frontiers in Blockchain

COMMUNITY CASE STUDY
published: 11 November 2021
doi: 10.3389/fbloc.2021.631640

3 Stages of a Pan-African Identity Framework for Establishing Self-Sovereign Identity With Blockchain

S. Solomon Darnell[1,2]* *and Joseph Sevilla*[2]

[1]Tint Right Colour Enterprise, Nairobi, Kenya, [2]@iLabAfrica, Strathmore University, Madaraka Estate, Kenya

The African continent (specifically its overwhelming in(animate) resources) is often referred to as the sleeping giant by magazines, blogs, research presentations and articles, and NGOs [such as World Bank]. Reasons for this moniker/title include the continent's plentiful natural resources, its large and quickly growing young population, and the young population's quick adoption and acclimatization to technology. Most countries on the continent are known as developing countries due to lack of access to safe drinking water, reliable electricity and roads, sanitation and hygiene, and a high number of people with tropical/infectious diseases. However, due to the usefulness of cellular phones and technology, several countries and companies within them have focused on cell phone proliferation (91% in Kenya). Smart phone usage allows Kenyans access to the world's information and potentially endless innovation. Given that a large number of Kenyans with smartphones use social media, coupled with the advent of Europe's GDPR (general data protection regulation), African identity and its associated data became an area of great interest. As the world is quickly progressing into a digital economy, a solution must be created that allows us to regain and control our identities, doing our best to ensure losing such is infinitely close to computationally and probabilistically impossible/improbable. Developing a blockchain-based identity backbone using biometrics and historical family information while allowing government-based identification documents is the best way forward. Three stages have been identified as necessities to accomplish the development of this system before opening it further beyond the pan-African worldwide community. The three stages are defined by systems that allow for biometric/demographic registration (stage 1), interoperability and security hardening (stage 2), and biometric modality data analysis/organization/association (stage 3).

Keywords: Africa, blockchain, biometrics, self-sovereign identity, pan-African, cancelable

OPEN ACCESS

Edited by:
Alan Sherriff, Consultant, London, United Kingdom

Reviewed by:
Larry C. Bates, AltMarket, United States
Richard Tighe, Oxfam, United Kingdom

***Correspondence:**
S. Solomon Darnell
sdarnell@strathmore.edu

Specialty section:
This article was submitted to Blockchain for Good, a section of the journal Frontiers in Blockchain

Received: *20 November 2020*
Accepted: *25 May 2021*
Published: *11 November 2021*

Citation:
Darnell SS and Sevilla J (2021) 3 Stages of a Pan-African Identity Framework for Establishing Self-Sovereign Identity With Blockchain. Front. Blockchain 4:631640. doi: 10.3389/fbloc.2021.631640

1 INTRODUCTION

For the last 6 years, identity in Africa has been put in the spotlight by several countries on the continent and organizations like World Bank, along with other NGOs (nongovernmental organizations). Sustainable development goals defined by the World Bank have helped lead to this focus (Bank-ID4D, 2017). Aside from external policy makers and institutions, Kenya has Vision

2030 which outlines world-class infrastructure facilities and services where "equality is entrenched, irrespective of one's race, ethnicity, religion, gender, or socio-economic status," and "nine governance principles shall be adhered to;" one of which is "decentralization" (Kenya, 2008). Decentralization is exceptionally important to Kenya as it is one of the nine governing pillars of Vision 2030. A companion idea supporting decentralization is "upgrading national ICT infrastructure," which includes the implementation of "public key infrastructure (PKI) to authorize and authenticate information systems in the country." Blockchain is a decentralized distributed computing platform that currently uses PKI to maintain security and privacy. PKI is not a technology unique to distributed ledger technology or blockchain but is used in several systems where privacy is of utmost importance, including distributed computing environments (Thompson et al., 2003) and many other areas including cards in Taiwan, electronic passport chips, certificates on USB keys, and many more (Wilson, 2005). Because PKI is a proven, often used, tried, and tested protocol whose security is based on the ownership and generation of a private key, it makes sense to use it with self-sovereign identity (SSI).

It makes sense for SSI in that for the scheme to work, the user must generate a key pair and only need to share the public key and never the private one. Since in PKI, the user generates the key pair, it seems to be a great component for a scheme referred to as "self-sovereign." Another companion idea, the one supported by this proposed framework, is "development of a national addressing system project to identify streets, buildings, plots, and other infrastructure and allocating them a street address" (Kenya, 2008). Currently, Kenyans in areas of low infrastructure can only describe where they live. Our system will allow for such a description to be added as demographic data, along with coordinates. This framework will be an aide to the street addressing system of Vision 2030 as global coordinates must correspond to physical addresses. This framework (containing a blockchain-based SSI) includes major features, such as "decentralization" and "PKI to authorize and authenticate information systems in the country," which are aligned with Kenya Vision 2030. This framework will serve as a model for African countries with existing citizen data infrastructures and for countries with limited identity systems.

In Naik and Jenkins (2020), the authors propose twenty governing principles of SSI, of which "sovereignty" is the first and refers to the creator of the identity having full control over the digital entity, in that no external person or organization has a say over management or usage. Centralization cannot, by definition, accomplish this goal as a central server managing the information for others can easily be manipulated. Distributed ledger technology (DLT) as in a distributed database that requires consensus voting to change a record is not good enough, specifically because a distributed ledger may or may not allow record deletion and modification. Blockchain is a better facilitator as it has the rule that data once written cannot be modified or deleted, allowing for a more assured trust. DLT and blockchain are technologies of the same family; however, as both technologies rely on a computational consensus mechanism, it becomes possible, in a general distributed ledger, for a record to be modified or deleted without proper intention, whereas the blockchain implementation of DLT does not allow data written to the ledger to be modified in any way once written. This speaks to the absolute necessity of a self-sovereign identity (SSI) system based on a decentralized, incorruptible ledger. As a pan-African self-sovereign identity framework, our proposal embodies the primary aspects of a foundational identity system.

2 MOTIVATION

Is there still a way to contribute to human digital infrastructure? As identity is one of the most fundamental and primary aspects of physical existence, is there an individually controlled trustworthy digital system that exists outside of governments and not completely controlled by an international conglomerate? How can we design, build, and set up such infrastructure to last beyond our generation and be created in such a way that it is not exploitative? Can we build an infrastructure that can be monetized but does not require people with the least resources to pay unless they desire it? Can we build digital infrastructure that can also be used by citizens in postcolonial countries who have so far been close to left out of the fourth industrial revolution? Can we design our addendum to the world's digital infrastructure that is different than what currently exists? Finally, can we build digital infrastructure that holds up in times of national and international tragedy, stress, and catastrophe?

The framework is meant to be paid for by governments, organizations, and companies while being free at the point of service for individual users. The development of the framework should be modular and easily updated while following the best software engineering development standards for testing, continuous integration, and deployment. A main purpose of the framework, to be free at the point of service, is designed to allow usage with minimal technological infrastructure and resource. Along with following the best software engineering development standards, continuous research will be carried out throughout development of the framework systems to ensure it solves or mitigates issues found with the existing systems. Decentralization, as a main tenet for the framework, will hopefully ensure the framework systems hold up in times of catastrophe.

2.1 Why Pan-African?

Within AI research, a common technique of calculating a "good enough" solution to an NP complete problem is to solve a similar problem of reduced complexity. In an attempt to create a robust self-sovereign identity system to satisfy all humans on the planet, it follows that attempting to create a robust identity system for the pan-African context (Du Bois, 1974) is a similarly challenging problem that when solved will be a "good enough" solution to the parent problem. Pan-Africa represents a segment of the population that is represented thoroughly throughout the

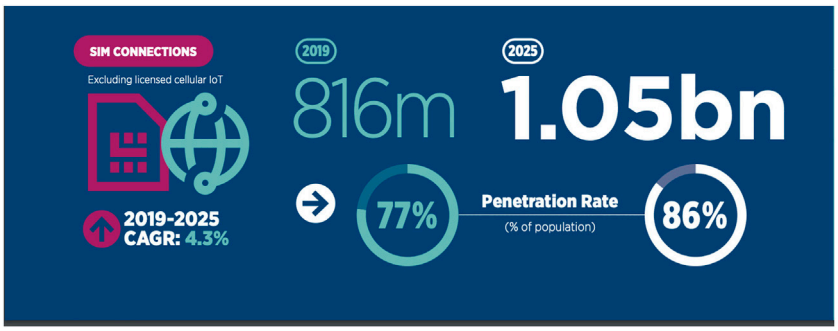

FIGURE 1 | Sub-Saharan SIM connections [GSMA Intelligence (Intelligence, 2020)].

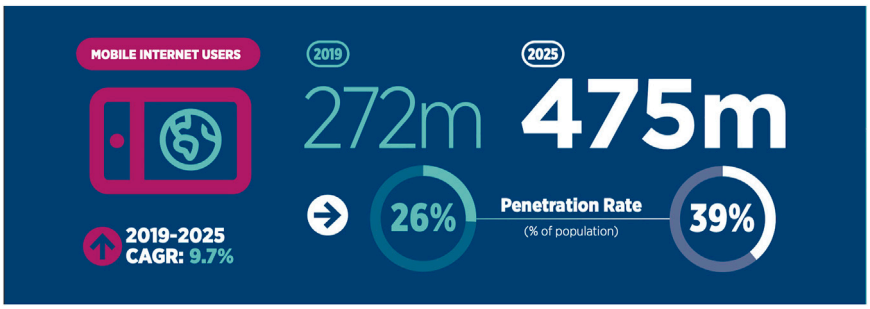

FIGURE 2 | Sub-Saharan mobile network users [GSMA Intelligence (10)].

world, often at the extremes of society. It seems when developing an identity system to serve everyone, we can design for the population that can approximate the full breadth and depth of humanity.

2.2 Young Mobile Population

Since the year 2000, the population of the African continent has nearly doubled, from around 815 million to 1.34 billion, based on figures from PopulationOf dot net (Africa population, 2020). With such a quickly growing population, it follows that the median age is not very high, at 24 years (Africa population, 2020). Mobile device usage is consistently growing on the continent, specifically in sub-Saharan Africa and is projected to continue (Intelligence, 2020). In Kenya alone, mobile phone proliferation surpassed 100% by the end of 2018 (Tanui, 2018).

Figure 1 shows the SIM connections in sub-Saharan Africa as a whole, which are at 816 million in 2019, and are projected to be just over 1 billion in 5 years.

Figure 2 shows that in 2019, approximately 26% of the population of sub-Saharan Africa is using mobile data.

Figure 3 shows that the mobile subscription rate is 45% of the sub-Saharan Africa's population.

With 24 years being the median age of the continent, the projections of SIM connections to grow by 9%, mobile data users to grow by 13%, and mobile subscribers to grow by 5% in sub-Saharan Africa, it shows us that the youth will be digital denizens. Creating a digital identity that will protect this population as it continues to grow is our aim. As this population is somewhat new to digital life, they lack an established mental paradigm for the concept of digital identity; this fact will possibly make adoption of self-sovereign identity paradigm and all it entails easier.

2.3 Contributing to Digital Infrastructure

We posit that one of the best ways to contribute to global digital infrastructure is to rebuild it, using decentralized system design (Henfridsson et al., 2013), from the World Wide Web technology level. However, such is a monumental undertaking and not the subject of the work at hand. Hence, on a small scale, as Kenya is bracing itself for the fourth industrial revolution (4IR) by the implementation of Kenya Vision 2030 (Kenya, 2008), the development of a cryptographically secure decentralized identity system can contribute positively to multiple areas, including ICT industry development, development and dissemination of digital content, creative industry development, and e-government systems.

Going the way of blockchain and DApps, we must evaluate the existing technologies in the area of interest. As digital identity is of interest, the different types of digital identity must be at least reviewed, so that we may put forth something we believe is an improvement on that which exists. Digital identity can be divided into three different

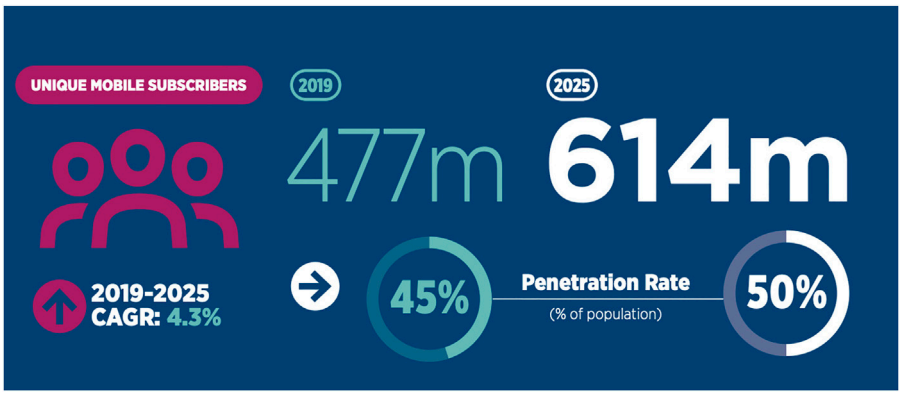

FIGURE 3 | Sub-Saharan mobile subscribers [GSMA Intelligence (Intelligence, 2020)].

categories: private provider (platform)–controlled, nation-controlled, and self-sovereign.

The private provider category has been in place since the establishment of the Internet. This category contains several types of private identity providers, some of which are dial-up provider identity, AOL identity, free Internet email (e.g., Hotmail, Yahoo, and Google), and membership-based sites (e.g., MySpace, Amazon, and Facebook). Social media sites are including other membership-based sites due to data usage protocols and purpose of identity management (Baars, 2016).

The second category under consideration includes digital identification initiatives by nation-states, some of the more significant initiatives include eCitizen (Kenya) (Ondego and Moturi, 2016), Aadhar (India) (Sen, 2019), WeChat (Plantin and de Seta, 2019), and Estonia Identity Suite (eID, Mobile, Smart, and Residency) (Id-card, 2019; Mobile-id, 2019; Smart-id, 2019; E-residency, 2019). WeChat could be placed in the first category as it is a membership-based identity for a social network; however, due to China's "markedly techno-nationalist media regulations and increasingly overt cyber-sovereignty agenda," it has gone from a private provider to a nationally controlled infrastructure service.

The final category is self-sovereign identity, which we posit is currently only possible by way of blockchain technology (van Wingerde, 2017). Because companies and governments require ownership of data they control, and hold on their servers, there is no way self-sovereign identity is possible through those entities. In fact, with blockchain, everyone can hold a copy of the ledger as everyone is a cooperator of the system. There is no sovereignty without supreme control of your data within a limited sphere, and that is impossible, by definition, if everything is controlled outside of the individual. Some of the people who need to be served by such a system do not have the resources necessary to maintain a full copy of such a ledger. Thankfully, due to the design of blockchain systems (Zheng et al., 2017; Gatteschi et al., 2018; Ul Hassan et al., 2020), at any time one does obtain such resources, one will be able to obtain the full ledger themselves and become a network node. The design of distributed ledger technology promotes inclusiveness (Allison et al., 2019) and security and hence is the only technology today that can realistically promise self-sovereign collective infrastructure for individuals in the digital world.

TABLE 1 | *Identity Service Comparison.*

Attribute	Private	Government	Self-Sovereign
Demographics	✓	✓	✓
Biometrics	▲	✓	▲
User owns data	⊖	⊖	✓
Share data profits	⊖	⊖	▲
Transparent data access	⊖	◆	✓
State system integration	▲	✓	▲
Transparent user audit	⊖	⊖	✓
National infrastructure	✓	✓	✓
Blockchain back-end	⊖	✓	✓
Data volunteering	⊖	⊖	▲

✓ = available, ⊖ = not available, and ▲ = partially available.

2.4 Pan-Africa Self-Sovereign Identity Qualifiers

Table 1 contains references as the entry to some of the table elements; in such a case, the reference denotes the possibility of that type of identity service having the attribute in question. Establishing a self-sovereign identity system with blockchain will need to have positive attenuation for every attribute listed in **Table 1** along with those outlined by Wingerde's master's thesis table 26 "Blockchain-enabled Self-sovereign Identity" (van Wingerde, 2017). Wingerde outlines a set of constraints in line with the General Data Protection Regulation (GDPR), the Revised Payment Services Directive (PSD2), and the electronic Identification, Authentication, and Trust Services regulation (eIDAS) (van Wingerde, 2017). Concerning being government infrastructure, like WeChat (Plantin and de Seta, 2019), the system should become so ubiquitous until it is necessary that government uses it as infrastructure.

2.5 Organization

The three stages of the framework will now be outlined by exposition of its registration processes, interoperability, and security, as well as its biometric-based longitudinal study.

3 STAGE 1: REGISTRATION

The first stage is the same for any identity system, and that is what and how information is stored in the system. What are the privacy tenets? How does one restore a lost or forgotten account? Can an individual register multiple accounts? If so, how are multiple accounts handled?

3.1 Demographics

The basic defining attributes of an individual form the bedrock of foundational identity. The framework must enable core attributes to be mapped to an identifier by which an individual is known. Demographic data is the first aspect of individual identification; this information is very important in designing supporting systems for essential services (e.g., financial inclusion/access, health-care access, and education). The context in which one finds oneself is a supplemental aspect of individual identification. Who you are varies depending on who asks. Your name may be Muthoni, but to your children, you are a parent, a resource to your employer, a student to your university, a taxpayer, and citizen to your government. Different contexts define who we are over our lifetime and how we identify ourselves. One may end up holding different forms of documentation to prove who they are to access and benefit from available services. Hence, functional identity is formed across myriad different contexts.

Some details vary on the different identifying documents, but some key details are constant. Common details include one's name, date of birth, gender, and image on an identifying document. One may hold a national ID card, a driver's license, a student's ID card, an employee card, a club membership card, and a health insurance card. Yet in reality, it is still the same person, regardless of interactions with differing authoritative bodies. In usage of any of the credentials, one only needs to show it and have its credibility checked before being granted access to a facility or services tied to the credential.

While many mundane tasks like money transfer have been successfully digitized, it has remained a hard task for the same to happen for exchange of identity credentials either due to poorly implemented standards or technology silos that hinder interoperability. Internet standards like the verifiable credentials spec and decentralized identifiers (DIDs) by W3C have evolved over time to support a standard version of credentials and credentials exchange when issuing and verifying claims held by an individual (A primer for decentralize, 2019). The digital identity revolution has been growing as seen in white papers published by the World Economic Forum highlighting the same (Nash, 2020; Community Paper, 2020).

3.2 Authentication

Authentication is an extraordinarily important component of every identity framework. Identity registration/verification has taken several forms, and one of the most often used today is multifactor authentication (Ometov et al., 2018). Multifactor authentication (MFA) refers to logging into a system using more than one verification step. A typical login is entering a user name and password on a page and getting access to personalized or private content. MFA uses various combinations of something you know (password), something you are (biometrics), and something you have/own (smartphone and pre-existing email) to perform more secure authentication (Ometov et al., 2018). The framework will initially use MFA, while the following section focuses on biometrics singularly. Biometrics is singularly focused upon because the technology is consistently being improved, and it is our vision that biometrics will be the only factors of authentication necessary at some point in the future.

3.3 Handling Biometric Data

Biometrics is the art and science of measuring life, and in computing practice, it uses sensors to record a physiological or behavioral marker to process and use for identifying and/or verifying individuals. Cancelable biometrics (Ratha et al., 2001) is a subfield created by Nalini Ratha, inspired by early one-time password (OTP) systems. Cancelable biometrics allows for a digital representation of one's biometric information to be transferred electronically without compromise. Changing one's physical biometrics permanently is unlikely; hence, we want a system that safeguards this information most stringently. Following that thought, unless special permission is given by the individual, the system will not require biometric templates to be sent directly for any operation. The system will utilize cancelable biometrics that builds a key representation from biometric information, similarly to a one-way hash (Merkle, 1989). When values/parameters that contribute to a cancelable functions output are compromised, biometric data are not. The aforementioned parameters to the function can be regenerated and updated with more attention to security. With normal biometric recognition upon registration, a template is generated. This template is stored in a biometric database to be used in the future for identification or verification. In such biometric systems, template theft is a common way to compromise the authentication process. Cancelable biometrics seeks to remedy this by never requiring a pure template of any biometric feature to be stored. With cancelable biometrics, at most a partial template is stored, and if the templates are compromised, it is part of the protocol to replace it with a different template. Because a true biometric template is never stored, template compromise does not compromise one's biometric signature (Ratha et al., 2006). Another issue present for cancelable biometric performance is biometric template degradation, or the fact that biometric features change with time (Fenker and Bowyer, 2012).

3.3.1 Where Biometrics Can Fall Short

Biometrics will not be initially used by itself as the framework should be open to people with only the most basic technological footprint, as in ownership of a feature phone. There are other issues with biometrics as in aging templates, chance of false positives (false accept rate—FAR), chance of false negatives (false reject rate—FRR), biometric spoofing, and challenges with "liveness" testing (Harakannanavar et al., 2019). For an example of an aging template, consider a picture of yourself at two years old and again at five years old, a biometric system would most likely categorize you as different people. In biometric systems, the FAR is the system saying you are not yourself, whereas an FRR is the system saying someone else is you. Biometric systems are purposefully designed and trained to

reduce false positives/negatives as much as possible, but these errors have not been eradicated in the field of study and practice. There are several other factors that cause biometrics to fall short; however, what have been outlined are the major categories. Cancelable biometrics is an area of biometric research that comes with some of its own issues. The foremost of which is the reduction of match confidence when using a transformed set of features, rather than the true biometric template. The one-way hash causes some information to be lost, which improves the match score. Some believe this makes cancelable biometrics untenable or not ready to be used in practice. Most biometric research includes testing biometric feature matching with a time lapse (single-day, multiday, multimonth, and multiyear) in between template capture (Harvey et al., 2018). To mitigate template aging, the cancelable parameters will be updated on a regular schedule based on the biometric feature(s) in use.

3.4 Serving Underserved Groups

One of the more meaningful reasons to build a blockchain SSI, starting with Kenyans as the inaugural population for the system, is the numerous challenges that present themselves with myriad groups in the country. Kenya's arid north is full of groups who are pastoralists, that is, who have no fixed address. There exists a tribe, the Maasai, who are pastoralists found throughout the country. Kenya is also home to many groups that live their entire lives on farmland far away from major cities and tech infrastructure. This system takes the needs and lifestyles of all of these different groups into account. Internally displaced persons (IDPs) are another group of individuals who can be aided by systems built using the defined framework. Internally displaced persons are those who have not fled their home country but have had to flee their homes due to terrorism (Nigeria's Boko Haram) and/or war (Owoaje et al., 2016). In the cases of many IDPs, they have lost all official claims of identity. After the collapse of the previous Somalia government, Canada's Department of Immigration and Refugees released a request for information explaining how identification documents could not be retrieved due to issues with civil management (C. Immigration, 2016). A self-sovereign identity (SSI) solution would help with all of the aforementioned cases, providing an identity that governments cannot erase and would be able to show whether or not the person ever had a verifiable identity from any government.

4 STAGE 2: INTEROPERABILITY AND SECURITY HARDENING

Today's world is changing rapidly and especially as we enter the fourth industrial revolution, the systems we build must be adaptable. History has shown us that the species that are most adaptable tend to have a higher survival rate than those that must cling to that with which it has always been familiar.

4.1 Interoperability

As the Internet and digital identity have progressed so has interoperability of differing types. New digital identity frameworks are being designed with an aspiration to achieve the efficiency of X-Road from Estonia. The X-Road government infrastructure supports a "once-only" approach to data access whereby no single piece of personal information should be entered twice (Saputro et al., 2020). Such an approach is possible due to individual servers being interlinked *via* end-to-end encrypted channels creating an X-like backbone that supports interoperability with relying systems. Secure access to the data is provided, given that a relying service cannot access personal data without approval by the owner of the information.

While backbone identity infrastructures exist in leading African economies, with the Integrated Population Registration Service (IPRS) in Kenya (Rading, 2019) and the NIMC Verification Service in Nigeria (KALU et al., 2018), they should have provision for the use of personal biometrics beyond enrollment of citizens into the systems. An additional layer that allows direct control by use of biometrics to access personal data would preserve information integrity, and an API first approach of state registries would be key in supporting interoperability of systems. As stated in the Authentication section 3.2, multifactor authentication (MFA) will be the system's initial way to manage authentication security. Biometrics is meant to be used as a part of MFA and later as the only way of authentication once the science (and technology affordability) reaches the proper stage of maturity for low-income individuals in postcolonial countries.

Interoperability has been achieved at different levels by some social networks and email providers, of most note is WeChat (Plantin and de Seta, 2019). WeChat is a Chinese digital infrastructure and platform for most things that can be accessed online in the country (Plantin and de Seta, 2019). Interoperability has already been solved by a few different approaches, of which X-Road (Saputro et al., 2020) and OAuth 2.0 (Hardt, 2012; Jones et al., 2015) are of most interest. Estonia's X-Road is of interest because it is the trusted Internet infrastructure for government entities (Saputro et al., 2020). Estonia's different identity systems and services run on it (Id-card, 2019; E-residency, 2019; Mobile-id, 2019). OAuth 2.0 is of interest because it is the protocol over which Javascript Web Tokens (JWTs) operate (Jones et al., 2015). The system will utilize OAuth 2.0 and JWT upon authentication to manage access to digital resources.

4.2 Security Hardening

In today's software practice, security patches have become quite the normal occurrence. Security patches apply to operating systems, developed by major companies and organizations, as well as mobile and computer applications. Common software engineering practice lends itself to security from compromise; however, in a world of humans where data are becoming more monetized and precious by the moment, we must design a system such that it keeps data safe from social engineering, biometric template theft, and general abuse/misuse.

Some security-hardening topics are not enumerated here as the cryptographic consensus–based distributed ledger manages to mitigate through its design, such as bad actors on the network (computers attempting to hijack the network), data intercepting (private data will be locked with encryption keys), and identity masquerading (transactions are signed). By using the blockchain, we introduce an ownerless distributed ledger that contains all historical system transactions. The distributed nature of the

blockchain is such that it allows *every* user of the system to view every transaction at any time. By using blockchain transactions, once they are submitted to the system, they cannot be modified in any way, including deletion. All of these blockchain attributes do a great job of keeping transactions and data secure.

Biometrics and system notifications will be used to help stymie social engineering approaches. Blockchain systems use private keys to manage data; however, an issue with such is that once a private key is lost, the certifications, claims, and assets related to that private key are forfeit. The pan-African system will use biometrics to aide in generation of the private key, so that it cannot be lost. Generation of cryptographic keys usually requires a random seed of some sort, and research exists that outlines how to use information from biometric templates to be that random seed. Such systems are referred to as Biometric CryptoSystems (Jin et al., 2016).

Template theft was addressed earlier along with the concept of cancelable biometrics (Ratha et al., 2001; Ratha et al., 2006). Part of security hardening is ensuring personal data cannot be shared without consent of the owner. To this end, we have to add smart contracts for the system that allow all personal data to be double-signed by the owner. Hence, when trying to move the data, a smart contract gives notice to the owner of someone's attempt to share their data. The smart contract will have to insist on approval by the data owner. If approval is not received after a certain time period and/or the data owner denies the operation, the network must cancel it while logging the transaction attempt. The system must automatically encrypt all personal data in personal claim repositories (user wallets). The wallets will be stored in a hybrid fashion on the cloud and on personal devices. Identity claims must be issued following a specific machine-readable format. The first signature is the data owners; the second is for transmission of data and consists of the public key of the recipient.

5 STAGE 3: LONGITUDINAL DATA STUDY

In biometric research, longitudinal studies are usually completed to prove assertions and learn more about a specific modality, as in evaluating the validity of a modality's persistence (Yoon and Jain, 2015). A longitudinal study is one in which the same group of participants are observed over an extended period of time, for example, 15 years. Such information, gleaned over time, has proven necessary for researchers and end users when making claims that can have legal ramifications.

In 2014, Yoon and Jain were able to perform such a study by using an "operational fingerprint database" (Yoon and Jain, 2015). This year, Mundnich et al. did a psychological and behavioral study utilizing data from "direct clinical providers in a hospital workplace" (Mundnich et al., 2020). One of the aims of our system is to obtain biometric and behavioral data without negative semblance. Speaking of negative semblances, we mean utilizing "records of repeat offenders apprehended by the MSP (Michigan State Police)" (U.S. citizen slave prisoners who have lost their human rights) (Yoon and Jain, 2015) and data sets of people who had to give away rights to certain data as an employment condition (Mundnich et al., 2020).

A reason a study is to be made with this framework is because of the current state of bias in biometric recognition systems (Buolamwini and Gebru, 2018). Machine learning models and scientists are majority Caucasian/Asian, and the major biometric face databases are of the same demographic. Buolamwini carried out studies and evaluations of face recognition corpi and systems of the largest providers of the technology in the United States. Buolamwini found "dark-skinned" women to be woefully underrepresented and dramatically misclassified, in comparison to lighter men (Buolamwini and Gebru, 2018).

Another reason a study to be made with this framework is to improve the system based on user feedback that will be completely optional. Biometric data are not the only information to be captured by the study but also various user sentiments, along with platform usefulness and usability. At each stage of the systems use, users will be able to provide feedback, at a granularity of their choice, which we will use to improve interactions, usability, partnerships, and more.

5.1 Participation Protocol

Participation in this study will follow strict guidelines to ensure participant privacy and secure their volunteered biometric data as much as possible. Our participation protocol has three components: fully informed self-sovereign volunteering (SSV), data obfuscation and usage, and self-sovereign control.

Fully informed self-sovereign volunteering (SSV) is the most ethical and responsible way to acquire information from people. SSV requires all data usage is logged to a blockchain network, and volunteers are notified as to how their data are being used. If their data are monetized, they will receive monetary reimbursement, using a model similar to that of Steem.com. Steem is a blockchain for the support of "community building and social interaction with cryptocurrency rewards" (STEEM, 2018). Concerning rewards for the monetization of the data of volunteers, a Steem-like system must be deployed on our network.

5.2 Data Handling

One-way hashing will be used to clean data of personally identifying information, such as names being attached to biometric signatures. The world is consistently moving forward with biometric research with every publication and new cell phone (Gelb and Clark, 2013). The data to be used along with registration in this system are multitudinous and by necessity will grow. As this is a framework intended to provide identity, in a complete sense, in a digital format only controllable by the owner of the identity, an exceptional amount of information can be gleaned from its proper study.

6 MOVING FORWARD

The requisite research and planning have been done for the implementation of the system to begin. Unstructured demographic data will be accepted into the system along with cancelable biometric templates. Acceptance of unstructured demographic data is to see what different populations deem as demographic data, populations that may not have much formal

education. Hyperledger Indy will be the first blockchain backbone component of the minimum viable product. As noted in research by Wingerde (van Wingerde, 2017) and Ferdous (Ferdous et al., 2019), the Sovrin platform, which uses Hyperledger Indy, is a popular blockchain identity system closer to being truly self-sovereign than others. Although Sovrin is the best system at the moment, it lacks a few features, those specifically outlined in van Wingerde (2017), which include the following:

- An individual needs another entity to generate a key pair (UC1-FR1).
- Identifiers are not generated on an open-source network not owned by a single entity (UC1-NFR2).
- Corresponding identifiers cannot stay the same upon loss of a private key (UC1-NFR3).
- Entities cannot associate an identifier with a human-readable name (UC2-NRF1).
- Not all data in personal data repositories are encrypted according to the highest industry standards (UC3-NFR1).

In order to reach the desired system, the blockchain on which Sovrin exists will require the addition of several smart contracts. More research is required to figure out the best way to fill in the gaps. Determination and full design of the longitudinal study must also be completed in order to have the study begin upon deployment of the system being built. The implementation and adoption of the system will lead us to a real conclusion of the efficacy of the ideas put forth.

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

SD is the main author, while JS is the main proof-reader.

REFERENCES

A primer for decentralized identifiers (2019). [Online]. Available: https://w3c-ccg.github.io/did-primer/.

Africa population (2020). "*Africa Population (Live)*". [Online]. Available: https://www.populationof.net/africa/.

Allison, J., Allison, P. J., Allison, M., and Allison, F. K. (2019). Blockchain Technologies: an Evaluation Using Digital Humanities as Search Light Revealing Nodes and Architectural Insights. *Res. Gate*. [Online]. Available: https://scholar.googleusercontent.com/scholar?q=cache:qZEN5_LfrV4J:scholar.google.com/+evaluation+of+inclusiveness+of+blockchain&hl=en&as_sdt=0,5.

Baars, D. (2016). *"Towards Self-Sovereign Identity Using Blockchain Technology," Master's Thesis*. University of Twente.

Bank-Id4D, W. (2017). *Principles on Identification for Sustainable Development : Toward the Digital Age*. [Online]. Available: http://documents.worldbank.org/curated/en/213581486378184357/Principles-on-identification-for-sustainable-development-toward-the-digital-age.

Buolamwini, J., and Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proc. Machine Learn. Res.* 81, 1–15. [Online]. Available: http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf.

C. Immigration (2016). *Response to information request som105248.e,*. [Online]. Available: https://irb-cisr.gc.ca/en/country-information/rir/Pages/index.aspx?doc=456434&pls=1.

Community Paper (2020). *Reimagining Digital Identity: A Strategic Imperative*. [Online]. Available: http://www3.weforum.org/docs/WEF_Digital_Identity_Strategic_Imperative.pdf.

Du Bois, W. E. B. (1974). The Pan-African Movement. in History of the Pan-African Congress. Pan African Congress.

E-residency (2019). *E-residency - E-estonia*. [Online]. Available: https://e-estonia.com/solutions/e-identity/e-residency/.

Fenker, S. P., and Bowyer, K. W. (2012). Analysis of Template Aging in Iris Biometrics. In 2012 IEEE Computer Society Conference On Computer Vision And Pattern Recognition Workshops. IEEE, 45–51.

Ferdous, M. S., Chowdhury, F., and Alassafi, M. O. (2019). In Search of Self-Sovereign Identity Leveraging Blockchain Technology. *IEEE Access* 7, 103 059–103 079. doi:10.1109/access.2019.2931173

Gatteschi, V., Lamberti, F., Demartini, C., Pranteda, C., and Santamaria, V. (2018). To Blockchain or Not to Blockchain: That Is the Question. *IT Prof.* 20 (2), 62–74. doi:10.1109/mitp.2018.021921652

Gelb, A., and Clark, J. (2013). Identification for Development: the Biometrics Revolution.*Cent. Glob. Dev. Working Paper*, 315, Center for Global Development Working Paper.

Harakannanavar, S. S., Renukamurthy, P. C., and Raja, K. B. (2019). Comprehensive Study of Biometric Authentication Systems, Challenges and Future Trends. *Ijana* 10 (4), 3958–3968. doi:10.35444/ijana.2019.10048

Hardt, D. (2012). *"The Oauth 2.0 Authorization Framework," RFC 6749*. Tech. Rep.

Harvey, J., Campbell, J., and Adler, A. (2018). Characterization of Biometric Template Aging in a Multiyear, Multivendor Longitudinal Fingerprint Matching Study. *IEEE Trans. Instrumentation Meas.* 68 (4), 1071–1079.

Henfridsson, O., Bygstad, B., and Bygstad, B. (2013). *"The Generative Mechanisms of Digital Infrastructure Evolution*," MIS quarterly, 37, 907–931. doi:10.25300/misq/2013/37.3.11

Id-card (2019). *Id-card - E-estonia*. [Online]. Available: https://e-estonia.com/solutions/e-identity/id-card/.

Intelligence, G. (2020). *"The mobile Economy Sub-saharan Africa 2020*. [Online]. Available: https://www.gsma.com/mobileeconomy/wp-content/uploads/2020/09/GSMA_MobileEconomy2020_SSA_Infographic.pdf

Jin, Z., Teoh, A. B. J., Goi, B.-M., and Tay, Y.-H. (2016). Biometric Cryptosystems: a New Biometric Key Binding and its Implementation for Fingerprint Minutiae-Based Representation. *Pattern Recognition* 56, 50–62. doi:10.1016/j.patcog.2016.02.024

Jones, M., Campbell, B., and Mortimore, C. (2015). *Json Web Token (Jwt) Profile for Oauth 2.0 Client Authentication and Authorization grants*. [Online]. Available: https://tools.ietf.org/html/rfc7523.

Kalu, M. I., David, N., and Nnaji, F. (2018). The Philosophy and Politics of National Identity Management in nigeria: A Case for Nation-Building. *Afr. J. Polit. Administrative Stud.* 11 (1).

Kenya, D. (2008). Deploying World Class Infrastructure Facilities & Services. *Kenya Vis.*

Merkle, R. C. (1989). One way hash functions and des. In Conference on the Theory and Application of Cryptology. Springer, 428–446.

Mobile-id (2019). *Mobile-id - E-estonia*. [Online]. Available: https://e-estonia.com/solutions/e-identity/mobile-id/.

Mundnich, K., Booth, B. M., L'Hommedieu, M., Feng, T., Girault, B., L'Hommedieu, J., et al. (2020). *Tiles-2018: A Longitudinal Physiologic and Behavioral Data Set of Hospital Workers*.

Naik, N., and Jenkins, P. (2020). Governing Principles of Self-Sovereign Identity Applied to Blockchain Enabled Privacy Preserving Identity Management Systems,. In IEEE International Symposium on Systems Engineering (ISSE). IEEE, 1–6.

Nash, J. (2020). *World Economic Forum Spells Out its Decentralized Biometric Travel Id Project*. [Online]. Available: https://www.biometricupdate.com/202003/world-economic-forum-spells-out-its-decentralized-biometric-travel-id-project.

Ometov, A., Bezzateev, S., Mäkitalo, N., Andreev, S., Mikkonen, T., and Koucheryavy, Y. (2018). Multi-factor Authentication: A Survey. *Cryptography* 2 (1). doi:10.3390/cryptography2010001

Ondego, B., and Moturi, C. (2016). Evaluation of the Implementation of the E-Citizen in kenya. *Int. J. Appl. Inf. Syst. (Ijais)* 10 (4). doi:10.5120/ijais2016451486

Owoaje, E., Uchendu, O., Ajayi, T., and Cadmus, E. (2016). A Review of the Health Problems of the Internally Displaced Persons in Africa. *Niger. Postgrad. Med. J.* 23 (4), 161–171. doi:10.4103/1117-1936.196242

Plantin, J.-C., and de Seta, G. (2019). Wechat as Infrastructure: The Techno-Nationalist Shaping of Chinese Digital Platforms. *Chin. J. Commun.* 12 (3), 257–273. doi:10.1080/17544750.2019.1572633

Rading, M. O. (2019). *Interoperability Framework for National Population Register a Case Study of Iprs*." Ph.D. dissertation, University of Nairobi.

Ratha, N., Connell, J., Bolle, R. M., and Chikkerur, S. (2006). Cancelable Biometrics: A Case Study in Fingerprints. In 18th International Conference on Pattern Recognition (ICPR'06), 4. IEEE, 370–373.

Ratha, N. K., Connell, J. H., and Bolle, R. M. (2001). Enhancing Security and Privacy in Biometrics-Based Authentication Systems. *IBM Syst. J.* 40 (3), 614–634. doi:10.1147/sj.403.0614

Saputro, R., Pappel, I., Vainsalu, H., Lips, S., and Draheim, D. (2020). Prerequisites for the Adoption of the X - Road Interoperability and Data Exchange Framework: A Comparative Study. in 2020 Seventh International Conference on eDemocracy eGovernment. ICEDEG), 216–222.

Sen, S. (2019). A Decade of Aadhaar: Lessons in Implementing a Foundational Id System. *ORF Issue Brief*, 292.

Smart-id (2019). *Smart-id - E-estonia*. [Online]. Available: https://e-estonia.com/solutions/e-identity/smart-id/.

STEEM (2018). *Steem: An Incentivized, Blockchain-Based, Public Content Platform*. [Online]. Available: https://steem.com/wp-content/uploads/2018/10/steem-whitepaper.pdf.

Tanui, C. (2018). "*Kenya's mobile Phone Penetration Surpasses 100% Mark*." [Online]. Available: https://kenyanwallstreet.com/kenyas-mobile-phone-penetration-surpasses-100-mark/.

Thompson, M. R., Essiari, A., and Mudumbai, S. (2003). Certificate-based Authorization Policy in a Pki Environment. *ACM Trans. Inf. Syst. Secur.* 6 (4), 566–588. doi:10.1145/950191.950196

Ul Hassan, M., Rehmani, M. H., and Chen, J. (2020). Differential Privacy in Blockchain Technology: A Futuristic Approach. *J. Parallel Distributed Comput.* 145, 50–74. doi:10.1016/j.jpdc.2020.06.003

van Wingerde, M. (2017). Tilburg University, School of Economics and Management."Blockchain-enabled Self-Sovereign Identity," Ph.D. Dissertation, Master's Thesis.

Wilson, S. (2005). The Importance of Pki Today. *China Commun.* 15.

Yoon, S., and Jain, A. K. (2015). Longitudinal Study of Fingerprint Recognition. *Proc. Natl. Acad. Sci. USA* 112 (28), 8555–8560. [Online]. Available: https://www.pnas.org/content/112/28/8555. doi:10.1073/pnas.1410272112

Zheng, Z., Xie, S., Dai, H., Chen, X., and Wang, H. (2017). "An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends," in IEEE international congress on big data (BigData congress). (IEEE), 557–564.

Conflict of Interest: SD was employed by the company Tint Right Colour Enterprise.

The remaining 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.