ESTABLISHING SELF SOVEREIGN IDENTITY WITH BLOCKCHAIN

EDITED BY: Alan Sherriff, Kaliya Young and Michael Shea
PUBLISHED IN: Frontiers in Blockchain

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-83250-121-4
DOI 10.3389/978-2-83250-121-4

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

ESTABLISHING SELF SOVEREIGN IDENTITY WITH BLOCKCHAIN

Topic Editors:
Alan Sherriff, Consultant, United Kingdom
Kaliya Young, Merritt College, United States
Michael Shea, Independent Researcher, United States

Citation: Sherriff, A., Young, K., Shea, M., eds. (2022). Establishing Self Sovereign Identity With Blockchain. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83250-121-4

Table of Contents

EDITORIAL
published: 19 August 2022
doi: 10.3389/fbloc.2022.955868

Editorial: Establishing Self Sovereign Identity with Blockchain

Alan Sherriff[1,2]*, Kaliya Young[3,4,5]* and Michael Shea[6,7]

[1]Independent Researcher, London, United Kingdom, [2]Manager, Time4Advice, Norwich, United Kingdom, [3]Merritt College, Oakland, CA, United States, [4]Independent Researcher, San Francisco, CA, United States, [5]Identity Woman, Kaliya Young, San Francisco, CA, United States, [6]Independent Researcher, Litchfield, CT, United States, [7]Consultant, Vienna, Austria

Keywords: self-sovereignty, digital identity, self-sovereign identity, decentralised identity, verifiable credentials, decentralised identifiers, DID, decentralisation

Editorial on the Research Topic

Establishing Self Sovereign Identity with Blockchain

Digital identity is a divisive topic. This is especially true of the Self-Sovereign Identity systems and standards being established today using decentralized architectures with blockchain technology foundations.

Self-Sovereign Identity (SSI) places individuals in control of their personal information; practitioners draw an analogy with physical wallets because a collection of digital identity credentials held in an SSI digital wallet are used within the digital realm in a similar way to paper and plastic credentials used in everyday life.

Fundamentally, the SSI approach serves to wrest control of digital identity from long-established centralized systems and authorities, democratizing and helping to rebalance the distribution of online power between individuals and institutions.

SSI represents a paradigm shift in the way digital identity is commonly managed and controlled today, presenting a serious alternative to many existing business models. This could provoke vested interests and opposition from global institutions and political structures that rely on their centralized digital identity infrastructures.

This Research Topic aims to provide a rich resource for identity practitioners, researchers, technologists, adopters, and policymakers to understand and advance the subject of SSI. While there is a wealth of publicly available material, this Research Topic provides the rigor of community peer review and the trust and confidence that this engenders.

The editors set out to curate a wide-ranging Research Topic of new academic research coupled with real-world experience and in-depth knowledge on the realities of implementing SSI. The topic brings together diverse perspectives from authors and reviewers invited from academia and industry, drawn from 12 countries and four continents, reviewed over a 2-year span. Several contributors have devoted their careers to SSI and been instrumental in driving SSI to where it is today, on the foothills of widescale adoption.

The 2-year curation period unintentionally incorporated the onset of the COVID-19 pandemic that hastened the use of online services, bringing debates around decentralized vs. centralized systems into the mainstream media. The pandemic created an unplanned pause in progress on this topic as most contributors were affected and many turned their energy to applying SSI technologies to the public health crisis, forming global collaboratives focused on helping global trade and travel to restart safely using privacy-preserving health credentials. Several papers were delayed and then revised in this light.

A common thread binding this topic together is the open standards that have emerged from the global SSI community. The Verifiable Credentials Data Model is now a W3C recommended standard. Coupled with Decentralized Identifiers (DID), these standards can ensure

OPEN ACCESS

Edited and reviewed by:
Richard Adams,
Cranfield University, United Kingdom

**Correspondence:*
Alan Sherriff
alan.sherriff@cboxx.com
Kaliya Young
kaliya@identitywoman.net

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

Received: 29 May 2022
Accepted: 15 June 2022
Published: 19 August 2022

Citation:
Sherriff A, Young K and Shea M (2022)
Editorial: Establishing Self Sovereign
Identity with Blockchain.
Front. Blockchain 5:955868.
doi: 10.3389/fbloc.2022.955868

interoperability between SSI ecosystems without reliance on centralized identity registries and authorities.

Identity systems have been evolving over centuries, accelerating recently with the advent of the internet, mobile digital devices, and population growth. This context is explored within several of the papers within this topic that are introduced below. The Research Topic provides new and experienced practitioners with a grounding on the development of identity systems, arriving at the rationale for SSI today within their own context. On a broad scale, a useful external reference on the historical development of identity systems and SSI can be read here Young et al.

The characteristics of SSI apply naturally to the *Blockchain for Good* forum inhabited by this Research Topic. This is highlighted by two case studies related to the African continent. The first describes a prototype system to carry out the initial steps of birth registration in an urban Kenyan setting (Freytsis et al.) using an identity system based on Verifiable Credentials and Decentralized Identifiers. Another outlines an ambitious solution to develop a blockchain based SSI backbone for Africa (Darnell and Sevilla) that seamlessly incorporates government issued identity documentation, providing a strategic vision that builds on research around the rate of mobile technology innovation within Kenya.

An underlying theme is the deep concern with the prevailing wind of global capitalist society and systems that prioritize profit and economic growth over well-being for people and the planet. The decentralized mindset of SSI engenders new possibilities for exploring change in our economic thinking and value accounting practice (Manski) and the need to design compelling value propositions to drive SSI adoption (Lockwood) that can play a part in addressing these concerns.

A common misconception of SSI is that the digital credentials are all self-certified and cannot be trusted in the way credentials issued by centralized trust authorities are. This is not true and, in fact, there is a sizable opportunity for today's trusted credential issuers to deploy their credentials into new decentralized channels that will provide security, privacy, and process efficiency benefits for all parties. Holders can take control of their data and eliminate the need for passwords. Issuers and relying parties dramatically reduce their exposure to data security risk by reducing their reliance on centralized identity silos that are a magnet for cybercrime.

This topic is a resource for those seeking to understand the building blocks and challenges of creating and growing SSI identity networks. Developing an SSI system is not straightforward; it takes a journey of collaboration and compromise. The Sovrin Network (Windley) identity metasystem is discussed and this itself is a deployment of Hyperledger Indy (Abramson et al.). Real-world lessons and recommendations are offered by the creators of a consortium-based approach to building an identity network for broad adoption across Canada (Boysen), based on SSI principles with blockchain. An insightful perspective is provided on high-profile projects in education, offering workable solutions to the key challenges within the blockchain-enabled, European digital credentials sector (Grech et al.). Arguably, SSI adoption still requires significant design-focused work at the human interface layer (Lockwood).

SSI relies on decentralized technology systems, governance, and compatible trust frameworks. These are complex concepts, and the successful establishment of SSI needs technologists and policymakers to work together (Chango) and appreciate their respective perspectives on digital identity.

At the time of writing, the DID specification stands on the brink of acceptance by the W3C as a new internet standard. Unsurprisingly, the key objectors to this milestone are a trio of global tech firms that control a large section of the web browser market and perhaps perceive a threat rather than recognizing a golden opportunity to embrace SSI for the benefit of all.

AUTHOR CONTRIBUTIONS

AS, KY, and MS are the topic editors. AS wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Conflict of interest: Author AS was employed by Time4Advice.

The remaining 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.

frontiers in Blockchain

REVIEW
published: 23 June 2020
doi: 10.3389/fbloc.2020.00029

Distributed Ledger Technologies, Value Accounting, and the Self Sovereign Identity

*Sarah Manski**

Department of Global Studies, University of California, Santa Barbara, Santa Barbara, CA, United States

Technological activists are designing blockchains and other distributed ledger technologies to challenge extractive value-accounting and identity management in global capitalism. This paper investigates how the new possibilities afforded through distributed ledger technology make possible an alternative future of generative value accounting and self-sovereign identity practices.

Keywords: self-sovereign identity, blockchain, holochain, distributed ledger technology, critical accounting, social movements, sociotechnical imaginary, value

> "If power is increasingly leveraged through online and mobile infrastructures-both on the part of movements and on the part of states-then some of the most important (and radical) movements will emerge around the use of those powerful technologies in societies." (Ilten and McInerney, 2019, p. 210).

INTRODUCTION

The problem with the logic of capitalism is that everything, including healthy social relationships, a stable climate, having meaning in life, etc. are only considered part of the value equation when it impacts profit. Technological activists are rejecting the logic of capitalism and insisting on creating a world where humans and living systems thrive, and therefore are developing new ways to recognize value.

Valuation is a social process, and accounting is a social practice (Callon, 1998; Boltanski and Thévenot, 2006; Callon et al., 2007; Knorr Cetina and Preda, 2012; Aspers and Dodd, 2015). Technological systems have shaped accounting in every setting, including the construction of markets, capital raising, algorithm pricing, digital platform services, and corporate organization. Some of these transformations have been the subjects of intensive study; research on others is lacking. This paper deals with new digital valuation technologies that could transform values and valuations within institutions in which valuation takes place. The same technologies will allow for the reclamation of our digital identities and real reputation, which is necessary for the trust required for online organizing. Technological activists are gaining momentum in their mission to design and use digital technologies for a world beyond capitalism. In this future, people, nature, and things are not valued by the market, but rather by their capacity to further human flourishing and account for planetary limitations. These efforts are part of three contemporary historical determinants recognized by technological activists: first, the need to evade state repression; second, the need to maximize limited resources; and third, the need to create effective institutional solutions despite past failures to do just that.

The construction of each accounting technology is mortared by ideology (Dillard, 1991). The dominant ideology of our age is capitalism. Everyday material technologies of accounting (written

OPEN ACCESS

Edited by:
Alan Sherriff,
Consultant, United Kingdom

Reviewed by:
Marc Rocas-Royo,
Open University of Catalonia, Spain
Gary Nuttall,
University of Nicosia, Cyprus

****Correspondence:***
Sarah Manski
Sarah@Manski.com

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

Received: *29 January 2020*
Accepted: *11 May 2020*
Published: *23 June 2020*

Citation:
Manski S (2020) Distributed Ledger Technologies, Value Accounting, and the Self Sovereign Identity. Front. Blockchain 3:29. doi: 10.3389/fbloc.2020.00029

reports, techniques, books of accounts, pictures, charts) make possible the practice of capitalist governance and corresponding modes of social control. Accounting technologies have material agency within large sociotechnical networks because they enable action at a distance (Robson, 1992), and they make "invisible" objects visible (MacKenzie, 2009). Inscriptions of accounts enable the modern state and institutions to "govern at a distance" and make present things, ideas, and people in "centers of calculation" (Latour, 1987; Miller, 1990).

> "… accounting cannot be independent of its social conditions. Under capitalism, the moving force of accounting lies in political economy—in class contradictions. Accounting is made, in part, by adjustment to the economic needs of the ruling class." (Catchpowle and Cooper, 1999, p. 712).

Tinker (1985) agrees that under capitalism accounting technology is a "logic for appropriating material production," "a way of rationalizing or explaining away the appropriation of the production of one social class by members of another" as "an intellectual and pragmatic tool in social domination" (p. 100). This understanding of accounting calls on scholars interested in building a world free of economic exploitation to understand how social movements and the technologists involved are creating new technologies of valuation and personal identity reflecting an emancipatory imaginary of the future beyond capitalism.

Dillard argues that a "fundamental change in the underlying economic structure must occur before change can occur in accounting technology" (p. 24), but what if technological activists within social movements can reverse this historical process and strategically radically redesign accounting technology; imbuing new accounting technology with favorable affordances that give it transformative material agency to fundamentally change the economic structure? The strategic design of technology has long been a part of activist repertoires. The use of value accounting to demonstrate exploitation and inequality against capitalist opponents is commonplace. It can be found among trade unionists and socialists (Gallhofer and Haslam, 2003), anti-sweatshop and fair-trade activists (Bartley and Child, 2014), anti-corporate globalization campaigners (Juris, 2007), and others. What is new in this historical moment are the emancipatory affordances of modern digital systems of value accounting and digital identity enabled by distributed ledger technologies or DLTs (i.e., blockchain and holochain).

An outline of this paper follows (see **Figure 1**). Part 1 describes the sociotechnical imaginary of a commons-oriented global social movement. Part 2 reviews what we know about how value is accounted for in capitalism, introduces a theoretical framework for understanding socio-economic objects within capitalist value accounting compared to commons value accounting, and includes a discussion on the tendency in capitalism toward increasing complexity. Part 3 discusses the affordances of blockchain technology, how we might begin to use the material agency of DLTs to shift the organization of value, and a discussion of self-sovereign identity's role in this process. Part 4 explores the possible futures of value accounting as glimpsed in the MetaCurrency project, Deep Wealth, Holochain, and the distributed application (hApp) Personas. This paper concludes with a vision of the future in which a global movement of entrepreneurs, cooperative members, and technological activists use a new form of value accounting to move beyond capitalism and create the next system for the benefit of all.

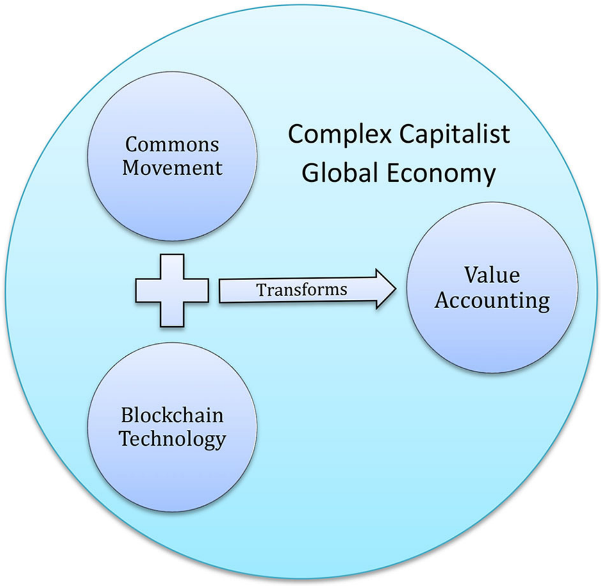


FIGURE 1 | Transformation occurring within capitalism as social movements combine with new technologies.

PART 1. THE SOCIOTECHNICAL IMAGINARY OF A COMMONS-ORIENTED GLOBAL SOCIAL MOVEMENT

Globally, hundreds of millions of people are rising and demanding that more than profit be valued (Della Porta et al., 2015). The determination of what is valuable is indicative of how societies can stay together, and what is valued demonstrates our collective social capacity and interdependence upon each other. People create technologies enabling their vision of the future and once created that technology does indeed expand what is possible in the future. Blockchain technology and the post-blockchain DLTs that followed the first Bitcoin blockchain (Nakamoto, 2008) are excellent examples of how sociotechnical imaginaries are put into practice through the design of new technologies.

The question here is, "How do distributed ledger technologies, a new set of technologies that include blockchain and post-blockchain systems, shape visions of the future, and how do these visions in turn influence the construction of new technologies?" Several approaches from the futures discourse could be taken to study the social and institutional practice of value accounting. For instance, causal layered analysis (CLA) could be useful for mapping and analyzing a number of competing discourses related to layers of worldview and metaphor (Inayatullah, 1998;

Inayatullah and Milojevic, 2015), or future empirical research could seek to quantify the multiple dimensions of trust within each technological accounting system and how that impacts user outcomes (Lander and Cooper, 2017). However, the advantage of making use of the concept of the sociotechnical imaginary in this theory paper is that it offers a framework for understanding how a technologist's vision of the ideal future influences their design choices in the present (Jasanoff and Kim, 2009).

Jasanoff and Kim label technologist's utopian vision of the future the "sociotechnical imaginary" (2009). This phrase incorporates the concept of the "sociotechnical" with that of the imaginary. In the field of Science and Technology Studies (STS) the term sociotechnical is used to indicate that technology is neither wholly socially determined nor deriving from an essential internal logic, "Technologies and technical practices are understood as durable (but not immutable) assemblages of social relation and technical artifacts" (Dunbar-Hester, 2019, p. 83). There is a lively discussion among technologists on how to use distributed ledger technology to realize a collective vision of a better future. The findings rely on grounded theory-based interpretations of numerous formal and informal interviews with technologists. Technologists shared an agreement on several standard components of a *global technological commonwealth*, the movement's emancipatory sociotechnical imaginary (Manski, 2017). This shared imaginary consists of a post-capitalist society where communities of mutual interest cooperate in the construction of institutions of regenerative economic relations. This movement of technologists has a strong faith in serendipity because they believe the necessary pieces will fall together if the correct intentions are directed outward and if the participants are mindful of the opportunities that can be pulled inward. These principles of technological design include:

→ Technological design should incorporate planetary boundaries
→ Technological design should be modeled on natural biological ecosystems
→ Technological design should enable the redefinition of value (ex. Distributed Value Accounting[1])
→ Technological design should enable radically democratic coordination and governance
→ Technological design should allow for the growth of a cooperative commons as the desirable future.

PART 2. HOW VALUE IS RECOGNIZED UNDER CAPITALISM

There is a renewal of interest among political economists and others in the role the imagination plays in shaping our vision of the future. Studies of future imaginaries have been conducted in the fields of climate engineering, body enhancement (Roco and Bainbridge, 2002), nanotechnology (Fiedeler et al., 2010), and synthetic biology (Giese et al., 2014). Through the use of our imagination and anticipatory thinking, we can build a bridge from our current present to the desired future present. When we make statements about the desired future, we are making an intervention in the present (Lösch, 2006), because future scenarios, once articulated, influence political debate and policy decisions (Selin, 2007). It is essential to recognize that people engaged in imagining the future bring to this process their ideology, interests, and positions of power within society (Brown et al., 2000). Every network architecture hides a power structure, "we can be a lot more nuanced in the design and usage of technologies by being explicit about the values we imprint in our economic systems" (DisCO.coop, 2019, p. 13). The dominant future imaginary is capitalist, but a commons movement is working on using blockchain technology to transform the nature of capitalist value accounting.

This section briefly reviews the literature on how value is currently accounted for under capitalism, including a discussion on the tendency in capitalism toward increasing complexity. As many researchers have observed, accounting is not neutral nor separate from prevailing economic ideology (Knights and Collinson, 1987; Catchpowle and Smyth, 2016). Critical to capitalism, new accounting and production technologies and organizational forms are invented to increase productivity, reduce the costs of production and manage the resulting processes and complexities (Cooper, 2015), "The only characteristics of concern are those associated with changes in the economic objects." (Dillard, 1991, p. 20).

The logic of capitalism derives from the drive to maximize profit (see **Table 2**). What is produced is driven by what can be profitably sold on the market, and production decisions are made by the quite small category of people—capitalists—who own and control the means of production. The labor of production is completed by wage laborers who must sell their labor to capitalists to survive, as they receive bank-credit money in return (McCarthy, 2018).

In Capital V1, Marx (2019) states what has value is only that which can be used to produce commodities that can be sold for profit in the market; this form of value is called *exchange value*. Such a market can only work with the existence of money as a material representation of value. It is the circulation of money as capital, the transformation of nature, and wage labor into commodities that have *exchange value* that drives capitalist economies. Marx envisioned a mechanization process that we now call modernization by which scientific knowledge and technology come to be more important factors in production. Competition inspires technological and organizational innovations that make value unstable and a "perpetually evolving inner connectivity (an internal or dialectical relation) between value as defined in the realm of circulation in the market and value as constantly being re-defined through revolutions in the realm of production." (Harvey, 2018), "Forces of production and social relations—two different sides of the development of the social individual—appear to capital as mere means and are merely means for it to produce on its limited foundation. In fact, however, they are the material conditions to blow this foundation sky-high." (Marx, 1993 [1857-8], p. 705-6). Technological innovations also

[1] See Manski, S. G., and Bauwens, M. (2020). Reimagining new socio-technical economics through the application of distributed ledger technologies. *Frontiers in Blockchain*. 2(29).

involve greater systemwide complexity, which carries its own challenges in part because defining the concept of complexity is a matter of debate (Pryor, 1996; Rosser, 1999). Hodgson (2003) defines complexity as systemically interconnected and interactive variety within a structured system, "By this definition, increasing economic complexity means a growing diversity of interactions between human beings and between people and their technology." (p. 472).

Early into the study of technology, Mumford (1996) recognized that technologies represent complex layers of objectified intentions that embody cultural artifacts into technical systems. As technological systems change over time, the original design choices gradually solidify and become viewed as timeless. These systems become interlocked and exert power over social systems and intuitions (Hughes, 1987). Our institutions are challenged by increasing complexity, and the digitization of the economy has accelerated this process (DeSanctis and Poole, 1994). The interconnectedness of complex systems makes outcomes more difficult to predict and causes negative consequences (Penow, 1986; Grabowski and Roberts, 1999). Massive amounts of information are available irrespective of geographic boundaries, and increasingly people have access to participation in a formal economy, which is governed by automated algorithmic systems communicating interdependently with each other. Humanity has attempted to solve coordination challenges in complex networks with systems of hierarchy, including monarchies, corporations, militaries, and representative democracies with layers of bureaucracy. Yet, current economic and governance patterns are proving inadequate (Duit and Galaz, 2008). Markets have been proposed as a solution, but current market approaches are proving inadequate, because markets tend to have limited or irregular communication patterns that do not contain information about all that is valued by society such as care work, environmental beauty, leisure time, etc. (Doane, 2002).

Price communicates across complex supply chains incredibly well, but the price of something is an oversimplified communicator of value. For example, when the price of copper goes up, the price of goods that use copper and the price of services that make use of those goods tend to go up as well. At the end of the line, a consumer can sense the difference between a supply chain that makes use of copper and one that makes use of a cheaper alternative because of the difference in sticker price at the point of purchase; the process by which "the invisible hand" functions (Hayek, 1945). However, other forms of information, such as the working conditions under which the copper was mined, or the environmental record of the mining company do not get communicated across the supply chain with nearly the same level of fidelity. This imbalance in the composability of price information vs. other forms of information leads to larger-scale effects that amount to a race to the bottom. The fact that price is the dominant form of information traveling with this level of efficacy is a challenge for technological activists and points toward potential technological solutions.

To overcome this problem, technological activists have asked, is it possible to increase the adaptive capacity of value accounting not just to single organizations but of markets more generally? The conclusion that many have arrived at is that what is needed is more rich and varied forms of information to be not only communicable but also composable. Currently, the use of dollars is the only value metric that is highly composable across contexts (Krafel, 1999; Harris-Braun and Brock, 2018). Activists argue that what is needed are other ways that individuals and communities can communicate about value in ways that can be composed across contexts because whether something is valuable depends on the context.

There is a tension here with the recognition that value judgments are always communicated within specific relationship contexts. And yet, it can be useful to have that information be composable beyond those specific contexts and can also end up altering the dynamic of that initial relationship in the future. For example, there is now pressure for restaurants to create photogenic food that will make a nice picture on Instagram. Matt Schutte, Holochain Director of Communications, argues, "In order to thrive we need to create value accounting systems that increase internal complexity." He is part of a movement of technologists using ideas drawn from the field of cybernetics to explore new technologically enabled protocol cooperativism accounting systems.

Organizational theory states that organized systems must adapt to their environment to survive (Lawrence and Lorsch, 1967; Aldrich, 1979). Ashby's (1961) Law of Requisite Variety, presupposes that "for any system to be stable, the number of states of its control mechanism must be greater than or equal to the number of states in the system being controlled." Activists argue that we need new ways to coordinate in an increasingly complex global system. Technological systems will enable advanced forms of social cooperation that form the principles of a new political economy, a global technological commonwealth. Their socio-technical imaginary of the economy is one in which the primary role of production is to meet the needs of the community; the productive assets are held in common under democratic control; people work because it provides meaning in their lives, and; money is a mutual credit system specific to the community's needs.

PART 3. ACCOUNTING FOR VALUE AND SELF-SOVEREIGN IDENTITY USING DISTRIBUTED LEDGER TECHNOLOGY

This section explores the social movements of the global technological commonwealth using new technologies strategically to shift value accounting to move beyond capitalism to a commons-based economic system that regenerates both people and the planet. The commons can be managed sustainably by local communities of peers when communities communicate to build standard protocols and rules that ensure their sustainability (Ostrom, 1990). Distributed ledger technologies can be designed for the creation of self-sustaining commons economies where all participants profit according to the value that they produce rather than trying to conform to the capitalist economy. These are the cyber-physical commons powered by blockchain networks, which are designed to align user incentives toward maintaining the system. Miners earn tokens, developers hold the tokens hoping their efforts will raise their value, and users purchase tokens creating demand and pay transaction fees.

Open shared ledgers are a key mutual coordination mechanism to shift open-source coordination from software to manufacturing. Blockchain and distributed ledgers generally enable open and contributive ecosystem accounting (such as practiced by SENSORICA (2019), REA (resource—event—agent), which let us see flows in shared circular economies involving multiple players, and biocapacity accounting, which is based on a direct vision of the flows of matter and energy. These types of contributory accounting systems promote fairness, openness, transparency, security, and environmental limits. The current state of the blockchain world is one of fragmentation, but the tools are in development for the creation of interoperable P2P ledgers.

For example, members of the Giveth team are using blockchain technology for good by building a toolkit for creating these new community economies. The project is called the Commons Stack and is a collaboration with BlockScience, a Complex Systems Engineering R&D firm. The Commons Stack is a project started in 2019 that aims to create community tools to improve decentralized coordination around shared goals. In these "community commons," blockchain technology is used to align economic incentives with each communities's values and scale these previously underfunded communal efforts into effective networks for good. They believe the growth of the commons will be accelerated through access to an open-source library of modular, customizable, and interoperable components enabling purpose-driven communities to unite around shared goals (**Figure 2**).

(**Figure 2**. The Commons Stack is building a library of tools for context-specific methods of governance, incentivization, accountability, monitoring, and initialization using holistic system simulations. Used with permission. Graphic by Jeff Emmet, published in "Commons Stack System Overview").

The Commons Stack project has identified components for what they term a "Minimum Viable Commons," to provide essential functionality in coordinating a group around raising and allocating funds, making decisions, and measuring impact. The first component is the "Augmented Bonding Curve," providing continuous funding for a commons initiative through community transvestment, with growing academic foundations for this new economic tool. The second is a transparent and accountable proposal service, which they call the "Giveth Proposal Engine." The third is a novel process for continuous decision making modeled off the mechanics of a neuron firing in the brain, called "Conviction Voting." And, finally, a means to monitor and measure the value produced in these communities, they term the "Commons Analytics Dashboard," which they see leading to a future of Computer Aided Governance. The most important aspect of the Commons Stack is their emphasis on Token Engineering, including the use of an open-source sophisticated system modeling and simulation tool called CAD.

The Commons Stack is using the emerging discipline of token engineering to design technological improvements to streamline community fundraising and decision making, lowering the barriers for groups with shared goals to operate as distributed protocol cooperatives. They are doing this by producing design patterns for community toolkits, a library of code specifications and reference implementations. These designs will be chain-agnostic and can be applied to data-centric and/or agent-centric architectures (see **Appendix 2**). However, most developer interest so far exists in the Ethereum ecosystem, so that is likely where they will see their designs first implemented.

The Commons Stack could be the technological evolution needed to enable the growth of the commons by enabling crypto-economic systems of cooperation and governance. This modular "cultural and technical stack for the commons," could help communities reach shared goals by giving them the tools to bootstrap necessary funding (often the main hindrance to launching), and empowering that community with proportionally weighted peer governance, real-time preference signaling, and monitoring systems that respect complexity. By creating a growing library of open source component blueprints for governance, funding, and other critical infrastructure, the Commons Stack enables communities to act as effective platform cooperatives, co-owning and co-managing shared funds as a commons. These components can be combined to create intentional, circular, community-driven economies powered by continuous funding streams and transparent decision making, which will enable the threefold coordination of the post-capitalist economy.

> "In materializing, objectifying, and displaying the value of acts, the publicity and formality of ritual approximate the way the market objectifies the value of work but making the consequences impossible to commoditize. One might even say that ritual *de-commoditizes* value." (Lambek, 2013, p. 154).

In the quote above, Lambek discusses how humans have used ritual to define community value. Ritual returns a sense of the sacred to human activity, while commodification alienates humans from their labor. Technological activists argue to move to a cooperative, post-capitalist planet; then on the societal level, we must decommodify human energies by treating our productive activity as sacred and ethical. Macpherson (1973) argued that human activity is moral when our internal and external motivations for performing acts are in alignment, "Man is not a bundle of appetites seeking satisfaction but a bundle of conscious energies seeking to be exerted" (p. 4–5). To the greatest extent possible, value must be *incommensurable*; meaning value must remain unique and unalienated.

How could we even begin this process? Technologies have material agency, defined as the structured set of relations enabling or constraining different sets of possibilities. On a global scale, technological activists are designing new technologies with the agency to open new pathways and foreclose others via the operation of technology's "material agency."

Distributed ledger technologies (DLTs), such as blockchains, are contributing to a wave of infrastructure distribution in industrial production. Such distribution is made possible because people place their trust in the software to accurately validate the

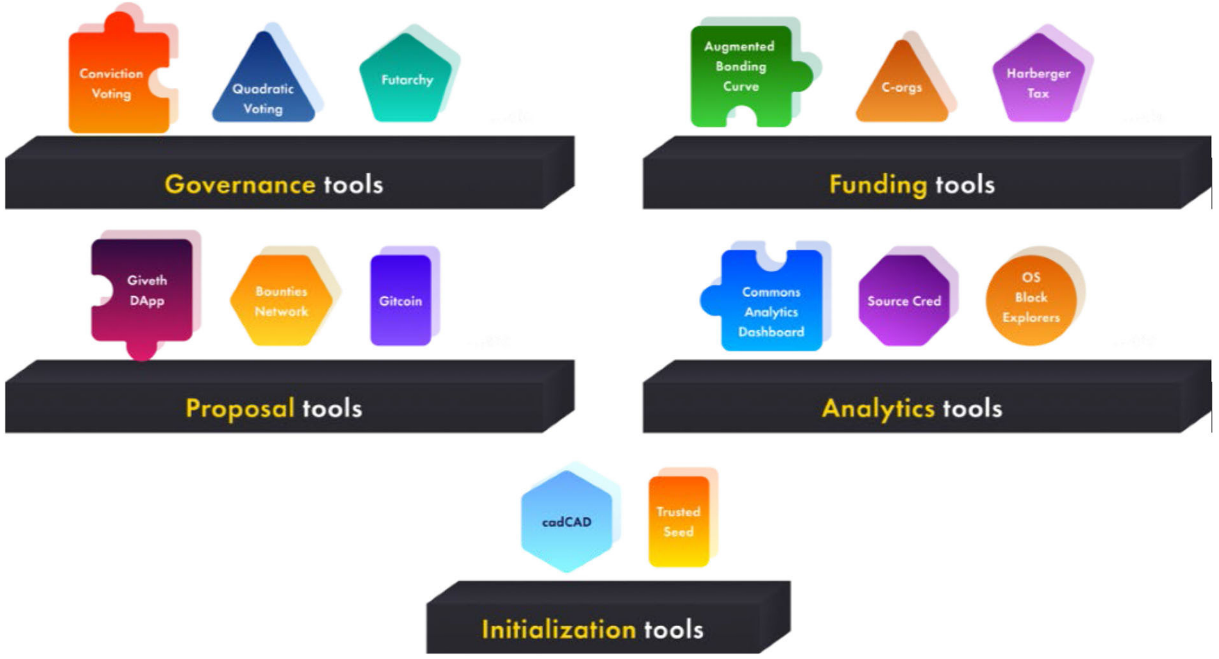


FIGURE 2 | A future stack for the commons (used with permission. Graphic by Jeff Emmet, published in "Architecting the Cyber-Physical Commons", https://medium.com/commonsstack/architecting-the-cyber-physical-commons-a294d88b5415).

transaction rather than trusting a bank or other intermediary. Some DLTs are being designed to be "unencloseable," meaning that no party can capture and control the communication occurring on the ledger. This is being done so that these DLTs are non-commodifiable in traditional capitalist markets, and in theory, they will enable more democratic forms of governance and organizational structures. Yet these radical possibilities will not be realized without strategic action to design systems that alter value in financial, service, and national infrastructures. Blockchain is an emergent technology created to enable the transfer of value with increased transparency, efficiency, and security (Nakamoto, 2008) that possesses a transformative material agency (Manski, 2017). The affordances of blockchain technology are directly available in its code, and seven such tendencies are listed in **Table 1**.

(Used with permission. The data from this **Table 1** are from: "No Gods No Masters No Coders? The Future of Sovereignty in a Blockchain World," by Manski, Sarah Grace, and Manski and Manski (2018), *Law and Critique*, 29:2, pp 151–162).

TABLE 1 | Seven tendencies of blockchain technology.

1. **Verifiability.** Transactions are assured through encrypted network consensus mechanisms in such a form that all transactions from the very first to the most recent are recorded in a ledger open to its maintainers, reducing information asymmetries.
2. **Globality.** Digital transactions and cultural information flows transcend geographic space and national borders.
3. **Liquidity.** Value liquidity is enhanced as the location of a store of value that does not depend on or is not under the direct control of a sovereign, central bank, or private corporation.
4. **Permanence.** The ledger of a transaction is immutable by design.
5. **Ethereality.** Transactions are conducted in a digital medium.
6. **Decentralization.** The ledger is widely distributed among many stakeholders and maintainers.
7. **Future Focus.** Found in newer developments of blockchain such as Ethereum, a stored autonomous self-reinforcing agency (SASRA) is formed in the temporal displacement of action through the use of smart contracts enabling the prefigurative recording of future transactions.

Used with permission. Published in Manski and Manski (2018).

The material agency of distributed ledger technology could enable "the construction of self-sovereign identity." The word sovereignty refers to "the receiving of a general recognition of exclusive domain and consequent possession of the capacity to establish the rules of conduct within a particular field of action" (Manski and Manski, 2018). We all have both offline and online identities. For anyone who uses digital systems, tied to our material identity are various digital identities. For the most part, these digital identities are not under our control, and often, we are not able to see what information is contained within each system. Problematically, if the information is incorrect, we cannot correct these errors, nor do we control what and with whom information is shared and sold. The self-sovereign infrastructure allows users to set boundaries regarding who has access to their data and maintain their privacy. It can also reward users for being contributors. This infrastructure thus will enable people to protect their autonomy while conducting joint work and collective action.

TABLE 2 | Socio-economic objects within capitalist value accounting compared to commons value accounting.

Socio-economic objects	Capitalist value accounting	Commons value accounting
Human labor	Commodity value	Reflects the species being
Time	Continuous made discrete	Experienced via natural body processes
Institutions	Embodiments of class hierarchies	Reflects individuals' perceptions of themselves
Transactions	Restricted to narrow prespecified attributes	Incorporates a broad range of social/environment attributes
Means of production	Capital dominates labor	Labor dominates capital

Our "'technical' technologies will not generate broad human gains unless we invest an equal amount of time, energy, and resources in the development of social and emotional technologies that drive how our whole society is organized and how we work together." Kaliya Identity Woman (2017) is making the argument any technology controlled by corporations and governments will always restrict human progress. Manski and Manski (2018) support this point by outlining five possible future scenarios of blockchain technology and conclude technology without direct social movement intervention always reinforces existing power relations. They take a contrasting position to those who believe individual self-sovereign identity is the most likely outcome of the use of new technology,

> "In blockchain's tendencies toward verifiability, globality, permanence, and future focus, state actors are finding greater capacities to intervene globally in the daily lives of individuals. These expanded capacities are making possible the emergence of new technological totalitarian forms of state sovereignty. To begin with, states cannot easily control what they cannot measure, and a blockchain-enabled Internet of Things (IoT) amplified by artificial intelligence furthers the degree with which states can monitor the material and social world. The rapidly expanding IoT is expected to more than triple in size by 2020 to nearly 21 billion devices (Stavridis and Weinstein, 2016). When there is a tiny blockchain-connected chip embedded in each material object with which we interact, state institutions will assuredly seek to monitor and discipline the personal, political, and economic activities of the many." (Manski and Manski, 2018).

PART 4. FUTURE OF VALUE ACCOUNTING: METACURRENCY, DEEP WEALTH, HOLOCHAIN AND PERSONAS

Holochain is a clear case of a new technology strategically created by social movement activists to achieve regenerative value accounting, which they call "holoptical" knowledge accounting. On New Year's Eve of 2016, Eric Harris-Braun and Arthur Brock started to build Holochain, "For me, what we need to create is a very rich multidimensional accounting. We need lots of feedback loops beyond the single dimension of price." (Brock, 2017). Holochain was created by the founders of the MetaCurrency Project to realize a part of their socio-technical imaginary. Holochain is the foundation for Holo, a cloud hosting market for dApps, and the future of the Internet.

> "Holo as a name was not pulled from a hat. It has roots. Back in the original collective intelligence of tribal communities, we had holopticism where everyone participates as part of a feedback loop of the whole. Holomidal instead of pyramidal, where, together, we sensed the whole. This isn't just voting or even decision-making, this is about an embodied integral experience of sensing together. In the original collective intelligence, there is co-creation and connectedness, even in a changing environment. We need less democratic debate in this form because holoptical clarity shapes individual actions. Holopticism doesn't mean you see everything; you see the whole from your perspective. And it, then, collectively becomes aperspectival by our sensing and communicating together." (Russell, 2018)

Holochain is a DLT platform-based ecosystem with affordances fostering co-production, open content, and co-ownership. Holochain, aims at facilitating interconnectivity among direct and indirect participants, such as those who install Holochain on their hardware devices to provide hosting space and those who access Holochain through a web browser. One of the advantages of this design is that it avoids the blockchain requirement for global consensus among maintainers and thus affords greater scalability, as well as 'self-sovereignty;' the user controls their data and identity information. The design of holochain is extremely distributed for a DLT[2]. Holochain activists call this design 'agent-centric' as opposed to corporate 'data-centric' models. The affordances of this design mean that users are given sovereign control over their data and are solely responsible for granting permissions.

Holochain does not have a built-in currency or token. However, the distributed internet architecture Holo does use a cryptocurrency HOT Fuel and Holochain was designed to make it easy to bid alternative cryptocurrencies in the form of distributed accounting applications (dApps). Holo Fuel, a mutual credit system (Manski and Bauwens, 2020), will cover the costs for data storage and Holochain development and maintenance. Holo Fuel is not a crypto-token or cryptocoin, but a mutual credit system issued within a double accounting system where one party holds a debit (the provider of goods and services) and the other party holds a credit (a debt to the provider of goods and services). On Holochain every transaction is countersigned on the local chains of both counterparties. Holo Fuel will be purchased as a token or received as a credit. This process occurs either through the exchange of fiat currency or another cryptocurrency into Holo Fuel or by setting credit limits; Holo Fuel to be paid later. The exchange of money and cryptocurrencies into Holo Fuel and the allocation of credit limits are done through the "Reserve Accounts," which is a facility provided by Harris-Braun and Brock (2018).

[2] For a comparison between Blockchain and Holochain, see **Appendix 1**.

> "MetaCurrency is the name for the infrastructure and protocols necessary for an open source economy, and free currencies to flow in an interoperable and standardized way" (Harris-Braun, 2018). The open source economy and free currencies are meant to function in a non-monopolizable manner by building protocols and platforms to 'open source' the next economy (Brock, 2009), "Building the core infrastructure for open sourcing money and currencies and developing projects that embody the values of Deep Wealth design." (Harris-Braun et al., 2018).

The concept of Deep Wealth (see **Figure 3**) shifts the value accounting incentive from accumulating material wealth to experiencing wealth through elements such as beautiful surroundings, friendship, capacity of being generous, leisure, travel, family and fun and perhaps, most importantly, deep connections with others (Brock, 2009). In the view of Metacurrency, there are three forms of wealth. The first is *tradable wealth*: food, time, energy, services, material resources, etc. The second is *measurable wealth*: performance, sustainability, physiological health, quality, etc. and the third is *acknowledgeable wealth*: fun, love, care, trust, beauty, etc. Each of these three forms of wealth is a subset of the other. For example, time is *tradable* wealth as well as *measurable* and *acknowledgeable* wealth; together, they create "integral wealth" (The MetaCurrency Project).

Technological activists involved with Metacurrency and Holochain refer to a "quantum leap" transition from a complex capitalist political economy to a post-capitalist society where information technology plays a significant role in fostering the creation of large-scale collective intelligence. The design of Metacurrency's technologies model the same organizational patterns as living systems. By "living systems," these activists refer to biological organisms, atoms, forests, languages and other continuously transforming systems, "the same kind of architectures of intelligence that makes it possible for trillions of cells to work together in an organism." (Harris-Braun and Brock, 2018). Within this architecture, "communication is virtually instantaneous (electronic), peered, decentralized, semantic and designed to evolve in response to rapidly changing needs" (Harris-Braun et al., 2018). Such communication parameters lead to effective, large scale, distributed collaboration that would remove "most of the power structures that underpin the social barriers to change and could make formerly intractable problems (such as climate change, species extinction, resource depletion, or poverty) quite readily solvable." (Harris-Braun et al., 2018).

Blockchains are token-centric, and by this, I mean that they are concerned with the history of token transactions and not necessarily with the people at the end of each transaction. In contrast, the creators of Holochain designed *Personas* as an agent-centric solution that allows individual users to maintain a reputation. This reputation will document their behavior within a community and across multiple applications that need a person's profile information. In this way, users will be able to trust those with whom they are transacting. Personas allow the user to store and edit their information in one account, similar to "log-in with Google[3]" and offer/revoke any applications' access to it[4] In addition, to control your data, Personas is designed to allow users to create multiple identities within each account so that users can have a different business, personal, government, medical, family, and friend personas. Each persona can also have an expiration date. It allows for the revocability of data as required by European Union law[5].

The definition of value has been changing for the past few decades, from market value to community value, and distributed ledger technologies are furthering this transformation by pushing out centralized identities in favor of self-sovereign identities. The widespread adoption of self-sovereign identity applications, such as Holochain-based Personas, is still yet to be realized, but the incredible interest in user-controlled identity makes it likely that some DLT application will make this a reality. Distributed ledgers are a critical piece of the puzzle of technologies including, smartphones, cloud computing, public key infrastructure (PKI), open standards for decentralized identifiers, directed identifiers, and open standards for verified claims (DIDs) fitting together to enable self-sovereign identities.

CONCLUSION

This article seeks to begin a dialogue on the topic of how distributed ledger technologies may transform our understanding of value and identity. Valuation is a social process, and distributed organizations of technological activists are utilizing new technologies to disrupt accounting and identity management in contemporary capitalism, and thus transforming global economics, the nature of work, and the distribution of wealth. This paper explores the radical generative accounting practices and ideological imaginaries underpinning this new form of social movement activism, and whether or not the development of new technologies of value accounting and self-sovereign identity may address the challenges of an increasingly complex global political economy of the future.

There is not a straight line between technological innovation and the increasing complexity of the political economy. As a society, we can decide to create technologies that will enrich humanity rather than commodify it. However, it is a certainty that if we continue to live on a planet where capitalism is the dominant determinant of value accounting and social identity, then expanding complexity and distorted value accounting will usher humanity to the edge of the collapse of democratic civilization.

Self-Sovereign Identity is a necessary but insufficient tool to deal with some aspects of growing complexity. Only a widespread popular global movement will have the power to snuff out the underlying drivers of capitalism. As a part of this process, we can use new forms of value accounting to reinforce and reify the social system under which we imagine we want to live. There is a growing movement of social entrepreneurs, cooperatives,

[3] Test out Personas' demo here: https://bit.ly/2SfhKIT.

[4] See how PayPall sells your data here: https://rebecca-ricks.com/paypal-data/.

[5] the General Data Protection Regulation 2016/679.

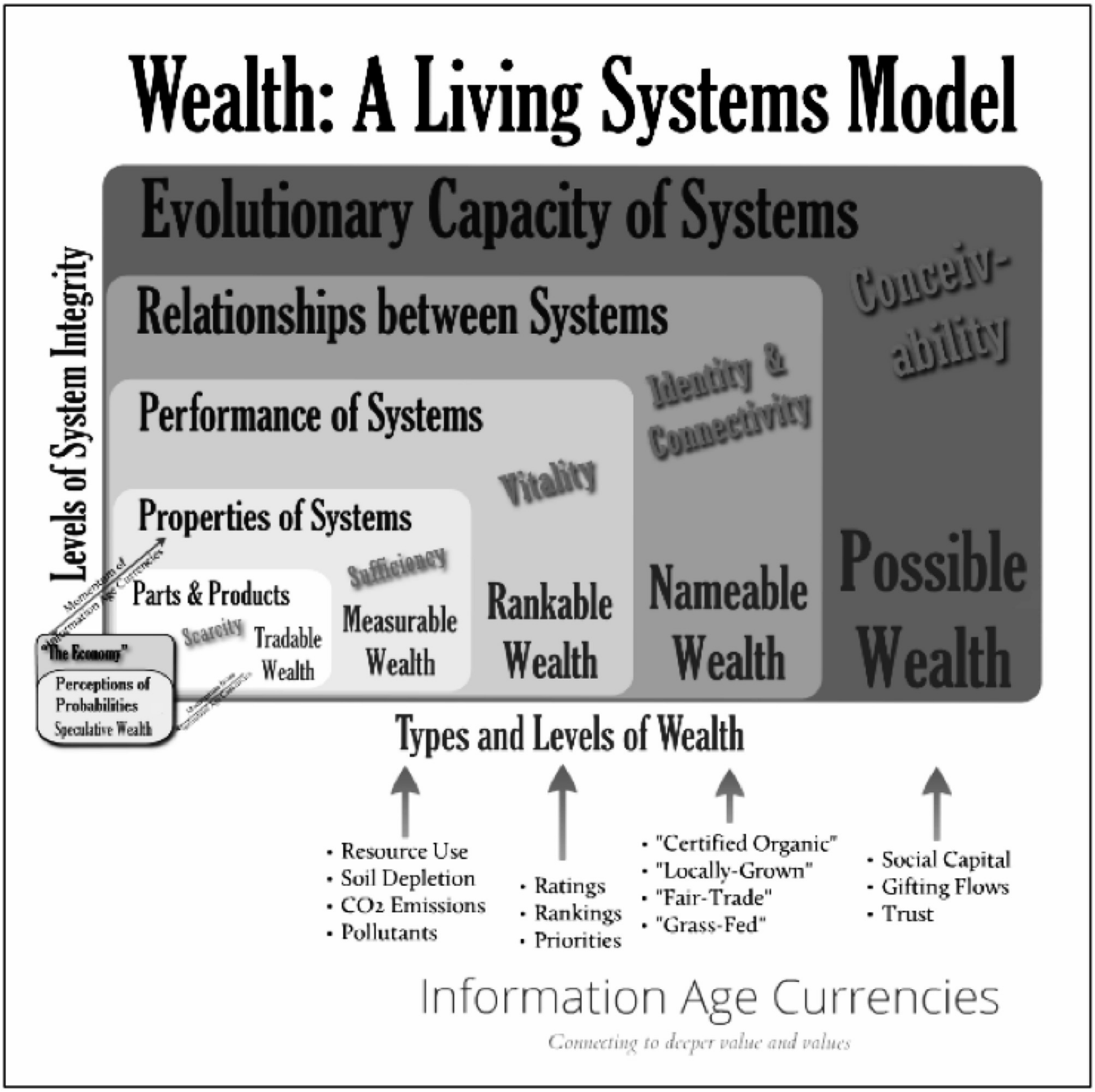


FIGURE 3 | Deep Wealth (used with permission. Graphic by Arthur Brock, published at https://metacurrency.org/portfolio-item/living-systems-model-of-wealth/).

and technological activists who are using these technologies in pursuit of cooperative ownership and management of wealth. It is in everyone's interest to pay attention to this development.

AUTHOR CONTRIBUTIONS

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

REFERENCES

Aldrich, H. E. (1979). *Organizations and Environments*. Englewood Cliffs, NJ: Prentice-Hall.

Ashby, W. R. (1961). *An Introduction to Cybernetics*. London: Chapman and Hall Ltd.

Aspers, P., and Dodd, N. (2015). *Re-imagining Economic Sociology*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780198748465.001.0001

Bartley, T., and Child, C. (2014). Shaming the corporation: the social production of targets and the anti-sweatshop movement. *Am. Soc. Rev.* 79, 653–679. doi: 10.1177/0003122414540653

Boltanski, L., and Thévenot, L. (2006). *On Justification. Economies of Worth. Translated by Catherine Porter.* Princeton, NY: Princeton University Press.

Brock, A. (2009). *New Economy, New Wealth.* Retrieved. Available online at: https://prezi.com/xmzld_-wayho/new-economy-new-wealth (accessed August 8, 2019).

Brock, A. (2017). *Intro to Currency Design. ArtBrock.com.* Available online at: https://www.artbrock.com/2017/02/27/intro-to-currency-design (accessed January 12, 2019).

Brown, J., Rappert, B., and Webster, A. (2000). *Contested Futures. A Sociology of Prospective Techno- Science.* Burlington, VT: Ashgate.

Callon, M. (1998). *The Laws of the Markets.* Oxford: Blackwell.

Callon, M., Millo, Y., and Muniesa, F. (2007). *Market Devices.* Oxford: Blackwell.

Catchpowle, L., and Cooper, C. (1999). No escaping the financial: the economic referent in South Africa. *Crit. Perspect. Account.* 10, 711–746. doi: 10.1006/cpac.1998.0257

Catchpowle, L., and Smyth, S. (2016). Accounting and social movements: an exploration of critical accounting praxis. *Account. Forum* 40, 220–234. doi: 10.1016/j.accfor.2016.05.001

Cooper, C. (2015). Accounting for the fictitious: a Marxist contribution to understanding accounting's roles in the financial crisis. *Crit. Perspect. Account.* 30, 63-82. doi: 10.1016/j.cpa.2014.08.002

Della Porta, D., Andretta, M., Calle, A., Combes, H., Eggert, N., Giugni, M. G., et al. (2015). *Global Justice Movement: Cross-National and Transnational Perspectives.* London: Routledge.

DeSanctis, G., and Poole, M. S. (1994). Capturing the complexity in advanced technology use: adaptive structuration theory. *Organ. Sci.* 5, 121–147. doi: 10.1287/orsc.5.2.121

Dillard, J. (1991). Accounting as a critical social science. *Account. Audit. Accountability J.* 4, 8–28. doi: 10.1108/09513579110143849

DisCO.coop (2019). *If I Only Had a Heart: A DisCO Manifesto.* Available online: https://disco.coop/wp-content/uploads/2019/11/DisCO_Manifesto-v1-1.pdf (accessed April 1, 2020).

Doane, D. (2002). *Market Failure: the Case for Mandatory Social and Environmental Reporting.* London: New Economics Foundation.

Duit, A., and Galaz, V. (2008). Governance and complexity—emerging issues for governance theory. *Governance* 21, 311–335. doi: 10.1111/j.1468-0491.2008.00402.x

Dunbar-Hester, C. (2019). "If "Diversity" Is the answer, what is the question? Understanding diversity advocacy in voluntaristic technology projects," in *digitalSTS: A Field Guide for Science & Technology Studies* 81.

Fiedeler, U., Coenen, C., Davies, S.R., and Ferrari, A. (2010). *Understanding Nanotechnology: Philosophy, Policy and Publics.* Heidelberg: AKA.

Gallhofer, S., and Haslam, J. (2003). *Accounting and Emancipation: Some Critical Interventions.* London: Routledge. doi: 10.4324/9780203986622

Giese, B., Pade, C., Wigger, H., and von Gleich, A. (2014). *Synthetic Biology: Character and Impact.* Heidelberg: Springer. doi: 10.1007/978-3-319-02783-8

Grabowski, M. and Roberts, K. H. (1999). Risk mitigation in virtual organizations. *Organ. Sci.* 10, 704–721. doi: 10.1287/orsc.10.6.704

Harris-Braun, E. (2018). *Metacurrency Project - P2P Foundation.* Available online at: http://wiki.p2pfoundation.net/Metacurrency_Project (accessed August 14th, 2019).

Harris-Braun, E., and Brock, A. (2018). *Grammatic Capacities and the Evolution of Complex Adaptive Systems.* Available online at: https://docs.google.com/document/d/12Jd7ZeUzT-k6h2Qxw9jmOhNVlkYD1RnO4t2_mp6g4kg/edit (accessed August 14, 2019).

Harris-Braun, W., Luck, N., Perrin, N., Russell, J., McGuire, E., Harris-Braun, E., et al. (2018). *Holo Green Paper.* Available online at: https://files.holo.host/2018/03/Holo-Green-Paper.pdf (August 8, 2019).

Harvey, D. (2018). *Marx's Refusal of the Labour Theory of Value.* Retrieved from http://davidharvey.org/wp-content/uploads/2018/03/MARX%E2%80%99S_REFUSAL_OF_THE_LTV.pdf

Hayek, F. A. (1945). The use of knowledge in society. *Am. Econ. Rev.* 35, 519–530.

Hodgson, G. (2003). Capitalism, complexity, and inequality. *J. Econ. Issues* 37, 471–478. doi: 10.1080/00213624.2003.11506595

Hughes, T. P. (1987). "The evolution of large technological systems," in *The Social Construction of Technological Systems,* eds W. E. Bijker, T. P. Hughes, and T. Pinch (Cambridge, Ma: The MIT Press), 51–82.

Ilten, C., and McInerney, P. B. (2019). "Social movements and digital technology a research Agenda," in *digitalSTS: A Field Guide for Science and Technology Studies* (Princeton, NJ: Princeton University Press), 198. doi: 10.2307/j.ctvc77mp9.18

Inayatullah, S. (1998). Causal layered analysis: poststructuralism as method. *Futures* 30, 815–829. doi: 10.1016/S0016-3287(98)00086-X

Inayatullah, S., and Milojevic, I. (2015). *CLA 2.0: Transformative Research in Theory and Practice* Tamsui: Tamkang University Press.

Jasanoff, S., and Kim, S. H. (2009). Containing the atom: Sociotechnical imaginaries and nuclear power in the United States and South Korea. *Minerva* 47:119. doi: 10.1007/s11024-009-9124-4

Juris, J. S. (2007). "Practicing militant ethnography with the movement for global resistance in Barcelona," in *Constituent imagination: Militant Investigations, Collective Theorization* (Chico, CA: AK Press), 11–34.

Kaliya Identity Woman. (2017). *Humanizing Technology. openDemocracy.* Available online at: https://www.opendemocracy.net/en/transformation/humanizing-technology/ (accessed March 23, 2019).

Knights, D., and Collinson, D. (1987). Disciplining the shopfloor: a comparison of the disciplinary effects of managerial psychology and financial accounting. *Account. Organ. Soc.* 12, 457–477. doi: 10.1016/0361-3682(87)90031-6

Knorr Cetina, K., and Preda, A. (2012). *The Oxford Handbook of the Sociology of Finance.* Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199590162.001.0001

Krafel, P. (1999). *Seeing Nature: Deliberate Encounters with the Visible World.* Hartford, VT: Chelsea Green.

Lambek, M. (2013). The Value of (performative) Acts. *HAU: J. Ethnographic Theor.* 3, 141–160. doi: 10.14318/hau3.2.009

Lander, L., and Cooper, N. (2017). *Promoting Public Deliberation in Low Trust Environments.* Australian Use Cases. doi: 10.2139/ssrn.3077474

Latour, B. (1987). *Science in Action.* Cambridge, MA: Harvard University Press.

Lawrence, P. R., and Lorsch, J. W. (1967). Differentiation and integration in complex organizations. *Adm. Sci. Q.* 12, 1–47. doi: 10.2307/2391211

Lösch, A. (2006). *Anticipating the Future of Nanotechnology: Some Thoughts on the Boundaries of Sociotechnological Visions.* Department of Sociology, Technical University Darmstadt.

MacKenzie, D. (2009). Making things the same: gases, emission rights and the politics of carbon markets. *Account. Organ. Soc.* 34, 440–455. doi: 10.1016/j.aos.2008.02.004

Macpherson, C. B. (1973). *Democratic Theory: Essays in Retrieval.* London: Oxford Press.

Manski, S. G. (2017). Building the blockchain world: technological commonwealth or just more of the same? *Strategic Change* 26, 511–522. doi: 10.1002/jsc.2151

Manski, S. G., and Bauwens, M. (2020). Reimagining new socio-technical economics through the application of distributed ledger technologies. *Front. Blockchain* 2:29. doi: 10.3389/fbloc.2019.00029

Manski, S. G., and Manski, B. (2018). No gods, no masters, no coders? The future of sovereignty in a blockchain world. *Law & Critique* 29, 151–162. doi: 10.1007/s10978-018-9225-z

Marx, K. (1993). *Grundrisse.* London: Penguin.

Marx, K. (2019). *Capital: Volume One.* Mineola, NY: Courier Dover Publications.

McCarthy, M. (2018). *Is Sociology Stuck in the Middle? The uses of Marxist General Theory. Marxist Sociology Blog.* Available online at: https://marxistsociology.org/2018/10/is-sociology-stuck-in-the-middle-the-uses-of-marxist-general-theory/ (accessed January 12, 2019).

Miller, P. (1990). On the interrelations between accounting and the state. *Account. Organ. Soc.* 15, 315–38. doi: 10.1016/0361-3682(90)90022-M

Mumford, L. (1966). Technics and the Nature of Man. *Technol. Cult.* 7, 303–317. doi: 10.2307/3101930

Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System.* Available online at: https://bitcoin.org/en/bitcoin-paper (accessed January 12, 2019).

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action.* Cambridge University Press.

Penow, C. (1986). *Complex Organizations.* New York, NY: Random House, third edition.

Pryor, F. (1996). *Economic Evolution and Structure: The Impact of Complexity of the US Economic System.* Cambridge; New York, NY: Cambridge University Press.

Robson, K. (1992). Accounting numbers as 'inscription': action at a distance and the development of accounting. *Account. Organ. Soc.* 17, 685–708. doi: 10.1016/0361-3682(92)90019-O

Roco, M. C., and Bainbridge, W.S. (2002). *Converging Technologies for Improving Human Performance*. Arlington, TX: Kluwer Academic. doi: 10.1007/978-94-017-0359-8

Rosser, J. (1999). On the complexities of complex economic dynamics. *J. Econ. Perspect.* 13, 169–192. doi: 10.1257/jep.13.4.169

Russell, J. (2018). *Beyond Democracy: Increasing the Capacity, Context, and Combinatorial Possibilities. Medium.com.* Available online at: https://medium.com/holochain/beyond-democracy-increasing-the-capacity-context-and-combinatorial-possibilities-7eebdd4ad079 (accessed December 27, 2018).

Selin, C. (2007). Expectations and the emergence of nanotechnology. *Sci. Technol. Hum. Values* 32, 196–220. doi: 10.1177/0162243906296918

SENSORICA (2019). *Homepage*. Available online at: http://www.sensorica.co/ (accessed February 1, 2019).

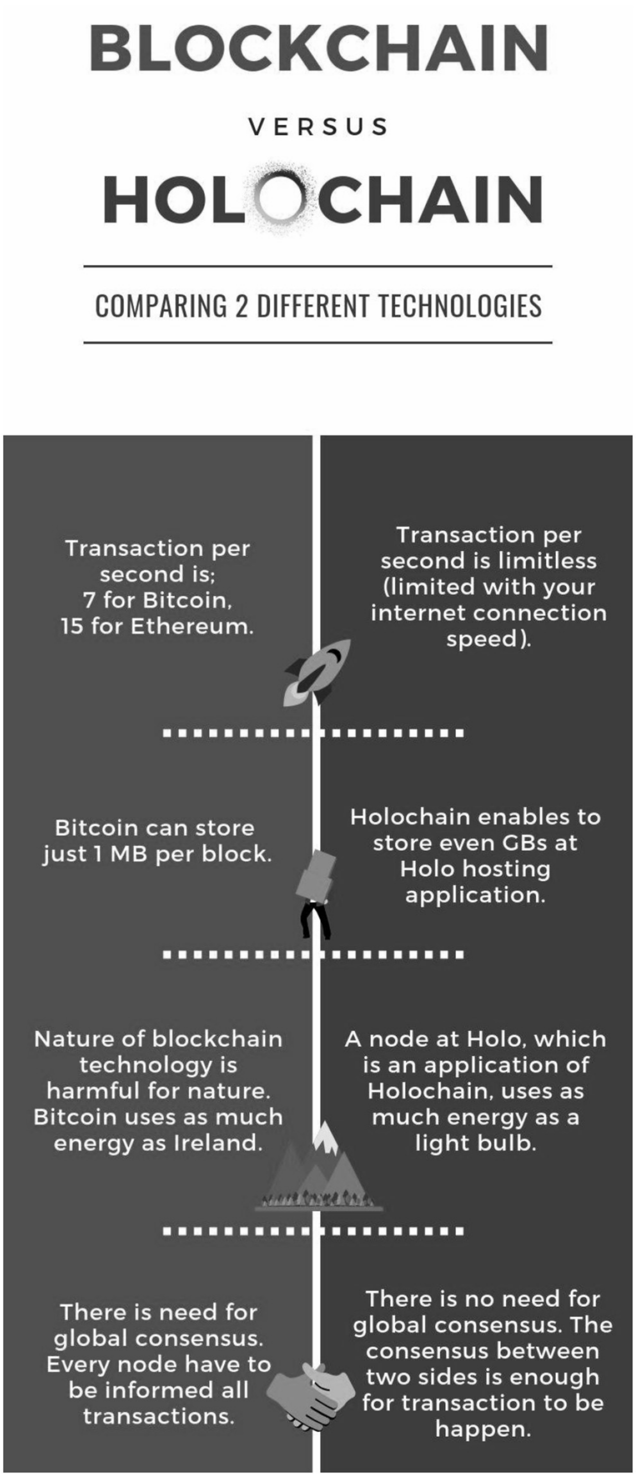
Stavridis, J., and Weinstein, D. (2016). The internet of things is a cyberwar nightmare. *Foreign Policy*. 3.

The MetaCurrency Project. (2019). *About - The MetaCurrency Project.* Available online at: http://metacurrency.org/about/ (accessed February 1, 2019).

Tinker, T. (1985). *Paper Prophets.* New York, NY: Praeger Press.

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.

APPENDIX 1



BLOCKCHAIN

VERSUS

HOLOCHAIN

COMPARING 2 DIFFERENT TECHNOLOGIES

Transaction per second is; 7 for Bitcoin, 15 for Ethereum.

Transaction per second is limitless (limited with your internet connection speed).

Bitcoin can store just 1 MB per block.

Holochain enables to store even GBs at Holo hosting application.

Nature of blockchain technology is harmful for nature. Bitcoin uses as much energy as Ireland.

A node at Holo, which is an application of Holochain, uses as much energy as a light bulb.

There is need for global consensus. Every node have to be informed all transactions.

There is no need for global consensus. The consensus between two sides is enough for transaction to be happen.

Blockchain versus Holochain (image used with permission. Created by Holochain).

APPENDIX 2

Agent Centric and Mutual Sovereignty Holochain Design Principles

- Anybody can try a new grammar (tweets, likes, rideshare requests, five-star ratings, etc.) without needing permission or support from others.
- Anyone that wants to communicate with them using that new grammar, can do so.
- If it proves useful, they can keep using it without requiring a business model that can extract value from the participants (they are using it due to intrinsic value).
- If it starts to prove too costly, annoying, or simply useless, they can alter in any way they see fit or abandon that grammar altogether.
- This enables far greater responsiveness by the participants in a community to the circumstances they face.
- Make it difficult for both corporations and states (powerful actors) from foreclosing possibilities.
- There are ecologically inspired patterns of organization that simply aren't possible with existing tools (http, dollars, incorporation, etc.). New tools like Holochain can enable coordination that is not dependent on access to or control over existing power structures (corporations, governments, etc.).
- It does not free people of the control that powerful entities might seek to wield. But it enables them to coordinate independently if they choose to.

frontiers
in Blockchain

COMMUNITY CASE STUDY
published: 15 February 2021
doi: 10.3389/fbloc.2021.631341

Development of a Mobile, Self-Sovereign Identity Approach for Facility Birth Registration in Kenya

Maria Freytsis[1]*, *Iain Barclay*[2], *Swapna Krishnakumar Radha*[3], *Adam Czajka*[4], *Geoffery H. Siwo*[3,5], *Ian Taylor*[2,3] *and Sherri Bucher*[5,6]*

[1]NeoInnovate Collaborative Consortium, New York, NY, United States, [2]School of Computer Science and Informatics, Cardiff University, Cardiff, United Kingdom, [3]Center for Research Computing, University of Notre Dame, Notre Dame, IN, Unitde States, [4]Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN, United States, [5]Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN, United States, [6]Indiana University School of Medicine, Department of Pediatrics, NeoInnovate Collaborative Consortium, Indianapolis, IN, United States

Birth registration is a critical element of newborn care. Increasing the coverage of birth registration is an essential part of the strategy to improve newborn survival globally, and is central to achieving greater health, social, and economic equity as defined under the United Nations Sustainable Development Goals. Parts of Eastern and Southern Africa have some of the lowest birth registration rates in the world. Mobile technologies have been used successfully with mothers and health workers in Africa to increase coverage of essential newborn care, including birth registration. However, mounting concerns about data ownership and data protection in the digital age are driving the search for scalable, user-centered, privacy protecting identity solutions. There is increasing interest in understanding if a self-sovereign identity (SSI) approach can help lower the barriers to birth registration by empowering families with a smartphone based process while providing high levels of data privacy and security in populations where birth registration rates are low. The process of birth registration and the barriers experienced by stakeholders are highly contextual. There is currently a gap in the literature with regard to modeling birth registration using SSI technology. This paper describes the development of a smartphone-based prototype system that allows interaction between families and health workers to carry out the initial steps of birth registration and linkage of mothers-baby pairs in an urban Kenyan setting using verifiable credentials, decentralized identifiers, and the emerging standards for their implementation in identity systems. The goal of the project was to develop a high fidelity prototype that could be used to obtain end-user feedback related to the feasibility and acceptability of an SSI approach in a particular Kenyan healthcare context. This paper will focus on how this technology was adapted for the specific context and implications for future research.

Keywords: self-sovereign identity, decentralized identifiers (DIDs), blockchain, birth registration, newborn health, mobile phones, Kenya

OPEN ACCESS

Edited by:
Kaliya Young,
Merritt College, United States

Reviewed by:
Ilene Speizer,
University of North Carolina at Chapel Hill, United States
Paula Braitstein,
University of Toronto, Canada

Correspondence:
Maria Freytsis
mhealthmidwife@gmail.com
Sherri Bucher
shbucher@iu.edu

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

Received: *19 November 2020*
Accepted: *08 January 2021*
Published: *15 February 2021*

Citation:
Freytsis M, Barclay I, Radha SK, Czajka A, Siwo GH, Taylor I and Bucher S (2021) Development of a Mobile, Self-Sovereign Identity Approach for Facility Birth Registration in Kenya.
Front. Blockchain 4:631341.
doi: 10.3389/fbloc.2021.631341

1 INTRODUCTION

The Sustainable Development Goals agenda (UN General Assembly, 2015), launched by the United Nations in 2015, spurred renewed focus on the high rates of neonatal mortality and low rates of birth registration in low and middle-income countries (LMICs). Neonatal mortality is a key indicator of the overall well-being of a society, and birth registration is central to ensuring not only newborn health, but continued well-being and access to essential rights and services throughout the lifespan (Solberg, 2015). Low rates of birth registration, the lack of a reliable system for vital statistics reporting and tracking of mother-baby pairs contribute to the problem of excess neonatal mortality by preventing health care systems from effectively delivering crucial evidence-based interventions (Hereward et al., 2019).

An estimated 2.5 million babies globally die during their first month of life, known as the neonatal period, and approximately 47% of all the deaths of children under 5 years old occurred in the neonatal period (Hug et al., 2019). Birth registration is almost universal in most high income countries where the vast majority of births take place in facilities. But in LMICs, where many births take place in the home, about 1 in 4 children under age five are not registered. Of the children who are registered, an estimated 237 million children under age five globally do not have proof of registration in the form of a birth certificate (Selim, 2019). Some of the highest rates of neonatal mortality and lowest rates of birth registration can be found in parts of Sub-Saharan Africa. SDG targets 16.9 and 17.9 set forth the goals of providing legal identity for all, including birth registration, by 2030, and supporting countries to establish robust civil registration systems (Hereward et al., 2019). The agenda calls for development of innovative technologies to help reach these targets, however, experts are calling for caution that new technologies do not increase the potential for human rights abuses or further disenfranchize marginalized populations. A lack of adequate data governance infrastructure across nations threatens the ability of identity technologies to protect the personal identifying data of both children and guardians (Privacy International, 2018; World Bank, 2018; Hug et al., 2019; Schoemaker et al., 2019).

Continued growth in mobile device and telecommunication service penetration, as well as a decreasing gender gap in digital literacy and device ownership in Sub-Saharan Africa has led to a proliferation of mobile-phone based programs targeting maternal and newborn health (GSM Association, 2013; Sahay et al., 2013; Kurth et al., 2015; Sondaal et al., 2016; Rowntree and Shanahan, 2020). In 2019 Kenya reported 51% penetration of unique mobile subscribers which was up from 29.9% in 2009. The penetration of mobile internet use in 2019 was 25.8% (GSM Association, 2020). For the last 10 years, the NeoInnovate Collaborative Consortium (NCC), a multidisciplinary group of academic collaborators from multiple universities, has been developing and studying technologies to reduce preventable newborn mortality in Sub-Saharan Africa, with a particular focus on Kenya. Employing a user-centered design approach with end-users in the Moi Teaching and Referral Hospital system, the NCC built and deployed its foundational technology called mobile Helping Babies Survive (mHBS). mHBS is a suite of mobile phone-based applications for training, clinical decision support, and data collection, developed to support health workers in the scale-up of Helping Babies Survive programs. The applications were built on the open source DHIS2 platform, which is also the national health data reporting system in Kenya (Manya et al., 2012). mHBS was developed for Android smart phones using an iterative process with multiple rounds of user testing (Bucher et al., 2020). Integration of birth registration and mother/baby linkage with mHBS/DHIS2 has been an area of interest to the collaborative as it could allow for tracking of mother/baby pairs to provide invaluable data on essential newborn care delivered and newborn health outcomes (Aluvaala and English, 2020).

In 2018, the Kenyan Ministry of ICT solicited stakeholder presentations addressing opportunities for use of blockchain technology in Kenya[1]. In response to this solicitation, NCC developed a vision for a birth registration and newborn health technology incorporating blockchain and self-sovereign identity (SSI) principles, as a proof of concept project called NeoLinkID. SSI describes the ability of an individual to have ownership of their personal data and to control who has access to that data, without the need for centralized infrastructure, or any control or authorization by any third party (Allen, 2016). Acknowledging that blockchain-based technologies were gaining increasing attention in a landscape of poor global data governance infrastructure, the project was envisioned as an opportunity to understand how the SSI approach can offer additional data protection from within the context of the centralized birth registration process in Kenya and the challenges involved in integrating the SSI layer into the existing system.

A partnership with the Evernym ID for Good accelerator[2], and support of Google Summer of Code[3], enabled NCC to rapidly develop a prototype that will allow end-user feedback on the feasibility and acceptability of this technology approach with both health workers and families in a facility-based birth registration use case in Kenya. While the majority of unregistered Kenyan births take place in the community, the technology platform's heavy reliance on connected environments made it clear that development for community settings would not be possible at the outset. However, developing this technology would allow for study of other possible benefits of this approach, such as facilitating privacy preserving digital linkage with the national birth certificate acquisition process, contribution of aggregate newborn health data to vital statistics and public health via DHIS2 integration, and linkage of mother-baby pairs for health tracking via a mobile personal health record for newborns held by guardians. A different research team within NCC is currently working to develop a solution to work in environments with limited internet connectivity.

[1] https://ict.go.ke/taskforce-on-distributed-ledgers-and-artificial-intelligence-presentation-schedule

[2] https://www.evernym.com/identityforgood/

[3] https://summerofcode.withgoogle.com

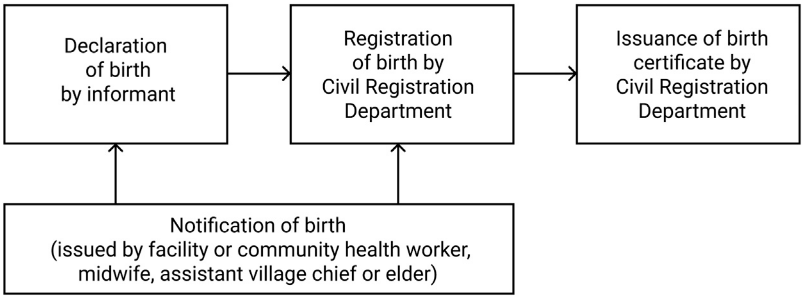

Declaration of birth by informant
Registration of birth by Civil Registration Department
Issuance of birth certificate by Civil Registration Department
Notification of birth (issued by facility or community health worker, midwife, assistant village chief or elder)

FIGURE 1 | Kenyan birth registration process.

2 KENYAN CONTEXT: BIRTH REGISTRATION AND DIGITAL INNOVATION

Currently, the neonatal mortality rate (NMR) in Kenya is 21/1,000 (UNICEF[4]); the SDG NMR target, by 2030, is 12/1,000. Despite significant changes in the health system over the past few years, including devolution to the County level (Barker et al., 2014), and elimination of user fees for facility-based births, rates of maternal and perinatal mortality have remained stagnant (Kunkel et al., 2019; Gitobu et al., 2018). The current birth registration rate in Kenya is 67%, with rural areas having 61% coverage, and urban areas at 79%[5]. However, rates as low as 20% and as high as 90% have been documented across regions (Gelb et al., 2016). In one study published in 2014 about half of the participating Kenyan children had birth certificates, with participants in urban areas more likely to have birth certificates than in rural locations (Apland et al., 2014).

In Kenya, as in many other LMICs, birth registration is a process by which "informants" who include health workers in facilities (primarily midwives), and assistant village chiefs/elders in the communities, acting on behalf of the Civil Registration Department (CRD), interact with "guardians" (primarily mothers and fathers) in order to provide documentation of birth for both the guardians and for the health and civil registration systems (**Figure 1**). Informants verify the parents' identities via legal identity documents and document birth details and other background information as required by Form B1[6]. Informants are legally authorized and obligated to do this in relation to their role as health care providers and assistant village chiefs/elders. The issuance of a birth notification document, which is the top portion of Form B1, by the informant to the guardian is the first step in birth registration. The remaining part of the form is sent to the local office of the CRD which allows for the birth to be counted in vital statistics. When parents take the next step of applying for the birth certificate at the local CRD office, they are required to present the birth notification document, which will be matched with the lower portion of the form submitted by the informant (Apland et al., 2014; MEASURE Evaluation, 2014).

Without a birth certificate, Kenyan parents may not be able to access services, ranging from health insurance to education, which would ensure a thriving childhood. The birth certificate is helpful but not required to obtain the current form of national ID (Apland et al., 2014). However, lacking a birth certificate can be a profound threat to children during periods of conflict or forced migration, and a life-long barrier to accessing essential services such as voting, obtaining a passport, opening a bank account, and mobile phone ownership which enables access to a growing range of essential digital services (Apland et al., 2014; Selim, 2019).

The barriers to birth registration are complex. Although there are common themes across countries and geographic areas, these barriers are also highly contextual. Over the last 20 years there have been multiple initiatives aimed at strengthening the Kenyan civil registration system including research and targeted interventions aimed at specifically understanding and ameliorating the barriers to birth registration. (Apland et al., 2014; MEASURE Evaluation, 2014; Gelb et al., 2016). Barriers fall into several categories including:

1. **Lack of awareness** regarding the importance of birth registration and obtaining a birth certificate.
2. **Physical and situational difficulties** contribute to the inability to complete the multi-step process. This may include traveling long distances to the registration authority and fear of penalties for late registration, which in Kenya can include not only monetary fees but prison time. Additionally, language barriers or illiteracy, or parents not having the personal identity documents required for registration, such as in the case of refugees or stateless groups, can also be contributing factors. There are also cultural beliefs within some groups that are not aligned with registering newborn births.
3. **Discriminatory laws and practices** which prohibit certain groups of people from accessing birth registration based on race, ethnicity, religion, gender, or other characteristics. This can include officials requiring bribes to process applications.
4. **Inadequate staff and infrastructure** to perform birth registration efficiently, as well as negative attitudes of some registration workers (Apland et al., 2014; Gelb et al., 2016; UNHCR, 2017; Kenya Human Rights Commission, 2019).

[4] https://data.unicef.org/country/ken/

[5] https://data.worldbank.org/indicator/SP.REG.BRTH.ZS?end=2014&locations=KE&start=2003&view=chart

[6] http://forms.co.ke/forms/41_Birth_Acknowledgement-of-Birth-Notification-For-Parents_Form-B1.pdf

Research in Kenya has also found a lack of motivation by parents to register their children in advance of any particular need and local officials not placing a high priority on incentivizing registration (Gelb et al., 2016). A 2014 study found Kenyan parents reported high awareness and low practical barriers and concluded that parents are making a deliberate informed choice not to pursue birth registration, weighing the cost and benefit. Suggestions of the researchers included increasing the use of information and communication technology to impact parental decision-making (Pelowski et al., 2015). Other interventions have focused on mass education campaigns, targeted education and incentivization of community based birth registration including linking opportunity for registration with other essential services such as immunization and school enrollment, and introduction of mobile technology to support the birth registration process (World Health Organization, 2013; Apland et al., 2014).

Kenya has an extremely active, innovative, and engaged digital health landscape representing an enabling environment for new technology. Known as the "silicon savannah" (Schoemaker et al., 2019), Kenya has led the world in regards to technological innovations related to mobile banking (e.g., mPESA[7]), crowdsourced, decentralized monitoring and reporting (Ushahidi), and development of vibrant innovation ecosystems (iHub[8]). Mobile phone-base health interventions have also proliferated over the last decade with many focusing on maternal and child health. However, evidence shows that many of these interventions have not been evaluated and few have scaled beyond pilot projects. Additionally, few projects had been implemented in marginalized areas with more health care needs (Njoroge et al., 2017).

The Kenyan context also presents myriad challenges for implementing identity technology including a high proportion of vulnerable and marginalized populations such as refugees and other groups of stateless persons, a fragmented identity ecosystem, increasing reliance on digital services with a private sector directly connected to the state, and a government with opaque operations, dense bureaucracy, and a history of corruption. At the same time, e-Government services are growing, and the entire health reporting system has been cloud-based since 2010. The current national debate about a new form of biometric national ID has brought issues of data protection to the forefront. The policy landscape related to data protection is evolving with significant advocacy efforts from community service organizations and numerous failed attempts at passing data protection legislation, culminating in the current Data Protection Act, introduced in 2019, making its way through the lawmaking process (India, 2013; Gelb et al., 2016; Kenya Human Rights Commission, 2019; Schoemaker et al., 2019).

3 USE CASE: TECHNOLOGY DESIGN AND DEVELOPMENT

3.1 Self-Sovereign Identity Background

SSI is built upon well established cryptography techniques (Preneel, 1994) where a securely held private key is used to sign documents, while a complimentary public key can be used to verify the signature and that the document has not been tampered with. Researchers have developed mechanisms (Sporny et al., 2019) for organisations and individuals to issue signed credentials to other parties, where each is identified by a unique decentralized identifier (DID). In this way, any party attesting something about another party can declare and sign their claim, using their DID and cryptographic protocols. The signed document is known as a Verifiable Credential (VC), and is held by the subject of the credential, or in the case of a child, by a guardian. At a later date, when the holder of a VC needs to enter into a transaction, a service provider can request proof of their status or entitlements. The holder of the credential can generate a Verifiable Presentation containing assertions from the VC document, to provide cryptographically verifiable proof of the claims being made. A level of privacy is provided by the principle of Selective Disclosure, which allows VC holders to provide presentations containing only selected elements of credentials, so that they retain control over the data shared in any individual transaction. These data models and protocols are implemented in software toolkits, which can be used by third party developers to add SSI capabilities to their solutions. The Sovrin network provides the foundation for many of the toolkits, including the Evernym platform adopted here, and uses the public-permissioned Hyperledger Indy blockchain ledger to keep a permanent and immutable record of the DIDs of public agencies, along with credential schemas. No personal information is written to the blockchain (Kondova and Erbguth, 2020).

3.2 Modeling Birth Registration Processes as an SSI System

Prototype design and development was preceded by a research phase including a literature review and interviews with Kenyan facility-based midwives to understand the current birth registration process. A set of personas representing the two groups of end users–guardians (parents, primarily mothers) and informants (facility-based health workers, primarily midwives), were developed. A high level use case description based on these personas was conceived and then evolved into user stories describing the workflows for the two applications. Assumptions regarding the selected context and users include consistent mobile device access with the possibility of shared devices, English language literacy, technology literacy with Android smartphones, reliable connectivity, and adequate mobile data. Interactions between informants and guardians were considered to take place at the facility where the birth occurred prior to the mother/baby dyad's discharge home. The prototype represented the process of creating a digital "copy" of the birth registration process alongside the current paper-based process. Personas for the roles of informant and guardian in the

[7] https://techcrunch.com/2015/07/23/the-rise-of-silicon-savannah-and-africas-tech-movement/

[8] https://www.dw.com/en/finding-digital-solutions-to-local-problems-kenyas-innovation-scene-is-no-one-hit-wonder/a-47119339-0

TABLE 1 | Credentials developed for Prototype.

Credential	Issuer	Holder	Notes
Informant	CRD (research team)	Informant	Allows informant access to system (not implemented in prototype)
Identity verification	Informant (for CRD)	Guardian	Asserts that guardian's identity has been verified
BND	Informant (for CRD)	Guardian	Birth notification document (one for each child)
Link credential	Informant (for CRD)	Guardian	Forms a connection between the guardian and each newborn
BCG vaccination	Informant (for CRD)	Guardian	Sample of one vaccination as a credential given to each child

community-based birth registration process were also developed to illustrate the challenges of extending this technology approach to those populations who may have lower language literacy, less proficiency with smartphones, unreliable connectivity, and less consistent device and data access.

Designing systems to use decentralized SSI protocols involves developing an understanding of the participants of the system in the real world, their differing goals, and the interactions and dependencies that they have on each other to fulfill those goals. In previous work (Barclay et al., 2020) the authors introduced an application of the iStar (Yu, 2011) conceptual modeling framework as a method of describing the actors and their interactions in an SSI system, in the context of the birth registration process. In order to meet their goal of "Get Birth Certificate", a newborn's guardian needs to receive a copy of the birth notification document (BND), which is issued to them by the informant. The prototype required that the informant issues the BND on provision of a suitable proof of the guardian's identity. In a deployed system, this type of requirement would be a government policy decision, and the technology would be adapted to meet the needs.

3.3 Architecture

The NeoLinkID prototype adopts three participant roles–CRD, as the credential issuing authority (and represented by the research team in the prototype design); the informant, or health worker; and the guardian, as the recipient and holder of the credential. In interactions, the informant is considered to be a representative of the CRD, which manifests in the architecture as the informant requesting the CRD to issue credentials. Credentials developed for the prototype include a credential to record that the informant has checked the guardian's identity, the BND and a linking credential which creates mother-baby linkage, as detailed in **Table 1**.

Issues related to the SSI concept of "guardianship"[9] were not addressed in credential development, but rather, the Link Credential was built to provide a mechanism to study context for the development of a local guardianship framework. Verification of credentials was not addressed in the first iteration but the design process allowed for considerations of various verification scenarios.

In the prototype system, the informant is issued with a mobile web application which accesses a server operated for the CRD. The server integrates with the core SSI platform, which provides verifiable credential structures, and populates and issues credentials. The informant's web application provides forms to collect information about guardians and newborns, and then requests the CRD server to sign and issue credentials to the guardian. An Android smartphone application was developed to receive and hold credentials for the guardian. No direct connection is established between the informant's web application and the guardian's application. The informant's application provides data to the server to populate and request credentials to be sent to the guardian's application. An informant will meet with many guardians during the course of the day. To issue credentials to the correct guardian, the informant application creates a connection between the server and the guardian they are currently meeting by presenting a QR code to be scanned by the guardian. When the appointment finishes the informant ends the session, and the connection is removed. The informant's application displays forms that are populated with the guardian and used to generate and issue credentials. The server sends credentials directly to the guardian's mobile application, where they are stored for later use. **Figure 2** shows the handshaking and credential issuing process flows and the interactions between the three participants. Screenshots of the applications developed for the informant and the guardian are shown in **Figure 3**.

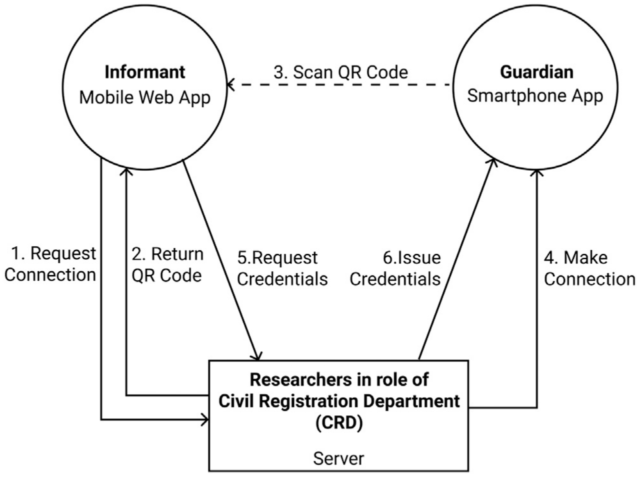
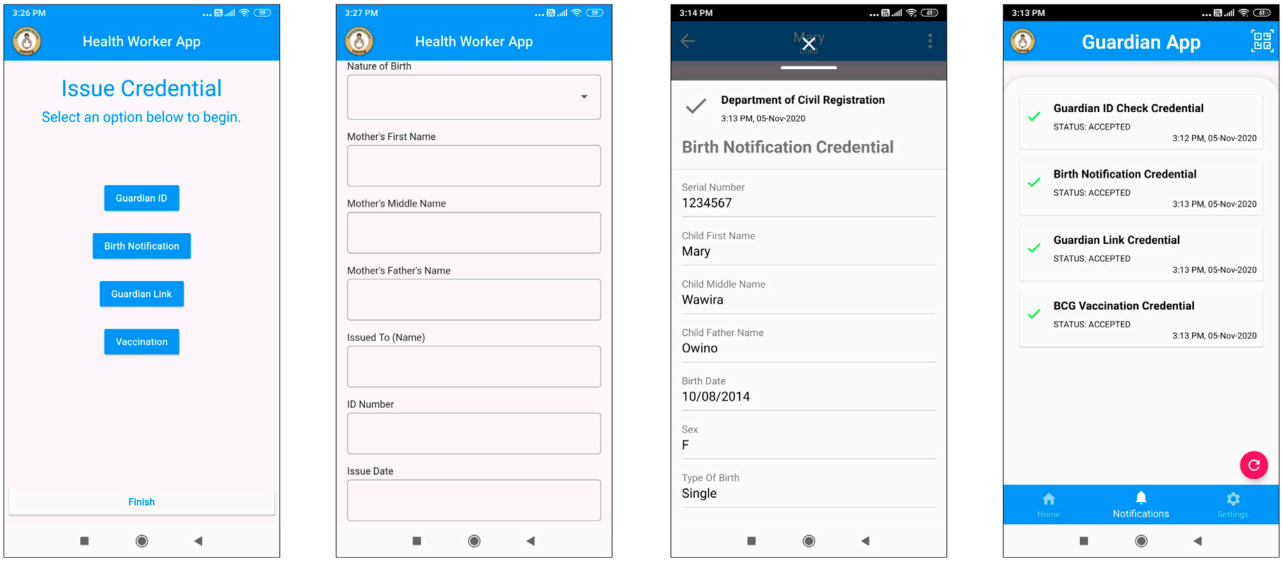


FIGURE 2 | Interaction flows between the participants in the system.

4 DISCUSSION

The goal of the project was to develop a prototype modeling the first steps of birth registration using SSI concepts that would allow the team to obtain user feedback from health workers and parents

[9] https://sovrin.org/library/guardianship-white-paper/



3:26 PM

Health Worker App

Issue Credential

Select an option below to begin.

Guardian ID

Birth Notification

Guardian Link

Vaccination

Finish

3:27 PM

Health Worker App

Nature of Birth

Mother's First Name

Mother's Middle Name

Mother's Father's Name

Issued To (Name)

ID Number

Issue Date

3:14 PM

Mary

Child

Department of Civil Registration
3:13 PM, 05-Nov-2020

Birth Notification Credential

Serial Number
1234567

Child First Name
Mary

Child Middle Name
Wawira

Child Father Name
Owino

Birth Date
10/08/2014

Sex
F

Type Of Birth
Single

3:13 PM

Guardian App

✓ Guardian ID Check Credential
STATUS: ACCEPTED
3:12 PM, 05-Nov-2020

✓ Birth Notification Credential
STATUS: ACCEPTED
3:13 PM, 05-Nov-2020

✓ Guardian Link Credential
STATUS: ACCEPTED
3:13 PM, 05-Nov-2020

✓ BCG Vaccination Credential
STATUS: ACCEPTED
3:13 PM, 05-Nov-2020

Home
Notifications
Settings

FIGURE 3 | Mobile applications for informants and guardians.

related to the feasibility and acceptability of an SSI-based technology approach for facility-based birth in Kenya.

Building on the Evernym SSI platform lowered the barrier to delivering an initial prototype capable of demonstrating a complete interaction between the issuer and holder of credentials. The project team also de-risked delivery, by adding a simulation mode to the prototype, which used representations of transitions between the applications to present pre-loaded credentials to the user of the guardian application. Ultimately, this has proven to be useful, as the COVID-19 pandemic resulted in an unexpected delay for the planned participatory design and user feedback activities in Kenya. The simulation mode in the applications will enable researchers to introduce the application and its operations to facility staff remotely, prior to any future facility-based user trials being conducted. The use of simulations for the SSI interactions was not considered prior to the project inception, but will be considered in future work, as this technique has been found to provide an efficient way of demonstrating concepts and interactions. The development process which has been utilized also allows for understanding of technology issues that, in turn, can guide future research, as described below.

4.1 SSI Within the Kenyan Healthcare Context

The prototype was developed as a standalone system, which enabled the team to focus on the SSI architecture and technology. Any practical future deployment will require integration into existing healthcare and civil registration technology systems for successful adoption, as shown in **Figure 4**, with a need for integration between the credentialing system and existing DHIS2 framework utilized by the Ministry of Health, as well as digital CRD systems. In this scenario, a mobile application would be used by informants to collect information about newborns from parents and guardians, which would be integrated into the centralized DHIS2 system. A digital copy of the BND would be issued to guardians, to maintain on their smartphone. The guardian would be able to use this digital copy of the BND when they needed to provide proof for any reason. A claimed benefit of holding the digital copy in a self-sovereign identity-based system would be that the guardian has control over access to their data, and can share it as needed. Selective disclosure means that the guardian

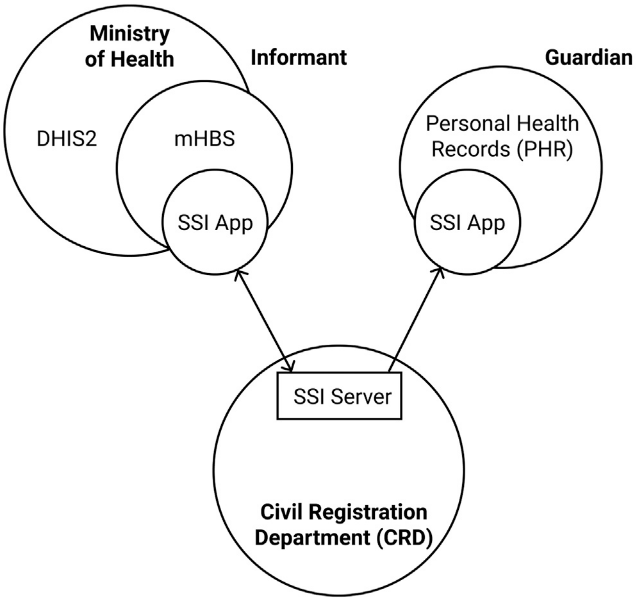


FIGURE 4 | The prototype in context as part of a larger infrastructure.

would, theoretically, have control over which parts of the digital document were shared with third parties. The decentralized architecture of the SSI system implies that presented credentials, when used in a SSI ecosystem, could be verified without the CRD having knowledge of the circumstances. This proposed architecture, however, does not eliminate the need for central databases held by government agencies, which can pose risks to human rights given lack of data protection frameworks (Dixon, 2017); however, it may provide greater privacy at the point of transactions.

The current prototype has two applications, one for the informant and one for the guardian, as shown in **Figure 3**. In a future deployment, the informant's application would need to be integrated with the existing mHBS application suite, which is already in use and familiar to the health worker community. User studies undertaken with the prototype system will help to identify the most effective point of integration with the mHBS app, such that it could be most efficiently adopted into the health worker's workflow.

The guardian's application was developed for use by parents and other guardians, and uses the guardian's own smartphone to receive and store credentials issued by the informant. This reduces the cost of deployment of the NeoLinkID prototype, but places a burden onto users to provide the necessary equipment, and to be comfortable installing the software onto their own devices. It is clear that this approach will not be applicable for those areas of Kenya that currently have lowest birth registration rates, however, it is possible that this approach could improve birth registration rates among smartphone owners who are a large and rapidly growing segment of the population (Rowntree and Shanahan, 2020). Typically in an SSI deployment this application is based upon the metaphor of a "wallet," and uses concepts of credentials and proofs. The NeoLinkID use case presents an opportunity to provide an application more suited to its target audience, and to research language and metaphors that could be readily adopted by that community. The future vision for this application is a personal health record design modeled on the paper Mother and Child Health Booklet currently used in Kenya. The World Health Organization (WHO) recommends the use of home-based records, as a complement to facility-based records, for the care of pregnant women, mothers, newborns, and children (World Health Organization, 2018). Adopting an SSI-based solution as part of nascent personal health records systems[10,11] has the potential to provide data security and privacy to parents and guardians, such that they are able to hold a cryptographically secured copy of their health data which can be used without involvement of a centralized authority.

4.2 Identity Verification and Authentication

In the NeoLinkID prototype, informants are required to check the identity of a guardian against existing identification, e.g., a photographic identity card. This identity source is referenced by its unique identifier in the "Identity Checked" credential, so that it can form a link between the credential and its holder, and can be used to provide proof that the credential is being legitimately presented. A credential is also issued to verify the birth of the newborn, based on the witness account of the informant. A further credential links the guardian and the newborn, and contains identifiers from both the guardian's identity credential and the newborn's credential, to form an inextricable connection between the two.

Authentication processes are needed to ensure that a verified identity is subsequently presented by the authorized person, i.e., that the presented identity matches the person. In a credential-based system, the verifying party needs to have assurance that credentials are being presented by the correct party. The prototype highlighted particular concerns about providing identity assurance where a phone is shared by multiple users. Any practical system would need to ensure that an application storing credentials is protected from unauthorized use with login protection, so it can only be accessed by the user of the phone that the "wallet" app belongs to, and provide assurance that the presenter of the credential is the authorized party.

For mother and baby pairs, authentication would be necessary to show that both the mother and the baby being presented match personal identifiers in their credentials. An architecture can be designed to include a biometric identifier within a credential, and used to authenticate the presenter of the credential (Hardman et al., 2019). This could be developed in a decentralized manner, without requiring a national or centralized biometric service, with the biometric template being used solely to verify the presenter of the credential. Formally, verification should be matched to identified assurance levels, as defined by national and/or international standards, though this may be complicated in some countries. The biometric mode must be carefully chosen with consideration for its specificity in newborns. The use of biometrics for newborns has not been widely studied and is a high priority research topic for this use case. The authors are exploring the possibility of applying iris recognition methods adapted to newborn eyes, due to demonstrated long-term stability of iris patterns. A slight drop in similarity between iris images as a function of time between enrollment and verification has been demonstrated by various research groups (Baker et al., 2009; Czajka, 2013). The impact of these observed time-related fluctuations on commercial iris recognition systems has been found, however, less important than other factors contributing to potential degradation of recognition reliability, as demonstrated by NIST in their IREX VI report (Grother et al., 2013). The research to fully understand iris aging is ongoing.

4.3 Interoperability and Open Source

Robust interoperability and open source development are critical elements of scalable health technology for LMICs[12]. It is an ongoing requirement for the SSI community at large

[10] https://news.vumc.org/2017/11/16/teams-mobile-app-helping-healthcare-workers-in-africa/

[11] https://www.muzima.org/

[12] https://id4d.worldbank.org/principles

which has led to development of open standards by W3C and non-profit organizations such as the Decentralized Identity Foundation and the Hyperledger Foundation, who are collectively bringing interoperability to the top of their agenda[13]. As well as open standards, open source implementations of these protocols are also becoming increasingly available[14]. Future NeoLinkID work will prioritize open standards, open source development, and interoperability of credentials as criteria for platform selection. Since mHBS and NeoLinkID are being incubated under the open source community of LibreHealth, there is an additional opportunity to leverage the expertize of the DHIS2 and OpenMRS developer communities to explore integration of SSI-based birth registration with both of these health IT platforms which have been adopted at the national level in Kenya.

5 CONCLUSION

The NeoLinkID project has successfully developed an SSI-based prototype modeling the first steps of birth registration based on the Kenyan process, using a personal health record approach to store information about newborns and their guardians. This development phase has provided some initial insights into the possibilities for improving data privacy, security, and portability, as well as the possible limitations of a decentralized approach to birth registration. The next phase of this work will be a feasibility, acceptability, and user design study, conducted virtually with Kenyan nurse-midwives. This paper represents an initial body of contextual knowledge for SSI-based birth registration in Kenya. SSI protocols continue to evolve and platforms and implementations are immature, presenting challenges for technology adoption outside of research and proof-of-concept deployments. Areas of high priority for research include authentication and verification, particularly integration of newborn biometrics, as well as integration with DHIS2 and the local framework for guardianship.

Increasing access to birth registration is a critical part of improving the survival and health of Kenyan newborns (Målqvist et al., 2008). Creating user-centered systems for birth registration that prioritize data protection and selective data sharing could mitigate some of the inherent risks and barriers of the birth registration process for segments of the Kenyan population. The experience of adapting SSI technology for this use case has illuminated multiple challenges that warrant further research for applying SSI to birth registration in Kenya and other LMIC settings. Effective solutions will require detailed understanding of the specific local context. Systems interoperability will be critical in order to derive benefits of this technology for increasing both birth registration rates and birth certificate acquisition, as well as the quality of data available for public health and vital statistics (Labrique et al., 2018; Wang and De Filippi, 2020). Research on SSI for birth registration should target the diverse local stakeholders whose collaboration will be essential for deployment success, including registration, justice, health, statistics, and civil society, as well as a broad representation of parent, family, and community stakeholders (AbouZahr et al., 2015). Investment in SSI-based birth registration solutions for LMICs should prioritize research to understand the unique needs and perspectives of all stakeholder groups, particularly those who are vulnerable to discrimination or exclusion based on their demographic, health, or political status.

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 authors.

AUTHOR CONTRIBUTIONS

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

FUNDING

In-kind support for prototype development was provided by The Evernym ID for Good Accelerator. Funding for engineering was provided by Google Summer of Code to LibreHealth (May–August 2020). LibreHealth provided generous in-kind hosting for the GSoC efforts. GS was supported in part by an IBM Faculty Award.

ACKNOWLEDGMENTS

We are grateful to the team at LibreHealth for incubating this project on their open source community platform and to Prajwal Belagavi for his dedicated development work through the Google Summer of Code program. Robby O’Connor, provided highly engaged and expert guidance for our team’s successful participation as NeoLinkID GSoC mentors. Dan Barclay developed the mobile web application for informants. We are also grateful for the technical support provided by the Evernym team throughout the participation in the accelerator. This work would not be possible without the bi-directional partnership that we are privileged to share with collaborators at Moi University, Alupe University College, and Moi Teaching and Referral Hospital, in Kenya. We are also grateful to the Helping Babies Survive Master Trainer Corps in Kenya, which, over the years, has provided invaluable feedback for concepts and technologies developed by the NeoInnovate Collaborative Consortium, including NeoLinkID.

[13] https://identity.foundation/interop/

[14] https://github.com/decentralized-identity

REFERENCES

AbouZahr, C., De Savigny, D., Mikkelsen, L., Setel, P. W., Lozano, R., Nichols, E., et al. (2015). Civil registration and vital statistics: progress in the data revolution for counting and accountability. *Lancet* 386, 1373–1385. doi:10.1016/S0140-6736(15)60173-8

Allen, C. (2016). The path to self-sovereign identity. Available at: http://www.lifewithalacrity.com/2016/04/the-path-to-self-sovereign-identity.html (Accessed April 24, 2016).

Aluvaala, J., and English, M. (2020). Implementing change for facility-based peripartum care in low-income and middle-income countries. *Lancet Global Health* 8, e980–e981. doi:10.1016/S2214-109X(20)30306-5

Apland, K., Blitz, B. K., Calabia, D., Fielder, M., Hamilton, C., Indika, N., et al. (2014). Technical Report. Birth registration and children's rights: a complex story. Available at: https://eprints.mdx.ac.uk/id/eprint/17346 (Accessed August 04, 2015).

Baker, S. E., Bowyer, K. W., and Flynn, P. J. (2009). "Empirical evidence for correct iris match score degradation with increased time-lapse between gallery and probe matches," in International conference on biometrics, Alghero, Italy, June 2–5, 2009 (Berlin, Heidelberg: Springer), 1170–1179.

Barclay, I., Freytsis, M., Bucher, S., Radha, S., Preece, A., and Taylor, I. (2020). Towards a modelling framework for self-sovereign identity systems. Preprint repository name [Preprint]. Available at: arXiv:2009.04327 (Accessed September 9, 2020).

Barker, C., Mulaki, A., Mwai, D., and Dutta, A. (2014). Devolution of healthcare in Kenya assessing county health system readiness in Kenya: a review of selected health inputs. *Facilities.* 16, 18. doi:10.13140/RG.2.2.36622.87363

Bucher, S. L., Cardellichio, P., Muinga, N., Patterson, J. K., Thukral, A., Deorari, A. K., et al. (2020). Digital health innovations, tools, and resources to support helping babies survive programs. *Pediatrics* 146, S165–S182. doi:10.1542/peds.2020-016915I

Czajka, A. (2013). Template ageing in iris recognition. *BioSignals* 1, 70–78. doi:10.5220/0004245800700078

Dixon, P. (2017). A failure to "do no harm"–India's aadhaar biometric id program and its inability to protect privacy in relation to measures in Europe and the us. *Health Technol* 7, 539–567. doi:10.1007/s12553-017-0202-6

Gelb, A. H., Anandan, V., and Cannata, A. (2016). "Identification for development (ID4D) country diagnostic: Kenya," in *Technical report*. Washington, DC: The World Bank, 1–82.

Gitobu, C., Gichangi, P., and Mwanda, W. (2018). The effect of Kenya's free maternal health care policy on the utilization of health facility delivery services and maternal and neonatal mortality in public health facilities. *BMC Pregnancy Childbirth* 18, 77. doi:10.1186/s12884-018-1708-2

Grother, P. J., Matey, J. R., Tabassi, E., Quinn, G. W., and Chumakov, M. (2013). Tech. rep. IREX VI-Temporal stability of iris recognition accuracy. Available at: https://nvlpubs.nist.gov/nistpubs/ir/2013/NIST.IR.7948.pdf (Accessed July 11, 2013).

GSM Association (2013). Mobile birth registration in sub-saharan africa a case study of orange Senegal and Uganda telecom solutions. Available at: https://www.gsma.com/identity/resources/mobile-birth-registration-in-sub-saharan-africa-a-case-study-of-orange-senegal-and-uganda-telecom-solutions (Accessed July 2, 2013).

GSM Association (2020). Mobile taxation in Kenya: accelerating digital development. Available at: https://www.gsma.com/publicpolicy/resources/mobile-taxation-in-kenya-accelerating-digital-development (Accessed March 23, 2020).

Hardman, D., Harchandani, L., Othman, A., and Callahan, J. (2019). Using biometrics to fight credential fraud. *IEEE Commun. Stand. Mag.* 3, 39–45. doi:10.1109/mcomstd.001.1900033

Hereward, M., Williams, C., Petrowski, N., and Cappa, C. (2019). Universal birth registration by 2030: progress and challenges. *Lancet* 394, 2211–2212. doi:10.1016/s0140-6736(19)33101-0

Hug, L., Alexander, M., You, D., Alkema, L., and For Child, U. I.-A. G. (2019). National, regional, and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. *Lancet Global Health* 7, e710–e720. doi:10.1016/S2214-109X(19)30163-9

India, H. (2013). Technical Report. Systematic review of eCRVS and mCRVS interventions in low and middle income countries. Available at: https://www.who.int/publications/i/item/systematic-review-of-ecrvs-and-mcrvsinterventions-in-low-and-middle-income-countries (Accessed October 19, 2013).

Kenya Human Rights Commission (2019). Report of the digital identification document (id) and citizenship consultative meeting. Available at: https://www.khrc.or.ke/mobile-publications/equality-and-anti-discrimination/198-report-of-digital-identification-citizenship-workshop-naivasha/file.html (Accessed January 18, 2019).

Kondova, G., and Erbguth, J. (2020). "Self-sovereign identity on public blockchains and the gdpr," in Proceedings of the 35th annual ACM symposium on applied computing, Brno, Czech Republic, March–April 30–3, 2020 (New York, USA: SAC) 342–345.

Kunkel, M., Marete, I., Cheng, E. R., Bucher, S., Liechty, E., Esamai, F., et al. (2019). "Place of delivery and perinatal mortality in Kenya," in *Seminars in perinatology*. Editor E. D. Mary (New York, NY: Elsevier), 43, 252–259.

Kurth, A., Inwani, I., Agot, K., Macharia, P., and Buttolph, J. (2015). mHEALTH compendium volume 5. Available at: http://www.africanstrategies4health.org/uploads/1/3/5/3/13538666/mhealthvol5_final_15jun15_webv.pdf (Accessed November 20, 2016).

Labrique, A. B., Wadhwani, C., Williams, K. A., Lamptey, P., Hesp, C., Luk, R., et al. (2018). Best practices in scaling digital health in low and middle income countries. *Glob. Health* 14, 103. doi:10.1186/s12992-018-0424-z

Målqvist, M., Eriksson, L., Nga, N. T., Fagerland, L. I., Hoa, D. P., Wallin, L., et al. (2008). Unreported births and deaths, a severe obstacle for improved neonatal survival in low-income countries; a population based study. *BMC Int. Health Human Rights* 8, 4. doi:10.1186/1472-698x-8-4

Manya, A., Braa, J., Øverland, L., Titlestad, O., Mumo, J., and Nzioka, C. (2012). "National roll out of district health information software (dhis 2) in Kenya, 2011–central server and cloud based infrastructure," in IST-Africa 2012 conference proceedings, Dares, Salaam, May 9–11, 2012 (Dublin, Ireland: International Information Management Corporation), 1–9.

MEASURE Evaluation (2014). Technical Report 14–110. County civil registration and vital statistics stakeholder forums: strengthening civil registration systems at the county level. Available at: https://www.measureevaluation.org/resources/publications/ (Accessed October, 2014).

Njoroge, M., Zurovac, D., Ogara, E. A., Chuma, J., and Kirigia, D. (2017). Assessing the feasibility of ehealth and mhealth: a systematic review and analysis of initiatives implemented in Kenya. *BMC Res. Notes* 10, 1–11. doi:10.1186/s13104-017-2416-0

Pelowski, M., Wamai, R. G., Wangombe, J., Nyakundi, H., Oduwo, G. O., Ngugi, B. K., et al. (2015). Why don't you register your child? a study of attitudes and factors affecting birth registration in Kenya, and policy suggestions. *J. Dev. Stud.* 51, 881–904. doi:10.1080/00220388.2015.1010156

Preneel, B. (1994). Cryptographic hash functions. *Eur. Trans. Telecommun* 5, 431–448.

Privacy International (2018). The sustainable development goals, identity, and privacy: does their implementation risk human rights? Available at: https://privacyinternational.org/long-read/2237/sustainable-development-goals-identity-and-privacy-does-their-implementation-risk (Accessed August 29, 2018).

Rowntree, O., and Shanahan, M. (2020). The mobile gender gap report 2020. Available at: https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2020/05/GSMA-The-Mobile-Gender-Gap-Report-2020.pdf (Accessed March 5, 2020).

Sahay, S., Nielsen, P., and Saebo, J. (2013). Systematic review of eCRVS and mCRVS interventions in low and middle income countries. Available at: https://webcache.googleusercontent.com/ (Accessed October 19, 2013).

Schoemaker, E., Kirk, T., and Rutenberg, I. (2019). *Caribou digital, Kenya's identity ecosystem*. Farnham, Surrey: Caribou Digital Publishing, 1–60.

Selim, L. (2019). What is birth registration and why does it matter? Available at: https://www.unicef.org/stories/what-birth-registration-and-why-does-it-matter (Accessed December 10, 2019).

Solberg, E. (2015). From mdgs to sdgs the political value of common global goals. *Harv. Int. Rev.* 37, 58. doi:10.4045/tidsskr.14.1445

Sondaal, S. F. V., Browne, J. L., Amoakoh-Coleman, M., Borgstein, A., Miltenburg, A. S., Verwijs, M., et al. (2016). Assessing the effect of mhealth interventions in improving maternal and neonatal care in low- and middle-income countries: a systematic review. *PloS One* 11, e0154664. doi:10.1371/journal.pone.0154664

Sporny, M., Noble, G., Longley, D., Burnett, D. C., and Zundel, B. (2019). Verifiable credentials data model. Available at: https://www.w3.org/TR/vc-data-model/ (Accessed March 1, 2019).

UN General Assembly (2015). Transforming our world: the 2030 Agenda for sustainable development (UN). Available at: https://sdgs.un.org/2030agenda (Accessed November 25–27, 2015).

UNHCR (2017). Ensuring birth registration for the prevention of statelessness. Available at: https://www.refworld.org/pdfid/5a0ac8f94.pdf (Accessed November 02, 2017).

Wang, F., and De Filippi, P. (2020). Self-sovereign identity in a globalized world: credentials-based identity systems as a driver for economic inclusion. *Front Blockchain* 2, 28. doi:10.3389/fbloc.2019.00028

World Bank (2018). Principles on identification for sustainable development : toward the digital age. Available at: https://id4d.worldbank.org/principles (Accessed March 15, 2017).

World Health Organization (2013). Move it: report on monitoring of vital events using information technology. Available at: https://www.who.int/publications-detail-redirect/move-it-report-on-monitoring-of-vital-events-using-information-technology (Accessed October 20, 2013).

World Health Organization (2018). WHO recommendations on home-based records for maternal, newborn and child health. Available at: https://www.who.int/publications/i/item/9789241550352 (Accessed January 1, 2018).

Yu, E. (2011). Modeling strategic relationships for process reengineering. *Social Model. Requir. Eng.* 11, 66–87. doi:10.7551/mitpress/7549.003.0005

frontiers in Blockchain

ORIGINAL RESEARCH
published: 02 March 2021
doi: 10.3389/fbloc.2020.609101

An Accessible Interface Layer for Self-Sovereign Identity

*Mick Lockwood**

University of Salford, Salford, United Kingdom

The mechanisms and evolving standards collectively known as self-sovereign identity (SSI) offer the prospect of a decentralized Internet by providing a central pillar for a human-centered data ecosystem (HCDE). Once established this technology promises to afford participants the same agency in the digital realm as individuals experience in the real world. Investigation suggests that the domain is now sufficiently mature to realize practically the principles of SSI, but in order to achieve sustainable adoption, significant design focused work needs to be undertaken at the interface layer. This paper presents recent practice-led research designed to project current SSI prototypes to scale through conceptual modeling, preliminary user interface, and critical analysis. This research introduces the term sovereign boundary mechanism (SBM), a standardized collection of SSI interactions, which can be described as a metaphorical ring of sovereignty between the participant and the wider network. Within this model, participants control identity, relationships, and data streams and access control. This research identifies the domains of interaction and the minimum required objects for a full-scale SSI engagement through an SBM. It defines the component parts and functionality of a wider HCDE which require further consideration, and it identifies emergent concepts for which a participant may lack mental model and understanding. The research considers human computer interaction (HCI) theory across internalized, external, and distributed cognition, arguing that the current trajectory of SSI requires significant internalized representations, prior knowledge, and participant responsibility. This research argues that these elements are problematic and pose a significant barrier to sustainable adoption. In conclusion, this research suggests that the decentralized community needs to recognize the obstacle potentially posed at the interface layer and engage in collective standardization, strategy, and design thinking to increase the probability of sustainable adoption.

Keywords: self-sovereign identity, human data interaction, human-centered data ecosystem, sovereign boundary mechanism, decentralized internet, interface layer, usability

OPEN ACCESS

Edited by:
Kaliya Young,
Merritt College, United States

Reviewed by:
Larry C. Bates,
AltMarket, United States
Tony Fish,
AMF Ventures, United Kingdom

***Correspondence:**
Mick Lockwood
m.g.lockwood@salford.ac.uk

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

Received: *22 September 2020*
Accepted: *24 December 2020*
Published: *02 March 2021*

Citation:
Lockwood M (2021) An Accessible Interface Layer for Self-Sovereign Identity.
Front. Blockchain 3:609101.
doi: 10.3389/fbloc.2020.609101

INTRODUCTION

The management of digital identities and personal data represents a formidable challenge for the 21st century; issues of privacy, inference profiling, surveillance capitalism, GAFA monopoly, democratic interference, and the lost opportunities of big data are significant. Many envisage a decentralized alternative to the centralized network, one that places the human at the center of data streams, facilitating transparency, agency, and negotiated access to personal data. Self-sovereign identity is a collection of concepts and standards that promises to emancipate the everyday user from the asymmetrical relationships observed across today's Internet. By establishing and controlling persistent digital identities and exchanging digital credentials, relationships and trust networks can be established. The decentralized nature of the technology and the subsequent requirement for the individual to manage and protect what is a complex decentralized key management system pose several interactive design challenges. This paradigm shift introduces unfamiliar concepts and raises issues of missing metaphor and mental model. If SSI is to become accessible and sustainable, we must investigate and understand SSI applications and case studies and interactions and identify the foundation design patterns needed to manage and transact personal data through digital identity.

This paper presents what is part one of a two-part publication of outcomes from doctoral research conducted between 2017 and 2020. The research poses a primary question: can a sustainable technology be established to allow for individual agency within a decentralized Internet? Two additional questions were then derived. The first considered usability at the interface layer and asks: can an interface layer for a decentralized Internet be designed to allow for accessible interaction? And the second considered value proposition and adoption and asks: how might a decentralized Internet provide value, emerge, and be adopted? As this work progressed, it became evident that the trajectory toward a decentralized Internet would require the development of a human-centered data ecosystem (HCDE): a central pillar of which could be provided by the emergent domain of self-sovereign identity (SSI). This paper presents the academic framework, method, findings, and recommendations relating to the investigation of usability and accessibility at the interface layer for SSI. A second paper presenting the findings relating to value proposition and adoption can also be found within this journal.

PRE-ASSUMPTIONS

It is not the intention of this research to advocate privacy concern or lobby for the adoption of sovereign identity decentralized technologies. The aim of this research is to consider and reflect the current proposals for SSI interactions and to extend current prototypes to scale in the context of the defined principles, required mechanisms, and evolving standards. The objective has been to enter the problem space as a designer, extend its current position, and then reflect on the outcome. That said, there are a number of accepted arguments on which this work is constructed. It is accepted that the broad concept of network centralization means giving up control of our personal information and identity (Moglen, 2013; Van Kleek and O'Hara, 2014) that the advent of the centralized model poses a significant threat to our collective and individual privacy (Solove, 2008) and that personal data are now exploited by capital in order to leverage influence over our daily lives (Schneier, 2015; Zuboff, 2015). It is accepted that there is a decentralized alternative to the centralized model which will offer a participant greater agency over elements of their personal data (Haddadi, 2015; Hornung et al., 2015; Mortier, 2014) that identity can be established through the ownership and control of personal data (IIW, 2019) and that an identity layer is integral to a sovereign engagement with the wider network and evolving Internet (Cameron, 2005; Allen, 2016; Tobin and Reed, 2016). There is an awareness of what could be considered to be a moderate bias, and every conscious effort has been made to prevent it from influencing the design of this research, the methods of data gathering, and the analysis and interpretation of results.

THEORETICAL FRAMEWORK

When exploring the interface layer for SSI, this research considers the wider academic context. The domain of human data interaction (HDI) (Mortier, 2014; Haddadi, 2015) recognizes the pervasiveness of computing in our data driven society. The theory argues that human computer interaction (HCI) has traditionally focused on interactions between humans and computers as artifacts. However, with the rapid evolution of human's interacting predominantly with data, a different academic perspective is required. Moritier (2014) defines HDI as "placing the human at the center of the flows of data, providing mechanisms for citizens to interact with these systems and data explicitly" (p. 1). The concepts of HDI illustrate the opaque mechanisms used to process personal data and the hidden inferences and subsequent feedback loops. This theory argues that a user requires legibility to understand the ambient ways in which data are processed and utilized, that agency is required to control, manage, and permit access to personal data, and that users require a means to negotiate the terms under which their data can be used. The concept of SSI provides the mechanisms required fulfilling the principles laid down by HDI and as such HDI can act as a suitable overarching academic domain.

As the interface layer for SSI interactions is investigated, it becomes evident that the paradigms of cognitive theory found within human computer interaction (HCI) are highly relevant (Harrison et al., 2007; Rogers, 2012). The current trajectory within SSI suggests that participants will manage their affairs independently through a digital wallet, engaging what is coined within this research as a sovereign boundary mechanism (SBM). These are the core interactions enabled through SSI, found at the center of a wider human-centered data ecosystem. An SBM represents a standardized set of concepts, tools, and user representations that allows for an interaction with a decentralized Internet, through a HCDE with SSI at its core.

Figure 1 illustrates the component parts of a sovereign boundary mechanism. The inner core consists of identity,

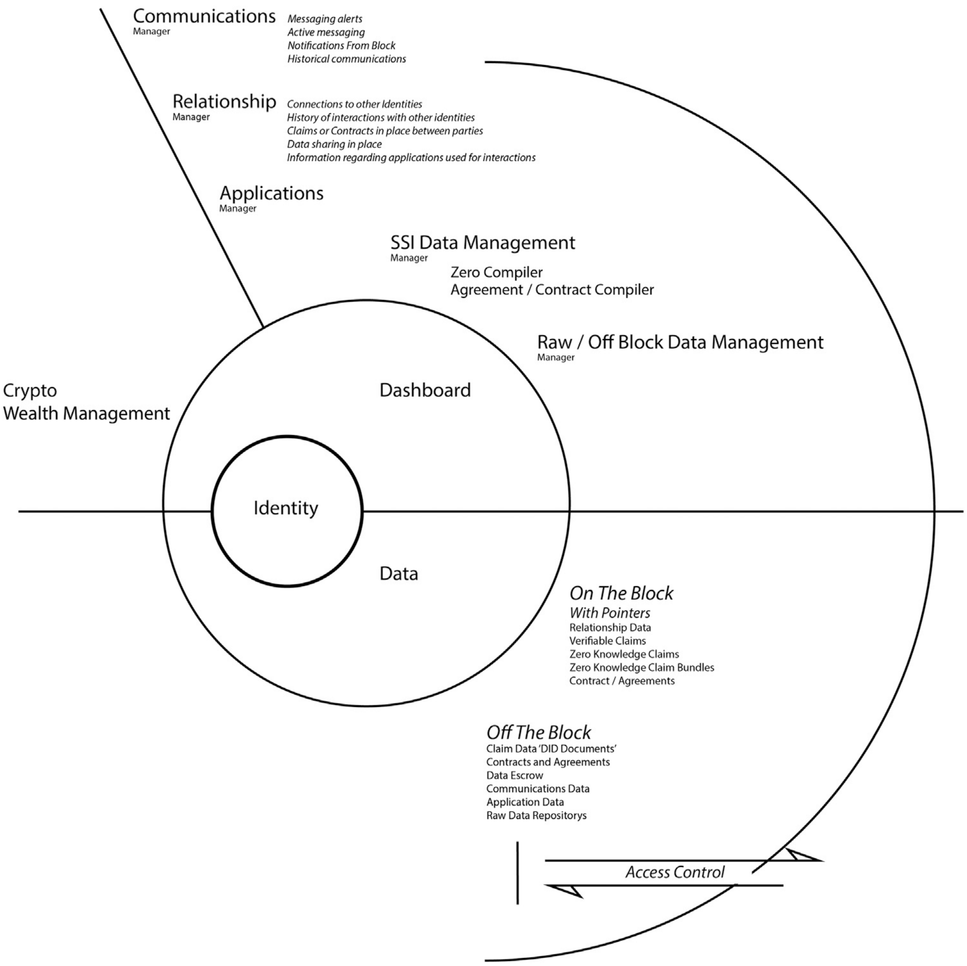


FIGURE 1 | The component parts of a sovereign boundary mechanism.

access to core functionality, and personal data. The second shell consists of the accessible applications and mechanisms required to manage personal identity, credentials, data, and contracts, together with the management of data storage and access control. The outer ring depicts the sovereign boundary between the participant and the wider network.

The SBM model represents a standardized collection of interactions that the participant independently controls. This is a strict task-based engagement that arguably presents new distinct ideas and original concepts. The notion of sovereignty at the core of this model equates to individualism and given the gravitas of the personal data being transacted, the assumption can be made that a strong internalized understanding of the domain will be required to engage initially and sustain participants. This research recognizes the value of the classical HCI theory of internalized cognition (Craik, 1943; Norman, 1986; Kirsh, 1997; Johnson-Laird, 2001), considering participant mental model, internalized understanding, and processing of information at both the interface layer and within the wider data ecosystem. There are long standing arguments that challenge the validity and value of specifically internalized cognitive psychology in HCI to inform the design of computer systems and interface (Carroll, 1991; Landuer, 1991). This research suggests that internalized HCI cognitive theory is still an important consideration, not in the pursuit of an overarching set of design rules for the development of general interactive experiences, but as a framework to map and design the required internalized understanding and core functionality of a specific and standardized domain. This position is supported by Payne (2003) who argues that internalized cognitive processes and mental model are valid when considered in a specific context.

As an SBM engagement progresses outwards toward transactions with the wider network, the required internalized cognitive understanding and processing of the core concepts migrate into the functional management of identity, relationships, credentials, and data across a spectrum of scenarios. When considering the design of the tools and mechanisms to facilitate this, the theories of externalized cognition are important. The notion of external cognition centers around the argument that "when individuals are solving problems, human beings use both internal representations stored in their brains and external representations, recorded on a paper, on a blackboard, or on some other medium" (Larkin and Simon, 1987, p. 66). Scaife and Rogers (1996) describe computational offloading, the way in which external representations change the amount of cognitive effort needed to carry out a task. They go on to describe graphical constraining, arguing that external representations can be designed in such a way as to limit the possible inferences that can be made in completing a task, reducing the load on internal memory allowing more space to plan the next move. There is the notion of re-representation, a consideration of how different external representations with the same abstract structure make the solving of problems easier or more difficult. They argue for the theory of cognitive tracing that an external representation should be interactive and that a user should be able to mark and annotate to aid understanding and to build external memory (Scaife and Rogers, 1996). O'Malley and Draper (1992) offer interesting arguments through which to consider the interplay between internalized and externalized representations when interacting with systems. They suggest that the internalized representation is not only knowledge of a systems function but also knowledge of where to look and how to find further information through the externalized representations.

It is the interplay between internalized and externalized cognition that needs to be understood and managed in the development of an interface layer for SSI, in what is arguably a significant paradigm shift in network engagement that by its very nature requires a degree of internalization. It is the balancing of these cognitive processes within the context of adoption theory (Rogers, 1962) and the technology life cycle (Moore, 1991) that will determine the probability of achieving accessible, sustainable technologies and tools for SSI within a wider human-centered data ecosystem.

When contemplating the cognitive HCI theory relative to SSI and a sovereign boundary mechanism, it is important to consider the interactions within the wider ecosystem. A human-centered data ecosystem has the potential to comprise a spectrum of participants, both human and machine, in a number of environments both virtual and physical. Users may engage with applications in emergent ways, as individuals or as a collective. This spectrum of activity and interaction involves interpretation and utilization of mechanisms relative to context, so the theories of situated action need to be considered within the equation (Suchman, 1987). As interaction across the wider network will include group and collaborative coordination, the theory of distributed cognition also needs to be considered (Hutchins, 2000).

This research argues that an accessible model of interaction for an SSI driven HCDE will contain internalized, externalized, and distributed cognitive processes and that understanding, mapping, and optimizing the relationships between these elements are critical for a sustainable, functional interface layer. This utilization of a range of cognitive theories across a border interaction utilizes what Harrison et al. (2007) describes as a phenomenological matrix, a collection of HCI theory applied pragmatically where and when it is required.

SELF-SOVEREIGN IDENTITY

The concept of self-sovereign identity solves one of the most challenging problems facing the Internet: the capability to establish, own, and control a persistent verifiable identity. SSI can be defined as the following: "a digital identity that is owned and controlled by an individual, company, or machine that has no reliance on any centralized authority. The identity is persistent and can never be taken from its owner. The identity is part of a wider ecosystem, where relationships can be built, trust can be developed, and identity attributes and data can be exchanged under the complete control of the sovereign identity." The concept of a wider ecosystem is important, as once a user can control an identity, the concepts of SSI then allow the user to establish independent unique relationships and private communication channels with peers across the network. It allows for the requesting, issuing, and distribution of verifiable credentials. The ecosystem allows for the development of trust networks that are judged appropriate, dependent on a given situation. The realization of SSI is considered to be Web 3.0 and to many is inevitable (Tobin and Reed, 2016). The core principles of SSI can be traced back to the work of Kim Cameron in his Laws of Identity (Cameron, 2005). These principles of digital identity were then evolved further in the context of SSI by Christopher Allen (Allen, 2016). Allen defined a number of principles that need to be satisfied in order for a technology to be considered self-sovereign.

- *Existence.* Users must have an independent existence.
- *Control.* Users must control their identities.
- *Access.* Users must have access to their own data.
- *Transparency.* Systems and algorithms must be transparent.
- *Persistence.* Identities must be long-lived.
- *Portability.* Information and services about identity must be transportable.
- *Interoperability.* Identities should be as widely usable as possible.

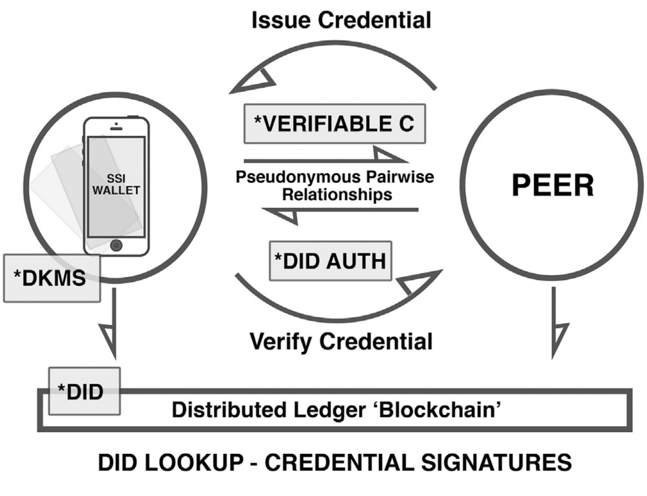

Issue Credential
*VERIFIABLE C
Pseudonymous Pairwise Relationships
SSI WALLET
PEER
*DKMS
*DID AUTH
Verify Credential
*DID
Distributed Ledger 'Blockchain'
DID LOOKUP - CREDENTIAL SIGNATURES

FIGURE 2 | The component parts of an SSI interaction.

- *Consent.* Users must agree to the use of their identity.
- *Minimalization.* Disclosure of claims must be minimized.
- *Protection.* The rights of users must be protected.

The guiding principles defined by Christopher Allan together with those established by the field of human data interaction (HDI) (Mortier 2014; Haddadi, 2015) have been considered alongside the wider objective of the decentralized movement and evolving technologies and standards to define the high-level functionality and interactions required at the SSI interface layer.

Figure 2 illustrates the component parts of a self-sovereign identity ecosystem, establishing relationships and engagement between two sovereign peers through a pseudonymous pairwise relationship, the exchange and authentication of credentials, the writing and retrieval of digital identifiers, and credential signatures from a dedicated blockchain. When contemplating an interface layer for SSI, it is important to consider it functioning at scale; **Figure 2** defines the core mechanics of SSI, but when developing this model within a sovereign boundary mechanism and indeed a wider human-centered data ecosystem, a practical engagement will require additional, peripheral functions and interactions.

THE NEED TO BUILD A CONCEPTUAL MODEL

Insight into usability issues surrounding personal data protection for the everyday participant can be found in previous research. The paper entitled Why Johnny Can't Encrypt (Whitten and Tygar, 2005) investigates email encryption software and argues that despite a well-designed interface, a lack of fundamental mental model and understanding prevents Johnny from successfully encrypting his communications. The research suggests that new methods of testing are required when developing interactions that contain novel concepts and ideas. It could be argued that this work draws striking contemporary resemblance to the design space surrounding SSI. Both involving a complex collection of novel interactions, an unfamiliar digital environment, and the transaction of information with a high personal value. When investigating the current decentralized domain, there is a sense that a technically focused development community is overlooking the inevitable user experience and accessibility issues. There is evidence supporting this assumption. In the research conducted by Dunphy and Petitcolas (2018) entitled A First Look at Identity Management Schemes on the Blockchain, the principles of Kim Cameron's Laws of Identity (Cameron, 2005) are used to evaluate several SSI projects. In conclusion, it is argued that none of the projects currently satisfy Cameron's 6th law, Human Integration, and in summary state that "there is a noticeable lack of contextual understanding relating to the user experience elements of the schemes we encountered" (Dunphy and Petitcolas, 2018). Developers of SSI applications need to consider the fundamental principles of the diffusion of innovation (Rogers, 1962) which recognizes the requirement for a participant to understand a new product offering, its function, and the value found within its concepts; this needs to be considered in the context of the persuasion of a new user when making a decision to adopt or reject a novel innovation.

This research has concluded that the technical components are now in place to build a functional human-centered data ecosystem with a central component of SSI at its core. Developing an accessible interface layer for such an ecosystem is now a design problem, one which needs to balance the cognitive load required for engagement, with the value proposition decentralized tools and services offer. When we consider the cognitive load, we need to understand the model of interaction, its component parts, and the journey users take to achieve their aims and objectives. We can then begin to map the required user understanding and scaffold for a mental model. The evolving structure can then be interrogated to examine the frictions, while iteratively evolving and improving the design. The first stage in this investigation is to establish a conceptual model (Johnson and Henderson, 2002), the details of which are communicated in the following section.

METHODS

The following section describes the methods employed to investigate the interface layer for SSI through practice-led research. This is a component part of a larger mixed methods design (Creswell, 2003) influenced by Design Theory for Mixed Methods in HCI (Turnhout, 2014). Prior to the design of this component, the pertinent literature was considered, and an investigation of historical and contemporary decentralized artifacts was undertaken. Semi-structured interviews with experts from the decentralized field were conducted. Also, interviewed were practitioners from the domain of user experience and user interface design. **Figure 3B** illustrates how this work sits within a wider research study.

Research Phases

This component of research was undertaken in 4 phases:

Phase One: Defining a Conceptual Model

The current position of SSI comprises of a clear description of principles, defined concepts, developing standards, and preliminary prototypes. The objective of this practice-led component is to extend this current model to a scalable analog interaction and then critically analyze the result. As such, phase one engaged a conceptual modeling method (Johnson and Henderson, 2002) which suggests that practitioners should "begin by designing what to design" (p. 1). The method describes a process which results in a structured text and table-based outcome. The first step is to define the application's purpose and high-level functionality. Once this is established, the process continues with the definition of the major concepts and vocabulary. The next stage considers the conceptual objects visible to the user through what is termed as an objects and operations analysis. This process investigates the objects users manipulate, their attributes and operations, and the relationships between them. The method then progresses to a task-

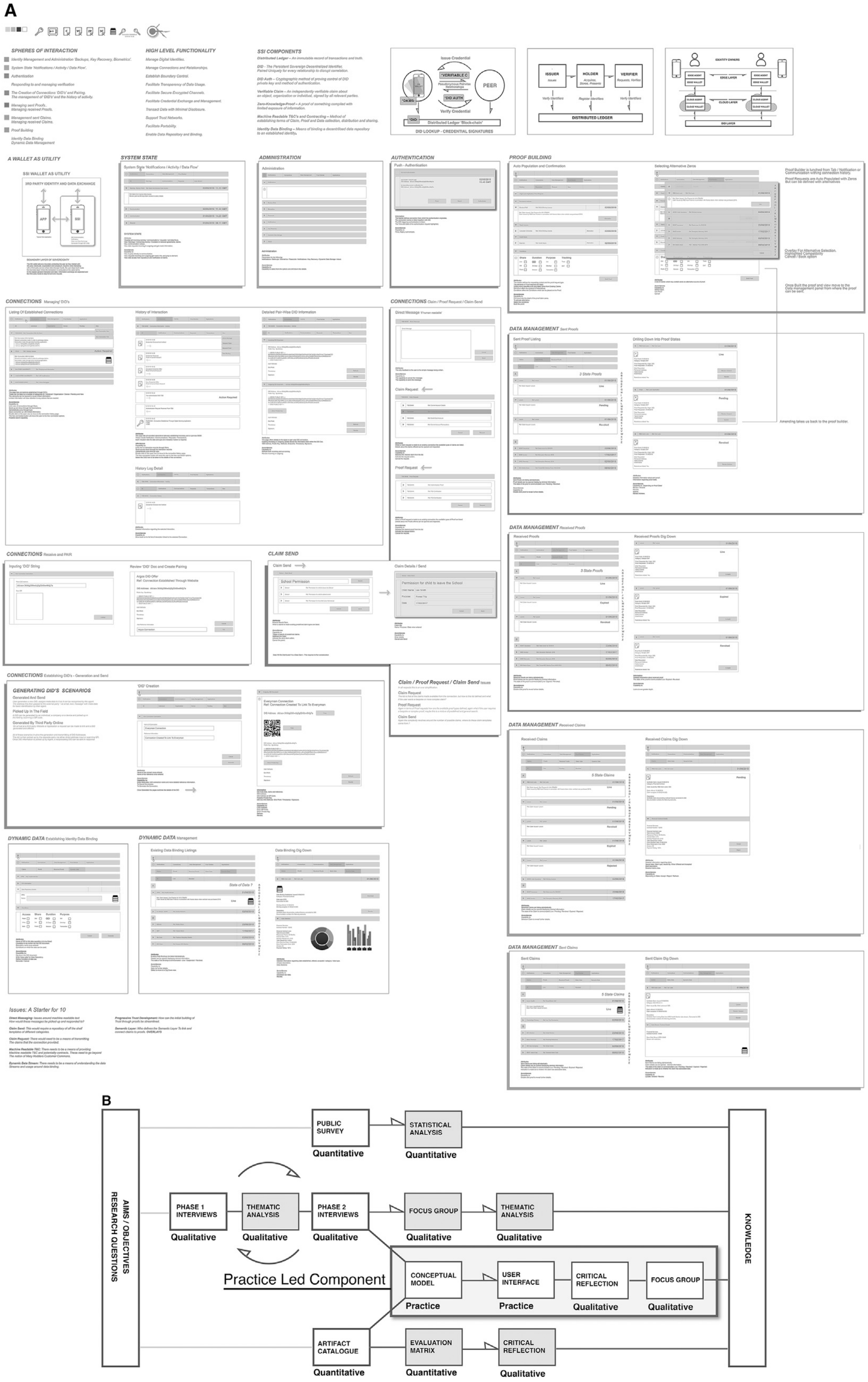


FIGURE 3 | (A) Initial User Interface and **(B)** All component parts of the wider research study. The highlighted section represents the practice-led component described within this paper.

to-tool mapping exercise. This considers how a participant uses an application to carry out tasks engaging the visible objects, attributes, and operations. Within this practice, the development of a preliminary user interface reflects this stage.

An important understanding within the conceptual modeling process is that it is agile and flexible. As the designer progresses to develop a user interface and evaluation methods are employed (Nielsen, 2005), the conceptual model is continually updated and refined.

Phase Two: Developing a Preliminary User Interface

With a defined conceptual model in place, the method progressed to establish a preliminary user interface in a wireframe format. The purpose of the wireframe is to visualize the defined objects and their relationships, while developing the minimal interface touch points to enable the required interactions. The guiding principle in developing the user interface is simplicity. There are no radical design choices and the process utilizes existing table-based interface design patterns.

Phase Three: Critical Reflection

Once a conceptual model and preliminary user interface had been established, a critical reflection was undertaken by the researcher. This was consciously conducted from the standpoint of the user. A user journey was considered across the full range of functionality at an introductory level. Efforts were made to uncover significant interactive friction, while identifying original concepts, metaphor, and the participant mental model. In addition, the critical reflection identifies elements of a wider human-centered data ecosystem, identifying missing components which may be required for a full spectrum of interaction.

Phase Four: Evaluation Focus Group

Once established the conceptual model and preliminary user interface can be used to communicate the concept of the domain, its components, the required objects, and the scope and scale of the required user interaction. The developed model was subjected to an additional critical analysis, this time through consultation with usability experts and interface designers. The objective of this evaluation was to consider a first cycle of development, investigate the integrity of the fundamental concepts, and provide an unbiased perspective from a practitioner's standpoint. In order to facilitate this evaluation, a workshop was designed incorporating a focus group. Guidance was taken from Finch and Lewis (2003) in its design and planning. Participants were selected locally, in the Manchester United Kingdom vicinity, for their reputation, experience, and expertise. A website was authored together with **Supporting Materials** which were then distributed to participants in advance of the workshop. The workshop began with a detailed presentation of the research, the concept of the decentralized Internet, and its principles and objectives. An overview was given of existing technologies, concepts, and standards. A detailed explanation was given regarding the practice-led component of research, including the conceptual model and initial UI. Following the presentation and Q and A, a focus group was conducted based on the following topics of discussion:

- Participant Understanding
- Views on the Complexity of the Interactions
- Mental Model and Metaphor
- Building Accessibility
- The Potential for Automation

RESULTS

The Conceptual Model

The following section presents results for each stage of the conceptual modeling process.

Purpose and High-Level Functionality

The following list defines the purpose and the high-level functionality of a standardized sovereign boundary mechanism, considered in the context of a wider human-centered data ecosystem, the principles of human data interaction and self-sovereign identity, and the evolving technologies and standards.

- Manage digital identities
- Manage connections and relationships
- Establish boundary control
- Facilitate transparency of data usage
- Facilitate secure encrypted channels
- Facilitate credential exchange and management
- Transact data with minimal disclosure
- Support trust networks
- Facilitate portability
- Enable data repository and binding

SSI Case Studies and User Scenarios

Table 1 presents an overview of case studies and user scenarios defined as part of the conceptual modeling process. Each has been explored through the development of a user journey and reflects core interactions which might be found in a broader human-centered data ecosystem.

Objects and Operations Analysis

The objects and operations analysis requires the identification of the conceptual objects that a participant can see and manipulate. Once the objects are identified, their attributes are specified and listed alongside their relationships to one another. As the conceptual model develops, outstanding and resolved issues are also listed. Currently, 27 objects have been defined and a brief overview of each is listed in **Table 2**.

Initial UI Development

With an objects and operations analysis defined, a full initial user interface has been developed. A low-resolution image of which can be seen within **Figure 3A**. This initial user interface is a wireframe utilizing a simple table-based layout. A panel for each of the identified objects has been created, supplemented by a description of the required attributes and functions. To aid in the reading of the wireframe, a colour key has been provided highlighting the identified spheres of interaction; these relate directly to the objects and operations analysis. A PDF version of the wireframe can be found as **Supplemental Material** of this paper. INITIAL_SSI_UI_DEV_LOCKWOOD__2020.pdf.

Critical Reflection

The conceptual model suggests that a minimum of 27 objects are required for a functional sovereign boundary mechanism. When observing the resulting UI, it is evident that this is complex, and in terms of a cognitive load, it can be debated that it is significant. The user experience requires a clear interpretation and understanding of the system state, within which the interactions are reliant on complex sequences. There is a number of sub-domains within the system and a general understanding of the majority of objects and their relationships would be necessary in order to enable confident engagement. In subsequent phases of development, efforts can be made to reduce complexity, and many of the system processes might be reordered, automated, or streamlined following usability testing. It remains, however, that a human-centered data ecosystem through self-sovereign identity, offering full agency through a sovereign boundary mechanism, presents a collection of original concepts and interactions, which may prove challenging in the context of initial and sustained adoption, interactive friction, and participant mental model.

This research suggests, based on the evidence derived from the conceptual model and subsequent UI, together with the defined principles of HDI and SSI, coupled with the value of personal data, that a sovereign boundary mechanism requires a considerable amount of internalized understanding before meaningful interaction can be achieved. The introduction of such a system introduces considerable friction and is a backwards step in the context of contemporary user interaction design. The notion of sovereignty and independence places a great weight of responsibility on the participant, which potentially results in what is discussed later in this paper as the paradox of the sovereign boundary mechanism.

It is important to separate what can be argued to be a high friction demanding user interface experience and the internalized knowledge and understanding of system and concepts that will be required in order to engage. Critical reflection suggests that the dominant issue in any future development of an analog self-sovereign identity system is not the physical interface design. Many office management tasks, media editing tools, and social networks require engagement with complex UI structures and interactions. This research concludes that in this context, the dominant issue is that of the understanding of concepts and mechanisms. Within a sovereign boundary model of interaction, there are potentially multiple novel concepts that lack precedent, existing mental model or metaphor to allow a participant to build a sufficient internalized understanding.

Below is a non-exhaustive list of potential original concepts which may prove alien to a new participant:

- Identity sovereignty
- Multiple identifiers
- Individual unique relationships with peers
- Peer to peer relationships and equality
- Establishing individual relationships to engage services
- The metaphorical boundary between the participant and the wider network
- The understanding of and the managing of static and dynamic data
- Data binding
- The blockchain as a source of truth
- Verifiable credential and the cascade of validity
- The issuing of credentials
- The concept of proving truths through partial data disclosure
- Finding faith in distributed storage
- Finding faith in one’s self and the responsibility of managing presence
- Understanding the relationships between digital and physical manifestations

Initial critical analysis of the model also raises some interaction challenges and potential incomplete mechanisms. These issues manifest around the edges of the core interaction and include the following:

- Direct messaging: issues concerning machine readable text. How can messages be interpreted, processed, and responded to at scale?
- Claim/credential sending: this would require a repository of standard templates across different categories.
- Claim/credential request: there would need to be a means of transmitting the claims that the connected entity provided.
- Machine readable T&C: there needs to be a means of providing machine readable T&C and potentially contracts. These need to go beyond the notion of Mary Hodders Customer Commons (Hodder, 2019).
- Dynamic data streams: there needs to be a means of understanding the data streams and usage around data binding.
- Progressive trust development: how can the initial building of trust through proofs be streamlined?
- Semantic layer: who defines the semantic layer to link claim elements to zero knowledge proofs?

Focus Group

The focus group included ten prominent individuals from the user experience and interface domain. Five predefined questions were posed to participants following prior communication of SSI technology and the results of the conceptual model and preliminary user interface through a purpose built website. There now follows a summary of the discussion and emergent themes.

From a Participant Perspective, How Understandable Are the Concepts Surrounding Self-Sovereign Identity?

Discussion was centered on the idea that the core concept of relationship building and the proving of something with a

TABLE 1 | Simplified table presenting SSI case studies and user scenarios.

Connection
Establishing a pairwise identifier with an individual in the field or generate and send.
Establishing a pairwise identity through a website.
Establish a pairwise relationship from a public DID.
Authentication
Sign into a website with a pairwise relationship.
Authenticate a credit card transaction online.
Authenticate a credit card transaction within a retail environment.
Authenticate identity within a physical space.
Authentication in the field to open a locked door.
Authenticating a ticket at a gate.
Sharing data
Establish a relationship and terms of data use when visiting a website.
Prove eligibility to hire a vehicle, age, license, capability to pay, and additional insurances.
Supply a number of proofs for an employment license.
Apply for credit providing proof of employment, address, and income.
Share realtime data from an IOT health data device with medic.
Share purchase history and financial position with an intent casting application.
Provide a claim for a personal reference.
Provide a claim to allow a child to attend a school trip.
Data gathering/management
Request a claim of educational record.
Establish a repository for IOT data.
Download and redistribute social network data.

verifiable credential was solid, but the peripheral mechanisms may prove problematic. Comments were made regarding the paradigm shift from asymmetric relationships to peer to peer engagement. Concerns were raised that the concept of having individual relationships for every transaction might prove difficult to grasp. First-hand experiences of the development of money sending applications were discussed, highlighting the problems encountered engaging users with one single new concept. It was argued that SSI posed an interesting challenge, as the overall system incorporated multiple new concepts and mechanisms. Arguments were made that these concepts need to be placed into context with clearly understandable use cases and value proposition. Concerns were raised with regard to the abstraction of many of the components of interaction and that in its current form, only the most committed privacy advocate or technology enthusiast would have the drive to fathom the conundrum. The consensus was that the system needed simplification, it needed to be placed in context, and that niches of value needed to be found in order to drive interest and engagement.

What Are Your Views on the Complexity of These Interactions?

The consensus to this question, in the context of the system presented, was that the current manifestation is over complex, that many of these issues can be resolved, but that a considerable reduction in friction needs to be achieved through systematic redesign and usability testing. It was argued that a balance between exposed and hidden interactions needs to be considered if a participant is to be able to initially engage with the system. An interesting debate ensued with regard to how much the user needed to see to comprehend the value, to understand the systems functions, to trust the system, and ultimately to develop and engage a mental model. This was considered to be a key set of variables that would need to be crafted within a further design cycle.

How Do We Build Something That Is Accessible?

The debate continued in the vein of the exposure to the under-laying mechanism, complexity of interaction, and motivation to engage. It was recognized that this would require a fine balance considering initial introduction of the technology through common place usage. The issue of exclusion and the prospect of a large proportion of society not being able to access SSI due to its complexity, initial friction, and the weight of responsibility were discussed in detail. The sentiment being "just because I can, does not mean I would want to," a conversation continued to discuss on boarding, with ideas being suggested that existing users of large public or corporate systems might be automatically enrolled into an SSI system, only realizing this once additional products and services were offered. This would be seeded by an existing base set of identity credentials. The debate moved to foundation identity credentials and how they could be established and the need for some kind of solid ground that all parties could trust and build upon. There was some interesting conversations regarding trust in one's own capabilities and the risk this entailed and from a user's perspective having different, sometimes compromising identities linked to one sovereign system. The consensus was that the interactive friction needed to be reduced, automatic migration to such systems might aid in adoption, and people needed to trust and understand the system, while having confidence in their own competence and capabilities.

Could You Share Your Thoughts on Mental Model and Metaphor?

The general consensus was that in its current form, the system would be difficult to comprehend and understand. There are potentially multiple original concepts all of which need to be considered and simplified relative to a holistic user experience. It was argued that a system of this kind needed to be standardized in terms of general concepts, language, and interaction, so the

TABLE 2 | Simplified table presenting objects and operations analysis.

Objects	Description
System state "notifications/activity/data flow"	A central object or dashboard from which other objects will be managed. The area will deal with notifications and warnings and act as a jumping-off point to other areas of interaction.
Admin and auth	
Administration	An object from which general administration configuration can be accessed.
Push authentication	An object from which pushed requests for authentication from established relationships can be accessed and actioned.
Connection "creation and distribution"	
DID creation	This object allows the creation of a new DID. At this point, metadata can be associated, so it can be recognized and managed by the agent application.
DID distribution	This object allows for the external distribution of a newly created DID.
Connection "incoming and pairing"	
Inputting DID "by string"	This object handles the inputting of initial or introductory DIDs.
Review DID doc and create pairing	Once a DID is entered into the inputting DID string object and verified, there is the option to pair with this DID to create a pairwise relationship.
Connection "established connections management"	
Listing of established connections	Within this object, all existing connections manifesting as DIDs can be viewed, filtered, and accessed.
History of interactions	Selecting an existing pairwise connection in the established connections object opens the history of interactions object. This displays a log of all interactions with this DID pairing. Messages/request claim/request proof/send claim/data binding.
Detailed pairwise DID information	Within this object, the details of the paired DIDs can be reviewed in detail. The DID document on both sides. Auth methods/endpoints/signature. There will be an option to refresh and revoke.
History log detail	If a specific element in a pairing history is selected, a detailed drill down of this interaction can be understood within this object.
Connection section "functions"	
Direct message	From here, a direct text-based message can be sent. At this stage, this is very basic and will evolve as the system develops.
Claim request	The predetermined claims available and offered by the pairwise relationship will be listed. The users can then select the claim they require and make a request.
Proof request	The proofs available and offered by the pairwise relationship will be listed. The users can then select the proofs they require and make the request.
Claim send	The claims available will be listed. The users can then select the claim they require, populate, and send.
Proof building	
Auto population and confirmation	This object is launched from a proof request. Once opened, the details of the proof requests are visible. The zeros are auto populated and can be seen. Within each element, there is the option to select an alternative.
Data management "proof/claim"	
Sent proof listing	Within this object, all existing sent proofs can be viewed filtered and accessed. The state of the proof is indicated. Live/pending/revoked. Selecting the proof opens the sent proof dig down object.
Sent proof dig down	Within this object, details of the proof can be understood. Dependent on the state, the proof can be sent/amended/revoked.
Received proofs	Within this object, all existing received proofs can be viewed, filtered, and accessed. The state of the proof is indicated. Live/expired/revoked. Selecting the proof opens the received proof dig down object.
Received proofs dig down	Within this object, details of the proof can be understood. From here, there might be an option to dig down further into the real details of the proof and the zero elements.
Received claims	Within this object, all existing received claims can be viewed, filtered, and accessed. The state of the claim is indicated. Live/pending/revoked/expired/rejected.
Received claims dig down	Within this object, details of the claim can be understood. Depending on the state, the claim can be either accepted or rejected.
Sent claims	Within this object, all existing sent claims can be viewed, filtered, and accessed. The state of the claim is indicated. Selecting the claim opens the sent claim dig down object.
Sent claim dig down	Within this object, details of the claim can be understood. Depending on the state, the claim can be updated/amended/revoked.
Data binding	This object is selected from the new option within the history of interactions object linked to the selected pairing. Once a pairwise connection is selected, the user can choose to allocate a data repository to it or bind data.
Existing data binding listings	Within this object, all existing data bindings can be viewed, filtered, and accessed. They can be scrolled and opened for limited information. The state of the claim is indicated.
Data binding dig down	Details of a data binding can be accessed from this object. From here, the raw data binding can be downloaded/suspended/revoked.

experience was consistent across contexts. There was agreement that the lack of understanding of the system might exclude certain types of individuals and that participants would need assistance from a trusted party to adopt. Comparison was made with crypto currency with it being argued that for many, this is not a technology that can be adopted independently. Comment was made that the next steps should include a systematic testing strategy in order to understand, simplify, and reduce friction and that a scaffold of the required mental model in terms of the concepts, interactions, and wider ecosystem should be mapped.

Can Any of This Be Automated?

The general consensus that emerged is that automation could solve many of the complexities and frictions identified that individuals are becoming accustomed to AI and that AI can work for the individual in a sovereign way. It was suggested that many interactions in the system are mundane and that repetitive actions could be driven by an overarching user policy. Discussion centered around two areas: firstly, the means for an individual to inspect the automated component and gain trust that it was acting in their interests, and secondly that we should not replace one kind of blackbox solution with another. There still needed to be visibility of the underlying processes and functions, so the participant can understand what was happening, in order to comprehend the value propositions and advantage.

DISCUSSION

The following sections cover the pertinent topics emerging from this research relative to the development of an interface layer for SSI, in the context of both a sovereign boundary mechanism and wider human-centered data ecosystem.

The Sovereign Boundary Mechanism

The development of a conceptual model in-line with the principles of the decentralized domain, following the trajectory of technologies and standards, realizes a system that allows participants to manage their data, information, communications and affairs independently through a digital wallet and agent. The concept of sovereignty in this context translates to individualism, and this in turn, given the complexity of the required interaction, poses several issues. Within this research, this independent domain of interaction has been titled a sovereign boundary mechanism. This means there is a clear boundary between the sovereign domain and the wider network. Within this domain participants manage identity, relationships, credentials, personal data, and access control, this is a strict task-based interaction, one which incorporates new distinct ideas and concepts. Given the gravitas and value of the personal data being transacted, this research suggests that a strong internalized understanding of the domain will initially be required to engage users. This requirement for internalized knowledge relates to the traditional notion of internalized cognition and mental model (Craik, 1943; Norman, 1986; Payne, 2003). Creating a situation where a participant is required to engage in significant internalized cognitive processes is counterintuitive to the evolution of HCI theory and accepted design thinking, where externalized cognition and computational offloading are considered best practice (Scaife and Rogers, 1996; Hutchins, 2000; Payne, 2003).

The required degree of internalized understanding and cognition in the context of adoption and the diffusion of innovation (Rogers, 1962) is arguably a primary consideration for the decentralized community. This research has clearly highlighted the complexity of a sovereign boundary mechanism, and further research and design practice needs to be undertaken to explore how complexity, friction, and internalized cognitive processes can be significantly reduced.

The Paradox of a Sovereign Boundary Mechanism

The objective of decentralization is to emancipate the participant from the centralized Internet. In doing so, the negative consequences are mitigated, and the missed opportunities presented by a decentralized alternative can be realized. However, this research suggests that the current trajectory may inadvertently replace one set of constraints with another. The complex landscape, isolation, internalized cognitive load, responsibility of managing one's own data, and generation of friction not found in centralized counterparts may replace one form of incarceration with another. It can be argued that these issues can be addressed and overcome, but the notion of the proverbial, out of the frying pan and into the fire, needs to be considered as future tools are conceived and developed.

The notion of a genuine decentralized Internet is predicated on the principle that the owner of the data should have control over it. The individual should have command over multiple immutable persistent identifiers, and they should have agency to decide who to share a relationship with and who on the network can observe their activities and transactions. The participant should have the capability to manage and redistribute their credentials or data to whom they see fit under their own terms and conditions (Mortier, 2014; Allen, 2016). This research has demonstrated what has been termed as a sovereign boundary mechanism, representing a participant who sits within a metaphorical boundary, defining identity, controlling relationships, and managing data streams. This is achievable through an analog model that potentially liberates the participant from the centralized Internet. Paradoxically, this model comes with a number of caveats which pose significant problems. This research has raised the issues of complexity, internalized cognition, mental model, friction, risk, responsibility, trust, and exclusion. By recognizing and considering these challenges through design thinking, coupled with an investigation of value proposition, the probability of mainstream SSI adoption can be significantly improved.

Back Peddling on Friction

A topic discussed within expert interviews and exposed through the practice-led component is that of increased interactive friction. This research suggests that the existing SSI analog model extended within this research exhibits a higher level of friction across interactions than that found within centralized counterparts. This relates to the required cognitive engagement, the understanding of original concepts, vague mental model and metaphor, and the shouldering of more personal responsibility. This research does not offer a metric on this assumption, but through developing and analyzing a sovereign boundary mechanism, the array of conceptual components for interaction and the collective required objects, and subsequent multilayered user interface, the friction level would appear to be substantial. If decentralized technologies are to find adoption, friction needs to be reduced not increased. Placing a number on this increased friction, potentially found in differing forms of decentralized interactions, is outside the scope of this research. However, this is a topic that warrants further investigation and should be prioritized within any continued endeavor.

A Missing Mental Model

When considering the underlying assumptions driving this research, the paper, Why Johnny Can't Encrypt, is cited

(Whitten and Tygar, 1999). The assumption is stated that the circumstances described within Whitten's work may emerge when considering decentralized endeavors. Within the cited research, it is argued that different methods of user evaluation are required when considering software, where a user lacks the understanding of the underlying concepts and mechanisms. In essence, the user lacks the mental model of the domain. The paper is concerned with sending emails with encryption software and argues that even with a well-designed interface, users struggle to complete what is a relatively simple task. A sovereign boundary mechanism represents a system which is arguably considerably more challenging. This research has demonstrated a user domain that is significantly complex. There are multiple concepts, processes, and interactions, which when taken individually, potentially lack the mental model for meaningful engagement. However, when these elements are combined as a whole, this research suggests that without considerable guidance, the objective of sovereign agency and utilization of personal data through a mechanism of this kind are impractical. It is accepted that the developed user interface within the conceptual model is preliminary and that further cycles of refinement will reduce complexity and potentially improve the mapping of interactions. However, refinement of the UI will not be enough and it has to be recognized that a full sovereign boundary mechanism, in this guize, presents considerable barriers with respect to forming an operational mental model for the participant.

Internalized Cognition

Leading on from the discussion regarding the complexity and potential missing mental models, even with sufficient understanding of the system domain, the degree to which a participant may rely on internalized cognition in order to engage is of concern.

This research identifies that the required internalized understanding for both the interaction and many of the broader concepts is significant. It can be argued that the cognitive load for initial engagement poses an issue for adoption. The very concept of sovereignty, and the metaphor of a secure boundary, suggests a degree of user isolation and internalization. The value and differing types of the personal data, the weight of being solely responsible, and the complex processes through which data must traverse to engage in meaningful transactions amounts to a sizeable load of internalized understanding and knowledge. If a mental model for this domain can be established, it can be argued that

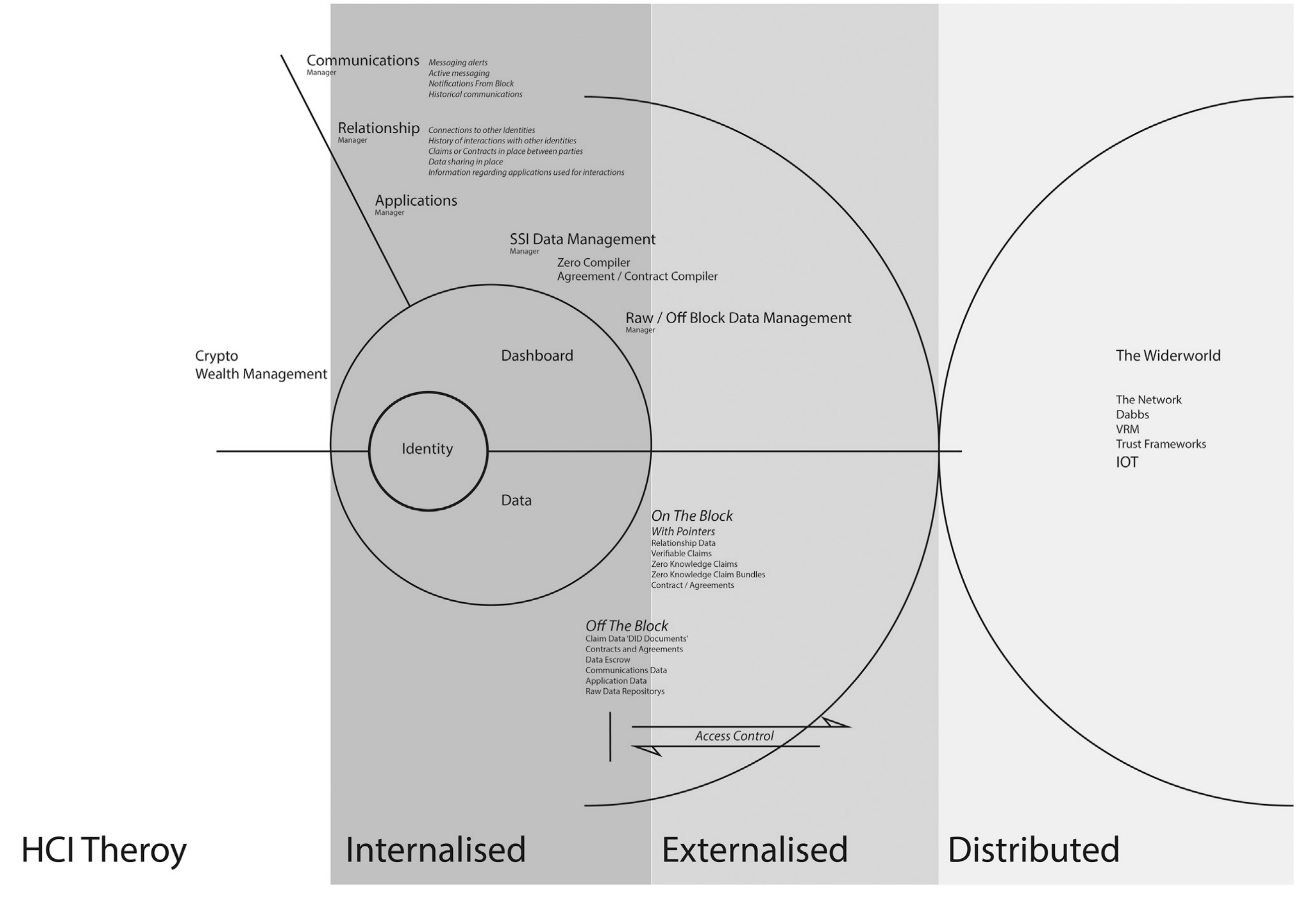


FIGURE 4 | The relevant HCI cognitive theory across a human-centered data ecosystem.

engagement will still require a considerable degree of internalized cognition. Within the expert interviews, comparison was made with crypto currencies and the difficulty individuals have in understanding and engaging with an ecosystem when asset value and responsibility fall into the equation. It can be argued that a sovereign boundary mechanism is considerably more complex. Within the product design process, the degree of internalized cognition needs to be accepted and measured. From there, every effort needs to be made to reduce the internalized understanding and decision making required. As discussed in the back pedaling on friction section, the increased friction posed by decentralized systems is a primary issue, and it is the internalized processes that are arguably responsible.

A Spectrum of Human Computer Interaction Theory

This research concludes that a human-centered data ecosystem through a sovereign boundary mechanism requires the consideration of a spectrum of HCI theory and paradigms (Rogers, 2012) and that any further innovation needs to recognize this in its deliberation. The graphic in **Figure 4** suggests the applicable theory across a model of a human-centered data ecosystem.

Internalized cognition and domain specific mental model (Payne, 2003) is relevant to the core of a sovereign boundary mechanism. Externalized cognition (Scaife and Rogers, 1996) is relevant between the core and the boundary. The notion of distributed cognition (Hutchins, 2000) (Payne, 2003) and situated action (Suchman, 1987) is applicable as engagement and transaction occur outside of the user boundary across the wider network. In addition, interactions and decision making within the wider ecosystem will be reliant on community-based templates and the development of trust networks, so societal and cooperative theories of HCI are also relevant (Schmidt, 2011). What is clear is that SSI falls into a number of HCI paradigms, and in identifying them, the application of Harrison's phenomenological matrix, where theories are considered and selected were deemed appropriate, would seem to be of relevance (Harrison et al., 2007).

This research suggests that the SSI community needs to recognize the importance of HCI cognitive theory and systematically address both issues of direct usability and the interaction and internalized understanding of concepts. This research argues that Payne's theory (Payne, 2003) of specific mental models for domains should be followed to map the required user understanding in detail. Any direct interaction should consider how the interface layer can push as much cognition as possible into the externalized realm. Finally, interaction and transactions need to be fully understood outside the sovereign boundary mechanism so that distributed cognitive relationships and situated actions can be defined. Above all, the consistency and cooperation across all stake holders are considered critical.

Types of activity recognized through this research, which may be reliant on distributed cognition and the consideration of situated action are as follows:

1. Validation/reputation/trust of individual actors
2. Collective decision making
3. Collective production activity
4. Collective data sharing
5. The construction of larger cognitive artifacts and systems

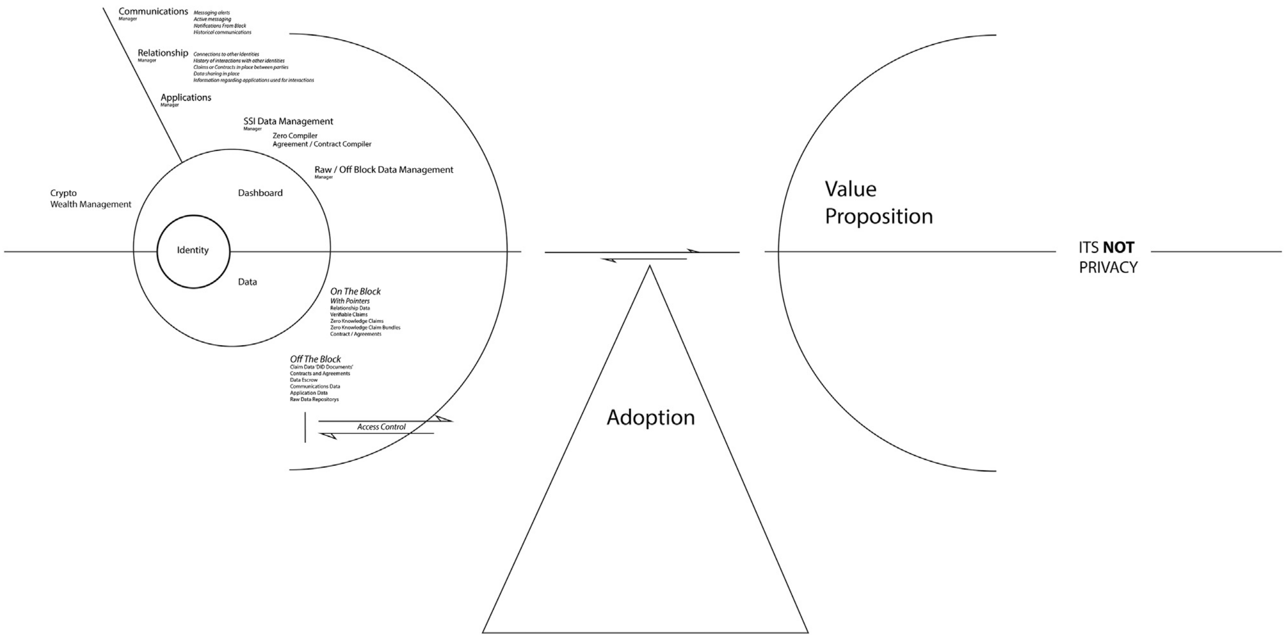


FIGURE 5 | Balancing the cognitive load against the value proposition.

A final notable HCI theory and ethical movement is that of human values, which argues that digital technology should strive to consider human values within its development process and ultimate application (Harper et al., 2008). The theory introduces the notion of empathy into the design process, as a point of reflection to consider the human values within any technology development.

Can AI or Collective Intelligence Reduce the Cognitive Load?

This topic of discussion derives from both consultation with experts and the critical reflection of the conceptual model. More conversation centers around the intelligent agent or personalized artificial intelligence, taking control of much of the mundane decision making and ongoing administration, involved in this type of decentralized system. This might be based on broad sweep criteria defined by a participant or derived through machine learning based on the participant's history. This concept has the potential to drastically reduce the cognitive load required, drawing the participant's attention only to critical and important decisions. The concept of AI supporting decentralized engagement is attractive but can also be seen as a misguided panacea. There are issues of trust, understanding, and value, which need to be considered as trade-offs against automation. If an objective of decentralization is transparency, to hide critical decision making within a blackbox may be contradictory. How can the participant trust the AI? Who defines the AI? And more importantly, how much of the inner workings of the system does the user need to see and understand, in order to both have faith and see value in the engagement? This is a significant topic for further research and debate.

Another solution to cognitive load reduction and decision making may be the establishing of group or tribe, a trust network that collectively makes decisions through shared values, for instance an environmental collective that marshals relationships and transactions through ethical reputations. This is part of the trust framework conversation, related to the theories of distributed cognition (Hutchens, 2000) and the emerging of technology in context (Suchman, 1987). This again offers a rich seem for future research.

An aside to the notion of collective decision making is that of democracy, political representation, and vote casting. An interesting discussion might be found around the concept of decentralized systems acting as a voting mechanism. The logic being if knowledge is power then sharing your data and subsequent inferred collective information may offer a new and dynamic means of democratic process.

Balancing the Cognitive Load Against the Value Proposition

The practice-based component of this research has demonstrated that there is considerable cognitive load, complexity, and participant responsibility that potentially manifests within an active human-centered data ecosystem. This friction might be mitigated through careful design considerations, but it can still be argued that even then the required engagement demands more effort on behalf of the participant than that currently found within existing centralized services. To this end, in line with many of the decentralized arguments around the communication of privacy, control, missed opportunity, and value proposition, any development of an interface layer needs to be balanced against the value that interaction serves to the user. Where this balance lies and how it manifests is a central conundrum in the delivery of a sustainable decentralized Internet and stands as a source of considerable further research. **Figure 5** Illustrates the balancing of Cogitative Load again the Value Proposition

A Starting Point for a Full Interface Layer Mapping

This research has argued for a standardized set of user interactions to facilitate engagement with a user-centered data ecosystem utilizing SSI at its core. The objective of this work has been to develop an initial conceptual model and preliminary user interface. Critical reflection has suggested that there are significant issues regarding internalized cognition, mental model, and metaphor. There are considerable opportunities for further practice-based research as this initial representation evolves through subsequent cycles. The preliminary conceptual model requires further testing, prototyping, and development. The mechanisms of the wider ecosystem need to be explored and metrics need to be drawn against a spectrum of models for a sliding scale of user engagement. There is a need to refine a detailed optimized mapping of this challenging user experience, as this work has taken only the first tentative steps toward that objective. This research has not resolved the challenge of establishing an accessible interface layer for self-sovereign identity, instead it acts as a contribution to knowledge and a jumping-off point for further research and development for what is a very important academic and practical domain.

CONCLUSION

This research concludes that the core technological infrastructure is now in existence to facilitate a genuine sovereign identity layer for the Internet, one which satisfies the principles of both human data interaction and self-sovereign identity. Investigation suggests that a dominant trajectory for a human-centered data ecosystem with the advent of a functional identity layer is progressing toward a sovereign boundary mechanism with self-sovereign identity at its core. By projecting forward the current trends through a conceptual modeling exercise, this research has demonstrated a potential interaction model that is complex and high in friction, requiring significant internalized cognitive processes and knowledge. Though the core technological infrastructure is in place, the development of a preliminary user interface suggests a number of mechanisms and interactions which still need to be developed to facilitate a full human-centered data ecosystem. This research concludes that in the development of an interface layer for SSI, a spectrum of HCI cognitive theory needs to be considered and any next steps should attempt to map the interplay between internalized, externalized, and distributed cognition. This research suggests that the SSI community needs to recognize the obstacle potentially posed at the interface layer and engage in collective standardization, strategy, and design thinking to increase the probability of the sustainable adoption of this revolutionary technology.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the author, without undue reservation

AUTHOR CONTRIBUTIONS

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

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbloc.2020.609101/full#supplementary-material.

REFERENCES

Allen, C. (2016). The path to self-sovereign identity. Available at: http://www.lifewithalacrity.com/2016/04/the-path-to-self-sovereign-identity.html (Accessed January 1, 2020).

Cameron, K. (2005). The laws of identity. Available at: https://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf (Accessed January 1, 2020)

Carroll, J. M. (1991). *Designing interaction: psychology at the human computer interface*. Cambridge, United Kingdom: Cambridge University Press.

Craik, K. J. W. (1943). *The nature of explanation*. Cambridge, United Kingdom: Macmillan University Press.

Creswell, J. W. (2003). *Research design: qualitative, Quantitative and mixed methods approaches.* 2nd Edn. London, United Kingdom: SAGE Publications Ltd.

Dunphy, P., and Petitcolas, F. (2018). A first look at identity management schemes on the blockchain. *IEEE Security and Privacy* 16 (4), 20–29. doi:10.1109/MSP.2018.3111247

Finch, H., and Lewis, J. (2003). *Focus groups, qualitative research practice*. London, United Kingdom: SAGE Publications Ltd.

Haddadi, H. (2015). Personal data: thinking inside the box. *Aar. Hu. Cen. Comp.* 1, 1. doi:10.7146/aahcc.v1i1.21312

Harper, R., Rodden, T., Rogers, Y., and Sellen, A. (2008). *Being human: human-computer interaction in the year 2020.* Cambridge, United Kingdom: Microsoft Research.

Harrison, S., Tatar, D., and Sengers, P. (2007). "The three paradigms of HCI," in Alt. Chi. Session at the SIGCHI conference on human factors in computing systems, San Jose, CL, January 2007 [abstract].

Hodder, M. (2019). Customer Commons. Available at: http://www. http://customercommons.org/author/mary-hodder/ (Accessed October 13, 2019).

Hornung, H., Pereira, R., and Baranauskas, M. C. C. (2015). Challenges for human-data interaction—a semiotic perspective. *Hum.-Comp. Inter.: Design and Eval.* 1, 37–48. doi:10.1007/978-3-319-20901-2_4

Hutchins, E. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Trans. Com.-Hum. Inter.* 7 (2), 174–196. doi:10.1145/353485.353487

IIW (2019). Internet identity workshop. Available at: https://internetidentityworkshop.com/ (Accessed April 7, 2019).

Johnson, J., and Henderson, A. (2002). Conceptual models: begin by designing what to design. *Interactions* 9 (1), 25–32. doi:10.1145/503355.503366

Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends Cognit. Sci.* 5 (10), 434–442. doi:10.1016/s1364-6613(00)01751-4

Kirsh, D. (1997). Interactivity and multimedia interfaces. *Instr. Sci.* 25, 79–96. doi:10.1023/A:1002915430871

Landuer, T. K. (1991). "Let's get real designing interaction," in *Psychology at the human computer interface* (Cambridge, United Kingdom: Cambridge University Press).

Larkin, J., and Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognit. Sci.* 11, 65–99. doi:10.1016/S0364-0213(87)80026-5

Moglen, E. (2013). The tangled web we have woven. *Commun. ACM* 56 (2), 20–22. doi:10.1145/2408776.2408784

Moore, G. (1991). *Crossing the chasm*. New York, NY: Harper Business Essentials.

Mortier, R. (2014). Human-data interaction: the human face of the data-driven society. Available at SSRN: https://ssrn.com/abstract=2508051 (Accessed September 10, 2018).

Nielsen, J. (2005). *Usability inspection methods*. New York, NY: John Wiley & Sons.

Norman, D. (1986). "Cognitive engineering," in *User centred system*, Editors S. Draper and D. Norman (Hillsdale, NJ: Lawrence Erlbaum Association).

O'Malley, C., and Draper, S. (1992). "Are mental models really in the mind? representation and interaction," in *Models in the mind: perspectives, theory and application*, Editors Y. Rogers, A. Rutherford, and P. Bibby (London, United Kingdom: Academic Press), 73–91.

Payne, S. (2003). "Users mental models: the very ideas," in *HCI models, theories, and frameworks: toward a multidisciplinary science* (Burlington, MA: Morgan Kaufmann).

Rogers, E. (1962). *Diffusion of innovations.* 5th Edn. New York, NY: Free Press of Glencoe.

Rogers, Y. (2012). *HCI theory, classical, modern, and contemporary*. San Rafael, CA: Morgan & Claypool Publishers.

Scaife, M., and Rogers, Y. (1996). External cognition: how do graphical representations work? *Int. J. Hum. Comput. Stud.* 45 (2), 185–221. doi:10.1006/ijhc.1996.0048

Schmidt, K. (2011). *Cooperative Work and coordinative practices: Contributions to the conceptual Foundations of computer-supported cooperative work (CSCW)*. Heidelberg, Germany: Springer. doi:10.1007/978-1-84800-068-1

Schneier, B. (2015). *Data and Goliath*. New York, NY: W. W. Norton & Company.

Solove, D. (2008). *Understanding privacy*. Cambridge, MA: Harvard University Press.

Suchman, L. (1987). *Plans and situated actions*. Cambridge, UK: Cambridge University Press.

Tobin, A., and Reed, D. (2016). The inevitable rise of self-sovereign identity Sovrin Foundation. Available at: https://sovrin.org/wp-content/uploads/2018/03/The-Inevitable-Rise-of-Self-Sovereign-Identity.pdf (Accessed April 20, 2020).

Turnhout, K. (2014). *Design patterns for mixed-method research in HCI*. Nijmegen, Netherlands: HAN University of Applied Science.

Van Kleek, M., and O'Hara, K. (2014). "The Future of Social is personal: the Potential of the personal data store," in *Social collective intelligence: combining the powers of humans and machines to build a smarter society*. Heidelberg, Germany: Springer. Available at: http://eprints.soton.ac.uk/363518/1/pds.pdf (Accessed September 28, 2017), 125–158.

Whitten, A., and Tygar, J. (1999). Why Johnny can't encrypt: a usability evaluation of pgp 5.0," in Proceedings of the 8th USENIX Security Symposium, Washington, DC, December 8, 1999.

Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *J. Inf. Technol.* 30, 75–89. doi:10.1057/jit.2015.5

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.

Exploring Value Propositions to Drive Self-Sovereign Identity Adoption

*Mick Lockwood**

Salford School of Arts, Media and Creative Technology, University of Salford, Salford, United Kingdom

This paper presents research exploring the balancing of interactive friction and value proposition in the context of Self-Sovereign Identity (SSI) technology adoption. This work extends a related investigation of a full agency engagement with a User-Centred Data Ecosystem utilising what is described as a Sovereign Boundary Mechanism (SBM). An SBM is a standardised collection of SSI interactions, which can collectively be described as a metaphorical ring of sovereignty between the participant and the wider network. Within this model participants control identity, relationships, credentials, data streams, and access control. This related work concludes that the developing trend poses significant interactive friction, and that clear and substantive value proposition would be required to drive and sustain participant adoption. This paper explores potential value propositions for SSI, considering theory relating to Privacy, Surveillance Capitalism, and Human Data Interaction; in parallel opinions are drawn from the thematic analysis of interviews with experts in the decentralised field and results from a public survey. This research concludes that the value proposition is unlikely to come from the direct perceived protection of privacy. Also, that the decentralised technologies cannot be marketed solely on the fact that it is decentralised. Instead, value will emerge from the capability of SSI functionality to supersede the centralised model, offering innovation and reduced transactional friction across individual, business and wider society. This research suggests that the SSI community needs to develop a cohesive design strategy, a clear narrative and vocabulary. Value needs to be defined across cultural context, while targeting accessible, high value niche opportunities to build momentum toward sustainable adoption.

Keywords: Self-Sovereign Identity, Human Data Interaction, Human-Centred Data Ecosystem, Sovereign Boundary Mechanism, Decentralised Internet, Value Proposition, Adoption

OPEN ACCESS

Edited by:
Kaliya Young,
Merritt College, United States

Reviewed by:
Larry C. Bates,
AltMarket, United States
Pratap Tambe,
Tata Consultancy Services, India

****Correspondence:***
Mick Lockwood
m.g.lockwood@salford.ac.uk

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

Received: *29 September 2020*
Accepted: *13 January 2021*
Published: *04 March 2021*

Citation:
Lockwood M (2021) Exploring Value Propositions to Drive Self-Sovereign Identity Adoption.
Front. Blockchain 4:611945.
doi: 10.3389/fbloc.2021.611945

INTRODUCTION

Within a separate paper published within this journal entitled *An Accessible Interface Layer for Self-Sovereign Identity,* the need to balance the significant levels of cognitive load found within SSI interactions with genuine value proposition is discussed at length. An interface layer for SSI engagement is a paradigm shift in the way individuals interact with the network. Concepts of identity management, relationship building and data sharing as part of a wider User-Centred Data Ecosystem (UCDE), present a problematic level of friction, when considered alongside the theories of adoption. The value proposition enabled through SSI has to offer more than decentralisation, more than a vague promise of privacy protection. It has to enable clear, sustainable advantages over its centralised counterparts, or the technology will fail to find widespread adoption. This paper presents a component part of wider doctoral research undertaken between 2017 and 2020. The research posed the central question: Can a sustainable technology be established to allow for individual agency within a decentralised Internet? Two additional questions were then derived. The first considered usability and accessibility at the interface layer and asked: Can an interface layer for a decentralised Internet be designed to allow for accessible interaction? And the second considered value proposition and adoption with the question: How might a decentralised Internet provide value, emerge and be adopted? This paper presents the investigation of the latter through the lens of SSI. It does this through an exploration of the literature, a public survey and thematic analysis of a series of semi-structured interviews with both experts from the decentralised field, and practitioners from the realm of usability and user experience.

Theoretical Framework

A literature review was conducted which focused on four pertinent areas: Surveillance Capitalism, Network Privacy, Human Computer Interaction, and the principles and supporting arguments for Human Data Interaction.

The review undertook a foundation investigation of classical surveillance theories, before investigating arguments concerning personal data gathering, aggregation and secondary use. It continued to investigate the historical narrative that has led to the status quo, and the relationship between large-scale data collection, and our digital economy. The review considered the notion of privacy, exploring the fundamental theory, cultural differences and social norms. It explores the economic, social and cultural value of personal data. It investigates the legal landscape, and the arguments for the granting and restriction of privacy rights. The review considers privacy in the digital realm, investigates the positive aspects, and potential harms of big data collection. The review considered Human Computer Interaction, exploring the domain's progression, with a focus on cognition, investigating theories most associated with individual interaction with both system and interface. Finally, the review considers the emergent domain of Human Data Interaction, charting its evolution, arguments for its realisation, and underlaying principles and trajectory. In the following paragraphs the relevant theories are surmised.

Surveillance Capitalism

Initial investigation considered the *Panopticon*, the design for a penal institution conceived by social reformist Jeremy Bentham (1791) in which inmates could be observed by a single guard, without ever knowing for certain that they were being surveilled. Investigation continued to explore more recent interpretations of Benthams philosophy, which saw the Panopticon model as social control by the capitalist (Himmelfarb, 1968). Michael Foucault's observations of the Panopticon are considered, alongside his arguments surrounding changes in western social control, were discipline is now metered in the mind as opposed to the body (Foucault, 1977). The review considered the transition of the Panopticon into the digital realm through the notion of *Cybernetic Capitalism* (Robins and Webster, 1988). Poster (1990) offers a profound prospective through *The Electronic Superpanopticon*, in which the individual has a second observable existence within the database. The *Social Sort* David Lyon (1993) describes the way individuals are profiled, targeted or excluded from communication and marketing materials. The *Panoptic Sort* explores the technology driven intelligence gathering of an individual's economic value (Gandy, 1996). The investigation of surveillance continues to consider its mechanisms with the concept of *Produsage* coined by Alex Bruns (2006) in which the participant is both producer and consumer of media and knowledge. Christain Fuchs (2012) extends this further with the *Prosumer Proletariat*, arguing that participants become part of *Marxist Class Theory* as they become productive labourers who produce surplus value. Shoshana Zuboff (2015) continues the line with the introduction of the term *Surveillance Capitalism*, arguing that each phase of capital requires a reinvention of the *Logic of Accumulation*. Jacob Silverman (2017) argues that we are entangled in these networks, and all but the most committed rebel or eccentric are resistant to its grasp. This rich seam of literature has been influential in this research as it provides a lens through which to understand the current landscape, while supporting the arguments of opaque exploitation and the notion that *'we are on the verge of eliminating forever the fundamental right to be alone in our thoughts'* (Moglen, 2013).

Network Privacy

Allen Weston (1967) defines privacy as *'the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent personal information is communicated to others'* (p. 7). This definition is clear, but when applied to the complexity of the real world, it becomes evident that privacy as a concept is not only incredibly complex, but poorly defined and misunderstood. Robert Post (2001) explains: *'Privacy is a value so complex, so entangled in competing and contradictory dimensions, so engorged with various and distinct meanings that I sometimes despair whether it can be usefully addressed at all'* (p. 2087). Judith Thomson (1975) observed of privacy that *'perhaps the most striking thing about the right to privacy, is that nobody seems to have any clear idea what it is'* (p. 272). Jeff Jarvis (2011) comments on public perceptions of privacy across the Internet as *'a confused web of worries, changing norms, varying cultural moves, complicated relationships, conflicting motives, vague feelings of danger with sporadic specific evidence of harm, and unclear laws and regulations made all the more complex by context'* (p. 101). Arguments have been made that attempts to locate the essence or core characteristics of privacy have led to failure. (Solove, 2008, p. 8). Contrasting this confused landscape of understanding are claims for the need for privacy: It is a fundamental part of our social structure. To have a society without a degree of non-disclosure of private thought, action, property or information would be impossible to achieve. Privacy is fundamental to our notion of self, to our independence and sense of dignity. It is part of our cognitive development, as we first understand that those around us do not have access to our inner thoughts and ideas. In choosing to disclose our emotions, our desires, our motivations or political positions, we develop complex social structures and intimate relationships. Privacy is a critical component of our democracy, and our western liberal society. (Gavison, 1984; Solove, 2008; O'Hara, 2016). When exploring the privacy literature, it is evident that defining the concept of privacy is complex, and that building clear value propositions around its essence for the decentralised domain might prove problematic.

This research continued to considered theories that may hold value when attempting to understand the domain of network privacy, while forming the basis for the exploration of application, value and communications strategy. The following sections highlight some of the prominent ideas.

Bruce Schneier (2015) makes strong arguments for the ephemeral and the right to be forgotten. He argues that the very nature of our social interactions are reliant on an ability to forget the happenings of the past. Our capacity to forget, and for painful memories to fade out of existence, is part of the process of healing. To lose the ephemeral in our cultural interactions is a paradigm shift. In addition, Schneier makes compelling arguments to counter the claims that automated Algorithmic Surveillance is not a privacy infringement until a human being enters the equation (Kessler, 2013). He argues that a computer can flag up at any time information it encounters and that a participant cannot be sure they won't be *'judged or discriminated against on the basis of what the computer sees'* (Schneier, 2015, p.153, p.153)

Paul Ohm (2010) argues that we are making a mistake in putting our faith in the anonymisation of personal data. He argues that data can be re-identified when cross referenced against other data sources, and that there is a motivation to limit anonymisation, as it decreases its utility and monitory value.

Daniel Solove (2008b) counters Eric Schmidt's argument (Huffpost, 2010) that an individual should not be fearful of surveillance if they have *Nothing to Hide*. He argues that hiding something is assumed to be about hiding bad things, when in reality privacy is a function of human development and a wider function of society. Max Van Kleek and O'Hara (2014) argues that data mining and aggregation of personal data can 'threaten our privacy, or our dignity, or our autonomy by '*diluting the privileged first-person access to our own experience*' (p. 5). Solove (2009) argues that aggregated information can reveal facts that the participant did not expect to be known when the original isolated data was collected. Perhaps the most powerful example of the potential for aggregated data and subsequent knowledge gleaned from inference is the work of Dr. Michal Kosinski et al. (2013). In his paper entitled, *Private Traits and Attributes are Predictable from Digital Records of Human Behavior*, Kosinski demonstrates a powerful method to develop accurate individual psycho-demographic profiles through the analysis of Facebook Likes. This method is broadly accepted to be the one used by Cambridge Analytica (2017) which sparked controversy and accusation of electoral manipulation (Rosenberg, 2018).

Patricia Norberg's et al. (2007) *Privacy Paradox* describes a disparity between attitudes and behavior concerning network privacy. It is claimed that individuals voice concern about their privacy online, only then to act in a way that demonstrates little concern for their private information, often releasing personal data for very little reward. Acquisti (2004) argued that individuals may not be able to act rationally in an economic transaction when it comes to personal data. He extended behavioral economics literature to describe what he termed *Immediate Gratification Bias* (p. 2), a term which suggests that individuals place higher value on immediate benefits rather than future risks.

In coining the term *Bounded Reality* Herbert (1955) argues that many economic predictions of an individual's behavior and decision making when forming choices are based on a capability to act rationally. Herbert argues that true rational decision making requires a complete understanding of alternative choices and their consequences and would require an infinite time to deliberate. Instead Herbert suggests that an individual's capability to act in a rational way is bounded by the individual's tractability, the cognitive limitations of the mind and the time available to make any decision. Herbert comments that an 'organism's simplifications of the real world for purposes of choice introduce discrepancies between the simplified model and reality' (p. 114). When considering the development of any decentralised system we cannot assume that an individual will act rationally in the classic sense, instead an individual may act in a way that reflects their own reality and understanding of the world.

Sharot (2011) describes *Optimism Bias* as a cognitive process by which an individual believes that they are less likely to experience a negative occurrence then is statistically probable. She explains, *'humans, exhibit a pervasive and surprising bias: when it comes to predicting what will happen to us tomorrow, next week, or fifty years from now, we overestimate the likelihood of positive events, and underestimate the likelihood of negative events'* (p. 941).

Danial Solove (2008) argues that the conceptualisation of privacy is of *'paramount importance for the information age because we are beset with a number of complex privacy problems that cause great disruption to numerous important activities of high social value'*. He suggests that instead of a top down approach we should come from the bottom up to *'understand privacy as a set of protections against a plurality of distinct but related problems'* (p. 171). The term 'privacy' then acts as an umbrella term to cover these protections. He argues that we should see privacy issues through the lens of the problem, adopting a pragmatic approach that resists universals and embraces specific solutions, and that we should *'understand privacy in specific contextual situations'* (p. 47).

Danial Solove's taxonomy is important to this research, as it offers a framework through which to explore real world privacy issues, user journeys and potential privacy harms. This research argues that the theory can be extended, not only to support law and policy makers, but also to inform the development of decentralised systems, tools and services, genuine value proposition, and communications strategies.

Human Data Interaction

The field of Human Data Interaction (HDI) (Mortier, 2014; Chaudhry et al., 2015) recognises the pervasiveness of computing in our data driven society. The theory argues that Human Computer Interaction (HCI) has traditionally focused on interactions between humans and computers as artifacts, but with the rapid evolution of humans interacting predominantly with data, a different academic perspective is required. Moritier (2014) defines the essence of HDI as *'placing the human at the center of the flows of data, providing mechanisms for citizens to interact with these systems and data explicitly'* (p. 1). The concepts of HDI illustrate the opaque mechanisms used to process personal data and the hidden inferences and subsequent feedback loops. The theory argues that a user requires legibility to understand the ambient ways in which their data is processed and utilised, that agency is required to control, manage and permit access to personal data, and that users require a means to negotiate the terms under which their data can be used. SSI as a standardised collection of interactions can form the core component of a UCDE. As this component provides the sovereign mechanisms and a metaphorical boundary between the participant and the wider network, it is described within this research as a Sovereign Boundary Mechanism (SBM). The exploration of these concepts together with the broader discourse surrounding transactional mechanics, economics, societal impact, identity and individual privacy can in the context of a data ecosystem, be represented within the academic domain of HDI.

As part of the exploration of HDI, theories have been considered which may form a scaffold for justification for adoption, the development of value proposition and the building of narrative and communications strategy. The following sections explore some of these concepts.

Adoption Theory

While exploring variables surrounding this problem space, adoption theory forms an important foundation. In this respect the *Diffusion of Innovation* (Rogers, 1962) and the subsequent *Technology Life Cycle* Theory (Moore, 1991) are considered most relevant.

According to Rogers (1962), the adoption of a new product, service or technology happens in five stages, known as the *Innovation Decision Process*.

- **Knowledge:** when the individual is exposed to the innovation's existence and gains an understanding of how it functions.
- **Persuasion:** the forming of a positive or negative attitude.
- **Decision:** when an individual engages in activities that lead to a choice to adopt or reject.
- **Implementation:** when a user commits and begins to use a product or service.
- **Confirmation:** the user seeks reassurance about a decision to adopt and may reverse that decision if exposed to conflicting messages.

This research suggests that SSI may encounter resistance at the *Knowledge Stage*, as participants are confronted with a system that is poorly defined and complex. This is also the case at the *Persuasion Stage*, as participants struggle to comprehend a clear value proposition and the benefits of adoption.

Geoffrey Moore (1991) expands Rogers theory with the *Technology Life Cycle*. Moore argues that cracks can appear in the adoption curve between innovators and early adopters, and a chasm can emerge between early adopters and the early majority, when a disruptive technology cannot be readily translated into a major new benefit. Moore argues that *'the enthusiast loves it for its architecture, but nobody else can even figure out how to start using it'* (p. 14).

When describing the concept of a chasm, Moore explains that *'when a product reaches this point in the market development, it must be made increasingly easier to adopt in order to continue being successful. If this does not occur, the transition to the late majority may well stall or never happen'* (p. 14). It can be argued that without clearly defined value proposition, SSI potentially represents a textbook case for Moore's adoption chasm. Moore defines a number of steps that need to be considered to avoid the chasm in the adoption curve.

- **Target the Point of Attack:** This step refers to the identification and focus on a specific market niche.
- **Assemble an Invasion Force:** This refers to the creating of the whole product, recognising the problem faced by a participant and providing everything necessary to solve the problem.
- **Define the Battle:** The identification of the competition, the development of a competitive claim, the formulation of the communication of that claim, and the capability to demonstrate its validity.
- **Launch the Invasion:** In the context of traditional sales of technology or product, this relates to distribution and pricing. Moore advocates a direct sales approach with a central consultative figure supported by application and technology specialists.

This research suggests that in any continued development of SSI technologies, adoption theory needs to become a critical variable in the overall consideration of the problem space.

The Complexity of Personal Data

An issue to consider in the context of personal data management and the design of decentralised systems is the complexity of personal data. Chaudhry et al. (2015) explains '*as soon as one begins to examine the requirements for a Databox, one thing becomes very clear: data is a dangerous word. In particular, personal data is so complex, and rich that treating it homogeneously is almost always a mistake'* (p. 3). SSI at present focuses on the generation of verifiable credentials, evolving collections of data that build elements of identity. These credentials can be authenticated through issuer signatures on a blockchain. A full-scale UCDE will require, static and dynamic data, data that is continually updated and data that is produced, utilised and controlled by multiple identities. Van Kleek and O'Hara (2014) comments: *'the task of identifying all of the kinds of data a person might need to keep, manage and use is complex and not easily scoped'* (p. 8). The title *Keeping Found Things Found* (Jones, 2010) offers a taxonomy of personal data, which may act as an excellent starting point when considering the format of data types required to drive a functional decentralised system across multiple contexts.

The Lost Opportunity of Big Data

A powerful argument for the adoption of SSI technology comes from the lost potential of *Big Data*. Wendy Hall (2016) commented on the value of personal data with the following statement: *'When I say value, I don't simply mean a nation of individuals being able to sell their data for monetary gain. I am talking about how vital the sharing of personal data is in technological, and specifically digital, innovation'* (p. 3). The term *Big Data* does not solely refer to a vast quantity of data which cannot be processed or made sense of, but rather, to a vast collection of valuable information, that offers great potential to a spectrum of society. Alex Pentland argues that Big Data offers huge opportunities, as it promises to reveal the underlying mechanisms of the world in real-time. We are only just beginning to understand through data science, the potential innovations and benefits to society that this rich knowledge resource can offer. Pentland argues: *'I believe that the power of Big Data, is that it is information about peoples' behaviors, instead of information about their beliefs'* (Pentland, 2012). Planning, health, business, security and personal interactions with the world, can be revolutionised as we move from knowledge based on averages and statistics, to real-time, real-world data at a micro level: *'With Big Data, we can begin to actually look at the details of social interaction, and how those play out and are no longer limited to averages like market indices or election results. This is an astounding change'* (Pentland, 2012). Pentland goes on to argue, that this prospect will only become a reality if people are willing to release their personal data, freely, confidentially, and on their own terms. Without this agency and trust, we risk stifling, restricting or losing altogether this promising capability.

The Economic Value of Personal Data

The value of personal data is a topic widely discussed in the literature. The direct sale of personal data by an individual for financial renumeration is questionable, as the dollar value in this context is very low. The interesting value can be found in the macro economic data. A report published in 2012, by *The Boston Consulting Group*, highlighted the huge current and future value that can be attributed to personal identity and personal data. Within the EU it equates to 8% of the EU-27 GDP. They predicted this to be worth €330 billion annually to organisations, and €670 billion to consumers by 2020 (BCG, 2012). This did though come with one significant caveat. The report explained: *'However, two-thirds of potential value generation, €440 billion in 2020, is at risk if stakeholders fail to establish a trusted flow of data'* (BCG, 2012, p. 3). The report continues to list areas of value for commerce as: process automation, user enablement, personalisation, enhanced delivery, personal data driven R&D, and secondary monetisation.

A Stifled Digital Economy

There are arguments regarding the current trajectory of the digital economy and the consequences of a model that locks in and constrains the customer. Chaudhry et al. (2015) states that *'increasing lock-in and network externalities are preventing formation of a truly competitive market'* (p. 1). The publishing of the *Cluetrain Manifesto* by Rick Levine (2000), communicated to business the profound change the Internet would have on established markets, and mechanisms for doing business. It likens the advent of the Internet, and its ability to facilitate conversation within the market, to that of an ancient bazaar, Levine explains, *'in sharp contrast to the alienation wrought by homogenized broadcast media, sterilised mass culture, and the enforced anonymity of bureaucratic organisations, the Internet connected people to each other, and provided a space in which the humans voice would be rapidly rediscovered'* (p. 6). The text argued that business had to adapt to this new reality of two-way conversation or die. Doc Searls extended his own contribution to the *Cluetrain Manifesto*, with *The Intent Economy* (Searls, 2012). This text incorporates many ideas and concepts derived from the twice-yearly *Internet Identity Workshops* (IIW, 2019) founded by Searls, Young, Hamilin and Windley in 2005, and Project VRM *'Vendor Relationship Management'* started by Searls at *Berkman University* (ProjectVRM, 2019). A central argument in *The Intent Economy,* is that in order for Digital Commerce to reach its true potential, the customer must be freed from the silo of *Customer Relationship Management* and *Captor of Choice*. It is argued that the liberation and communication ability that the Internet brings, makes obsolete, or at least inefficient the industrial revolution type business model of mass production, mass marketing, and mass media. That the *Contract of Adhesion*, or *Adhesionism*, where establishing asymmetric contracts is the only option when dealing with large numbers of unknown customers and users, is out-dated. The current models of marketing through the amassing and secondary use of personal data is unsustainable. It is argued that there are many opportunities, for those who can be first to market, or who empower the user to communicate their intent into the marketplace. We are beginning to see the breakdown of the existing models, and a growing awareness that we have built our digital economy on a foundation that is ethically questionable and potentially finite. As individuals become more aware, and begin to employ privacy enhancing technologies, such as Ad and Cookie Blockers, VPM's and Tunneling, the ability of marketers to gather quality data and marketing intelligence diminishes. The advent of GDPR in the European Union, has the potential to disrupt the current practices, and it is argued that there needs to be a new approach that recovers the digital economy from a race to the bottom.

This research suggests that there is value proposition in many of the developed concepts of VRM for both the network participant and vendor. It remains to be seen if the advent of SSI and its ability to establish an identity layer for the Internet, can move any of the existing models of VRM from concept through to the mainstream.

The Risk to Our Democracy

When commenting on political campaigns, Cathy O'Neil (2016) argues that *'they can target micro-groups of citizens for both votes and money, and appeal to each of them with a meticulously honed message, one that no one else is likely to see. Each one allows candidates to quietly sell multiple versions of themselves, and its anyone's guess which version will show up for work after* *inauguration'* (p. 160). Within a traditional democratic political campaign, the objective is to appeal to as many voting groups as possible, spreading your policies widely, while being able to defend each of them in the public domain. If voters can be profiled and influenced directly away from the public sphere, without scrutiny, the model of a western liberal democracy is jeopardised. Monbiot argues that: *Our model of democracy is based on public campaigning followed by private voting. These developments threaten to turn this upside down, so that voting intentions are pretty much publicly known, but the arguments that influence them are* made in secret, *concealed from the wider world, where they might be contested'* (Monboit, 2017). Indeed, a powerful argument for HDI is the risk posed to the democratic system. Data inference and pattern recognition offer the prospect of micro targeting of an individual's political persuasion, in a narrow cast and unaccountable manner. Monbiot argues that, '*micro-targeted ad campaigns are by their nature private or narrowcast. They never reach outside their target audience. Thus, they can contain falsehoods or insinuations that are never challenged because they are never brought to light'* (Monboit, 2017).

In recent times, insight into a possible future comes from the Cambridge Analytica episode. This company specialised in targeted campaign intelligence, based on establishing psychological profiles through behavioral science and big data analysis. In an article entitled *The Data That Turned The World Upside-down* published by Swiss publication *Das Magazine* (Grassegger and Krogerus, 2016), it is claimed that by using a profiling technique called *'OCEAN, an acronym for Openness, Conscientiousness, Extroversion, Agreeableness, Neuroticism—we can make a relatively accurate assessment of the kind of person in front of us'* (Grassegger, 2016). Coupling psychological profiles with tailored advertising allowed micro targeting of the voting public in the US 2016 presidential election. This method is said to be a version of that developed by Dr. Michal Kosinski (Kosinski, et al., 2013). The real impact of Cambridge Analytica's methods have been countered and unpicked by Martin Robbins, who disputes the claims based on the numbers presented. He argues that *'there's no evidence of this voodoo marketing in action, and we have plenty of anecdotes pointing to less than stellar use of data by campaigns'* (Robbins, 2017). Leonid Bershidsky also points out his doubts of the claims made, based on his own experience of the poorly targeted messages he received during the campaign (Bershidsky, 2016). Both counter arguments claim that Cambridge Analytica's capabilities have been over-hyped, and that their involvement and media coverage, has more to do with the members of its board, than its actual ability. Whatever the depth of influence, it demonstrates a trajectory that may not be desirable, and that threatens to undermine democratic systems. As Mondiot explains: *'the Cambridge Analytica story gives us a glimpse of a possible dystopian future, especially in the US, where data protection is weak'* (Monboit, 2017).

The surveillance, classification and monitoring of individuals and groups to profile politically is nothing new. However, the advent of *Big Data* analytics allows mass surveillance and inference to be drawn across every participant who engages with the network. The advent of this capability potentially removes the privacy component that allows democracy to function, allowing clandestine micro targeting of political messages. It must also be considered that the Cambridge Analytica story involved a third-party company who received their data second hand. Facebook however, has a vastly larger reservoir of real time data and considerable data analytic expertise. O'Neil (2016) questions *'by tweaking its algorithm and molding the news we see, can Facebook game the political system?'* (p. 145). Facebook also has the capability to enact echo chamber. A great proportion of current affairs and general news is now ingested by way of the Internet and through social media. The echo chamber metaphor suggests that news and ideas will be tailored for the individual, relative to a profile constructed from personal data. In essence they are telling the individual what they want to hear, reinforcing their expressed views, without ever being exposed to the ideas and opinions of others. The message of a threat to democracy and manipulative control is powerful and can be woven into a clear value proposition and communications strategy for both SSI and a wider UCDE.

METHOD

The following section describes a combination of research components from a wider doctoral study designed to explore potential value propositions for SSI. Together with an exploration of the literature, the investigation draws on two strands of primary research, a public survey and a series of expert interviews. These components are part of a broader mixed methods design (Creswell, 2003) influenced by design theory's for Mixed Methods in HCI (Turnhout, 2014) **See Figure 1**.

Public Survey

The Public Survey investigated attitudes toward Internet usage, data privacy, the disclosure and secondary use of personal data, and engagement with activities and opportunities to protect and control personal information. Analysis of the data gathered provided a detailed picture of public perceptions and attitudes at a descriptive level. Latent considerations were designed into the survey to uncover signifiers relating to Catalyst for Adoption, Value Proposition, and potential Development Strategies. The survey was made up of 52 questions consisting of Likert Items and Forced Binary. The questions were designed to function in two forms. Firstly, as individual Likert Elements targeting specific desired information and Second, collections of Likert Elements designed to generate Likert Scales (Likert, 1932). The resulting data is presented in two forms, basic descriptive statistics of individual questions and correlation and comparisons of Likert Items and Forced Binary scales. The full listing of survey questions has been provided as a **Supplementary Material** to this paper.

Expert Interviews

Primary data was gathered through three phases of semi structured interviews. The first phase explored the board decentralised domain with the objective of understanding the

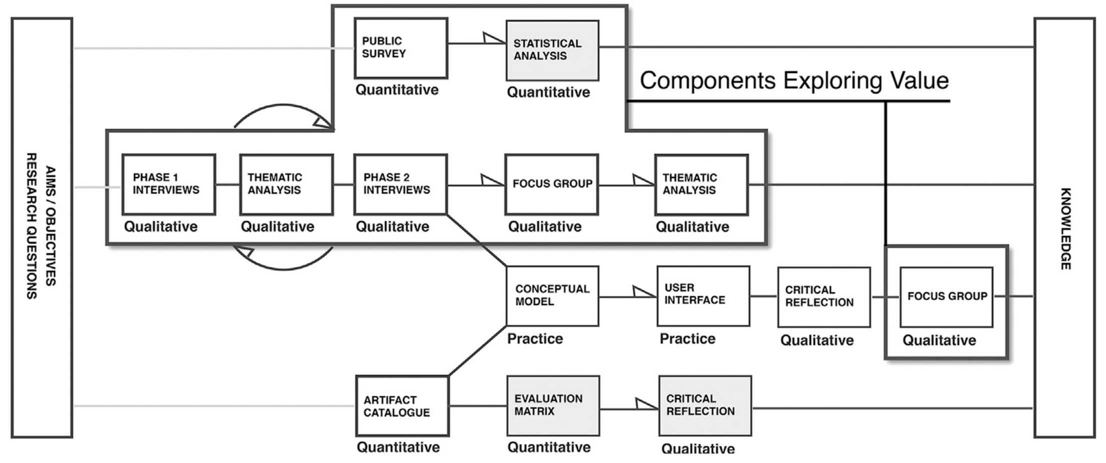


FIGURE 1 | Diagram mapping the components of the broader doctoral research highlighting elements relative to value proposition..

trend and direction of travel, the technological usability barriers and obstacles, and views around sustainable adoption. A second phase of interviews focused on individuals from the decentralised domain with an active interest in Self-Sovereign Technologies. These interviews are narrower in scope and focused specifically on user interaction and adoption. A third phase related to data gathered from a focus group conducted as part of the practice led component of the wider doctoral research. As this data had value in the context of this analysis, it was subjected to, and included in the same analytical process.

Thematic Content Analysis

A qualitative analytic method was required to make sense of the data gathered through semi-structured interviews. Thematic Content Analysis was selected as it offers an accessible and theoretically flexible approach (Braun and Clarke, 2006). The method generally consists of the *'identifying, analysing, and reporting patterns (themes) within data'* (p. 6), and requires the development and application of codes to the data. The coding develops through convergence and grouping into defined themes. Braun & Clarke describe two levels of themes: Semantic and Latent. Semantic themes emerge through the analysis of the data without drawing inferences beyond what a participant has said. Latent themes are developed by moving the analysis beyond the surface, examining and interpreting the data at a deeper level. Braun and Clarke state the importance of defining the theoretical framework through which the data can be considered, the theoretical framework section of this paper, highlights the main discourse around which the thematic analysis has been formed.

RESULTS

The following section first presents the pertinent results of the public survey, the section then communicates the results of the thematic content analysis of Expert Interviews and Focus Group.

Significant Survey Results

A public survey was administered through an Internet mediated questionnaire in line with the defined methodology and survey method plan. In total n 295 surveys have been completed. 62% of participants were male, while 34.6% were female. The age of participants resulted in 52.5% aged 21 and under, 20.3% aged 22–34, 12.9% ages 35–44, and 9.8% being aged 45–54, and 3.4% being 55 or above. Participants were drawn from both a varied student population, and professional and non-professional occupations.

Descriptive Statistics of Significance

Q28 What concerns you most about sharing your personal data?

The results of this individual question are significant, with 68.5% of participants citing concerns that they don't have control over how their personal data is shared. The concept of *Control*, as a means of communicating privacy harms and the risks associated with the sharing of personal data, has been highlighted repeatedly across this research. The notion of Control is powerful, and this result supports the argument that the narrative of *Being Controlled*, should form part of a communication strategy to drive adoption of decentralised technology.

Q37 Which sector do you trust the most with your personal data?

The results of this individual question are significant, with 38% of participants voicing Financial and 34.1% Public Sector. This result is similar to that found within the *Catapult, Digital Trust in Personal Data* survey (Catapult, 2016) which resulted in Public Sector 43% and Financial Services 28%. It can be claimed that both areas are favourable focal points for initial product development and adoption strategy.

Q38 Which one of the following would most convince you to share your personal data?

The results of this individual question are significant, with 58.2% of participants citing *Improving Society* as a motivational driver. This result is similar to that found within the *Catapult Digital Trust in Personal Data* survey (Catapult, 2016) which resulted in 42% opting for societal gain. Arguments for the affordance of privacy rights and the benefits of data sharing for society are a central argument for decentralisation (Solove, 2008; Pentland, 2012; Van Kleek and O'Hara, 2014; Schneier, 2015; O'Neil, 2016; Monbiot, 2017). The academic arguments aligning with the position of the general public, present a primary direction for product development, and a strong narrative for adoption strategy.

Q43 Have you ever been a victim of what you would consider a fraud, breach or an abuse of personal data?

With a result of 71% of participants answering 'No', a central justification for the adoption of decentralised technology may be absent. The argument that unless a serious data breach has ever been experienced, participants are unlikely to be interested in decentralised technologies has been made on a number of occasions. This is compounded further when we consider that the consequences of the majority of data breaches are financial, for which there is a common understanding that insurances are in place to rectify. Adding to this is the general confused picture held by participants with regards risks and harms, which for many will never become a reality (Jarvis, 2011). This supports arguments around the communication of the positive advantages of decentralisation rather than the negative consequences that the majority may never experience. There is though, the hidden exploitation of personal and collective data, individuals are not aware of, gathering, inference and secondary use (Van Kleek and O'Hara, 2014). The communication of this type of unconscious self-inflicted data disclosure, running alongside the positive advantages of decentralisation, potentially provides a compelling argument for adoption.

Scales of Significance

Understanding the Value of Personal Data: resulted in M = 3.64, from a maximum potential of 5. This suggests a general population with a high perceived understanding of the value of personal data. This result has been derived through a number of questions that explore the process of data collection and the value of data not only to the individual, but also as a broader commodity. The results suggest that the population understands that data is bought, sold, processed and ultimately exploited by capital, and that there is a general awareness of *Surveillance Capitalism*.

Comfort Level with Network Engagement: resulted in M = 2.28, from a maximum potential of 5. The results across the elements of this scale are consistent. Participants expressed views regarding the fairness of personal data exchange for services provided, the amount of control the participant felt, the trust that data would be kept secure, the perception of inferred data, and over all opinion of the practice of data collection. The results would suggest a tolerant population who are marginally disaffected with the current centralised system.

Perception of the Importance of Personal Data: resulted in M = 3.95, from a maximum potential of 5. This suggests a population that is highly conscious of the importance of different data types shared across the network. The consistency of results across elements is split, with perception being high in data disclosure which might be obvious. For example, email, file download, location information and online chat. However, a lesser perception was recorded within engagement which might be argued to be more inferred, browsing patterns, search terms, downloaded applications and times of day online. These results suggest a population who perceive their personal data as important at a surface level, but potentially lack an appreciation of the deeper methods of data analysis. This result is interesting when considered against the arguments made during expert interviews questioning the statistical literacy of the general population.

Effort Made to Protect Privacy: resulted in M = 0.331, from a maximum potential of 1. This is considered to illustrate a low level of engagement by participants to protect their personal data. Other than clearing cookies and browser history, and deleting or modifying Internet posts, little effort would appear to be made. It could be argued that participants are unaware of the spectrum of more obscure methods available but equally, it could be argued, contrasted with the *Understanding the Value of Personal Data* results, that this is evidence of the Privacy Paradox (Norberg, et al., 2007). This is further supported by the results of **Q31** and **Q32,** which both signify that individuals have a strong interest in controlling personal data and an interest in engaging with emergent decentralised technology. However, when asked within **Q33** if current concerns about data privacy would sufficiently motivate participants to actively manage part, or all of their personal data, the answer is contradictory, with 68.9% of participants answering 'No'. Further support is found in the results from, **Q42** When asked: In all honesty, how concerned about the disclosure of personal data are you? Participants concern level seemed to be moderate at M = 2.77, from a maximum potential of 5.

Willingness to Engage Third Parties: combined **Q34** and **Q35** to define a result which indicates the participants' willingness to allow either third party or AI management of personal data. The results indicated a low comfort level with this prospect at M = 2.44 from a maximum potential of 5. This is an important statistic as the efficient management of personal data within a UCDE may ultimately require a degree of automation.

Thematic Content Analysis

In total 26 individuals participated in semi-structured Interviews. A process of *Thematic Content Analysis* was then undertaken (Braun and Clarke, 2006), supported by a clearly defined theoretical framework and informed by the results of the public survey. All three stages of data gathering have been transcribed. Transcriptions were then coded through a number of cycles of generation and combination. In total 48 codes were generated. Once coded a process of memoing was undertaken. Collections of interview quotations associated with codes were printed, and the process was conducted manually. Through this endeavor a significant number of themes and sub-themes were identified. Themes have been categorised into three core areas, Adoption, Interface and Broader Themes. In total, 64 themes have been defined and are listed in **Figure 2**.

Each theme is supported by a description. It is impractical to convey the detail within the confines of this paper. The full list of themes and descriptions have been provided as a **Supplementary**

Material. It is recommended that the reader considers this document before proceeding to the following discussion section.

DISCUSSION

The following section endeavors to distil the results of the public survey, the thematic analysis of interviews, and relevant literature, to establish the pertinent topics relating to value proposition and adoption of decentralised technologies through Self-Sovereign Identity.

Marketing Privacy is Not Enough

A dominant theme throughout the expert interviews, and indeed a seminal pillar of this research, is the value of decentralised technology and how this is embedded within artifacts and communicated to participants. The communication and understanding of value are critical to the preliminary stages of the *Diffusion of Innovation* (Rogers, 1962). The dominant narrative for the adoption of decentralised technology is privacy. Expert opinion has clearly stated a position that the decentralised Internet cannot be marketed solely on the fact that it is decentralised. It can in turn be argued that individuals don't perceive the value or context of privacy, and subsequently don't see the advantages of switching to technology that offers little more, or indeed less functionality than their centralised counter parts. The literature describes privacy as a complex and misunderstood concept. It is clearly difficult for individuals or indeed academics to define and contextualise as an overarching concept, and this is repeatedly argued in the literature (Thomson, 1975; Post, 2001; Solove, 2008). Jeff Jarvis (2011) describes concerns regarding privacy on the Internet, as a *'confused web of worries, ill-conceived, and unjustified'* (p. 9). Danial Solove argues that privacy is an umbrella term for intrusions in a myriad of contexts across a spectrum of cultures and social norms (Solove, 2008). Solove suggests a bottom up approach based on a taxonomy of privacy harms, through the notion of family resemblance, in order to clearly define and understand privacy concerns within the digital domain. It would appear that this theory offers a starting position from which to consider the specific domain of network data privacy, through which one might identify privacy infringements, emergent advantages, and the potential benefits and innovations of a decentralised model. There are many other factors compounding the participants' perception of privacy harms in the context of a decentralised Internet. As a participant commented during the expert interviews *'In the West, we have just enough privacy'*. Meaning direct individual privacy infringement is either misunderstood or tolerated and has not yet reached a point of comprehendible harm. There are arguments concerning changing social norms. Campbell and Carlson (2010) suggest an acceptance and apathy toward privacy issues, and Cohen (2012) has argued that the concept of privacy is becoming old fashioned. Zuboff (2015) argues that an acceptance of Surveillance Capitalism is now seen as necessary in order to achieve an effective life. Ian Brown (2013) argues that *Immediate Gratification Bias* and the *Privacy Paradox*, are demonstrations of individual actions and cognitive biases that lead to *'non-optimal privacy decisions by individuals'* (p. 13). The evolving landscape is arguably perpetuated and indeed orchestrated by those holding power. O'Hara's (2013) rebuttal of *Zuckerbollocks* shines light on the power of influence, as arguments are made for the justification and disruption of social norms relating to privacy.

This research highlights the excepted position that privacy is a vague concept that is generally misunderstood and poorly defined. This research suggests that privacy in the decentralised domain is no different and that a systematic analysis following the principles defined by Danial Solove (2008) should be undertaken. In doing so it is expected that a deeper understanding of the decentralised domain can be established, that the real privacy issues within it can be defined, that solutions can be developed, and that it may lead to clearly defined value propositions.

Privacy, A Primary or Secondary Concern?

Throughout the expert interviews, there is a sense that the dominant concept of privacy, as a justification for engaging with decentralised technologies, may be masking other potential value propositions and positive narratives. Indeed, privacy may become a secondary concern or positive consequence of decentralisation. If Danial Solove's (2008) position is to be considered and privacy is seen as an umbrella term instead of a definitive catchall definition, arguments might be built through the taxonomy of privacy to communicate specific privacy problems and the solutions offered by decentralisation. At the same time it recognised the benefits offered through decentralised innovations. It can be argued that this is not an issue of whether privacy is relevant or not, rather this is an issue of semantics in the communication of value proposition. In some situations, the narrative will be focused around privacy protection, but in others, the narrative will be framed around positive innovation, opportunity, friction reduction and new interaction models.

Building A Message

When considering the communication of value within the decentralised domain, research suggests that this falls into two categories: arguments against privacy infringement, and arguments defining the advantages and potential innovations decentralisation supports.

Interviews suggested a need for a consistent narrative, to communicate the justification of decentralisation. A significate theme is that of control, that people don't understand or indeed care little for the concept of privacy, but that when people realise, they are being controlled, it is something very different. The literature provides a foundation for the further exploration of the mechanisms and methods of control. This is evident in the concept of the *Panopticon*, (Bentham, 1791; Himmelfarb, 1968), the concept of control being metered in the mind (Foucault, 1975), and the notion of *Social* and *Panoptic Sort*, (Lyon, 1993; Gandy, 1996). These arguments of control and subsequent exploitation are drawn into the digital realm, and to the depths of Marx's theory, through the *Prosumer Proletariat*, with notions of class, exploitation and surplus value (Fuchs 2012). The narrative of resisting being controlled offers a clear means of expressing a rationale for adoption, which may potentially strike more resonance with the average participant then the notion of privacy.

An additional powerful message is that of failing to benefit from the innovations and opportunities decentralisation potentially offers. This is supported in the literature. Hall argues *'how vital the sharing of personal data is in technological, and specifically, digital innovation'* (Hall, 2016, p. 03). Van Kleek argues that we are jeopardising the realisation of Web 3.0 technologies (Van Kleek and O'Hara, 2014). Pentland highlights the potential, positive societal impacts, if we can move from data based on beliefs, to data based on behaviors (Pentland, 2012). This research suggests that the decentralised community should be looking positively forward to the innovation's decentralisation offers, to identify the emergent value through which to build a positive narrative. Indeed, interviews highlighted great frustration that the *'Decentralised Brigade'*, have to a degree highjacked the argument, focusing primarily on a vague battle for privacy with the objective of reversing the status-quo.

In summary this research suggests two core strands for a decentralised communication strategy, the notion of being controlled, and the significant benefits and missed opportunities of decentralisation.

Finding Value in Decentralisation

Throughout the expert interviews, there has been significant debate, regarding what decentralised innovation may offer. The themes generated from these conversations are valuable, as they act as an inspirational catalyst for innovation. In addition, they form compelling narratives through which value can be established to promote adoption. The themes are broadly divided into three areas: the individual, commerce, and society.

For the Individual

It is argued that decentralised models, which provide agency through reusable and verifiable personal data, offer considerable advantages. A prominent theme is that of streamlining and acceleration of daily transactions, reducing friction, and making it easier to complete tasks. Gaining control over federated identity currently controlled by third parties, is another notable example. The Identity that you invest in, that is developed and refined over time has great value and should belong to its subject and not indefinitely held by a third party. The power of federation, or redistribution of personal information, on the user's terms, is a powerful mechanic of decentralisation.

The concept of empowerment is a compelling idea. Participants' controlling their digital presence, using the validation of identity, verifiable credential and mechanisms of negotiation and contract, form a powerful message that a decentralised Internet delivers the same agency in the digital realm, as that experienced in the real world. This empowerment manifests from the capability to communicate with anonymity, through to the means to avoid echo-chamber and political manipulation, the concept of a Sovereign Boundary Mechanism, and the metaphorical ring of steel between the participant and the network. Collectively these ideas can be woven into persuasive metaphors and value statements.

A significant digestible example of empowerment is Vendor Relationship Management (ProjectVRM, 2019). The principles of VRM are predicated on the rebalancing of the current asymmetric relationships between participant and vendor, freeing the participant from contracts of adhesion across a spectrum of transactions. This is a powerful narrative, re-decentralising through a peer-to-peer model goes beyond privacy protection, and arguably presents an array of opportunities for individuals to transact independently within a rebalanced landscape.

The cost savings for a free agent on the network is another notion that might build a persuasive message. During an interview the comment was made: *'individuals simply don't understand just how much surveillance capitalism is costing them'*. If this could be quantified, in real terms, it would constitute an immediate understandable value proposition.

In summary, the notion of streamlining, the ownership of identity, and the power of federation, the prospect of empowerment and the rebalancing of relationships with vendors, offer a collection of themes around which to build individually focused value proposition. If this is wrapped in the narrative of emancipation from a controlling and manipulative dominant force, it provides a powerful argument, more so than the vague prospect of privacy protection alone.

Societal Gain

Societal gain, as an understandable justification for adoption, is a central narrative that was discussed at great length during expert interviews and focus groups. The importance of privacy for the well-being of society is well documented in the literature (US-Gov, 1973; Gavison, 1984; Solove, 2008; O'Hara, 2016). Our ability to protect the vulnerable, improve health and social care, as well as education and the efficiency of public services are all components of a functional society that will benefit from open sharing of personal data. Silverman expresses concerns about our trajectory of travel and our lack of understanding regarding the social benefits of privacy (Silverman, 2017). At a macro level, the argument that we need to safeguard our democracy (Grassegger, 2016; O'Neil, 2016; Monbiot, 2017), build a healthier society and support adolescent development by maintaining the ephemeral (Schneier, 2015), offer a further dimension for '*the benefit to society*' argument. Indeed, the concept of societal gains aligns with the arguments of Danial Solove (2008) that any granting of privacy rights should be afforded if it benefits society. The results of the Public Survey have illustrated the favored motivation for the sharing of personal data as societal gain. It can be argued that the rewards for a functional, open, decentralised mechanism are clear, and a narrative can be framed in terms of the missed opportunities facing a society locked into a centralised model.

For Business

Positive sentiment was held across the majority of experts consulted with regards the potential benefits to commerce decentralisation offers. A functional Human-Centred Data Ecosystem is considered to offer significate opportunities for new business models and efficiencies. Chaudhry et al. (2015)

argues that the locking in of network participants is *'preventing the formation of a truly competitive market'* (p. 1). Levine expresses a view that the Internet could provide an environment which resembles the vitality of an ancient bazar (Levine, 1999). Searl's (2012) argues that the internet makes *'obsolete, the Industrial Revolution business models of mass marketing, and mass media'* (p. 159). In a relatively short period of time, the Internet has gone from an open marketplace of thousands of individual businesses, to businesses that are forced to engage with, and or go through one of four major players. There would seem to be a great appetite to break these monopolies, and release commerce from being forced to operate through controlled mechanisms. It is argued that this provides opportunities for established larger organisations, but more importantly, acts as a leveller for smaller operations and entrepreneurial endeavor. Indeed, many of the potential models for innovative business through decentralisation have previously been conceptualised and developed, to a degree through the principles of VRM (Vendor Relationship Management). With the advent of a functional identity layer, many of these concepts would now seem to be within grasp. During interviews, a number of specific ways decentralisation might offer value to commerce were voiced. These include: the removal of back room costs, reduction in friction, off-loading the responsibility of data holding, the prospect of real-time high-quality data marketing intelligence, and the competitive advantage of direct trusted relationships with customers. As well as clear advantage for business, the related notion of emancipation from the current centralised model, and the cost savings, offers a valuable marketing message for both vendor and consumer.

The Cultural Context and Niche Pockets of Value

This discussion falls into two strands, the cultural context of decentralisation and the recognition of niche pockets of value. The cultural context is important, and in any effort to design, build and disseminate decentralised technology, the consideration of the cultural dimension and its relevance to any overarching strategy is critical. The notion of strategy in this context, relates to designing decentralised tools and services, that are aligned with the requirements and worldview of a recognised culture. This research suggests that identifying a cultural niche, may offer an opportunity to realise adoption. If the overall community objective is to achieve a critical mass for a global ecosystem, identifying genuine cultural need, with lower barriers to entry, and targeting these domains first, raises the probability of realising a sustainable ecosystem. This notion aligns itself with Moore's Technology Lifecycle Theory (Moore, 1991) where in order to gain adoption, identification of niche markets is required.

During the expert interviews, the argument was made that in a western liberal democracy, we currently enjoy just enough privacy, and care little enough to see the value in decentralised services. This is supported by the theories of the *Privacy Paradox* (Norberg, et al., 2007), and *Instant Gratification Bias* (Acquisti, 2004). But equally, other arguments are made, with German society identified as a group that values privacy highly in a family context. Points have been made regarding community groups that sit outside the mainstream, countries that don't enjoy the same levels of democracy and freedoms, peoples who are without recognised identity and documentation, the unbanked, refugees and asylum seekers, or those that simply don't proscribe to the established social norms. This research concludes that there is a great deal of work to do in identifying cultural groups, that might benefit from a decentralised Internet outside of the western vain. When considering the varied cultural contexts, a signal standardised ecosystem maybe suitable, but the developed services and applications, and the targeting for adoption is varied.

Unforeseen Barriers of Decentralisation

Pertinent insights emerged through the theme of *Barriers to Adoption* and suggested a number of issues that could be argued to be unforeseen consequences of decentralisation. These issues centered around conceptual barriers, which may emerge once interaction with the network becomes enabled through a Sovereign Boundary Mechanism.

The issue was raised of decentralisation working both ways, meaning once access to extensive personal data becomes normalised, third parties may begin to demand more of it, in order to provide transaction and services. There is a sense that the concept could rebound, leaving individuals increasingly exposed. Debate did not reveal specifics, but this is an interesting angle which requires further study.

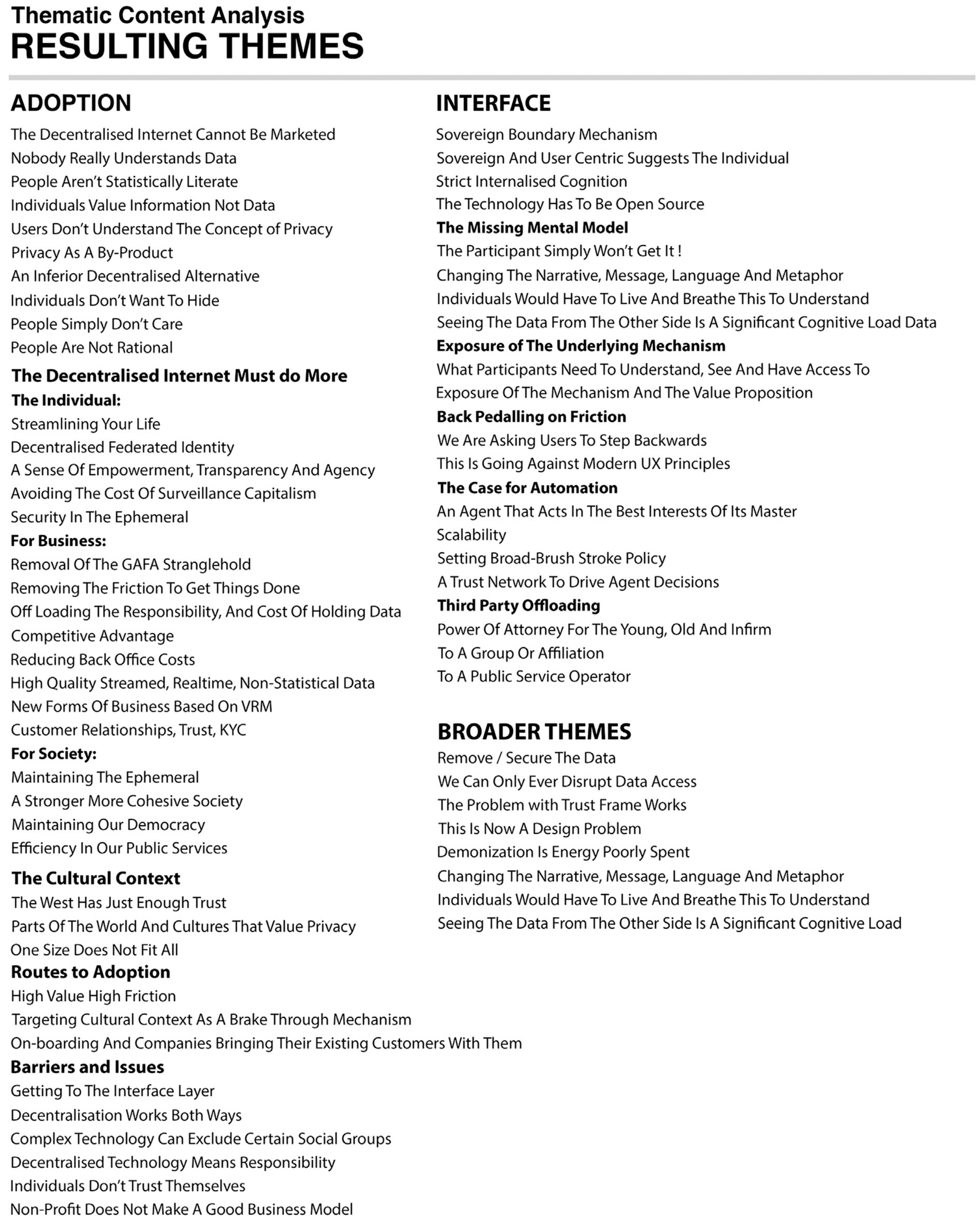
Differing user groups who do not understand the technological concepts or struggle with the mental models may find themselves excluded from the benefits. This topic was heavily debated during the focus groups and is a theme that required serious further consideration. In parallel debate, the concept of responsibility was raised. The issue that taking control over personal data through a Sovereign Boundary Mechanism, defining relationships, making judgments of trust, the monitoring of dynamic transactions, and being ultimately responsible for backup and fail safe, represent a significant on-going responsibility and potential isolation. This was considered to pose considerable friction and potential anxiety. The risk that the participant may lack trust in their own capabilities and competence represents a potential adoption obstacle.

It is important to consider that outside of the primary focus around value proposition and functionality at the interface layer, there are many nuanced variables across differing user groups which need to be further investigated and fully understood.

The Trust Framework

A central component of a Human-Centred Data Ecosystem is a Trust Framework, indeed, a driving organisation behind decentralisation is known as *Rebooting the Web of Trust.*

WOT (2017). There has to be some solid ground so that peers can trust one another over the network. At present trust is facilitated across a string of usernames and passwords, issued



Thematic Content Analysis

RESULTING THEMES

ADOPTION

The Decentralised Internet Cannot Be Marketed
Nobody Really Understands Data
People Aren't Statistically Literate
Individuals Value Information Not Data
Users Don't Understand The Concept of Privacy
Privacy As A By-Product
An Inferior Decentralised Alternative
Individuals Don't Want To Hide
People Simply Don't Care
People Are Not Rational

The Decentralised Internet Must do More

The Individual:
Streamlining Your Life
Decentralised Federated Identity
A Sense Of Empowerment, Transparency And Agency
Avoiding The Cost Of Surveillance Capitalism
Security In The Ephemeral

For Business:
Removal Of The GAFA Stranglehold
Removing The Friction To Get Things Done
Off Loading The Responsibility, And Cost Of Holding Data
Competitive Advantage
Reducing Back Office Costs
High Quality Streamed, Realtime, Non-Statistical Data
New Forms Of Business Based On VRM
Customer Relationships, Trust, KYC

For Society:
Maintaining The Ephemeral
A Stronger More Cohesive Society
Maintaining Our Democracy
Efficiency In Our Public Services

The Cultural Context
The West Has Just Enough Trust
Parts Of The World And Cultures That Value Privacy
One Size Does Not Fit All

Routes to Adoption
High Value High Friction
Targeting Cultural Context As A Brake Through Mechanism
On-boarding And Companies Bringing Their Existing Customers With Them

Barriers and Issues
Getting To The Interface Layer
Decentralisation Works Both Ways
Complex Technology Can Exclude Certain Social Groups
Decentralised Technology Means Responsibility
Individuals Don't Trust Themselves
Non-Profit Does Not Make A Good Business Model

INTERFACE

Sovereign Boundary Mechanism
Sovereign And User Centric Suggests The Individual
Strict Internalised Cognition
The Technology Has To Be Open Source

The Missing Mental Model
The Participant Simply Won't Get It !
Changing The Narrative, Message, Language And Metaphor
Individuals Would Have To Live And Breathe This To Understand
Seeing The Data From The Other Side Is A Significant Cognitive Load Data

Exposure of The Underlying Mechanism
What Participants Need To Understand, See And Have Access To
Exposure Of The Mechanism And The Value Proposition

Back Pedalling on Friction
We Are Asking Users To Step Backwards
This Is Going Against Modern UX Principles

The Case for Automation
An Agent That Acts In The Best Interests Of Its Master
Scalability
Setting Broad-Brush Stroke Policy
A Trust Network To Drive Agent Decisions

Third Party Offloading
Power Of Attorney For The Young, Old And Infirm
To A Group Or Affiliation
To A Public Service Operator

BROADER THEMES

Remove / Secure The Data
We Can Only Ever Disrupt Data Access
The Problem with Trust Frame Works
This Is Now A Design Problem
Demonization Is Energy Poorly Spent
Changing The Narrative, Message, Language And Metaphor
Individuals Would Have To Live And Breathe This To Understand
Seeing The Data From The Other Side Is A Significant Cognitive Load

FIGURE 2 | Theme category's and subthemes.

through various degrees of verification, centralised organisations federating loaned identifiers, and a pyramid of certificate providers. These centralised mechanisms, combined with secure payment services offering a degree of insurance, establish an acceptable level of trust that allows interaction and transaction. If the Internet is to move to a decentralised model, the evolution and mechanisms of trust need to be considered carefully, to establish what is an acceptable and functional level of anchorage across differing kinds of transaction. The distributed ledger is one part of the equation, providing a means to prove control over encryption keys and identifiers: It is a way of verifying credentials through digital signatures and establishing agreements through smart contracts. But where is the anchor? How does one verify a credential, an identity or a reputation? One answer is to seed identity from state or corporate sources. Such as a personal credential issued by a commonly known root identifier, for example the driving licence association or a passport issuer. Identity may be seeded by corporation or financial institution, such as a public service provider or bank. It may be that biometrics come into play, for example physical identity shops, an early exemplifier of which is Arkhive (Arkhive, 2016). How does a centralised anchor relate to a decentralised objective? Is this still a centralised model? If the central anchors on which the verification of an identity is built can be retracted without notice, this contradicts the principles of *Existence and Persistence* defined by Christopher Allen (Allen, 2016). An identifier can be persistently controlled by the participant, but the potential verification of that identity is ultimately reliant on a third party. Are there other methods of building trust? Perhaps in the same way as centralised identities are developed overtime, through content, ranking and reputation? Are there existing models for this elsewhere? And is trust even needed when smart contracts can lock in agreement through the notion of *Code as Law?* Many of these questions are yet to be resolved or explored, and there would seem to be a rich stream of research materialising within this area.

Looking Past the Technology, Turning to Design

Throughout this research, supported by conversation during expert interviews, there is a sense that the objective of a decentralised Internet has now moved out of the realm of the purely technical, into the domain of design thinking. Investigation has concluded that the majority of the technical stack layers are now available, and the mechanisms for interaction with a full UCDE are evolving rapidly. This research concludes that the balance of development has now moved into the realm of design. The crafting of value propositions, digital services, interaction, and underlaying narrative, are all elements that can be considered, and resolved through design thinking. The problem space can be considered systematically, and processes can be engaged to develop solutions. It is telling that at the time of writing, December 2019, if we consider the strands published for the MyData.org (2019) conference, there is a great deal of opportunity to hear speakers discuss technology, computer science, ethics, law, and commerce. But there is a clear lack of a dedicated design strand, exploring and identifying the fundamental questions that need to be resolved. Indeed, a contribution to knowledge within this research, is a body of work that will help the design community to understand better the decentralised domain, the opportunities it presents, and the variables and constraints within which new products and services could be developed.

Getting to the Interface Layer

A powerful argument that warrants further discussion is that of *Getting to The Interface Layer*. Any attempts to decentralise the Internet face the issue of access to the literal screen space, that many of the dominant forces have monopolised to a greater or lesser degree. The barriers to overcome are significant. *'Apple'* devices and operating systems are closed and controlled, *'Android'* is in essence open source, but the influence of Google is significant. Most web portals are under the control of the dominant Internet forces, and the power or search and targeted marketing may favor centralised offerings. With the normalisation of network activity moving to smart handheld devices, accessing this interface layer in a sustainable way, needs to be considered in any strategic planning by decentralised advocates. Indeed, anecdotally, a detailed conversation was had during MyData.org (2019) with a senior designer at a globally recognised telecoms provider, who claimed, *'without access to the hardware and the interface layer, without a fundamental change to the interaction model within mobile devices, the prospect of decentralisation is limited'.*

Community Agendas

The conducted interviews, together with conference attendance and the reading of the literature, reinforces the inevitable camps of political perspective, and motivation within the decentralised community. It is interesting to observe these differing, and potentially problematic positions, as attempts are made to define manifesto and realise collective cooperation. For many, the resistance to the dominant Internet forces is almost militant in nature, arguably driven by a negative world view toward capitalism, or an anti-disestablishment and incredulous position toward the state and surveillance. This is contrasted by individuals and organisations, who see the commercial opportunities of decentralisation, and are focused on capitalising from models of limited sovereignty with a semi open ecosystem. There are other groups who see the missed opportunities of Big Data and the social advantages a data driven society has to offer. And there are those with a passion for technology, who are motivated through the building of new innovations, standards and infrastructures. The following examples illustrate a selection of these positions.

The MyData organisation defines its objective as: '*To empower individuals with their personal data, thus helping them and their communities develop knowledge, make informed decisions, and interact more consciously and efficiently with each other as well as with organisations'.* (MyData.org, 2019). The MyData position is reasonably neutral, but might be argued to be more activist led, with a focus toward social responsibility. In contrast *BlockStack*, is a company that is clearly focused on a market share. It aims to be first to the table with a semi open ecosystem, offering Identity, Distributed Storage, and a DAPP 'Decentralised Application' marketplace (BlockStack.org, 2018). *Sovrin* and its associated company *Evernym*, would seem to be focusing on the bigger picture, publicly building infrastructure, while at the same time developing peripheral business models through commercial tools and agent and wallet software that participants will later require (Sovrin, 2017; Evernym, 2018). Finally, projects *'Veres One'* (2018) and *'Uport'* (2018), would seem to be purely technology and developer focused, with little evidence yet of practical application.

This research suggests that the realisation of a sustainable Human-Centred Data Ecosystem, is unlikely to be achieved by one organisation or individual, and will instead require coordination, and collective effort. But this may prove challenging in a community of tribes with conflicting agendas. This research does not take a position on this issue, nor does it offer a solution. This is an observation that we may need to be mindful of, when considering overall strategy, and offers an interesting landscape for further research.

The Need for A Cohesive Strategy

Following on from the discussion concerning community agendas, the need for a cohesive strategy would seem to be evident. There are a great many stakeholders who believe in the benefits of a decentralised Internet. The first wave of concepts, applications and the technology infrastructure are beginning to materialise, many are driving to be first to market with solutions through semi decentralised architectures. Others are attempting to develop a full ecosystem, which once established, provides a foundation for commercial opportunities. In trying to develop something which is arguably a paradigm shift against a powerful monopoly, it could be argued that a cohesive decentralised community strategy is required. To rely on individual break through, or a serendipitous moment is not enough. A cohesive strategy, standardised methods, seeded trust frameworks, targeted opportunities and establishing consistent narrative, are all examples of how collective endeavors will increase the probability of achieving a sustainable ecosystem.

CONCLUSION

This research concludes that the concept of privacy, in the context of a decentralised Internet is poorly defined and miss-understood. That participants desire privacy, but struggle with it as a concept and fail to see its value across context and cultures. Privacy as a justification for adoption should not be seen as the primary message and instead the privacy benefits of decentralisation are potentially a second order consequence. This research concludes that privacy should be considered as an umbrella term, and that innovations should identify and focus on the specific problems and frictions posed by the centralised model. A decentralised Internet facilitated through Self-Sovereign Identity cannot be marketed on the fact that it is decentralised. Instead the innovation needs to supersede the centralised model in order to raise the probability of adoption. This research concludes that value can be developed by looking progressively forward, exploring concepts that go beyond a centralised model, focusing on the advantages and innovations that will emerge through a functional identity layer and its peripheral mechanisms. A preliminary investigation has highlighted potential pockets of value based around the individual, society and commerce.

This research concludes that the current trajectory of Self-Sovereign Identity results in a standardised collection of interactions defined as a Sovereign Boundary Mechanism. It argues that a major barrier to the adoption of an SBM is the proportion of internalised cognitive process and understanding needed for initial engagement, coupled with a number of additional unforeseen frictions. If adoption is to be realised this friction needs to be recognised, analyzed and systematically reduced.

This research suggests that a cohesive strategy is required by the SSI community in order to achieve widespread adoption. It needs to be one which collectively identifies and develops offerings of value through design thinking, while defining a consistent narrative and language to deliver targeted solutions within cultural contexts. Ultimately, adoption will require the balancing of cognitive load at the interface layer with genuine value proposition, and if this can be achieved, the raise of Self-Sovereign Identity, the development of the Sovereign Boundary Mechanism and the realisation of a Human-Centred Data Ecosystem is indeed inevitable.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research, Innovation and Academic Engagement Ethical Approval Panel. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

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

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbloc.2021.611945/full#supplementary-material.

REFERENCES

Acquisti, A. (2004). "Privacy in electronic commerce and the economics of immediate gratification," in Proceedings of the ACM conference on electronic commerce (EC '04), New York, NY, 21–29.

Allen, C. (2016). The path to self-sovereign identity. Available at: http://www.lifewithalacrity.com/2016/04/the-path-to-self-sovereign-identity.html (Accessed January 1, 2020).

Arkhive. (2016). The world's first identity shop Available at: https://medium.com/mydex/press-release-arkhive-by-timpson-the-worlds-first-identity-shop-powered-by-mydex-7c0afd347ebc (Accessed: 8 Feb 2021).

BCG (2012). The value of our digital identity. The Boston Consulting Group. Available at: https://2zn23x1nwzzj494slw48aylw-wpengine.netdna-ssl.com/wp-content/uploads/2017/06/The-Value-of-Our-Digital-Identity.pdf. (Accessed March 22, 2017).

Bentham, J. (1791). *Panopticon: or the inspection house*. London: Kessinger Publishing.

Bershidsky, L. (2016). No, big data didn't win the U.S. Election. Bloomberg. Available at: https://www.bloomberg.com/opinion/articles/2016-12-08/no-big-data-didn-t-win-the-u-s-election (Accessed April 20, 2020).

Braun, V., and Clark, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3 (2), 77–101. doi:10.1191/1478088706qp063oa

Brown, I. (2013). "The economics of privacy, data protection and surveillance,"in *Handbook on the economics of the internet*. Available at: SSRN: https://ssrn.com/abstract=2358392. (Accessed: 06 Feb 2021).

Burns, A. (2006). "Towards produsage, futures for user-led content production," in Proceeding of the 5th international conference on cultural attitudes towards technology and communication. Editors C. Ess, F. Sudweeks, and H. Hrachovec. Australia: School of information technology, 275–284.

Cambridge Analytica. (2017). Available at: https://cambridgeanalytica.org/ (Accessed September 30, 2017).

Campbell, J. E., and Carlson, M. (2010). Online surveillance and the commodification of privacy. Available at: Panopticon.com (Accessed January 9, 2015).

Catapult (2016). Trust in personal data: a UK review. Available at: https://www.digicatapult.org.uk (Accessed January 9, 2018).

O'Neil, C. (2016). *Weapons of math destruction*. Crown Publishing Group.

Cohen, J. (2012). What is privacy? *Harvard Law Review*. 1–24. Available at: http://www.businessdictionary.com/definition/privacy.html. (Accessed September 28, 2017).

Creswell, J. W. (2003). *Research design: qualitative, quantitative and mixed methods approaches*. 2nd Edn. London, UK: Sage.

Evernym. (2018). Evernym | the self-sovereign identity company. Available at: https://www.evernym.com/ (Accessed: 13 Oct 2019).

Foucault, M. (1977). *Discipline and punish: The birth of the prison. (Translated by Allan Sheridan)*. London: Penguin Books.

Fuchs, C. (2012). *Internet and surveillance: the challenges of web 2.0 and social media*. New York, NY: Routledge.

Gandy, O. (1996). "Coming to terms with the panoptic sort," in *Computers, surveillance, and privacy*. Editors D. Lyon and E. Zureik (Minneapolis, MN: University of Minnesota Press), 132–155.

Gavison, R. (1984). "Privacy and the limits of law," in *Philosophical dimensions of privacy. An anthology*. Vol. 89, Cambridge University Press, 346–402.

Grassegger, H. (2016). The data that turned the world upside down. Motherboard. Available at: https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win (Accessed September 30, 2017).

Grassegger, V., and Krogerus, M. (2016). Ich habe nur gezeigt, dass es die bombe gibt. Available at: Das Magazinhttps://www.dasmagazin.ch/2016/12/03/ich-habe-nur-gezeigt-dass-es-die-bombe-gibt/ (Accessed April 20, 2020).

Chaudhry, A., Crowcroft, J., Howard, H., Madhavapeddy, A., Mortier, R., and Haddadi, H. (2015). Personal data: thinking inside the box. Critical alternatives, 1 (1) Available at: https://doi.org/10.7146/aahcc.v1i1.21312.

Hall, W. (2016). Trust in personal data.*A UK review*, London: Digital Catapult.

Herbert, A. (1955). A behavioral model of rational choice. *Q. J. Econ.* 69, (1), 99–118.

Himmelfarb, G., (1968). "The haunted house of Jeremy Bentham," in *Victorian minds*. New York, NY: Alfred A. Knopf. 32–81.

Huffpost. (2010). Google CEO on privacy. Huffington Post. Available at: http://www.huffingtonpost.com/2009/12/07/google-ceo-on-privacy-if_n_383105.html (Accessed September 22, 2017).

IIW (2019). Internet identity workshop. Available at: https://internetidentityworkshop.com/ (Accessed April 7, 2019).

Jarvis, J. (2011). *Public parts: how sharing in the digital age improves the way we work and live*. New York: Simon & Schuster.

Jones, W. (2010). *Keeping found things found*. Burlington, MA: Morgan Kaufmann Publishers.

Kessler, G. (2013). *James clapper's "least untruthful" statement to the Senate*. The Washington Post. (Accessed September 28, 2017).

Kosinski, M., Stillwell, D., and Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proc. Natl. Acad. Sci. U.S.A.* 110, 5802. doi:10.1073/pnas.1218772110

Levine, R. (2000). *The cluetrain manifesto: the end of business as usual*. Cambridge, Massachusetts: Perseus.

Likert, R. (1932). A technique for the measurement of attitudes. *Arch. Psychol.* 22 140, 55.

Lyon, D. (1993). An electronic panopticon. A sociological critique of surveillance theory. *Socio. Rev.* 41, 653–678. doi:10.1111/j.1467-954X.1993.tb00896.x

Moglen, E. (2013). The tangled web we have woven. Vol. 56 No. 2. Communications of the ACM. Available at: http://dl.acm.org/citation.cfm?doid=2408776.2408784 (Accessed January 19, 2014).

Monboit, G. (2017). Big data's power is terrifying. That could be good news for democracy. *The Guardian*. Available at: https://www.theguardian.com/commentisfree/2017/mar/06/big-data-cambridge-analytica-democracy (Accessed September 28, 2017).

Moore, G. (1991). *Crossing the chasm*. New York, US: Harper Business Essentials.

Mortier, R. (2014). Human-data interaction: the human face of the data-driven society. Available at SSRN: https://ssrn.com/abstract=2508051 (Accessed September 10, 2018).

Norberg, P., Horne, D., and Horne, D. (2007). The privacy paradox: personal information disclosure intentions verses US behaviour's. *J. Consum. Aff.* 41 (1), 100–126. doi:10.1111/j.1745-6606.2006.00070.x

O'Hara, K. (2013). Are we getting privacy the wrong way round? *IEEE Internet Comput.* 17, 89–92. doi:10.1109/MIC.2013.62

O'Hara, K. (2016). The seven veils of privacy. *IEEE Internet Comput.* 20 (2), 86–91. doi:10.1109/MIC.2016.34

O'Neil, C. (2016). *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York, NY: Crown Publishers.

Ohm, P. (2010). Broken promises of privacy: responding to the surprising failure of anonymization. *UCLA Law Rev.* 57 (6), 1701–1777. Available at: http://www.uclalawreview.org/broken-promises-of-privacy-responding-to-the-surprising-failure-of-anonymization-2/.

Pentland, A. (2012). Reinventing society in the wake of big data. Edge.org. Available at: https://www.edge.org/conversation/alex_sandy_pentland-reinventing-society-in-the-wake-of-big-data (Accessed September 28, 2017).

Post, R. C. (2001). Three concepts of privacy. Faculty scholarship series. Paper 185. Available at: http://digitalcommons.law.yale.edu/fss_papers/185 (Accessed February 15, 2019).

Poster, M. (1990). "Foucault and databases: participatory surveillance," in *The mode of information*. Chicago, IL: The University of Chicago Press. 69–98.

ProjectVRM. (2019). Available at: https://cyber.harvard.edu/projectvrm/Main_Page (Accessed April 7, 2019).

Chaudhry, A., Crowcroft, J., Howard, H., Madhavapeddy, A., Mortier, R., and Haddadi, H. (2015). Personal data: thinking inside the box. Critical alternatives, 1 (1) Available at: https://doi.org/10.7146/aahcc.v1i1.21312.

Robbins, M. (2017). The Myth that British data scientists won the election for trump. *Little Atoms*. Available at: http://littleatoms.com/news-science/donald-trump-didnt-win-election-through-facebook (Accessed September 30, 2017).

Robins, K., and Webster, F. (1988). "Cybernetic capitalism: information, technology, everyday life," in *The political economy of information*. Editors V. Mosco and J. Wasko (Madison, WI: The University of Wisconsin Press). 44–75.

Rogers, E. (1962). *Diffusion of innovations*. 5th Edn. New York, NY Free Press of Glencoe.

Rosenberg, M. (2018). How trump consultants exploited the facebook data of millions. *New York times*. Available at: https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html. (Accessed: September 2019).

Schneier, B. (2015). *Data and goliath*, New York, US: W. W. Norton & Company.

Searls, D. (2012). The intention economy: when customers take charge, *Linux J.*

Sharot, T. (2011). The optimism bias. *Curr. Biol.* 21 (23), R941–R945. doi:10.1016/j.cub.2011.10.030

Silverman, J. (2017). Privacy under surveillance capitalism. *Soc. Res.: Int. Q.* 84 (1), 147–164.

Solove, D. (2008a). *Understanding privacy*. Harvard University Press.

Solove, D. (2008b). *I've got nothing to hide, and other misunderstandings of privacy*. Solove Post. 745–772.

Sovrin. (2017). Sovrin foundation. Available at: https://sovrin.org/ (Accessed: 13 Oct 2019).

Thomson, J. J. (1975). The right to privacy. *Philos. Publ. Aff.* 4 (4), 295–314. Available at: http://www.jstor.org/stable/2265075.

Turnhout, K. (2014). *Design patterns for mixed-method research in HCI*. New York, USA: HAN University of Applied Science.

US-Gov (1973). *Records computers and the rights of citizens*. Washington, DC: Department of Health, Education and Welfare.

Van Kleek, M., and O'Hara, K. (2014). The Future of Social is personal: the Potential of the personal data store. Social Collective Intelligence: Combining the Powers of Humans and Machines To Build A Smarter Society. 125–158. Available at: http://eprints.soton.ac.uk/363518/1/pds.pdf. (Accessed September 28, 2017).

Weston, A. (1967). *Privacy and freedom*, New York, US: Ig Publishing.

WOT (2017). Web of trust. Available at: https://www.weboftrust.info/ (Accessed April 7, 2019).

Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *J. Inf. Technol.* 30, 75–89. doi:10.1057/jit.2015.5

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.

frontiers in Blockchain

PERSPECTIVE
published: 30 March 2021
doi: 10.3389/fbloc.2021.616779

Blockchain, Self-Sovereign Identity and Digital Credentials: Promise Versus Praxis in Education

Alex Grech[1]*, *Ira Sood*[2]* *and Lluís Ariño*[3]

[1]Department of Media & Communications, University of Malta, Valletta, Malta, [2]Faculty of Management and Business, Tampere University, Tampere, Finland, [3]Chief Information Officer, Universitat Rovira i Virgili, Tarragona, Spain

Blockchain's versatility is primarily due to its immutable and almost indestructible nature. These attributes have caught the attention of researchers and developers interested in applications and environments where the need for the integrity of identity and content are as paramount as the safe delivery and record of transactions. Self-sovereign digital identity in particular is often cited as a human right that nation states need to embrace with as much conviction as education and lifelong learning are considered to be a public good. Although the blockchain has long been identified as an opportunity for driving much-needed change in the core processes of the education sector, use cases to date have been limited in scope and execution, with blockchain advocates and education policy makers seemingly disconnected on fundamental issues such as governance, self-sovereignty, interoperability, choice of blockchain platforms and overall trust in standards and the integrity of the infrastructure. This article is primarily interested in the affordances of the technology as a public good for the education sector. It levers on the lead author's perspective as a mediator between the blockchain and education sectors in Europe on high-profile blockchain in education projects to provide a snapshot of the challenges and workable solutions in the blockchain-enabled, European digital credentials sector.

Keywords: blockchain, self-sovereign identity, digital credentials, DLT, policy, education, identity

OPEN ACCESS

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

Reviewed by:
Prerna Tambay,
Consultant, London, United Kingdom
Richard Tighe,
Oxfam, United Kingdom
Andreas Freund,
Independent Researcher, San Diego, CA, United States

****Correspondence:***
Alex Grech
alex.grech@um.edu.mt
Ira Sood
ira.sood@tuni.fi

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

Received: *13 October 2020*
Accepted: *28 January 2021*
Published: *30 March 2021*

Citation:
Grech A, Sood I and Ariño L (2021) Blockchain, Self-Sovereign Identity and Digital Credentials: Promise Versus Praxis in Education. Front. Blockchain 4:616779. doi: 10.3389/fbloc.2021.616779

INTRODUCTION

One of the goals of United Nations Sustainable Development 2030 agenda is "ensuring inclusive and equitable quality education and promote lifelong learning opportunities for all[1]." With a global pandemic in the mix, capable of effectively delaying and even diluting the existing progress made in this direction, the turn to technology as the salvation of global education systems is palpable. A 2018 UNESCO report addressing the challenges in digital credentialing and recognition fittingly called out the lack of an efficient "one stop shop" universal system with the ability to collect, store, verify and connect educational credentials in a comparable manner across national contexts (Chakroun and Keevy, 2018). Following the European Commission's publication of a JRC[2] report in 2017 (Grech

[1] https://sdgs.un.org/goals/goal4

[2] The Joint Research Center (JRC) is a department (Directorate-General, DG) of the European Commission providing independent scientific and technological support for EU policy-making.

and Camilleri, 2017) that proposed the use of blockchain technology as a potent tool to achieve that goal, a wide array of publications and pilot programs have since gathered steam. Once exclusively circumscribed to the fintech sector, blockchain technology is now identified as a force of change in multiple realms of operation including public sector services such as healthcare, voter identity registration, asylum process management[3] and higher education. In a recent report, the American Council on Education (Lemoie and Soares, 2020) identified three key themes emerging from the intersection of blockchain technology and education: personal data agency, lifelong learning, and the power of connected ecosystems. While lifelong learning has been a recurring theme in the education sector for several years (Ates and Alsal, 2012; Volles, 2016), the concept of personal data agency as achieved via self-sovereign identities is still gaining momentum in academic and policy circles (Wang and De Filippi, 2020). Despite the technological promise of blockchain, several barriers remain that have limited the practical applications to proof of concepts and pilots so far. COVID-19 may yet be the watershed moment in the education sector that will accelerate the drive toward a system of self-sovereign, ubiquitous, affordable, and verifiable credentials powered by blockchain technology. Efforts in this domain are shifting from theory to practice, largely due to the fruition of multiple initiatives, emerging from both public and private sectors.

This article unpacks these concepts and how the "blockchain in education" ecosystem has unfolded in recent years. It focuses on the lessons learnt from case studies where the blockchain has been deployed to re-imagine digital credentials in high-profile pilots in Europe. Whether deployed as experiments or nation-state initiatives, what these pilots have in common are prescribed objectives to enhance learners' self-sovereignty and agency and improve the options for issuers looking for more cost-effective, secure, democratic and trustworthy solutions than those currently available. The article attempts to bring clarity to ongoing discussions on whether decentralized credentialing ecosystems contribute to more robust, scalable and flexible systems than centralized systems; and whether policy makers and citizens should continue to wait for the technologies to mature or look elsewhere for pragmatic technological solutions to long-standing governance issues specific to the education sector, including the interoperability and recognition of learning credentials across Europe.

DECENTRALIZATION, BLOCKCHAIN TECHNOLOGY AND THE PROMISE OF SELF-SOVEREIGN IDENTITY

Decentralization is defined as the ability of an ecosystem to shift power and control from a centralized host to a distributed network (Anderson, 2019). The World Wide Web or Web 1.0 was originally developed as a decentralized platform. Control was soon appropriated by Web 2.0 behemoths who turned it into a two-sided client-server model, with a business hosting an application and users (Soghoian, 2010; Toledano, 2013). Recent experiences of data appropriation and surveillance capitalism have left idealists yearning for a Web 3.0 underpinned by decentralized ecosystems on open platforms.

Blockchain technology[4] went mainstream in 2008 after its elusive founder/s Satoshi Nakamoto conceptualized it in a white paper and later used it to implement the cryptocurrency Bitcoin (Nakamoto, 2008). As one of the first large scale applications of decentralization, the implications of the technology go far beyond its use as the backbone of a cryptocurrency (Wu and Tran, 2018). As a distributed ledger technology with a decentralized protocol that allows the *network* to validate a transaction (as opposed to some central authority), the blockchain holds the same socially empowering promise of the early internet. Our often-misplaced trust in centralized platforms, databases and protocols could perhaps be addressed by a technology that is *trustless by design*—yet allows varying degrees of trust to be built in at the transaction and communication level of the infrastructure itself.

Technologies without a central, controlling authority also tend to be associated with a compelling social value proposition (e.g., M-Pesa). The social value proposition of the blockchain is a composite of a number of intertwined principles (Grech and Camilleri, 2017; Grech, 2018). Of these, *Self-sovereignty*, *Identity* and *Trust* have particular resonance in these uncertain times: self-sovereignty is frequently associated with the right of individuals to own and control their own identity online and be the final arbiter of who can access and use their data and personal information.

[3] https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/2020/06/10/Three+new+CEF+Blockchain+Use+Cases

[4] In its simplest form, a blockchain is a *type* of distributed ledger technology (DLT) where transactions are recorded with an immutable cryptographic signature called a "hash," and then "grouped in blocks." Every new block includes a hash of the previous one, chaining them together—hence a "blockchain." Data in the block cannot be altered or removed, so every transaction exists in perpetuity while the blockchain exists. The distributed electronic ledger functionality also provides a mechanism for a community to record and exchange information. In this community, each member maintains his or her own copy of the information and all members must validate any updates collectively. The information could represent transactions, contracts, assets, identities, or practically anything else that can be described in digital form. Entries are permanent, transparent, and searchable, which makes it possible for community members to view transaction histories in their entirety. Each update is a new "block" added to the end of a "chain." A protocol manages how new edits or entries are initiated, validated, recorded, and distributed. With blockchain, cryptology replaces third-party intermediaries as the keeper of trust, with all blockchain participants running complex algorithms to certify the integrity of the whole. A *distributed ledger* is a decentralized database, distributed across several computers or nodes, managed by multiple participants, without the participation of a central authority. Each node has equal status in terms of authority, without a central authority or server managing the database, so each node can independently maintain and update the ledger and any of the nodes will verify its existence. A *blockchain* is a usually distributed, usually cryptographically assured chain of blocks (the technical term is Merkle tree), whereas a distributed ledger is a database that exists on (i.e., is distributed over) multiple locations (but not necessarily secured on an actual blockchain). Technically, the git version management system is a blockchain, and a RAID 1 hard drive array is a distributed ledger (Basu and Gabbay, 2021).

In a knowledge-based global economy, an ability to state, verify and prove digital identity in a seamless fashion in a hyperconnected webspace is a vital human right, "the key to survival" according to Wang and De Filippi (2020). According to a World Bank report[5] as of 2016, over 1.5 billion people around the globe have no means to prove their identity. According to UNHCR[6], by 2018 70.8 million people had been forcibly displaced due to persecution, conflict, violence, or human rights violations. Voluntary and regular migrations for the purpose of employment and education also require a portable and dynamic identity that can be unequivocally associated with fairly earned credentials (Toth et al., 2003). The lack or loss of identity credentials inevitably subjects minority groups to a unique form of socio-economic exclusion.

The above implies an urgent need for citizens to secure complete ownership over their identities. In 2016, the United Nations launched the multi-stakeholder ID2020 Alliance[7] with the objective of ensuring universal and ethical digital identity for all within the UN's 2030 Sustainable Development Goals, satisfying the four P's: private, portable, persistent and personal. In 2019, ID2020 proposed a "Certification Mark" to those companies that channel their efforts into technologies capable of providing solutions that meet the 4P's criteria.

In practice, there is rarely any provision to create an identity without relying on a third-party provider. Should that provider cease to exist, so do all the identities of all users in that domain: this defeats all the foundational properties of a self-sovereign identity, such as existence, autonomy, ownership and access, and the principle that the user, *and only the user*, must have full control over their identity data in a user-controlled data management facility (Ferdous et al., 2019).

In a much-cited post in 2016[8], cryptography pioneer Christopher Allen described self-sovereign identity as "the next step beyond user-centric identity [where] the user must be central to the administration of identity" (Allen, 2016). Compare this to a traditional identity management scenario, where a user's identity is defined from the perspective of the provider for a specific purpose and is therefore only valid within the domain of that specific provider within that purview (Wagner et al., 2018). Smolenski (2016, 2020) considers self-sovereignty as an attempt to answer long-standing philosophical questions about social personhood. People have all sorts of identities conferred on them in various forms (passports, proof of employment, diplomas) and by various third parties operating as sources of authority (e.g., credentialing bodies). None of these forms can revoke the fact that individuals are the ultimate source of data about themselves: a citizen's identity pre-exists before the conferral of an identity by any third party. The Sovrin Foundation[9], a not-for-profit global consortium working toward building and governing a network of globally acceptable self-sovereign identity, has stated that in any such network the three core tenets of individual control, security and full portability must be met.

The blockchain is frequently cited by ID2020 and the Sovrin Foundation as a prime candidate for decentralized, tamper-free digital identity solutions since several characteristics of the technology comply with the key properties of self-sovereign identity. Blockchain provides a decentralized domain which is not controlled by any single entity, and where data stored in any blockchain is readily available, as "availability property" to any authorized entity or "access property" (Ferdous et al., 2019). An owner of a particular data (an identity data such as Personally Identifiable Information or PII) has full control over it and dictates how such data can be shared with other users within the blockchain domain, thereby satisfying the disclosure property. The discussion around self-sovereign identities and DIDs[10] has also become one of the key areas in generating momentum toward personal data agency (Lemoie and Soares, 2020). A Digital Identity report[11] concludes that "there is enough evidence available to predict increased adoption of DLT/Blockchain digital identity, including schemes developed around Self-Sovereign Identity (SSI) principles" (Goode, 2019). The report predicts that by 2025, 20 percent of total digital ID will be built using DLT/Blockchain technology, increasing from 5 percent in 2020.

BLOCKCHAIN IN EDUCATION: DIGITAL CREDENTIALS COME OF AGE

In February 2020, the U.S. Department of Education's Blockchain Initiative posted a playful post with the title: "Education has a problem? Put a blockchain on it!"[12] The potentially symbiotic relationship between blockchain and education owes much to the self-sovereign affordances of the technology. Blockchain has been associated with the unbundling of higher education (Sood et al., 2020); and from a praxis perspective, with the basic building block of education, the credential. Definitions of credentials tend to be associated with power—with evidence of authority, status, rights, entitlement to privileges, or the like, usually in written form[13]. According to Gallagher (2019), the reputation of an educational institution is dependent on the market value of this credential. Pittinsky (2015) considers the credential as the only form of non-negotiable currency in the higher education

[5] World Bank's 2016 ID for Development (ID4D) report.

[6] UNHCR's 2018 report on global trends in forced displacement is a record high in human history.

[7] https://id2020.org/digital-identity

[8] In his 2016 paper, Allen identified ten principles of self-sovereignty: existence; control; access; transparency; persistency; portability; interoperability; consent, minimization and protection. Also see: https://github.com/WebOfTrustInfo/self-sovereign-identity

[9] See https://sovrin.org/

[10] A DID or a Decentralized Identifier is a globally unique identifier developed specifically for decentralized systems as defined by the W3C DID specification. DIDs enable interoperable decentralized Self-Sovereign Identity management. More info: https://w3c-ccg.github.io/did-primer/

[11] The Digital Identity Report—The Global Opportunities for Verified Citizen & Consumer Digital ID: Market & Technology Analysis and Forecasts 2020-2025. Published in November 2019

[12] See https://medium.com/designing-the-future-of-education-and-workforce/education-have-a-problem-put-a-blockchain-on-it-bc2574826752

[13] https://www.dictionary.com/browse/credential

ecosystem; in practice, this statement applies to all forms of lifelong learning.

Within the education ecosystem, the credential is a representation of the different types of learning acquired by an individual; a composite of accredited formal, informal and non-formal learning outcomes—the set of knowledge, skills and/or competences acquired or demonstrated by an individual after completing a formal, informal and/or non-formal learning process that tends to include an issuing institution (Chakroun and Keevy, 2018; Grech, 2018). The term "credential" is also used to refer to the qualification (transcripts, diplomas, certificates, assessments, badges etc.) that a learner receives from an educational institute after fulfilling a set of pre-defined criteria (Seymour et al., 2015). Although the majority of credentials remain paper-based, electronic or digital credentials are now part of the education vernacular. This turn to digital is also associated with alternatives and the need for latent change in the modus operandi of education institutions; as if digitization is making it possible to transcend the limits of traditional credentials, and address many of the concerns raised by students and employers about education institutions (Chakroun and Keevy, 2018). Digital credentials are therefore not mere functional elements—a form of skill/qualification—but tangible proof of identity or self-sovereign identity (Stokkink et al., 2020), with the "value-added" significance of an educational credential unlocked when it can be *effectively linked* to the sovereign identity of an individual.

Education and identity, both termed as "undeniable human rights" by the UN need to be turned into the cornerstones of resilient, inclusive and equitable systems that ensure these rights for all. The evolution of identity models over the years has been a metamorphic process. Here lies the dilemma. According to Ferdous et al. (2019), the most commonly used identity model at present is the *SILO* or *Isolated User Identity* model wherein each service provider gives the user unique credentials to access their services, which means that each time users want to access a particular service, they need to verify their credentials. This can be seen in effect with most of the internet service providers such as Google, Facebook, Twitter etc. The *Federated model* is employed by private organizations such as HEIs (Higher Education Institutions) or Tax authorities and the *User Centric* model where a dominant service provider (such as Google) can provide access to a host of other service providers pending verification of credentials. None of these models provide the kind of functionalities that would satisfy the conditions of portability, interoperability and user data ownership rights that allow a data owner to own, control and manage their identity without any intermediary. Moreover, with the massification of higher education and increased student mobility, the demographical composition of students has changed rapidly, challenging the notion of 'the traditional student' (Mintz, 2015).

DID is the key element that enables entities (natural persons, legal entities, or things) to interact with services provided by other entities. One entity may have more than one DID, and it will be the owner who will choose with which specific DID he/she wants to interact with other entities (avoiding profiling). A DID by itself says nothing about its owner since it is just an identifier; it is not an identity.

Once an entity has a DID, different data in the form of verifiable credentials provided by third parties can be linked to it. Some of those verifiable credentials—*Verifiable IDs* (VID) - may describe the DID owner's identity attributes (national identifier, name, surname, etc.); while others—*Verifiable Attestations* (VA) - may be just data issued to DID owner (accredited education, university membership, etc.). Plastic credit cards, library cards, driving licenses, national ID, passports, or any other membership cards are daily physical examples of Verifiable Credentials (VCs). Holders are able to share existing selected claims from their wallets to third parties (in the form of *Verifiable Presentations or VPs*). The decentralized native features of the blockchain, without a single point of control, can nevertheless provide an authoritative source of data that different parties may trust. The blockchain can therefore be used to both register and resolve DIDs and public keys that, in turn, allow digital wallets and their owners to communicate and exchange verifiable credentials in a secure way. Registering DIDs will enable natural persons or legal entities to utilize VCs and VPs.

The blockchain infrastructure is ideal for a digital credential ecosystem that supports the issuance, security, storage and verification of learning credentials over time, and across different professional, cultural and geographical contexts (Smolenski, 2016; Grech and Camilleri, 2017; Chakroun and Keevy, 2018). In a truly self-sovereign ecosystem, recipients should be able to control every aspect of their credentials: where they are stored, with whom they are shared, and how they are identified as individuals in the credential. Since personal data and identity is to be shared online, they should own, manage and have the option to choose to share all or parts of their digital credential records in return for access to services they want—without the need of constant recourse to a third-party intermediary to validate or correlate such data or identity to other data[14]. The ability to provide "a single secure record of educational attainment, accessible and distributed across many institutions" is particularly compelling (Sharples and Domingue, 2016), although Grech and Camilleri (2017) assert that the benefits of blockchain in education are best addressed through open implementations of the technology, which utilize open-source software and open standards for data and implement self-sovereign data solutions.

In 2021, praxis in the blockchain and education sector is about pilots in credentials and infrastructure[15]. Blockcerts[16] was the first open standard specifically developed to create, issue, view and

[14] The advantages of a blockchain credentials system over a traditional, centralized, proprietary system include: the co-ownership of records by issuers and recipients; vendor-independent verification; the ability to issue to multiple blockchains; portability; privacy; interoperability; ease of use and scalability (Grech, 2018).

[15] The US Government's Office of Educational Technology manages a Directory of Blockchain Efforts in Education at: https://usedgov.github.io/blockchain/directory

[16] https://www.blockcerts.org/ The initial design for Blockcerts was based on prototypes developed in collaboration by the MIT Media Lab and Learning Machine (now Hyland Credentials)

verify blockchain-based certificates. From its inception, Blockcerts was meant to facilitate a set of common standards for blockchain certification from which interoperability would emerge. Since 2017, high-profile blockchain certification pilots developed on the Blockcerts standard include a nation-state project by the Government of Malta, the Caribbean Examinations Council, the Federation of State Medical Boards (FSMB) and by MIT Media Lab[17]. Open University's Knowledge Media Institute (KMI) is a partner in a number of large-scale projects with practical use of blockchain based credentials in the education and identity domain. Qualichain is a flagship KMI project supported by the European Commission to understand the intersection of blockchain technology with semantics and data analytics performing a dual function of storing and issuing credentials as well as providing a set of more advanced services, including career counseling, intelligent profiling, and competency management (Kontzinos et al., 2020). The university-led Digital Credentials Consortium (DCC) aims "to create a trusted, distributed, and shared infrastructure that becomes the standard for issuing, storing, displaying, and verifying digital academic credentials; [and its] focus is the design of the standard and development of a transparent governance model that keeps the learner's rights at the center" (Digital Credentials Consortium, 2020).

There are a handful of state-funded digital credential initiatives (such as Diplome[18]) as well as private collaborative initiatives (such as Sony Global Education[19], ODEM[20], IBM's Learning Credential Network[21]) that range between being in nascent stages to piloting stages. The majority of blockchain based pilots are taking place seem to be centered around small nation states such as Estonia (eEstonia), Malta (Nationwide Blockcerts) and Switzerland (Blockstack) (Campbell et al., 2018). In 2016, Verbert et al. suggested that blockchain can be used to 'open up the system of scholarly reputation currently associated with academics, and a number of institutions have reported experimenting with blockchain including United Kingdom NARIC (National Academic Recognition Information Centers), PESC (Postsecondary Electronic Standards Council), AACRAO (American Association of Collegiate Registrars and Admissions Officers), CHESICC (China Higher Education Student Information and Career Center), Mozilla and Deakin University.

The European Commission (EC) is investing in the development of techno-legal frameworks suitable for self-sovereign identity between member states. The Connecting Europe Facility (CEF) program is funding a set of generic and reusable digital service infrastructures (DSIs) also known as building blocks[22]. A CEF building block is a collection of reusable specifications, software and services structured in a service offering that serve general concerns of digital (public) services across EU borders and sectors. Europass 2.0, the related European Digital Credentials Infrastructure (EDCI)[23] and eiDAS[24] fall in the scope of creating a space in the higher ed ecosphere where learners may secure, own and share their digital identity credentials in a trusted, distributed, and shared infrastructure.

Probably the most ambitious blockchain infrastructure initiative in Europe is the European Blockchain Services Infrastructure[25] (EBSI) project. Launched in 2019 by the EC together with governments from member states and the European Court of Auditors (having come together as part of the European Blockchain Partnership), EBSI is being built for cross-border government services. The longer-term roadmap is to make EBSI interoperable with other government and commercial blockchain platforms. At face value, EBSI represents an attempt by policy makers to engage with the technology and learn how to regulate it through the simple expedient of using it themselves[26].

EBSI is a public permissioned blockchain, which means that only reputable entities will be able to write to the chain, but everyone will be able to read/verify. Thus, for public permissioned blockchains a governance model will be required (*see Self-Sovereign Identity and the Interoperability of Digital Credentials on the Blockchain section*). EBSI includes a "Diploma Use Case" as one of the four foundation use cases, with cryptographic proofs of digital diplomas stored in a blockchain network. The Use Case is based on the European Self-Sovereign Identity Framework (ESSIF), a pure SSI framework extended and adapted to European values and regulatory frameworks - in practical terms, the eIDas trust framework and the GDPR directive. Under this new SSI paradigm, digital credentials will be issued directly to citizens for storage in wallets that citizens own and control. In the process, recipients secure full control of their identities and data. No personal data will be stored on chain, other than the attestation of the issuance or any other relevant digital credential status changes. Any third party with whom the citizen has shared any credential, will be able to verify both, provenance (for the holder and issuer) and status (valid, revoked, suspended, expired) for the issued digital credential.

The combination of ESSIF principles and mechanisms with ESSIF ensures both *consent and privacy by design*. It will always be the owner (holder) of the digital credential who will: start

[17] Detailed information on Hyland's official website: https://www.hylandcredentials.com/

[18] http://www.cimea.it/en/diplome-in-breve.aspx

[19] https://www.sonyged.com/

[20] https://odem.io/odem-trust-network/

[21] https://www.ledgerinsights.com/education-orgs-ibm-blockchain-credentials/

[22] See https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/CEF+Digital+Home

[23] https://ec.europa.eu/futurium/en/europass/europass-digital-credentials-infrastructure

[24] https://ec.europa.eu/digital-single-market/en/trust-services-and-eid

[25] How the EU is using blockchain to build a citizen-centric European Internet. Access at: https://www.ledgerinsights.com/how-the-eu-is-using-blockchain-to-build-a-citizen-centric-european-internet/

[26] Provisioned as a service (a set of services), EBSI is made up of two main layers: the Core Infrastructure layer and the Use Case Application layer. The Core Infrastructure layers include the Infrastructure (compute, storage and network systems), the Chain and Storage layer (initially provisioned with two concrete blockchain implementations—Hyperledger Fabric and Hyperledger Besu—and data storage capabilities) and the Core Services and Interfaces layer (providing interfaces for on-chain and off-chain services). The Use Case Applications layer provides the business domain contents for specific use cases.

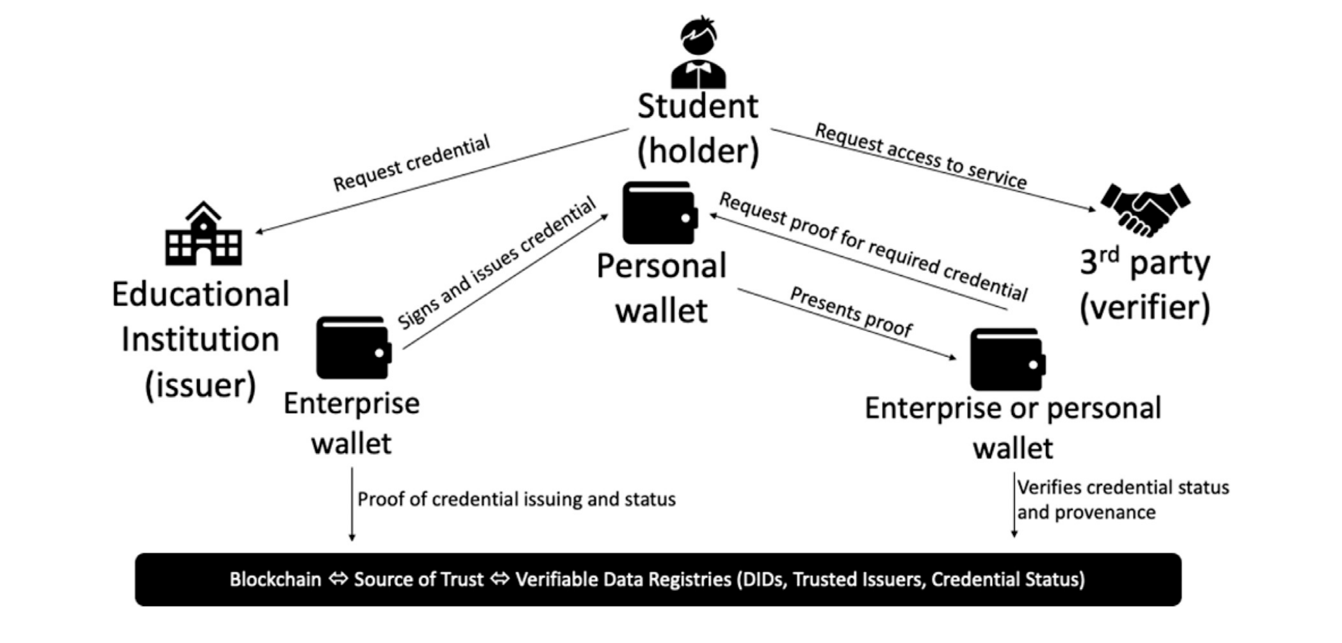

Student (holder)
Request credential
Request access to service
Educational Institution (issuer)
Signs and issues credential
Personal wallet
Request proof for required credential
3rd party (verifier)
Presents proof
Enterprise wallet
Enterprise or personal wallet
Proof of credential issuing and status
Verifies credential status and provenance
Blockchain ⇔ Source of Trust ⇔ Verifiable Data Registries (DIDs, Trusted Issuers, Credential Status)

FIGURE 1 | Main components and flows for an enabled SSI and Blockchain scenario.

interactions with third party services; accept data (in form of verifiable attestations); or share data (in the form of verifiable presentations).

Levering on the design architecture of EBSI, **Figure 1** illustrates the main components and flows when self-sovereign identity principles are applied in a blockchain scenario.

SELF-SOVEREIGN IDENTITY AND THE INTEROPERABILITY OF DIGITAL CREDENTIALS ON THE BLOCKCHAIN

Smit (2020) believes that the most significant benefit of SSI is interoperability, which she states has two different dimensions—a philosophical and a technological dimension. In practice, within an education context, there are four distinct dimensions that need to be managed if interoperability is to be achieved to prescribed standards:

1. *Technical dimension*: Verifiable credentials enable information to be packaged, issued or shared in a standardized format. The following *de facto* and formal standards should be considered to ensure true interoperability: W3C-VC, W3C-VC-EDU, W3C-DID, W3C-JSON-LD Wallet, DIF and the corresponding IEEE and ISO working groups.

2. *Legal dimension*: The two main aspects for consideration are identity and data. For example, in Europe, the eIDas trust framework should be considered to facilitate identity and cross-border validity. In the case of data, GDPR regulation, educational jurisdiction rules and national legislation have to be navigated.

3. *Semantic dimension*: Standardization extends beyond technological interaction and the transmission of data. Interoperability has to do with the seamless receipt of the data package, its opening, and a common understanding of how the fields that make up the data can be read. Both the sender and the receiver need to be using the same semantic model. This is a challenge since there are different (perhaps too many) semantic models to describe a student's learning pathway (for instance PESC or EMREX/ELMO). There is a need for a clearly defined, common model for the accreditation of learning achievements to ensure the portability of both the identity and the record of a student throughout her life. This may well represent an opportunity to differentiate between describing the learning route and the accreditation of the learning achievements obtained during the route. There is a need for a common schema that may describe the accreditation of learning outcomes. This must in turn be capable of: describing any kind of learning (formal, informal, non-formal); recognizing accredited and non-accredited credentials (including micro-credentials); and supporting different learning contexts (from higher education and technical and vocational education and training or TVET to modular learning). In Europe, the Europass Learning Model (ELM)[27] is the data model able to accredit any type of learning outcome achievement, and a correspondence between ELM and ELMO has been provided. The following *de facto* and formal standards should be considered to ensure true interoperability: W3C-VC-EDU, and the corresponding IEEE, ISO working groups.

4. *Governance dimension*: This may be further tabulated as follows:

1) Overall governance dimension: Aspects like the purpose, ownership and responsibilities, decision flows,

[27] The EC also refers to ELM as "the Europass EDCI Data Model."

accountability, communication roles and responsibilities, exit conditions, accommodation for existing solutions, technology standards to be applied, or the type of blockchain deployed etc. (see below).

2) Technical governance dimension: Once there is a decision on the type of blockchain model to be deployed, there will be technical governance aspects to be considered which are conditional on overall governance. For instance: aspects related to the concrete implementation of the blockchain (e.g., Hyperledger Fabric or Hyperledger Besu, etc.); the consensus protocol to be fixed; the minimum nodes required; the level of data segmentation and encryption between nodes; etc.
3) Educational governance dimension: roles adopted by legal entities will enable the educational governance for accredited education (see below).

The governance of digital credentials is also dependent on two critical set of decisions related to the type of blockchain deployed, and accreditation taxonomies:

Decisions on type of blockchain to use:

This is a critical aspect to define for the governance of digital credentials on the blockchain, and very much indicative of the way governance is managed, or perceived to be managed, in specific socio-economic contexts. The type of blockchain selected for digital credentials has much to do with the trust that decision-makers vest in the type of blockchain being used. The choice of blockchain will be made from the following types:

1) A *public blockchain* (so anyone who is connected to the internet can join and become a part of it),
2) A *private blockchain* (so a restrictive blockchain that operates in a closed network),
3) *A permission-based blockchain* (so some accreditation/authorization mechanism to enable roles should exist)
4) *A permission-less blockchain* (so anyone is able to write/update)
5) A *hybrid* blockchain (combination of private and public blockchain than can also be permissioned or permission-less)[28].

Decisions on accreditation taxonomies:

1) Accredited education: in situations where permission attributes need to be issued (such as in the case where a higher education institution issues a degree title to a student). A clear business governance model must be defined, along with the related type of blockchain that is chosen to suit the best business model/requirements.
2) Non-accredited education: in situations where there is no need for permission attributes to be issued (such as in the case of the HR department of a company issuing certificates for completion of an internal professional training course for employees)

The issuing of accredited educational credentials requires clear governance rules for a set of variables. These are likely to include decisions on:

1) An entity that is qualified to host nodes (that is, who hosts mining and verifying nodes).
2) An entity that can authorize legal entities to become Trusted Issuers. To become a Trusted Issuer a legal entity will require authorization from another accreditive source (usually a national quality accreditation agency) that will "accredit" the requesting legal entity to issue certain types of verifiable credentials; and the “accreditive source entity” will be a Trusted Accreditation Organization (TAO). In this context, "accreditation" simply means "to make authoritative, creditable, or reputable".
3) An entity qualified to be a Trusted Issuer (TI): A Trusted Issuer is a legal entity that is accredited to issue certain types of verifiable credentials (such as a Higher Education institution accredited to issue qualifications as defined in the level 7 from the European Qualification Framework).
4) Supporting rules for definition of identity and levels of assurance for entities (natural persons and legal entities).
5) Data schemes to ensure semantic and technical interoperability. The blockchain will provide the source of trust containing at least the following trusted registries to enable business domain governance:

– DIDs registry: contains DIDs and public keys;
– Trusted Accreditation Organization registry: details of the trusted accreditation organization and the “authorisations” it may accredit;
– Trusted Schema Registry: Data schemes;
– Trusted Issuer Registry: Trusted Issuers details and accreditations;
– Revocation & Endorsement Registry: verifiable credentials status (valid, revoked, suspended, expired).

In principle, by addressing all of the four interoperability dimensions and the related issues highlighted in this section, the blockchain solution should be able to support the accreditation of any kind of learning. However, following this process alone will not necessarily fast-track the adoption of digital credentials.

DECENTRALIZED RESILIENT MODELS FOR EDUCATION? TOO EARLY TO CELEBRATE

The impact of the COVID-19 pandemic on the education sector has been devastating. By April 2020, 94% of learners in 200 countries were adversely affected (United Nations, 2020), with

[28] Blockchains will become increasingly nuanced. For instance, Corda is an open source blockchain project, designed for business, with one key differentiator: it does not periodically batch up transactions needing confirmation into a block and confirm them in one go. Instead, Corda confirms each transaction in real-time. There is therefore no need to wait for other transactions to come along or a “block interval.” Transactions are confirmed immediately. This means that the transaction is not dependent on any others, increasing both privacy and scalability. So, Corda is both a blockchain and not a blockchain.

the UN Secretary General deeming school closures "a generational catastrophe" (Farzan and O'Grady, 2020). In countries with already fragile education systems, there are fears that discontinuation might lead to a permanent removal of education services[29]. The pandemic has also exacerbated a latent crisis within education institutions, suddenly exposing precarious business models and resistance to change. Years of debate about the merits of online and blended learning models and OER (Open Educational Resources) vanished in the wake of the crisis, with universities shifting overnight to emergency, remote online teaching (Baker, 2020; Grech, 2020; Mitchell, 2020).

The pandemic has triggered an overall digital transformation and rapid, large-scale change in most higher education institutions. Yet almost 10 years since the technology's inception, with the exception of fintech, there is no industry where blockchain has secured a foothold. Blockchain credentials have not gone mainstream; the blockchain university as envisaged in the Woolf University white paper is stillborn (Gerard, 2019)[30]. The revolution has not quite happened (Baraniuk, 2020). We may tabulate a few reasons[31] for this state of play in the education sector, based on first-hand experience[32]:

Lack of large use case studies: According to Sindi (2019), research on the diffusion of blockchain innovation has not progressed enough due to a lack of use cases within the higher education community. It is not an accident that the "nation-state" initiatives have been piloted by small states with a legacy of trialing emerging technologies, and with ready access to policy makers, ensuring speed in decision-making and political will to cut through "red tape". Gartner observed that most blockchain applications seemed to be stuck in the experimentation mode at the end of 2019[33]. A peer review study published by the Center for Evidence Based Blockchain concluded that "almost half of the blockchain firms show no explicit evidence of the problem to be solved. Approximately one-third fail to cite a comparison and intervention analysis, and less than 2 per cent demonstrate evidence of outcomes backed by filtered (critically appraised, peer reviewed) information" (Naqvi and Hussain, 2020).

Interoperability is rarely just about technology: The real obstacles to the implementation of emerging technologies such as the blockchain "for the public good" lie in the socio-technical integration of rules-based, autonomously operating DLT systems in complex social environments. This is not just about whether end users become data controllers (Van der Bergh, 2018), but often whether a project can deliver the same value across borders and nation state jurisdictions. Technologists tend to develop solutions in ideological silos, with little understanding of the barriers systemic to socio-political environments or the need to secure the buy-in of policy barriers to overcome such barriers. Taking EBSI as a technologically-driven project and idealistically meant to be taken up by EU member states as a public good: for EBSI blockchain credentials to become the EU-standard for education credentials, interoperability authentication and mechanisms need to be determined at the outset with existing EU member state projects and quality assurance and accreditation institutions. That implies seamless technology and member state policy interoperability on issues such as education accreditation and quality assurance and portability of formal and non-formal credentials. The pandemic has led to more nation-state insularity, as opposed to solidarity. Digitally secure educational credentials to facilitate international student mobility are not necessarily on the agenda of nation states.

Self-sovereign identity does not entail individuals certifying their own identity. As long as societies are structured in non-anarchical political systems with well-defined government structures that guarantee and enforce laws while allowing for the establishment of public and private trust frameworks, public administrations will still have the final sovereignty of the identification of citizens. The best self-sovereignty that technology can propose to individuals is not in the issuance, but in the management of their identity (Allende López, 2020).

Resistance from central governance: The inherent resistance to change demonstrated by mainstream institutions is symptomatic of an overall governance and structural issue associated with the hegemonic brick and mortar model of the university (Caruth and Caruth, 2013; Dans, 2020). Fear of decentralization is rife, both at nation state level and particularly in a higher education sector: the blockchain for many higher education institutions implies a threat to "central governance," business models and a loss of power vested in legacy systems and in the HR or Registry departments. When digital credentials have been registered on blockchains, they are not being claimed or used very often: hiring managers and registrars have yet to trust or understand how to evaluate them (Lemoie and Soares, 2020). The same resistance may be found in central governments: it is not to every nation state's liking to trust the trustless public blockchain, open standards et al. The much-lauded Estonian blockchain model is a centralized, militarized version of the technology, not some variant based on open standards and a public blockchain. The analogy of trusted, centralized paper credentials vs. mis-trusted, decentralized, permissionless, digitized counterparts will unfortunately continue to resonate with policy- and decision-makers, until there is a tipping point whereby the interoperability issue described above is 'resolved' by some higher authority—say through prescriptive regulation from bodies such as the European

[29] A recent article from the Economist has cited cases where following lockdowns and quarantine, young girls are consistently being forced into marriage or withdrawn altogether, placing them at risk of never returning to school, available at https://www.economist.com/international/2020/07/18/school-closures-in-poor-countries-could-be-devastating

[30] Woolf envisaged a business model whereby academics worldwide can create and manage a borderless, geographically-agnostic, collaborative university with cross-cultural curricula using some variant of blockchain tokens and smart contracts.

[31] Although we refer specifically to the higher education sector in this paper, most of the reasons we cite could apply to an overall resistance to the adoption of the blockchain in almost any education sector.

[32] The lead author was the architect of the nation state Blockcerts pilot in Malta, and currently a partner in a Horizon 2020 project looking at the impact of emerging technologies on digital education, and a consultant to the European Commission on the EBSI project.

[33] https://www.gartner.com/en/newsroom/press-releases/2019-10-08-gartner-2019-hype-cycle-shows-most-blockchain-technologies-are-still-five-to-10-years-away-from-transformational-impact

Commission or national governments. In our view, the success or otherwise of EBSI in the European education credentials sector is critically dependent on the buy-in of a core set of policy-makers and national quality assurance stakeholders in EU member states, particularly those prepared to explore the implementation of high-profile EBSI pilots, as opposed to the actual technical interoperability of the EBSI infrastructure.

Lifelong learning is still at odds with the hegemony of universities. The discourse about the blockchain's "potential" to disrupt education systems worldwide has been a feature of academia since 2016. Beyond the "low-hanging fruit" of verifiable, tamper-proof education credentials, most studies concur on the opportunity for decentralized systems to facilitate the interoperability of education worldwide, and the mobility, self-sovereignty and lifelong learning aspirations of citizens. Yet many universities cling to the paper credentials as symbols of centralized power. In Malta, for instance, the University of Malta continues to resist joining the nation-state Blockcerts initiative, despite its launch in 2017, citing administration challenges. On the basis of use case studies, there is more enthusiasm for blockchain credentials from the TVET sector and from professional and non-formal institutions (including those exploring micro-credentials) than orthodox higher education institutions.

Open standards are at odds with business returns. Many blockchain credentials initiatives masquerade as "open" but need to be closed for their promotors to secure a return on their investment. Commercial blockchain credential solutions have been developed on the Blockcerts open standard, and then closed to ensure lock-in with the end user. Trust in open standards have frequently been associated with mistrust in the security of the public bitcoin standard. Governments continue to wait for others to take up the baton of blockchain self-sovereignty.

CONCLUSION

The analogy that the blockchain is a hammer looking for a nail continues to resonate, even if the education sector seems to be an obvious nail (Herd, 2019; Singer 2020). Gartner's prediction in 2019 that blockchain technology-oriented solutions could create more than $176 billion worth of business value by 2025 and $3.1 trillion by 2030 seems optimistic in 2021. Blockchains may be presented as the verification of identity across adversarial networks, but the promise of a global, interoperable identity ecosystem is dependent not just on trust in decentralized infrastructure, but on the willingness of nation states to collaborate for the common good. Despite the best intentions of centrally-driven digital identity projects such as EBSI and eiDAS[34], cross-border interoperability needs buy-in from third parties, including standards bodies and policy makers in member states, and within different entities within the Commission itself.

Covid-19 has led to a huge range of human activities migrating online, and far more smoothly than anticipated;[35] for the education sector, the pandemic represents an organic crisis which may drive latent change in the education sector (Heitz et al., 2020). The move to technology-enabled education appears to be as inevitable as more discerning learners questioning the return on their investment in orthodox higher education. The credential will not disappear as long as citizens need to demonstrate identity and skills sets to others. If skills, as opposed to degrees, will really shape the future of work, there may be greater possibilities of the labor market attributing value to digital repositories and mutual recognition of blockchain credentials than traditional bricks and mortar universities trying to cling to outdated business models.

According to Lemoie and Soares (2020), blockchain technology can be applied to advance social equity through personal data agency, lifelong learning, and the power of connected ecosystems. Optimists such as John Domingue at the KMI believe the time has arrived for blockchain to underpin a new resilient decentralized model for lifelong learning where all of the diverse educational experiences available to modern students are tracked, verified and stored as immutable records (Hayward, 2020)[36]. Students will have a self-sovereign student identity where all of their educational certifications are completely owned, controlled, and managed by them, without the need to invoke the support of an intermediary. In troubled times like today with severely contracted economic activity leading to large scale job losses[37], blockchain-backed educational credentials could ultimately create access to job opportunities which would otherwise go unacknowledged. It might be possible to directly impact an individual's ability to find employment; for example, recruitment sites could match vacancies to candidates based on a broader range of experiences as reflected in their student experience collected from a multitude of resources (Kalla et al., 2020; Marbouh et al., 2020).

Technology history indicates that an organic crisis frequently leads to significant innovation and social change. The burst of the Internet bubble and the emergence of social media platforms is a pertinent analogy. Technology and a pandemic are a whole new ball game, and the blockchain can hardly be considered to be a placebo for the ongoing challenges of the global education sector. A return to that most mundane of applications, a "better, self-sovereign education record" (Griffin, 2020) may be a by-product of these troubled times. To regenerate the blockchain project requires much work in the three inter-related areas of regulation, interoperability and human trust frameworks. The technology affordances of the blockchain alone will not suffice.

[34] https://www.biometricupdate.com/202009/european-digital-identity-vision-outlined-by-ec

[35] https://www.economist.com/leaders/2020/09/05/covid-19-strengthens-the-case-for-digital-id-cards

[36] Also see del4all.eu

[37] The impact of the COVID-19 pandemic on jobs and incomes in G20 economies: Report by International Labor Organization (ILO) and Organization for Economic Co-operation and Development (OECD) accessed at: https://www.ilo.org/global/docs/WCMS_753607/lang–en/index.htm

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 authors.

AUTHOR CONTRIBUTIONS

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

REFERENCES

Allen, C. (2016). The path to self-sovereign identity. Available at: http://webcache.googleusercontent.com/search?q=cache:ljAmoHj_bBIJ:www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html+&cd=2&hl=en&ct=clnk&gl=in (Accessed April 25, 2016).

Allende López, M. (2020). *The future of identity: self-sovereignty, digital wallets, and blockchain*. Washington, D.C., United States: Inter-American Development Bank. Available at: http://dx.doi.org/10.18235/0002635.

Anderson, M. (2019). Exploring decentralization: blockchain technology and complex coordination. *J. Des. Sci.* Available at: https://jods.mitpress.mit.edu/pub/7vxemtm3 (Accessed February 7, 2019).

Ates, H., and Alsal, K. (2012). The importance of lifelong learning has been increasing. *Procedia Social Behav. Sci.* 46, 4092–4096. doi:10.1016/j.sbspro.2012.06.205

Baker, K. (2020). Covid-19 is changing education for the better. *Financial Times*. Available at: https://www.ft.com/content/51496fde-98e7-11ea-871b-edeb99a20c6e (Accessed May 18, 2020).

Baraniuk, C. (2020). Blockchain: the revolution that hasn't quite happened. Available at: https://www.bbc.com/news/business-51281233 (Accessed February 11, 2020).

Basu, D., and Gabbay, M. J. (Forthcoming 2021). "Karl Marx and the blockchain," in *Media, technology and education in the post-truth society*. Editor A. Grech (Emerald Publishing).

Campbell, R., Thompson, G., Ferry, P., and Rudman, H. (2018). *Distributed ledger technologies in the public sector: learnings on the application of distributed ledger technologies across international public services and their role in realising Scotland's full potential in a digital world*: The Scottish Government.

Caruth, G. D., and Caruth, D. L. (2013). Understanding a resistance to change: a challenge for universities. *Turk. Online J. Dist. Educ.* 14 (2), 12–21.

Chakroun, B., and Keevy, J. (2018). Digital credentialing: implications for the recognition of learning across borders. UNESCO, Technical Paper ED-2018/WS/29

Dans, E. (2020). The Coronavirus pandemic has unleashed a revolution in education: from now on, blended learning will be the benchmark. *Forbes*. Available at: https://www.forbes.com/sites/enriquedans/2020/04/13/the-coronavirus-pandemic-has-unleashed-a-revolution-in-education-from-now-on-blended-learning-will-be-the-benchmark/#8fadf21536fd (Accessed April 13, 2020).

Digital Credentials Consortium (2020). Building the digital credential infrastructure for the future. white paper. Available at: https://digitalcredentials.mit.edu/wp-content/uploads/2020/02/white-paper-building-digital-credential-infrastructure-future.pdf (Accessed February 8, 2020).

Farzan, A. N., and O'Grady, S. (2020). Closing schools around the world could cause a "generational catastrophe," U.N. secretary-general warns. Available at: https://www.washingtonpost.com/world/2020/08/04/closing-schools-around-world-could-cause-generational-catastrophe-un-secretary-general-warns/ (Accessed August 4, 2020).

Ferdous, M. S., Chowdhury, F., and Alassafi, M. O. (2019). In search of self-sovereign identity leveraging blockchain technology. *IEEE Access* 7, 103059–103079. doi:10.1109/access.2019.2931173

Gallagher, S. (2019). How the value of educational credentials is and isn't changing. *Harvard Business Review*.

Gerard, D. (2019). Woolf, the university on the blockchain — or not. Available at: https://davidgerard.co.uk/blockchain/2019/05/26/woolf-the-university-on-the-blockchain-or-not/.

Goode, A. (2019). The digital identity report—the global opportunities for verified citizen & consumer digital ID: market & technology analysis and forecasts 2020–2025.

Grech, A. (2018). "Blockchain's early promise: self-sovereignty for holders of verified credentials," in *Connections, November 2018* (Vancouver, BC, Canada: Commonwealth of Learning, Vol. 23, No. 3, 16.

Grech, A., and Camilleri, A. (2017). Report No.: EUR 28778 EN. Blockchain in education. joint research centre, European commission. Available at: https://publications.jrc.ec.europa.eu/repository/bitstream/JRC108255/jrc108255_blockchain_in_education(1).pdf (Accessed July 26, 2019).

Grech, A. (2020). Teaching online in the time of Coronavirus. Available at: https://connectedlearning.edu.mt/teaching-online-coronavirus/ (Accessed April 6, 2020).

Griffin, A. (2020). Can blockchain technology provide A better learner record?. *Forbes*. Available at: https://www.forbes.com/sites/alisongriffin/2020/09/16/can-blockchain-technology-provide-a-better-learner-record/#5569479114c1 (Accessed September 16, 2020).

Hayward, A. (2020). Coronavirus is driving education onto blockchain. with the traditional learning model disrupted, a bright idea for decentralized, digital certificates is gaining traction. Available at: https://decrypt.co/24785/coronavirus-is-driving-education-onto-blockchain (Accessed April 13, 2020).

Heitz, C., Laboissiere, M., Sanghvi, S., and Sarakatsannis, J. (2020). Getting the next phase of remote learning right in higher education. *McKinsey Insights*. Available at: mckinsey.com/industries/public-and-social-sector/our-insights/getting-the-next-phase-of-remote-learning-right-in-higher-education (Accessed April 23, 2020).

Herd, C. (2019). Why blockchain is a hammer looking for a nail and where it might find it. Available at: https://medium.com/swlh/why-blockchain-is-a-hammer-looking-for-a-nail-and-where-it-might-find-it-bc15faf11e21 (Accessed July 8, 2019).

Kalla, A., Hewa, T., Mishra, R. A., Ylianttila, M., and Liyanage, M. (2020). The role of blockchain to fight against COVID-19. *IEEE Eng. Manag. Rev.* 48 (3), 85–96. doi:10.1109/emr.2020.3014052

Kontzinos, C., Kokkinakos, P., Skalidakis, S., Markaki, O., Karakolis, V., and Psarras, J. (2020). "Using blockchain, semantics and data analytics to optimise qualification certification, recruitment and competency management: a landscape review," in Proceedings of the 12th international conference on mobile, hybrid, and on-line learning (eLmL 2020), Valencia, Spain, November 21–25, 2020. Editors A. Mikroyannidis, M. Chang, and S. White (IARIA, 2020), 44.

Lemoie, K., and Soares, L. (2020). *Connected impact. Unlocking education and workforce opportunity through blockchain*. Washington, D.C., United States: American Council on Education (ACE). Available at: https://www.acenet.edu/Documents/ACE-Education-Blockchain-Initiative-Connected-Impact-June2020.pdf.

Marbouh, D., Abbasi, T., Maasmi, F., Omar, I. A., Debe, M. S., Salah, K., et al. (2020). Blockchain for COVID-19: review, opportunities, and a trusted tracking system. *Arab. J. Sci. Eng.* 45, 9895–9911. doi:10.1007/s13369-020-04950-4

Mintz, S.. (2015). Who are our students?. *Inside Higher Ed.* Available at: https://www.insidehighered.com/blogs/higher-ed-beta/who-are-our-students (Accessed June 26, 2020).

Mitchell, N. (2020). What have we learned and how will HE change post COVID-19?. *University World News*. Available at: https://www.universityworldnews.com/post.php?story=20200626100952531 (Accessed June 26, 2020).

Nakamoto, S.. (2008). Bitcoin: a peer-to-peer electronic cash system. Available at: http://www.Bitcoin.Org.

Naqvi, N., and Hussain, M. (2020). Evidence-based blockchain: findings from a global study of blockchain projects and start-up companies. *J. Br. Blockchain Assoc.* 3 (2), 16795. doi:10.31585/jbba-3-2-(8)2020

Pittinsky, M. (2015). Credentialing in higher education: current challenges and innovative trends. *Educause Review*. Available at: http://er.educause.edu/articles/2015/3/credentialing-in-higher-education-current-challenges-and-innovative-trends (Accessed March 2, 2015).

Seymour, D., Everhart, D., and Yoshino, K. (2015). *The currency of higher education: credits and competencies*. Washington, D.C., United States: American Council on Education and Blackboard website. Available at: https://events.educause.edu/annual-conference/2015/proceedings/the-currency-of-higher-education-credits-and-competencies (Accessed October 29, 2015).

Sharples, M., and Domingue, J. (2016). The Blockchain and Kudos: a distributed system for educational record, reputation and reward. *Adaptive and adaptable learning. EC-TEL 2016. Lecture Notes in Computer Science*. Editors K. Verbert, M. Sharples, and T. Klobuĉar Cham: . Springer), Vol. 9891, 490–496. doi:10.1007/978-3-319-45153-4_48

Sindi, A. F. (2019). Adoption factors of a blockchain digital identity management system in higher education: diffusing a disruptive innovation. Doctoral dissertation. Los Angeles (CA): California State University.

Singer, A. (2020). Blockchain can disrupt higher education today, global labor market tomorrow. Available at: https://cointelegraph.com/news/blockchain-can-disrupt-higher-education-today-global-labor-market-tomorrow (Accessed September 17, 2020).

Smit, A. (2020). Blockchain and self sovereign identity. Available at: https://open.spotify.com/episode/6NzG33IRM0KH3V4CDLIXhm?si=deSDZ9wATcGpxJvN7a2p6g.

Smolenski, N. (2020). Digital credentials and the pursuit of self-sovereignty. Available at: https://www.youtube.com/watch?v=Os859aPLWNQ&feature=emb_logo&ab_channel=CommonwealthCentreforConnectedLearning. Webinar.

Smolenski, N. (2016). Identity and digital self-sovereignty. a new paradigm for sovereignty on the high seas. Available at: https://medium.com/learning-machine-blog/identity-and-digital-self-sovereignty-1f3faab7d9e3#.3jcgvnbok (Accessed September 19, 2016).

Soghoian, C. (2010). Caught in the cloud: privacy, encryption, and government back doors in the web 2.0 era. *J. Telecomm. High Tech. L.* 8, 359. doi:10.31228/osf.io/9qg75

Sood, I., Pirkkalainen, H., and Camilleri, A. (2020). “Can blockchain technology facilitate the unbundling of higher education,” in Proceedings of the 12th international conference on computer supported education (CSEDU 2020). SCITEPRESS, May 2–4, 2020, Vol. 2, 228–235. 978-989-758-417-6. doi:10.5220/0009339202280235

Stokkink, Q., Epema, D., and Pouwelse, J. (2020). A truly self-sovereign identity system. arXiv preprint [Preprint]. Available at: https://arxiv.org/pdf/2007.00415.pdf (Accessed July 1, 2020). arXiv:2007.00415.

Toledano, C. A. (2013). “Web 2.0: the origin of the word that has changed the way we understand public relations,” in Barcelona international PR conference, Barcelona, Spain, July 2–3, 2013.

Toth, K., Subramanium, M., and Chen, I. (2003). Persona concept for privacy and authentication. *Int. Business Econ. Res. J.* 2 (6). doi:10.19030/iber.v2i6.3810

United Nations (2020). UN policy brief: education during Covid-19 and beyond. Available at: https://www.un.org/sites/un2.un.org/files/sg_policy_brief_covid-19_and_education_august_2020.pdf (Accessed August, 2020).

Van der Bergh, R. (2018). Paradigm shifts for the decentralized web. Available at: https://ruben.verborgh.org/blog/2017/12/20/paradigm-shifts-for-the-decentralized-web/ (Accessed December 20, 2017).

Volles, N. (2016). Lifelong learning in the EU: changing conceptualisations, actors, and policies. *Stud. High. Educ.* 41 (2), 343–363. doi:10.1080/03075079.2014.927852

Wagner, K., Némethi, B, Renieris, E., Lang, P., Brunet, E., and Holst, E. (2018). *Self-sovereign identity. a position paper on blockchain enabled identity and the road ahead*. Berlin, Germany: Blockchain Bundesverband, 57.

Wang, F., and De Filippi, P. (2020). Self-sovereign identity in a globalized world: credentials-based identity systems as a driver for economic inclusion. *Front. Blockchain* 2, 28. doi:10.3389/fbloc.2019.00028

Wu, J., and Tran, N. (2018). Application of blockchain technology in sustainable energy systems: an overview. *Sustainability* 10 (9), 3067. doi:10.3390/su10093067

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.

GLOSSARY

AACRAO, American Association of Collegiate Registrars and Admissions Officers.

CHESICC, China Higher Education Student Information and Career Center.

DCC, Digital Credentials Consortium.

DID, Decentralized Identifier.

DIF, Decentralized Identity Foundation.

DLT, Digital Ledger Technology.

EBSI, European Blockchain Services Infrastructure.

EDCI, European Digital Credentials Infrastructure.

EDU, Education

eID, Electronic Identification

eIDAS, Electronic Identification, Authentication and Trust Services.

ELM, Europass Learning Model.

ELMO, European Learning Mobility.

EMREX, Easy Mobility on Recognition of External[38]

IEEE, Institute of Electrical and Electronics Engineers.

ISO, International Standards Organization.

ESSIF, European Self-Sovereign Identity Framework.

FSMB, Federation of State Medical Boards.

GDPR, General Data Protection Regulation.

HEI, Higher Education Institution.

ID, Identity.

JRC, Joint Research Center.

JSON, Java Script Object Notation.

NARIC, National Academic Recognition Information Center.

ODEM, On Demand Education Marketplace.

OER, Open Education Resources.

PESC, Postsecondary Electronic Standards Council.

PII, Personally Identifiable Information.

SSI, Self-Sovereign Identity.

TAO, Trusted Accreditation Organization.

TI, Trusted Issuer.

TVET, Technical, Vocational and Education Training.

UN, United Nations.

UNESCO, United Nations Educational, Scientific and Cultural Organization.

UNHCR, United Nations High Commissioner for Refugees.

VA, Verifiable Attestation.

VC, Verifiable Credential.

VID, Verifiable Identity.

VP, Verifiable Presentation.

W3C, World Wide Web Consortium.

[38] An electronic data exchange solution empowering individual to control their own student data and exchange throughout lifespan, across borders for various purposes.

frontiers in Blockchain

COMMUNITY CASE STUDY
published: 29 April 2021
doi: 10.3389/fbloc.2021.624258

Decentralized, Self-Sovereign, Consortium: The Future of Digital Identity in Canada

Andre Boysen[1,2]*

[1] *SecureKey Technologies Inc., Toronto, ON, Canada,* [2] *Centre for International Governance Innovation, Waterloo, ON, Canada*

This article introduces how SecureKey Technologies Inc. (SecureKey) worked with various network participants and innovation partners alongside government, corporate, and consumer-focused collaborators, in a consortium approach to create a mutually beneficial network of self-sovereign identity (SSI) principles with blockchain in Canada. These principles are based on giving users ownership and control over all of their digital identity attributes as an alternative approach to the current *status quo* of centralized digital identity, which focuses on discrete identities are made within individual online properties. Blockchain is used as the foundation for its strong security protocols to prevent information from being identified, accessed, or misused and uphold SSI principles. This article will consider the current *status quo* of digital identity known as centralized digital identity and comparisons to the case study's emphasis on the alternative thinking of SSI with principles with blockchain, which prioritizes a decentralized, self-sovereign, consortium approach as opposed to discrete identities within individual online properties. Each of these principles will be explained in detail before highlighting the practical implications, lessons learned for future applications, and how both the Canadian and global identity landscapes should proceed for wider acceptance of SSI with blockchain. The case study detailed – that of Verified.Me – will demonstrate how blockchain developers can actively work to help partners transition from current identity silos to instead collaborate across varied industries and create a cohesive, secure service and digital identity network that benefits users through SSI principles and the benefits of blockchain. We also offer recommendations for how both the Canadian and global identity landscapes should proceed for wider acceptance of SSI with blockchain, the benefits of doing so, and anticipated barriers affecting the adoption of future decentralized identity initiatives.

Keywords: digital identity, identity, blockchain, self-sovereign, decentralized, identity verification, data, data privacy

OPEN ACCESS

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

Reviewed by:
Iain Barclay,
Cardiff University, United Kingdom
Ian Taylor,
University of Notre Dame,
United States
Uyen Trang Nguyen,
York University, Canada

****Correspondence:***
Andre Boysen
andre.boysen@securekey.com

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

Received: *31 October 2020*
Accepted: *02 March 2021*
Published: *29 April 2021*

Citation:
Boysen A (2021) Decentralized, Self-Sovereign, Consortium: The Future of Digital Identity in Canada. Front. Blockchain 4:624258. doi: 10.3389/fbloc.2021.624258

INTRODUCTION

The increased prevalence of today's data breaches and cyber security incidents, the detriments of data silos, and the benefits of proper protocols enforcing security and usability have been important considerations amid the heightened interested in and developments of modern digital identity systems. Our rapidly growing digital world, with subsequent increases in fraud and privacy concerns, requires evolved efforts and advancements in thinking to keep up with these threats and take advantage of opportunities as they develop. The approach of ever-increasing vigilance on the part of users and online properties has well past the peak of diminishing returns. A different approach is needed, an approach of simplification for users that removes the "user-sophistication" requirement of understanding the security model in order to keep data safe.

People need methods to establish the same or better levels of trust for online interactions than we have with in-person transactions. For example, Smits and Hulstijn (2020) detail that a blockchain application may affect the decision to enter the network and engage in a transaction in four ways:

1. The actor believes the institution(s) offering the blockchain-based platform to have properly implemented the blockchain, and for each transaction, to faithfully represent the agreement on the blockchain (party-based trust).
2. The actor believes the blockchain-based network can be monitored and subsequently that the blockchain application helps to reduce transaction risks (control-based trust).
3. The actor sees potential gains because of the blockchain application in the business network. More potential gains enhance engaging in business network transactions.
4. The actor sees transaction risks in the original business network and believes that a blockchain application may reduce those risks, through blockchain-based controls.

The most important principles to establish this trust and increase adoption are security and usability. Consumers want to know their data are safe with proper cyber security measures while having the ease of use required to access services in a way that is not prohibitively complicated. The ability to access different services with the same credentials while staying protected has similarly been a priority for people to increase convenience.

The Commission on Enhancing National Cyber Security established six main imperatives to secure and grow the digital economy (Commission on Enhancing National Cybersecurity, 2016):

1. Protect, defend, and secure today's information infrastructure and digital networks.
2. Innovate and accelerate investment for the security and growth of digital networks and the digital economy.
3. Prepare consumers to thrive in a digital age.
4. Build cyber security workforce capabilities.
5. Better equip government to function effectively and securely in the digital age.
6. Ensure an open, fair, competitive, and secure global digital economy.

This set of principles is full recognition of the inextricably intertwined interests of everyday consumers, the businesses they interact with and the wider economy including government in its own online transactions, as well as in its national cyber-security strategy. User passwords, widespread data breaches, and national cyber security are all facets of the same problem. In short, you cannot have a digital economy without digital identity. Each of these principles is evidence that the lack of digital identity infrastructure is holding the economy back – for commerce, for employers, and for government.

Before going further, it is worth noting what does work well in "street identity" – the identity that is used for in-person transactions today. This allows for a plurality of providers where every business can make its own rules, individuals all make their own choices about what to bring to join the service, and there is some inherent privacy with today's digital identity methods. With street identity documents today, the issuer is blind to where and when the user chose to present the document to a service destination – this is a good thing we want to preserve as we forge ahead.

This plurality exists because there is more than one provider of identity information and more than one sector providing it – driver's licenses, bank statements, and utility bills are all accepted in some transactions but with different providers. Every business can make its own decisions about what is sufficient for them to achieve the requisite level of trust to proceed. The issuing authority is blind to the transaction when the driver's license or bank statement is used in a transaction, which adds to the level of privacy for the user.

Street identity takes a village to make it work. Neither private nor public sector solves the whole street identity problem individually – they solve it together in the hands of and under control of the user. This emphasizes a high level of commitment between the public and private sectors that requires cooperation and collaboration between the two to enhance the state of national cyber security, which is especially important considering that the digitalization of government services includes the need for a safe, portable, and easily accessible digital identity (Zwitter et al., 2020). The advancement of technology continues to outpace security – as such, changes are required in how these sectors approach and implement cyber-security strategies and practices while preserving innovation and ease of use.

The main requirement to satisfy these conditions, eliminate potential obstacles, and increase the benefits of self-sovereign identity (SSI) with blockchain is communication between blockchain organizations, partners in the public and private sectors, and consumers to show the value-add and viability of existing solutions to bring digital ID to its full potential. Transitioning from online identity silos to full collaboration in digital identity that works across the economy requires each to recognize that the benefits of participating in a scheme outweigh the perceived benefits of owning and controlling the whole identity management technical stack to the exclusion of any partnership.

SecureKey developed Verified.Me, a blockchain-based and privacy-centric digital identity verification network – along these imperatives and SSI principles to meet these requirements to provide strong authentication while protecting individual privacy (**Figure 1**).

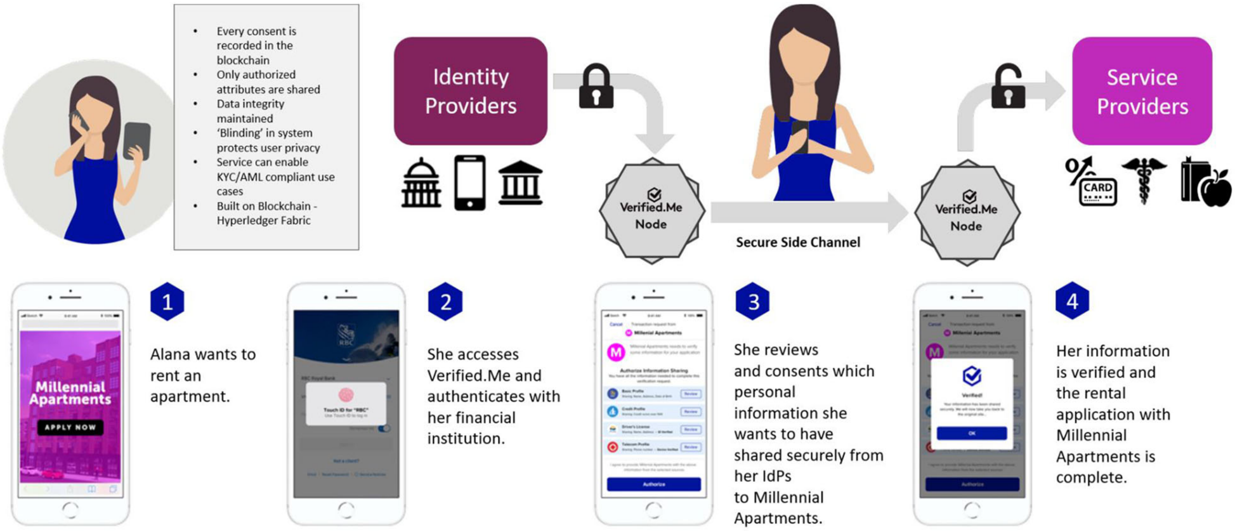

Every consent is recorded in the blockchain
Only authorized attributes are shared
Data integrity maintained
'Blinding' in system protects user privacy
Service can enable KYC/AML compliant use cases
Built on Blockchain - Hyperledger Fabric
Identity Providers
Verified.Me Node
Secure Side Channel
Verified.Me Node
Service Providers
CARD
Millennial Apartments
APPLY NOW
1
Alana wants to rent an apartment.
2
She accesses Verified.Me and authenticates with her financial institution.
3
She reviews and consents which personal information she wants to have shared securely from her IdPs to Millennial Apartments.
4
Her information is verified and the rental application with Millennial Apartments is complete.

FIGURE 1 | A use case example of Verified.Me to verify identities for a rental application.

The initial work toward the eventual launch of the Verified.Me service was first backed by an applied research-focused partnership of the Digital ID and Authentication Council of Canada (DIACC) and Rutgers University Command, Control and Interoperability Center for Advanced Analysis, concentrating on model definition, business analysis, and applied research. Upon completion of the research leg, the Verified.Me service was formally developed in cooperation with seven of Canada's major financial institutions – BMO, CIBC, Desjardins, National Bank of Canada, RBC, Scotiabank, and TD – with additional partners from other industries continuing to join over time after its launch in May 2019.

CONTEXT

Currently, the state of the art in digital identity focuses on centralized models. Discrete identities are made within individual online properties, such as social media accounts, government identity issuance, and corporate management systems. In general, these typically feature one set of credentials, such as a username and password combination, which allows users to access and use platforms, services, and software. While passwords have to be secure, the recovery mechanisms to reset the complex password tend to be trivially simple, thereby defeating the purpose of a complex password. The administrative effort needed to use identities is one of the core challenges of SSI to offer solutions that help users' comfort levels (Der et al., 2017).

Fundamentally, this creates a fragmented identity experience requiring different sets of credentials for different platforms and uses. Centralized digital identity results in sensitive personal data to be stored by each platform in order to operate, which increases security and privacy risks due to how much personal data are stored on their servers (van Wingerde, 2017). The user burden to manage all of this complexity is too high, and the data required to undermine all of these services are stored across all of them. Effectively, if one service is breached, then they are all breached because the breachers replay the data at every endpoint in both password resets and credential-stuffing attacks.

This form of digital identity also lacks the ability to verify the data against the source or with the person presenting it – the system simply knows that the person accessing the system knows certain login credentials (SecureKey, 2020). The combination of a fake driver's license photograph and a real person's driver's license data (name, address, and birthdate) is effective for identity theft in both street and online identity. In street identity use cases, the destination service cannot verify the document against the issuing source, so it falls victim to the real data, fake photo document. The current online trend of taking a selfie and sending it alongside your driver's license also does not solve this problem.

Centralized digital identity also results in the oversharing of data. The documents that are available to choose from in order to verify one's identity may provide required proofs such as name and address, but they also display other personal data rather than what is required by the transaction. A bank statement verifies a name and address while also displaying bank account information and other data such as shopping and spending habits. Consumers are forced to participate in fragmented identity systems where the net benefit and authority over data sharing skew far in favor of the organization with whom they interact. Users are giving up more data than they need to, and this oversharing is a downstream risk when data breaches use the extra information to conduct replay attacks.

This is the essence of the flaw of the existing identity architecture we have today – it is a double-diffusion issue. Neither users nor business can tell what is real and what is not because there is so much fraud noise caused by too many endpoints. The business remedy to fraud noise is to ask for lots of user data to mitigate risk. Thus, crooks then harvest ample user data because they can make money from the data by pretending to be real people. The best way to shut down identity fraud is to make the identity data worthless – mere possession of user should not be sufficient to mount a masquerading or takeover attack of real people. An additional benefit of this approach is that synthetic fraud will also go away because only real people will possess the requisite tools to transact.

Here is an alternative to today's approach, a trusted network approach to digital identity that has demonstrated the ability to solve these negatives (**Figure 2**). Rather than forcing a counter visit, where the documents cannot be verified anyway, a network approach to digital identity with a user-controlled sharing mechanism to present trusted and verified data would serve both the user and the service they want to connect to.

As an example, the registration application can be completed online through the financial institution's online systems by invoking a trusted network service. Street identity verification, while it can be cross-checked with online services, requires an in-person process to confirm that the owner of the credentials is the one giving his/her own information. Network services attempt to solve this limitation of requiring in-person visits by allowing users to collect and present trusted and verified digital assets to and from network participants. Trusted and verified data mean that the data come from an existing, known source that does this already for street identity today.

In that context, these networks have the potential to secure the following (SecureKey, 2019):

- A user's right to privacy of activity.
- A user's right to decide when and what information about themselves is shared between organizations.
- Cryptographic protection of digital assets for confidentiality and integrity.
- That all digital asset exchanges and transactions are cryptographically auditable.
- No central point of failure or trust: a distributed network of trusted organizations runs a cryptographically protected consensus protocol that collectively determines the state of the networks, the participants, the digital assets, and the users.
- Permissions, authentications, and auditability of network participant activities.

DETAIL TO UNDERSTAND KEY PROGRAMMATIC ELEMENTS

Self-sovereign identity is a digital identity philosophical perspective that emerged based on providing users with ownership and control of their digital identity information. This allows them to retain sole control over the management of their digital identity. In comparison to the current philosophy used by centralized digital identity methods, this shifts decision authority to the user through secured distributed ledger – blockchain – technology. It also means that data-replay attacks that are prevalent with user data today are much harder to mount.

While the 10 principles defined by Christopher Allen (Allen, 2016) are abstract and arguably require further development and operationalization (van Wingerde, 2017), these attempt to better conceptualize standards for SSI. Most digital identity projects will not meet all of these criteria, but the 10 principles serve as a preliminary benchmark to assess existing SSI solutions (Wang and De Filippi, 2020):

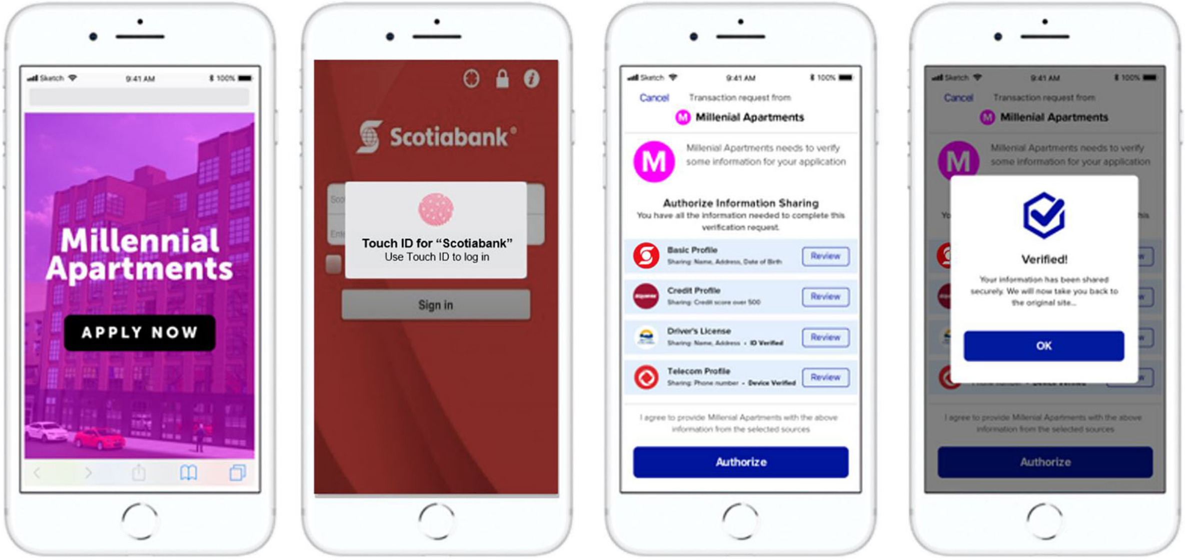
1. *Existence*: Users must have an independent existence.
2. *Control*: Users must control their identities.
3. *Access*: Users must have access to their own data.
4. *Transparency*: Systems and algorithms must be transparent.
5. *Persistence*: Identities must be long-lived.
6. *Portability*: Information and services about identity must be transportable.
7. *Interoperability*: Identities should be as widely usable as possible.
8. *Consent*: Users must agree to the use of their identity.
9. *Minimalization*: Disclosure of claims must be minimized.
10. *Protection*: The rights of users must be protected.

Understanding the current state of digital identity and alternatives to it requires understanding federated identity management. Federated identity uses one system or organization as the main source of managing user authentication as a platform for a group of organizations that offer many different services. Users in this group of organizations can then leverage the same credentials and data to access resources from every organization within the group for the repurposing of identity credentials. One of the biggest challenges of siloed approaches of central and federated systems is overburdening users with identity management (Zwitter et al., 2020). Compared to conventional centralized digital identity models, these credentials allow for access to more than one system as opposed to being limited to one organization per credential. Federated identity management requires the group to trust the one organization designated to manage the user authentication.

Eighty-eight percent of United States consumers have used social logins such as Facebook or Google to conduct authentication through an existing user account (Gigya, 2015), representing the most prominent examples of federated identity management. This information and data are used by an array of other organizations for their own login and authentication processes with the responsibility of managing identities held by Facebook or Google.

Verified.Me takes a hybrid approach, expressing SSI principles within a federated and decentralized identity management system for digital identity verification. Multiple participants work together within a common ecosystem to securely and privately verify the identities of users across the participating organizations with others within the group. SecureKey manages the underlying network to ensure Verified.Me is safe, private, and useful, while upholding the SSI principles.

Federated identity means one identity provider with lots of service destinations. Hybrid means many identity providers with many service destination bound together in a scheme – or trust framework. Hybrid also relates to the method of data sharing.



Millenial Apartments needs to verify some information for your application

You have all the information needed to complete this verification request.

I agree to provide Millenial Apartments with the above information from the selected sources

Your information has been shared securely. We will now take you back to the original site...

FIGURE 2 | The screens a user will see while accessing Verified.Me to verify their identity.

From a data sharing perspective, SSI operates as a store-and-forward model. Users gather claims from identity sources and store the claim in a wallet they control and later share those claims with the service destination. These claims are static as they are created and dated.

While Verified.Me accommodates these types of claims, it also supports real-time claims. With real-time claims, the exact location of the phone the user uses to transact on can be provided, which is harder in a store-and-forward approach. Proof that the user has logged to the user's registered bank account, without divulging the bank details, combined with proven location of the phone, adds additional integrity to static claims and mere possession of the phone. Hybrid also allows for privacy innovation. SSI is a double-blind sharing scheme – the source does not know where the data are sent, but the receiver knows where it came from. With Verified.Me, there is a triple-blind sharing mechanism that blinds the source and destination to each other while also blinding the network to the contents of the transaction. This helps properly account for one of the main caveats associated with decentralization with everyone's interactions typically being made visible to all network nodes (De Filippi, 2016).

"Making Sense of Identity Networks" is an expanded white paper discussing different types of identity networks that has been authored by DIACC[1]. It discusses different approaches, including the approach taken by Verified.Me. What is salient is that it enumerates and discusses the different stakeholder interests in creating digital identity and provides guidance on how to properly balance the different interests with a focus given to user agency in conducting transactions. The key success ingredients are identity portability, stakeholder collaboration, and network governance.

DISCUSSION

One of the most important elements of SSI and federated identity management that is important to consider as a relatively new philosophy in a long-standing digital landscape is the essential role of trust and the fact that the design of a blockchain application influences the trust induced (Smits and Hulstijn, 2020). Given the amount of control users will have over their data, the number of other organizations required, and the principles dictating that the freedoms and rights of users should be preserved over the needs of the network, all the parties involved placing a large deal of responsibility on organization managing the user authentication – "the data is real, but is it the real user?" In the real world, the practical implication of this is that the initial coordination process and gathering funding can be a significant undertaking to prove the organization's capabilities, while also having enough partners involved to showcase the value-add of choosing this model over the current DIY model that may be immediately more convenient, but is more problematic over time.

It is important to state the obvious in order to overcome it. The investment made in identity security today is uneven across online properties – the money available and the skill to administer user identity are not uniform. Yet, startups, internet giants, governments, utilities, and healthcare providers all possess the same essential user data required for crooks to mount successful fraud attacks at all the other online destinations. So money and skill are not a complete remedy to the problem for any online destination. Smarter investment is required in collaborative approaches.

When Verified.Me launched in May 2019, SecureKey worked with various network participants and innovation partners, alongside government, corporate, and consumer-focused collaborators, in a consortium approach to create a

[1] https://diacc.ca/2020/05/13/making-sense-of-identity-networks/

mutually beneficial network that upholds the principles of SSI in Canada. This was developed in cooperation with seven of Canada's major financial institutions – BMO, CIBC, Desjardins, National Bank of Canada, RBC, Scotiabank, and TD – as part of a collaborative blockchain-based approach to help bring about the benefits of decentralized digital identity across the public and private sectors. This process, taking 4 years from the initial design phase and gathering the network of collaborators over time, was a prime example of how blockchain developers can actively work to help partners transition from identity silos to SSI principles and collaborate to create a cohesive, secure digital identity blockchain service. It is important to note that while Verified.Me took 4 years to introduce, it was against the backdrop of the existing successful federated authentication scheme called SecureKey Concierge that launched in 2012; the banks and governments had already learned from that experience.

In order to be successful, a wide variety of public and private sector organizations must be actively involved and act together in close collaboration. For example, the financial institution identity and data providers involved with Verified.Me, namely, the financial institutions listed previously, are responsible for hosting core components of the network and verifying users to service providers, also known as relying parties. Additional roles within the Verified.Me service are played by Canadian organizations that facilitate desired transactions by asking users to provide certain information through the service. Existing and anticipated service providers include, but are not limited to, financial institutions, insurance companies, telecommunications providers, online merchants, healthcare solutions, credit bureaus, legal professionals, sharing economy, online gaming, governments, and educational institutions.

The emphasis on control, privacy, data minimization, and user consent as dictated by SSI principles are incorporated into Verified.Me and advocated for by the network owner, SecureKey. These tie into the advancement of decentralized identity standards in Canada and globally through active collaboration with major institutions for an identity service used by millions of Canadians. SecureKey plans to register Verified.Me for the Decentralized Identifiers specification as set by the Decentralized Identity Foundation. In addition, SecureKey is a founding member and active contributor to the DIACC, as well as the World Wide Web Consortium, UK. Verify, OIX, Kantara, Open ID, and European eIDAS standards. As such, Verified.Me is set up for interoperability with other decentralized identity systems that adhere to these standards.

ACKNOWLEDGMENT OF ANY CONCEPTUAL OR METHODOLOGICAL CONSTRAINTS

There are still a number of challenges for widespread adoption of decentralized identity and SSI principles despite the opportunities available. Although the Commission on Enhancing National Cyber Security was a mentioned component earlier in the article for creating six main imperatives to secure and grow the digital economy, the regulatory landscape is uncertain. As awareness and adoption increase, the attention given to more definitive regulation is expected to increase as well. In particular, encouraging every online service delivery organization to see beyond the perceived safety of complete control over the user ID and password stack they have today is no small feat.

At the time of writing, North America lacks specific regulatory restrictions on SSI and decentralized identity, but private organizations must comply with data privacy regulations and industry-specific requirements (SecureKey, 2019). Decentralized digital identity is understandably and greatly impacted by data privacy regulations. Recent regulatory developments, such as the Personal Information Protection and Electronic Documents Act, the General Data Protection Regulation, and the California Consumer Privacy Act, rightfully seek to manage data portability and place great emphasis on user consent – particularly around data collection and ultimate usage (SecureKey, 2019). While regulations do not specifically prohibit digital alternatives, there are few regulations that acknowledge and encourage better digital alternatives to street identity.

The lack of governance frameworks and agreements between identity providers and service providers has resulted in limited liability assurance and hesitancy by organizations to embrace decentralized digital identity (SecureKey, 2019). One example of uncertainty resulting from a lack of regulation is that financial institutions are required to conduct customer due diligence to prevent fraudulent actions. If a bad actor is permitted into the network by another party, it is unclear who would be held accountable (SecureKey, 2019). As a result, these processes cannot purely rely on SSI and decentralized digital identity until further frameworks are developed and adopted.

More recently, DIACC alongside SecureKey and more than 20 of DIACC's members officially launched and began testing for the Pan-Canadian Trust Framework – a model that will make it easier for Canadian users and businesses to interact online with a high degree of confidence and trust. This initiative sets a streamlined framework of digital ID standards and requirements in place that will guide identity innovation moving forward.

This uncertainty in liability required additional processes to be taken by Verified.Me in drafting new agreements for each network participant on the network to mandate certain performance levels, security requirements, and compliance with privacy and other laws (SecureKey, 2019). Agreements between SecureKey and service providers prohibit the use of subject information for purposes other than the approved sharing transaction (SecureKey, 2019), which also helps satisfy the SSI principle of minimalization. Trust frameworks are both procedural and contractual, but support network effects that eliminate pairwise service and contract negotiation.

In addition to the six main imperatives from the Commission on Enhancing National Cyber Security, it was also stated that preserving innovation and ease of use should be a priority moving forward, which countered the prevalent approach of pushing security to the edges of the network. The ease of implementation for other identity and relying parties, user adoption challenges, and interoperability between organizations and different decentralized identity systems are additional challenges for any decentralized digital identity to be effective for all parties involved (SecureKey, 2020).

From a programmatic perspective, the requirement for all network participants to coordinate, align, and execute on a single launch date was an important undertaking. The planning and execution complexity required in partnership with organizations and within each of those organizations and lines of business were important considerations for the program management team to guide all of the technical, business, and operations teams from all partners. A strong project management office is essential for managing the launch and for any potential crises or detriments that occur in the prelaunch and launch periods (SecureKey, 2020).

For the postlaunch period, the necessity of ongoing management of the ecosystem is another potential constraint for the implementation of a decentralized digital identity system. Adding new parties, monitoring, managing changes and incidents, and end user support are all required elements. Designing, testing, and operationalizing these will be a long-term driver of user and partner satisfaction (DIACC, 2020). The SSI principles guiding this decentralized digital identity network must also be maintained throughout this process, requiring the commitment of all partners on the network to ensure the ongoing success of the network.

While not every digital identity ecosystem will be developed on Verified.Me's scale, it is worth noting that bringing a new system to market based on new blockchain technology will present another set of challenges given the lack of existing resources, references, and lessons to learn from in comparison to centralized digital identity networks (DIACC, 2020). The requirement for this infrastructure and new technology to be scalable to accommodate additional partners over time and resilience to cyber threats is another concern. For Verified.Me, the baseline plan was constantly adjusted to accommodate the additional time required to manage evolving operational, infrastructure, and compliance requirements (DIACC, 2020), and similar efforts will require similar flexibility.

As service delivery organizations gain further knowledge of blockchain technology, and established legal and governance frameworks are developed, there is an anticipation that the technology's prominence and level of participation will increase for businesses, as well as a need for information technology professionals to understand how to use it. The more prominent SSI initiatives with blockchain become, the more likely it will be for organizations to observe and adopt.

CONCEPTUAL BLOCKCHAIN IMPLEMENTATION

Before providing detail on blockchain is being used in the approach presented, it is important to understand that no personally identifiable information (PII) is being stored on chain. Storing PII on chain is privacy degrading in the first instance because the data would be replicated across the verifier nodes, the number of which may increase over time. Getting advanced consent is problematic in that you would be asking the user to agree to share with a party not yet identified. Second, if a user asserts the manifest right to be forgotten, the only way to honor their wish is to delete the whole blockchain.

Blockchain fulfilled three key requirements in a network approach to digital identity:

1. A method to provide triple-blind data sharing under user control and consent while maintaining high business integrity (making it trustworthy to the relying party).
2. A method to compute and record integrity proofs about the data shared.
3. A method to mitigate distributed denial of service attacks owing to the larger number of service endpoints that can provide stand-in processing.

Triple-blind data sharing allows the data to move from the source to the destination service the user chose while mutually blinding the source and destination from each other. The network functions as a blind postal service that delivers the hash address and half of the decryption key, and network address to pick up the payload. The second half of the decryption key is delivered directly from the user agent on the user's mobile phone. The relying party can retrieve the payload and decrypt the payload by assembling the two keys together. This means neither the source, destination, nor the network operator receives a complete picture of the user transaction.

Of integrity proofs, there are three key computations:

1. User chose to have a payload computed and held by the source.
2. The user directed the payload to be sent from the source to the destination.
3. The destination was retrieved and decrypted the payload (to activate the license to the user data).

There is a method for the relying party to compute a hash of the data payload and compare it to the hash that was recorded by the source of the data on chain at creation-time.

This methodology meets the three requirements of trusted data as described above. Trusted means:

(1) a known and trusted source because only trusted sources can write on chain,
(2) knowing that data have not been altered since it was issued by that source because the hashes computed by the source and destination match, and
(3) that data belong to the person presenting them because only the user agent could cause delivery of the payload to the destination.

CONCLUSION

This article introduced how SecureKey worked with various network participants and innovation partners, alongside government, corporate, and consumer-focused collaborators, in a consortium approach to create a mutually beneficial network of SSI principles with blockchain in Canada, a network based on triple-blind privacy, designed to work across the economy under the control and direction of the user with higher integrity, lower cost, and customer experience benefits for businesses. Through Verified.Me, arguments for the usage of blockchain-based services that bake the SSI philosophy into their foundation were presented to demonstrate its benefit to organizations and users alike. Despite the challenges associated with adoption, implementation, and the current lack of regulatory restrictions, decentralized digital identity continues to increase in usage in Canada while the global identity landscape shifts to a wider acceptance of SSI with blockchain and a better understanding of the benefits of doing so.

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.

REFERENCES

Allen, C. (2016). *The Path to Self-Sovereign Identity*. Available online at: http://www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html (accessed September 15, 2020)

Commission on Enhancing National Cybersecurity (2016). *Report on Securing and Growing the Digital Economy*. Available online at: https://www.nist.gov/system/files/documents/2016/12/02/cybersecurity-commission-report-final-post.pdf (accessed September 15, 2020)

De Filippi, P. (2016). The interplay between decentralization and privacy: the case of blockchain technologies. *J. Peer Prod.* 7:hal–01382006.

Der, U., Jähnichen, S., and Sürmeli, J. (2017). Self-sovereign identity $-$ opportunities and challenges for the digital revolution. *arXiv [Preprint]* doi: arXiv: 1712.01767

DIACC (2020). *Consumer Digital Identity Leveraging Blockchain*. Available online at: https://diacc.ca/wp-content/uploads/2020/03/DIACC-White-Paper_Consumer-Digital-Identity-Leveraging-Blockchain_Feb2020.pdf (accessed September 15, 2020).

Gigya (2015). *The 2015 State of Consumer Privacy & Personalization*. Available online at: https://www.slideshare.net/Gigya/white-paper-the-2015-state-of-consumer-privacy-personalization (accessed September 15, 2020)

SecureKey (2019). *Consumer Digital Identity Leveraging Blockchain*. Available online at: https://verified.me/wp-content/uploads/2019/05/DIACC_Phase2_SK-2019-FINAL.pdf (accessed September 15, 2020).

SecureKey (2020). *A Primer and Action Guide to Decentralized Identity*. Available online at: https://securekey.com/wp-content/uploads/2020/07/VerifiedMe_OWIWhitepaper_APrimertoDecentralizedIdentity.pdf (accessed September 15, 2020).

Smits, M., and Hulstijn, J. (2020). Blockchain applications and institutional trust. *Front. Blockchain.* 3:5. doi: 10.3389/fbloc.2020.00005

van Wingerde, M. (2017). *Blockchain-enabled self-sovereign identity* Doctoral dissertation, Master's thesis. Tilburg: Tilburg University.

Wang, F., and De Filippi, P. (2020). Self-Sovereign identity in a globalized world: credentials-based identity systems as a driver for economic inclusion. *Front. Blockchain.* 2:28. doi: 10.3389/fbloc.2019.00028

Zwitter, A., Gstrein, O., and Yap, E. (2020). Digital identity and the blockchain: universal identity management and the concept of the "Self-Sovereign" individual. *Front. Blockchain.* 3:26. doi: 10.3389/fbloc.2020.00026

Conflict of Interest: AB is employed by the company SecureKey Technologies Inc.

frontiers in Blockchain

DATA REPORT
published: 30 April 2021
doi: 10.3389/fbloc.2021.622090

Evaluating Trust Assurance in Indy-Based Identity Networks Using Public Ledger Data

Will Abramson [1]*, *Nicky Hickman* [2] *and Nick Spencer* [3]

[1] Blockpass Identity Lab, Edinburgh Napier University, Edinburgh, United Kingdom, [2] Sovrin Foundation, Provo, UT, United States, [3] Incudeas Limited, Thatcham, United Kingdom

Keywords: self-sovereign identity, distributed ledger analytics, verifiable credentials, cryptography, verifier, trust, assurance, risk

OPEN ACCESS

Edited by:
Michael Shea,
Independent Researcher, Litchfield,
United States

Reviewed by:
Andreas Freund,
Independent Researcher, San Diego,
United States
Meghana Kshirsagar,
University of Limerick, Ireland

Correspondence:
Will Abramson
will.abramson@napier.ac.uk

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

Received: *27 October 2020*
Accepted: *22 March 2021*
Published: *30 April 2021*

Citation:
Abramson W, Hickman N and Spencer N (2021) Evaluating Trust Assurance in Indy-Based Identity Networks Using Public Ledger Data. Front. Blockchain 4:622090. doi: 10.3389/fbloc.2021.622090

1. INTRODUCTION

Self-sovereign identity (SSI) encapsulates a set of technologies, tools, and governance models designed to outline and facilitate the transition to a new paradigm for digital identity systems. One where individuals, organisations, and things are able to actively participate as peers in the digital relationships they establish and maintain over time. The evolving ideology around this movement, initially articulated as 10 principles (Allen, 2016), focuses on empowering the individual, providing them with an independent digital existence that is usable and useful across contexts.

The technical architecture that is emerging defines three distinct transactional roles that entities within an SSI system can engage in; issuer, verifier, and holder. An issuer signs a set of attributes they are attesting to about an entity then presents a data object containing this signature and attributes, a credential, to the entity in the role of holder for this interaction. A holder can then present these attributes along with a cryptographic proof to any number of entities in future interactions. The entity receiving this proof and verifying its integrity is defined as the verifier for that interaction (Sporny et al., 2019a).

To support this architecture a number of open standards are under development. The most mature is the Verifiable Credential Data Model, a W3C recommended standard for the structure of the credential data object that issuers sign (Sporny et al., 2019a). Decentralised Identifiers (DIDs) are another key specification currently going through standardisation in the W3C DID Working Group. This specification defines a new type of identifier designed to facilitate this verifiable, decentralised architecture for digital identity (Reed et al., 2020). DIDs enable entities to provision and manage their own identifiers using a decentralised system and public key cryptography rather than external parties (Allen et al., 2015). These identifiers must be resolvable to a DID Document which contains public keys and authentication mechanisms that support the cryptographic verification of signatures made by the entity in control of the associated private keys.

The W3C DID specification is designed to be technology and protocol agnostic, instead defining a common syntax that can be used to understand all DIDs and a generic set of requirements for create, read, update, and deactivate operations of DID Documents (Reed et al., 2020). Implementers of DID methods select an infrastructure they trust to store these identifiers and their related documents. A distributed ledger, as an append only, immutable, highly available decentralised data storage system is ideal for this infrastructure (Allen et al., 2015; Evans-Greenwood et al., 2016).

This paper focuses on a specific type of distributed ledger designed to support this technical architecture, Hyperledger Indy. The data contained within this ledger are analysed from the perspective of a verifier attempting to assess the risk associated with accepting a credential presentation they have received.

2. METHOD

Hyperledger Indy is an open source code base for the instantiation of a ledger to support the creation of public identifiers, DIDs, able to issue and revoke cryptographic credentials using an RSA based scheme first published by Camenisch and Lysyanskaya (2001, 2002). Anyone with read access to the ledger can verify signatures made by issuers on credentials, or their presentations. As Indy has been designed solely for the purpose of identity management and supports anonymous credential cryptography, it stores unique data in contrast to other ledgers that store decentralised identifiers, such as the Bitcoin or Veres One ledgers (Allen et al., 2019; Sporny et al., 2019b). These data are written to the ledger in a number of different transaction types. These are:

- **NYM**—These transactions write a new DID and related DID Document to the ledger.
- **ATTRIB**—Transactions that update existing DID Documents on the ledger, such as rotating keys or changing service endpoints. These must be authored and signed by the DID that identifies the DID Document being updated.
- **SCHEMA**—These transactions define a schema name, version, and list of attribute names for a specific credential. The schema name must be unique on the ledger, but can be altered by writing a schema with the same name and different version number. Versioning a schema must be done by the original author of the schema transaction.
- **CLAIM_DEF**—Often referred to as a credential definition, these transactions write the public key from a generated key pair of an CL-RSA signature for a specific credential schema (Camenisch and Lysyanskaya, 2002). Only DIDs with CLAIM_DEF transactions for specific schema included in the ledger can issue credentials of this schema that are publicly verifiable. Many DIDs can author CLAIM_DEF transactions referencing the same schema.
- **REVOC_REG_DEF**—Transactions that define a revocation registry for a certain credential definition transaction (CL-RSA public key) meaning that credentials signed by this public key can be revoked. Currently, these registries use cryptographic accumulators defined in a 2009 paper by Camenisch et al. (2009).
- **REVOC_REG_ENTRY**—Whenever an issuer issues or revokes a credential, they must author a transaction that updates the revocation registry keeping them up to date so they can be used to construct and verify proofs of non-revocation.

Only NYM and ATTRIB transactions are analogous to other ledgers storing and maintaining DIDs. The reason Indy ledgers include SCHEMA and CLAIM_DEF transactions is likely determined by the need to efficiently support CL-RSA signatures. They have public keys that grow linearly in size with the number of attributes being signed and can take seconds to generate (Camenisch and Lysyanskaya, 2002; Pointcheval and Sanders, 2016). This is too long to be generated at verification time from a single key, hence they are pre-generated by issuers who specify the number of messages to be signed by identifying the schema they intend to issue. This is then stored on the ledger improving verification efficiency. The revocation transactions are similarly unique to Indy ledgers, to our knowledge the only ledger attempting to support anonymous revocation of credentials. These design choices, heavily influenced by the cryptographic primitives the ledger supports, present a richer source of transaction data than other ledgers used to support SSI interactions. As new, more efficient cryptographic protocols, such as BBS+ are supported by Indy, it is expected that the design choices of these ledger will not be so dependent on these protocols (Camenisch et al., 2016).

All transactions include the time they were authored and a unique identifier that can be used to reference and resolve data from within them. Transactions must be signed by the public key associated with a DID already stored on the ledger before it is accepted by the nodes maintaining the ledger state. This leads to a hierarchical structure whereby all DIDs must first be authored to the ledger in a nym transaction signed by the key of another DID before they can themselves write transactions to the ledger. Any Indy-based ledger is initiated with a number of genesis nym transactions and all other nym transactions can be traced to a nym transaction signed by one of these DIDs. This structure of signed transactions allows any entity to verify the validity of the ledger state by starting from these genesis transactions. It also ensures rules around which DID has the authority to update a DID Document, schema, or revocation registry can be cryptographically enforced.

The dataset under analysis in this paper are the transactions from a specific instantiation of an Indy based distributed ledger, the Sovrin MainNet. A ledger that has been running since July 2017 that supports some of the most mature deployments SSI systems today. The ledger includes 448 nym transactions, including 16 genesis nyms representing the Sovrin board of trustees, 88 schema, and 356 credential definitions. While other Indy ledgers include far more transactions, such as the Sovrin StagingNet with almost 20,000 nym transactions, the Sovrin MainNet is for production deployments of SSI so provides a more realistic dataset. Despite this focus on the MainNet, the analysis should be at least partially applicable to any Indy-based distributed ledger.

A major difference between the Sovrin MainNet and other Indy networks is that it is a public-permissioned network governed by the Sovrin Governance Framework that defines the roles and responsibilities of different actors within the network (Sovrin Governance, 2019a). A permissioned network adds additional constraints around who can write to the ledger. In the Sovrin MainNet only DIDs with the role of transaction endorser are able to write nym transactions to the ledger and all subsequent transactions these DIDs author must be additionally signed by a transaction endorser (Sovrin Governance, 2019c). This presents interesting opportunities for analysis as we discuss later in the paper.

The nodes within the Sovrin MainNet are run by Sovrin Stewards, organisations that volunteer time and resources to maintain the network. The network is administered and managed by the Sovrin Foundation which also acts as a Governance Authority (Sovrin Governance, 2019a). Stewards are selected by the Governance Authority to ensure maximal distribution of hardware, domain, and geographic location limiting the threat vector of malicious takeover and promoting resilience. All stewards agree to the requirements specified by the Governance Framework and sign the Sovrin Stewards Agreement (Sovrin Governance, 2019a,b). Nodes accepted into the network then engage in a consensus protocol named plenum based on redundant byzantine fault tolerance (Aublin et al., 2013). As such, assuming the Sovrin Foundation and a subset of the stewards can be trusted then the transactions stored within the ledger can be trusted with a high degree of confidence.

The data discussed within this paper can be accessed through the public hyperledger indy transaction explorer, IndyScan. For more detailed analysis, it is also possible to clone the github repository for this explorer and visualise the data from the ledger through a Kibana dashboard or similar. Alternatively, the ledger data can be fetched using indy-vdr a hyperledger repository designed for querying indy nodes. This paper uses visualisations of a subset of MainNet transactions retrieved using the IndyScan API.

3. ANALYSIS

Analysis of the data held within a Hyperledger Indy network may be useful for answering questions from many different perspectives within an SSI system. This paper focuses on one in detail, that of a verifier attempting to determine whether to accept a proof of a set of attributes presented by a credential holder. While this decision will be tied to the semantic context of the interaction and is largely subjective for each verifier, we focus our analysis specifically on the syntax, the information contained within the ledger that might influence the decision of a verifier. Either alerting them to increased risk, or giving them a greater degree of assurance.

The presentation of indy-backed credentials is specified by Aries-rfc-0037 (Khateev, 2019), a protocol involving two entities, a holder and a verifier, that have previously exchanged peer DIDs to establish a DIDComm channel across which encrypted, digitally signed messages can be exchanged, authenticated, and decrypted. The holder then constructs a proof object from a set of credentials that have previously been issued to them and sends this to the verifier. From this proof, the verifier is able to learn:

- The attribute values presented
- The identifiers of the scheme the attributes were issued in
- The identifiers of a set of claim definitions
- The mathematical proof of the integrity of the attributes
- The mathematical proof of a common master secret attribute known to the holder and signed by the issuer of each credential involved in the presentation
- The identifiers for the revocation registries of credentials if applicable

The verifier can then query the ledger for the CLAIM_DEF transactions to return the public keys of the issuers of each of the credentials used to construct the proof. Using these keys the proof object can be mathematically verified such that the verifier can have high confidence that the attributes presented were issued to the same master secret, the holder knows this secret and the attributes presented have not been tampered with since issuance. Additionally, resolving the REVOC_REG_DEF transactions allows for verification of any proof of non-revocation, if this has been included in the presentation. However, in addition to the fidelity of the information contained within the presentation, a verifier must assess its provenance (Windely, 2020).

This paper suggests Indy transaction data can provide insights into the question of provenance by using the SCHEMA and CLAIM_DEF transaction identifiers as a starting point for inquiry. By querying the ledger dataset for these transactions, the verifier learns the DIDs of the transaction author and transaction endorser for both transactions. Depending on the context, different comparisons may be appropriate here. A verifier may expect both of these transactions to have been endorsed by the same DID. In the future, this may present a mechanism to associate a presentation with a specific governance domain that the credentials were issued under, where the endorser represents a governance authority. In contrast, when comparing the DID that authored the SCHEMA with that of the CLAIM_DEF, a difference here might give the verifier greater assurance.

Another potentially useful insight can be gained from the ledger by querying all CLAIM_DEF transactions that reference the schema used within the presentation. See the dotted lines between blue nodes (SCHEMA) and green nodes (CLAIM_DEF) in **Figures 1**, **2**. Through this, the verifier learns how many distinct issuers are able to issue this credential, giving some indication of its value and adoption. This analysis can be extended further by including the transaction endorsers of these CLAIM_DEF transactions and, further still, to include the endorser of the NYM transactions for the DIDs that authored these CLAIM_DEFs. A visualisation of this analysis can be seen in **Figure 2**.

This approach effectively graphs the roots of trust associated with a particular credential schema. In this instance, a single endorser used for all transactions might indicate a strong governance domain, particularly where there are many issuers involved. The analysis of these patterns can be derived from the SCHEMA transaction identifier, information that is included in a presentation request so available to all verifiers. Additionally, by placing the CLAIM_DEF and NYM transactions of the issuer within this pattern it may be possible to spot anomalies alerting them of potentially untrustworthy issuers. For example, if these transactions had been endorsed by a different DID in a schema pattern that has a common endorser for all other transactions. Such patterns can clearly be seen within the Sovrin MainNet, as the visualisations in **Figures 1**, **2** show.

Querying the ledger for information about a DID could be worthwhile for certain verifiers as it would enable them to see all the transactions they have authored over time. The importance of the author of the NYM transaction that initially wrote this DID to the ledger has already been emphasised, however, other information may be equally useful. For example, how long ago the NYM transaction was authored, how many CLAIM_DEF transactions they have written to the ledger, and which credential schema are they for.

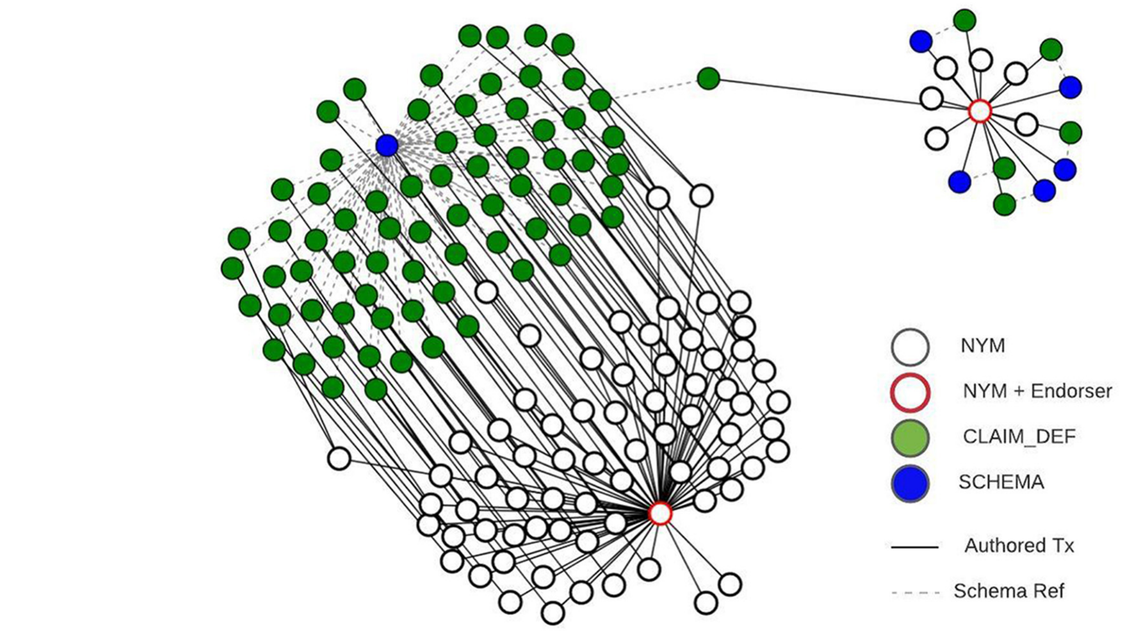


FIGURE 1 | Visualisation of authored transactions linked to a single schema identifier.

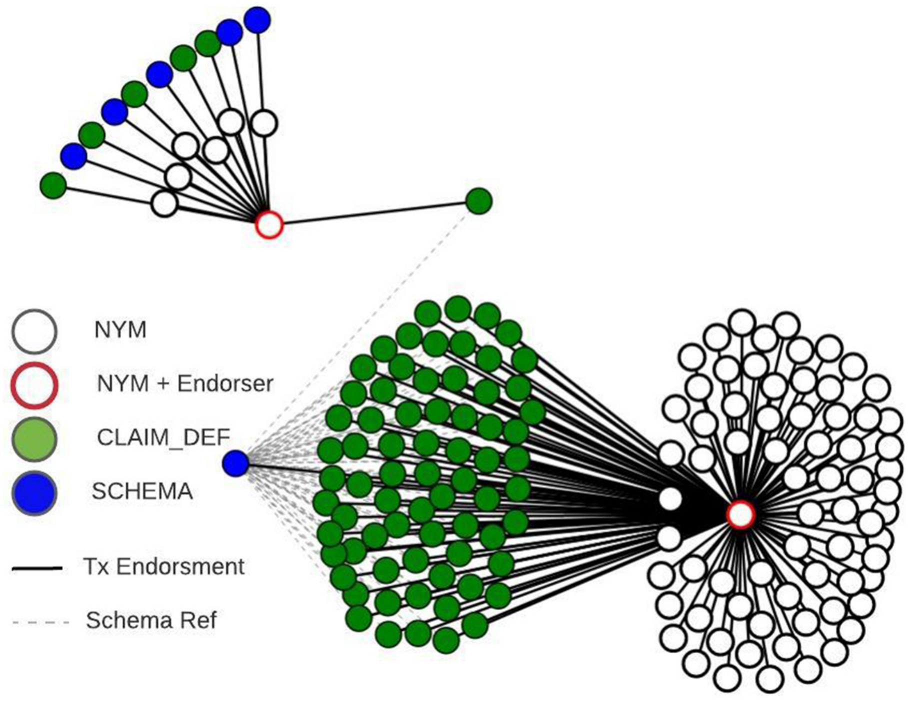


FIGURE 2 | Visualisation of transaction endorsement for the same transactions shown in **Figure 1**.

The analysis presented has focused only on the ledger data, following a logical pathway of inquiry a verifier might take when presented with a proof object from an entity containing SCHEMA and CLAIM_DEF transaction identifiers. It has been described to illustrate what it is possible to learn from this data independently of any contextual information that can be inferred from the interaction or provided by the verifying entity itself. This additional information may determine which questions are appropriate to ask from the data, as well as the acceptable answers a verifier expects. An example of this might be the expectation that issuers NYM and CLAIM_DEF transactions were endorsed by a specific DID that is meaningful to the verifier.

4. CONCLUSION

This paper takes an indepth look at the data available within Hyperledger Indy-based ledgers, focusing particularly on the Sovrin MainNet, an established public ledger designed for production use cases. This specific instantiation has well-defined governance processes and legally binding agreements for all actors within the network. Assuming trust is placed in these processes then the information within the ledger can be trusted to a high degree of assurance. In the future, it is expected that many more public networks based on Hyperledger Indy will emerge for production use cases, as this happens the ability to assess the trust placed in the specific ledger itself will become increasingly important. This work is already underway within the Sovrin community to define a set of common metrics with which to evaluate different Indy nodes, ledgers, and networks (Foundation, 2020; Indy, 2020).

For now though, it is important to recognise that the ledger within an SSI network is designed to be a highly assured source of information. Wherever there is data, there are insights that can be drawn from this data. This paper puts forward an initial attempt to describe exactly what these insights might be and how they could be useful from the perspective of a verifier. Within SSI, there are many perspectives that could adapt the approaches described within this paper to answer their own questions. Implications of this research could be built into the governance framework's assurance policies as well as verifiers' business logic and user experience design. Equally, this suggests that information from a public Indy ledger has potential privacy and security implications for issuers. Further research is required here, but it may be that for certain use cases and industries, this is unacceptable.

We emphasise that this report is focused primarily on the structure of the transaction data found within Indy ledgers and the potential patterns that might emerge when these transactions and their relationships are graphed. While the use case visualised in **Figures 1**, **2** are of real transaction data on the Sovrin MainNet from an advanced pilot within healthcare known to the authors, it has been presented to illustrate the kinds of relationships and patterns we think are useful to pay attention to. It is our hope that this work stimulates further research into the patterns found across a statistically meaningful sample of SSI applications, so that reliable conclusions can be drawn.

In addition to this, there are many other DID methods that resolve identifiers against other distributed ledgers, such as Bitcoin, Ethereum, and Veres One. These are all permissionless ledgers that support decentralised identity systems without storing schema or credential definitions on the ledgers, a quirk of Indy-based ledgers due to the anonymous credential cryptography they support. This means that DIDs will not be so directly correlated with the schema they can issue, or schema with the DIDs that can issue them. Furthermore, since anyone can write a DID to permissionless ledgers, different mechanisms will need to be implemented to determine a DIDs provenance. Finally, credential systems using non-Indy ledgers often require holders to record DIDs on the ledger in order to be able to authenticate as the credential subject to a verifier. The advantages and disadvantages of these differences and their implications for potential ledger analysis deserve further attention.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: https://indyscan.io/home/SOVRIN_MAINNET, https://github.com/wip-abramson/sovrin-network-vis, https://github.com/wip-abramson/sovrin-network-vis/blob/master/src/graphdata.json.

AUTHOR CONTRIBUTIONS

WA, NH, and NS have been collaborating on broader research into metrics for SSI systems for over 6 months. This work informed all ideas presented in this paper. WA analysed the ledger data, created the visualisation, and wrote up the first draft. NH and NS provided the feedback and review. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors would like to acknowledge Patrik Stas, the creator of the IndyScan transaction explorer who helped us access and interpret the data analysed in this paper.

REFERENCES

Allen, C. (2016). *The Path to Self-Sovereign Identity*. Life with Alacrity. Available online at: http://www.lifewithalacrity.com/2016/04/the-path-to-self-sovereign-identity.html

Allen, C., Brock, A., Buterin, V., Callas, J., Dorje, D., Lundkvist, C., et al. (2015). *Decentralized Public Key Infrastructure*. Group Report. Rebooting the Web of Trust (RWoT). Available online at: https://danubetech.com/download/dpki.pdf.

Allen, C., Duffy, Kim, H., Grant, R., and Pape, D. (2019). *BTCR DID Method*. Technical report, World Wide Web Consortium.

Aublin, P. L., Mokhtar, S. B., and Quéma, V. (2013). "RBFT: redundant byzantine fault tolerance," in *2013 IEEE 33rd International Conference on Distributed Computing Systems* (IEEE), 297–306. doi: 10.1109/ICDCS.2013.53

Camenisch, J., Drijvers, M., and Lehmann, A. (2016). "Anonymous attestation using the strong Diffie Hellman assumption revisited," in *International Conference on Trust and Trustworthy Computing* (Springer), 1–20. doi: 10.1007/978-3-319-45572-3_1

Camenisch, J., Kohlweiss, M., and Soriente, C. (2009). "An accumulator based on bilinear maps and efficient revocation for anonymous credentials," in *International Workshop on Public Key Cryptography* (Springer), 481–500. doi: 10.1007/978-3-642-00468-1_27

Camenisch, J., and Lysyanskaya, A. (2001). “An efficient system for non-transferable anonymous credentials with optional anonymity revocation,” in *International Conference on the Theory and Applications of Cryptographic Techniques* (Springer), 93–118. doi: 10.1007/3-540-44987-6_7

Camenisch, J., and Lysyanskaya, A. (2002). “A signature scheme with efficient protocols,” in *International Conference on Security in Communication Networks* (Springer), 268–289. doi: 10.1007/3-540-36413-7_20

Evans-Greenwood, P., Hillard, R., Harper, I., and Williams, P. (2016). *Bitcoin, Blockchain & Distributed Ledgers: Caught Between Promise and Reality.* Deloitte.

Foundation, S. (2020). *SSI Metrics Dashboard*. Available online at: https://github.com/hyperledger/indy-node-monitor

Indy, H. (2020). *Indy Node Monitor*. Available online at: https://github.com/hyperledger/indy-node-monitor

Khateev, N. (2019). *Aries RFC 0037: Present Proof Protocol 1.0*. Technical report.

Pointcheval, D., and Sanders, O. (2016). “Short randomizable signatures,” in *Cryptographers’ Track at the RSA Conference* (Springer), 111–126. doi: 10.1007/978-3-319-29485-8_7

Reed, D., Sporny, M., Longely, D., Allen, C., Sabadello, M., and Grant, R. (2020). *Decentralized Identifiers (DIDs) v1.0*. Technical Report. World Wide Web Consortium.

Sovrin Governance, F. W. G. (2019a). *Sovrin Governance Framework v2*. Provo, UT: The Sovrin Foundation.

Sovrin Governance, F. W. G. (2019b). *Sovrin Stewards Agreement*. Provo, UT: The Sovrin Foundation.

Sovrin Governance, F. W. G. (2019c). *Sovrin Transaction Endorser Agreement*. Provo, UT: The Sovrin Foundation.

Sporny, M., Longely, D., and Chadwick, D. (2019a). *Verifiable Credentials Data Model 1.0*. Technical report, World Wide Web Consortium.

Sporny, M., Longely, D., and Christopher, W. (2019b). *Veres One DID Method 1.0*. Technical report, World Wide Web Consortium.

Windely, P. (2020). *Origins and Principles of SSI*. Breaking Silos.

Conflict of Interest: NS was employed by company Incudeas Limited.

The remaining 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.

frontiers in Blockchain

ORIGINAL RESEARCH
published: 28 July 2021
doi: 10.3389/fbloc.2021.626726

Sovrin: An Identity Metasystem for Self-Sovereign Identity

*Phillip J. Windley**

Office of Information Technology, Brigham Young University, Provo, UT, United States

Solving the problems of digital identity in a holistic manner requires that we rethink how we architect identity online. This paper presents the architecture of an identity metasystem called the Sovrin Network that aims to improve the user experience, increase flexibility, and reduce overall costs while supporting better privacy and security. We discuss the problems of online identity on the modern internet, discuss the nature of digital relationships, explore the architectures of identity systems, and detail the combination of these concepts into a comprehensive metasystem for solving the problems of online identity.

Keywords: self-sovereign identity, identity, cryptography, Sovrin, identifiers, verifiable credentials

OPEN ACCESS

Edited by:
Michael Shea,
Independent researcher, Litchfield,
Connecticut, United States

Reviewed by:
Prerna Tambay,
Consultant, London, United Kingdom
Maryline Laurent,
Télécom SudParis, France

****Correspondence:***
Phillip J. Windley
pjw@byu.edu

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

Received: *06 November 2020*
Accepted: *14 July 2021*
Published: *28 July 2021*

Citation:
Windley PJ (2021) Sovrin: An Identity Metasystem for Self-Sovereign Identity.
Front. Blockchain 4:626726.
doi: 10.3389/fbloc.2021.626726

INTRODUCTION

The internet was designed without an identity layer, at least for people (Cameron, 2005). At the time, any network user was identified by proxy through the machine they used to connect and whatever access control system it had. Personal computers and the web led to an internet where many people are online without any sponsoring organization. But the administrative model was so entrenched in the architecture of the internet that we simply perpetuated it with a different administrative identity system, username, and password for every relationship on every site and app.

The internet is a metasystem—a system of systems. The internet is not so much a communications system as it is a system for building communication systems. Metasystems employ protocols, governance, and convention to provide decentralized interoperability between the systems they comprise.

Naturally, an identity system for the internet should be a metasystem as well since no single system can meet the needs of every digital relationship. An identity metasystem is a system for building interoperable identity systems. The concept of an identity metasystem was first introduced by Kim Cameron in 2005 (Cameron, 2005). In describing this system, Cameron said:

> We need a unifying identity metasystem that can protect applications from the internal complexities of specific implementations and allow digital identity to become loosely coupled. This metasystem is in effect a system of systems that exposes a unified interface. . .

An identity metasystem provides the building blocks and protocols necessary for others to build identity systems that meet the needs of any specific context or domain.

This paper explores the architecture of an identity metasystem called the Sovrin Network. An identity metasystem like Sovrin is a prerequisite for an online world where identity is as natural as it is in the physical world. An identity metasystem can remove the friction, decrease cognitive overload, and make online interactions more private and secure.

THE PROBLEMS OF ONLINE IDENTITY

The internet's missing identity layer has resulted in a mishmash of one-off identity systems because every web site, service provider, and application has solved the problem in a unique way (Simmonds, 2015). As a result, people and organizations who use the internet are subject to cognitive overload, friction, increased costs, loss of privacy, and even outright fraud.

Fixing the internet's identity problem is hard. There have been numerous systems, protocols, and standards proposed over the past 20 years (Naik and Jenkins, 2016). While most of them have provided improvements and fixed specific problems, none have offered a holistic solution.

To see why digital identity is so hard, consider the following specific problems that make identity online different from the physical world.

Proximity—Because we are not interacting with people physically, our traditional means of knowing who we are dealing with are useless. None of the familiar signals of the physical world are present. Consequently, it is difficult to reliably recognize and remember people and organizations online (Andrieu 2018). Organizations have built administrative identity systems to serve their own needs in recognizing and remembering their customers, but people do not have the same capabilities. Consequently, we are mired in myriad, incompatible systems built for narrow purposes.

Autonomy—Each of these administrative systems is built for the convenience of the organization who controls it. Design choices for these systems are made to maximize the legibility of people to the organization for its purposes, skewing the balance of power toward the organization (Windley 2020). Consequently, people have very few natural rights and little leverage online. Current online identity systems significantly reduce individual freedom and autonomy.

Flexibility—Closely related to the autonomy problem is one of flexibility. Current online identity systems are built for very narrow purposes. But real life is messy, with billions of use cases (Windley 2018). People are innovative and infinitely diverse. None of us presents the same picture of ourselves to everyone and everything—how we recognize, remember, and respond to others is highly dependent on the context.

Privacy—No one will be surprised to learn that computers are very good at pattern matching. But a consequence of this is that online identity has very different implications for privacy than physical world interactions (Hardman 2019a). When you hand your driver's license to the bartender to establish your legal age, you would be surprised if she could remember all the detailed information it contains, like your address, and do that for every customer she encountered. Computers, on the other hand, retain a perfect memory of all the information they are presented with until they are told to forget.

Anonymity—Anonymity is closely related to privacy. In real life, we do without identity systems for most things. You do not have to identify yourself to the movie theater to watch a movie or log into some system to sit in a restaurant and have a private conversation with friends. Many of our interactions in the physical world are naturally anonymous because they are ephemeral. The ticket taker at a movie theater does "identify" you momentarily for purposes and checking your ticket, but that connection is short-lived and thus anonymous for most purposes. Many online interactions could make use of ephemeral relationships as well to better support privacy.

Interoperability—A consequence of myriad identity silos is that we are unable to carry context from system to system (Simmonds, 2015). Your friend in one system might have a different identifier in another. Consequently, your ability to recognize and remember varies from system to system.

Scale—There are billions of people online. Each of them has dozens, even hundreds of relationships. The internet of things promises to increase that by several orders of magnitude. Consequently, a general-purpose identity system needs to account for trillions of relationships between the many billions of people, organizations, and things that make up the online world. No single, centralized system can do it.

Solving these problems requires building something more abstract and general than the one-off, context-specific identity systems of the past.

RELATIONSHIPS

Identity systems exist to support online relationships. Managing identity information is merely a means to an end. In *Identities Evolve: Why Federated Identity is Easier Said than Done* (Wilson, 2011), Steve Wilson argues that the goal of using federation schemes to create a few identities that serve all purposes is deeply flawed. Wilson's point is that we have hundreds, even thousands, of online identities because we have lots of relationships. The identity data for a given relationship is contextual and highly evolved to fit its specific niche.

Each relationship has a common root, the person being identified, but it is highly contextualized. Some relationships are long-lived, some are ephemeral. Some are personal, some are commercial. Some are important, some are trivial. Still, we have them. The information about ourselves, what many refer to as identity data, that we share with each is adapted to the specific niche that the relationship represents. Once you realize this, the idea of creating a few online identities to serve all needs becomes preposterous.

Because of the proximity problem, we are not interacting with people physically and so our natural means of knowing who we are dealing with are useless. Joe Andrieu defines (Andrieu, 2018) identity as "how we recognize, remember, and respond to" another entity. I add "rely on" to the list.

These activities depend on three properties that any digital relationship must have to overcome the proximity problem:

Integrity—we want to know that, from interaction to interaction, we are dealing with the same entity we were before. In other words, we want to identify them so that we can recognize and remember them.

Lifespan—normally, we want relationships to be long-lived, although we also create ephemeral relationships for short-lived interactions.

Utility—we create online relationships in order to use them within a specific context.

Relationship Integrity

Integrity allows parties to a relationship to recognize each other. Consequently, all identity systems manage relationship integrity as a foundational capability. Federated identity systems improve on one-off, often custom, identity systems by providing integrity in a way that reduces user management overhead for the organization, increases convenience for the user, and increases security by eliminating the need to create one-off, proprietary solutions. An identity metasystem aims to establish relationship integrity with the convenience of the federated model but without relying on an intervening identity provider (IdP) in order to provide autonomy and privacy.

A relationship has two parties, let us call them P1 and P2[1]. P1 is connecting with P2 and, as a result, P1 and P2 will have a relationship. P1 and P2 could be people, organizations, or things represented by a web site, app, or service. Recognizing the other party in an online relationship relies on being able to know that you are dealing with the same entity each time you encounter them.

In the identity metasystem represented by the Sovrin Network, a relationship is initiated when P1 and P2 exchange decentralized identifiers (DIDs) (Decentralized Identifiers, 2020). For example, when a person visits a web site or app, they are presented with a connection invitation. When they accept the invitation, they use a software agent to share a DID that they created. In turn, they receive a DID from the web site, app, or service. We call this a "connection" since DIDs are cryptographically based and thus provide a means of both parties mutually authenticating. The user experience does not necessarily surface all this activity to the user[2].

In contrast to the federated model, the participants in the metasystem mutually authenticate and the relationship has integrity without the intervention of a third party because the identifiers are self-certifying (Smith, 2020). By exchanging DIDs, both parties have also exchanged public keys. They can consequently use cryptographic means to ensure they are interacting with the party who controls the DID they received when the relationship was initiated. Mutual authentication based on self-certifying DIDs provides SSI relationships with inherent integrity. P1 and P2 are peers since they both have equal control over the relationship.

In addition to removing the need for intermediaries to vouch for the integrity of the relationship, the peer nature of relationships in the Sovrin Network also means that neither party has access to the authentication credentials of the other. Mutual authentication means that each party manages their own keys and never shares the private key with another party. Consequently, attacks, like the recent attack on Twitter accounts (Conger and Popper, 2020) cannot happen because there is no administrator who has access to the credentials of everyone using the system—there is no trove of high-value data.

Relationship Lifespan

Relationships have lifespans. Some relationships are long-lived, some are short-term, and others are ephemeral, existing only for the duration of a single interaction. We typically do not think of it this way, but every interaction we have in the physical world, no matter for what purpose or how short, sets up a relationship. So too in the digital world, although our tools have been sorely lacking in support for anything by long-lived relationships.

The administrative identity systems we have built to service online relationships usually fail to recognize that some relationships are not permanent. Imagine that if whenever you stopped in the convenience store for a cup of coffee, you had to create a permanent relationship with the coffee machine, the cashier, the point of sale terminal, and the customers in line ahead and behind you? Sounds ridiculous. But that is what most digital interactions require. At every turn, we are asked to establish permanent accounts to transact and interact online.

There are several reasons for this. The biggest one is that every web site, app, or service wants to send you ads, at best, or track you on other sites, at worst. Unneeded, long-lived relationships have come to define the modern online experience and are the foundation of the surveillance economy that Shoshana Zuboff describes (Zuboff, 2020).

Relationship Utility

Relationships are established to provide utility. A university wants you to register for classes. An ecommerce site wants to sell you things. A social media site wants to show you ads. Thus, their identity systems, built around the IAM (identity and access management) system, are designed to do far more than just establish the integrity of the relationship. They want to store data about you and your activities.

Thus, any identity system is much larger and more specialized than the IAM portion. All of the account or profile data these companies use are properly thought of as part of the identity system that they build and run. Returning to Joe Andrieu (Andrieu, 2018):

> Identity systems acquire, correlate, apply, reason over, and govern (the) information assets of subjects, identifiers, attributes, raw data, and context.

Regardless of whether or not they outsource the integrity of their relationships using federation, companies still have to keep track of the relationships they have with customers or users in order to provide the service they promise. They cannot outsource this to a third party because the data in their identity system have evolved to suit the needs of the specific relationship. We will never have a single identity that serves all relationships because their unique contexts demand their own identity data. Change the identity system in a Netflix or Amazon and it will not be the same company anymore.

This leads us to a simple, but important conclusion: You cannot outsource a relationship. Online apps and services

[1] For simplicity, we limit the discussion to two-party relationships, but the model can be generalized to multi-party relationships.

[2] To get a feel for the user experience, see the demo at https://try.connect.me.

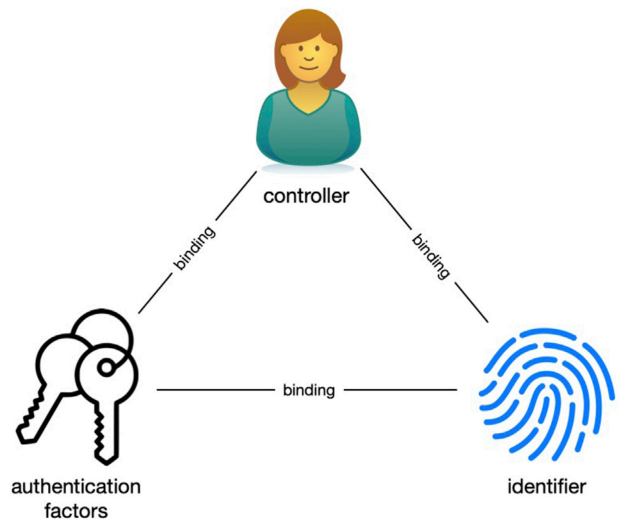


FIGURE 1 | Binding of controller, authentication factors, and identifiers in identity systems.

decorate the relationship with information they observe and use that information to provide utility to the relationships they administer. Doing this and doing it well is the foundation of the modern web.

Consequently, the bad news is that an identity metasystem does not reduce the need for companies to build, manage, and use identity systems. Their identity systems are what make them what they are—there is no "one size fits all" model. But the identity metasystem does make the relationships they form richer, provides a more balanced relationships by providing symmetric value to all parties, and increases flexibility and privacy.

THE ARCHITECTURE OF IDENTITY SYSTEMS

To understand how an identity metasystem like Sovrin Network supports better online relationships, it is useful for clearly understanding the architectures of identity systems.

As we said, identity systems provide the means necessary for remembering, recognizing, and relying on the other parties to the relationship. To do so, they use identifiers, convenient handles that name the thing being remembered. Identifiers are unique within some namespace. The namespace gives context to the identifiers since the same string of characters might be a phone number in one system and a product ID in another.

As shown in **Figure 1**, identifiers are issued to or created by a *controller* who by virtue of knowing the *authentication factors* can make authoritative statements about the identifier (e.g., claiming it by logging in). The controller might be a person, organization, or software system. The controller might be the subject that the identifier refers to, but not necessarily. The authentication factors might be a password, key fob, cryptographic keys, or something else. The strength and nature of the *bindings* between the controller, authentication factors, and identifier determine the strength and nature of the relationships built on top of them.

To understand why that is so, we introduce the concept of a *root of trust*. A root of trust is a foundational component or process in the identity system that is relied on by other components of the system and whose failure would compromise the integrity of the bindings. A primary root of trust cannot be replaced, while a secondary root of trust can be. Together, the roots of trust form the trust basis for the system.

The trust basis enabled by the identity system underlies a particular *trust domain*. The trust domain is the set of digital activities that depend on the binding of the controller to the identifier. For example, binding a customer to an identifier allows Amazon to trust that the actions linked to the identifier are authorized by the controller. Another way to look at this is that the strength of the binding between the identifier and customer (controller) determines the risk that Amazon assumes in honoring those actions.

The strength of the controller–identifier binding depends on the strength of the binding between the controller and the authentication factors and between the authentication factors and the identifier. Attacking either of those bindings reduces the trust we have in the controller–identifier binding and increases the risk that actions taken through a particular identifier are unauthorized.

We can place all identity systems into one of three broad architectural categories based on their structure and primary root of trust:

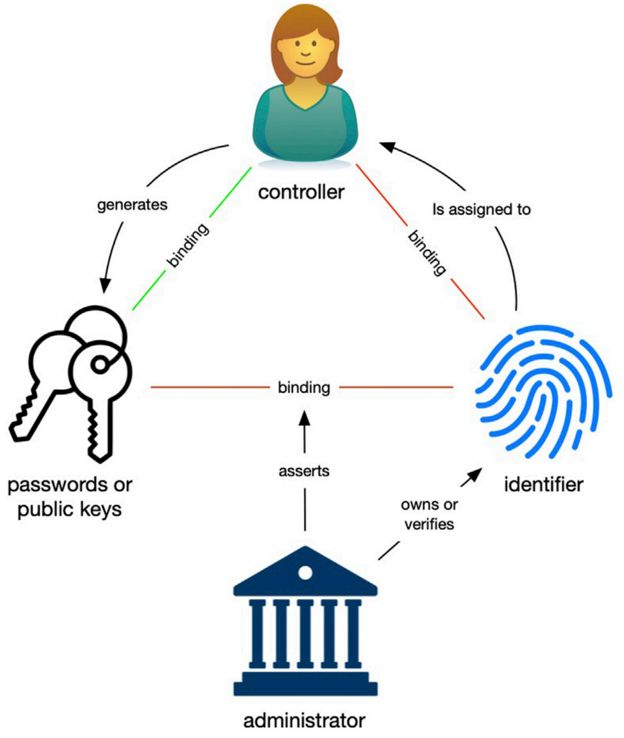


FIGURE 2 | The trust basis in administrative identity systems.

- Administrative
- Algorithmic
- Autonomic

These architectures differ in who controls what. Knowing the locus of control is the primary factor in determining the basis for trust for each. We call this *control authority*. The entity with control authority takes action through operations that affect the creation (inception), updating, rotation, revocation, deletion, and delegation of the authentication factors and their relation to the identifier. How these events are ordered and their dependence on previous operations is important. The record of these operations is the *source of truth* for the identity system.

Administrative Architecture

Identity systems with an administrative architecture rely on an administrator to bind the identifier to the authentication factors. The administrator is the primary root of trust for any domain with an administrative architecture. Almost every identity system in use today has an administrative architecture and their trust basis is founded on the administrator.

Figure 2 shows the interactions between the controller, identifier, and authentication factors in an administrative identity system, the role of the administrator, and the impact these have on the strength of the bindings.

The controller usually generates the authentication factors by choosing a password, linking a two-factor authentication (2FA) mechanism, or generating keys. Even though the identifier might be the controller's email address, phone number, public key, or other ID, the administrator "assigns" the identifier to the controller because it is their policy that determines which identifiers are allowed, whether they can be updated, and their legitimacy within the identity system's domain. The administrator "owns" the identifier within the domain.

The administrator also asserts the binding between the identifier and the authentication factors. An employee's mistake, a policy change, or a hack could affect the binding between the identifier and authentication factors or the identifier and the controller. Consequently, these bindings are relatively weak. Only the binding between the controller and authentication factors is strong because the controller generates them.

The administrator's primary duty is to authoritatively assert the binding between the controller and identifier. Authoritative control statements about the identifier are recorded in the administrator's database, the source of truth in the system, subject to retroactive change by employees and hackers. The administrator might be an ecommerce site that maintains an identity system as the basis for its customer's account. In this case, the binding is private, and its integrity is of interest only to the web site and the customer. Alternatively, the administrator might provide federated login services. In this case, the administrator is asserting the controller–identifier binding in a semi-public manner to anyone who relies on the federated login. A certificate authority is an example of an administrator who publicly asserts the controller–identifier binding, signing a certificate to that effect.

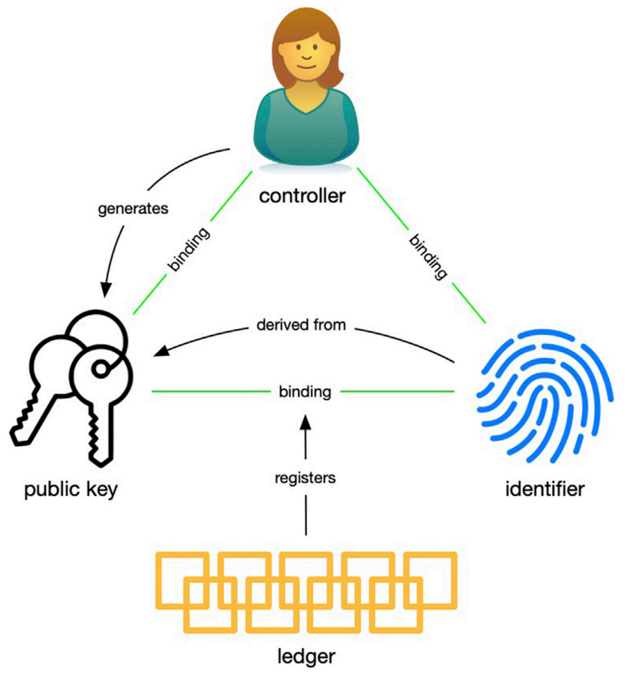


FIGURE 3 | The trust basis in algorithmic identity systems.

Because the administrator is responsible for binding the identifier to both the authentication factors and the controller, the administrator is the primary root of trust and thus the basis for trust in the overall system. Regardless of whether the binding is private, semi-public, or public, the integrity of the binding is entirely dependent on the administrator and the strength of their infrastructure, policies, employees, and continued existence. The failure of any of those can jeopardize the binding, rendering the identity system unusable by those who rely on it.

Algorithmic Architecture

Identity systems that rely on a ledger have an algorithmic architecture. I'm using "ledger" as a generic term for any algorithmically controlled, distributed-consensus-based datastore including public blockchains, private blockchains, distributed file systems, and others. Of course, it is not just algorithms. Algorithms are embodied in code, written by people, running on servers. How the code is written, its availability to scrutiny, and the means by which it is executed all impact the trust basis for the system. "Algorithmic" is just shorthand for all of this.

Figure 3 shows how the controller, authentication factors, identifier, and ledger are bound in an identity system with an algorithmic architecture. As in the administrative identity system, the controller generates the authentication factors, albeit in the form of a public–private key pair. The controller keeps and does not share the private key. The public key, on the other hand, is used to derive an identifier (at least in well-designed SSI systems) and both are registered on the ledger. This registration is the inception of the controller–identifier binding since the controller can use the private key to assert her control over the identifier as

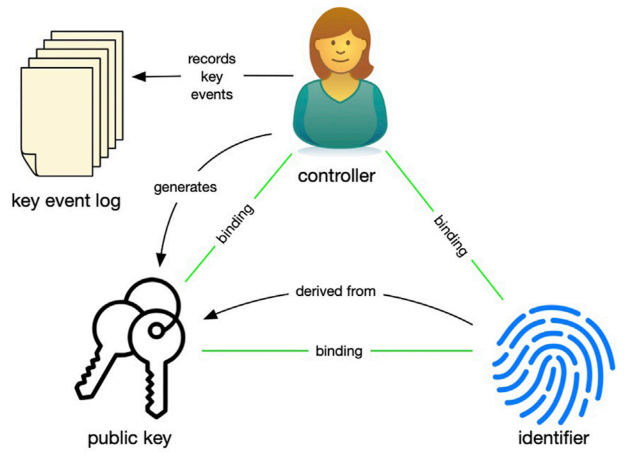


FIGURE 4 | Trust basis in autonomic identity systems.

registered on the ledger. Anyone with access to the ledger can algorithmically validate the controller–identifier binding.

The controller makes authoritative control statements about the identifier. The events marking these operations are recorded on the ledger, which becomes the source of truth for anyone interested in the binding between the identifier and authentication factors.

In an identity system with an algorithmic trust basis, computer algorithms create a ledger that records the key events. The point of the ledger is that no party has the power to unilaterally decide whether these records are made, modified, or deleted and how they are ordered. Instead, the system relies on code executed in a decentralized manner to make these decisions. The nature of the algorithm, the manner in which the code is written, and the methods and rules for its execution all impact the integrity of the algorithmic identity system and consequently any bindings that it records.

Autonomic Architecture

Identity systems with an autonomic architecture function similarly to those with an algorithmic architecture. As shown in **Figure 4**, the controller generates a public–private key pair, derives a globally unique identifier, and shares the identifier and the currently associated public key with the party she wishes to create a relationship with.

The controller uses her private key to authoritatively and non-repudiably sign statements about the operations on the keys and their binding to the identifier, storing those in an ordered key event log[3]. One of the important realizations that make autonomic identity systems possible is that the key event log must only be ordered in the context of a single identifier, not globally. So, a ledger is not needed for recording operations on identifiers that need not be publicly validated. The key event log can be shared with and verified by anyone.

The controller also uses the private key to sign statements that authenticate herself and authorize use of the identifier. A digital signature also provides the means of cryptographically responding to challenges to prove her control of the identifier. These self-authentication and self-authorization capabilities make the identifier self-certifying and self-managing, meaning that there is no external third party, not even a ledger, needed for the controller to manage and use the identifier and prove to others the integrity of the bindings between herself and the identifier. Thus, anyone (any entity) can create and establish control over a personal identifier namespace in a manner that is independent, interoperable, and portable without recourse to any central authority. Autonomic identity systems rely solely on self-sovereign authority.

Autonomic identifiers have a number of advantages:

- Self-Certification—autonomic identifiers have no reliance on a third party.
- Self-Administration—autonomic identifiers can be independently administered by the controller without reliance on a third party.
- Low Cost—autonomic identifiers are virtually free to create and manage.
- Security—because the keys are decentralized, there is no trove of secrets that can be stolen.
- Regulatory—autonomic identifiers need not be publicly shared or stored in an organization's database, and consequently reduce regulatory concern over personal data.
- Scale—autonomic identifiers scale with the combined computing capacity of all participants, not a central system.
- Independent—autonomic identifiers are not dependent on any specific technical system or even being online.

CREDENTIAL EXCHANGE AS THE FOUNDATION FOR ONLINE IDENTITY

In the physical world, people collect and manage credentials from various sources including governments, financial institutions, employers, schools, businesses, family, colleagues, and friends. Individuals also assert information themselves. These various credentials serve different purposes. We have credentials that we use often and carry around with us. We have important credentials we file away and even some we keep in safe deposit boxes. Some, like boarding passes, we use once, then throw away. Others, like birth certificates, we keep for our entire life.

We use credentials, alone or in concert with other credentials, when we need to prove something about ourselves. We present credit cards to prove we are authorized to charge an account. We present a driver's license to prove we are of legal age at a bar. We present letters from our employer to prove our salary when applying for a loan. The credential verifier is free to determine whether to trust the credential or not.

[3] A number of cryptographic systems are trivially self-certifying (e.g., PGP, Ethereum, and Bitcoin). What sets the autonomic identity systems described here apart is the key event log.

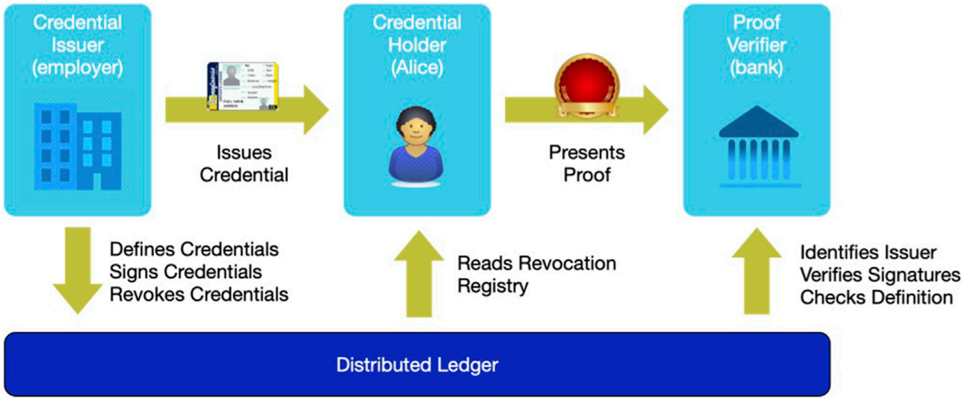


FIGURE 5 | Parties to credential exchange and their interactions with each other and the ledger.

Identity does not work that way online. As we have discussed, online identity has traditionally been administrative, centralized, and built for specific purposes. Various, so-called "identity providers" authenticate people using usernames and passwords and provide a fixed, usually limited, set of attributes about the subject of the identity transaction. The identity information from these systems is usually used within a specific, limited context. For example, federated login (e.g., Log in with Google) allows a login to be used across contexts, but the kind of information shared is limited and its provenance is often difficult to determine. These various administrative identity systems are not interoperable, making it hard to combine attributes from one with those of another. Consequently, online identity is one-dimensional and has limited value.

In credential exchange, there are three parties: the credential issuer, the credential holder (sometimes called the identity owner), and the credential verifier (also known as the relying party).

A credential is a collection of claims (i.e., attributes) that is signed by the issuer and held by the identity owner. Credentials conform to the Verifiable Credential specification (Sporny et al., 2019). While the word "credential" conjures images of formal documents, almost anything representable in JSON that needs to be attested can be a credential. So, while things like passports and driver's licenses fit this bill, so do things like membership cards, boarding passes, school report cards, invoices, purchase orders, and store receipts.

Figure 5 shows how credential exchange works. Suppose Alice (the identity owner) is applying for a loan at her local bank (the credential verifier). The bank requires proof that Alice is employed and makes at least $70,000 per year. Alice's employer (the credential issuer) has issued an employment credential that includes her employment status and her current salary. The credential might also include many other attributes related to Alice's job. Alice holds the employment credential and can present it to prove to the bank that she is employed and makes more than $70,000.

When Alice proves her employment status to the bank online, she does not present the entire credential since doing so would reveal more information than is necessary. Instead, Alice presents just the information the bank needs using a cryptographic technique known as "zero knowledge proof." The ability to limit the information presented from a credential is important to maintain privacy through the principle of minimal disclosure.

The online model for verifiable credentials has five important characteristics that mirror how credentials work in the offline world:

- Credentials are decentralized and contextual. There is no central authority for all credentials. Every party can be an issuer, a holder (identity owner), or a verifier. Verifiable credentials can be adapted to any country, any industry, any community, or any set of trust relationships.
- Credential issuers decide on what data are contained in their credentials. Anyone can write credential schemas to the ledger. Anyone can create a credential definition based on any of these schemas.
- Verifiers make their own decisions about which credentials to accept—there is no central authority who determines what credentials are important or which are used for what purpose.
- Verifiers should not need to contact issuers to perform verification. Credential verifiers do not need to have any specific technical, contractual, or commercial relationship with credential issuers.
- Credential holders are free to choose which credentials to carry and what information to disclose. People and organizations are in control of the credentials they hold (just as they are with physical credentials) and determine what to share with whom.

In addition to these five characteristics, credential presentment *via* zero-knowledge proofs offers important privacy protection to the credential holder (Lodder, 2018). ZKP presentment.

- increases the cost of correlation since the identifier of the holder is blinded and other data can be excluded if the verifier does not need it;

- reduces the parties with whom the data are shared;
- supports incremental disclosure as a relationship becomes more trusted;
- restricts the attributes that are shared to just the needed subset of what is contained in the credential; and
- empowers the holder to restrict resharing.

These powerful protections against correlation increase privacy in a structural way and make possible more effective regulation of verifiers.

RISK AND TRUST: CREDENTIAL FIDELITY AND PROVENANCE

Trust is a popular term in the identity community. Some people rightly ask about risk whenever someone in the identity community talks about trust. Because of the proximity problem, digital relationships are potentially risky. One of the goals of an identity system is to provide evidence that can be used in the risk calculation.

In their excellent paper, Risk and Trust (Nickel and Vaesen, 2012), Philip Nickel and Krist Vaesen define trust as the "disposition to willingly rely on another person or entity to perform actions that benefit or protect oneself or one's interests in a given domain." From this definition, we see why crypto proponents often say, "To trust is good, but to not trust is better." The point being that not having to rely on some other human, or human-mediated process is more likely to result in a beneficial outcome because it reduces the risk of non-performance.

Relationships imply a shared domain, context, and set of activities (Wilson, 2011). We often rely on third parties to tell us things relevant to the relationship. Our vulnerability, and therefore our risk, depends on the degree of reliance we have on another party's performance. Relationships can never be "no trust" because of the very reasons we create relationships. Bitcoin, and similar systems, can be low or no trust precisely because the point of the system is to reduce the reliance on any relationship at all.

The architecture of the identity metasystem significantly limits the ways we must rely on external parties for the exchange of information *via* verifiable credentials and thus reduces the vulnerability of parties inside and outside of the relationship. The design of the identity metasystem clearly delineates the parts of the system that are low trust and those where human processes are still necessary.

Our basis of trust in the physical world is other humans We interact with people directly, recognizing, remembering, responding to, and relying on them (Andrieu 2018). As we pointed out in *The Problems of Online Identity*, in the digital realm, we are not proximate to the parties we have a relationship with. As a result, credential exchange replaces the human basis of trust in the physical world with algorithmic and autonomic bases of trust.

Returning to the example in the last section, Alice's bank needs two levels of trust: first it needs to know the credential is authentic. Second, the bank wants to verify the veracity of the contents of the credential.

With respect to credential authenticity, the bank wants to know:

1. Who issued the credential,
2. That the credential was issued to Alice,
3. That the credential has not been tampered with, and
4. That the credential has not been revoked.

The metasystem provides these properties cryptographically (Hardman 2018). We call the properties that the metasystem provides *credential fidelity*. Fidelity is cryptographic. The bank can verify these four properties by looking at the credential definition on the ledger, retrieving the issuer's public DID from the definition, resolving the DID to get the public key of the issuer, and using the public key to check the signature of the credential to ensure it has not been tampered with. The bank can also cryptographically verify that the credential was issued to Alice. As part of making her proof from the credential, Alice also proves that it has not been revoked by referencing a revocation registry on the ledger. The ledger ensures that the bank can do all of this without contacting the employer, helping preserve Alice's privacy.

Fidelity allows the bank to verify the credential as a container, but fidelity does not prove the veracity of the statements within the credential. Generally, credential veracity depends on the reputation of the issuer. More specifically, we establish it through *credential provenance*.

In this example. the bank wants to know that the issuer identifier in the credential is associated with a legitimate business[4], the details of that business, and what others have said about that business so they can judge the veracity of the statements made in the credential. The bank has several options depending on their internal policies.

- They could use an out-of-band method to validate the identifier of the issuer by, say, looking up the public DID of the issuer on the issuer's web site.
- They could ask that the bank prove things to them by establishing a direct DID-based relationship with the bank and requesting data from the credentials the bank holds (e.g., their FIDC membership).
- The banking industry could create an industry-specific governance framework and list the public DIDs of its members in a public registry that anyone could access.

Determining the provenance of the credential's content cannot be done through purely technical means. Clearly, technology can help, but unlike credential fidelity where cryptography alone can prove credential authenticity, provenance is a matter of human process, policy, regulation, and law.

[4] I am saying business, but in fact this could apply to any entity that can issue credentials including people and things.

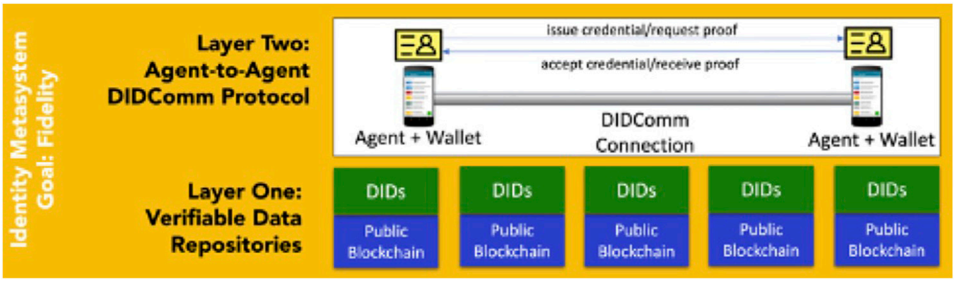


FIGURE 6 | The identity metasystem.

THE IDENTITY METASYSTEM

The identity metasystem embodied in the Sovrin Network provides three primary capabilities that allow it to be used as the basis for any context-specific identity system that is needed:

- Relationships—the architecture must allow people, organizations, and things to have relationships with each other.
- Messaging—the architecture must support messaging between the parties to those relationships.
- Trustworthy Attribute Exchange via Verified Credentials—parties to relationships must be able to reliably exchange information about attributes (often called claims by identity professionals).

The architecture for the identity metasystem supplies these features using layers that build on each other as depicted in **Figure 6**.

The metasystem is a hybrid architecture, using both algorithmic and autonomic identifiers to provide these capabilities.

Autonomic Identifiers Support Relationships and Messaging at Layer 2

When Alice forms a relationship with her friend, her colleague, her doctor, her employer, an ecommerce web site, or even her thermostat, she uses an autonomic identifier based on the Peer DID specification (Peer DID, 2020). Alice and other parties use agents based on the Hyperledger Aries open-source code (Hyperledger Aries, 2020). The user interface to these agents is called a wallet.

To form a relationship, Alice and the other party each generate a new peer DID and send it to the other. Peer DIDs need not be publicly resolved since both parties know about the other. The result is a network of peer-to-peer relationships between agents under the control of the people and organizations forming relationships. This forms Layer 2 of the identity metasystem in **Figure 6**.

Because the parties have exchanged DIDs, each party can authenticate the other. Mutual authentication allows the relationship to have integrity without an intervening third part (Young, 2020).

The relationship created by exchanging Peer DIDs is useful for more than mutual authentication. The mutually authenticated channel supports a uniform and democratic protocol for secure interaction called DIDComm (DID Communications, 2020). The DIDComm protocol allows parties to a relationship to securely and privately share authenticated messages. The security and authority of a DIDComm channel are rooted in DIDs and their associated authentication factors. DIDComm can be used over a wide variety of transports.

One of the primary uses of the DIDComm channel is to support the verification of key events following a key rotation. Whenever one of the parties needs to rotate their keys, they make an entry in their key event log (called "deltas" in the Peer DID specification) that records the relevant operations on the keys in a cryptographic manner. The key event log is a chain of signed change records that can be cryptographically verified. The parties in a DID-based relationship share (using a CRDT) key event logs for each identifier. If either party updates the keys associated with the DID, the other is informed of change.

But beyond that core functionality, DIDComm support message exchanges for many purposes. Aries RFCs (Hyperledger Aries RFCs, 2021) describe protocols for several core use cases for DIDComm including:

- Establishing peer DID Connections
- Requesting and issuing credentials
- Presenting a credential proof

Future use cases could include protocols for the following:

- Payments
- Interactions with IoT devices
- Buying and selling

Vic Cooper likened DID-based P2P messaging to the Batphone (Windley, 2020b). When Batman picks up the Batphone to talk with Commissioner Gordon, Commissioner Gordon does not start off the conversation with "Who am I speaking to?," "Can you give me your account number?," "What's your date of birth?," or "What street did you live on in Junior High?" When Commissioner Gordon picks up the Batphone, he knows it is Batman on the other end. Only Batman can call on the Batphone.

DIDComm-based messaging is like having a Batphone for every digital relationship you have. You and they know they are communicating with the right party. All the messages are authenticated and protected from eavesdroppers.

DID Messaging could revolutionize how we talk to each other and how we communicate with businesses.

- We no longer have to rely on a correlatable identifier like an email or phone number, to identify, discover, or connect to the other party.
- We no longer have to use centralized systems to talk to other parties with the attendant risk of the system being down or the conversation not being private.
- We save time and money using frictionless, direct communications with companies we need to work with.
- We can verify who is at the other end by asking them to prove things to us.
- We can sever one relationship without affecting others since everyone has a different identifier for us.

Agents exchange attributes over the channel created in Layer 2 using a flexible, decentralized system of credential exchange as discussed in *Credential Exchange as the Foundation for Online Identity*.

Algorithmic Identifiers Support Credential Exchange

The metasystem's algorithmic identifiers also take the form of DIDs. But rather than the peer DIDs used at Layer 2, DIDs at Layer 1 are public DIDs. DIDs have a number of important properties that make them ideal as identifiers in an algorithmic system. Specifically, they are non-reassignable, resolvable, cryptographically verifiable, and decentralized.

As algorithmic identifiers, DIDs allow the controller to make cryptographically authoritative statements about the identifier and the keys it is bound to. Those statements are recorded on a ledger to provide a record of the key events that anyone with access to the ledger can evaluate.

The DID specification provides for many DID methods such that DIDs may be recoded on a variety of data stores. There is nothing in the DID specification itself that requires that the data store be a blockchain or ledger, but that is the primary use case. The collection of ledgers supporting the binding of public DIDs to their authentication factors forms Layer 1 of the metasystem shown in **Figure 6**.

The record on the ledger is public since the purpose of putting DIDs on a ledger is to allow parties who do not have an existing relationship to evaluate the identifier and its linkage to the controller and public keys. The ledger provides several important features:

- The ledger creates a circuit breaker so that issuers do not know when and where credentials are being used, increasing the privacy of the transaction. Consequently, the metasystem structurally supports the privacy of participants.
- The ledger enables offline exchange of credentials. This not only supports verification of a credential when the issuer is offline, but support for state proofs in the ledger allows exchange to occur when all the parties are offline but the holder and verifier can connect over some local network (e.g., Bluetooth).

The metasystem makes use of the resolvability of DIDs to support credential exchange. Issuers (in the role of controller) register DIDs on a public ledger and issue credentials using that identifier. When the credential holder proves attributes to a verifier, she also proves the identifier of the issuer. The verifier can resolve the DID for the issuer from the ledger as part of ensuring the fidelity of the credential exchange.

BUILDING IDENTITY SYSTEMS ON THE METASYSTEM

The capabilities of the identity metasystem provide a sure foundation for creating identity systems that are secure and support the autonomy and privacy of people and organizations. The goal of an identity metasystem, like Sovrin Network, is to connect individual identity systems and allow them to interoperate since no single system meets the needs of every digital identity scenario.

As we discussed in *Relationships*, the goal of the metasystem is to support relationships between parties online and provide a secure, private means of exchanging verified credentials. The metasystem uses credential exchange on top of DIDComm messaging at Layer 2 as the unifying protocol for exchanging identity information. In credential exchange, an issuer issues a credential to a person or organization called the holder. The holder holds one or more credentials and uses the protocols provided by the metasystem to prove things about themself to a verifier who needs trustworthy attributes. **Figure 7** shows the layers of this system.

The blue box on the top of **Figure 7** represents an identity system built on top of the metasystem. There is more than one identity system. In fact, there are tens of millions, maybe more. Every credential definition represents a new identity system created for a specific context. Anyone can define a credential for any purpose. And even though each identity system stands alone for its own purpose, they are interoperable because they are built on top of the metasystem and employ common protocols.

For example, Alice may have a credential representing her driver's license and one representing her employee ID. These are designed for a specific purpose by the DMV and the employer. Yet, because they are based on a metasystem and use a common protocol, she could go to the bank and use those in concert to prove that she is employed (employee ID) and her date of birth (driver's license) in one operation.

The two systems shown in **Figure 7** have different properties. The identity metasystem (orange box) provides important assurances about the fidelity of the credential. A credential verifier who receives a proof is concerned about credential fidelity, but they are also concerned with the credential's provenance. The fidelity provided by the identity metasystem, combined with the credential provenance provided by the

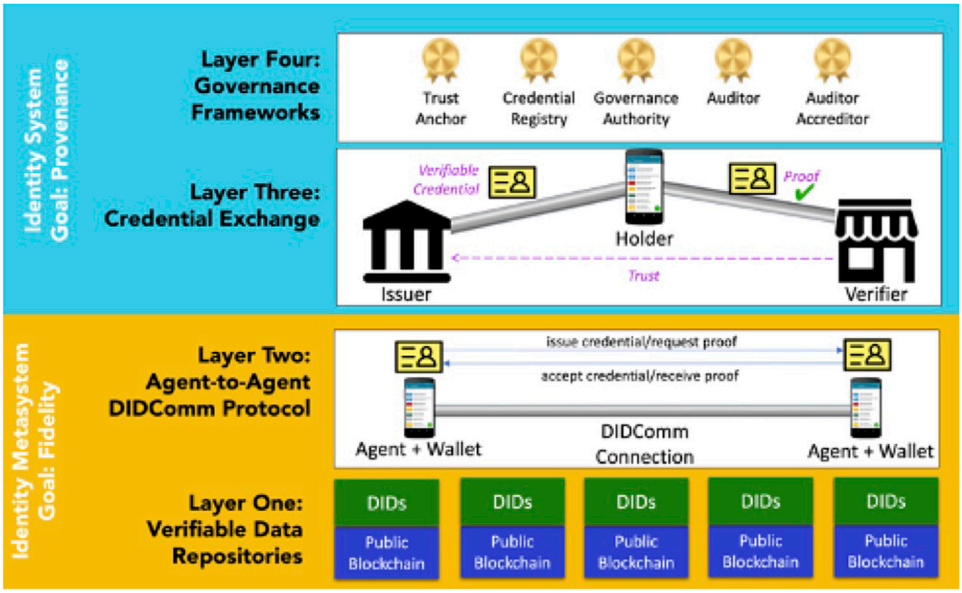


FIGURE 7 | Identity on the metasystem.

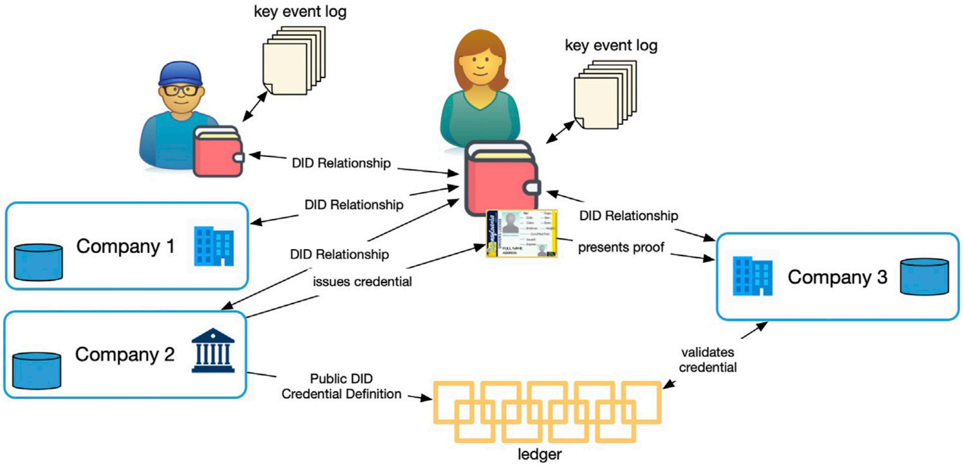


FIGURE 8 | Relationships and interactions in the Sovrin Network.

context-specific identity system operating on top of it, provides the basis for trusting the information that the holder has conveyed through credential exchange.

Operationalizing Digital Relationships

Figure 8 shows the relationships and interactions in the Sovrin Network. In the figure, Alice has an SSI wallet[5]. Alice's SSI wallet is like other wallets she has on her phone with several important differences. First, it is enabled by open protocols, and second, it is entirely under her control. She uses the wallet to manage her relationship with Bob as well as a host of organizations.

This diagram has elements of each architectural style described in *The Architecture of Identity Systems*. Alice has relationships with four different entities: her friend Bob and three different companies. These relationships are based on autonomic identifiers in the form of peer DIDs.

Company 2 has an algorithmic identifier in the form of a public DID that has been recorded on the ledger along with a credential definition. Company 2 has, based on that credential definition and its associated public DID, issued a credential to

[5] I am using the term "wallet" fairly loosely here to denote not only the wallet but also the agent necessary for the interactions in an SSI ecosystem. For purposes of this article, delineating them is not important. In particular, Alice may not be aware of the agent, but she will know about her wallet and see it as the tool she uses.

Alice. The contents of that credential are based on the information Company 2 knows from its relationship with Alice, stored in its *internal* administrative identity system.

Alice has presented a proof based on the credential to Company 3 who can validate its fidelity using the credential definition on the ledger. Company 3 likely has its own internal administrative identity system where it stores information about its relationship with Alice.

The peer DIDs that Alice presented to Company 2 and Company 3 are different. Nevertheless, the cryptographic procedures of the zero-knowledge proof (ZKP) that Alice presents to Company 3 ensure that Company 3 can know that the credential used as the basis of the proof was issued to the same person who they have a relationship with. More generally, Company 3 knows that the same entity controls the keys for the DID Alice shared with Company 2 and the DID she shared with them.

Company 2 does not issue the verifiable credential to the peer DID Alice gave them. In Hyperleger Aries credential proof, Alice creates a blinded link secret and sends it to Company 2 in response to a credential offer. The verifiable credential contains the blinded link secret. When Alice uses ZKP to prove attributes from her credentials, the blinded secret is what proves Alice is the same Alice to whom all the credentials she used were issued. The proof contains a special predicate showing that the link secret in the credential, if unblinded, would be the same as the link secret Alice shared with Company 3, if unblinded. No unblinding actually happens. Since the credential is not linked directly to the peer DID, but indirectly through the blinded link secret, Alice is free to rotate the DID-associated keys underneath the credential without invalidating it. And the DID continues to serve its purpose of identifying Alice to Company 2 (Hardman 2018).

Because Alice uses different peer DIDs for Company 2 and Company 3, they cannot correlate data they have about her through the identifier independently. They need Alice, who controls the link secret, to correlate the information for them. That ensures Alice is in control of what information is shared and correlated based on the peer DID relationships.

Identity Systems

When we say "digital identity system", most people probably think of just one thing: authentication. The digital identity systems we have built over the last 30 years are so anemic that it is difficult for us to imagine the kind of rich identity systems that exist in the physical world being available online.

In the offline world, we use credentials to prove things about ourselves to others. Each of these credentials constitutes an identity system, designed and built for a specific purpose in a given context. For example, businesses frequently give employees ID cards. I have one for Brigham Young University (BYU), my employer. I can use it to open doors, get a discount at the bookstore, get a car from the motor pool, and even ride a local bus or train. This flexible identity system allows the university to add new functionality over time as needs change. The university sets the rules about who gets an ID card and what it means. Of course, it also has use outside the context of the university, say, for example, at a store that gives discounts to university employees and is willing to accept the ID card as proof of employment.

Businesses are full of credentials. Each one represents an identity system designed and built for a specific context. Every form or official piece of paper is a potential credential. Every bundle of data transmitted in a workflow is a potential credential. Here are a few examples of common credentials:

- Employee badges
- Driver's license
- Passport
- Wire authorizations
- Credit cards
- Business registration
- Business licenses
- College transcripts
- Professional licensing (government and private)

Here are some others that may not be typically thought of as credentials, but fit the definition:

- Invoices and receipts
- purchase orders
- Airline or train ticket
- Boarding pass
- Certificate of authenticity (e.g., for art, other valuables)
- Gym (or any) membership card
- Movie (or any) tickets
- Insurance cards
- Insurance claims
- Titles (e.g., property, vehicle, etc.)
- Certificate of provenance (e.g., non-GMO, ethically sourced, etc.)
- Prescriptions
- Fractional ownership certificates for high value assets
- CO2 rights and carbon credit transfers
- Contracts

Since even a small business might issue receipts or invoices, have customers who use the company web site, or use employee credentials, most businesses will define at least one credential, and many will need many more. There are potentially tens of millions of different credential types. Many will use common schemas but each credential from a different issuer constitutes a different identity credential for a different context.

With the ongoing credential work in Hyperledger Aries (Hyperledger Aries 2020), these use cases expand even further. With upcoming "redeemable credentials" feature, issuers can double-spend-proof proving credential possession without a ledger. This works for all kinds of redemption use cases like clocking back in at the end of a shift, voting in an election, posting an online review, or redeeming a coupon.

You might notice that many of the things listed above are solutions some people advocate building entire blockchains for. That is overkill when you can use a credential to get the job done. Especially when that credential is interoperable with others in a ubiquitous identity metasystem. By double-spend-proofing credentials, you create a system capable of representing value of all sorts. An identity metasystem for trustworthy credential exchange has uses far beyond what we might typically think of as an "identity system."

A Marketplace for Credentials

Many credentials will be created for internal or non-commercial purposes (like the employee credential). But some will have a supporting business model. This is exactly what happens offline where many credentials are exchanged for money. The metasystem should support credential business models to achieve ubiquity. Daniel Hardman discusses this in his excellent blog post about Categorizing Verifiable Credentials (Hardman, 2019b).

Credentials may intersect with payment in different ways. Some may be issued and used for free; others may be purchased; still others may incur a fee with every use. And while payment could be viewed as entirely independent from credentials, the binding is actually more interesting. This is because economics and levels of assurance are intertwined. For example, a top-secret security clearance may require thousands of dollars of field work and investigation and bump its holder's salary by even more. Thus, business models that allow economic value to be harvested in credential interactions are important.

With non-free credentials, who pays whom is interesting. The most straightforward model is holder-pays-issuer; we already expect to pay a fee when we apply for a passport. But other variations are equally possible, and they represent potential innovation that is impractical with physical credentials. For example, a holder who is applying to a university might pay the university a fee to verify their academic credentials. A potential employer with stringent security requirements might pay an issuer to achieve assurance that an applicant has a government security clearance. A medical researcher might pay a holder for the privilege of verifying genetic information from credentials, as part of a study they are conducting.

While it is impossible to anticipate every possible credential use case that includes a reciprocal exchange of value, looking at a few use cases is instructive. The following use cases are just for the Holder-Pays-Issuer pattern, but other patterns, like Verifier-Pays-Issuer, are possible.

Driver's License—Driver's licenses are an excellent example of a credential people pay for. There are 112 million licensed drivers just in the US. If we assume each license costs $30 and is renewed every 5 years, almost $700 million is paid per annum for driver's licenses.

Memberships—Memberships in gyms are just one example of a membership credential where the credential holder pays the issuer. Gym membership revenues in the US in 2018 was $32 billion according to Wellness Creatives[6]. There are many more membership types that could be built on top of Sovrin Network.

Movie Tickets—Movie tickets are another credential that is bought. In 2018, 1.3 billion movie tickets were sold in the US[7]. At $10 per ticket, that is $13 billion.

Airline Tickets—Airline tickets are a special kind of credential that is purchased. According to IATA, there were 4.1 billion airline passengers in2017[8]. The US Department of Transportation Bureau of Transportation Statistics reports that average airfare was $347 that same year[9]. We can estimate that worldwide airfare was about $1.4 trillion in 2017.

Online Sales—Online sales could be accomplished using Holder-Pays-Issuer credential exchange. By paying for the receipt (a credential) equal to the amount of the order, we can view all of ecommerce as a form of paid credential issuance. Linking payment to a credential and placing it inside a wallet that emphasizes relationships and credential management may make credential-related payments an important component of online retail. US online retail sales were $519 billion in 2018[10].

These are just a few potential use cases where credentials and value are exchanged. While not all of these will necessarily come to pass, it is easy to conclude that the potential marketplace for credentials is in the trillions of dollars. The identity metasystem, with its mutually authenticated messaging protocol, is an excellent platform for supporting commercial credential exchange. These workflows, with built-in value exchange, can be developed on the identity metasystem.

An identity metasystem like the Sovrin Network provides the foundation for creating tens of millions of interoperable identity systems for every conceivable context and use. By virtue of being built on the metasystem, these identity systems share a common protocol and similar user experience. The metasystem is available to all and is decentralized, allowing each participant to make their own decisions about what identity systems they will build and participate in to support their goals and ambitions.

CONCLUSION: LIFE-LIKE DIGITAL IDENTITY

We use identity in the physical world without thinking about it. And when we do, there are patterns that are so ingrained in our ways of interacting that we do not give them a second thought. If we are to move more and more of our lives to the digital realm while also preserving agency and autonomy, we must create a digital world that allows us to jump the trust gap we inevitably have with people, organizations, and things when our interaction is digital.

An identity metasystem provides the long-missing identity layer for the Internet that will allow this to happen. The metasystem can be incorporated into every digital tool and system providing a consistent, trustworthy experience that feels as frictionless and natural as identity in the physical world.

The identity metasystem overcomes the problems of digital identity described in *The Problems of Online Identity*. We have described a system that carefully uses cryptography to overcome the problems introduced by distance while providing autonomy

[6] https://www.wellnesscreatives.com/gym-market-statistics/.

[7] https://www.statista.com/statistics/187073/tickets-sold-at-the-north-american-box-office-since-1980/.

[8] https://www.iata.org/en/pressroom/pr/2018-09-06-01/.

[9] https://www.bts.gov/content/annual-us-domestic-average-itinerary-fare-current-and-constant-dollars.

[10] https://www.digitalcommerce360.com/article/us-ecommerce-sales/.

and flexibility for people and organizations without compromising strong privacy and workable anonymity. The nature of credential exchange based on an interoperable protocol specification introduces a system for building myriad identity systems that provide a more life-like experience than current, disconnected administrative identity systems.

Decentralized, self-sovereign identity depends on an identity metasystem and is the foundation for a decentralized web—a web that flexibly supports the kind of ad hoc interactions people have with each other all the time in real life. We will never get an online world that mirrors real life and feels frictionless and life-like until we do.

Consequently, the arguments for creating the identity metasystem provided by Sovrin Network are not narrow or technical issues. Sovrin Network does not merely provide narrow technical benefits. Rather, the identity metasystem is vital for personal autonomy and ultimately human rights. Computers are coming to intermediate every aspect of our lives. Our autonomy and freedom as humans depend on how we architect this digital world. Unless we put digital systems under the control of the individuals they serve without intervening administrative authorities, the internet will undermine the quality of life it is meant to bolster. The identity metasystem is the foundation for doing that.

AUTHOR'S NOTE

Parts of this article have appeared previously on the author's blog at https://www.windley.com.

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

PW is solely responsible for the content of this publication.

REFERENCES

Andrieu, Joe. (2018). Five Mental Models of Identity. Available at: https://github.com/WebOfTrustInfo/rwot7-toronto/blob/master/topics-and-advance-readings/five-mental-models-of-identity.md (Accessed Nov 5, 2020).

Cameron, Kim. (2005). The Laws of Identity. Available at: https://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf (Accessed Nov 5, 2020).

Conger, Kate., and Popper, Nathaniel. (2020). *Florida Teenager Is Charged as 'Mastermind' of Twitter Hack*. New York Times. Available at: https://www.nytimes.com/2020/07/31/technology/twitter-hack-arrest.html (Accessed Nov 5, 2020).

Decentralized Identifiers (DIDs) specification v1.0 (2020). Decentralized Identifiers (DIDs) Specification v1.0. W3C Working Draft 08 October 2020, Available at: https://www.w3.org/TR/did-core/ (Accessed Nov 5, 2020).

DID Communication Working Group (2020). Available at: https://identity.foundation/working-groups/did-comm.html (Accessed Nov 5, 2020).

Hardman, Daniel. (2019b). Categorizing Verifiable Credentials. Available at: https://www.evernym.com/blog/categorizing-verifiable-credentials/ (Accessed Nov 5, 2020).

Hardman, Daniel. (2018). How DIDs, Keys, Credentials, and Agents Work in Sovrin, Sovrin Technical Report. Available at: https://sovrin.org/library/how-dids-keys-credentials-and-agents-work-in-sovrin/ (Accessed Nov 5, 2020).

Hardman, Daniel. (2019a). The Dangerous Half-Truth of "We'll Be Correlated Anyway". Available at: https://www.evernym.com/blog/well-be-correlated-anyway/ (Accessed Nov 5, 2020).

Hyperledger Aries Project (2020). Available at: https://www.hyperledger.org/use/aries (Accessed Nov 5, 2020).

Hyperledger Aries RFCs (2021). Available at: https://github.com/hyperledger/aries-rfcs (Accessed June, 2021).

Lodder, Mike. (2018). The Sovrin Network and Zero Knowledge Proofs. Available at: https://sovrin.org/the-sovrin-network-and-zero-knowledge-proofs/(Accessed Nov 5, 2020).

Naik, N., and Jenkins, P. (2016). "An Analysis of Open Standard Identity Protocols in Cloud Computing Security Paradigm," in 2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, Auckland, New Zealand, 8-12 Aug. 2016 (IEEE), 428–431. doi:10.1109/DASC-PICom-DataCom-CyberSciTec.2016.85

Nickel, Philip. J., and Vaesen, K. (2012). "Risk and Trust," in *Handbook of Risk Theory*. Editors Sabine Roeser, Rafaela Hillerbrand, Martin Peterson, and Per Sandin (Springer).

Peer DID Method Specification (2020). Peer DID Method Specification. W3C Document 25 August 2020, Available at: https://identity.foundation/peer-did-method-spec/index.html (Accessed Nov 5, 2020).

Simmonds, P. (2015). The Digital Identity Issue. *Netw. Security* 2015, 8–13. doi:10.1016/s1353-4858(15)30069-6

Smith, Samuel. (2020). Key Event Receipt Infrastructure (KERI). Available at: https://arxiv.org/abs/1907.02143 (Accessed Nov 5, 2020).

Sporny, M., Noble, G., Longley, D., Burnett, D., and Zundel, B. (2019). Verifiable Credentials Data Model 1.0. Nov 19 2019. Available at: https://www.w3.org/TR/2019/REC-vc-data-model-20191119/.

Wilson, Stephen. (2011). "Identities Evolve: Why Federated Identity Is Easier Said Than Done," in AusCERT 2011 Conference: "Overexposed" Gold Coast, Australia. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2163241 (Accessed Nov 5, 2020).

Windley, Phillip. J. (2020). *Authentic Digital Relationships*. Lindon, UT, United States: Technometria. Available at: https://www.windley.com/archives/2020/08/authentic_digital_relationships.shtml (Accessed Nov 20, 2020).

Windley, Phillip. J. (2020b). *DIDComm And the Self-Sovereign Internet*. Lindon, UT, United States: Technometria. Available at: https://www.windley.com/archives/2020/11/didcomm_and_the_self-sovereign_internet.shtml (Accessed Nov 20, 2020).

Windley, Phillip. J. (2018). *You've Had An Automobile Accident: Multi-Source Identity To the Rescue*. Lindon, UT, United States: Technometria. Available at: https://www.windley.com/archives/2018/08/youve_had_an_automobile_accident_multi-source_identity_to_the_rescue.shtml (Accessed Nov 5, 2020).

Young, Kaliya. (2020). Understanding DIDComm. Available at: https://medium.com/decentralized-identity/understanding-didcomm-14da547ca36b (Accessed Nov 5, 2020).

Zuboff, Shoshana. (2020), The Age Of Surveillance Capitalism: The Fight For a Human Future At the New Frontier Of Power, (PublicAffairs) January 2019.

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.

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

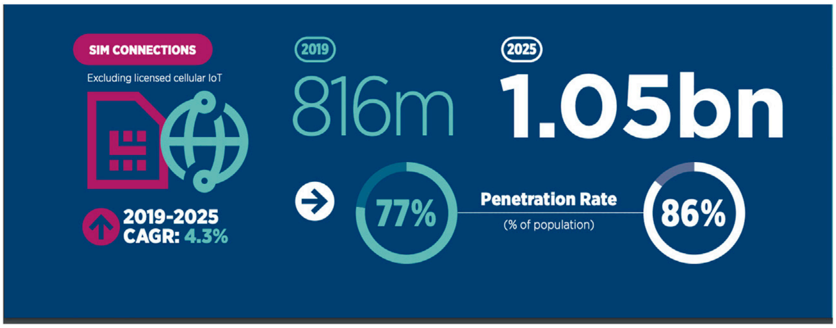


FIGURE 1 | Sub-Saharan SIM connections [GSMA Intelligence (Intelligence, 2020)].

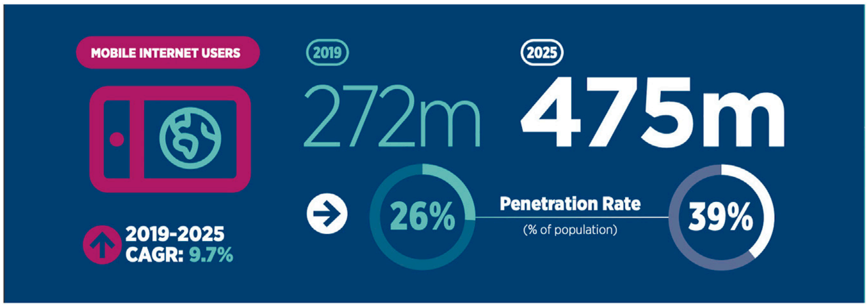


FIGURE 2 | Sub-Saharan mobile network users [GSMA Intelligence (10)].

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

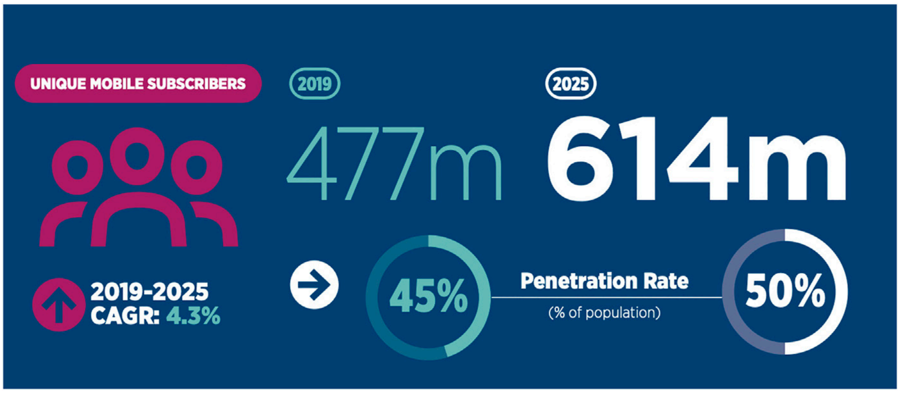


FIGURE 3 | Sub-Saharan mobile subscribers [GSMA Intelligence (Intelligence, 2020)].

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.

frontiers in Blockchain

ORIGINAL RESEARCH
published: 14 February 2022
doi: 10.3389/fbloc.2021.629790

Building a Credential Exchange Infrastructure for Digital Identity: A Sociohistorical Perspective and Policy Guidelines

Mawaki Chango *

DigiLexis Consulting, Lome, Togo

Credential Exchange Infrastructures based on open standards are emerging with work ongoing across many different jurisdictions, in several global standards bodies and industry associations, as well as at a national level. This article addresses the technology advances on this topic, particularly around identification mechanisms, through the Self-sovereign identity model. It also tackles necessary institutional processes and policy concerns relating to their implementation. Rooted in a sociohistorical culture and practice of inquiry, the goal of the article is to bring emerging digital identity systems within the grasp of a wider public as well as to contribute to mutual understanding across stakeholder groups (technical community, governments, international cooperation entities, civil society and academia) about what is at stake. This is expected to enhance their capacity to better navigate across the pitfalls of this transition period from paper to digital systems and the full adoption of the latter, with each of these stakeholders playing a part in enabling trust around digital identity infrastructure and transactions, both within related ecosystems and in the broader society. This article makes contributions around three axes. First axis is conceptual and analytical. The article outlines three conceptualized phases in the evolution of identity practices in history with the hypothesis that the availability of new record-creation methods invites changes in, and expansion of, the existing identification processes. This helps make a stronger case for why the Internet needs an identity capability. In addition, the article defines or elaborates on key concepts including identity, credential and trust. The second axis of the article is a case study on self-sovereign identity as instantiated by the Sovrin network. The case study presents the technology and its design with a view to enabling a non-technical public to understand what it is and how it works, while highlighting the fact that the technology still needs institutional processes to make it work as intended. The final axis of this article provides guidelines to policy actors potentially facing the need to enable large scale implementations of these emerging technologies, as they mature. Policy-makers approaching this material may want to read this section first and then return to the rest of the paper.

Keywords: digital identity, self-sovereign identity, identity systems, credential exchange, decentralized identifiers, verifiable credentials, governance frameworks, policy

OPEN ACCESS

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

Reviewed by:
Iain Barclay,
Cardiff University, United Kingdom
Maryline Laurent,
Télécom SudParis, France

****Correspondence:***
Mawaki Chango
digilexis.consulting@gmail.com

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

Received: *16 November 2020*
Accepted: *28 December 2021*
Published: *14 February 2022*

Citation:
Chango M (2022) Building a Credential Exchange Infrastructure for Digital Identity: A Sociohistorical Perspective and Policy Guidelines.
Front. Blockchain 4:629790.
doi: 10.3389/fbloc.2021.629790

INTRODUCTION

Most of the population in the industrialized countries and at least city dwellers over the rest of the world are familiar with situations where, for all intents and purposes, they have to present identifying documents before they can proceed with the business at hand. From a basic standpoint, that is an identification process which is enabled by some administrative artifacts we generically call identity documents. Without those documents, individuals will, in the best-case scenario, have to spend a lot more time resolving their identity for the person or institution they are faced with, or they just might not get anything done as they intended. For that reason and for several other benefits, we go through the process of getting those documents and we carry them around with us so that we can use them as needed. For the same reasons, some other people may find incentives to forge those documents where they cannot or do not want to get proper ones. Therefore, authenticating those documents themselves, as well as authenticating the link between them and their holder (checking the accuracy of their identity function), has been a critical need and endeavor throughout their multi-century history.

For the most part of that history, those administrative artifacts have been made in paper or in paper-like material. Over time and given the above-mentioned risks, various techniques and technology have been used to make them more reliable and tamper-proof as much as possible, improving their identification capability overall. In recent times, that challenge has been taken up by digital technology using biometric data to bind the body of the identity holder (subject) to those artifacts. However, we cannot address digital identity exclusively just as the latest form of identity on-land. The "land of origin" for the digital itself is the Internet, not just from its native protocol stack but also as popularized by the Web and today's mobile apps. Solving the identity problem on the Internet is of critical value once we realize that the digital economy is here to stay and that online identification loopholes and malpractice are a major hindrance.

The purpose of this paper is to introduce to stakeholders other than the small group of technologists involved in building solutions to address this issue—meaning governments, international cooperation entities, civil society and academia—but particularly to policy-makers, one of the fundamental ways in which the Internet identity problem is being solved today using a framework known as Self-Sovereign Identity (SSI). The Sovrin Network, an implementation of that framework using blockchain technology, will serve as a case. Following that exposé focusing on the technology, institutional aspects of this implementation will be teased out by examining its governance mechanisms. And finally, a number of recommendations are formulated for policy-makers, particularly in the countries less familiar with, or less engaged in, these fast-paced developing technologies, at this point in time. Before we get to the empirical part however, the paper sets the stage with a theorizing view of a historical account of the evolution of identity practices, following an overview of the epistemological and methodological context in which this approach is rooted.

METHODS

From a methodological[1] standpoint, there are two prongs to this research article. First, it develops a conceptualized narration of the evolution of identity (the way people have come to handle the process of identification over time) and the available enabling tools. We make the case that digital technology, particularly the Internet, is still in search of its own version of identity which it will inevitably find—or humanity will not fully enter the digital era.

The second prong in our methodology is the use of a case study to illustrate what is shown to be predictable from the first prong: that great strides are being made, by necessity, towards achieving a viable solution for digital identity. The case selected is one of the most current, indeed still emergent, technologies for digital identity to show how the problem of identity on the Internet can be solved and how close we might be to solving it.

The value of the method used in this paper is grounded in sociohistorical practices of inquiry (Somers 1994; Somers 1998; Hall 1999; Tilly 2006; Tilly 2008). First of all, according to Tilly (2008), "transactions, interactions, social ties, and conversations constitute the central stuff of social life." That postulate characterizes the epistemological stance he calls "relational realism." Reinforcing the same idea, Somers (1998) notes that the basic units of social analysis are "neither individual entities (agent, actor, firm) nor structural wholes (society, order, social structure) but the relational processes of interaction between and among identities." Furthermore, the notion of relation in this framework also has theoretical implications. On the one hand, society is a bounded set of "numerous matrices of patterned relationships, social practices, and institutions mediated not by abstractions but by linkages of political power, social practices and public narratives" (Somers 1998). On the other hand, theory in this context is mostly a generalization about observable facts treated as effects of unobservable factors which are only inferred, to the extent that they appear to be compellingly necessary to explain certain outcomes and, in that regard, relational realism is also and particularly based in that link posited between observable facts and non-observable ideas with an explaining power about those facts.

Within this framework, theory does not depend only on the capacity of the rational mind applying universal and a-temporal rules of formal logic, which would imply that a theory remains eternally true as long as those rules obtain; rather, relational realism acknowledges by anticipation that theory is "historically provisional" (Somers 1998); it is time-bound and subject to change. Concurring with that, George and Bennett (2005) further emphasize a distinctive trait of theory in social sciences, pointing to the fact that theory is not exclusively devoted to enabling prediction but also to explaining social phenomena or patterns. While doing the latter, cumulative and progressive advances into theorizing may be

[1] This section is of interest mainly for academics. For the policy or technology reader not interested in the ins-and-outs of social science methodology, please skip ahead to the next section.

TABLE 1 | The role of formative discourses in inquiry practices of configurational history and analytic generalization, with their common ordering discourse and its roles in italic bold [based off Hall (1999), Tables 7.1 and 8.1].

Formative discourses	A particularizing practice of inquiry	A generalizing practice of inquiry
	Configurational history	Analytic generalization
Values	Grounded in social theoretical configuration	Knowledge by "bounded generalizations"
Narrative	Focus of "break point" analysis	Basis for analytic comparison
Social theory	***Extrinsic analysis of development***	***Tests hypotheses by comparison***
Explanation/Interpretation	Identifies "accidents"	Controls or accounts for extraneous variation

accomplished, notably through the strengthening and the wider applicability (to multiple settings or to different phenomena) of analytical frameworks that may have proven more heuristic than others.

To complete this methodological overview, we turn to Hall (1999) who frames sociohistorical research as a practice whereby different research communities make claims to knowledge using different types of discourse which they develop in an effort to sustain their claims. Hall calls those types formative discourses in that, ultimately, they in turn help form different practices of inquiry. He distinguishes four such discourses, each one having its role across the different practices of inquiry: those include value discourse, narrative, social theory, as well as explanation and interpretation discourse.

On the other hand, Hall identifies mainly eight "alternative and yet interdependent methodological practices of inquiry" (p.169) split over two orientations, four particularizing practices being one orientation, and four generalizing practices the other. None of the practices of inquiry is discursively pure; rather, each one of them is "an ordered hybrid of discourses" (p.216) only with a predominant role of methodological significance for one particular discourse. In other words, each one of the four discourse types is formative for one particularizing and for one generalizing practices of inquiry, while playing a minor role for the other practices.

For instance, in what I call below a theoretical reduction, we are guided by social theory discourse. Social theory discourse is formative to the particularizing practice of configurational history by enabling extrinsic analysis of development, and to the generalizing practice of analytic generalization through the testing of hypotheses by comparison. The first of the parts of this paper addressing the subject matter (*A Theoretical Reduction: History and Concepts* section) falls under the latter: I am extrinsically analyzing historical periods, the delineation and the connection of which is only based on the focus of external observers (us) on a particular problem of interest (identity). That focus is not necessarily that of the actors contemporaneous to, or even involved in, the events and phenomena that are covered by this account. Such theoretical delineation or periodization of history around identity gives perspective, both retrospectively and prospectively, and allows one to see the scale of the challenge and explore what potential solutions may look like.

Later on, when different digital identity solutions are fully deployed and effective, this work may help us elaborate hypotheses to be tested with regard to which ones of the solutions might prevail and under which social and other non-technological conditions. In that possible future scenario, we will be inquiring for analytical generalization using social theory discourse (**Table 1**). For now, let us expound our proposed social theory-oriented configuration of the history of identity practices, starting with the underlying theoretical view.

A THEORETICAL REDUCTION: HISTORY AND CONCEPTS

By theoretical reduction I am abstracting and conceptually assembling a storyline or simply a narrative account from empirical phenomena. In this case, I am linking historical events or processes, which certainly are more variegated in their actual occurrence, so as to offer a picture of theoretical significance or to generate a theoretical statement. In the following, such theoretical reduction is applied to the way identity has been historically addressed from merely using humans' natural senses to using digital technology as a means of making and keeping records[2]. But let us start with the statement of our theory which provides the basis for this way of thinking about the evolution of identity practices through the lens of historical periodization.

Theory Formulation

Record-making techniques enable or augment human agency[3]. More precisely, new record-creation techniques bring about new forms of mediated human agency; new ways for humans to be present, to decide, act and change things at a distance. With a new widespread record-creation technique comes a significant extension of human agency, supported by a number of accompanying mechanisms.

By new, we do not just mean a technique that is chronologically more recent in existence, but a technique that allows to do significantly more than its last predecessor or to do really new things which its last predecessor couldn't do, in such an amount that it can be considered a life-changer for people in

[2] In this article, I am using the term "record-making" in the same sense that Geoffrey Yeo uses it in his 2021's book *Record-Making and Record-Keeping in Early Societies*. This is not only about record-keeping practices but first and foremost about the way records are created, the resource that enables them to be made records and, only subsequently, to be kept as such. I therefore speak of record-making or record-creation techniques interchangeably across the article.

[3] See next paragraph for clarifications.

need of using that type of techniques. Agency is defined in the online Sociology Dictionary as the "capacity of an individual to actively and independently choose and to affect change; free will or self-determination"[4]. It is the capacity of human individuals to exercise their free will, to reflect or deliberate, form an intention or choose a purpose of their own, make a decision and act on their own behalf. Our need to resort to the concept of agency, which is borrowed from institutional theories, particularly institutional sociology, is not commanded by a collective action problem where the free will of individuals is faced with collective structures, as usually the case. Rather, the main focus in our context is on the identity subjects who are individual entities (here humans) and to whom we are applying the concept of agency hence, the apparent emphasis on individuality in our definition. But given that theoretical background of the concept, let us clarify further some of the implications of using it in this context with our formulation of its definition.

We do not think "social phenomena result from the actions of atomized (socially unconnected) individuals" but rather, that "human agency is both constrained and enabled" (Emirbayer and Mische 1998; Abdelnour et al., 2017). While individuals with agency are free to the extent which they, and the society as a whole, can conceive of freedom, they exist and evolve within physical and social settings and as such, they are not completely foreign to pre-existing norms and commitments that prevail in those settings. As a consequence, acknowledging agency for individuals does not mean we think pure and absolute individualism is possible and that such individualism prevails over social structures. Most likely, social phenomena are an outcome of an open interaction between agency and structure[5].

Extending human agency through record-making techniques then means that the above-mentioned multi-faceted capacity by which we define agency for the actual physical individual can be fully projected and maintained through the type of records at hand, whether paper-written (using human language alphabet, numbers and humanly created symbols) or digital (using a wider array of characters and symbols, numbers, and various codes based on machine languages as well as encoding schemes, etc.).

One particular type of mechanism which does that is called a credential[6]. Credentials are not just any assertion of claims; rather they are meant to be trusted (to be accorded the status of truth) and, as a result, they need to meet a number of requirements that make them credible and reliable in the relevant context. In effect, credentials modify the boundaries of human agency only to the extent that others[7] trust what is being asserted through them. Such extension raises the need to address identity within the new scope of agency, using the very means of that record-making technique which enables it in the first place. It would be self-defeating to allow the claims about somebody, or something the credential was meant to warrant, to be misattributed to somebody else or be taken for something else.

Any given record-making technique fosters the development of corresponding practices and institutions. In other words, every new record-making technique enables new practices as well as new institutions or institutional processes. The new record-making technique must clearly provide an added value compared to the older techniques; it has to make business and life easier, in one way or another, while improving institutional processes and overall performance. This means at least one of the following: it can significantly extend the pre-existing scope of human agency; or it can significantly reduce the cost, or take much of the friction out, of exercising human agency under the pre-existing scope; or it can do both. The potential or actual value to be added by the new record-making technique, including the extended scope of agency it may enable, dictates the need and interest to embrace such technique as well as to address identity within that scope using the resources availed by that technique.

Questions that arise include:

- How can we make sure the new technique is reliable, trustworthy, in the various ways it extends human agency?
- How can we make sure it accurately represents personhood as well as the reliable attribution and discovery of the actual roles, rights, liabilities, privileges and authorities which any given individual instance of personhood may bear?
- How can we avoid falsehood in such representations?

In generalized terms, these are and will always be the identity challenges at every turn of significant change in the nature of records and the affordances of the means by which they are made, due to related technological change or evolution.

A Three-phase Evolution

Phase I, Face to Face

At the beginning, there are people living together. They go about doing whatever they need to do to live and survive, to keep going with their life and to thrive. That necessity generates all sorts of behaviors including interactions with others as well as transactions. Conceptualizing the evolution of identity, one may describe the first phase as follows. Mostly, individuals' behaviors and actions are performed and can only be performed when they are physically involved, either themselves or by another representing individual. And anybody who would witness such behaviors or actions can only rely on the capacity of their own senses and human memory to identify the person who was involved in those actions, interactions or behaviors as someone they have already seen, met or someone they knew. This is all the more feasible that the chances of having to deal with people popping in, out of nowhere, hailing from humanly unreachable distances are very low and, as a consequence, such rare occurrences are easily manageable for the human memory and, if necessary, by

[4] See The Open Education Sociology Dictionary at https://sociologydictionary.org/agency/.

[5] As a case in point, technology infrastructures and their design provide such a structure with some non-negotiable parameters within which the user has to evolve while using the infrastructure. Also, let us note that the infrastructure itself is designed by people sharing some fundamental values and commitments with the average or the enlightened user, which explains how the user can still exercise their free will within the confines of conceivable freedom in the larger social setting.

[6] Sovrin Glossary defines a credential as an "assertion containing a set of Claims made by an Entity about itself or another Entity." *See Clarifying Key Concepts* section for an amended definition and more discussion on this concept.

[7] Any parties other than those making the assertion or those about whom the assertion is made.

mobilizing the community's attention (collective memory). It may be noted that, already in this phase, identity is ascertained—authenticated, I might say, by one's own means and for one's own intents and purposes—through the ability to match incoming information (exhibited by the person appearing now before us) with an information record we are already familiar with which is generally stored in human memory[8].

In sum, to a great extent during this phase, human senses and memory are enough for people to be able to attribute to their fellow community members whatever they need to for practical purposes, in a consistent manner, over whatever period of time may be needed. Such empirical capacity to make attributions and to make them consistently is also what makes human beings able to attribute and recognize roles, rights and responsibilities (duties)[9] in relation with any given individual in their social environment.

Phase II, Paper

With the thirteenth century paper revolution—accelerated by diffusion of Gutenberg's printing press by the fifteenth century—documentation practices evolved to integrate paper and written records including documentation of identity.

For paper to have a meaningful impact on the things humans do as well as on how they do those things (their behaviors), on the state of anyone's roles, rights or responsibilities in the society, it will need to be used in ways that can be trusted enough by all key stakeholders, including anyone who might have claims that could interfere with existing roles, rights or responsibilities as well as on the community's common resources. This implies that those written records will have to be endowed with some authority—in relation to their ability to accurately reflect the outer world order. Such world order is shaped by, among other things, people's decisions and choices which re-order the distribution of rights and obligations. The way that reordering is done and the result has to be acceptable in the eyes of the key stakeholders (and beyond them, the community overall), and that is achieved by following certain protocols and using certain symbols and signs—which is facilitated by the sharing of the same beliefs. To trust this type of mechanism means that all key stakeholders accept it as a valid way to represent people who may then use such representations to enact decisions and choices, to assert or alter their roles, rights and responsibilities, and possibly those of others, provided that protocol and format requirements are met.

Historically, particularly in the West, those tools included seals, handwritten signatures, bureaucratic procedures plus, later on, agreements among nations-states and the continuous integration of evolving techniques, notably in more recent times, some degree of technology into paper-based record-making methods. All of that is done while keeping an eye on the need to prevent or mitigate the risks of tampering. The Church, the King and then the Government, or other accepted authorities (banks, schools, hospitals) backed or regulated by any of the first three, vouched for representations made through those systems. In their respective setting and at the best of their authority, those institutions along with their system of governance have served as the source of trust in this type of identity mechanism[10, 11].

As shown in Chango (2012), it took a long historical process to get from the time when, as a document, the passport started crystalizing in its core components and functions, in the 15th century, to a place where it became an internationally accepted and effective standard credential for all border-crossing travelers, in the 20th century. In effect, it is only after the First World War that the first international conference was ever convened, by the League of Nations, to agree on international guidelines for the passport; that was the "Conference of Passports, Customs Formalities and Through Tickets" held in Paris in October 1920. Other follow up conferences include the "Conference on safety and viability of international travel" held in Chicago in 1944 which gave birth to the International Civil Aviation Organization (ICAO)[12]. Ever since, the task of refining passport standards so as to tackle the challenges to its efficacy, has fallen to ICAO.

Basically, identity through paper-based written records is essentially made by describing observable attributes and known facts about the identity subjects. That information and relevant data are collected at enrollment and kept in paper files which are classified using some bureaucratic physical scheme, with a view to easing their manual retrieval at any point in the future if need be. A subset of that information and data, including particularly most observable attributes, is captured on a handy document which is given to the identity subject (making them a holder, indeed the only legitimate holder, of that document.) We have left the first phase, the Face-to-Face identification process where enrollment is random and authentication is done live, based on personal memories. Now the identifying entity is impersonal—it is an institution, e.g., the state, the state bureaucracy, the government—and so is their memory made of all the information and data they retain about the identity subject, in paper files in the back-end office. The memory is objectivized through a file system, a sort of paper database, and several potential individuals with the proper authorization may check out the content of that memory. Distinctive features in this model include photography and inked fingerprint, both of which are anthropometric data or data source but may arguably be considered as early biometrics. Authentication is

[8] In the history of western literature, there are numerous tales of this instance of the identity puzzle, from Ulysses to Martin Guerre, etc. See: Dimock (1956); Davis (1983); and Vernant and Ker (1999). On naming, see Wilson (1998). For further discussion on the meanings of personal identity outside administrative documentation, see: Perry (1975); Parfit (1984); and Noonan (1989).

[9] I am referring to those three things (which one may call the R3: roles, rights and responsibilities), knowing full well that there are plenty of roles and also responsibilities (duties) of various levels, some of which don't require that they be established *via* some form of material records, even still today.

[10] In today's terms, one would say that those institutions bootstrapped the said identity system by providing it with a trust framework.

[11] On the transition from memory to written records, see Clanchy (1993); on authority, trust or rights as well as various aspects of the mechanisms at play: Grant (1946); Kantorowicz (1951); Kantorowicz (1955); Fraenkel (1992); Burns (1988); Ekelund et al. (1996); Wolter (1997); Bedos-Rezak (2000); MacNeil (2000); Sassen (2006); Ekelund et al. (2011); on passports: Torpey (2000); Caplan and Torpey (2001); Lloyd (2003); Robertson (2010); on national ID card: Piazza (2004); and more discussion on the state or institutional mechanisms of control in Foucault (1988a) and Foucault (1988b).

[12] See the Conference documents at https://www.icao.int/ChicagoConference/Pages/proceed.aspx and also the international conference proceedings and official documents from the list of references at the end of this paper (Doc.LN, 1920; Doc.LN, 1922; Doc.UN, 1947; Doc.UN, 1956; Doc.UN, 1959; Doc.UN, 1961; Doc.UN, 1963; Doc.UN, 1966; Turack, 1968) Also see Stanton et al., 2007.

done by looking at the content of the document and observing the identity holder in order to check the observable information in the document,[13] including the photography, against its living source. In this phase, regular authentication still relies widely on human eye and visual observation capacity. At most, law enforcement would use a magnifying glass to scrutinize the ID photography details or to parse the inked fingerprint they have on file, trying to match them with the living face of an identity holder or with another specimen of fingerprint which they just collected from a suspect, for instance[14].

Phase III, Digital

Digital technology opens up two main paths for further progress. The first is the use of digital technology as an additional step to increase security and trustworthiness within the paper-written records paradigm[15]. I would call this a linear path, the path of incremental improvement (within the same paradigm).

The second one is a path of a paradigm shift or a qualitative leap; it introduces a completely new way of expressing and sharing identity information which would be commensurate with fully digital record-making settings. This path appears inevitable because, among other things, the Internet already allows people to conduct a sizable amount of their daily life operations online—while adding new capabilities to the previous two phases (the phase of physical presence-based agency and the phase of paper records mediated agency). Furthermore, many of those operations can be fully completed and validated without any physical presence or interactions during the process, neither for the person conducting those operations nor for the party on whose behalf they are conducted.

The question now is, can we conduct any of those operations requiring a proof of our identity without sending around, on the Internet, an electronic copy of our limited and monolithic physical credentials or some sensitive identity-related information? In other words, are we merely going to transpose analog methods to electronic environments, while applying them to electronic versions of physical stuff (thereby deemed digital), or are we going to shift to digitally doing digital stuff? Clearly, there is tremendous value to be gained, at scale, if we could do the latter and do it well—and that is the challenge many dedicated technologists have been working on for almost two decades[16].

Those two paths may be recognized as that of 1) digital identity in the form of a digitized physical credential, and 2) that of digital identity in the form of a fully digital (online) credential. It must be noted though, that under some circumstances, the first one may also help operate online. As a matter of fact, these need not necessarily be two different things. Digital identity may associate a physical token with online digital records and systems, both enabled by the same digital technology, making it possible to use or to refer to the same identity offline and online. Either way, it is the capability of online operations afforded to the identity holders themselves which brings about the full value of a new extension of human agency. In any case, the state of the technology today clearly allows us to think of digital identity as something of its own, based only on digital components, totally operable online in a digital environment. And that is our primary concern in this article: whatever happens outside the networks, how can that lead to digital identity solutions that work over the networks?

Clarifying Key Concepts

Identity and Credential

The community mobilized around the Sovrin Foundation has put together a Glossary which defines identity as "Information that enables a specific Entity to be distinguished from all others in a specific context. Identity may apply to any type of Entity, including Individuals, Organizations, and Things. Note that Legal Identity is only one form of Identity." Back in 2005, Kim Cameron in his Seven Laws of Identity[17] offered the following definition for digital identity: "a set of claims made by one digital subject about itself or another digital subject." This definition was then embraced by a cross-section of software industry players plus various other stakeholders[18]. From the same Sovrin glossary, a credential is "A digital assertion containing a set of Claims made by an Entity about itself or another Entity. Credentials are a subset of Identity Data. A Credential is based on a Credential Definition."

Before we get into discussing those concepts and some corollaries, I propose to consider the following reformulations or alternative definitions.

> Identity is basic information about any individual entity, in a given context, 1) which said individual entity can use to support the validity of a claim they might need to make relating to themselves, or 2) which a legitimate party needs to verify, and can do so, in order to make a necessary decision about said individual entity, in the context at hand.

As an informational resource, identity often is in a structured format (especially when it comes in the form of a credential: see

[13] Even the date of birth may be useful for authentication by observation, within some margins: for instance, if the date of birth indicates that the identity holder is 28, but the person presenting the document looks like a person in their 50's.

[14] Those are typical processes that characterize the paper era record-making and identification techniques. But as I implied before, one should expect that in the transition periods between two eras, arguably, there might remain territories, after a long period of time into the next era, where the tools and resources defining the two different eras will intersect.

[15] Most digital identity instances being promoted by the World Bank, particularly in developing countries, are of that type first of all, although the Bank uses a lot more the phrases "Identification Systems" or "ID Systems" (in the digital technology context) than it uses "Digital Identity," which it also does. See: World Bank (2018) and World Bank (2019), and their webpage https://id4d.worldbank.org/research.

[16] The following, among many others, discuss the digital transition in record-making, digital evidence and digital identity: Bolter (1991); Duranti et al. (2002); Solove (2004); Kerr et al. (2009); Rannenberg (2009); Blanchette (2012); World Bank (2018); Sovrin Foundation (2019); World Bank (2019); López (2020); Preukschat and Reed (2021) and Yeo (2021). Also, see the Internet Identity Workshop from 2005 to present: https://internetidentityworkshop.com/

[17] See https://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf.

[18] Gathering around the Internet Identity Workshop, they dubbed themselves "Identity Gang." The remainders of their lexical work can be found at http://wiki.idcommons.net/Lexicon. Commenting further, this group noted that any given entity or digital subject may have multiple digital identities and that a digital identity may be created on the fly for a one-time or short-lived purpose or it can be made persistent so as to be continually referenced back as the unique representation of the same digital subject in applicable contexts.

definition below) but it may also be any piece of information fulfilling either one of the two requirements in this definition. It is basic information in the sense that it generally is part of the primary information that is used, or is most relevant, to define the concerned entity in the context at hand. Here, the term "individual" doesn't necessarily refer to an individual human being, but to an individual instance of any entity. An individual entity in the physical world is a physically discrete thing which can be counted as one among its kind. A corporate entity may be one legal entity but instantiated through several different branches; each one of those branches may qualify as an individual entity (even if some assertions can also be made about the corporate legal entity as a whole; that is because identity is contextual[19]). In other, non-physical, environments, an individual entity is whatever is structured, whether through syntax or other means, to perform as a unit of its kind.

Starting from the Sovrin Glossary's definition of credential, we shall note that there is a difference, generally, between a credential and simple information as part of an assertion: a credential is a specific type of assertion which exhibits characteristics that make it trustworthy for most stakeholders who are ready to consider it as a proof for the assertion it is making.

> A credential is a document, an object or a data structure designed or intended to make any kind of assertion about an entity, according to a method that qualifies it as proof of what is being asserted. As a result, it may also serve as proof for any number of claims one may directly derive from such assertions[20].

In that sense, the two functions enumerated in the above reformulation of the definition of identity are concretely achieved using appropriate credentials. A corollary of the two definitions (of identity and credential) is that it is only by way of a credential that identity becomes a concrete, usable, portable and effective tool, in the form of some sort of artifact whether physical or digital, which is thus recognized by a variety of stakeholders as an identity credential. In the expression "identity credential," the notion of credential adds more of a dimension of proof to the simple notion of identity. A second corollary is that identity credentials are, a priori, a subset of credentials, a specific type of credentials while, arguably, credentials in general may be used as well for a variety of other things not intended for identification[21].

Any informational resource that can fulfill either of the two functions outlined in our definition of identity, or both, is enough to be referred to as identity in the practical context of identity management. The information needed to achieve those two basic functions may include all of the attributes on an identity credential, or just one of them, although in the latter case, the identity subject in the physical world will still have to show the whole credential. In the digital world however, the technology allows the credential holder to select and present only the one relevant attribute or even to derive a lower-definition claim from a pre-defined, higher-definition attribute, as opposed to presenting the original attribute itself in a transparent manner (e.g., "Age 21 or older" as a claim, instead of "Born on August 30, 2000" for instance, as an attribute).

Overall, at the very basic level, identity management processes need the following:

1) An individual entity who will be the one whom the identity information is about (also referred to as identity subject);
2) The registration of said individual entity by collecting and storing data about them so that the data can be discovered or retrieved later on, for verification and authentication purposes;
3) The subsequent issuance or attribution of some token, potentially with authenticating capabilities (which is known as a credential), so that it can serve as proof of registration as well as proof of a number of facts about the registered individual entity, including the ones collected at time of registration.

There might be other requirements depending on the technology being used. But in the absence of any of those three things, there can't be a reliable process of identification and thus, there is no identity management system[22].

[19] At the legal level for instance, dealing with the legal existence of organizations, such corporation registered as one will count as an individual entity regardless of the fact that, at the physical level, it has several branches which are not registered separately as legal individual entities of their own. In other words, the notion of individual entity depends on the relevant level of definition (or granularity) for the type of entity being dealt with, which depends on the level of agency concerned, keeping in mind that the entity types include human beings, organizations and things. Referring back to the concept of agency as defined in *Theory Formulation* section, the individual entity is where agency is located or manifested in the context at hand; it is, in a sense, a source unit or a subject unit of agency. Moreover, whereas agency is defined as something proper to human beings, it may be activated as per delegation in things that human beings build, be it organizations or other performing stuff such as a piece of software.

[20] That entity about whom the assertion is being made is inevitably an identity subject in that the credential has to clearly spell out its identity in a reliable manner, since the proofing is not of an abstract statement but of something being said about someone or something else (the entity). The truth that is being alleged in the assertion lies in that link and clearly, it can't hold if the issuer and the subject are not properly and reliably identified. However, we don't insist here on the individual dimension of the entity (as we did in the definition of identity), as the purpose of all credentials is not to identify a specific unit of an entity. Moreover, entities of any type and any dimension, including collective and geographically distributed ones, may hold credentials.

[21] Although it may also be argued that all credentials are identity credentials, at least when they apply to an individual entity, in light of two things. First, to the extent identity is made of valid claims about an individual entity, and the entity an assertion (a potential claim) is being made about through a credential is clearly referenced in said credential as it should, inevitably such proofing also applies to that entity's identity as referenced in the credential. Second, with the meaning that the concept of identity has taken in the digital context, it applies not just to people and animals but also to organizations and all sorts of things, both movable and motionless, including a piece of software as well as a piece of land, etc. Therefore, any credentials about all those types of entity may qualify as identity credentials, again, particularly when they apply to individual entities as opposed to a group of entities.

[22] The first element assumes a population of individual entities, and that's why identity is handled *via* a management system.

About Identity and Uniqueness

The two definitions of identity given above, from Sovrin Glossary and from Kim Cameron's Laws of Identity, may appear to show some tension in the way they are formulated: the first definition makes of the property of uniqueness (capability to distinguish a specific entity from all others) its central point, while the other doesn't mention it but rather focuses on claims. Why is that? And if uniqueness is actually involved, where do we locate and how do we apprehend it?

A long historical track of mathematical elaborations as well as philosophical debates around identity have probably prepared the ground for the compelling notion and inclination to think that oneness and permanence are constitutive dimensions to the concept of identity (Perry 1975; Parfit 1984; Noonan 1989). Moreover, the notion of authoritative identity credentials as a monopoly of the government has also instilled over time some sense of requirement for identity to be unique in order to be true. For multiple generations, the only identity which nearly all stakeholders regard as authoritative, that is, as the "real and true" identity, is the one that the government vouches for through a national credential[23]. Even any other identity credential, most of the time, relies on a government-issued credential, that is, the government-defined identity. In the resulting mental model, people would have a hard time with the idea that one individual can have multiple, alternative national or legal, or simply valid, identities within the same nation-state.

We know from experience that identity verification encounters generally are trivial and do not involve proofing of uniqueness at any level. Even authentication is more about accuracy than it is about establishing proof of uniqueness of anything, as that process normally deals with one identity subject and one credential at a time. But the whole process works because uniqueness is implied, and enabled at some point the whole identity value chain. How does that work? Identity verifiers are mostly concerned with checking for the following:

1) The identity holder is actually the subject to whom the credential was intentionally issued.

   As a consequence, only the intended subject of any given credential must be able to control it, under normal circumstances[24].

2) The source of the credential, its issuer, is clearly and reliably identifiable from the credential.

   This helps assess the value (particularly in terms of pertinence to the context and potential for truthfulness) to ascribe to the assertions or attributes contained in the credential and, subsequently, the veracity or level of confidence to accord to the claims enabled by those assertions or attributes.

3) There are no other conditions, either originally included or having occurred since the credential was issued, which invalidate it at the time, or for the context, of use.

   There are no restrictions added to the assertions which may exclude a specific use case or may not apply to the context at hand; at the time of use, the credential is not materially distorted or deteriorated, possibly transforming it into something the original issuer wouldn't endorse, or it hasn't expired or hasn't been revoked,[25] etc.

Those requirements are general, also applicable to non-digital, physical credentials. However, if we were to spell out the same set of requirements applying them specifically to the digital realm, the third requirement is better split into two. That is because typically, physical credentials not involving any digital technology are tampered with only physically or materially, so that the result can generally be spotted by expert human beings (through naked eye or possibly with the help of a simple piece of equipment such as some special electric light.) In any case, the possibilities for tampering with digital credentials are potentially endless compared to physical credentials. For that reason, an exclusively digital context would have requirement "***3***" above split into a new requirement "***3***" which will simply read: "The credential is not restricted or has not been revoked," and the following requirement:

4) The credential has not been tampered with.

   The credential has not been altered or compromised by third-party's malicious manipulations either to misappropriate it or to make other false assertions.

Focusing here on the digital context, requirement "2" above addresses the provenance of the credential, while the remaining three requirements address the fidelity of the credential[26]. The provenance requirement is mixed in that, while it may use cryptographic functions relating to the issuer's identifier, the

[23] In programs promoting governments' approach to digital identity, identity is considered as unique and ideally unvarying just as pre-digital identity credentials were. And in the supporting literature for the World Bank digital identity programs—such as the West Africa Unique Identification for Regional Integration and Inclusion (WURI) Program—it is qualified as legal or "foundational identity" (World Bank, 2018; World Bank, 2019).

[24] It shall not be possible to replicate a credential so that a non-intended holder can effectively use it just like the originally intended holder.

[25] Example of revocation for a physical credential: The passport regulations in some countries don't allow you to add extra pages to the passport but only to fully renew it if, for some reason, it can no longer serve as originally issued. In a given such country, when a holder's passport has run out of empty pages while still within the validity period, and yet there still is a multi-year valid entry visa for a foreign country in said passport, a new passport is delivered and the old one returned to the holder with all the pages punched except the one with the valid visa. Such an older passport is basically revoked, although the original validity dates are still current: the holder has to show the new passport along with the visa in the old passport for that visa to be considered valid, and the customs' stamp for entry in that country, as well as any other stamps for future traveling, will be affixed into the new passport as long as it is valid and usable.

[26] See Phil Windley's "Technometria" (blog) at https://www.windley.com/archives/2019/10/fidelity_provenance_and_trust.shtml.

full assessment of the authority and credibility of the issuer to make the assertions conveyed by the credential is enabled through governance processes including some knowledge of the outer world environment. The fidelity requirements ensure that the credential is true to itself, that it fully works and is appropriately used, as designed. The fidelity requirements are fully enabled by cryptography and they represent the "What you see is what you get" part in the credential exchange—here meaning, what you see through cryptography.

Practically meeting requirement "1" from the list above is what brings up the uniqueness dimension in the conceptualization of identity as a subject of management. That is not to say that an entity can only have one identity. The same attribute may be claimed by millions of people but if proof of that attribute is needed for any particular individual, it will have to be part of a credential which said individual can show for a simple verification and which, at best, can be authenticated. The farther we are from authenticating an assertion or an attribute about a particular entity, the lesser we can be certain that said assertion or attribute is true about that particular entity. Then we cannot build any commitment on that identity (which links a particular subject to an assertion about them), since some other entity foreign to it can wrongfully claim it and (mis-)use it. Clearly, the reason for this requirement is that typically, an identity credential can only realize its true purpose and full utility when it is tied to only one subject at a time as a single source of agency[27]. Binding any given identity to a unique individual entity as the intended subject of that identity is the condition that makes it possible later on to verify whether the presenter of said identity is its rightful holder or not. And the best way to attest to that binding is to include authentication as part of the verification process (which is not always done with physical credentials).

Depending on the context in the physical world, the verification of requirement "1" is done in various ways with varying levels of certainty or assurance about the result. Historically, at the beginning of the rise of identity documents, law enforcement relied on the good will of identity holders to dutifully use only the documents that were intended to them by the authority, not someone else's. At that phase, the proof was the weakest, assuming that can even be called a proof. After that, the credential-subject binding method was based on what I call anthropometrics,[28] along with other evidentiary features of uniqueness, including photography and ink fingerprint, affixed to identity credentials in order to bind the actual identity subject to the document. Then, most of the time, verifiers would just look back and forth at the picture on the credential and at the face of the identity holder; they also have the possibility to use the date of birth in order to assess its plausibility as compared to the estimated age range that could be imputed to the physical subject. That is still a weak correlation method. Only when the subject is submitted to verification at a law enforcement office where fingerprint can be taken again and compared with what is on the identity holder's document or on file, only then a strong case can be made based on evidence supporting that requirement "1" is met.

With digital identity, there is the model of a physical credential enabled by digital technology but which to a large extent operates as a more sophisticated version of the previous type of credential, whether it is used as a stand-alone token or in interface with online systems, still in the physical presence of the subject holding the credential. Here, the enabling elements with regard to requirement "1" include biometrics (electronic fingerprint, iris scan, etc.) which is encoded, that is, translated into a machine language, affixed to the credential, and will be shown to a machine reading equipment connected to an electronic database during the verification or authentication process. At that point, relevant biometrics is captured anew from the identity subject and is matched with the biometrics the machine reads from the credential which is matched with the biometrics previously stored in the database, in order to authenticate the credential, the data contained in it as well as the binding of the subject to that credential. This brings us to the model of a totally digital identity online. With this model, the physical subject does not interface directly with the system but through computer networks, and therefore cryptographic keys—which are secret information that is supposed to be known or possessed only by the identity holder—are the key element that enables the proof of binding between the credential and the user presenting it, that is, the identity subject (as we will see in *The Trust Over IP Technology Stack* section, particularly at Layer 3).

In any case, establishing the uniqueness of an identity subject in relation with the credential being presented is, in a sense, done by proxy: the uniqueness of the correlated subject derives from, and is supported in proportion of, the strength of the evidence supporting the binding of the credential at hand with said subject. The stronger the evidence available to support that binding, the more certain we can be that the current holder[29] is the rightful identity subject and therefore she or he is unique in that position, since it is a feature of the system that (by design) a credential is bound only to one subject who is the unique legitimate holder.

To accomplish the identity functions (as per our definition), some information about the identity subject first needs to be

[27] As per our definition in *Theory Formulation* section . Also, see note 19.

[28] Which I define as the process of measuring the size and proportions as well as detecting and reporting distinctive and even unique physical traits of a person's body, all done manually or by mechanical tools, as a means of recording or confirming identity; the recorded collection of the data thus generated. In that sense and when it comes to identification processes, I take anthropometrics as a predecessor of biometrics with the difference that the tools have changed since, with respect to their capabilities and scope (they can penetrate and read the human body deeper) as well as to what is considered knowledge (it is no longer considered that any reliable finding can systematically be inferred from the size of the skull of a person or the width of their temple or the length of their nose, as European powers did when they wanted to record the identity of the adult population under their colonial rule in Africa: See the example of the Belgians in Rwanda).

[29] Note that the phrase "identity holder" is used here as synonymous with "identity subject." In that sense, when a guardian (see Sovrin Glossary), acting on behalf of a dependent, presents the identity of the latter, said dependent still remains the identity holder in that context. The phrase 'identity owner' has also been used in the same glossary, at least at some point.

recorded in some fashion, somewhere. That first recording is also known as registration. Registration is the key procedure in the whole identity value chain which provides the basis for uniqueness and for meeting requirement "1." In effect, one fundamental role for any registration scheme is to reflect a basic truth about the state of the world with regard to the existence of the things to be registered. One of the basic laws (or facts) that structure the world as perceptible and comprehensible by human beings is that all the things which humans can naturally and materially observe as such exist in their original form in one instance only. For instance, a human being is located in one physical body only. Therefore, if the relevant things to be registered exist or are observable in their original form as discrete single units, then each one has to be registered only once, and has to be instantiated only as one in any given register of those things. Those entities can only exist as one in the modeled world of their kind through registration, as they exist in the original physical world: otherwise stated, to every relevant individual entity corresponds only one registration entry, under the same registration scheme[30]. Registering any entity more than once in the same register or database, under the same rules, as if the different instances of registration represent distinct, unrelated entities, would defeat the purpose of representing the world as it is; it would be a flawed representation of the world where the things being registered belong, which will lead to a flawed system (fraught with risk of impersonations and other fake representations)[31].

As we can see, the relevant notion of uniqueness to keep in mind when defining the concept of identity is that of relational uniqueness: the uniqueness of the individual entity which any given identity is bound to as its rightful subject, and therefore the uniqueness of the relation between an identity credential and the individual entity it correlates with. In Cameron's definition,[32] the key word is "about." How do you know for certain it is about this one and not that one, from among many potential identity subjects? By making sure you build it in such a way that it cannot be, at the same level of clarity and certainty, linked with or bound to more than one entity.

We know that, in the physical world and when it comes to the identity of human beings particularly, verification and authentication processes involve, more often than not, the physical presence of the identity subject. On the other hand however, the digital realm is characterized by the absence of the physical subject as part of the same processes and basically at all points where credentials need to be presented or claims need to be made and supported. Consequently, the challenge for identity in the digital world is broader. We must accomplish things requiring identity through digital information alone, in a sea of other digital information. And under the right conditions, a lot of those things can be done in a manner that can be as effective as, if not more effective than, what the direct action of an actual human being would accomplish in the same situation on land, and at a lower cost. As a result, identity no longer concerns only natural, embodied entities with agency; it is also, in a way, the identity of information itself.

Identity: "Who you are" vs. "What you are"

The trouble with overemphasizing uniqueness in matching identity attributes to an identity subject is that it opens the door to the confusion that leads to conflating the two, having us thinking of identity as a monolithic and complete informational representation of the identity subject. Such misleading perception then shapes expressions that lead to unreasonable expectations and inadequate mental models. The expression "Identity credentials prove who you are" sums up that misleading notion. We are now going to examine that conception as well as its assumptions, implications and limitations from different angles[33].

"Who you are"—The question "Who are you?" is often used as a prompt to elicit a response that is considered to be the identity of the respondent. However, the pronoun "Who" in this context suggests an essentialist or at least a monolithic view of identity: a person is always the same as self, they are who they are, with all their facets at once, regardless of context. From that standpoint, the "who" identity is as unique as the identity subject. Based on this conception, I should strive for a single definition of who I am, of my identity, which will contain every significant aspect of my whole self, no matter how lengthy that definition may turn out. However, there is no single identity that can comprehensively represent the self, fully provide the outlines of the actual self, including meaningful dimensions of self-identity (since the above question is normally addressed to the identity subject). As a result, and contrary to the common belief whereby identity credentials prove who you are, a person's identity credential doesn't tell who they are, overall or in the absolute.

"What you are"—Instead of "Who you are," we contend that your identity is rather "What you are." The pronoun "What" here introduces a clear rift between the actual subject and their identity. Humans don't naturally see themselves as a "What," that is, as being intrinsically a collection of things, so it is clear that the "What" (identity) is of a different nature from the "You" (the identity subject.) For that reason, identity does not have to be as unique as the identity subject, and it isn't (**Table 2**). Moreover, contrary to the phrase "Who I am" which may suggest that I have a single and universal identity, the phrase "What I am" is more

[30] That is just a necessity logically deriving from empirical conditions, which is fundamental, but that has nothing to do with a preordained necessity to define identity as a representation of uniqueness.

[31] And the day humans can naturally apprehend things that may appear at the same time in two separate places while still being one and the same thing, then uniqueness may no longer have to be at the heart of their notion of identity.

[32] Which is: "a set of claims made by one digital subject about itself or another digital subject."

[33] The ideas developed in this section build on an insight I already shared in my dissertation (Chango, 2012: section 6.2.1). While this won't change the way people speak and write about identity, the value of clarifying this is analytical and will make the experts and the technologists more careful in using such paraphrases as "Who you are" to explain identity or even to build identity systems on assumptions deriving from that view. As a matter of fact, we have come to discover the following piece written (in October 2021) by one of the notable technologists in the field and making the same point, as we were wrapping up the writing of this section: "Token-based Identity" by Phil Windley at https://www.windley.com/archives/2021/10/token-based_identity.shtml.

TABLE 2 | Unique or not unique: Who you are vs. What you are.

Identity subject	**“Who you are”**	**Identity**
	Unique	
Unique	**“What you are”**	*Multiple*
	Multiple	

apt to suggest the need for a context. Because I cannot reduce my whole self to things, exclusively, I will have to think of the things that are more relevant to represent such self of mine with, here and now or in any given context.

“Who you are, 2.0”—Under some circumstances, it could make sense to ask the question “Who are you?” as the adequate prompt to elicit identity. Those circumstances are not generally invoked when people paraphrase identity as being or proving “who you are,” which is why I am labeling this iteration of the phrase as version number 2.0. The predicate that identity credentials prove who you are can only be accurate with the following caveat: here, the pronoun “Who” does not refer to the first-order instance of the identity subject (i.e., for example, the embodied human being for human subjects). In fact, the question “Who are you?” becomes the equivalent of “Which one are you from among this group of entities we already have some knowledge about?”[34] Put another way: “We know something about each one of this set of people or entities, and to the extent that you are one of them, tell us: which one is you? That is, who, based on the few things we know about you already”. (What we know about them, for us that is who they are). Only when the question addresses an entity that is supposed to be part of a collection of entities about whom some amount of identifying information is already known does the who-question become not only relevant but adequate. Moreover, the only optimal scenario for this is that the question is being asked by the identity authority with whom the subject has been registered or maintains an account, or by its agents or any other entity authorized to interoperate with the concerned registration system, either in order to gain a full view of the information that defines who the subject is within the concerned system or in order to verify just a piece of information needed to make a decision about the subject. Normally, only transactions that would need to be recorded for whatever reason or would require an update with the subject’s account or file should call for the who-question.

The mental model stemming from the who-question might also, to some extent, be explained by the following fact. Historically, identity verifiers have been first and foremost agents of the issuing authority (law enforcement officers, civil servants and other public administration agents, etc). For those verifiers, an identity holder is only a collection of information they keep on the actual entity holding that identity, that is, as an information record, a file, or a database entry, along with the respective contents of those artifacts, including relevant historical data such as past changes, plus whatever else is required by design, based on the purpose of the identity system at hand. To the extent that there is a tool or a mechanism (part of which is a token put in the hands of every registered individual) which enables the issuer and subsequent authorized verifiers to find and retrieve the proper record pertaining to every individual whose claims they might need to assess for veracity in order to make a decision, then such a tool or mechanism qualifies as having an identity capability, as it enables them to find “who you are” in the system or in the mass of several records—in the sense of which record is yours, which one represents you in there.

In other words, under conditions where there is no prior contact with potential identity subjects[35] and where the purpose is to offer a general explanation of the notion of identity, the who-question does not work adequately. However, from the empirical standpoint of identity management,[36] it works in contexts where identity subjects have first been registered; the “Who” is adequate for a system of accounts, or registered individuals bound to existing accounts by authenticating procedures. Short of that pre-requisite, and keeping it simple while still striving for accuracy, the right question translating identity from the general, theoretical standpoint of identity management is

[34] First, saying “which one” (as opposed to “what”) is possible because we already have some knowledge about the concerned entities; second, “which one” is specific enough to translate as “who.”

[35] This also includes contexts where the subject is known in some existing domain, but the entity asking the question “Who are you?” has no relationship with that domain.

[36] Empirical standpoint of identity management refers to a context where an actual identity management system is being built and the question is to be confronted from that standpoint, whereas the theoretical standpoint of identity management refers to a context where one is just thinking about, analyzing and explaining identity management systems, the way they work or are supposed to work, and related concepts.

"What are you?" rather than "Who are you?" Overall, "Who you are" is either the actual you, meaning the physical life identity subject, or the registered version of you, that physical life identity subject, in a given system. Either way, "Who you are" is unique. Whereas "What you are" is multiple, depending on the context, just as is identity (**Table 2**).

A number of corollaries can be drawn from this. First, identity properly understood as "Who you are" (version 2.0) is built out of "What you are" which implies "potentially all what you may be"; one is made of a subset of the elements of the other. In other words, the "Who" is a function of the "What" in such a way that the scope of the "Who" is directly proportional to the wider scope of the "What" it is made of.

Second, the what-question has the advantage of dual relevance: it can be used in contexts with prior registration or existing accounts as well as (and even more appropriately so) in contexts where there are none. On the one hand, in contexts where knowledge of the subject is available prior to the current encounter, the things that will be sought after with the what-question, the things to be discovered in the instance of the subject at hand, would consist of the specific values taken in that instance by the parameters of the scheme used to form that prior knowledge about the subject. On the other hand, the what-question indicates that we already acknowledge that we are in the realm of representations, although that is the only thing we know. We know nothing, specifically, about the model of representation applicable in the context at hand—either because the subjects have never registered with us or we are completely foreign to any accounts they may have anywhere else. And in that case, the context will dictate what is relevant to defining the identity subject candidate, the parameters needed for the model of representation relevant to the entity seeking to know (and that is what happens at any registration with a new system).

And lastly, an important corollary of that distinction between mental models pertaining to "Who" and to "What" is that many identity transactions may be conducted without prior registration or setting up accounts for the identity subject. Tokens (including identity credentials used as such) are enough to handle some transactions with an individual, as those transactions don't require reading or capturing every piece of identity information available nor do they need to be recorded but just to be carried out to conclusion at once. In the physical world for example, there are many situations where we conduct identity transactions only based on "what we are" of pertinence in the situation at hand, without the need to open an account. As a customer in a place of public accommodation, the staff might need to identify me at some point in the process, not at the level of who I am but simply at the level of what I am. For instance, I have booked an alley at a bowling facility: "customer for lane X just requested an extension for their game time and we have received the extra payment and confirmed the extension." Or when I want to get some liquor at the store while in the United States, I show my identity credential to the cashier just so they can check my age status—that I am at least 21 and, as a consequence, they are authorized by regulations to sell me liquor (while I show an ID, here it only plays the role of a token for the proof that I am at least 21, nothing more.) No account needs to be set up or maintained in either case, because those entities do not need to care about who I am; we simply need to exchange necessary information transactionally and the business is done. In digital environments, because everything is done through exchange of information, it is even more critical to recognize the importance of the what-model and to enable related scenarios to be handled as such, as opposed to treating every transaction or interaction that requires the slightest bit of identifying information as if it pertains to the who-model.

Trust

Just like identity, trust is a concept of notable interest to both philosophy (Baier 1986; Baier 1994) and management (Barney and Hansen 1994; Wicks et al., 1999). It is a recurring theme in discussions relating to identity management systems such as the Sovrin Network, particularly with regard to the governance mechanisms that surround the technical system, which are designed to nurture trust,[37] at least in part.

To begin, let us be clear about one thing. There is no sense to a human being trusting a thing like, say, a stone (and obviously, a stone can't trust anyone, or anything, for that matter). Trust cannot apply to something that is not capable of any behavior. And something that behaves, one way or another, is either endowed with at least its own volition, or is made by or of other beings endowed with at least their own volition,[38] who then enable or shape the behavior of that thing. Either way, trust may apply. In the end—and at least in the context of identity management systems—trust comes from human beings and applies to human beings, or to something human beings are involved in one way or the other. Let us consider these two orientations in turn.

The first relates to trust from the standpoint of interpersonal relationships. Human beings get trained to trust, or to reserve their trust, mainly through these relationships; that is the context where most people first experience trust, as a personal state of mind or sentiment. To trust a person, one has to make the determination or decide for oneself whether that person is worthy of trust, based on any available information deemed useful for that purpose, including their own or other people's previous experience with the person to be trusted. Here, trust fully is a human sentiment and a subjective experience.

Drawing from that experience, to trust a person is to be inclined to believe that they will behave as we expect. However, expecting a villain to behave badly, and then they do, does not quite imply that the villain is a trusted fellow, in the way people think of a person they trust. Trust does not just result from a recurring confirmation of what is expected of someone; it implies a positive valence in that it is supposed to result in positive

[37] See *Governance Frameworks* section.

[38] Without entering into philosophical debates as to whether other beings, such as animals, are endowed with own volition or whether that is the exclusive province of human beings, we will only focused on human beings in this context, as there wouldn't probably be any identity problem for human beings to solve if it were not for the scope of all what human beings are capable of doing (their behaviors). Furthermore, here our notion of volition points to free-will and agency, as it requires the capacity to choose a course of action from among several others one is aware of.

outcomes from the point of view of the trusting person. People we trust are not just consistent and somewhat predictable regarding the issues we trust them on, but their consistent and predictable behavior generally goes in the direction of what is sound, good or desirable from our point of view (which may not necessarily be what is good in the absolute or what is commonly good). We trust someone when we think they will consistently do what we believe is the right thing to do in a given set of circumstances, even though they might be aware of other options available for a different course of action. The implication is that if they were to have total control over something of significant interest or value to us and they know how that thing is supposed to be handled or how best to handle it (from the standpoint of that interest or value), we do not worry because we are confident that they will handle it properly. Therefore, in cases of interpersonal or direct relationships, we expect that, at least in normal circumstances, they will most likely behave in a way that aligns with our interest as long as they are aware of that interest.

At this point of the discussion, we might want to acknowledge that the notion of trust implies some amount of risk, as it transpires from Barney and Hansen (1994) definition of trust as “the mutual confidence that no party to an exchange will exploit another’s vulnerabilities.” There is an aspect of the prisoner dilemma here, facing the risk of having one’s vulnerabilities being exploited by the other party while we bet on the contrary by protecting their vulnerabilities. The same idea of risk attached to trust has been elaborated on by Nickel and Vaesen (2012).

Sovrin Network uses blockchain technology which is claimed to enable us to do away with having to rely on third parties in order to successfully conclude transactions over the network. In that sense, blockchain is reputed to enable systems that do not need to resort to trust at any point, and yet they work reliably well (Antonopoulos 2014; Werbach 2016). From the perspective of the cryptocurrency world where blockchain is foundational and algorithm reigns, trust is like a last-resort device. People would trust only because they have to—when there is no better solution available to them. That would be better if they could avoid trusting, for trusting still implies that we rely on someone else’s moral compass, consistency of character, sense of duty or sheer discretion to rise to the level of the trust we are placing in them and related expectations. And of course, there always is a risk they might not rise to the occasion.

And as part of the response to that claim about blockchain as a technology for trust-ridden systems, distinction has been made between trust and confidence whereby blockchain qualifies as “a confidence machine” (De Filippi et al., 2020) while it would be, in a way, trust-incompatible. From that angle, trust is based on a personal belief or on a value judgement and as such, it cannot be objectively assessed. There is nothing deterministic about trust, whereas confidence stems from some deterministic mechanism, the workings of which can be objectively controlled. For instance, algorithms and computation methods involved in a blockchain-based process will yield the same result every time they apply to the same inputs, all things being equal. Anyone with the adequate knowledge (which is publicly available) can check the process and verify that it has followed the appropriate methods and rules, or use available evidence to the contrary to challenge the result. Short of the latter, people can have confidence in the system and its outputs. In such a context, no one needs to trust anyone, as trusting any entity would imply that the latter has an exclusive, superior access or knowledge, which places such entity in a unique position to both attend to the system and address the concerns of the parties to the exchange as well as any issues that may arise between them.

The dimension of mutual care in Barney and Hansen (1994) definition above may imply that trust happens among equal parties (i.e., peers), or parties that can stand on equal footing. However, while trust is typically a personal sentiment, it turns out to be a human inclination that can be extended from trusting other humans to trusting human collectives, such as organizations and even institutions, as well as to trusting technical systems with less human involvement on a direct and continuous basis.

This leads us to the second orientation of trust which is applicable to “something human beings are involved in one way or the other.” Organizations, institutions and technical systems are designed by human beings to operate in a certain way. Moreover, the business and operations of those structures are also conducted with the participation of human beings who strive to cooperate with one another by following agreed-upon procedures, all of which involves their worldview or belief system shaping, in turn, their intentions and behaviors. As a result, those collective entities and systems can more or less be trusted, or not at all, depending on their features and the way the people in charge handle their business, etc.

In the context of technical systems, particularly identity management systems or infrastructures such as the Sovrin Network, we start from a place of power imbalance, as the end user is an individual facing the system. Under normal conditions of use, there is no balance, not even close, between the vulnerabilities of the individual user facing the system (including those who run it and the way it is ran) and the vulnerabilities of the system facing the user[39]. Differences include the fact that the system-side:

- Operates at an impersonal (institutional) level while the user operates at a personal level;
- Is a steward of user’s resources, some very personal ones at that, with no comparable reciprocal function;
- Has the capacity to adversely impact resources and interests of the user;
- Has more control, more leverage over the relationship;

As a result, one may conclude that there is a vulnerability asymmetry: no equivalence can be established between the two sides with regard to the extent of vulnerability they are exposed to in the relationship. Therefore, if trust is of any relevance here, that can only be asymmetrical trust (which would be a different concept altogether.) One party doesn’t particularly need to trust the other, only the other does—either because the former has no vulnerabilities or if they do, it doesn’t take being in a relationship with them to exploit those vulnerabilities (they are

[39] See also Solove (2003) about the “architecture of vulnerability.”

potentially available to anyone with some capacity to exploit, as it happens by computer hacking and virus attacks). In such cases, it is normal that the system-side uses other levers and takes additional steps to create and foster trust from the user-side in the relationship. Two sets of elements contribute to addressing that asymmetry:

1) Policy provisions and rules that take care of the interest of the user and which the system-side commits to abide by;[40]
2) The system-side is public-facing, meaning it is potentially accountable to the public, even if such accountability is voluntary or done under a regime of self-regulation.

Regarding the case presented in this paper, it might be useful to note that the Sovrin Governance Framework was initially called a "Trust Framework" which might indicate that the main virtue of having these governance frameworks is to foster trust, to bring the users to trusting or, if you will, to being confident in the system as well as to bring the stakeholders to trusting each other. How?—In other words, how is the point 1) above addressed in the context of the Sovrin infrastructure?

First, by developing the SSI principles and letting the user know the values that have guided the design of the system, are infused into it, and shape its operations. The goal is to help the user recognize that those principles and underpinning values (along with the features they lend to the system) lead to outcomes that align with the user's best interest. Second, by demonstrating through experience and over time that the system is working as designed and as expected, according to requirements that derive from its guiding principles and values.

All that is true, except that it is incomplete: the reader should read again and systematically replace the word "system" by "ecosystem," as it isn't just that the technical system needs to be designed (by people) following requirements and bringing about features that inspire trust. People, along with the institutions they enact, intervene on a continuous basis beyond design, from implementation to operations, including by using a host of non-technical mechanisms, in order for the system to achieve its goal and produce desired outcomes while meeting customers' and users' expectations. Beyond the technical system, that is what we mean by ecosystem.

Eventually, De Filippi et al. (2020) reach the same conclusion that trust needs to be brought back in blockchain-based systems, as they always involve human components. In any case, whether trust is needed or not is not a "either . . . or" question; we might just need to identify and distinguish the elements that lend themselves to confidence and the elements that might use trust[41].

This concludes our review of the key concepts of identity, credential and trust. In the next section, we will expound on self-sovereign identity as an architectural level view of which the Sovrin Network will later be studied as a case.

SELF-SOVEREIGN IDENTITY ARCHITECTURE

Self-sovereign identity is not a particular digital identity technology. Rather, it is a vision that is captured, at best, through principles which ultimately outline a model for the technology to instantiate[42]. The term "sovereignty" here (which, in fact, should never be separated from "self") does not interfere with the sovereignty of nation-states or any similar authority, in any way. The phrase expresses the need to fill a gap, which is: regardless of any external authority and whatever administrative identity they may claim to define for individuals they can control, every human being should enjoy the right to hold an identity, including the capability to make one for themselves if need be (especially if no other option is available to them). SSI doesn't provide a particular solution as to what technology or system to use; it simply ensures that identity capability is available to anyone who needs or wants to use it, without trade-offs on their agency or autonomy, particularly in the digital realm. The result is a set of tools that enable identity management to be truly decentralized in order to empower the autonomy of the identity subject. They empower every identity owner to have control over their identifiers and their identity data, and to be able to securely share any of that with legitimate verifiers or any other party they may decide to transact with. This structurally puts them on par (making them peers) with the other party in any identity-related transaction, whoever that is, and in subsequent decision-making processes regarding the use of their identity data[43]. This decentralization comes with a conceptual dislocation—and a rebalancing—of authority as an exclusive source of truth and decision from the identity system toward the identity subject.

While it is true that "user-centric identity" design has already shifted the focus on the user in the recent past, it still did not give the user much autonomy and agency within the data exchange mechanism, particularly because the portability of the credentials was relatively limited as it required some level of pre-arrangement (e.g., federation) between issuers and potential relying parties, which the user has no control over. SSI further shifts toward full portability and, subsequently, toward user autonomy in their identity transactions. Through the latter, users can build relationships around their identity as it suits them. With SSI,

[40] Although there are also rules for the users, through the conditions of use, policies and other related tools are the place where the system-side engages on commitments that they volunteer to be held accountable against, with a view to enabling a trusting relationship from the user.

[41] "The design of the identity metasystem clearly delineates the parts of the system that are low trust and those where human processes are still necessary" (Windley 2021).

[42] For more about the genesis of SSI, see "The Path to Self-Sovereign Identity" by Christopher Allen, posted on April 25, 2016 at http://www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html (accessed on December 27, 2021). For more recent developments on SSI in general, see Preukschat and Reed (2021) and López (2020).

[43] In other words, identity subjects or owners now have the capability to fully be counterparts in identity transactions, alongside issuers, verifiers or relying parties, etc.

users hold and manage their identity credentials using digital wallets, vaults or any other secure data store, and use them to prove a variety of claims to legitimate verifiers (legitimate in the eye of the claim-maker, that is, the credential holder), whenever they deem the circumstances warrant it. In this context there is no authority that is the sole source of validity for the user's digital identity. Rather, validity stems from an interplay between the credential holder, the issuer of the credential and the cryptographic infrastructure which contributes to enable trust.

The infrastructure design that has been developed to achieve this is referred to as Trust over Internet Protocol (ToIP). In the remainder of this section, we will examine the technical aspects of the infrastructure as well as the human and social processes which are intended to enable trust over this infrastructure. The ToIP stack includes four layers (Res.ToIP, 2020; Res.GitH.0289 2019) along two dimensions which I am calling the Technology Lane and the Governance Lane. The entire architecture is structured along the four layers in the Technology Lane (the technology stack). Let us first examine the layers in the technology lane before turning to the governance lane.

The Trust Over IP Technology Stack

The technology lane of the Trust over IP architecture assembles four layers starting with the utilities layer at the bottom, each one enabling the next one at the top (**Figure 1**). They are all described as follows.

Layer 1: Public Utilities

In the ToIP framework, "utility" is the name given to the system used to anchor a cryptographic root of trust. That system can be any type of distributed database or file system, or any other system which can fill that function (such as a distributed hash table (DHT), a blockchain or distributed ledger, etc.). The technical generic name for those utilities is "verifiable data registry" systems. The W3C defines verifiable data registry as follows[44]:

> A role a system might perform by mediating the creation and verification of identifiers, keys, and other relevant data, such as verifiable credential schemas, revocation registries, issuer public keys, and so on, which might be required to use verifiable credentials. Some configurations might require correlatable identifiers for subjects. Example verifiable data registries include trusted databases, decentralized databases, government ID databases, and distributed ledgers. Often there is more than one type of verifiable data registry utilized in an ecosystem.

In addition, this layer includes the methods for generating and verifying decentralized identifiers (DIDs). As a W3C standard, DIDs, are a new type of globally unique identifier which is adapted to the distributed systems at the foundation of the ToIP stack. DIDs hold four core properties: they are permanent (once assigned to an entity, the DID is a persistent identifier for that entity and cannot be reassigned); resolvable (it resolves to a DID document which is a data structure describing the public keys and service endpoints necessary to engage in secure interactions with the DID subject); cryptographically verifiable (the content of the DID document enables a DID subject to prove cryptographic control over a DID); and decentralized (being cryptographically generated and verified, a DID does not require a centralized registration authority like other resource identifiers such as phone numbers, IP addresses, or domain names)[45].

Layer 2: DIDComm Peer-to-Peer Protocol

DIDComm is a protocol providing a collection of secure messaging standards. These standards cryptographically enable secure communication between two software agents[46] either directly edge-to-edge or *via* intermediate cloud agents, which is why DIDComm protocol is also referred to as agent-to-agent protocol. Sovrin identity owners, for instance, must have an agent in the cloud and one on any personal device they use for their Sovrin identity transactions. Agents are the basis for peer-to-peer relationships in the infrastructure. Credentials are not stored in the registries at Layer 1. Rather, software agents are used to provide identity owners with a place (such as a digital wallet) to hold and manage their credentials and private keys, either directly by themselves or in a delegated fashion (e.g., in the case of guardianship). Agents communicate with other agents directly for DID and credential sharing, using signed and encrypted messaging.

Layer 3: Data Exchange Protocols

Layer 3 determines how the issuer's agent issues credentials to the credential holder, how the credential verifier requests information from the credential holder, and how the credential holder presents a proof of information from their credentials that the verifier can trust. However, before all of this happens, the issuer must register a credential definition and a public DID to the data registry so that a verifier can look up the definition and collect the cryptographic bits that will enable the verifier to ascertain the fidelity[47] and the provenance of the credential[48]. The issuer may also add revocation registries and schema definitions to the utilities in

[44] World Wide Web Consortium (W3C), "Verifiable Credentials Data Model v1.0." W3C Recommendation, 19 November 2019, https://www.w3.org/TR/vc-data-model/.

[45] GitHub, 0289: The Trust Over IP Stack: https://github.com/hyperledger/aries-rfcs/tree/master/concepts/0289-toip-stack.

[46] It should be noted that the use of the term "agent" when discussing the technology, is unrelated to the theoretical concept of agency we elaborated on in *Theory Formulation* section. According to the Sovrin Glossary, an agent is "a software program or process used by or acting on behalf of an Entity to interact with other Agents or with the Sovrin Ledger or other distributed ledgers. Agents are of two types: Edge Agents run at the edge of the network on a local device; Cloud Agents run remotely on a server or cloud hosting service. Agents require access to a Wallet in order to perform cryptographic operations on behalf of the Entity they represent." And agency here is not more than "A service provider that hosts Cloud Agents and may provision Edge Agents on behalf of Entities."

[47] Note that sometimes we also use the term "integrity" as synonymous with "fidelity," as both point to the notion that the credential being presented is exactly as issued, without any alteration in the parameters that define the conditions of its validity.

[48] Only for knowing who the issuer is. Trusting the provenance (the issuer) is another matter which is dealt with through governance provisions complemented by "real world" experience relatively to the ecosystem of the transaction.

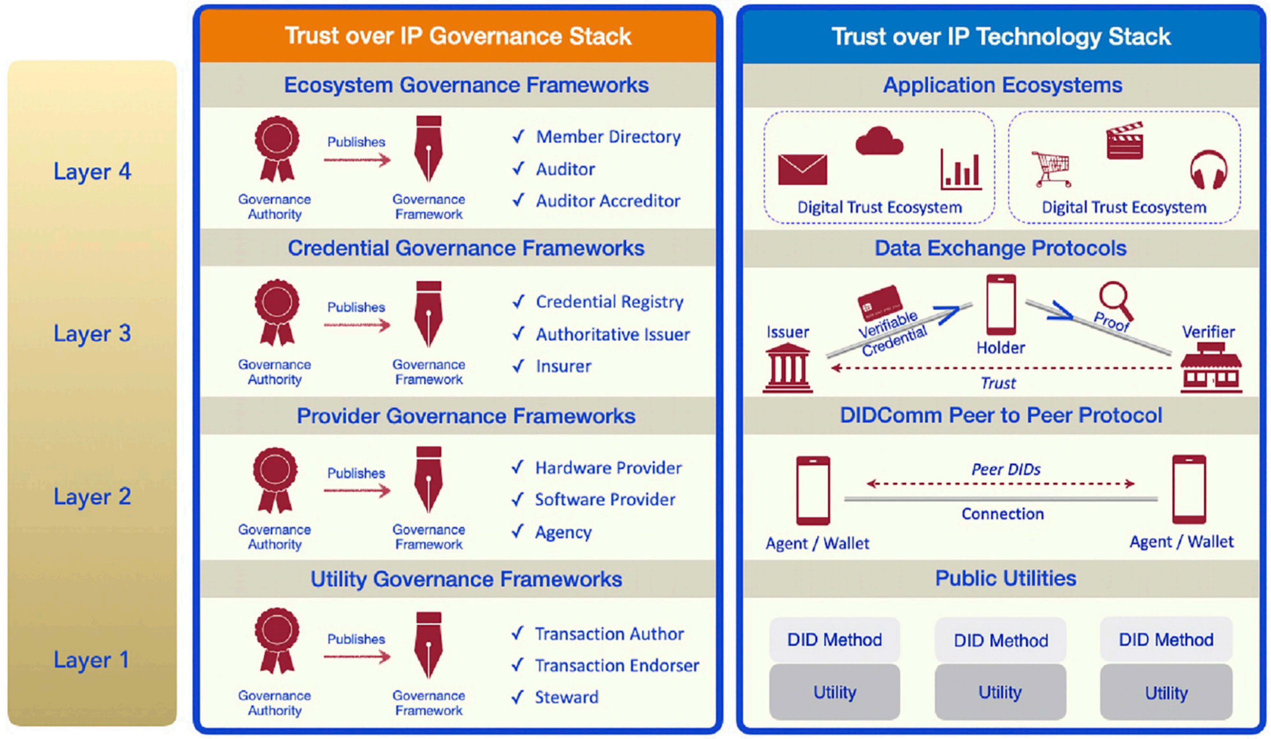


FIGURE 1 | Architecture of the Trust over IP (ToIP) stack. (Source: Trust over IP Foundation https://trustoverip.org/).

Layer 1 which are used in the credential exchange. Layer 3 is where humans use the system and create the trusted interactions that are only technically enabled in the first two layers.

Layer 4: Application Ecosystems

This is the layer where ecosystems of trust may form around applications, involving their owners or operators, their user base and their data apparatus[49]. These ecosystems are fostered by appropriate work processes, policies and governance mechanisms. Humans interact with applications for purposes that concern their business, their personal daily life, or other roles they may play in the broader society. With the appropriate infrastructure enabling the exchange of verifiable credentials, they might accomplish more with those applications, depending on the trust they actually experience as human beings.

The Trust Over IP Governance Stack

In **Figure 1**, the governance lane comprises layers that are perfectly aligned with the ones in the technology lane, each one in the former addressing the governance framework for the corresponding layer in the latter. Governance frameworks ensure that at every layer, the infrastructure orderly operates according to collectively agreed upon rules and procedures as well as applicable regulatory and legal provisions, the goal of which is to shape expectations, create regularity and maximize trust in the ecosystems. Some governance frameworks may simply serve to enact existing rules and relevant authorities in the context at hand, depending on the type of credentials to be supported and their purpose; some others may have to erect new authorities and rules. Whatever the case, governance arrangements and processes do not only serve to ensure that the rules of the ecosystem itself are set and upheld (for instance, preventing censorship and ensuring portability) but they are also critical in producing systems that can meet governmental and jurisdictional requirements including any applicable rules from higher-level authorities (for instance on data security and privacy protection).

The governance frameworks specify the purpose, principles, and policies that apply to all governance authorities and participants in that ecosystem. Based on its purpose, each layer has specific functions and standard roles that the governance frameworks must define while outlining a governance model suited to the constraints of the business model, legal model, and technical architecture of that layer. The governance frameworks also elaborate on the principles and values that need to guide the technical design and the human behavior which would be optimal to help achieve the purpose of the concerned layer.

At Layer 1, the governance frameworks will support the standard roles related to different types of utilities as well as interoperability and transitive trust,[50] including "transparent

[49] By that I mean all the pieces of equipment and infrastructure that the application designers, owners or operators have in place to collect, store and process data.

[50] A quality by which an authorized user (trusted) in a domain is automatically authorized (trusted) in a new domain originating from the first.

identification of the governance authority, the governance framework, and participant nodes or operators; transparent discovery of nodes and/or service endpoints; and transparent security, privacy, data protection, and other operational policies"[51].

At Layer 2, the primary governance focus will be on establishing "interoperability testing and certification requirements, including security, privacy, data protection, for the standard roles involved as per the governance framework."

Layer 3 is the first layer where the technically-enabled trust at lower layers starts transitioning to human-experienced trust. Consequently, credential governance frameworks become a critical component for interoperability and scalability of digital trust ecosystems. The frameworks can be used to specify credential schema definitions; requirements for authoritative credential issuers; the policies those issuers must follow to issue and revoke credentials; applicable business models, liability provisions, and insurance models.

Layer 4 is where humans will directly experience the ToIP Governance stack, manifested by provisions in the ecosystem governance frameworks that shape user experience through the applications available in related ecosystems.

Taken together, all these governance rules, mechanisms and tools critically complement the technological support tools (such as the cryptographic ones in this context) to make trust a reality. In other words, governance is indispensable to trust in the ecosystem—to the point that the two phrases "governance framework" and "trust framework" are often used synonymously.

As shown above, the SSI network infrastructure requires a number of features, both technical and institutional, designed to enable and maximize trust in the infrastructure so that it works as intended to provide a high level of confidence in the accuracy and effectiveness of the results, both in regard to what is intended and what is performed. Users must have such confidence in order to trust the system to deliver the value it is designed for without causing significant harm. This model architecture for identity can be implemented in various ways, with infrastructure components based on different technology solutions. In the next section, we will focus on a case that uses distributed ledger technology also known as blockchain for Layer 1.

SOVRIN NETWORK: A CASE OF BLOCKCHAIN-BASED SSI

Sovrin Network is one early instance of SSI that uses the distributed ledger technology (blockchain) in Layer 1. Like we saw in the general SSI model presented above, the Sovrin Network solution relies on technological components in addition to what we may broadly refer to as "social components." These are brought together into the ecosystem governance frameworks (including principles to guide stakeholders' behavior).

The Technological Components of the Sovrin Infrastructure

The technology components in Sovrin Network include both hardware and software, namely the devices used by each type of player to enable or use the systems running on the infrastructure, plus applications, standards, protocols and cryptographic keys. The Sovrin Network is built on three open-source projects developed by the Hyperledger community[52]. For the Sovrin infrastructure to operate reliably, it must be ensured that the paths and mechanisms by which credentials and data are exchanged across the systems are secured from unwanted and unwarranted interference to prevent tampering, and that the cryptographic operations yield accurate results. There are a total of four requirements for enabling trust in the infrastructure,[53] but only three of them fall under the technical dimension (Res.SF, 2019a), meaning they are fully enabled through cryptography:

1) The credential was issued to the presenter;
2) The credential has not been tampered with;
3) The credential has not been revoked.

Before we can address these requirements, we need to have a standard way to verify digital credentials (Res.E.SF, 2018). Two main standardization activities have been critical in achieving that, including:

1) Standardization of the format of digital credentials; and
2) Standardization of the way to verify the source and the integrity of digital credentials.

Before even standards for decentralized identifiers (DIDs) and verifiable credentials (VCs) were developed by W3C, Sovrin Identity community (members of which were instrumental in initiating within W3C the workstreams that led to those standards) anticipated and developed the layered technology stack which has since evolved to become part of the Trust over IP stack[54] (**Figure 1**).

In the previous description of Layer 1 which provides the critical foundation of this infrastructure for trust, we saw that the

[51] This summary about the ToIP governance stack is based on GitHub 0289: "The Trust Over IP Stack" where this quote and the next are taken from. See https://github.com/hyperledger/aries-rfcs/tree/master/concepts/0289-toip-stack.

[52] Hyperledger is an open-source global collaborative effort designed to advance blockchain technologies across industries. It is hosted by Linux Foundation. The three projects developed around the Sovrin code are Hyperledger Indy, Hyperledger Aries and Hyperledger Ursa.

[53] These are the same four requirements enumerated in *About Identity and Uniqueness* section above. The fourth, dealing with provenance and being more dependent on governance, belongs in the next *Social and Institutional Dimensions: The Ecosystem Governance Frameworks* section. In effect, the credential issuer (the provenance) is known by its identifier which allows referencing the DID and getting the public key to validate the credential.

[54] The first three layers of the initial Sovrin technology stack were identical to the first three of the new ToIP stack, while its fourth layer at the top addressed the required governance frameworks. In this new model, the Application Ecosystems layer emerges at the top of the stack, moving governance concerns into a separate, parallel stack. The Trust over IP Foundation is the entity that was set up to take over the work of defining the architecture of trust at the Internet scale, not only on the machine side but also on the human side.

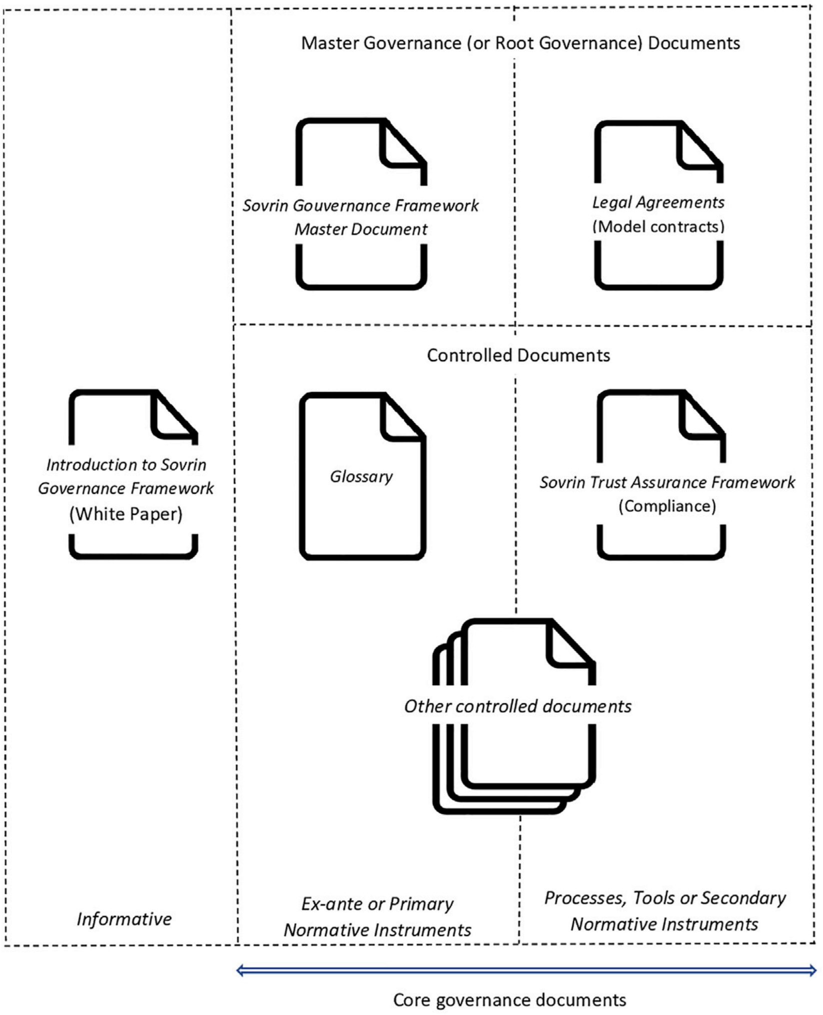

Master Governance (or Root Governance) Documents
Sovrin Gouvernance Framework Master Document
Legal Agreements (Model contracts)
Controlled Documents
Introduction to Sovrin Governance Framework (White Paper)
Glossary
Sovrin Trust Assurance Framework (Compliance)
Other controlled documents
Informative
Ex-ante or Primary Normative Instruments
Processes, Tools or Secondary Normative Instruments
Core governance documents

FIGURE 2 | Categorization of Sovrin governance framework resources.

whole edifice is rooted in a verifiable data registry of some sort. Blockchains can be used for such a system, as Sovrin does. Blockchain, or distributed ledger technology, has emerged over the last decade and, as far as digital identity is concerned, appears to afford the opportunity to develop solutions that could potentially complement the Internet itself, as is, taking us much closer to a networking experience that would flow from the Internet protocol stack itself, augmented with an identity layer. Just like the original seamlessly and globally distributed network that is the Internet, this solution avoids the risk of a single point of failure based on a distributed infrastructure for identifiers and cryptographic keys (Res.SF, 2016; Res.E.SF, 2018; Res.SF, 2018; Res.SF, 2019a) while showing a much stronger potential for data protection and security.

With Sovrin, the utility is a decentralized, public but permissioned ledger specifically designed to support identity transactions through a network of globally distributed nodes. Being a public ledger means that anyone can read from and write to it. However, it is also a permissioned ledger because using an open process, a number of entities from around the world are vetted by the Sovrin Foundation to serve as Stewards: they run the globally distributed nodes and validate transactions written to the ledger in order to enable proof-of-authority consensus whenever required.

As has been illustrated with crypto-currencies for several years, blockchain is a technology that uses cryptography to enable a kind of trust that is different from human-to-human trust: blockchains provide confidence through cryptography (De Filippi et al., 2020). In a sense, blockchain is a practical way of using technology to scale up trust to a large number of actors where trusted relationships cannot depend on personal and human-built records of past interactions as a prerequisite. In a blockchain, each transaction is digitally signed with the private key of its originator; each transaction creates a new state of transactions or a new record (block) that is logically linked to the previous one in the system, forming a chain; and once validated a transaction is replicated across all the machines on the network, using a consensus algorithm. As a result, a record for a transaction can be changed only by creating a new one (i.e., a new block). This makes blockchain transactions immutable, a property that is crucial for accountability, as every change is immutably recorded and auditable at any time thereafter.

Authenticating the author of a transaction requires knowing the public key associated with the author's signing key. The information enabling the discovery of the public key on the ledger is included in the DID document which is referenced in the credential by the issuer's DID. This information, which is the actual DID that is associated with the transaction, serves as a resource locator to discover the sender's public key. All these credential transactions are made through an encrypted peer-to-peer connection. This architecture makes it possible to do away with a central authority such as the certificate authority in the traditional PKI.

DIDs are the first globally unique identifiers that require no registration authority. They are used to assign an address to any public key and, most importantly, they enable key rotations without changing the associated DID. An SSI solution using DIDs enables the mapping of these unique identifiers to any given entity involved in credential transactions, be it a person, an organization or a connected device. With a public blockchain for DIDs, anyone can issue a digitally-signed credential, and anyone else can verify it. Public DIDs, written to a blockchain are resolvable to their DID Document, which contains public keys and service endpoints. In effect, any participant in the network can now create their own unique DIDs, attach their public keys and write them to the public ledger. Any person or entity that can locate these DIDs will be able to gain access to the associated public keys in order to verify the signing private key. Because every DID has an associated public-private key pair, anyone with a DID can digitally issue and sign verifiable claims and other documents.

For all of its above-described features, blockchain appears to be well-suited as a decentralized self-service registry for public keys[55]. Lastly, it is also worth noting that no verifiable credentials nor any personally identifiable information (PII) are stored on the ledger in the Sovrin Network. Only cryptographic resources are.

The above describes the Sovrin Network's digital credential exchange infrastructure from technological standpoint, with the components and features that will enable trust in the systems which will be built on it. Those components include edge agents and wallets, cloud agents and wallets, the standards and protocols enabling the connections and exchanges, as well as the distributed registry system (in this case, blockchain) at the root of trust, along with all the hardware devices on which all those software elements operate. However, these technical components, while necessary, are not enough to fully establish trust in the ecosystem. Trust is a human thing in that, ultimately, it has to be experienced and assessed by humans, and as such it can also be altered by human behaviors. As a result, in addition to the technical components, the trust ecosystems in this infrastructure must address social and institutional components as well, keeping in mind that the overall goal of this infrastructure is trust.

Social and Institutional Dimensions: The Ecosystem Governance Frameworks

As we've seen, digital identity transactions are not made trustworthy by technology alone. An enabling institutional environment is needed, as human decisions and behaviors may shape identity ecosystems toward either optimal or sub-optimal outcomes. As in any endeavor of societal import which depends on people's behavior, successfully building and operating this infrastructure will require governance mechanisms and authorities agreed upon by the concerned community, whether it is at global level, at nation-state level or at local level.

Governance Frameworks

The Sovrin Glossary defines a governance framework as a "set of business, legal, and technical definitions, policies, specifications, and contracts by which the members of a Trust Community agree to be governed in order to achieve their desired Levels of Assurance . . . A Governance Framework is itself governed by a Governance Authority. A Governance Framework is also known as a Trust Framework." The Pan-Canadian Trust Framework Overview,[56] an initiative that comprises many government actors, defines a trust framework as "a general term to describe a set of auditable business, technical, and legal rules that apply to the identification, authentication, and authorization of accessing resources across organizations"—or across ecosystems or whatever level of social settings the framework is referring to.

The Sovrin Governance Framework has several components which can make it look complex; however, many of those parts may evolve separately, making it modular. Setting aside informational resources, there are two sets of core documents (**Figure 2**)[57] detailing the governance requirements and arrangements for the Sovrin Infrastructure; they include the following[58].

Master Governance or "constitutional" order documents

[55] Although it is not the only one, nor are we claiming it is overall the best for that. Only time will tell.

[56] Authored by the DIACC Trust Framework Expert Committee (DIACC: Digital ID and Authentication Council of Canada). The Overview and other components of this Framework may be found here.

[57] Note that the current version of the Master Document ("Sovrin Governance Framework V2") uses the terms "constitutional" and "legislative" to categorize what it refers to as the normative documents of governance. Even though we are mentioning those terms here while describing our own categorization, both schemes don't totally match. Only the master governance category here matches exactly the constitutional category there; the legislative documents there are only a subset of all controlled documents here (see our **Figure 2** and the next footnote below for further clarifications).

[58] Except wherever otherwise indicated, the descriptions that follow are based on the content of the "Sovrin Governance Framework V2" (the current version of the so-called "Master Document"). Although in the Introduction section of that document, particularly Figure 2, Sovrin Foundation distributes the governance documents into four types or domains (informational, constitutional, legislative and compliance), we see only two meaningful categories of documents as stated here. One has to wonder whether, beyond a desire for symmetry in said figure, there is any reason of substance for the four-part grouping, since the "Sovrin Glossary" (legislative domain) and the "Sovrin Trust Assurance Framework" (compliance domain), both of which appear on their own in that grouping scheme, are elsewhere (Appendix A of the Master Document) also classified, and more accurately so, as Controlled Documents.

TABLE 3 | Self-sovereign identity core principles.

Earlier version: Core principles of SGF (2019)	Latest version: SSI principles (2020)
Self-Sovereignty Guardianship	Representation Verifiability and authenticity Control and agency Portability
Inclusive by design	Equity and inclusion
Collective best interest	
▨	Usability, Accessibility, and Consistency
Openness & Interoperability	Interoperability
Decentralization by design	Decentralization
Privacy by design	Privacy and minimal disclosure
Security by design	Security
Data protection by design and default	
Transparency	Transparency
Accountability	▨
Sustainability	▨
▨	Participation

1) The "Sovrin Governance Framework Master Document" or SGF Master Document mainly addresses core principles, core policies, and the rules applying to the revision of all governance documents. The SSI Principles as formulated by the Sovrin community in its latest update (December 2020) will be detailed below. The core policies are elaborated rules in keeping with some of the main principles and they address topics such as stewardship, guardianship, inclusion, trust assurance and the economics of the Sovrin Infrastructure and the Foundation's finances. Note that the governance framework as described in the Master Document serves as a reference and a foundation on which domain-specific governance frameworks may further be built as needed for the purposes of different use contexts. In that sense, one may refer to this material as the "Master Governance" or "Root Governance."
2) The Legal Agreements are model contracts written as generic contract templates. The three templates are between the Sovrin Foundation on the one hand and, on the other, all stewards who operate nodes on the Sovrin Ledger ("Sovrin Steward Agreement"), all identity owners writing transactions on the Ledger ("Transaction Author Agreement") and any organizations using permissioned write access to the Ledger ("Transaction Endorser Agreement"). Later on, two more were added which are the "Steward Data Processing Agreement" and the "Transaction Endorser Data Processing Agreement," both of which establish the responsibilities of the two contracting parties for complying with GDPR and other data protection regulations[59].

Controlled Documents, including Documents of "legislative" order The Controlled Documents are subdocuments to the Master Document in which they are referenced as normative components of the governance framework. They may include technical specifications, standards, and policies that are independently maintained and versioned either by the Sovrin Foundation (e.g., the Sovrin DID Method) or external standards bodies (e.g., W3C, OASIS[60]). The following two documents are also controlled documents although they are sometimes mentioned separately from that group of documents (as explained in footnote 58).

1) The "Glossary" provides definitions for the terminology (about 250 entries in alphabetical order) used in all publications of the Sovrin Foundation in connection to subjects such as digital identity, the Sovrin Infrastructure, its operations and its governance, etc.

[59] See https://sovrin.org/library/sovrin-governance-framework/.

[60] The World Wide Web Consortium https://www.w3.org/and the Organization for the Advancement of Structured Information Standards https://www.oasis-open.org/.

2) The "Sovrin Trust Assurance Framework" defines criteria and processes for assessing conformance of Sovrin actors, including the Foundation itself, to the policies of the Sovrin Governance Framework.

One last governance document referenced in the current version of the Master Document is "An Introduction to the Sovrin Governance Framework V2" which is a white paper, an informational resource intended to serve as an overall guide to the governance framework.

Basically, the first category of documents above, which contains the core elements of the governance mechanisms, is meant to be more stable in its content, as they require community consultation and consensus before they can be modified; while the second category of materials, the Controlled Documents, may more easily be revised through less demanding simple administrative procedures, just as much as their contents are more likely to evolve as the Sovrin Infrastructure grows and its environment evolves.

One last feature of this categorization scheme, which needs to be accounted for, is the vertical compartmentalization of the core documents, as per **Figure 2** (see the labels at the bottom of the figure). The Ex-ante or primary normative instruments are the stronger normative documents which may be commanded by fundamental values (e.g., SGF Master Document) or by some objective constraints (e.g., Glossary),[61] all of which are normative a priori and are hardly negotiable, albeit still subject to change. The other category, Processes, tools or secondary normative instruments, comprises a number of things, but let us start with the last part of this label. For instance, a contract or legal agreement—particularly a signed one—has a normative dimension. However, one may argue that the source of its normative force is not the contract itself but a higher-level normative instrument, such as the legal system that backs it up, for instance. In addition, generic contract templates are tools that can be crafted in advance because there are principles and policies (part of the normative sources) that direct which clauses need to be in there and their wording. They thus qualify as secondary normative instruments. Lastly, the same category also includes various tools or resources that may help conduct or document a process, such as verifying compliance.

After this overview of the Sovrin ecosystem governance frameworks, let us now turn the focus on the core principles which were initially spelled out as part of the governance framework documents but are considered important enough by the Sovrin actors for them to be recently updated and published as a stand-alone document[62].

Principles of Self-Sovereign Identity

Over recent years, the Sovrin Foundation and its community have built consensus on twelve principles to guide their technical architecture, their services and their practices. The latest version (Res.SF, 2020) of those twelve principles is somewhat different from the earlier version provided in the Sovrin Governance Framework[63] (Res.SF, 2019b). In the new version, some of these principles are further broken down into sub-principles or values to be observed in practice. In turn, those core principles inform core policies that should guide all the actors in their respective roles in the ecosystem. The core policies address the following topics: stewardship, guardianship, inclusion, trust assurance and economics.

Table 3 compares the two versions of principles (aligning each principle in the latest version with the closest principle from the previous version), reinforcing the continuity of the most fundamental ideas behind those principles. Underlying these principles are a number of essential rights, norms and values. They affirm the autonomy of identity holders and acknowledge the need to empower them to exercise such autonomy to the maximum extent possible in the SSI space. These range from the right to seek and obtain any number of digital representations needed as verifiable and provable identities[64] to the right, along with the technological capability, to control any consequential use of one's own identity data by any party, a right which they can exercise directly themselves or delegate to agents or guardians of their choice. The principles carrying or enabling the value of autonomy thus understood include: representation; verifiability and authenticity; control and agency; as well as, to some extent, decentralization[65] and security. But autonomy can only be fully experienced if a number of other rights and freedoms are available to the subject population. These include the right to keep one's personal business only to oneself (right to privacy), as well as the digital equivalent of the freedom of movement. The latter implies identity rights holders can move around unfettered with their digital identity data, credentials and related cryptographical accessories. The principles that cover these are: privacy and minimal disclosure; interoperability; and portability.

While all the principles have a foundation in a set of values, they range on a spectrum between technology design and governance. For instance, the interoperability principle (just like portability) is somewhat based on the belief that the digital world would be a better place if people can enjoy—as

[61] Words have basic, collectively-accepted ("objective") meanings and the Glossary entries have to reflect the concepts and terms needed to describe at best the subject and processes at hand for all participants to speak the same language and to understand each other.

[62] Note that if the SSI Principles were to be considered as a separate document, the latter would be sitting in the same category corner as the SGF Master Document that initially included the Principles.

[63] See the Sovrin Governance Framework V2 dating from December 4, 2019, including sub-principles and core policies, whereas the latest version which is, at this point in time, available on the Foundation's website https://sovrin.org where the principles are simply formulated without further elaboration, dates from December 10, 2020.

[64] This also highlights another central tenet of the SSI worldview which is inclusiveness. Not only shall an SSI ecosystem avoid any form of discrimination or exclusion toward any potential identity holder, it must also proactively seek to facilitate access to and usability of all its components.

[65] One may also note that decentralization is a technical and design requirement for these essential rights and norms to be applicable. At any rate, it is a necessary and an overarching requirement for the SSI architecture.

they do in the physical world—the freedom to move around with any pre-existing identities that they may have. Making that possible heavily depends on technical components and the way the system is designed. Whereas the equity and inclusion principle is also value-based, and more directly so, its implementation leans more heavily on governance than on technical components.

The ten SSI principles in the color shaded cells (with at least one corresponding principle on the earlier version side) of **Table 3** might be seen as the ones with the more enduring core ideas and values for a Self-sovereign identity culture. Although some of them—such as privacy and security—are relatively common principles in information systems and networks, their level of impact on SSI ecosystems is uncommon by contrast with other comparable systems; in other words, these principles along with their level of requirement are characteristic of SSI.

In the context of SSI, the human and institutional requirements for trust are spelled out in a number of resources and tools, from fundamental principles and core policies to contractual agreements and other business and administrative procedures, which the governance frameworks particularly encapsulate[66]. In this context, the role of the governance frameworks is to create rules that make sense to human beings and will regulate their behavior, making it more predictable. This allows for every participant to know what is expected in their role and what to expect from the other roles in the ecosystem—and as a result, this enables all participants to act accordingly and predictably in the best interest of all. The governance resources and tools are needed to organize relationships in the ecosystem and steer it effectively as it grows. They create the conditions for a shared understanding and therefore, they are a critical component for the trust needed for the ecosystem to work with little to no friction.

In the final section on the subject matter of this paper, I outline different scenarios as possible pathways for governments on the way forward with regard to digital identity on the Internet. In addition, I formulate a few recommendations which they may want to consider while making decision to engage.

PATHWAYS AND BASIC GUIDELINES FOR POLICY-MAKERS

Government Pathways for Identity on the Internet

Government-issued identity credentials are considered authoritative by all stakeholders. They provide for an identity that qualifies as legal identity because it must have, and has, the capability to enable legal accountability, whether negative (e.g., attribution of liabilities) or positive accountability (e.g., attribution of assets and properties.) There are a couple of reasons for that.

First, the primary identity subjects of interest, the human subjects, are embodied living beings, evolving in physical settings. Governments rule the physical world by legislating and enforcing the laws they make within their geographical jurisdiction. Those laws and any legally enforceable rules governments make are the most objectively binding of social-ordering tools, applicable to people in their physical settings which, for most part of the world, are under the jurisdiction of a government. Those laws and legally enforceable rules generally apply to the society as a whole—this includes most members, who are law-abiding citizens—and, as such, the society as a whole has interest in accountability.

Second, in terms of accountability, the material "price to pay" for infringing those laws and rules is usually the highest compared to other applicable rules in the public sphere, sourced from any other authority. In other words, every regular person has a stake when it comes to legal accountability and that stake is significant. Most people wouldn't want to incur the actual cost of such infringement, which validates the deterrence function of those social-ordering tools.

The above is, from our analysis (Chango, 2012), the rationale that played out at the beginning of the era of identity papers. In effect, the history of paper-based identity credentials shows that in the early days of issuing those credentials to the broad public and throughout the 19th century, there first was a wide range of variety of identity papers created independently from any common standards or reference model by a whole host of collective entities (companies to employees; places of public accommodation to customers and users; associations, clubs, or other membership groups to members, etc.). Then progressively they made way for the government-issued identity document to emerge as a singular source of authoritative identity, and eventually piggybacked onto it.

On the other hand, on the Internet or any open, public digital network:

1) There is no ruling entity, no single entity is in charge of the network space;
2) No entity makes law or any legally enforceable rule on the whole network and its users;
3) no single entity exists to provide the network-based equivalent of legal or foundational identity for the whole of the network.

However, relevant Internet technical communities and various user stakeholders have shown they can work together and reach consensus to formulate protocols and deliver technical standards so as to enable the ascertainment of the provenance, the integrity and the validity status at any given point in time of identity data.

A thought experiment building on historical and socio-political experiences of identity credentials from the pre-digital era, as well as on contemporary experiences with Internet governance, leads to the following pathways to map out the possible future of the government response to the Internet identity challenge:

[66] While the Sovrin Network is decentralized, it still operates under a community-driven governance framework whose goal is to maximize trust in Sovrin as a global identity network. As of October 2020, the Sovrin Governance Framework is defined in a set of documents including three primary documents, three legal agreements and six controlled documents. The Sovrin Network enforces rules through a mixture of open-source code and an active, open governance process for rulemaking starting from the development of the rules.

1) Build nothing really new in terms of online digital identity system—only digitally enabled physical credentials (biometrics, QR codes, etc.) Mobile or Web applications, plus any other hardware accessories as necessary, may enable people to use those credentials in online transactions.
2) Individual nation-states collaborate with the Internet technical community in order to establish a single, national foundational ID system for digital credentials. All interested institutions and Web services operating from within that country's ccTLD namespace on the Internet would be required to use this for identification in applicable online transactions involving the citizens of the concerned nation-states.
3) A collection of nation-states gets together and develop their own specifications, awarding grants to, or procuring from, the technical community, academia or the private sector in order to build their own system as per their requirements for use only in the adhering countries.
4) The Internet technical community develops a set of technically robust solutions taking into account most governments' concerns as well as other stakeholders', setting a framework for solutions that are open enough in their design so as to accommodate interoperability, evolution and further improvements while meeting government standards and expectations of security. More and more governments adopt solutions based on that framework, enabling their respective legal digital ID systems to interoperate and their related credentials to be recognized and accepted in online transactions based on that framework, regardless of national boundaries, beyond ccTLD namespace and across the gTLD namespace[67].

Self-sovereign identity has the potential to realize the latter scenario which will require the use of standards, certainly more likely so than the first three options. The Internet technical community, along with interested stakeholders, has taken the lead for developing the necessary and appropriate standards and writing open-source code libraries. The challenge now is to bring policy-makers onboard, first by translating the critical capabilities of the technology into meaningful policy language, while highlighting potential comparative advantages.

Policy Recommendations

With an SSI infrastructure in place, no industry, sector or group of actors seeking to enable trusted credential exchange online needs to build the technology from scratch. Their priority, instead, will be to elaborate their governance frameworks (defining the business, legal and technical rules for their operations), and make sure they are in alignment with the law and regulations of the jurisdiction(s) to which they must be accountable. An SSI network infrastructure, such as the Sovrin Network, is not designed to offer any one particular identity system, or a definite set of systems, directly to Internet users or any subset thereof (e.g., the nationals of a country), but rather to provide the infrastructure needed for identity issuers, owners and verifiers to securely engage in credential exchanges using identity systems of their choice;[68] in that sense, it is an identity metasystem (Cameron 2005; Windley 2021). The only requirement is that those systems operate by the principles, rules and agreements defined through the governance framework at the metasystem level, as applicable to the domain at hand and to the roles of the participants. Those rules are collectively defined or agreed upon by the ecosystem participants for their collective best interest and for an optimal outcome. A notable benefit of this architecture is making specific SSI solutions potentially scalable across the Internet.

For the purpose of deploying an SSI solution at national level, policy-makers may choose to develop the country's own governance framework,[69] or review and adapt existing ones such as the Sovrin Governance Framework or the Pan-Canadian Trust Framework (PCTF). A network infrastructure such as the Sovrin Infrastructure (among other SSI solutions) presents a good opportunity for governments seeking innovative solutions for identity management and related cybersecurity concerns for the delivery of their e-government services. On the other hand, it requires a lot of time and a great deal of collective, yet specialized wisdom and skills, to be developed, maintained and continuously improved. That task is better left to technology professionals dedicated to building and running the infrastructure for such networks. Governments may however start discussions with those actors in order to define the terms of a partnership addressing their specific concerns and requirements, including the development of appropriate governance frameworks guided by the applicable laws and regulations in their country.

In any case, for government-backed credentials, the government is obviously well suited to be one of the governance authorities, either directly or through a delegation of authority, depending on government choice and capabilities. Even in those cases, given the nature of the technology as well as the complexity of its implementation setting, the governance authority would be better carried out through a public-private partnership. For while law enforcement still remains the responsibility of the government, there are domain-specific governing rules which are equally binding for participants, although initially subscribed to voluntarily.

With the insights gained through experience and this research, we close this paper with the following guidelines for policy-makers and other interested policy stakeholders, particularly but not exclusively in countries which are the farthest from the places where the technology is actually emerging.

[67] The two larger categories of names in the Internet domain name system include: 1) the country-code top-level domain (ccTLD) where the suffix of the domain name is a two-character code identifying a country (such as .tg for Togo and .us for the United States), and 2) the generic top-level domain (gTLD) where the suffix of the domain name is a generic, transversal identifier such as .com or .org. Other categories of top-level domains have emerged over the years but those two remain the historical ones and still the most largely used.

[68] It is at the level of these identity systems where particular identity and credential definitions are provided.

[69] The Sovrin Foundation has anticipated the need for itself to further develop Domain-Specific Governance Frameworks (DSGFs) in addition to its primary Governance Framework.

Building Trust-Enabling Governance Frameworks

For their national digital identity solution, policy-makers may choose to develop their country's own governance framework, taking into account applicable laws and regulations in their national jurisdiction as well as basic SSI principles and best practices. Doing so promotes trust within their national ecosystem. Alternatively, they may choose to review and adapt an existing framework. Partnerships may be developed with technology professionals and communities that have developed governance frameworks while building and running similar network infrastructure.

Multistakeholder Governance

Governments working with other stakeholders might want to put in place at least one multi-stakeholder structure (possibly including global membership or liaisons to relevant global processes or groups) to monitor the implementation of their governance and trust framework, to deliberate on critical decisions to make, and recommend best practice solutions for any issues their SSI project and operations may encounter. This could have a particular focus on security and rights within the confines of applicable law, regulations and policies. We may generically refer to that multistakeholder structure as the Digital Credential Exchange Council. It should use open decision-making processes including public consultations whenever relevant.

No National Boundaries for SSI-Interoperable Solutions

A national or a country-bound ecosystem should not be understood as an instantiation of national territories and boundaries—and whatever this entails—in the digital realm. Here, an ecosystem is a defined set of actors sharing the same set of rules and procedures around the same infrastructure and for the same purpose. Beyond that, some identity features we are accustomed to in the physical world still obtain: the identity issuer does not define whom I can present my credential to, nor does she/he need to know whenever or wherever I present my credential. It is up to the verifier or relying party to decide whether my credential is an acceptable proof for their purpose. Therefore, citizens who own or hold digital credentials from any national or government-backed ecosystem are still allowed to use them, in digital interactions and transactions where the counterparty is not a participant in the issuing ecosystem—provided that the technology components in that ecosystem be based on interoperable specifications and standards as relating to SSI. This makes it possible for citizens of a given nation holding SSI-compliant or SSI-compatible government-issued digital credentials to both enjoy the access to, and the use of, their e-government services and to conduct business online globally with any entities that operate under the SSI framework.

No Digital Identity for Developing Countries vs. Developed Countries

More particularly in developing countries and also emerging economies, it is important that governments avoid running to solutions intended only for that group of countries. In their deliberations and decision-making on this issue, and while retaining their right to adapt existing solutions to their needs, these countries need to take into account the gains made anywhere with these evolving identity technologies and practices, including in the more advanced digital economies. Likewise, solution packages pushed through public international institutions or bilateral state-to-state relations, should not be embraced without vetting them against the backdrop of the global technology developments outlined above. The true digital economy will be global or it won't be.

Preference to Interoperable Solutions Using Adopted Technical Standards

Governments should refrain from being quickly sold on any specific turn-key digital identity solution in the market, more particularly proprietary ones, without carefully considering interoperability and long-term value. Preference should be given to solution components that have been developed and tested by a broad base of the technical community. For instance, governments should be informed of the standardization processes, notably with the W3C's activities on digital identifiers (DID) and verifiable credentials (VC), and favor the use of those standards wherever warranted in developing solutions for their digital records and identity needs (including for instance the digital vital records of their citizens).

An SSI Bill of Rights?

At a global level, SSI is based on a set of principles and values. Each government should consider issuing one form or another of a Bill of Rights for their SSI space. Or alternatively, they may issue a comprehensive "Declaration of Rights and Obligations", aiming at making the SSI principles—among possible other regulations and legal provisions—enforceable in their ecosystem and at the level of their national jurisdiction.

CONCLUSION

As outlined in *A Three-phase Evolution* section of this article, we are at Phase III in the conceptualization of the historical evolution of identity mechanisms where digital technology is redefining the boundaries of the self in so many ways that we cannot fully address digital identity without addressing it for the Internet, the largest, most common, and mother of all digital networks. At this point, we cannot simply renew the paper-based logic with digital plugs or on digital surfaces, by generating electronic copies of physical credentials and pushing them through digital transmission channels or storing them in digital databases, all of that with the same analogical mindset and way of handling credentials. The digital playing field[70] holds its own logic, methods and forms which need to be brought to bear on all the different ways the society used to handle and leverage

[70] Where humans' digital existence and agency unfold, across all the activities they need to conduct through digital representations in order to sustain or entertain their life, including their business.

credentials, plus more ways it might still need to use them, in order for the digital to unleash its full potential in that regard.

As we have argued, identity is not a monolithic informational representation of the self. The Internet identity challenges have helped us understand that identity is required wherever any claim whatsoever is made by any entity endowed with agency through digital networks. And because anything that is done through the Internet, indeed through any digital network, boils down to an exchange of information, claims will always be made about something or another in the course of a transaction, starting with the entities that are part of the transaction. Since there is not a central digital authority governing for all the ins and outs of digital transactions, ensuring how claims are made and ascertained in digital networks (thus, digital identity) is paramount to enabling and securing transactions of any sort across any digital networks. By providing sound analytical arguments for a useful distinction between identity as "What you are" vs. "Who you are," a wider range of identity-based interactions is shown to be possible online, without even the burden of a registration or of an account. We thus realize more clearly that digital identity may bring in new challenges (which are being resolved one after the other) but it certainly also opens up a much broader scope for effective agency than identity in the physical world. This, in addition to the fact that we can obviously reach farther and more rapidly through digital networks (wherever they are available) than we have ever done using any other record-making technique along with the applicable communication capabilities, verifies in this instance our theory as formulated at the beginning of this paper.

The SSI model presents a good opportunity for governments and other institutions seeking innovative solutions to identity management online, while improving security and preserving privacy, particularly with regard to the delivery of their e-government services. Furthermore, this emerging technology, including decentralized identifiers and verifiable credentials, does more than just provide digital identity to individuals. It is also critical to organizations, companies, institutions whose assets also need to be digitally and securely represented in the digital economy. In more general terms, this technology allows putting a workable structure on piles of user-generated data mostly scattered across silos and in a variety of heterogeneous formats over the Internet and related networks. The technology, and the relationships which it helps foster in various ecosystems, make it possible to assign data where data belongs, to bind data to their legitimate subject as well as to most relevant and trustworthy sources, while enabling its secure and rapid exchange. As digital assets broadly become more manageable, this will open the gates to a thriving digital economy.

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.

ACKNOWLEDGMENTS

I would like to acknowledge with thanks the reviewers for taking the time to engage with the earlier versions of the manuscript and help improve it with their comments and suggestions, as well as the Handling Editor and the whole editorial team for their patience during the long process that has led to this result. I am specially indebted to my "peers in identity," Phil Windley and Kaliya Young (a.k.a "Identity Woman"), who have generously accepted to assist with proofreading and eventually did much more than that: through their thoughtful reading and insightful comments they have tremendously helped clarify my thinking and improve this written piece. If, despite all the assistance I have received from them all, there are still flaws and errors to be found in the final result, they are mines alone.

REFERENCES

Abdelnour, S., Hasselbladh, H., and Kallinikos, J. (2017). Agency and Institutions in Organization Studies. *Organ. Stud.* 38 (12), 1775–1792. doi:10.1177/0170840617708007

A. Preukschat and D. Reed (Editors) (2021). *Self-Sovereign Identity: Decentralized Digital Identity and Verifiable Credentials* (Shelter Island, NY: Manning Publications).

Antonopoulos, A. (2014). Bitcoin Security Model: Trust by Computation. A shift from trusting people to trusting math. O'Reilly Radar: February 20, 2014. Available at: http://radar.oreilly.com/2014/02/bitcoin-security-model-trust-by-computation.html (Accessed Dec24, , 2021).

Baier, A. (1994). *Moral Prejudices*. Cambridge, MA: Harvard University Press.

Baier, A. (1986). Trust and Antitrust. *Ethics* 96 (No. 2), 231–260. doi:10.1086/292745

Barney, J. B., and Hansen, M. H. (1994). Trustworthiness as a Source of Competitive Advantage. *Strat. Mgmt. J.* 15, 175–190. doi:10.1002/smj.4250150912

Bedos-Rezak, B. (2000). Medieval Identity: A Sign and a Concept. *Am. Hist. Rev.* 105 (No. 5), 1489–1533. doi:10.2307/2652028

Blanchette, J.-F. (2012). *Burdens of Proof: Cryptographic Culture and Evidence Law in the Age of Electronic Documents*. Cambridge, Mass: MIT Press.

Bolter, J. D. (1991). *Writing Space: The Computer, Hypertext, and the History of Writing*. Hillsdale, N.J.: Lawrence Erlbaum Associates Publishers.

Cameron, K. (2005). The Laws of Identity. Available at: https://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf (Accessed Dec 3, 2021).

Chango, M. (2012). *Becoming Artifacts: Medieval Seals, Passports and the Future of Digital Identity* (Syracuse, NY: Syracuse University). thesis.

Clanchy, M. T. (1993). *From Memory to Written Record: England 1066-1307*. Cambridge, Mass: Blackwell Publishers.

Davis, N. Z. (1983). *The Return of Martin Guerre*. Cambridge: Harvard University Press.

De Filippi, P., Mannan, M., Reijers, W., and Reijers, W. (2020). Blockchain as a Confidence Machine: The Problem of Trust & Challenges of Governance. *Techn. Soc.* 62, 101284. doi:10.1016/j.techsoc.2020.101284

Dimock, G. E., Jr. (1956). The Name of Odysseus. *Hudson Rev.* 9 (No. 1), 52–70. (Spring 1956),. doi:10.2307/3847614

Doc.LN (1922). "League of Nations (Advisory and Technical Committee for Communications and Transit)," in *Replies of the Governments to the Inquiry on the Application of the Resolutions Relating to Passports, Customs Formalities and through Tickets*. Geneva: League of Nations. Doc.C.183.M.101.1922.VIII.[71].

Doc.LN (1920). "League of Nations (Provisional Committee on Communications and Transit)," in *Conference on Passports, Customs Formalities and through Tickets: Resolution Adopted by the Conference*. (Paris: League of Nations) October 21st 1920.

Doc.UN (1947). *United Nations, Economic and Social Council. Official Records*. New York, NY: United Nations. Second Year, 5th Session, Supplement No. 1, 1947. United Nations Doc. E/436 (Recommendations of the Committee of Experts from Geneva meeting, 14-25 April 1947 in the Appendix).

Doc.UN.1956 United Nations. Doc. E/2933, 23 November 1956 and Addenda. United Nations Publication.

Doc.UN.1959 United Nations. Doc. E/CN.2/190, 1 May 1959. United Nations Publication.

Doc.UN.1961 United Nations. Doc. E/3438/Addendum 1, 27 February 1961 and Addenda. United Nations Publication.

Doc.UN (1963). *Recommendations on International Travel and Tourism. United Nations Conference on International Travel and Tourism*. Rome: United Nations Publication. 21 August – 5 September 1963. Doc.E/CONF.47/18.

Doc.UN.1966 United Nations. 1966. Report of the Secretary-General. Doc. E/4145, 5 January 1966 and Doc.E/4145/Add.1, 8 June 1966. United Nations Publication.

Duranti, L., Eastwood, T., and MacNeil, H. (2002). *Preservation of the Integrity of Electronic Records*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Emirbayer, M., and Mische, A. (1998). What Is agency?. *Am. J. Sociol.* 103 (No. 4), 962–1023. doi:10.1086/231294

Foucault, M. (1988a). "Technologies of the Self," in *Technologies of the Self: A Seminar with Michel Foucault*. Editors L. H. Martin, H. Gutman, and P. H. Hutton (Amherst: The University of Massachusetts Press), 16–49.

Foucault, M. (1988b). "The Political Technology of Individuals," in *Technologies of the Self: A Seminar with Michel Foucault*. Editors L. H. Martin, H. Gutman, and P. H. Hutton (Amherst: The University of Massachusetts Press), 145–162.

Fraenkel, B. (1992). *La Signature: Genèse D'un Signe*. Paris: Gallimard.

George, A. L., and Bennett, A. (2005). *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

Grant, M. (1946). *From Imperium to Auctoritas*. Cambridge: Cambridge University Press.

Hall, J. (1999). *Cultures of Inquiry: From Epistemology to Discourse in Sociohistorical Research*. Cambridge: Cambridge University Press.

J. Caplan and J. Torpey (Editors) (2001). *Documenting Individual Identity: The Development of State Practices in the Modern World* (New Jersey: Princeton University Press).

J. H. Burns (Editor) (1988). *The Cambridge History of Medieval Political Thought c.350-c.1450*. (New York and Cambridge: Cambridge University Press).

J. Perry (Editor) (1975). *Personal Identity*. second edition. (Berkeley, Calif: University of California Press).

Kantorowicz, E. H. (1951). 1951. Pro Patria Mori in Medieval Political ThoughtAmerican Historical Association. *Am. Hist. Rev.* 56 (no. 3), 472–492. doi:10.2307/1848433

Kantorowicz, E. H. (1955). "Mysteries of State: An Absolutist Concept and its Late Medieval Origins," in *The Harvard Theological Review* (Cambridge, Mass: Cambridge University Press), Vol. 48, 65–91. doi:10.1017/s0017816000025050

Lloyd, M. (2003). *The Passport: The History of Man's Most Travelled Document*. Gloucestershire: Sutton Publishing.

López, M. A. (2020). *Self-Sovereign Identity. The Future of Identity: Self-Sovereignty, Wallets, and Blockchain*. Inter-American Development Bank.

MacNeil, H. (2000). *Trusting Records: Legal, Historical and Diplomatic Perspectives*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Nickel, P. J., and Vaesen, K. (2012). "Risk and Trust," in *Handbook of Risk Theory*. Editors S. Roeser, R. Hillerbrand, M. Peterson, and P. Sandin (Berlin: Springer).

Noonan, H. W. (1989). *Personal Identity*. Second edition. New York: Routledge.

Parfit, D. (1984). *Reasons and Persons*. New York: Clarendon/ Oxford University Press. Reprinted with further corrections, 1987.

Piazza, Pierre. (2004). *Histoire de la Carte Nationale d'identité*. Paris: Odile Jacob.

Res.E.SF (2018). Evernym, Inc. And Sovrin Foundation. Sovrin: What Goes on the Ledger. (September 2018, first published in April 2017). Available at: https://sovrin.org/wp-content/uploads/2018/10/What-Goes-On-The-Ledger.pdf (Accessed in June, 2021).

Res.GitH.0289 (2019). The Trust over IP Stack (Hyperledger Aries RFC 0289). Available at: https://github.com/hyperledger/aries-rfcs/tree/master/concepts/0289-toip-stack (Accessed June 26, 2021).

Res.SF (2016). How Sovrin Works. A Technical Guide from the Sovrin Foundation (3 October 2016). Available at: https://sovrin.org/wp-content/uploads/2018/03/How-Sovrin-Works.pdf (Accessed June 12, 2021).

Res.SF (2018). Sovrin: A Protocol and Token for Self-Sovereign Identity and Decentralized Trust. (A White Paper for the Sovrin Foundation, version 1.0, January 2018). Sovrin Foundation. Available at: https://sovrin.org/wp-content/uploads/Sovrin-Protocol-and-Token-White-Paper.pdf (Accessed in June, 2021).

Res.SF (2019a). Sovrin Foundation. Sovrin Governance Framework V2. December 4, 2019. Available at: https://sovrin.org/wp-content/uploads/Sovrin-Governance-Framework-V2-Master-Document-V2.pdf (Accessed in June, 2021).

Res.SF (2019b). How DIDs, Keys, Credentials, and Agents Work in Sovrin. Sovrin Foundation. Available at: https://sovrin.org/wp-content/uploads/2019/01/How-DIDs-Keys-Credentials-and-Agents-Work-Together-in-Sovrin-131118.pdf (Accessed June 12, 2021).

Res.SF (2020). The Principles of SSI. Available at: https://sovrin.org/principles-of-ssi/ (Accessed June 26, 2021).

Res.ToIP (2020). Introducing the Trust over IP Foundation. Available at: https://trustoverip.org/wp-content/uploads/sites/98/2020/05/toip_introduction_050520.pdf (Accessed in June, 2021).

Robertson, C. (2010). *The Passport in America: The History of a Document*. Oxford University Press.

Solove, D. (2003). "Identity Theft, Privacy, and the Architecture of Vulnerability," in *Hastings Law Journal 54*. San Francisco: Calif.

Solove, D. (2004). *The Digital Person: Technology and Privacy in the Information Age*. New York: NYU Press.

Sassen, S. (2006). *Territory, Authority, Rights: From Medieval to Global Assemblages*. Princeton: Princeton University Press.

Somers, M. R. (1994). The Narrative Constitution of Identity: A Relational and Network Approach. *Theor. Soc.* 235, 605–649. doi:10.1007/bf00992905

Somers, M. R. (1998). We're Not Angels": Realism, Rational Choice, and Relationality in Social Science. *Am. J. Sociol.* 104 (No. 3), 722–784. doi:10.1086/210085

Stanton, J., Chango, M., and Owens, J. (2007). "ICAO and the Biometric RFID Passport: History and Analysis," in Paper presented by first author at the Research Workshop on National ID Cards at Queen's University. June 7–9, 2007.

Tilly, C. (2006). "Why and How History Matters," in *The Oxford Handbook of Contextual Political Analysis*. Editors R. E. Goodin and C. Tilly (Oxford and New York: Oxford University Press), 417–437.

Tilly, C. (2008). *Explaining Social Processes*. Boulder, CO: Paradigm Publishers.

Torpey, J. (2000). *The Invention of the Passport: Surveillance, Citizenship and the State*. New York: Cambridge University Press.

Turack, D. C. (1968). Freedom of Movement and the International Regime of Passports. *Osgoode Hall L. J.* 6 (2), 230–251.

Vernant, J.-P., and Ker, J. (1999). Odysseus in Person. *No.* 67, 1–26. doi:10.2307/2902884

Werbach, K. D. (2016). *Trustless Trust. Paper Presented at the TPRC Conference on TelecommunicationsVA*. Arlington: Information and Communications Policy.

Wicks, A. C., ShawnBerman, L., and Jones, T. M. (1999). The Structure of Optimal Trust: Moral and Strategic Implications. *Acad. Manage. Rev.* 24 (No. 1), 99–116. doi:10.5465/amr.1999.1580443

Wilson, S. (1998). *The Means of Naming: A Social and Cultural History of Personal Naming in Western Europe*. Bristol, PA: UCL Press.

Windley, P. (2021). Sovrin: An Identity Metasystem for Self-Sovereign Identity. *Front. Blockchain* Vol. 4. Article 626726. doi:10.3389/fbloc.2021.626726

Wolter, U. (1997). "The *Officium* in Medieval Ecclesiastical Law as a Prototype of Modern Administration," in *Legislation and Justice*. Editor A. Padoa-Schoppa (New York: Clarendon Press).

World Bank (2019). *ID4D Practitioner's Guide: Version 1.0 (October 2019)*. Washington, DC: International Bank for Reconstruction and Development. World Bank License: Creative Commons Attribution 3.0 IGO (CC BY 3.0 IGO).:

World Bank (2018). *Technology Landscape for Digital Identification*. Washington, DC: International Bank for Reconstruction and Development. orld Bank License: Creative Commons Attribution 3.0 IGO (CC BY 3.0 IGO).

Conflict of Interest: The corresponding author is the sole proprietor of DigiLexis Consulting.

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership