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

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-sovererign-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.

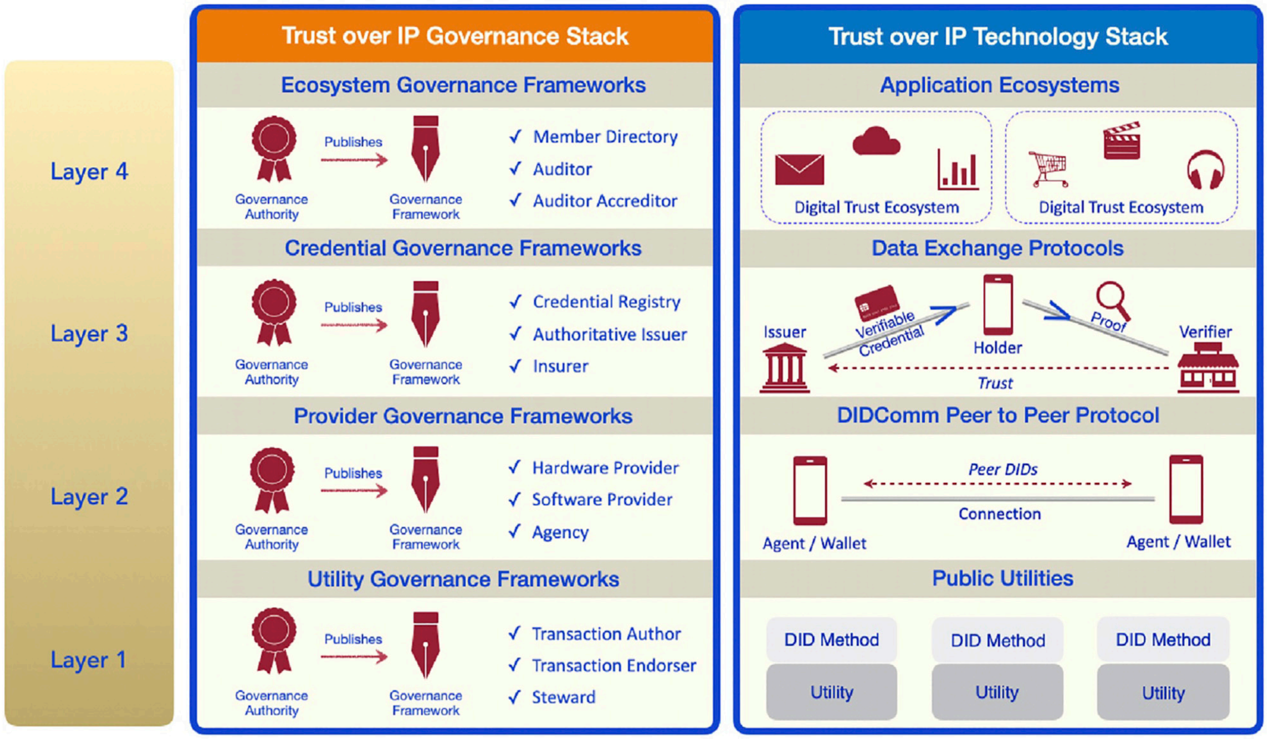

FIGURE 1 | Architecture of the Trust over IP (ToIP) stack. (Source: Trust over IP Foundation https://trustoverip.org/).

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.

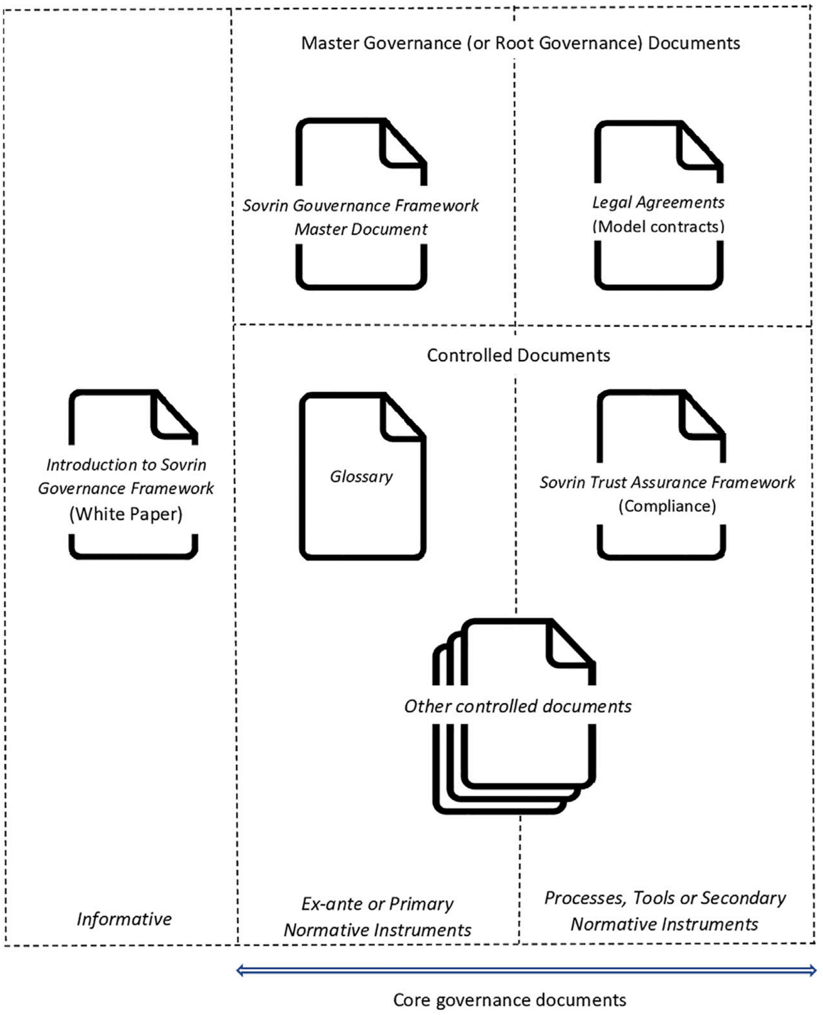
Master Governance (or Root Governance) Documents
Sovrin Gouvernance Framework Master Document
Legal Agreements (Model contracts)
Controlled Documents
Introduction to Sovrin Governance Framework (White Paper)
Glossary
Sovrin Trust Assurance Framework (Compliance)
Other controlled documents
Informative
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.