed to quantitative application (Naidoo, 2023a). How is an irreducible ambiguity constructed? The solution that Naidoo (2023a) presented is that of *meta-difference*. This will be explored in detail in further work; however, ambiguity does express itself in the following ways: (1) the Kantian transcendental apperception/imagination (Naidoo, 2023d); (2) the incomplete Absolute as explained above; (3) Darwin's utter extinction (see "Utter extinction" in Supplementary Material); and (4) the Kantian infinite/sublime judgement (below).

12 Ambiguity, evolution, and novelty

Context-independent constraints are preset configurations concerning dimensions and possibilities. They (context-independent constraints) are limits to multivariate/multidimensional landscapes. The boundaries and contours form the initial conditions that represent the context-independent constraints. These context-independent constraints bias the direction of energy flow, but factually, they do not *strictly determine from the outset which, of the multiply realizable alternatives within that possibility space, will be realized*. It is thus *not true* that context-independent constraints *are determinative of identities or outcomes*. For example, context-independent constraints mapped one-to-one with a given phenotype would render it impossible for any variation, specification, or individualization to occur other than "by-chance" mutations. Correlatively, the cosmos would have never been able to evolve in the past or at present. The cosmos would never reach the complexity we know it to reflect today (Juarrero, 2023).

For the kind of complexity we observe today, the initial constraints must be vague/ambiguous in terms of their scope/reach. In other words, the initial constraints must not be fixed and must rather be flexible or in flux. Context-independent constraints are thus flexible and contain "feasibility regions" wherein subsequent constraints can interact—similar to a stage whereby a variety of different narratives or plotlines play out/simulate (Juarrero, 2023).

Furthermore, maintenance of any kind of dynamic equilibriums or homeostasis (in terms of biological bodies) requires a continuous balancing and re-adjustment of bodily properties *in lieu* of the context (Juarrero, 2023). A balancing of this sort requires flexibility. Lastly, even though context-independent constraints bring systems into non-equilibrium, they do not produce complexity or persistent structures or dynamics. They also cannot transmit complex messages (although they can aid fidelity of the transmission of communication systems). Hence, they cannot be foundational for complexity formation. How then does complexification arise, and what does flexibility even mean?

First, flexibility simply means that context-independent constraints must be *ambiguous, vague*, and *allow for multiple realizability* (of the various hills and ridges of the epigenetic landscape). What is necessary then to achieve flexibility are *context-dependent constraints* that enable context adaption, maintenance of dynamic equilibrium, and specific realization (Juarrero, 2023). In terms of complexity, the same reasoning applies. Complexity in biology, cosmology, or social systems requires two kinds of constraints, i.e., context-independent and context-dependent ones. Both operate concurrently and do interact with one another and as a unit.

Within the evolutionary discourse, this is represented by Waddington's landscape with hallows and hills, which arise from different, multiply realizable constraints.

According to Waddington, who based his epigenetic topographic-landscape theory on dynamical system theory, different dimensions of the landscape correspond to different physical quantities and qualities. Ontogenetic developments and differentiations are like a ball rolling down from a ridge into a valley, which are representative of context-independent constraints (Walsh, 2015). The trajectory of the ball is directed by the contours of the walls of the valley. Any perturbations to the system would result in the ball moving up on the side of the walls and then back down toward its regular path—in this way, the topography of the epigenetic landscape secures a robust development

of forms (Walsh, 2015). In the presence of significant perturbations, the ball can scale and overcome the walls, which form the epigenetic buffer development against systemic shock. The results produced by the ball overcoming these buffering valleys are strange phenotypes, known as "phenocopies." The production of phenocopies, according to Waddington, can alter epigenetic landscapes and reshape them in such ways that novel phenotypic traits become increasingly canalized (Walsh, 2015). These are exceptional phenotypes/traits, which are unstable but can become stable through processes of genetic assimilation.

In short, the motion of the ball thus influences and changes the landscape; they are mutually constitutive. In the words of John Wheeler—*space tells matter how to move, and matter tells space how to curve*. Novelty too can be introduced by environmental perturbations as well, thus resulting in the obtaining of a novel, exceptional phenotype/trait (in the form of a novel path/valley created), which can be stabilized by a genetic system. In this way, the previous trajectory/path, including the determined outcome, is negated, with a new outcome *in the process of being obtained*.

This system proposed by Waddington demonstrates that even initial fundamental constraints are not deterministic and are in fact *vague* (Juarrero, 2023). Hence, even initial constraints allow for the enactment of later local time-dependent contextual constraints, which can add complexity to systems while delaying heat death in the form of the second law of thermodynamics. As mentioned earlier, events, as perturbations, require co-constitution via subjective interpretation/registration. In this way, there is an agency afforded to any system, relative to the system's ability to respond to and interpret perturbations, thus facilitating their own evolution, as proposed by Naidoo (2023f). In cryptography, ambiguity is secured through by using a cryptographic tool known as indistinguishability obfuscation, which renders programs unintelligible, while still preserving functionality (Jain et al., 2020).

13 Ontic-state spaces and encoding

If the information observers have about causal networks is limited to macroscopic data, this information must be treated statistically. This means that the information an observer has about a given causal network *will always* take the form of an ensemble of microstates called the ontic (like ontology)-state space (Powers and Stojkovic, 2023). If events are separated in spacetime, then each event must have a separate ontic-state space. The ontic-state spaces are associated with two different observers or with the same observer who observes at two different times. Each different kind of observer will encode observed macroscopic information about an event available to them, at the site of the event. Quantum theory here *describes the discrete evolution of statistical ensembles of causal networks* (by counting the path between each state space) (Powers and Stojkovic, 2023). It has already been demonstrated that discrete formulisms can support physics models (Powers and Stojkovic, 2023). Hence, this is a model for a *non-deterministic system/nature*. To be more precise than the study, this is an *asynchronous* system, i.e., a system that is non-linear or ordered by causality, temporality, or physicality. The organization of the system is that of selective synchronization, with parts working independently and coordinating through non-linear and non-physical/logical means, such as information transmission/propagation, beliefs (Naidoo, 2023g), or other forms of subjective registration/recognition. Thus, the full extent of the synchronicity of the system is contingent on subjective recognition. In the presence of a *lack of recognition* (ignorance) or an *intentional misrecognition* (an intentional negligence—also known as the Kantian infinite judgement, as explained below), there is no mutual or dialectic recognition and, thus, constitution, thus leading to the event not being constituted. This is also akin to the plan recognition problem as mentioned above. For example, in South Africa, the "open society" is an underlying constitutional principle, tethered to other constitutional values such as freedom and dignity, which are implicit. These values require subjective registration to be actualized (see "Sedimentation and memory" in Supplementary Material). For example, a provision is not unconstitutional, or not constitutionally aligned with the values and ethos of a constitution, until it is demonstrated to be—through subjective demonstration/interpretation or construction. A state constitution represents a 1:1 mapping with any system constitution (or mind).

14 Inferences and variables

The non-local emergent properties mentioned are those that drive inference. Events, as measurements, can take the form of two separate observers or the same observer at different times (known as the split-brain [Naidoo, 2023a]). These are two sources of information that would *fully characterize an outcome* (Powers and Stojkovic, 2023). Each ontic space contains the encoded information of each observer, measured in their respective physical contexts.

The physical context can be identified using quantum numbers and divided into four categories: (1) random variable; (2) conditioning variable; (3) local nuisance variable; and (4) non-local nuisance variable. The conditioning variable can be described as a continuity parameter in models that enables probabilities to become arbitrarily smooth (see "(Self) entrenchment and flexibility" in Supplementary Material). These variables are selected or controlled in model experiments. Random variables, on the other hand, are physically observable variables that are neither selected nor controlled (explanation closure axioms). They can take on any value permitted by the conditioning variables. Nuisance variables are physical quantitates, which are not observed in the experiment being performed.

This can be translated into the legal lexicon, being that of the four statutory interpretative methods used. The conditioning variable is the literal interpretative method, given that the immediate language of a statute typically implies a bounded rationality (thus smoothing the probabilities into a set of possibilities). Random variables are akin to the contextual interpretative method, given that contextual information or knowledge oscillates rapidly, which is not controllable, but rather subject to any particular time, place, and epoch. Random variables are akin to *subjective* purposive interpretation (instances of agency). Local nuisance variables are akin to the constitutional interpretative method, speaking to the immediacy aspect of constitutionalism, being the wording of the constitutional documents themselves. *Non-*

local nuisance variables are the histories, values, and philosophies (like the open society) underlying specific constitutions, which are merely implicit and not explicitly stated. These non-local instances require subjective actualizations to come into being, given that they are merely implicit. This renders non-local (and local) nuisance variables, largely contingent on subjective construction.

15 Probability

Within this formulism, probabilities arise because of *hidden information*, which implies that *probability is epistemic in origin* (Naidoo, 2023a; Powers and Stojkovic, 2023). *Ambiguity* is thus epistemic in origin. State spaces that are associated with distinct choices of observables *are always disjoint* (a Hegelian and Freudian concept), which means that no single-base 16 sequence will ever appear in state spaces associated with two different sets of observables (see "Freud" in Supplementary Material). The frequentist interpretation of probabilities states that the size of the conditioning variable (n) will have a significant impact on the size of the state spaces and, hence, on the continuity of probabilities.

If (n) is small, Alice and Bob (as two different observers) will be able to perform enough experiments to have observed all possible ontic states (if we assume that no ontic state will ever occur more than once). Upon doing so, the predicted probability and their measurement results will match exactly (which is in line with the frequentist approach to probability). Another consequence of a small (n) is the loss of statistical independence (assuming once more that no state occurs more than once). Hence, Alice and Bob's past measurement outcomes would influence what they will know about future experiments or measurements (*memory*). This is a reciprocal/dualistic approach to self-constitution, which leads to each observer to secure (or "predict") the results of their future observations/experiments. This leads to a possibility for completely deterministic experiments (see "Negation and prediction" in Supplementary Material). In computer science, "fuzzy logic" is used to model human decision under vague information (Zadeh, 1999). Recently, important links have been established between fuzzy logic and Bayesian inference (Gentili, 2021).

Relating back to contextual import, both iteration and recursion interweave the subject's/system's own history, paths, choices, and outcomes back into the subject/system. *Feedforward processes* can *modify themselves by using contextual information* as described to anticipate expected conditions and events. Feedforward loops are those in which the *attractors are anticipatory (predictive).* It is important to understand that self-modification or anticipation describes *probability distribution changes relating to events within a possibility space* (Naidoo, 2023g; Juarrero, 2023). More information on securing a desired state, instead of trying to predict it, is given in the study by Naidoo (2023g). This also circumvents the is–ought false dichotomy.

16 An ontology of reality

Events that are "fixed" are directly related to their ontological status. All observed properties of these fixed events are taken to be definite states of reality. Any base symbols within an event, whose counts are observed during an experiment, *cannot* vary in configuration within the associated observer state space. When the four base-4 symbols are held fixed in either observer state space, the assumption is that all four quantum numbers associated with each of their events are either random or conditioning variables. This experimental design requires *two sources* (duality) of information to fully characterize an outcome; hence, the *product of both observer state spaces* is fundamental to calculate probabilities. The *joint state space* (a *dialectic*) represents that both observers have reached a consensus (Naidoo, 2023a) on all local quantum numbers for each experimental outcome (Powers and Stojkovic, 2023) (see "Reciprocal self-constitution" in Supplementary Material).

In this formulism, the information that observers are permitted to have regarding a causal network is *limited* to local quantum numbers, which results in indistinguishable ontic states. Indistinguishable ontic states are linked to permutations and variance. The first type of the indistinguishable ontic state arises from permutations in local quantum numbers; *the invariance* of both observers' base-4 sequence is maintained. This symmetry arises because the information hidden in the ordering of the symbols constituting a sequence is hidden from observers.

The second type arises from permutations that lead to *variations in the non-local quantum numbers*, while the local quantum numbers remain unchanged. Here, the symmetries of local quantum numbers are maintained, but not that of the non-local quantum numbers (Powers and Stojkovic, 2023). These symmetries arise because the numerical values of the non-local quantum numbers are hidden from observers. Both symmetries describe the nature of hidden information—if information was not hidden from observers, then no two ontic states would be indistinguishable. Hence, there would not be any meta-degrees of freedom, for inferences or interpretation. In other words, the system of reality is determined to be open (Naidoo, 2023b; Naidoo, 2023e; Naidoo, 2023g).

17 Nihilism: linearity and optimism

Linearity results in feelings of nihilism, which is a lack of intrinsic purposiveness or optimism, since life becomes nothing more than a transient stop from the abyss to the void of nothingness instead of being purposive. This inherent groundlessness can be liberating since the absence of an underlying or guaranteed truth/objective can result in the unrestricted maximization of freedom. Ultimately, everything can be upturned or changed since nothing is grounded (Naidoo, 2023e). However, the focus on the point of nothingness, as being the origin, and endpoint/destination reduces optimism in individuals and society. The typical attitude of "get to the point" in all social and professional spheres of life demonstrates an inherent linear nihilistic tendency as if the point (or nothingness) itself is the purposive of things. This kind of thinking is *prevalent in academic culture.* Society fails to realize that information only exists *if there is no point*—that is why it is called in-formation (or evolving). *It is that which is currently in formation and, thus, unfinished or incomplete.*

The dopaminergic system is a double-edged sword (Sapolsky, 2018). Dopamine is about reward. Dopamine is created in multiple regions such as the *ventral tegmental* (VT) area

located near the old brain stem. Certain stimuli activate the VT area, which then activates the release of dopamine. Chronic stress or pain decreases dopamine and the sensitivity of the NA, leading to depression. The pleasure system is primarily probability-based; the higher the probability, the less dopamine released, the lower the probability, the more reward released. Thus, *the dopamine reward system is not absolute and is instead relative to the reward value of alternative outcomes* (Sapolsky, 2017). Habituation results in less dopamine release. It is the *surprise* factor that determines how much dopamine is released. Thus, dopamine's motivational function ("feeling good feeling") is not concerned with the reward or outcome, but rather, it serves to drive action to obtain an outcome based on the *anticipation of a reward and not the reward itself* (Sapolsky, 2017). This anticipation builds up and requires learning—which comes from contexts—and *is not inherent*. It is this modality of motivation that *builds optimism* and drives action (binding the reward to action) to pursue an *uncertain* or *unpredictable* eventuation of a reward. From this, we can surmise the following:

- Human reward systems function to reward failures to correctly predict (an unexpected surprise). Incorrect predictions thus lead to a reward—there is pleasure in pain!
- Rewards and pleasure are linked to fundamental biological drives, including reproduction, social status, and organization, but they can also be ontogenetically programmed into the brain. They can be developmentally/contextually based.
- Rare states, things, or achievements require more reward for anticipatory motivation because they require more energy expenditure; less rare states release less reward because of the converse.
- Human reward systems motivate the pursuance of rarer states, things, and experiences. Humans thus want what they cannot have or what is difficult to have.
- Thus, the human reward system sets us on a path that has a great risk of failure because rarity is a risky pursuit.

Paradoxically then, *the linear movement of efficiency, also known as optimization, functions inversely to dopamine release and, thus, enjoyment*. The reward system seeks out and functions on non-linearity or *uncertainty*. The certainty inherent to linearity is harmful to human optimism and enjoyment. This begs the question of whether there is any scope for freedom, autonomy, or enjoyment? Both Kant and Freud have answers for this, which, given the above, have now been validated.

18 Kantian aesthetics: the Sublime

In the *Critique of Power Judgement*, Kant begins by discussing the difference between the "beautiful" and the "Sublime." For Kant, beauty is connected to the form of an object, which has boundaries or containment, whereas the sublime is formless; to be more precise, the Sublime is described as boundless or formless (simply formlessness is ugly and produces displeasure) (Doran, 2015). Both are concerned with *feeling*. In describing the sublime in *Observations on the Feeling of the Beautiful and Sublime*, Kant (1987) says

> "For what is sublime, in the proper meaning of the term, cannot be contained in any sensible form but concerns only ideas of reason, which, though they cannot be exhibited adequately, are aroused and called to mind by this very inadequacy, which can be exhibited in sensibility. Thus the vast ocean heaved up by storms cannot be called sublime. The sight of it is horrible; and one must already have filled one's mind with all sorts of ideas if such an intuition is to attune it to a feeling that is itself sublime, inasmuch as the mind is induced to abandon sensibility and occupy itself with ideas containing a higher purposiveness."

The Sublime presents the Kantian mind with a special kind of dissatisfaction. On the necessity of the Sublime, Kant (2015) says

> ". . .that which excites in us, without any reasoning about it, but in the mere apprehension of it, the feeling of the sublime, may appear as regards its form to violate purpose in respect of the Judgement, to be unsuited to our presentative faculty, and, as it were, to do violence to the Imagination; and yet it is judged to be only the more sublime."

Kant argues here that the rational faculty of the mind paradoxically requires the contra-purposiveness of the Sublime (hence making the Sublime a purposive–contra-purposive necessity for the whole mind itself). The Sublime is thus that without a purpose, which is a requisite for purposiveness as the Beautiful (of the Understanding) itself. For Kant, subjective purposiveness of the mind is produced through the unison of the Kantian Imagination and the Understanding (when judging the beautiful) (Doran, 2015). In that same light, the unison of the Imagination and Reason produces their own subjective purposiveness *through conflict*.

We experience the Sublime as a sort of transcendence as freedom from sensible constraints because of the asymmetrical conflict between Reason (as the higher faculty) and the Imagination as the lower faculty. Reason imposes its superiority over the Imagination (Doran, 2015). This is a transcendence not by harmony in unison, *but rather by agonism*. The Imagination is this overwhelming capacity of the mind; Reason is this rational capacity that steps in when the Understanding is overwhelmed by the magnitude or force of something (Kant, 2015; Žižek, 2020).

Orgasms release a huge amount of dopamine (as mentioned above). Importantly, as Sapolsky noted, during orgasm, *the amygdala deactivates* in both men and women. Both sex and aggression activate the sympathetic nervous system (SNS), and there is no distinction in the heart rate in states of orgasm and states of murderous rage (Sapolsky, 2017). In other words, as Kant noted, an experience of the Sublime involved the subjectification of the Understanding (experience) to the higher faculty of Reason.

Kant splits the Sublime into the Mathematical and the Dynamical Sublime. In brief, the Mathematical Sublime is that which is the judgment related to the esthetic estimations of magnitude (Doran, 2015). It concerns itself with ideas of boundlessness and formlessness, especially in reference to "totality." The Dynamically Sublime is the esthetic judgement of nature as a power (which derives from the need for autonomy as opposed to totality) (Doran, 2015). The transcendence produced by agonism is a *felt pleasure in pain*. The pain is produced by the overbearing grandeur of the Mathematical Sublime or the

dominating forces of the Dynamical Sublime of nature (Doran, 2015). The pleasure that follows this pain arises through Reason's own transcendence into the supersensible (Doran, 2015). The Sublime is a thing of the mind itself, a mental elevation that comes from the transcendence of sensibility, i.e., the abandoning of sensibility in the pursuit of ideas that contain a higher purposiveness (Kant, 2015). *Reason introduces the idea of infinity, which only it contains* (Žižek, 2020). This Kantian idea is supported by renowned mathematician David Hilbert (1925) in his discussion on the mathematical concept of the infinite in his paper called *On the Infinite*. Hilbert argued that the infinite was fundamental to all thought and reason including mathematics itself. He even went insofar as expressing explicitly that Kant was right. The infinite thus was a constitutive necessity for all thinking and arose within thinking exclusively.

The idea of infinity is much larger than any magnitude an object can present to our understanding, hence diminishing the overwhelming sensations magnitude presents. In this way, we shift from our sensory experiences to a recognition of the higher transcendental powers of Reason that can ideate about infinity (Žižek, 2020). *This is a transcendence through a power of resistance*. It is a pleasure in overcoming. The faculty of Reason *transcends through failure* (by positing the infinite or the *an sich*—something that it can only circumscribe through failure). This is transcendence through de-sublimation.

There are two kinds of objects, namely, the beautiful and the ugly. Ugly objects are divided into contra-purposiveness and purposefully contra-purposive (Doran, 2015). The contra-purposive objects are those that are plainly ugly, while those that are purposefully contra-purposive (in the sense of being boundlessly formless) can be used to create the transcendence mentioned. This is a contingent judgment; in other words, one in which the mind itself reveals its own purposiveness by using "nature" as a means to an end (thus giving it its relative status) (Doran, 2015). This modality of judgement, of which the it purpose is to create a *transcendence or super-sensible feeling*, is not based on the concept of the object itself, *but rather a subjective purposiveness of the mind itself* (Doran, 2015).

In the *Critique of Pure Reason*, Kant (1890) presented the antinomies that result when pure reason tries to access the noumenal world. These antinomies are the (1) mathematical antinomies and (2) dynamical antinomies. The Kantian *ding an sich* could only be circumscribed through failures produced by the antinomies. These indeterminate concepts of Reason (Reason ideating) are the Sublime. They are *fruitful failures* because although one cannot have a positive knowledge of them, *one can know what it is not*. Hence, one can circumscribe the Sublime through failures (sort of like bumping into boundaries of a thing without knowing what the thing is itself). The Sublime is thus defined by its very *indefinability*. It is an aesthetic judgement, which refers not to objects *but to the mind itself* (Doran, 2015). The Sublime is that which could not be contained in any sensuous form, but rather speaks to ideas of reason itself. For example, the expression "I cannot express how much I love you," by its very impossibility or failure, constitutes an expression of the love. *Kant knew that the only way to access essence is to create it through failing to access something* (the fall). Failing to depict essence is paradoxically essence itself.

19 Kant, Gödel, Bartleby, Hegel, and Žižek

Kant (1890) introduced a third category of judgment, i.e., *infinite judgment*. The purpose of this judgement was to explain concepts and judgements from pure understanding (Guyer and Wood, 1998), meaning concepts and judgements that do not arise from perceptions or the empirical senses. Pure judgements of this kind, of the pure understanding, are logical forms of judgements, as pure concepts, which are the logical categories. In terms of *quality*, there are two forms, namely, *affirmation* and *negation*. One can explain this using zombies. One can affirm a predicate, an example of which is that *something is alive*. One can also negate the predicate, meaning that something is *not alive*—being dead. The infinite judgement, on the other hand, is the in between of affirmation and negation. The infinite judgement is the *affirmation of a non-predicate*. Something can be alive, or it can be dead, or it can be *un*dead (a zombie). What is affirmed in the latter is a non-predicate, being the "un." Thus, Kant toppled the traditional binary of affirmation and negation by introducing this third, in-between category. This category is precisely defined by the way it cannot be defined, and it does not require any sensory/empirical information.

The "un" is important, and it relates to the German language. In English, we typically understand finite as being the opposite of infinite. However, this is not the case in German, which is dialectic in nature. The German word for infinity is "*unendlichkeit*." The "un" predicate indicates an uncertainty from within the subject (*endlichkeit*). Unlike its English counterpart, the German concept of infinity does not describe something endless—it describes the concept of the finite as something which contains within itself its own negation. This self-negation then creates something else. It is similar to the Freudian concept of "*unheimlich*," which is translated into uncanny (Freud, 1919). This term means that which undermines itself from within. The German infinity thus describes a negation of negation, which is the affirmation of a non-predicate, or the Kantian infinite judgement. *Infinite judgements are Sublime*.

Turing's halting problem and Kurt Gödel's incompleteness theorem are of relevance. In brief, Turing demonstrated that computers cannot produce completely self-referential statements about themselves; computers cannot reveal truths about all computer programs. Gödel then demonstrated that mathematical systems do express truths about their own logics; mathematical systems express *what they can and cannot prove*. This is only true for mathematical systems based on computational logics. Hence, if axioms are computable, according to Gödel's theorem, mathematical systems *cannot simultaneously be consistent and complete*. Hence, a computational system can only (1) prove a false statement or (2) the computational system would fail to prove a true statement. Hence, all computational systems would contain true statements, which it cannot prove. If the system does prove the statement, then the system is proving that the statement is false (Scientific American, 2006). Hence, computational logics can only (1) *fail* to prove a truth or (2) uncover that the truth itself is of a *failure*. Most importantly, the Gödel theorem demonstrated that mathematics and numbers are just quantitative measures, which

would otherwise not exist without a quality to which they could attach. In other words, mathematics is derived from quality (or identity) and not the other way around (Naidoo, 2023b).

In Bartleby, the Scrivener: A Story of Wall Street (Melville, 2011), a clerk is hired by a wall-street lawyer to perform administrative tasks. After a bout of hard work, when the clerk is asked to perform another task, the clerk responds with "I would prefer not to." This refusal then results in the workplace being sent into disarray. As noted by Žižek (2008),

> "Sometimes doing nothing is the most violent thing to do."

Although Bartleby *does nothing*, this nothing is destructive and turns out to be much more effective than doing "something." For example, corruption is commonly understood as a hidden act. Most commonly, however, corruption involves hiding the fact that *nothing has been done*, as opposed to something. Corruption is an *intentional negligence*. As Žižek pointed out, "I would prefer not to" is an example of the infinite judgement, which would take the form of a refusal to accept false given, false dichotomies (as ideological oppositions), or choices. The use of "I would prefer not to" is, thus, a strong form of destructive autonomy, which highlights the self-undermining truth of all logical propositions.

20 The owl of Minerva

"When Philosophy paints its grey in grey, a shape of life has grown old, and it cannot be rejuvenated, but only recognized, by the grey in grey of philosophy; the owl of Minerva begins its flight only with the onset of dusk" (Hegel and Woods, 1991).

Hegel's argument here is that philosophy (rationality, reason, and coherence) *occurs only after* the occurrence of events. Rationalizations, as stories, narratives, or explanations, are those created by implementing causality on events and information. These rationalizations are of the past, but they form part of the present, and they proceed forward into the future, unless that presupposed rationalization is altered. This reduction in the dimensions of information is necessary for efficiency (biasing) purposes for processing. "Un" information requires a higher energy expenditure to compute; hence, memories perform abstraction processes to create efficiency. *Memories are modes for theorizing about the world*, with explanations following this process of theorization. These narratives require impositions of biased causal and logic links, thus creating relations between informational points—creating a logical sequence (Taleb, 2007).

21 An open future and "the end of history"

> "Men make their own history, but they do not make it as they please; they do not make it under self-selected circumstances, but under circumstances existing already, given and transmitted from the past." (Marx, 1852).

Meta-freedom, as Naidoo (2023a) noted, is qualitative freedom, which is not simply the freedom to choose between a selection of presented choices but to construe the concept of freedom, or freedom of choice itself, differently. A degradation of meta-freedom, or qualitative freedom, is exemplary of an unhealthy, totalitarian, static, risky, closed, dead, and deterministic society/system (including a mind). In other words, this is an efficient society. Imagination, as Naidoo (2023a) noted, is lacking in closed and efficient systems, as there is no place for autonomy, contradiction, differences, or freedom of thought/expression.

Hegelian teleological historicity has often been mistakenly criticized as being closed, or leading to totalitarianism (Popper, 1945; Naidoo, 2023a). The idea of any pre-determined teleological unfolding of history typically gives rise to naïve notions of strong determinism, a lack of free will, and thus, a lack of autonomy and freedom. The notion of fate is thus always attached to determinism and unfreedom. However, as Hegel (2010) have pointed out, both freedom and free will are only possible *because* of fate, not in spite of fate. Free will is not a concept that describes being able to do what one wants. Properly construed, free will is an inverse (counterfactual) relation, which is contingent on alternative histories. Free will requires a situation, wherein there is an exact copy of the actor and their universe elsewhere. The actor is free, only insofar as they can enact an outcome, which their corresponding copy, in the alternative universe, could not do. In other words, they follow another path to their copy. Free will thus depends on the degrees of meta-freedom (Naidoo, 2023a), which one would counterfactually not have if one did not have free will.

It is not that we must imagine that Sisyphus can do what he wants; we must imagine that Sisyphus *wants what he does*. *Freedom lies in the narrative interpretations that people create*. Fate is the enabling condition for both the concepts of freedom and free will because both concepts can only have meaning if they are contrasted to another concept as their contradiction being fate. Hence, without the concept of fate, there can be no freedom or free will.

Any historical process contains within it the overlapping of necessity and contingency. However, as Hegel (2010) demonstrated in his extrication of contingency and necessity, it is not that an underlying deeper necessity is realized through a set of contingent actualizations. It is, instead, the contingent actualizations that determine the fate of necessity itself. This means that necessity, as a concept itself, only arises through retroactive interpretation of contingent events. This is how subjective narratives are formed by the brain, which are semantic relations between events and facts, which are biased toward the observing subject (Naidoo, 2023a) (see "Sedimentation and memory" in Supplementary Material).

Thus, true freedom (and openness) is, for Hegel, the ability to change previously accepted presuppositions because all presuppositions are groundless (Naidoo, 2023e). As such, all presuppositions are intrinsically open to change and re-interpretations through the processes of self-referential reasoning and abduction (Naidoo, 2023f) (see also "(Self)entrenchment and flexibility" and "Temporality and generative entrenchment" in Supplementary Material). A history that is not open to re-interpretation, or stagnant, is one that is dead. For example, an interpretation of a work constructed in 2023 *must be different* to an interpretation *of the same work* constructed in the 1950s. For any future to be open, it is necessary for any past narrative to be subject to a modern re-interpretation of society.

Freedom is the ability to constitute and interpret (or negate) events, which is to write and re-write events (Naidoo, 2023e). Information, which is repeated, surprising, or recent tends to be prime for storage and usage because this enables the brain to better predict later temporal occurrences (Hawkins and Dawkins, 2021). Recalling information involves the recollection of subjectively imposed narratives of events, which, upon each recollection, are slightly altered (Taleb, 2007). Charles Baudelaire was the first to theorize this, where he compared our memories to palimpsests that could be written and re-written on continuously. This has neurochemical backing—when new memories are formed, the brain actively "breaks" DNA to store the new memories. Alzheimer's occurs when the repair process degrades that would "fix" this (Miller, 2021). This is a physical and violent re-ordering of the past (see "(Self)entrenchment and flexibility" in Supplementary Material).

Neurobiologist Robert Sapolsky demonstrated that there is fundamentally *no biological difference between love and hate*. Alertness is a function of the amygdala in conjunction with other regions. The amygdala activates a part of the brain stem (1) called the *locus coeruleus* (LC) (Breton-Provencher et al., 2021), which is like the brain's very own SNS. This sends norepinephrine projections throughout the brain—including the cortex. If the LC is not "excited," then the human is calm and unalert. If it demonstrates high activation, then this is a massive state of alertness in which *perception is amplified*. Importantly, this means that *the autonomic emotional patterns influence the intensity of feeling*/state, *but it does not determine the content of what one feels*. Both love and anger (positive and negative) work fundamentally in the same way, i.e., heightening or lowering feeling. If I "love" something very much, I tend to have a state of high alertness for that something, whether it is a person or observing the color blue of the sky. If I hate something equally, like the blue sky, I will have the same state and intensity of experience. As Sapolsky (2017) recalled, *the opposite of love is not hate, but rather indifference*.

Oxytocin is vastly considered to be the "love" drug (or chemical, to be more accurate). However, despite its benefits, there is also a negative side to oxytocin (Azar, 2011; Northwestern University, 2013; Badcock, 2016). Although oxytocin assists in the formation of mother–infant and monogamous pair bonds, lessens anxiety and stress, increases trust and social affiliation, and causes people to be more cooperative and generous, it only enhances pro-sociality towards the "us." Oxytocin presence in interactions with "them" causes ethnocentricity and xenophobia. That which fosters love and sociality is also that which *causes divide. It is not dissimilar to how we justify war under the auspices of peace.*

22 Todestrieb: saving the death-drive

The Freudian notion of the death drive has a long tradition of being misconstrued as a literal death drive. However, Freud's *todestrieb* described the view that man held about death, being something, which is staged, within life. The death drive describes the process of de-sublimation or de-subjectivization. This development, or ones *becoming*, occurs through this process of constitutive negativity or the Hegelian negation of negation (self-relating negativity). This is a pure form of agency, whereby one's unfolding is a process of incremental "deaths," which allows for development and the occurrence of heightened feelings (Naidoo, 2023b). Freud described that feeling humans obtain when closest to death (like rollercoasters).

There are four propositions that are important when considering the death drive (Hook, 2016): (1) biological instinct; (2) a cosmic principle; (3) the Nirvana-like release of tension; and (4) the impulse to self-annihilation. On (1), Žižek (1989) noted that

> "[W]e have to abstract Freud's biologism: 'death drive' is not a biological fact but a notion indicating that the human psychic apparatus is subordinated to a blind automatism of repetition beyond pleasure-seeking, self-preservation, accordance between man and his milieu. Man is – Hegel dixit – 'an animal sick unto death', an animal excoriated by an insatiable parasite (reason, logos, language). In this perspective, the 'death drive', this dimension of radical negativity… defines la condition humaine as such…. All 'culture' is in a way a reaction-formation, an attempt to limit, canalize – to cultivate this imbalance, this traumatic kernel, this radical antagonism through which man cuts his umbilical cord with nature, with animal homeostasis."

The above statement is a Lacanian proposition that can be found in the Seminar on *The Purloined Letter* (Lacan, 2006), wherein Lacan equates the death drive with a form of symbolic constitutive repetition (automatic repetition). Thus, the death drive is not a biological instinct to return to pre-life (inanimation). This repetition is a form of obsessive compulsive disorder. In his 1964 seminar *The Four Fundamental Concepts of Psychoanalysis*, Lacan (1979) highlighted how the death drive is inherent to the Freudian unconscious and memory (Hook, 2016). The unconscious, in this way, is not able to satisfy itself other than by re-finding an object that has forever been lost to it (Hook, 2016). Thus, for Freud, repetition was not something *determined by humans but something which determined humans.* Repetition is an instance of "more-making" that serves to preserve stability (Eldredge, 2015). As a context-independent constraint, repetition increases the magnitude of a value within state spaces. Large increases in magnitude correlate to an increase in density, which can then *deform a state space*, thus potentially driving systems even further away from equilibrium and thermodynamic heat death. Repetition is also an insurance mechanism that ensures that valuable traits are not lost when perturbations occur, like the repeated nucleotide bases in the genome (Juarrero, 2023). This is the process of creating redundancy. Redundancy ensures that there are additional components in a system, thus introducing "fail-safe" measures, which preserve systematic functionality, even if individual components fail (rendering systems *robust*). Repetition is also important for communication as it improves the fidelity of transmission within noisy mediums as repetition communicates information relating to regularity, which is rare. Thus, repetition and redundancy preserve and transmit information and keep systems further away from equilibrium while preserving coherence and metastability (Juarrero, 2023).

The Lacanian death drive involves *a death in form* (the mortification of the Symbolic) rather than in content (morality

and death) (Hook, 2016). The death drive in this way compels the subject, using antagonism, to transcend from the natural (the animal or the human) to the denaturalized subject (culture, identity, and the symbolics). This also involves a fundamental denaturalization of human sexuality (agent like sexuality that is not aimed at reproduction) as opposed to animal sexuality (instinctual coupling) (Hook, 2016). In this way, the death drive is the reason why the human animal is denaturalized and *not subject to the normal course of evolutionary adaptation*.

In *The Selfish Gene,* Richard Dawkins (1976) coined the term "meme" that described a transmissible piece of information, which can be sociocultural or symbolic–linguistic units, which circulate among entities who can share and use these units (Johnstone, 2008). They (memes) establish boundaries, co-ordinates, and attractors in mental, physical, and social spaces (Juarrero, 2023). They serve to frame and bias cognitive possibility spaces, thus spreading mental and social attractors. They are contagious, in that they are easily socially transmitted. The more predominant a meme is, the more they alter social and mental attractors, thus reconfiguring mental and physical spaces (Juarrero, 2023), including cultures, languages, and laws. Memes co-evolve with their contexts, and they display an inertia to change, even when conditions no longer require them or when they may be harmful.

Exaptation, an idea coined by Stephen Gould describes situations where the original evolutionary function of a feature is different to its current usage (Gould and Vrba, 1982). Wings, for example, originally evolved to increase surface area, thus enhancing thermoregulation. Flight as an ability was only exploited by organisms after the evolution of wings. Hence, features can have different functions, or no useable functions in one context, but *new/novel functions can emerge* in other contexts. Dawkins also noted the concepts of "de-aptation" or "de-aption." These concepts describe instances whereby a meme can override the interests of those who created them (or genetic programming) and can change aspects of its founder. Although memes can arise from genes/biology, they are not necessarily genetic/biological, and they can obtain an independence from the material substrate from which they arise—a transcendence of sorts. Memes can hegemonize the biological substance of humans (Johnston, 2008), subjecting them to "non-biological" or denaturalized structures. This kind of indirect adaptation is described in "Negation and prediction" in Supplementary Material.

It was Schelling (1802) who first described the Universe as biological. He realized that some things could not be derived logically but could only be narrated (Žižek, 2020). Schelling's conception of the Real as the primordial drives is meant to demonstrate a move from *logos to mythos*. The ancient Greeks believed (Schelling too) that the orgasm is the height of human experience because it symbolized the unification of the Two into the One. This unification was thought to be the Absolute, or the perfections, wherein harmony is achieved (no conflict) and differences are reconciled. However, there is no harmonious unity, only a unity-in-difference (a failure to unify, or an antagonistic gap), which enables a true brush with the Absolute (see "Temporality and generative entrenchment" in Supplementary Material). This is constitutive of human sexuality (Žižek, 2020) in two ways. The first, as an expression of autonomy, is through denaturalized sexual interactions and the second is through the introduction of novelty. Sexual reproduction involves a vertical genetic transferral (not duplication) from organisms to their offspring and a horizontal genetic transferral between bacteria and unicellular eukaryotes (Juarrero, 2023). The latter introduces novelty through *contextual import*, which is a multiple realizable constraint (non-random) that results in an expanded possibility space, with novelty and variation. Through both vertical and horizontal mechanisms, *novelty and change* are introduced. Sexual reproduction thus introduces genetic shuffling, which preserves a species lineage while also introducing novel variants and combinations of traits (Juarrero, 2023). This allows for a larger set of traits and behaviors while remaining true to type (see "Sex and novelty" in Supplementary Material). What this means is that *it is antagonism, contradiction, and failure to unify that is raised to the level of the Absolute*. Love is the death drive in a dance of de-subjectification. It is visceral and violent, carving away at one's past and an emptying out of one's content to de-subjectivize oneself for the other. This is also why "love" is commonly known as a truth event of pure freedom/autonomy as it involves a subjective interpretation and self-entrenchment (see "(Self)entrenchment and flexibility" in Supplementary Material). In other words, *the death drive is not biological, but ethical* (Hook, 2016). This forms the starting point for solving the is–ought dilemma, which will be explored in proceeding works.

To prove Kant right once more, the PFC *silences* nonpathologically during states of orgasm, which produces incredible amounts of emotion (Sapolsky, 2017). Hence, it is through its own *desublimation (silencing) of the superego as the PFC and ultimate reason and logic do we experience orgasm as the site of the most heightened experience*. The orgasm, as the work of the death drive, enables a brush with the Absolute (Žižek, 2020).

On (2), the death drive is not to be understood as the conflict between two opposing forces, but as an inherent *blockage* of the drives (Naidoo, 2023a; 2023f). This blockage then creates the appearance of two opposing cosmic forces (Eros and Thanatos), whom need to be chosen among and reconciled for there to be harmony. This will be dealt with in future works but relates to the concept of *semantic closure*. The death drive is the inner inconsistency of the psychical apparatus, *a constitutive gap that distinguishes drive from instinct* (Hook, 2016).

> "There is only one drive; and the libido which aims for enjoyment and the death drive is the curved space of its formal structure" (Žižek, 2010; Hook, 2016).

As the drives involve blind psychically repetitive behaviors in pursuit of their own satisfaction (Freud, 1920), they obtain pleasure from beyond the pleasure principle. This is pleasure in pain; humans *seek out pleasure beyond the pleasure principle, i.e., in pain* (Freud, 1920). In other words, this is excessive or surplus enjoyment that links to (3). Going beyond the pleasure principle involves developmental excesses, which arise when systems are antifragile (Taleb, 2004). Antifragile systems are those that benefit from failure by incorporating failure into their constitutions. These are "safe-to-fail" systems, as opposed to "fail-safe" robust systems. Antifragile systems can self-ratchet and self-modify through self-reference, thus ensuring their stability. *Post-traumatic growth* in muscles perfectly

describes this (see "Temporality and generative entrenchment" in Supplementary Material). Importantly, the concept of the death drive is fundamental to non-linear distributed networked communications, which will be explored in further work.

In terms of (3), the death drive is not the nirvana principle that seeks equilibrium or balance. Rather, it is *a mode for possibility* (Hook, 2016). The death drive is a mechanism of *expressing autonomy*. Properly conceived, the death drive then is not a compulsion to return to the void of pre-life that can be obtained through death; it is a movement in the opposite direction. It is the movement to lifeless life (the undead), also known as *the Kantian infinite judgement* (this judgement involves the affirmation of a non-predicate, like the undead, for example). *It is a movement away from death and toward immortality* (Hook, 2016). This is the source of human enjoyment, being the seeking and obtaining of surplus. Surplus enjoyment is something that is beyond repetitive biological life. It is a break with repetition, which is more than both life and death. *It is an excess of life*. *Jouissance* is the excess of pleasure that there is in pain, the perverted pleasure of the pain of repeatedly missing one's goals (a misrecognition, or a *failure to unify/complete*) (Žižek, 2000; Hook, 2016). Importantly, in linking the death drive with the unconscious, Žižek noted that

> "The unconscious intervenes when something 'goes wrong' in the order of causality that encompasses our daily activity: a slip of the tongue . . . a failed gesture . . . However . . . psychoanalytic interpretation does not simply fill in this gap by way of providing the hidden complete network of causality that explains the slip: the cause whose "insistence" interrupts the normal functioning of the order of causal is not another positive entity . . . it belongs rather to the order of the nonrealised or thwarted . . . that is in itself as a gap, a void insisting indefinitely on its fulfilment . . . The psychoanalytic name for this gap, of course, is the death drive, while its philosophical name in German Idealism is "abstract negativity," the point of absolute self-contraction that constitutes the subject as the void of pure self-relating" (Žižek, 2005; Hook, 2016).

This intervention (slippages in causality or where there is a misprediction or misrecognition) enables consciousness to arise (Hawkins and Dawkins, 2021). Importantly, Žižek highlighted the unconscious intervention as being a non-realized or thwarted gap instead of being a positive entity. This is not dissimilar to the simulative processes in the PFC. Undesirable outcomes that do not exist other than mere probabilities as alternative histories are thwarted. Decisions are made based on the thwarting of undesirable outcomes based on non-existing probabilities (not dissimilar to the collapse of the wavefunction [see Quantum theory in Supplementary Material]).

In terms of (4), the death drive is the movement both toward immortality as I mentioned above and the movement toward the destruction of the *metaphysical conception of immortality* (as life that persists beyond death) (Hook, 2016). Thus, it is the destruction of life, but not the destruction of that which is in life, more than life itself. It is a movement against moderation, which is in one's best biological interests. This is a form of self-relating negativity *a la Hegel*.

23 Conclusion: a brush with the Absolute?

A brush with the Absolute entails a brush with the explanation closure axioms of reality, which is a brush with a true *form of agency or freedom*. This is because explanation closure axioms specify (1) which events may/may not occur; (2) specification of which events can change a property (thus reducing the logic to simply reasoning about what events may or may not occur); and (3) axioms themselves encode solutions (however, not all solutions need to be coded within them, some can be generated in action). In other words, explanation closure axioms specify *that which they do not specify* (an incomplete Absolute). This is non-formulaic and enables representation, interpretation, and causation. Plan recognition includes instances whereby the pre-programmed epistemological axioms/assumptions "fail" in the sense that they did not predict an event that may have occurred or the solution is one that is generated on the fly through *acts of true freedom or agency* since these acts would be free from the formulism.

Importantly, these epistemological assumptions are evidence of pre-programmed *planning*. Both recursion and iteration are only possible after planning wherein temporal dependencies/sequences are formed and then used (Juarrero, 2023). Explanation axioms *do not need to encode for all events* like those caused by natural forces. This means that natural forces, as events, can interfere with the prediction process, and it would be *impossible* to prove that an event did not occur (a constraint). Kant described this perfectly when he painted a picture of the dynamically Sublime, which describes incredibly massive forces of nature that, when recognized in its representative form, does not threaten our lives, which instead fills us up with power of *resistance* to it (enables acts of freedom and heroism) (Doran, 2015; Žižek, 2020). In other words, *a failure of the plan produces an otherwise unprovable truth*. It may even be so that subjective interpretation *is itself a form of natural force, which alters the course of events*. This will be argued in proceeding works.

In this way, one can circumscribe the "existence" of a plan through a proof by contradiction. This was also the modality used to prove the most difficult mathematics theorem from the 17th century, Fermat's Last Theorem. Andrew Wiles constructed a proof by contradiction wherein Fermat's Last Theorem was proved to be unsolvable *and, thus, true* (Klarreich, 2020). Proofs of this nature (indirect proofs) establish the validity or truth of a proposition by demonstrating that if the proposition were to be false, it would lead to a contradiction. It is thus a proof by assuming the opposite of what one intends to prove as true—by relying on purposefully bringing about a contradiction. If the contradiction arises, the assumption is incorrect, and the conclusion is true. In this way, one can prove the existence of a kind of object, without providing an example of it (this is known as a non-constructive or existence proof) (MathWorld, n.d.). There are existing truths, which are not computationally identifiable (since they are epistemological). From the perspective of the planned for (or *plan*), it would be an instance of the Sublime, because there is a *lack/failure to obtain knowledge of the plan itself* (the epistemic constraint mentioned above). However, one can "identify" the contours, or existence of the plan itself, by circumscription (through failure or impossibility). Hence, the Absolute are the framing axioms of humans (our epistemology), and the brush with these axioms is a proof by contradiction of a programmer. Thus, truth

is obtained through an identification of "failures" or "gaps" or "limits" of initial planning axioms (their ambiguity). The brush with the Absolute, as the failure to obtain truth, is constitutive for individual (as the collective human) autonomy or freedom, as Kant described. A plan, however, does not negate freedom; *a plan enables it*.

The discussions of the various epistemological ontologies, including the scientific ones, served to highlight the invariant cross-cutting theme of radical, intrinsic openness. Each thus demonstrates the intrinsic property of openness as being the ability to write and rewrite (determine, re-determine, constitute, and reconstitute) narratives, knowledge, and systems. In other words, one is always able to go back to the drawing board, or as Hegel put it, necessity is just an instance of actualized contingency posited retroactively. Rationalizations occur in hindsight, and rationalizations *must always remain* open to be re-rationalization. Thus, degrees of freedom are afforded by this radical openness, which enables freedom of subjective interpretation and, ultimately, *agency*. Building on this, a truly open society and system is one that is measured by its ability to re-write itself and whether it actively creates conditions where its components can *challenge/change* their own presupposed truths/rationalizations. This is *semantic/imaginative freedom*.

Unfortunately, philosophy got lost along the way. Philosophy is the study of thought, logic, and rationalization itself. Philosophy *requires* taking the impossible topological (external, or Bartlebyian) view. Philosophy *is* knowledge; knowledge *is* coherence; coherence *is* trust; trust *is* subjective order; subjective order *is* embodied health. Unfortunately, current times seem to separate philosophy from other sciences, forgetting that the natural/empirical sciences used to be called "natural philosophy" until recently. Remember, there is a philosophy of science, but there is no science of philosophy. This attitude toward philosophy, or critical thinking largely emanates from capitalism, and its fair maiden, the analytic and empirical sciences (like physics). The domination of the latter schools of natural philosophy requires that philosophy must always be attached to another discipline and provide "practical applications." Practicality, properly speaking, is the *efficient cultivation of objects rather than persons*. Moreover, the schools of natural philosophy tend to repress the fact that they are philosophy, so they can repress other schools of philosophical thought and secure more resources (funding and primacy) for themselves at the cost of others (Naidoo, 2023a). The schools of natural philosophy would do well to remember that claiming a depoliticized, or objective zone, free from bias or subjectivity, *is the most political act*. Subjective ideology is at its strongest where it is universalized or not experienced as such. The most political act *is to claim a space of non-politics* (ultra-politics). In other bodies of work (Naidoo, 2023a; 2023b; 2023c; 2023e), it has been demonstrated that the logics underlying the empirical sciences come from the German Idealists, most notably, Hegel, who obtained such from Christian theology. This is a truth that must be rendered.

Implicit in the denigration of thought are the following assumptions: (1) thinking is not practical and (2) thinking is not valuable. In terms of (1), this is a false separation between theory and practice, arising from Kant (Naidoo, 2023a). There can be no practice without theory, and practice is blind without theory. All quantity is derived from a quality. This is the common mistake that mathematicians make as well; magnitude or counting cannot exist without having at first a *quality*, which is to be counted. Quality is difference or identity. Identity is theory, or a rationalization, or an epistemology, which is used as a framing device. Proper thinking is speculative processes, which thinks *quality and quantity together*; the *good and the bad are thought together*.

In terms of (2), the denigration of thinking as not being valuable unless it has a market application only serves to entrench terrible presuppositions rather than creating an environment where presuppositions can be tested, validated, and continually replaced. Viewing thinking and theorizing as not being valuable is harmful to the embodied health of persons and societies since it prevents both from actualizing novel rationalizations, which are more suited to remaining both viable and open. In societal terms, this attitude can lead to totalitarianism, pessimism, rigidity, and societal fragility. It is *contra-evolution*. Moreover, as mentioned, memories are modes for theorization. Thus, if meta-degrees of freedom for theorization are reduced, so too would be memory. The loss of individual or societal memory (semantics) is the reason why past mistakes are repeated and is a *sign of decline* (in terms of normal health, and especially embodied health). Memory is vital for understanding, creativity, production, and the continuance of life.

When one considers value in terms of a singular outcome, being a market application, one only serves to create a linear and nihilistic subject, which becomes a servant to a single truth; this truth is typically in the interests of someone other than the subject. This is the movement of efficiency, being a servant or a subject, of a single truth, instead of serving as many different truths as one wills. One cannot will different truths, if one is not able to construct different qualities. Value must remain ambiguous because this would enable persons to construct their own agency in relation to what is valuable to them. Ultimately, what is of *value to society* is the cultivation of individuals who can *think critically*. Critical thinking (*taking the topological view*) requires the use of metaphysics and meta-reasoning (Naidoo, 2023a), which is the ability to construct *qualitative rationalizations or new theories*. Why is this necessary? The importance of meta-thinking, or epistemological questioning, is that one can use it to ensure that a system/ontology remains open. In terms of health, the said methods allow one to critically question whether the ways in which problems are framed *actually reproduce those problems instead of solving them*. One needs theory to determine where practice is failing and to jury-rig new practices. This is also the importance of the violent act of love (or the death drive); it is love that is the enabling condition for rewriting one's past, steering the course of oneself, or society, into a more optimistic direction when previous paths lead to failure. Of course, the denigration of love today can be traced back to natural philosophy (Descartes, in particular), which seeks to cut people off from their emotions and feelings, and the self-empowerment associated with being able to steer the course of one's emotions. This happens only so that the schools of natural philosophy can maintain their entrenched positions.

Society today is one that could not produce valuable and creative thinkers, even if it willed. This is incredibly pessimistic for maintaining an open, liberal, optimistic, and free society. This is

terrible for embodied health as is the only position critical thinkers are afforded today:

> "Better to do nothing than to engage in localized acts whose ultimate function is to make the system run more smoothly (acts like providing space for the multitude of new subjectivities, and so on). The threat today is not passivity but pseudo-activity, the urge to "be active," to "participate," to mask the Nothingness of what goes on. People intervene all the time, "do something"; academics participate in meaningless "debates," and so forth, and the truly difficult thing is to step back, to withdraw from all this. Those in power often prefer even a "critical" participation, a dialogue, to silence—just to engage us in a "dialogue," to make sure our ominous passivity is broken" (Žižek, 2006).

Society tends to forget that people theorize and rationalize to make sense of the randomness, chaos, and suffering that surrounds them so that they may survive. In thermodynamic terms, they do this to reduce their entropies by creating order. People theorize to keep their futures undetermined, open, and optimistic. A healthy society is one that *provides resources* to people so that they may cultivate their own embodied health, as opposed to objects. As Kant noted, transcendence, or optimism, relies on *subjective purposiveness and not objects*. The human dopamine system supports this, wherein it is the subjective path, which is prime, not the object itself. Moreover, people simply enjoy and, thus, are optimistic about creating theory because it is the act of subjective purposiveness.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

MN: writing–review and editing and writing–original draft, visualization, validation, resources, methodology, investigation, formal analysis, data curation, and conceptualization.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was supported by the US National Institute of Mental Health and the US National Institutes of Health (award number U01MH127690) under the Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) program. The content of this article is solely the author's responsibility and does not necessarily represent the official views of the US National Institute of Mental Health or the US National Institutes of Health. In addition, support was also obtained from the National Research Foundation (NRF) under the Doctoral Innovation Scholarship.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer MB declared a shared affiliation with the author(s) NN to the handling editor at the time of review.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2024.1331237/full#supplementary-material

References

Allen, J. F., and Ferguson, G. (1994). Actions and events in interval temporal logic. *J. Log. Comput.* 4, 531–579. doi:10.1093/logcom/4.5.531

Azar, B. (2011). *Oxytocin's other side*. American Psychological Association. Available at: https://www.apa.org/monitor/2011/03/oxytocin (Accessed March 4, 2023).

Badcock, C. (2016). The dark side of oxytocin. *Psychol. Today*. Available at: https://www.psychologytoday.com/intl/blog/the-imprinted-brain/201610/the-dark-side-oxytocin (Accessed March 4, 2023).

Breton-Provencher, V., Drummond, G. T., and Sur, M. (2021). Locus coeruleus norepinephrine in learned behavior: anatomical modularity and spatiotemporal integration in targets. *Front. Neural Circuits* 15, 638007–638011. doi:10.3389/fncir.2021.638007

Camus, A. (2000). *The Myth of Sisyphus and other essays (translated by J. O'Brien)*. London: Penguin Books.

Caves, C. M., Fuchs, C. A., and Schack, R. (2002). Quantum probabilities as Bayesian probabilities. *Phys. Rev.* 65, 022305. doi:10.1103/PhysRevA.65.022305

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Doran, R. (2015). *The theory of the sublime: from longinus to kant*. Cambridge: Cambridge University Press.

Freud, S. (1919). The "uncanny". Available at: https://web.mit.edu/allanmc/www/freud1.pdf (Accessed November 18, 2022).

Freud, S. (1920) *Beyond the pleasure principle*.

Gabriel, M. (2009). The dialectic of the absolute: Hegel's critique of transcendent. *J. Philosophy A Cross-Disciplinary Inq.* 4 (8), 38–15. doi:10.5840/jphilnepal2009485

Gabriel, M. (2011). "The dialectic of the absolute: Hegel's critique of transcendent metaphysics," in *The dimensions of Hegel's dialectic*. Editor N. G. Limnatis (New York: Continuum), 76–96.

Gabriel, M. (2011). *Transcendental ontology: essays in German idealism*. New York: Continuum.

Gentili, P. L. (2021). Establishing a new link between fuzzy logic, neuroscience, and quantum mechanics through bayesian probability: perspectives in artificial intelligence and unconventional computing. *Molecules* 26 (5987), 1–18. doi:10.3390/molecules26195987

Gleason, S. B. (2021). *Hegel and schelling: the emptiness of emptiness and the love of the divine*. Florida: University of South Florida. Unpublished PhD thesis.

Gould, S. J., and Vrba, E. S. (1982). Exaptation - a missing term in the science of form. *Paleobiology* 8 (1), 4–15. doi:10.1017/S0094837300004310

Green, C. (1969). "An application of theorem proving to problem solving," in *Proceedings of the first international joint conference on artificial intelligence*. Editor D. E. Walker, 741–747.

Guyer, P., and Wood, A. W. (1998). *Immanuel kant: Critique of pure reason*. Cambridge: Cambridge University Press.

Hawkins, J., and Dawkins, R. (2021). *A thousand brains: a new theory of intelligence*. New York: Basic Books.

Hegel, G. W. F. (1969). *Science of logic (translated by miller, A. V.)*. London: Allen and Unwin.

Hegel, G. W. F. (1991). "Elements of the philosophy of right," in *Translated*. Editor H. B. Nisbet (Cambridge: Cambridge University Press).

Hegel, G. W. F. (2010). *Science of logic (translated by G. Di giovanni)*. Cambridge: Cambridge University Press.

Hilbert, D. (1925). *On the infinite*.

Hook, D. (2016). Of symbolic mortification and "undead life": slavoj Žižek on the death drive. *Psychoanal. Hist.* 18 (2), 221–256. doi:10.3366/pah.2016.0190

Jain, A., Lin, H., and Sahai, A. (2020). Indistinguishability obfuscation from well-founded assumptions. Available at: https://eprint.iacr.org/2020/1003.pdf.

Johnston, A. (2008). *Žižek's ontology: a transcendental materialist theory of subjectivity*. Unites States: Northwestern University Press.

Juarrero, A. (2023). *Context changes everything: how constraints create coherence*. Cambridge: MIT Press.

Kant, I. (1890). *Critique of pure reason (translated by J. M. D. Meikelejohn)*.

Kant, I. (1987). *Critique of Judgment (translated by W. S. Pulhar)*. Indianapolis/Cambridge: Hackett Publishing Company.

Kant, I. (2015). *Kant's Critique of Judgement (translated by J. H. Bernard)*. London: Macmillan Company.

Klarreich, E. (2020). Amazing' math bridge extended beyond Fermat's last theorem. *Quanta Mag*. Available at: https://www.quantamagazine.org/amazing-math-bridge-extended-beyond-fermats-last-theorem-20200406/ (Accessed June 5, 2023).

Lacan, J. (1979). *The Seminar of jacques lacan, book XI: the four fundamental Concepts of Psychoanalysis (translated by A. Sheridan)*. W. W. Norton and Company.

Lacan, J. (2006). "Seminar on the purloined letter," in *The first complete edition in English*. Editors B. Fink and J. Lacan (W. W. Norton and Company), 6–48.

Malabou, C. (2005). *The future of Hegel: plasticity, temporality and dialectic (translated by L. During)*. London and New York: Routledge.

Marx, K. (1852). The eighteenth brumaire of louis bonaparte. Available at: https://www.marxists.org/archive/marx/works/1852/18th-brumaire/ch01.htm (Accessed October 26, 2023).

Marx, K. (1869). *The eighteenth brumaire of louis bonaparte*.

MathWorld (2023). Nonconstructive proof. Available at: https://mathworld.wolfram.com/NonconstructiveProof.html#:~:text=A%20proof%20which%20indirectly%20shows,are%20also%20called%20existence%20proofs (Accessed June 5, 2023).

Melville, H. (2011). *Bartleby, the scrivener: a story of wall street*. Auckland: The Floating Press.

Miller, R. (2021). Brain cells destroyed as neurons learn faster; DNA breaks may provide answer to genetics. *Sci. Times*. Available at: https://www.sciencetimes.com/articles/33264/20210905/brain-cells-destroy-dna-learn-faster-study-finds.htm (Accessed February 5, 2023).

Moyar, D. (2017). *The oxford handbook of Hegel* (Oxford University Press).

Naidoo, M. (2023a). The open ontology and information society. *Forthcom. Front. Genet*, 1–21. doi:10.3389/fgene.2024.1290658

Naidoo, M. (2023b). Batman and the joker: Hegel (em)bodying kant. *Forthcom. Am. Psychol. Assoc. J. Theor. Philosophical Psychol*.

Naidoo, M. (2023c). Hegel among the sciences: logical development. *Forthcom. Humanit. Soc. Sci. Commun*.

Naidoo, M. (2023d). Descartes and German idealism: the mind. *Forthcom. Humanit. Soc. Sci. Commun*.

Naidoo, M. (2023e). Hegel among the sciences: "Nothing". *Forthcom. Humanit. Soc. Sci. Commun*.

Naidoo, M. (2023f). What does it mean to be an Agent? *Front. Psychol.* 14, 1–16. doi:10.3389/fpsyg.2023.1273470

Naidoo, M. (2023g). Artificial inventorship: prying "open" patent law with magic. *Forthcom. Front. Psychol*.

Northwestern University (2013). "Love hormone" is two-faced: oxytocin strengthens bad memories and can increase fear and anxiety. *ScienceDaily*. Available at: www.sciencedaily.com/releases/2013/07/130722123206.htm (Accessed March 3, 2023).

Popper, K. (1945). *The open society and its enemies*. London: Routledge.

Pouget, A., Beck, J. M., Ma, W. J., and Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nat. Neurosci.* 16, 1170–1178. doi:10.1038/nn.3495

Powers, S., and Stojkovic, D. (2023). An event centric approach to modeling quantum systems. arXiv [Preprint]. doi:10.48550/arXiv.2306.14922

Sapolsky, R. M. (2017). *Behave: the biology of humans at our best and worst*. New York: Penguin Press.

Sapolsky, R. M. (2018). Doubled-edged swords in the biology of conflict. *Front. Psychol.* 9, 2625–2629. doi:10.3389/fpsyg.2018.02625

Schelling, F. W. J. (1802). *Presentation of my system of philosophy*.

ScienceDirect (2023) *Plan recognition*. Available at: https://www.sciencedirect.com/topics/computer-science/plan-recognition#:~:text=Plan%20recognition%20is%20the%20task,that%20best%20explain%20observed%20actions (Accessed June 4, 2023).

Scientific American (2006). What is Gödel's proof? Available at: https://www.scientificamerican.com/article/what-is-goumldels-proof/ (Accessed December 3, 2022).

Taleb, N. N. (2004). *Fooled by randomness: the hidden role of chance in life and in the markets*. Penguin Books.

Taleb, N. N. (2007). *The black swan: the impact of the highly improbable*. Penguin Books.

Walsh, D. M. (2015). *Organisms, agency, and evolution*. Cambridge: Cambridge University Press.

Zadeh, L. A. (1999). From computing with numbers to computing with words. From manipulation of measurements to manipulation of perceptions. *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.* 46, 105–119. doi:10.1109/81.739259

Žižek, S. (1989). *The sublime object of ideology*. London: Verso.

Žižek, S. (2000). *The ticklish subject: the absent centre of political ontology*. London: Verso.

Žižek, S. (2003). *The puppet and the dwarf: the perverse core of christianity*. Cambridge: MIT Press.

Žižek, S. (2005). "Lacan between cultural studies and cognitivism," in *Interrogating the real*. Editor S. Žižek (London: Bloomsbury Publishing). chapter 5.

Žižek, S. (2006). *The parallax view*. Cambridge: MIT Press.

Žižek, S. (2008). *Violence*. New York: Picador.

Žižek, S. (2010). *Living in the end times*. London: Verso.

Žižek, S. (2014). *Absolute recoil: towards a new foundation of dialectical materialism*. London: Verso.

Žižek, S. (2020). *Sex and the failed absolute*. London: Bloomsbury Publishing.

Žižek, S., Crockett, C., and Davis, C. (2011). *Hegel and the infinite: religion, politics, and dialectic* (New York: Columbia University Press).

frontiers | Frontiers in Pharmacology

TYPE Policy and Practice Reviews
PUBLISHED 27 August 2024
DOI 10.3389/fphar.2024.1332700

OPEN ACCESS

EDITED BY
Lijun Shang,
London Metropolitan University,
United Kingdom

REVIEWED BY
Davit Chokoshvili,
Luxembourg National Data Service,
Luxembourg
Rania Mohamed Hassan Baleela,
University of Khartoum, Sudan

*CORRESPONDENCE
Lee Swales,
SwalesL@ukzn.ac.za

RECEIVED 03 November 2023
ACCEPTED 25 July 2024
PUBLISHED 27 August 2024

CITATION
Swales L, Gooden A and Thaldar D (2024) The anatomy of a data transfer agreement for health research.
Front. Pharmacol. 15:1332700.
doi: 10.3389/fphar.2024.1332700

The anatomy of a data transfer agreement for health research

Lee Swales[1]*, Amy Gooden[1] and Donrich Thaldar[1,2]

[1]School of Law, University of KwaZulu-Natal, Durban, South Africa, [2]Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics, Harvard Law School, Cambridge, MA, United States

In a data-driven era, the exchange and safeguarding of personal information has become paramount. Data transfer agreements (DTAs) serve to guard privacy, defining the rules for sharing and protecting sensitive data. Yet, the complexities surrounding issues such as data privacy, intellectual property, and dispute resolution within these agreements pose challenges that demand careful consideration. Through a scoping review of twenty-four publicly available, English language DTAs relevant to health research, this article undertakes a comprehensive analysis, examining common clauses, their vital components, and charting a course for responsible data sharing through the provision of insights and practical guidance for drafting DTAs. The article underscores the need for attention to detail and an understanding of data protection legislation in order to ensure that DTAs align with the law and maximize legal certainty.

KEYWORDS

data, data transfer agreement, research, scoping review, South Africa

1 Introduction

In a scientific and research context, the transfer of personal information has become routine. One of the key tools used in data protection compliance, and as part of a holistic data management strategy, is a data transfer agreement (DTA). Lawfully managing and strategically sharing data will arguably become more important than it is at present, where already, for at least the past decade, society has recognized that data has value, and the mantra: "*data is the new oil*" has become an oft-repeated line (Parkins, 2017; Swales, 2022). For example, the proliferation of artificial intelligence technologies, such as OpenAI's *ChatGPT*, that rely on data to produce meaningful output, has further fast-tracked discussions around data transfer, ownership, and management. Additionally, techniques and strategies relating to sharing data are evolving rapidly and should always recognize the value in *some* scientific and academic output. Even where data that does not contain personal information is shared—and, as such, data protection legislation will not apply—it is imperative that this is done intentionally and, in all circumstances, with an eye on the legal consequences (and with consideration for its ownership and value). In most scenarios where data is shared with others, it should be done via a DTA or similar instrument.

What is a DTA? It is a written agreement that facilitates the lawful transfer of data between parties. Typically, an agreement of this type will seek to comply with applicable legislation. Additionally, a DTA will regulate other important legal issues such as ownership of data, intellectual property, the terms of the agreement and how it will terminate, liability, dispute resolution, and whether any consideration is payable (Swales et al., 2023a).

This article presents an empirical study of twenty-four DTAs relevant to health research, which were examined to identify the specific clauses contained therein to tease out key trends and differences. This scoping review facilitates the main part of this article—an *anatomy*, or dissection, of a DTA, where we examine key features of this

type of agreement, and make recommendations on critical inclusions together with insight on why these clauses are necessary. This novel scoping review will animate parts of our discussion and assist in providing the guidance set out herein. Accordingly, the purpose of this contribution is to provide academics, researchers, scientists, ethicists, research managers, and all interested stakeholders with guidance on steps to take prior to executing a DTA, and insight into what to include in their own DTAs. Each case will no doubt have nuances and turn on its own facts. To be clear: There is no "one size fits all" template that can be uniformly applied without thought. However, there are many elements of a DTA that will be similar, and the holistic purpose of this discussion is to identify typical features of such an agreement, review best practice, and make recommendations for stakeholders going forward.

2 Scoping review

2.1 Methodology and results

Although there is no one standard DTA template and the contents of each will depend on certain factors, there are several clauses common to many DTAs. In order to ascertain and understand the key clauses used in DTAs that are contained in the health research space, we conducted a scoping review of publicly available and English-language DTAs. The search terms used are recorded in Supplementary Material 1. The inclusion criterion was that a DTA had to relate to data generally (and hence be inclusive of health research), or specific to health research. To ensure a broad, yet manageable spread of samples, our aim was to collect at least twenty samples, and to ensure that there are at least two samples from the Global South. When we reached twenty-four samples (B3 Africa, 2018; Bristol Myers Squibb, 2017; Clinical Study Data Request Consortium, 2015; Department of Health Western Australia, 2021; Dkfz German Cancer Research Center, 2020; FDP, 2017; Fred Hutch, 2020; GREGoR Consortium, 2022; Growing Up in New Zealand, 2014; Health Data Coalition, 2017; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; Infectious Diseases Data Observatory, 2021; Information Commissioner's Office, 2022; Johns Hopkins University, 2022; Kawartha Lakes OHT, 2020; KEMRI Wellcome Trust Research Programme, 2019; National Center for Advancing Translational Sciences, 2021; National Institute for Medical Research, 2020; National Health Service England, 2018; ONDC, 2024; Swiss Personalised Health Network, 2021; University of Newcastle, 2024; Utrecht University, 2024) containing two samples from Africa (B3 Africa, 2018; KEMRI Wellcome Trust Research Programme, 2019) and one from India (Indian Society of Critical Care Medicine, 2024), we decided that we had reached saturation point. Each DTA was examined to identify the specific clauses contained therein. The most frequently occurring clauses across the twenty-four DTAs were identified. These clauses were tabulated and categorized to facilitate a comprehensive comparison (see Supplementary Material 2). To gain insights into the prevalence and consistency of these clauses, their content was compared across all DTAs. Through this comparison, the common features shared by the clauses were identified—and encompassed language, structure, and substantive content of the provisions. The common clauses that we found were:

- Introduction (preamble/recitals), definitions, and parties.
- Purpose.
- Term and termination.
- Obligations on parties.
- Reporting and auditing.
- Intellectual property (and licensing).
- Data ownership.
- Publication and attribution.
- Confidentiality.
- Limitation of liability.
- General provisions (or miscellaneous).
- Governing law.
- Dispute resolution.

2.2 Limitations

Our study does have limitations. First, the sample size of twenty-four DTAs, while broad and representative of five continents, was intentionally kept to a size that we perceived as manageable. Second, our scoping review was confined to agreements available in the English language and freely accessible online. As such, the results may not fully capture the global landscape of DTAs. A further caveat is that terminology and definitions that are used in DTAs may vary across jurisdictions, and that the substantive provisions found in DTAs may cater for specific institutional needs or reflect domestic (national) legal requirements. Nevertheless, we suggest that the results of our scoping exercise are informative and useful. In the next section, we discuss the results in more detail.

3 Discussion: key features of a data transfer agreement

3.1 Introduction, definitions, and parties

Most commercial agreements begin with an introduction, also known as a preamble, or recitals (also referred to as "whereas" clauses). Like any good story or piece of writing, the introduction provides exactly that: An introduction to what is about to come. As Murray (2018) points out, this clause identifies the "who, what, when, and why" in the agreement. As noted by an English court in Toomey Motors v Chevrolet (2017), the fact that this clause is introductory in nature, does not mean its provisions are not binding, and these clauses may contain "operative provisions."

However, it is a matter of style and personal preference in deciding which clause comes first, and the order that follows. One might also see a definitions clause coming first, and that clause being followed by the introductory clause. The definitions clause is usually accompanied by an interpretation clause. This is a technical legal clause that provides a list of definitions and legal interpretative clauses. Usually, words used in the agreement with a capital letter will be defined terms and will be included in the definitions list—a definition is included in the agreement to assist with flow, and to aid the reader. For example, if a word or term has a long and/or complicated meaning, it is usually included in the definitions list (for example, "Intellectual Property" or "Processing Purpose").

Typically, near the start of the agreement, there is a clause that fully describes the parties to the agreement. This, as is the case with many parts of a contract, can be achieved in a multitude of ways: A clause on its own, or as part of the introduction, included in the definitions, or even on the cover page.

Given the importance of introductions and explanations, most of the DTAs that we analyzed included some form of introduction, definitions, and information about the parties—although these did vary depending on the DTA. Some provided an introduction or background (Bristol Myers Squibb, 2017; Clinical Study Data Request Consortium, 2015; Department of Health Western Australia, 2021; Fred Hutch, 2020; Growing Up in New Zealand, 2014; Health Data Coalition, 2017; Human Cell Atlas, 2019; Infectious Diseases Data Observatory, 2021; Information Commissioner's Office, 2022; National Center for Advancing Translational Sciences, 2021; ONDC, 2024; University of Newcastle, 2024), while others contained a recital (Dkfz German Cancer Research Center, 2020; Fred Hutch, 2020; Kawartha Lakes OHT, 2020; National Institute for Medical Research, 2020; Swiss Personalised Health Network, 2021; Indian Society of Critical Care Medicine, 2024). Some of the DTAs included a definitions section (Bristol Myers Squibb, 2017; Clinical Study Data Request Consortium, 2015; Department of Health Western Australia, 2021; Dkfz German Cancer Research Center, 2020; Fred Hutch, 2020; GREGoR Consortium, 2022; Growing Up in New Zealand, 2014; Health Data Coalition, 2017; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; Infectious Diseases Data Observatory, 2021; Kawartha Lakes OHT, 2020; KEMRI Wellcome Trust Research Programme, 2019; National Institute for Medical Research, 2020; Swiss Personalised Health Network, 2021; Utrecht University, 2024.), although in some it appeared as an appendix or glossary (GREGoR Consortium, 2022; Information Commissioner's Office, 2022; National Health Service England, 2018). All twenty-four DTAs provided information about the parties or a blank space in which information could be added.

3.2 Purpose

A purpose clause sets out the primary intention of the parties and articulates the nature of the agreement. This clause provides additional context, and sets out rights, responsibilities, and restrictions. In the context of a DTA, it is important to record the data transfer, the reason for the transfer, and note any important restrictions and obligations on the parties.

Ten of the DTAs that we examined contained a specific purpose clause (Department of Health Western Australia, 2021; Dkfz German Cancer Research Center, 2020; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; Information Commissioner's Office, 2022; Kawartha Lakes OHT, 2020; National Health Service England, 2018; ONDC, 2024; Swiss Personalised Health Network, 2021; University of Newcastle, 2024). Although some DTAs may not have a specific "purpose" clause, information relevant to the purpose was nevertheless included in other clauses (B3 Africa, 2018; Bristol Myers Squibb, 2017; Clinical Study Data Request Consortium, 2015; FDP, 2017; Fred Hutch, 2020; GREGoR Consortium, 2022; Health Data Coalition, 2017; Infectious Diseases Data Observatory, 2021; Johns Hopkins University, 2022; KEMRI Wellcome Trust Research Programme, 2019; National Center for Advancing Translational Sciences, 2021; National Institute for Medical Research, 2020; Utrecht University, 2024).

There is significant variance in how narrow or broad, general or detailed purpose statements are formulated, which may be a reflection of the legal tradition in the relevant jurisdiction. This is demonstrated by the following example:

> Improvements in information sharing, translate into many tangible benefits. Repeat diagnostic tests can be avoided. Medical errors are reduced and outcomes improved with quicker access to complete information. Time is saved by physicians, staff and patients. With less manual processing of information and fewer phone calls for results, patients can be cared for quicker.

> Ultimately patients will be more engaged in their care by leveraging the technology where providers and patients can securely access necessary PHI.

> Participants may include hospitals, healthcare organizations and healthcare providers involved in the circle of care that or who have direct involvement in the delivery of patient care, which requires the communication and sharing of patient information.

> This data sharing agreement is entered into by the Participants to enable more effective and efficient patient information sharing that then will translate into better patient care (Kawartha Lakes OHT, 2020).

By contrast, the University of Newcastle (2024) DTA contains a checklist of the various purposes for which the data is to be used:

> The Data is to be used only for the study of eating behaviours. Please indicate from the options below how you intend to use the Data:
>
> - Training and evaluation of new machine learning models for the detection of eating behaviours
> - Benchmarking existing machine learning models for the detection of eating behaviours
> - Creating and/or analyzing metrics of eating behaviors (e.g., eating pace and duration)
> - Other. Please specify (University of Newcastle, 2024):

An insightful drafting note is included in the National Health Service England (2018) template agreement in the purpose clause to assist those that use the template (the advice should be heeded in developing any purpose clause). An excerpt of it is below:

> Document the detail to explain the purpose and objectives of the information sharing . . . ensure that all parties affected by the information sharing are clear about why the information may be used. . . National Health Service England (2018)

We suggest that all the purposes of the sharing should be listed. It should be made clear which organization is processing the data

and for which purpose. It is important to specify this in sufficient detail and exactness, as DTAs typically limit the processing of the data by the recipient to the defined purpose. In other words, should the data recipient process the data for any purpose other than the defined purpose, it would be in breach of contract.

3.3 Term and termination

An important feature of any agreement is its term, and the manner of its termination. One must also be aware of the agreement's effective date (the date the agreement is binding from). With a DTA, usually there is a fixed term, with the ability for either party to give notice to the other to terminate (also known as cancellation for convenience—or no-fault termination—where one party does not need to give a reason for termination). Linked to this clause, one will usually also see a termination for fault or cause (a breach clause), and a clause which sets out how termination for convenience should be achieved. Some of the DTAs that we analyzed dealt with term and termination under one clause (Clinical Study Data Request Consortium, 2015; Bristol Myers Squibb, 2017; Fred Hutch, 2020; Infectious Diseases Data Observatory, 2021; Indian Society of Critical Care Medicine, 2024), while others either dealt with term and termination separately, or combined them with another clause (B3 Africa, 2018; Department of Health Western Australia, 2021; Dkfz German Cancer Research Center, 2020; FDP, 2017; GREGoR Consortium, 2022; Growing Up in New Zealand, 2014; Health Data Coalition, 2017; Human Cell Atlas, 2019; Information Commissioner's Office, 2022; Johns Hopkins University, 2022; Kawartha Lakes OHT, 2020; KEMRI Wellcome Trust Research Programme, 2019; National Center for Advancing Translational Sciences, 2021; National Institute for Medical Research, 2020; National Health Service England, 2018; Swiss Personalised Health Network, 2021; University of Newcastle, 2024; Utrecht University, 2024).

Above all, parties should know: (1) when the agreement is effective from; (2) how long it lasts for; and (3) how they can terminate the agreement, and under what circumstances. Below are two examples of this type of clause:

> Term and Termination. 22.1 This Agreement shall be effective as of the Effective Date and, unless cancelled or terminated earlier in accordance with the terms hereof, shall continue in effect until 30 September 2002 (the "Initial Term"). Thereafter, this Agreement shall continue in force and effect unless and until cancelled or terminated as provided in this Agreement (Law Insider, 2024).

> Termination for Convenience. Either party may terminate this Agreement without cause and at any time upon giving 30 days' prior written notice to the other party (each, a termination for "Convenience"). Such termination will be effective on the date stated in the notice (NetDocuments, 2024).

The first example displays a fixed term agreement clause where the agreement comes to an end on a specific date. Parties would also be able to terminate for cause on the basis of a clause found elsewhere in that agreement. The second example shows a termination for convenience clause where either party can terminate the agreement on notice without any fault and without having to give a reason. This type of clause provides maximum flexibility. Typically, where research institutions are involved, for the protection of both parties, one would want to see a termination for convenience clause so that a party is not forced to stay in a relationship that does not suit it. However, there may be economic or other factors that require the contract to exist for a long period, and for no termination for convenience to exist. Each case will turn on its own facts and this is a point parties must consider carefully.

Below are two examples of term and termination clauses found in the DTAs that we examined:

> 7.1. This Agreement shall come into force on the Effective Date and will remain in effect for a period of one (01) year from the Effective Date or on the expiration of a thirty (30) days' written notice by either party.
>
> 7.2. This Agreement will terminate immediately upon any breach of the provisions of this Agreement by the Recipient or by any of the Registered Users.
>
> 7.3. In the event that this Agreement is terminated in accordance with this Clause 7.1 or 7.2, the Recipient shall return or destroy all Data at the direction of the Provider (Indian Society of Critical Care Medicine, 2024).

And:

> This Agreement will expire on the completion of the Research and completion of the publications included in the Publication Plan but in no event later than three (3) years from the Effective Date. BMS may terminate this Agreement for Institution's material breach of its terms, where the breach is not cured within thirty (30) days following receipt of written notice of same. Upon termination or expiration of this Agreement the rights and obligations of the Parties which have accrued hereunder shall survive in accordance with their terms, and Institution's right to use BMS Confidential Information shall immediately cease. The terms of Section 3 (Term and Termination), 4 (Institution Representations, Warranties and Covenants), 5 (Confidentiality), 6 (Publication), 7 (Inventions), 8 (Miscellaneous) shall survive the expiration or termination of this Agreement (Bristol Myers Squibb, 2017).

3.4 Obligations on parties

The clause (or clauses) that set out the main obligations of the parties can be drafted in many ways, and different headings can be used. Twenty-one of the DTAs that we examined contained a clause (or information) detailing the obligations or duties of the parties to the agreement (B3 Africa, 2018; Clinical Study Data Request Consortium, 2015; Department of Health Western Australia, 2021; Dkfz German Cancer Research Center, 2020; FDP, 2017; Fred Hutch, 2020; GREGoR Consortium, 2022; Growing Up in New Zealand, 2014; Health Data Coalition, 2017; Human Cell Atlas,

2019; Infectious Diseases Data Observatory, 2021; Information Commissioner's Office, 2022; Johns Hopkins University, 2022; Kawartha Lakes OHT, 2020; KEMRI Wellcome Trust Research Programme, 2019; National Center for Advancing Translational Sciences, 2021; National Institute for Medical Research, 2020; National Health Service England, 2018; Swiss Personalised Health Network, 2021; University of Newcastle, 2024; Utrecht University, 2024).

Below is an excerpt of a DTA clause which lists the obligations (we have only reproduced part of the clause because of its length) of the parties:

> It is hereby agreed that the following conditions to the Agreement shall be binding on the RECIPIENT:
>
> (a) The RECIPIENT agrees to use, store or dispose of the DATA in compliance with all applicable laws including those relating to research involving the use of human and animal subjects.
>
> (b) The DATA shall remain the property of the PROVIDER and PROVIDER hereby consents to the DATA being made available as a service to the research community.
>
> (c) The RECIPIENT shall use the DATA for teaching or academic research purposes only.
>
> It is hereby agreed that the following conditions to the Agreement shall be binding on the PROVIDER:
>
> (a) The PROVIDER agrees to transfer, store or dispose of the DATA in compliance with all applicable laws
>
> (b) The PROVIDER shall transfer immediately the DATA upon receipt of one of the two copies duly signed by the RECIPIENT (National Institute for Medical Research, 2020).

This clause creates contractual obligations (or duties) on both parties. Usually, one would expect to find the key responsibilities of the parties in this clause. In the context of a DTA, primarily, one should ensure the clause places obligations on the parties to comply with the conditions of lawful processing set out in South Africa's Protection of Personal Information Act 4 of 2013 (POPIA), 2013 (or equivalent international legislation). As can be seen in the example above, both the provider and recipient have a duty to ensure compliance with "all applicable laws"—one could craft this to specifically refer to data protection legislation, such as POPIA.

Typically, one would also see obligations on the parties in relation to dealing with data after the relationship ends (in other words, to return or delete it), and in terms of how to *use* the data (such as for teaching or academic research purposes only). If there are specific requirements or nuances to a project, this is the clause that will list those requirements. We suggest that parties give careful thought to what the project entails—simply put, what is it each party needs to do in order to achieve a successful outcome, and then to ensure these obligations are listed in this clause.

Holistically, we suggest that a DTA can be a useful tool to facilitate compliance with data protection legislation. In this context, parties may consider including provisions that relate to the following:

- The ground of justification for the transfer;
- The manner in which the data was collected, how it will be processed, transferred, stored, and disposed of;
- Data subject access rights;
- Appropriate technical and organizational measures are taken, and that adequate safeguards are in place;
- Measures in place in relation to cross border data flows;
- Conditions and restrictions in place in relation to further processing of data beyond.

Parties should also ensure that the details and mechanics of the data being transferred are included in the agreement. As all of the agreements that we examined are DTAs, they all mention the transfer of data in some form. However, not all DTAs described the mechanics of such transfers (Growing Up in New Zealand, 2014; Clinical Study Data Request Consortium, 2015; Bristol Myers Squibb, 2017; Health Data Coalition, 2017; National Health Service England, 2018; KEMRI Wellcome Trust Research Programme, 2019; Dkfz German Cancer Research Center, 2020; Kawartha Lakes OHT, 2020; Department of Health Western Australia, 2021; Indian Society of Critical Care Medicine, 2024; University of Newcastle, 2024; Utrecht University, 2024). Twelve DTAs provided more detailed guidance relating to transfers of data (B3 Africa, 2018; FDP, 2017; Fred Hutch, 2020; GREGoR Consortium, 2022; Human Cell Atlas, 2019; Infectious Diseases Data Observatory, 2021; Information Commissioner's Office, 2022; Johns Hopkins University, 2022; National Center for Advancing Translational Sciences, 2021; National Institute for Medical Research, 2020; ONDC, 2024; Swiss Personalised Health Network, 2021). For practical reasons, this could be an annexure. Only six DTAs provided for the transfer of data in an annexure (FDP, 2017; Human Cell Atlas, 2019; Fred Hutch, 2020; National Institute for Medical Research, 2020; Swiss Personalised Health Network, 2021; Johns Hopkins University, 2022).

3.5 Reporting and auditing

An example of an a-typical clause in a DTA relates to auditing and reporting. Only two of the DTAs in our scoping review contained an audit clause (Growing Up in New Zealand, 2014; Kawartha Lakes OHT, 2020), which appear as follows:

> The Privacy Officer of each Participant shall audit access to PHI for which the Participant is the Custodian, including without limitation access by its Authorized Users (Kawartha Lakes OHT, 2020).

And:

> A representative of UniServices will be permitted access by the Institution, at all reasonable times, to the results and analyses obtained from the use of the Data Set together with any records and documents relating thereto for the purpose of verifying compliance with the conditions of this Agreement. The Institution will provide UniServices with any information which UniServices reasonably requests in relation to the Institution's compliance with this Agreement (Growing Up in New Zealand, 2014).

The primary purpose of a clause such as this is to allow the provider to ensure that the recipient is taking adequate steps to

comply with its obligations. Despite the importance of this clause, very few of the DTAs that we analyzed contained specific clauses relevant to reporting and auditing (Growing Up in New Zealand, 2014; Kawartha Lakes OHT, 2020; Infectious Diseases Data Observatory, 2021). None of the DTAs examined contained a specific reporting clause, and in eight of the DTAs reporting is instead mentioned either generally throughout the agreement or under another clause (Health Data Coalition, 2017; National Health Service England, 2018; Kawartha Lakes OHT, 2020; Infectious Diseases Data Observatory, 2021; Swiss Personalised Health Network, 2021; GREGoR Consortium, 2022; Indian Society of Critical Care Medicine, 2024; University of Newcastle, 2024).

One will also see clauses that require one party to report to the other in relation to, for example, processing activities with the data, and safeguards in place—and in some cases this type of obligation may be found in the main obligations clause discussed above in 3.4. Parties should consider what best suits their needs in the context of the data involved. However, we do suggest parties should have some ability to assess whether the other party is complying with the agreement.

3.6 Intellectual property and licensing

A specific intellectual property (IP) clause was present in twelve of the DTAs that we examined (B3 Africa, 2018; Clinical Study Data Request Consortium, 2015; Department of Health Western Australia, 2021; Dkfz German Cancer Research Center, 2020; Fred Hutch, 2020; Growing Up in New Zealand, 2014; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; Infectious Diseases Data Observatory, 2021; Swiss Personalised Health Network, 2021; University of Newcastle, 2024; Utrecht University, 2024), with another four dealing with IP under other clauses (Bristol Myers Squibb, 2017; KEMRI Wellcome Trust Research Programme, 2019; Kawartha Lakes OHT, 2020; National Institute for Medical Research, 2020).

IP clauses are like Janus, with one face looking back and one face looking forward. It looks back in the sense that it recognizes pre-existing IP rights, often termed as "background" IP. It also looks forward, and provides for rights in any new IP that is created by the Recipient using the Project Data. Typically, the Recipient will own the IP that it creates using the Project Data, but this can be negotiated. For example, the Recipient can grant a perpetual nontransferable use-license to the Provider in the IP that it creates, or the parties can be joint owners of the IP. Here is an example of a simple IP clause:

> Except for the rights explicitly granted hereunder, nothing contained in this Agreement shall be construed as conveying any rights under any patents or other intellectual property which either Party may have or may hereafter obtain (Human Cell Atlas, 2019).

Licensing is often dealt with under the IP clause. Ten of the DTAs from our scoping review include licensing within IP (Clinical Study Data Request Consortium, 2015; Bristol Myers Squibb, 2017; Human Cell Atlas, 2019; Dkfz German Cancer Research Center, 2020; Fred Hutch, 2020; Department of Health Western Australia, 2021; Infectious Diseases Data Observatory, 2021; Swiss Personalised Health Network, 2021; Indian Society of Critical Care Medicine, 2024; University of Newcastle, 2024). Examples of licensing provisions (within an IP clause) is as follows:

> Provider grants to Recipient the non-exclusive, worldwide, perpetual, sub-licensable, royalty-free, fully paid up license to use all Data for Recipient's non-commercial, research and educational purposes (Indian Society of Critical Care Medicine, 2024).

And:

> Subject to any pre-existing rights, obligations, options to license, or licenses granted by the Provider and/or Recipient to a third party, the Recipient and Provider retain or are granted a non-exclusive royalty-free license to use an Invention developed under the Purpose for their own research, educational, patient care purposes but not for Commercial Use unless otherwise outlined in the Implementing Letter (Fred Hutch, 2020).

Licensing is mentioned in relation to ownership as well as commercialization, as can be seen below:

> The University grants the Recipient Organisation a non-exclusive, non-transferable, fee-free licence to use the Data for the Purpose only.

> If the Recipient Organisation wishes to commercialise or have commercialised any Results or Data IP, or otherwise deal in the Data or Derivatives for any commercial purpose, it must first enter into an appropriate licence agreement with the University (University of Newcastle, 2024).

Next, we consider data ownership. It is important to note that although both data ownership and IP pertain to incorporeal objects, data ownership and IP are distinct legal concepts and are governed by different legal rules.

3.7 Data ownership

The Project Data would presumably consist of one or more computer files—i.e., digital objects. Each of these digital objects has an independent existence in the digital world, has value and usefulness, and can be controlled by humans. As such, in legal systems that have a basis in Roman Law, the Project Data should be susceptible of being owned (Thaldar et al., 2022). Yet, data ownership remains controversial in the West. By contrast, China is leading the way with the adoption of a policy on the commercialization of data, released in 2022 (Xiong et al., 2023). This policy provides for various property rights modules in data. If the data contains personal information, a privacy module applies to the data in addition to the property rights modules. With China officially endorsing data as legal property, we suggest that it would be unwise for the rest of the world to remain in data ownership purgatory.

It is essential to address and dispel the primary objection to data ownership, especially concerning personal data. This argument is structured as follows:

- Premise 1: In certain situations, the ownership rights of a data generator (like a university) might conflict with the privacy rights of data subjects.
- Premise 2: Political and legal policies underscore the importance of data privacy rights, as evidenced by the growing body of global legislation on the matter.
- Conclusion: Therefore, data ownership is viewed as politically and legally untenable.

While the premises are true, the conclusion does not necessarily hold. Thaldar et al. (2022) argue that ownership is always encumbered in some way, depending on the nature of the object and the circumstances. In the context of personal data, ownership is encumbered by privacy rights, allowing for a reconciliation between data ownership and data privacy. This perspective aligns with China's approach that provides that if data is personal data, the property rights in such data are superseded by the privacy rights of the data subjects. In a recent article, Thaldar (2024) turns the anti-data-ownership argument on its head by showing that research institutions can only properly fulfil their statutory duties to protect the personal data in their care if they actively claim ownership in such data. Thaldar (2024) uses an example of a person who has lawful access to the data at a research institution, such as a research collaborator or a student, who makes a copy of the file containing the relevant data on her own memory stick and deletes the original file from the research institution's system. Subsequently, the person declares herself the owner of the data contained in the file on the memory stick. If the research institution shunned data ownership, it has none of the well-established civil and criminal remedies of an owner available. It will have to rely on its contractual relationship with the person who took the data, which places it in a significantly weaker position. As Thaldar (2024) concludes, data ownership is a precondition for being an effective data custodian.

In agreements like DTAs, we propose that while ensuring the protection of individuals' data privacy rights through contractual obligations is crucial, as discussed above under Section 3.4, it is equally important to explicitly articulate ownership rights. This dual focus can harmonize the protection of privacy with the recognition of data as a valuable and ownable asset.

Let's now consider the results of the scoping review. Sixteen of the DTAs that are part of our scoping review mention "ownership." However, on closer inspection, only six of these DTAs unambiguously provide for *data ownership*—i.e., where the object of ownership is *data per se*, as distinct from *rights in data*, such as IP rights in data (B3 Africa, 2018; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; National Institute for Medical Research, 2020; Swiss Personalised Health Network, 2021; University of Newcastle, 2024). This is an important distinction. Claiming only IP rights in data and remaining silent about the data itself, means that ownership of the data itself—which is independent of any IP rights in the data—remains unresolved. Yet, this is the case in the majority of the DTAs that we reviewed. Two DTAs even conflate the objects of ownership (Dkfz German Cancer Research Center, 2020; Utrecht University, 2024). For example, one DTA provides:

> The RECIPIENT recognizes that nothing in this Agreement shall operate to transfer to the RECIPIENT or its RECIPIENT SCIENTISTs any INTELLECTUAL PROPERTY rights in or relating to the DATA, i.e., ownership of DATA remains unchanged (Dkfz German Cancer Research Center, 2020).

This kind of conceptual confusion should be avoided. A clear data ownership provision, such as the following simple provision should be included in any DTA:

> As this is an ISCCM initiated project, the entire ownership of the data will be with the ISCCM (Indian Society of Critical Care Medicine, 2024).

It is important that data ownership exists independently and distinctly from ownership of rights in the data, such as IP rights. As such, it makes sense to deal with these two kinds of objects of ownership in under separate headings. However, it can also be successfully combined in a single clause, provided that the concepts are not conflated, as illustrated by the following provision:

> The Receiving Institute will own all Research Data, results, inventions, copyright in datasets, sui generis database rights, and all associated rights, which arise which arise under the Research Project described in Appendix A (Human Cell Atlas, 2019).

An argument that is sometimes heard in academic circles is that because there is legal uncertainty about data ownership in a given jurisdiction, referring to data ownership should best be avoided as a component of a DTA. This argument is mistaken. If there is still a dearth of caselaw on data ownership in a given jurisdiction, resulting in the issue not yet being settled law, this fact is good reason to *explicitly* provide for data ownership in a DTA—in this way, the parties are bound to the agreed position. For example, if a recipient agreed that the provider is the owner of the project data (*qua* well-defined digital object), the recipient could be estopped from later asserting in court that the provider is not the owner. Accordingly, including an explicit data ownership provision in a DTA creates legal certainty—even in an environment of general uncertainty.

3.8 Publication and attribution

Typically, in data transfers involving universities or research institutions, one can expect to see a clause regulating the publication of results and/or academic publications. Only one DTA in our scoping review contained a specific attribution clause (Human Cell Atlas, 2019), but thirteen DTAs included publication clauses (B3 Africa, 2018; Bristol Myers Squibb, 2017; Clinical Study Data Request Consortium, 2015; Department of Health Western Australia, 2021; Dkfz German Cancer Research Center, 2020; Fred Hutch, 2020; Health Data Coalition, 2017; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; Infectious Diseases Data Observatory, 2021; Swiss Personalised Health Network, 2021; University of Newcastle, 2024; Utrecht University, 2024). In eight of the DTAs, publication was mentioned under another clause (Growing Up in New Zealand, 2014; FDP, 2017; KEMRI Wellcome Trust Research Programme,

2019; National Institute for Medical Research, 2020; National Center for Advancing Translational Sciences, 2021; GREGoR Consortium, 2022; Johns Hopkins University, 2022; ONDC, 2024).

As a starting point, we recommend that no results are released unless the other party consents. However, it is not unusual to expect that the party who provided the data would want the right to stipulate whether or not the results are published, and to retain the right to derive benefit from academic publications.

Further, given obligations imposed by data protection legislation, it is prudent to insert a provision regulating how results are made public. An example may appear as follows:

> As SPHN projects are funded with public money, the Parties strive to make the resulting scientific publications publicly accessible and available through Open access as far as possible according to publishers rights (Swiss Personalised Health Network, 2021).

And:

> The Receiving Institute must endeavour to publish results in an open access academic journal or database (Human Cell Atlas, 2019).

One would also expect to see something here, including an obligation to make acknowledgments. Fifteen DTAs required acknowledgements to be made in publications arising from the provider's data (B3 Africa, 2018; Bristol Myers Squibb, 2017; Department of Health Western Australia, 2021; Dkfz German Cancer Research Center, 2020; Fred Hutch, 2020; Growing Up in New Zealand, 2014; Health Data Coalition, 2017; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; Infectious Diseases Data Observatory, 2021; KEMRI Wellcome Trust Research Programme, 2019; National Center for Advancing Translational Sciences, 2021; National Institute for Medical Research, 2020; Swiss Personalised Health Network, 2021; University of Newcastle, 2024). An example of such a provision reads as follows:

> Publications: Unless directed otherwise, HDC must be acknowledged in any publication or presentation using HDC data, and the following disclaimer must appear on any materials developed for public distribution with data used under this DSA: "The views expressed herein do not necessarily represent the views of HDC (Health Data Coalition, 2017)."

And:

> Recipient will acknowledge the Provider as the source of the Data in any publication reporting on its use, unless requested otherwise by the Provider (Indian Society of Critical Care Medicine, 2024).

And:

> The Institution will ensure that all outputs that are intended for publication, including (but not necessarily limited to) reports, journal papers, working papers, conference and other public presentations, and other documents, contains an acknowledgement that the Data Set has been sourced from The University of Auckland, Growing Up in New Zealand: Longitudinal Study of New Zealand Children and Families, together with an appropriate acknowledgement of the funders of the study, all of which must be approved by the Data Access Committee in writing prior to the publication (Growing Up in New Zealand, 2014).

3.9 Confidentiality

A confidentiality provision is a standard clause in any commercial agreement, and a DTA is no exception. As with any other clause, there are many ways to draft this—typically, the clause stipulates that each party will keep all information (which will be broadly defined) confidential, and will not, without the prior written consent of the other party, disclose to any person any of the confidential information. This prohibition on disclosure of confidential information will usually not preclude any party from making any disclosure to its professional advisors (provided that the advisors ensure the information remains confidential). Further, it will preclude a party from making any disclosure which it is required to make by law (such as in the course of an investigation around a data breach).

The importance of confidentiality can be seen in the fact that thirteen of the DTAs in our scoping review contained a dedicated clause dealing with confidentiality (B3 Africa, 2018; Bristol Myers Squibb, 2017; Clinical Study Data Request Consortium, 2015; Department of Health Western Australia, 2021; Dkfz German Cancer Research Center, 2020; Fred Hutch, 2020; Health Data Coalition, 2017; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; Kawartha Lakes OHT, 2020; Swiss Personalised Health Network, 2021; University of Newcastle, 2024; Utrecht University, 2024). An additional five DTAs, although not including a dedicated confidentiality clause, mentioned confidentiality—in some form or another—throughout the DTA (National Health Service England, 2018; KEMRI Wellcome Trust Research Programme, 2019; Infectious Diseases Data Observatory, 2021; National Center for Advancing Translational Sciences, 2021; GREGoR Consortium, 2022). An example of a confidentiality clause is as follows:

> Either PARTY shall treat the CONFIDENTIAL INFORMATION confidential for the duration of this Agreement, including any extension thereof, and thereafter for a period of five (5) years following termination or expiry of this Agreement. Excluded from this obligation of confidentiality shall be any CONFIDENTIAL INFORMATION of which one PARTY can reasonably demonstrate that it (a) was previously known to them, or (b) is, and/or becomes, publicly available during said five (5) year period through no fault of a PARTY, or (c) is independently and lawfully developed by one PARTY. This obligation of confidentiality shall not apply to any disclosure required by law, provided that the RECIPIENT shall notify the PROVIDER of any disclosure required by law in sufficient time so that the PROVIDER may contest such requirement, if the PROVIDER so chooses. Subject to mandatory law, upon the expiration or

> termination of this Agreement for whatever reason, or at the earlier request of a PARTY, the other PARTY shall, at its own costs, return or destroy all originals and copies of CONFIDENTIAL INFORMATION, or, in case of CONFIDENTIAL INFORMATION stored in electronic, magnetic or digital media, shall erase or render unreadable all materials furnished (including without limitation, working papers containing any CONFIDENTIAL INFORMATION or extracts therefrom) which contain CONFIDENTIAL INFORMATION (Swiss Personalised Health Network, 2021).

And—note that Human Cell Atlas (2019) defines "Research Materials" to include, *inter alia*, "Research Data collected for the Research Project":

> 8.1. The Information may include confidential information of the Providing Institute. Accordingly, if and to the extent that any such Information is marked as "confidential," the Receiving Institute shall during the Term of this Agreement and for a period of [*insert period*] following its termination, treat such Information as confidential and only disclose it under like obligations of confidentiality and Restrictions on Use as those contained herein. The Receiving Institute shall be deemed to have fulfilled its obligation if it [*insert local criteria applicable to confidentiality standards/ requirements*].
>
> 8.2. The above-mentioned obligations of confidentiality shall not apply to Information which:
>
> 8.2.1. [*If contributing derived Research Data to the HCA:* Is identified as Research Data to be contributed to the HCA by the Providing Institute/Receiving Institute, as listed in Appendix A]; or
>
> 8.2.2. Can be shown to have been known to the Receiving Institute at the time of its acquisition from Providing Institute; or
>
> 8.2.3. Is acquired from a third party, not in breach of any confidentiality obligation to the Providing Institute; or
>
> 8.2.4. Is independently devised or arrived at by, on behalf of, or for the Receiving Institute without access to the Information; or
>
> 8.2.5. Enters the public domain otherwise than by breach of the undertakings set out in this Agreement.
>
> 8.3. In some cases, the Research Materials may also incorporate confidential Information pertaining to research participants or donors having provided the Research Materials. The Research Materials provided to the Receiving Institute have been [*enter information related to de-identification processes applied to the data, e.g., coded, double-coded, anonymized, anonymous (provide description of de-identification measures)*]. If the Receiving Institute inadvertently receives Information that identifies individual research participants or donors, the Receiving Institute will take all reasonable and appropriate steps to protect the privacy and confidentiality of such Information. This may require immediate destruction of the Research Materials on request of the Providing Institute. The Receiving Institute agrees to make no intentional attempt to re-identify research participants or donors, through linkage of data or otherwise. The Receiving Institute will immediately report any identification of research participants or donors to the Providing Institute (Human Cell Atlas, 2019).

3.10 Limitation of liability

Another of the "boilerplate" clauses (those clauses which you see in almost every commercial agreement, irrespective of what the agreement regulates), is a clause limiting the liability of the parties—sometimes, this may be coupled with indemnities. In larger, more complicated commercial deals these two clauses will be separated, but for purposes of a data transfer, it may well be that one can combine them. Nine of the DTAs in our scoping review contained a liability clause (which is often combined with warranties) (B3 Africa, 2018; Department of Health Western Australia, 2021; Growing Up in New Zealand, 2014; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; Information Commissioner's Office, 2022; Kawartha Lakes OHT, 2020; Swiss Personalised Health Network, 2021; Utrecht University, 2024). In eleven DTAs, liability is mentioned under another clause, such as limitations and exclusions (Infectious Diseases Data Observatory, 2021), data sharing (Clinical Study Data Request Consortium, 2015), disclaimer (GREGoR Consortium, 2022), terms and conditions (FDP, 2017; National Center for Advancing Translational Sciences, 2021; Johns Hopkins University, 2022), indemnification (Fred Hutch, 2020), warranty and indemnities (University of Newcastle, 2024), legal statement (Dkfz German Cancer Research Center, 2020), remedies and no waiver (National Health Service England, 2018), and obligations of provider and recipient (National Institute for Medical Research, 2020).

In terms of the various limitation of liability provisions, for the most part, where the clause exists, it attempts to protect the provider of the data, and ensure that it will not be liable for damages relating from the use or transfer of the data (B3 Africa, 2018; Dkfz German Cancer Research Center, 2020; FDP, 2017; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; Infectious Diseases Data Observatory, 2021; Johns Hopkins University, 2022; National Institute for Medical Research, 2020; Utrecht University, 2024).

An example of this clause is as follows:

> Providing Institute will not be liable for damages related to the provision of Research Materials to the Receiving Institute. This includes but is not limited to damages in relation to inaccuracies, lack of comprehensiveness, or use of the Research Materials, or any delays or break in supply by the Providing Institute (Human Cell Atlas, 2019).

Interestingly, two of the DTAs required that parties take out, and maintain, liability insurance for the duration of the agreement (Kawartha Lakes OHT, 2020; Department of Health Western

Australia, 2021)—with one DTA specifying the value of the insurance (Kawartha Lakes OHT, 2020).

We suggest that the limitations, as far as possible, should be reciprocal, and that both parties indemnify each other from unlawful conduct. Importantly, both parties should identify a figure that represents the entire amount any party could claim from another. The context will determine the appropriate figure, and this will be informed by the level of risk, insurance cost, and benefit derived from the project.

Parties should also ensure that neither party will be liable for loss of profits or consequential damages arising out of the project.

Further examples of liability clauses are as follows:

> 11.1 Providing Institute makes no warranty, either express or implied, of the fitness for purpose of the Research Material. However, to the best of Providing Institute's knowledge, the use of the Research Materials within the Purpose of Use shall not infringe on the proprietary rights of any third party.
>
> 11.2 Providing Institute will not be liable for damages related to the provision of Research Materials to the Receiving Institute. This includes but is not limited to damages in relation to inaccuracies, lack of comprehensiveness, or use of the Research Materials, or any delays or break in supply by the Providing Institute. The Receiving Institute acknowledges that the Providing Institute makes no guarantee that the Research Materials are free of contamination from viruses, latent viral genomes, or other infectious agents. The Receiving Institute agrees to treat the Research Materials as if they were not free from contamination, to ensure that appropriate biosafety training is provided to research personnel, and to implement appropriate biohazard containment measures.
>
> 11.3 The Receiving Institute agrees that, except as may explicitly be provided for in this Agreement, the Providing Institute has no control over the use that is made of the Research Materials or the Information by the Receiving Institute in accordance with the terms of this Agreement. Consequently, the Receiving Institute agrees that Providing Institute shall not be liable for such use.
>
> 11.4 The Receiving Institute will not be liable for damages incurred by the Providing Institute in providing the Research Materials to the Receiving Institute. This includes but is not limited to damages incurred through the Providing Institute's breach of contract or statute, its breach of institutional policy, research ethics requirements, as well as any tortious or extracontractual liability incurred (Human Cell Atlas, 2019).

And:

> Except to the extent prohibited by law, the Recipient assumes all liability for damages which may arise from its use, storage, disclosure, or disposal of the Data. The Provider will not be liable to the Recipient for any loss, claim, or demand made by the Recipient, or made against the Recipient by any other party, due to or arising from the use of the Data by the Recipient, except to the extent permitted by law when caused by the gross negligence or willful misconduct of the Provider (FDP, 2017).

And:

> 6.1 Nothing in this Agreement excludes or limits the liability of either Party:
>
> 6.1.1 for death or personal injury caused by that Party's negligence; or
>
> 6.1.2 for fraud or fraudulent misrepresentation; or
>
> 6.1.3 to the extent that such liability cannot be limited or excluded by law.
>
> 6.2 Subject to Clause 6.1, in no event will the University of Oxford or the Data Contributor(s) be liable for any use of the Dataset by the Recipient, whether in contract, tort (including negligence or breach of statutory duty) or otherwise howsoever arising (Infectious Diseases Data Observatory, 2021).

3.11 General provisions (miscellaneous)

Fourteen of the twenty-four DTAs that we examined contained a heading for general provisions, or sometimes called "Miscellaneous" (B3 Africa, 2018; Bristol Myers Squibb, 2017; Department of Health Western Australia, 2021; GREGoR Consortium, 2022; Growing Up in New Zealand, 2014; Health Data Coalition, 2017; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; Infectious Diseases Data Observatory, 2021; Information Commissioner's Office, 2022; National Health Service England, 2018; Swiss Personalised Health Network, 2021; University of Newcastle, 2024; Utrecht University 2024). The general clauses serve as a backbone to the overall contract, addressing various fundamental legal, operational, and administrative aspects that govern the relationship between the parties involved. These clauses are pivotal for ensuring clarity, legality, and fair practice in data transfers. The components (or sub-clauses) commonly found in these clauses are as follows:

- Waiver: This provision clarifies that the failure or delay in enforcing any part of the agreement does not constitute a waiver of rights.
- Assignment and Novation: This provision dictates the conditions under which parties can transfer their rights and obligations under the DTA to another party.
- Relationship of the Parties: It clarifies that the DTA does not create a partnership, joint venture, or agency relationship between the parties.
- Amendment: This specifies that changes to the DTA must be made in writing and signed by all parties.
- Severability: If any part of the DTA is found to be invalid or unenforceable, this provision allows for that part to be removed without affecting the remainder of the DTA.

- Entire Agreement: This provision states that the DTA constitutes the full and complete agreement between the parties, superseding all prior discussions and agreements.

Less common, but very useful components of general clauses are:

- Survival Clause: This provision specifies which provisions of the agreement will continue to be effective after the termination or expiry of the agreement. For example, the Department of Health Western Australia (2021) DTA specifies that certain clauses will survive the termination or expiry of the agreement.
- Counterparts: Some DTAs, like those of National Health Service England (2018), allow the agreement to be executed in counterparts, meaning separate copies can be signed and assembled to form the complete agreement.
- Contact Points and Notices: This provision specifies how formal communications related to the DTA should be made, often requiring written notices, as seen in the University of Newcastle (2024) DTA.
- Electronic Signatures and Form: With the advancement of technology, some DTAs, like the Swiss Personalised Health Network (2021), acknowledge electronic signatures and communications.

Two provisions that are sometimes found as sub-clauses under the general clause, but also frequently as self-standing clauses, are governing law and dispute resolution. We discuss these two provisions next.

3.12 Governing law

The inclusion of a governing law provision is a fundamental aspect of a DTA, as it establishes which country's law will govern the interpretation of the DTA. Typically, a governing law provision will also provide which court within the relevant country has jurisdiction to adjudicate disputes that arise from the DTA. In our analysis, it was observed that almost all the DTAs reviewed incorporate a governing law provision. Only six DTAs (FDP, 2017; KEMRI Wellcome Trust Research Programme, 2019; National Center for Advancing Translational Sciences, 2021; GREGoR Consortium, 2022; Johns Hopkins University, 2022; ONDC, 2024) eschew this essential element.

The DTAs that contain a governing law provision typically specify the country whose law will govern the DTA. However, in two cases, the Human Cell Atlas (2019) and B3 Africa (2018), the choice of jurisdiction is left open for the parties to decide.

Interestingly, among the eighteen DTAs that do include a governing law provision, ten delineate it as an independent clause (Clinical Study Data Request Consortium, 2015; Dkfz German Cancer Research Center, 2020; Fred Hutch, 2020; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; Infectious Diseases Data Obervatory, 2021; Information Commissioner's Office, 2022; Kawartha Lakes OHT, 2020; National Institute for Medical Research, 2020; Swiss Personalised Health Network, 2021), seven integrate it within the general or miscellaneous provisions (B3 Africa, 2018; Bristol Myers Squibb, 2017; Department of Health Western Australia, 2021; Growing Up in New Zealand, 2014; National Health Service England, 2018; University of Newcastle, 2024; Utrecht University 2024), and one defines the governing law under its definitions/interpretations section (Health Data Coalition, 2017). This differentiation in presentation underscores the varied approaches to structuring DTAs.

An example of a governing law clause is found in the Utrecht University (2024) DTA. It reads as follows:

> This agreement will be governed by the laws of Netherlands and disputes concerning its execution will be put before the competent district court of Utrecht (Utrecht University, 2024).

We suggest that this concise example is worth emulation. The absence of such a governing law clause means that resolving disputes could become complicated, potentially necessitating judicial intervention to ascertain applicable laws. Such situations could lead to unforeseen legal entanglements and protracted disputes, which could counteract the purpose of the DTA.

3.13 Dispute resolution

Most DTAs in our scoping review dealt with dispute resolution in some form. Of the twenty-four DTAs that we examined, four contained a dedicated dispute resolution clause (Growing Up in New Zealand, 2014; National Health Service England, 2018; Human Cell Atlas, 2019; Kawartha Lakes OHT, 2020). Twelve of the DTAs dealt with (or simply mentioned) dispute resolution under another clause (B3 Africa, 2018; Bristol Myers Squibb, 2017; Clinical Study Data Request Consortium, 2015; Department of Health Western Australia, 2021; Dkfz German Cancer Research Center, 2020; Fred Hutch, 2020; Health Data Coalition, 2017; Indian Society of Critical Care Medicine, 2024; Information Commissioner's Office, 2022; National Institute for Medical Research, 2020; Swiss Personalised Health Network, 2021; Utrecht University, 2024)—most commonly the governing law clause or the general provisions clause. Those DTAs that dealt with disputes under the governing law or general provisions clauses referred to the jurisdiction and the laws that will apply (Bristol Myers Squibb, 2017; Clinical Study Data Request Consortium, 2015; Dkfz German Cancer Research Center, 2020; Indian Society of Critical Care Medicine, 2024; Information Commissioner's Office, 2022; Swiss Personalised Health Network, 2021; Utrecht University, 2024). Others mentioned alternative dispute resolution mechanisms, such as arbitration, negotiation, and mediation (B3 Africa, 2018; Department of Health Western Australia, 2021; Fred Hutch, 2020; Growing Up in New Zealand, 2014; Health Data Coalition, 2017; Human Cell Atlas, 2019; Indian Society of Critical Care Medicine, 2024; Information Commissioner's Office, 2022; Kawartha Lakes OHT, 2020; National Institute for Medical Research, 2020; National Health Service England, 2018).

Holistically, it is important to ensure that the clause provides clarity on how a dispute will be managed—and, in our view, a tiered approach is best in this type of relationship. What do we mean by a tiered approach? The parties should be obliged to try and meet first to find a solution to the dispute by negotiation (usually senior representatives from both sides), failing that, a

formal mediation, and then an arbitration using well known rules. However, parties should include a provision that acknowledges that either party may be able to approach a court of law on an urgent basis. In some cases, parties may need urgent or interim relief pending the outcome of the negotiations, mediation, or arbitration, and it is wise to ensure that a party is not prevented from seeking such urgent, interim relief.

We also suggest including a provision to stipulate that the mediation or arbitration will be held via video conferencing, unless the parties agree otherwise—this will likely assist from a cost saving perspective, and should also expediate matters. Following the COVID-19 pandemic, video conferencing, such as Zoom and Microsoft Teams, is commonplace.

Ultimately, a dispute resolution clause should provide the parties with an efficient, pragmatic, and cost-effective manner to resolve any dispute. An example of a dispute resolution clause is as follows:

> 16.1 All disputes arising out of or in connection with this Agreement shall be settled under the Rules of Arbitration of the International Chamber of Commerce by one or more arbitrators appointed in accordance with the said Rules.
>
> 16.2 The Parties agree, pursuant to Article 30 (2) (b) of the Rules of Arbitration of the International Chamber of Commerce, that the Expedited Procedure Rules shall apply, provided the amount in dispute does not exceed US$ [specify amount] at the time of the communication referred to in Article 1 (3) of the Expedited Procedure Rules.
>
> 16.3 The Parties agree that arbitration shall be conducted in [CITY] at [PLACE].
>
> 16.4 Legal proceedings brought by a Party while this Agreement is in force, and legal proceedings brought by a Party arising out of or in connection with this Agreement may only be brought in the courts of [JURISDICTION] at [JUDICIAL DISTRICT]. This clause shall only have effect if, for any reason, a dispute cannot be brought to arbitration pursuant to the preceding clauses (Human Cell Atlas, 2019).

4 Conclusion

In a rapidly evolving data-driven landscape, DTAs stand as foundational instruments governing the exchange of data across various sectors, from scientific research to commercial partnerships. This comprehensive analysis sheds light on the critical clauses that underpin DTAs, highlighting their importance in facilitating secure, efficient, and legally compliant data-sharing relationships and providing guidance on the drafting of such clauses. Drafting DTAs requires attention to detail, a nuanced understanding of data protection regulations and, often, legal expertise. DTAs are pivotal, not only for safeguarding data, but also for fostering collaboration, innovation, and responsible data sharing. With the guidance and insights provided in this article, stakeholders can navigate the complexities of data transfers, maximize legal certainty, and adhere to evolving data protection laws.

We should mention that the findings of the scoping review formed the basis for an open-source DTA template that was developed for the South African research community (Swales et al., 2023a; Swales et al., 2023b).

Author contributions

LS: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing–original draft, Writing–review and editing. AG: Methodology, Writing–review and editing. DT: Formal Analysis, Investigation, Methodology, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. Work on this article was supported by the U.S. National Institute of Mental Health and the U.S. National Institutes of Health (award number U01MH127690) under the Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) program. The content of this article is solely our responsibility and does not necessarily represent the official views of the U.S. National Institute of Mental Health or the U.S. National Institutes of Health.

Acknowledgments

The authors wish to thank their research assistants, Roasia Hazarilall, Kiara Munsamy, Jodie de Klerk, and Naseeba Sadak. OpenAI's ChatGPT-4 was used to improve language and readability.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2024.1332700/full#supplementary-material

References

B3 Africa (2018). Data transfer agreement. Available at: https://www.researchgate.net/publication/357269852_DATA-TRANSFER-AGREEMENT (Accessed September 12, 2022).

Bristol Myers Squibb (2017). Data sharing agreement. Available at: https://www.bms.com/assets/bms/us/english-documents/independent-research/sample-data-sharing-agreement.docx (Accessed September 12, 2022).

Clinical Study Data Request Consortium (2015). Data sharing agreement. Available at: https://www.clinicalstudydatarequest.com/Documents/DATA-SHARING-AGREEMENT.pdf (Accessed September 12, 2022).

Department of Health Western Australia (2021). Material and data transfer agreement. Available at: https://rgs.health.wa.gov.au/Documents/Material%20and%20Data%20Transfer%20Agreement%20v1%2012%20Nov%202021.docx (Accessed September 12, 2022).

Dkfz German Cancer Research Center (2020). Data transfer agreement of human data for research purposes. Available at: https://www.dkfz.de/en/CanEpi/EGA/Data-Transfer-Agreement-of-Human-Data-for-Research-Purposes-Jan.-2020.pdf (Accessed October 3, 2023).

FDP (2017). Data transfer and use agreement. Available at: https://www.uthscsa.edu/sites/default/files/Services/forms/duatemplate.pdf (Accessed September 12, 2022).

Fred Hutch (2020). Multi-party data and material transfer agreement. Available at: https://research.fredhutch.org/content/dam/stripe/specimen-acquisition-network/COL210703_Umbrella%20Data%20and%20Material%20Transfer%20Agreement_FH_FINAL%20no%20org%20list.pdf (Accessed September 12, 2022).

GREGoR Consortium (2022). Data sharing agreement. Available at: https://gregorconsortium.org/sites/default/files/policies/GREGoR%20Consortium%20Data%20Sharing%20Agreement%20V1.0_Fillable.pdf (Accessed October 5, 2023).

Growing Up in New Zealand (2014). Data access agreement. Available at: https://cdn.auckland.ac.nz/assets/growingup/access-to-data/External%20DA%20Agreement_1003a_V2.pdf (Accessed September 12, 2022).

Health Data Coalition. (2017). *Data sharing agreement*.

Human Cell Atlas (2019). Template: material/data transfer agreement (MDTA). Available at: https://docs.google.com/document/d/1MpUXC7ayY7RAYeZrEMbEdoYn2l-rIBDR/edit (Accessed October 3, 2023).

Indian Society of Critical Care Medicine (2024). Data transfer agreement. Available at: https://isccm.org/isccm_v1/pdf/data_transfer_agreement.pdf (Accessed March 5, 2024).

Infectious Diseases Data Observatory (2021). COVID-19 data platform: data transfer agreement. Available at: https://www.iddo.org/sites/default/files/publication/2021-03/COVID-19%20Data%20Transfer%20Agreement%2016MAR21.pdf (Accessed September 12, 2022).

Information Commissioner's Office (2022). International data transfer agreement. Available at: https://ico.org.uk/media/for-organisations/documents/4019538/international-data-transfer-agreement.pdf (Accessed October 3, 2023).

Johns Hopkins University (2022). Data transfer and use agreement. Available at: https://jhura.jhu.edu/wp-content/uploads/2022/03/DUATemplate_Updated.pdf (Accessed October 5, 2023).

Kawartha Lakes OHT (2020). Data sharing agreement. Available at: https://pub-kawarthalakes.escribemeetings.com/filestream.ashx?DocumentId=37434 (Accessed September 12, 2022).

KEMRI Wellcome Trust Research Programme (2019). Application form for access to KWTRP data. Available at: https://kemri-wellcome.org/zp-content/uploads/2021/02/KWTRP_Dataverse_Data_Request_Form_2019.docx (Accessed October 3, 2023).

Law Insider (2024). Term and termination sample clauses. Available at: https://www.lawinsider.com/clause/term-and-termination (Accessed October 2, 2023).

Murray, J. (2018). What is a contract: preamble/recitals - let's begin at the beginning. *AMM*. Available at: https://www.ammlaw.com/blog/what-is-a-contract-preamble/-recitals-lets-begin-at-the-beginning.html (Accessed October 2, 2023).

National Center for Advancing Translational Sciences (2021). Data transfer agreement. Available at: https://ncats.nih.gov/files/NCATS_Data_Transfer_Agreement_508.pdf (Accessed September 12, 2022).

National Health Service (NHS) England (2018). Template data sharing agreement. Available at: https://www.england.nhs.uk/wp-content/uploads/2019/04/appendix-4-sample-data-sharing-agreement.docx (Accessed October 3, 2023).

National Institute for Medical Research (2020). Data transfer agreement for researchers/organizations. Available at: http://reims.nimr.or.tz:8010/guides/DTA.pdf (Accessed September 12, 2022).

NetDocuments (2024). Termination for convenience clause samples. Available at: https://afterpattern.com/clauses/termination-for-convenience#:~:text=TPI%20COMPOSITES%2C%20INC)-,Termination%20for%20Convenience.,date%20stated%20in%20the%20notice (Accessed October 2, 2023).

ONDC (2024). Data sharing agreement. Available at: https://www.datacommissioner.gov.au/sites/default/files/2022-07/ONDC_Legislation_Agnostic_DSA_Template.doc (Accessed March 5, 2024).

Parkins, D. (2017). The world's most valuable resource is no longer oil, but data. *Econ.* Available at: https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data (Accessed October 2, 2023).

Protection of Personal Information Act 4 of 2013, 2013 (POPIA) [South Africa].

Swales, L. (2022). The Protection of Personal Information Act 4 of 2013 in the context of health research: enabler of privacy rights or roadblock? *PER/PELJ* 25, 1–32. doi:10.17159/1727-3781/2022/v25i0a11180

Swales, L., Botes, M., Donnelly, D., and Thaldar, D. (2023a). Towards a data transfer agreement for the South African research community: the empowerment approach. *S. Afr. J. Bioeth. Law*. 16 (1), 13–18. doi:10.7196/SAJBL.2023.v16i1.827

Swales, L., Ogendi, P., Botes, M., Townsend, B., Donnelly, D. L., Abdulrauf, L., et al. (2023b). *A data transfer agreement template for South Africa*. doi:10.5281/zenodo.7537396

Swiss Personalised Health Network (2021). Data transfer and use agreement. Available at: https://sphn.ch/wp-content/uploads/2021/11/2018DEV21_DTUA_SPHN_SwissPKcdw_V3.0.pdf (Accessed October 3, 2023).

Thaldar, D. (2024). The wisdom of claiming ownership of human genomic data: a cautionary tale for research institutions. *Dev. World Bioeth.*, 1–8. doi:10.1111/dewb.12443

Thaldar, D. W., Townsend, B. A., Donnelly, D. L., Botes, M., Gooden, A., van Harmelen, J., et al. (2022). The multidimensional legal nature of personal genomic sequence data: a South African perspective. *Front. Genet.* 13, 997595–997611. doi:10.3389/fgene.2022.997595

Toomey Motors v Chevrolet (2017) EWHC 276 (comm) [United Kingdom].

University of Newcastle (2024). Data transfer agreement. Available at: https://www.newcastle.edu.au/__data/assets/pdf_file/0010/631954/Data_Transfer_Agreement_OREBA1.pdf (Accessed March 5, 2024).

Utrecht University (2024). Data transfer agreement. Available at: https://www.uu.nl/sites/default/files/data_transfer_agreement_youth_template.pdf (Accessed March 5, 2024).

Xiong, B., Ge, J., and Chen, L. (2023). Unpacking data: China's "bundle of rights" approach to the commercialization of data. *Int. Data Priv. Law* 13, 93–106. doi:10.1093/idpl/ipad003

Frontiers in Pharmacology

Explores the interactions between chemicals and living beings

The most cited journal in its field, which advances access to pharmacological discoveries to prevent and treat human disease.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact