



OPEN ACCESS

EDITED BY

Sandra MacFadyen,
Stellenbosch University,
South Africa

REVIEWED BY

Maria M. Romeiras,
University of Lisbon,
Portugal

Glenn R. Moncrieff,
South African Environmental Observation
Network (SAEON),
South Africa

*CORRESPONDENCE

Brenda Daly
✉ B.Daly@sanbi.org.za

SPECIALTY SECTION

This article was submitted to
Environmental Informatics and Remote
Sensing,
a section of the journal
Frontiers in Ecology and Evolution

RECEIVED 15 December 2022

ACCEPTED 29 March 2023

PUBLISHED 26 April 2023

CITATION

Daly B and Ranwashe F (2023) South Africa's
initiative toward an integrated biodiversity data
portal.
Front. Ecol. Evol. 11:1124928.
doi: 10.3389/fevo.2023.1124928

COPYRIGHT

© 2023 Daly and Ranwashe. This is an open-
access article distributed under the terms of
the [Creative Commons Attribution License
\(CC BY\)](#). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

South Africa's initiative toward an integrated biodiversity data portal

Brenda Daly* and Fhatani Ranwashe

Biodiversity Information and Planning Directorate, Division of Biodiversity Information and Policy Advice, Kirstenbosch Research Centre, South African National Biodiversity Institute, Cape Town, South Africa

Researchers and policymakers have called on the South African National Biodiversity Institute (SANBI), in its role as the statutory biodiversity organisation of South Africa, to develop a coordinated and integrated biodiversity informatics hub. While biodiversity information is increasingly available from several providers, there is no platform through which to access comprehensive biodiversity information from a single source. In response, SANBI is redeveloping the Biodiversity Advisor platform, which will integrate geospatial, species and ecosystem data, literature and other data made available by a wide variety of data partners. To do so it has adopted a Service Orientated Architecture, whereby existing, independent biodiversity datasets are integrated. Consolidating such an extensive and varied set of databases, however, introduces some significant operational challenges. Solutions had to be found to address limited infrastructure, the complexity of the system, the lack of taxonomic identifiers, as well as the need for access and attribution. Solutions had to be pragmatic, given limited financial resources and limited capacity for information technology. The emerging outcome is a system that will easily allow users to access most biodiversity data within South Africa from a single, recognised platform.

KEYWORDS

infrastructure, services, system, attribution, biodiversity data integration, information resources

1. Introduction

Currently, there is a global impetus toward an interconnected network to link other sources of biodiversity and environmental data and in this way provide interdisciplinary information (Hardisty et al., 2022). Hardisty et al. (2022) use specimens as a digital anchor connecting other discipline-specific data. A recent example is modelling, understanding, and preventing potential pandemics, following COVID-19 (Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services [IPBES], 2020). Conservation efforts such as modelling invasive species, food and water security, and restoration are among the major forces behind data-driven prioritisation in many countries and organisations. New opportunities have become possible with the availability of big datasets and the advances in artificial intelligence technologies and their use in different fields of study (e.g., image and text recognition, systematic conservation planning) is an upcoming innovation (Silvestro et al., 2022).

With the increase in global and local online platforms that offer biodiversity data, such as iNaturalist, Global Biodiversity Information Facility (GBIF), Plants of Southern Africa and Biodiversity Geographic Information System (BGIS), access to biodiversity data has become easier (MacFadyen et al., 2022). However, it is still difficult to obtain all this biodiversity information from just one source. The South African National Biodiversity Institute (SANBI)

was maintaining disparate information systems that required a broad range of skills to support and management costs were escalating (Daly et al., 2013). It was at this point that SANBI started working to recreate the Biodiversity Advisor (URL)¹, an interoperable biodiversity data portal, that will provide comprehensive biodiversity information to a wide range of users. Users will have access to geospatial data, plant and animal species distribution data, ecosystem-level data, literature, images, and metadata. The newly developed system promotes a shift from tactical information systems, which deliver products and services for individual projects, to a strategic system that builds capacity within organisations and networks.

The overarching goal of the new system is to integrate available biodiversity data by unifying information resources across SANBI and its data partners, to improve quality and use, and thereby transform data into knowledge. Joining these data infrastructures will give researchers a collective overview that better facilitates answering research questions and will provide policymakers with the necessary information to make more informed decisions. This paper describes how biodiversity information sources, systems, and services in South Africa are being integrated into a national information system, as part of a project called the National Biodiversity Information System (NBIS). The range of different data platforms that are being brought together presents significant operational challenges that have required expedient and resource-efficient solutions.

SANBI recognised the success of similar international initiatives and to avoid reinventing existing solutions, a comparison of eight national research infrastructures was completed during the scoping phase of the NBIS project. These included the Atlas of Living Australia (ALA), SiB Colombia, National Biodiversity Data Centre (Biodiversity Ireland), National Biodiversity Network Atlas (NBN Atlas), LifeWatch Marine Virtual Research Environment, Conabio, Zoo Universe and Catchments. Most of the systems investigated were bespoke solutions. The ALA, NBN Atlas and Biodiversity Ireland showed the greatest fit (20%) with SANBI's requirements in developing its biodiversity informatics infrastructure. The criteria that made these systems more similar to what was required were how adaptable these systems would be to the unique existing source repositories at SANBI, the data types made available online and the requirements identified. The ALA architectural model consisted of numerous modular tools and software suites (e.g., Sensitive Data Service, Image Service, BioLink, etc.) linked together via a micro-services architecture (Chapman et al., 2016). Several of these modules were later made available to other organisations to use as open-source software reusable modules. The existing international systems would therefore be used as exemplars for the South African system, with the necessary deviations to account for the unique local context.

Biodiversity information management is not just about creating new methods or tools, it is about the coordination of stakeholders (e.g., data partners, communities of practise, etc.), standards, digitisation processes, integration, processing and using data effectively to support decisions. Biodiversity data and its processed products such as the Red List of Ecosystems and Species, routinely inform spatial planning, environmental authorisation, and protected area expansion through established channels (Botts et al., 2019, 2020).

When compared to other countries, South Africa is ahead of many others in the global context because it covers the whole spectrum.

This perspective article is targeted toward institutions that are starting on the journey of developing biodiversity informatics infrastructure. It highlights aspects of NBIS technical design that are particularly challenging and solutions that have proven successful.

2. The biodiversity information architecture

2.1. Strategy and building blocks

The South African National Biodiversity Institute (SANBI) is a statutory organisation established under the (National Environmental Management: Biodiversity Act, No.10 of 2004, 2004). South Africa is one of the few countries in the world to have a statutory entity with a dedicated biodiversity focus. In fulfilment of its mandate, SANBI leads and coordinates research, monitors and reports on the state of biodiversity in South Africa, gives planning and policy advice, engages in ecosystem restoration, and has a variety of managed collections of preserved and living specimens, seed banks, biological samples (BioBank), literature and library records. SANBI has responded to identified needs over time and developed a range of systems, tools, and policies, however, the value of these resources has been undermined as they are not integrated.

Despite its extensive data and information holdings, SANBI is not the only biodiversity organisation in the country that collects and serves biodiversity data. SANBI recognises that it does not have the capacity to achieve its mandate single-handedly and has adopted a Network of Partners Model where partners, through formal agreements, can contribute toward delivering on the SANBI mandate. Partnerships are not established with individual consultants or organisations working purely for profit and there is a set of criteria that each institution must meet. There are legal and non-legal mechanisms for implementation, for example, data sharing agreements, collaboration agreements, secondments, etc. (South African National Biodiversity Institute [SANBI], 2017). For example, the South African Environmental Observation Network (SAEON) collects long-term environmental observation data in South Africa (such as weather, soil moisture and temperature, etc.), so SAEON is a data partner.

Due to the complexity of the source repositories and the significant investments in developing large biological information resources, a more streamlined technical and operational model was needed to integrate all information resources. The challenge was to combine the existing information environment despite limited financial resources and in-house information technology expertise within SANBI. Consequently, decisions made during the development of the Biodiversity Advisor sought pragmatic but innovative ways to achieve more in a resource-constrained setting.

The NBIS project, therefore, began with the existing set of established information resources that were largely independent. It made no sense to go to the significant effort of migrating data into the available open-source ALA software suites when these functional components already existed in the existing infrastructure. Instead, to accomplish the data synthesis required, a Service Oriented Architecture (SOA) was implemented, which is a style of software design that integrates distributed, separately deployed and maintained

¹ <https://biodiversityadvisor.sanbi.org>

software components that may be controlled by various owners (Reference Architecture Foundation for Service Oriented Architecture Version 1.0, 2012). The basic tenet of SOA is that it is independent of vendors, products, and technologies.

The benefits of following an SOA architecture are the ability to assemble services (functionality and data) that leverage existing investments. Another benefit is that, although software and application upgrades are required to ensure compliance, there has been minimal impact within the source datasets and information resources landscape, which has meant that users continue to use the existing software and applications. Independent data storage has meant that each application (authoring layer) or service is independently changeable and deployable and can use a different technology stack. The web application, however, will need to be modified to accommodate this change.

2.2. User needs analysis

A survey, as part of a thesis project, was completed to understand who the user community is and what their needs are. These findings are currently being built into the Biodiversity Advisor and will help inform the development of products and services through a clear understanding of user needs. Using the initial needs analysis, the following user-level functionality was highlighted (Daly, 2020):

- The ability to aggregate information from other relevant fields of study (social, political, and economic) for more informed decision-making.
- Presentation of useful (solve a problem or decision) case studies.
- Tailored information views (consider the viewpoint of the information seeker).
- Include intuitive navigation as users are often unfamiliar with the content of the website.
- An advisory section on emerging science and policy topics.
- Focus on the information most in demand (distribution, ecological and threatened species data).
- Provide sources of environmental change information.
- Crowd-source data deficient species.
- Increase accessibility to peer-reviewed research outputs.

2.3. The service orientated architecture model

The basic structural elements of an SOA model are: (1) the underlying source datasets accompanied by their independent authoring layers, (2) an index and services layer that catalogues the information in the source datasets and acts as a bridge between them, (3) the front-end website and app that users will interact with, and (4) a search engine option that offers the ability to navigate the information (Figure 1). The authoring layer is the ready-made, often commercial application that supports business activities.

The source datasets that will be integrated within the Biodiversity Advisor system hosted by SANBI include:

- BODATSA – Botanical Database of Southern Africa, official plant names and descriptions (taxonomic backbone), specimen, living and seed collection, medicinal plant data, national vegetation database, and invasive species data.
- ZODATSA – Zoological Database of Southern Africa, official animal names and descriptions.
- Institutional repository – document repository to store all SANBI historical collections, library services and publications.
- Invasive Species Management System – tracks invasive species locations, abundance, and control efforts.
- Ecosystem database – ecosystem type, description, threat status, protection level assessment, distribution, and extent.
- BGIS – Biodiversity Geographic Information System (BGIS), a stakeholder website hosting products of various biodiversity plans (conservation plans) and other related initiatives.
- Metadata portal.
- IPT – Integrated Publishing Toolkit, publish biodiversity datasets from data partners.
- SEIS – SANBI Enterprise Image System, specimen and other digital images.

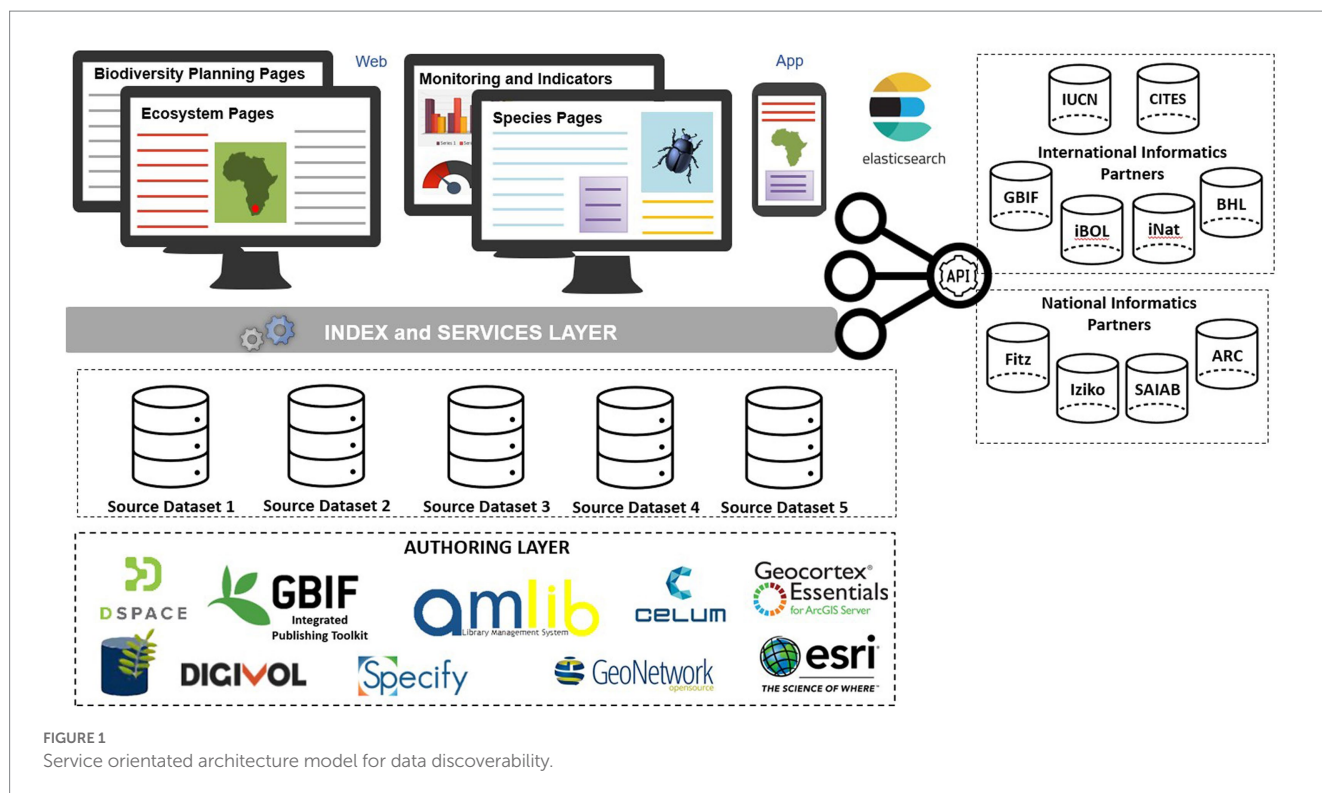
National Informatics Partners datasets:

- Fitz – FitzPatrick Institute of African Ornithology, hosts several biological resources.
- Iziko – Iziko South African Museum, museum specimens.
- SAIAB – South African Institute for Aquatic Biodiversity, fish specimens.
- ARC – Agricultural Research Council, conduct research in the agricultural sector.

International Informatics Partners datasets:

- GBIF – Global Biodiversity Information Facility, a global aggregator of species occurrence records.
- iBOL/Genbank – International Barcode of Life and Genbank, annotated collection of available DNA sequences.
- iNat – iNaturalist, Citizen Scientists can capture and upload sightings.
- BHL – Biodiversity Heritage Library, biodiversity literature.
- IUCN – International Union for Conservation of Nature, develops and promotes international standards for evaluating the conservation status of plant and animal species.
- CITES – Convention on International Trade in Endangered Species of Wild Fauna and Flora, a global agreement that ensures international trade does not threaten species' survival in the wild.

The transition to the SOA centred on the creation of an indexing system, which is a highly ordered set of lists of frequently searched data, coupled with the “ElasticSearch” search engine. The index and search engine are what allow calls to be made to the respective systems for data. In instances where data partners do not have application programming interfaces (APIs), data will be moved into the “index and services layer” with an extract-transform-load (ETL) process. An ETL is where data is extracted, transformed, and loaded into an output data container.



3. Operational challenges and solutions

3.1. Lack of infrastructure and tools when integrating data

The greatest challenge at the start of the project was the limited infrastructure concerning the resources assigned at the server level so the server specifications, poor network bandwidth, and separating servers to reduce response times by spreading the computational load. This task's resource considerations were underestimated, meaning it took up to 4 weeks to index the various source datasets. These technological challenges were overcome by procuring additional infrastructure and scaling to meet demands. Going forward, once the source datasets are indexed, incremental indexing can be used when changes are made to the source datasets. This will bypass the need for a complete resource-intensive reindex. Resulting updates will be run separately and published to the live portal once complete.

An API is a data interchange tool that allows applications to communicate and is most often developed and deployed by the vendor. An API means access to the data without having to understand all the system detail such as the database schema, functionality, etc. In some cases, web-based APIs were developed for the data sources, however, many software applications do not have a stable API, which meant the data was indexed directly from the backend database. The disadvantage of this solution is if any major changes are made to the software database it means changes need to be made to the platform. Therefore, it is essential to consider how future versions or changes in application products will fit with the current architecture and account for the time and resources to maintain the system. APIs can often also

be a constraint in a project as they only unlock certain data depending on the intended use case.

Many national informatics partners do not have the necessary resources to manage and provide data. System maintenance is also resource intensive and the sustainability of these projects is a risk. The lack of suitable mechanisms and infrastructure is often a barrier to data partners publishing their data. To overcome this challenge SANBI has offered support with data management and set up an Integrated Publishing Toolkit (IPT) instances used to publish biodiversity datasets (Robertson et al., 2014). This is working toward recommendations made by Costello et al. (2014) on strategies for the sustainability of datasets being the integration of datasets into a collaborative information system such as the IPT, within an institute with a suitable mandate.

3.2. Complexity of the system

With different data platforms being brought together various data categories are integrated under one architecture and the challenge is integrating these heterogeneous data. Source repositories have data categories ranging from geographic or spatial data (BGIS), key biodiversity areas data, occurrence records (BODATSA, SANBI IPT), ecosystem data (Ecosystem Descriptions Database), checklists (accepted and synonym names; protologue citations; type information; classification), distribution and residency status (BODATSA, ZODATSA), specimen data (BODATSA), descriptive information (BODATSA, ZODATSA), literature (SANBI Institutional Repository, SANBI library catalogue), metadata (SANBI Metadata), genetics (BioBank), images (SANBI Enterprise Image System (SEIS)), and threatened species and ecosystem data (Red List Assessment Systems), National Biodiversity Assessment (NBA) data to taxonomic descriptions (BODATSA and ZODATSA) and indicator data

(Figure 2). Interlinking information systems in interoperable ways in a consistent manner is challenging and much time needs to be allocated to tailoring data structures and query processing.

Another challenge faced by organisations is related to people or skills, the first implication of this skills shortage is that organisations often need service providers to fill this resource gap. It is vital with outsourcing to select the right service provider that understands the core business and time to be able to develop specifications for the project correctly.

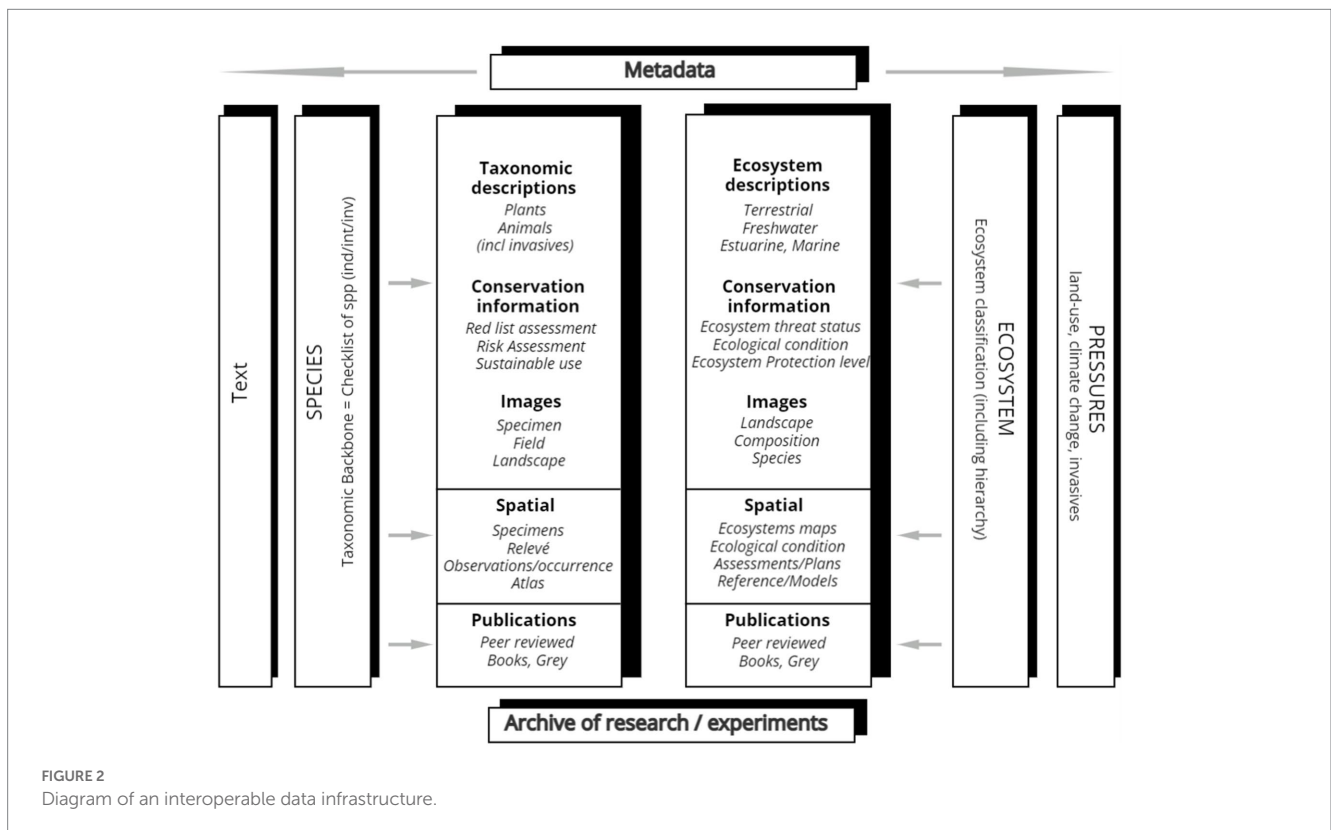
3.3. Taxonomic service backbone

SANBI is mandated to maintain and provide an up-to-date South African National Plant and Animal Checklist with accurate taxonomic information. This is achieved by publishing a consolidated national checklist of plants and animal species in South Africa yearly, with updates happening throughout the year as taxonomic changes are made available in the literature. Updating the checklist involves monitoring published literature. South Africa has over 67,000 described species of animals (Skowno et al., 2019), 21,467 species of plants (Klopper and Winter, 2022), and 1,422 alien plant and animal species (van Wilgen et al., 2020). As with any biodiversity data integration process, species names are often used as the common identifier, however, the limitations are that they are not unique or stable (Page, 2008). An enabling feature is assigning persistent identifiers (long-lasting references consisting of letters or numbers to a digital resource often machine generated) that will support linkages between data sources and allow for global compatibility. Historically, due to a lack of persistent identifiers, users have used primary keys

from the authoring layer which has resulted in obsolete numbers being migrated back into the system, as systems and software are updated. The lack of a common identifier has created a barrier to making data publicly available, accentuating shortcomings in available data, generating, or integrating any other types of data and ultimately the conservation and management of the species (Ely et al., 2017). Therefore, linking Globally Unique IDentifiers (GUIDs) in an authoritative taxonomic resource when integrating biodiversity data and using these GUIDs to link other source datasets is imperative (Guralnick et al., 2015).

3.4. Anonymous usage

SANBI's mandate is clear that as a public organisation, the biodiversity information it provides must be openly accessible (South African National Biodiversity Institute [SANBI], 2010a). The management of data is covered by the Biodiversity Information Policy Framework (South African National Biodiversity Institute [SANBI], 2010a), the Intellectual Property (IP) policy (South African National Biodiversity Institute [SANBI], 2010b) and the Protection of Sensitive Taxa policy (South African National Biodiversity Institute [SANBI], 2010c). Therefore, it is essential that the Biodiversity Advisor is available for anonymous usage and that data is free to download. However, to better understand users, manage the system and determine if the information being provided is having an impact, it is also necessary to monitor use. Several types of activity, such as the download of spatial data, must take the user through a confirmation process that makes it clear who owns the data and the terms of use. In addition, sensitive data, such as the location of species that are



vulnerable to collecting or over-exploitation, requires a process of access request and approval. Some data owned by partners and made available as part of data sharing agreements includes embargoes, redactions, or restrictions. Data sharing agreements within the framework of attributions ensure the data is suitably shared.

In response, a secure authentication mechanism with no restriction to register is being implemented. The authentication mechanism helps to manage the data attribution and can be used to manage access to some projects and functionality (such as authorising access to documents or download links). It also helps to analyse usage patterns and capture business intelligence data when required. A role-based security module was developed, known as the Biodiversity Passport, where functionality and datasets are available only to users who have been authenticated and authorised, such as allowing authorised officials of government conservation agencies to access content for conservation management purposes.

Access to literature is often as essential as raw data, however, copyrights and paywalls often stifle necessary access to information in the conservation of species and ecosystems. There is a push for open access when linking literature associated with biodiversity data. The solution here is to provide Uniform Resource Identifiers (URI) to the physical resource or a Uniform Resource Locator (URL) to the payment gateway to ensure the researcher's work is recognised. SANBI's Institutional Repository allows users to request a copy from the SANBI author as an alternative to buying the paper.

3.5. Ensuring attribution

Data citation and attribution seem to be a consistent struggle in building biodiversity informatics infrastructure (Reichman et al., 2011; Patterson et al., 2014). Attribution is defined as assigning appropriate credit for an organisation or individual's contribution (Haak, 2014; Franz and Sterner, 2018) perspective is that it holds authors accountable for data accuracy and potential criticism. Ensuring a consistent user experience across multiple channels and still preserving attribution is an underrated challenge. The issue then confounds when to increase data accessibility, users can reuse subsets of data by downloading .csv files linked to the data behind any user interface. In this case, any data record downloaded needs to include the contributor's name.

Metadata is considered a form of attribution. However, the challenge here again is ensuring the metadata is accurate and up to date as often the metadata-related changes are not documented. Metadata describes the origin and tracks dataset changes (Biodiversity Conservation Information System, 2000). It is essential to ensure that every record includes the author's name on the website and within any downloaded data. Another solution is an acknowledgements page or listing contributors or editors of data. A solution used by Figshare is to cite datasets using formatted references (Haak, 2014).

4. Conclusion

In the past, the numerous disaggregated and disparate systems, tools, and policies made it difficult to leverage biodiversity information to support research, policy development and decision-making. The Biodiversity Advisor is a service-orientated data management system, built largely from contributing systems. These data platforms are being

brought together to ensure the information resource is more adequate for national planning and management. By addressing limited informatics infrastructure, obtaining necessary human resources and skills, and establishing a system of unique identifiers, the complexity of the system can be overcome. Developing authentication systems and mechanisms for assigning credit can navigate the balance required between accessibility and attribution. The reimagined Biodiversity Advisor is thus a milestone in establishing a fully integrated data information system for South Africa.

The Biodiversity Advisor² is scheduled to launch in 2023. At the time of launch, the BODATSA (plant) and ZODATSA (animal) will be fully indexed, providing comprehensive species pages, including specimen collection records, iNaturalist observational records and occurrence records from various data partners. The SANBI library catalogue and institutional repository will also be indexed providing numerous literature resources. Systematic integration of the remainder of the systems and data will follow as datasets become available.

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

BD took the lead in writing the manuscript. FR compiled the presentation abstract and gave the presentation. All authors contributed to the article and approved the submitted version.

Acknowledgments

Thank you to Emily Botts for her helpful comments on this manuscript. Thank you to SANBI and Data Partners for their participation in the NBIS project.

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

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.

² <https://biodiversityadvisor.sanbi.org>

References

- Biodiversity Conservation Information System (2000). Framework for information sharing: Principles. Handbook Series. 1–8.
- Botts, E. A., Pence, G., Holness, S., Sink, K., Skowno, A., Driver, A., et al. (2019). Practical actions for applied systematic conservation planning. *Conserv. Biol.* 33, 1235–1246. doi: 10.1111/cobi.13321
- Botts, E. A., Skowno, A., Driver, A., Holness, S., Maze, K., Smith, T., et al. (2020). More than just a (red) list: Over a decade of using South Africa's threatened ecosystems in policy and practice. *Biol. Conserv.* 246:108559. doi: 10.1016/j.biocon.2020.108559
- Chapman, A. D., Berzin, L., Smith, R., and Tann, J. (2016). Atlas of living Australia infrastructure implementation. Atlas of living Australia (Issue October). Available at: <https://www.ala.org.au/wp-content/uploads/2017/01/ALA-Infrastructure-Implementation-overview-October-2016-final.pdf>
- Costello, M. J., Appeltans, W., Bailly, N., Berendsohn, W. G., de Jong, Y., Edwards, M., et al. (2014). Strategies for the sustainability of online open-access biodiversity databases. *Biol. Conserv.* 173, 155–165. doi: 10.1016/j.biocon.2013.07.042
- Daly, B. (2020). Building biodiversity data infrastructure for science and decision-making: Information needs and information-seeking patterns in South Africa. Faculty of Humanities, Department of Knowledge and Information Stewardship. Available at: <http://hdl.handle.net/11427/32635>
- Daly, B., Hathorn, P., Roberts, R., and Willoughby, S. (Eds.) (2013). Proceedings of an Information Architecture Workshop. South African National Biodiversity Institute.
- Ely, C. V., De Loreto Bordinon, S. A., Trevisan, R., and Boldrini, I. I. (2017). Implications of poor taxonomy in conservation. *J. Nat. Conserv.* 36, 10–13. doi: 10.1016/j.jnc.2017.01.003
- Franz, N. M., and Sterner, B. W. (2018). To increase trust, change the social design behind aggregated biodiversity data. *Database* 2018:bax100. doi: 10.1093/database/bax100
- Guralnick, R. P., Cellinese, N., Deck, J., Pyle, R. L., Kunze, J., Penev, L., et al. (2015). Community next steps for making globally unique identifiers work for biocollections data. *ZooKeys* 494, 133–154. doi: 10.3897/zookeys.494.9352
- Haak, L. L. (2014). Persistent identifiers can improve provenance and attribution and encourage sharing of research results. *Inf. Serv. Use* 34, 93–96. doi: 10.3233/ISU-140736
- Hardisty, A. R., Ellwood, E. R., Nelson, G., Zimkus, B., Buschbom, J., Addink, W., et al. (2022). Digital extended specimens: Enabling an extensible network of biodiversity data records as integrated digital objects on the internet. *Bioscience* 72, 978–987. doi: 10.1093/biosci/biac060
- Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services [IPBES] (2020). “Workshop Report on Biodiversity and Pandemics of the Intergovernmental Platform on Biodiversity and Ecosystem Services” in . eds. P. Daszak, J. Amuasi, C. G. Das Neves, D. Hayman, T. Kuiken and B. Rocheet al. (Bonn, Germany: IPBES Secretariat)
- Klopper, R.R., and Winter, P.J.D. (2022). *South African National Plant Checklist statistics, version 1.2022*. South African National Biodiversity Institute, Pretoria.
- MacFadyen, S., Allsopp, N., Altwegg, R., Archibald, S., Botha, J., Bradshaw, K., et al. (2022). Drowning in data, thirsty for information and starved for understanding: A biodiversity information hub for cooperative environmental monitoring in South Africa. *Biol. Conserv.* 274:109736. doi: 10.1016/j.biocon.2022.109736
- National Environmental Management: Biodiversity Act, No.10 of 2004 (2004). *Government Gazette*. 467(700). 7 June. Government notice no. 26436. Cape Town: Government Printer.
- Page, R. D. M. (2008). Biodiversity informatics: The challenge of linking data and the role of shared identifiers. *Brief. Bioinform.* 9, 345–354. doi: 10.1093/bib/bbn022
- Patterson, D. J., Egloff, W., Agosti, D., Eades, D., Franz, N., Hagedorn, G., et al. (2014). Scientific names of organisms: Attribution, rights, and licensing. *BMC. Res. Notes* 7:79. doi: 10.1186/1756-0500-7-79
- Reference Architecture Foundation for Service Oriented Architecture Version 1.0 (2012). OASIS committee specification 01. Available at: <http://docs.oasis-open.org/soa-rm/soa-ra/v1.0/cs01/soa-ra-v1.0-cs01.html>
- Reichman, O. J., Jones, M. B., and Schildhauer, M. P. (2011). Challenges and opportunities of open data in ecology. *Science* 331, 703–705. doi: 10.1126/science.1197962
- Robertson, T., Döring, M., Guralnick, R. P., Bloom, D., Wiczorek, J. R., Braak, K., et al. (2014). The GBIF integrated publishing toolkit: Facilitating the efficient publishing of biodiversity data on the internet. *PLoS One* 9:e102623. doi: 10.1371/journal.pone.0102623
- Silvestro, D., Gorla, S., Sterner, T., and Antonelli, A. (2022). Improving biodiversity protection through artificial intelligence. *Nat. Sustain.* 5, 415–424. doi: 10.1038/s41893-022-00851-6
- Skowno, A.L., Poole, C.J., Raimondo, D.C., Sink, K.J., Van Deventer, H., Van Niekerk, L., et al. (2019). National Biodiversity Assessment 2018: The status of South Africa's ecosystems and biodiversity. Synthesis Report. South African National Biodiversity Institute, an entity of the Department of Environment, Forestry and Fisheries, Pretoria. pp. 1–214.
- South African National Biodiversity Institute [SANBI] (2010a). Biodiversity information policy framework principles and guidelines. Available at: <http://hdl.handle.net/20.500.12143/7450>
- South African National Biodiversity Institute [SANBI] (2010b). Biodiversity information policy framework–intellectual property rights policy. Available at: <http://hdl.handle.net/20.500.12143/7449>
- South African National Biodiversity Institute [SANBI] (2010c). Biodiversity information policy framework-digital access to sensitive taxon data. Available at: <http://hdl.handle.net/20.500.12143/7087>
- South African National Biodiversity Institute [SANBI] (2017). *Network of partners policy*. Pretoria: South African National Biodiversity Institute.
- van Wilgen, B. W., Measey, J., Richardson, D. M., Wilson, J. R., and Zengeya, T. A. (2020). “Biological invasions in South Africa: An overview” in *Biological invasions in South Africa. Invading nature-springer series in invasion ecology*. eds. B. Wilgen, J. Measey, D. Richardson, J. Wilson and T. Zengeya (Cham: Springer).