Sovrin™: A Protocol and Token for Self-Sovereign Identity and Decentralized Trust

A White Paper from the Sovrin Foundation

Version 1.0

January 2018
Abstract

Digital identity is one of the oldest and hardest problems on the Internet. There is still no way to use digital credentials to prove our online identity the same way we do in the offline world. This is finally changing. First, the World Wide Web Consortium is standardizing the format of digitally-signed credentials. Secondly, public blockchains can provide decentralized registration and discovery of the public keys needed to verify digital signatures. These two steps pave the way to establish a global public utility for self-sovereign identity—lifetime portable digital identity that does not depend on any central authority and can never be taken away. The Sovrin Network has been designed exclusively for this purpose, including governance (the Sovrin Foundation and the Sovrin Trust Framework), scalability (validator and observer nodes and state proofs), and accessibility (minimal cost and maximum availability). Most importantly, Sovrin implements Privacy by Design on a global scale, including pairwise pseudonymous identifiers, peer-to-peer private agents, and selective disclosure of personal data using zero-knowledge proof cryptography. The emergence of this infrastructure can transform at least four major markets: identity and access management, cybersecurity, RegTech, and data integration. To provide economic incentives for credential issuers, owners, and verifiers, the Sovrin protocol will incorporate a digital token designed expressly for privacy-preserving value exchange. The Sovrin token should enable a global marketplace for digital credentials of all types and value levels together with ancillary markets for digital credential insurance and permissioned first party data (direct from the customer).

Table of Contents

Part 1: The Problem ................................................... 3
Part 2: The Solution .................................................... 8
Part 3: Identity for All .................................................. 12
Part 4: Privacy for All .................................................. 19
Part 5: The Impact ...................................................... 24
Part 6: The Token ....................................................... 30
Conclusion .............................................................. 40
Next Steps .............................................................. 41
Acknowledgements ..................................................... 41
PART ONE

The Problem
Digital identity is one of the oldest and hardest problems on the Internet

And it is only getting worse.

Although this famous New Yorker cartoon was first published in 1993, it remains true even today. Despite a quarter-century of advances in Internet technology, there is still no easy way to prove online that you are not a dog, are over 18, live at a certain address, graduated from a certain school, work at a specific company, or own a specific asset. These kinds of assertions about ourselves (the identity owner), known in the digital identity industry as claims, are difficult to trust because they are nearly impossible to verify.

In the physical world, we use the physical credentials in our wallet to prove our identity.

Each time we board an airplane, rent a car, reserve a hotel room, or take out a library book, we prove claims about ourselves simply by opening our wallet and showing one or more credentials containing claims issued by a trusted authority (called the issuer) to another human being or company who needs to trust the claim (called the verifier).

Why don’t we have an equivalent solution on the Internet? What is preventing it?

Where is the digital equivalent of a passport, driver’s license, or birth certificate that we can just “show” to a website to register, login, or verify our rights and privileges? Why do we instead have dozens or even hundreds of usernames and passwords to manage, and our personal details scattered across a multitude of databases guarded by companies who demonstrate almost daily they are incapable of keeping it safe?

---

1 This cartoon is also the inspiration for the logo of the Internet Identity Workshop, a twice-yearly event that just celebrated IIW #25.
The heart of the problem is that we have no standard way to verify digital credentials

A physical credential is relatively easy to verify: a human makes a judgment about a paper document.

Hotel clerks, car rental agents, librarians, and security guards all know the basic procedures for verifying a physical credential from your wallet, with varying degrees of certainty.

But this same process is not easy to duplicate online. To begin with, on the other end of an Internet connection you don’t have a human—you have a machine. And the credential you are showing them is not a physical document they can inspect, but a digital document.

To verify a digital credential, we need to solve two problems. First, we need to standardize the format.

Because a digital credential is read by a machine, it needs to be in a format that machines can understand. We’re already seeing this today with some paper credentials that must be verified all around the world, such as passports. Even though it is a physical document, a passport includes sections that are machine-readable in a standardized format.

Second, we need a standard way to verify the source and integrity of these digital credentials.

Digital signatures are already legally valid in most jurisdictions around the world. However they require two keys. The first key—the private key or signing key—is used to sign the document, and is kept secret by the issuer. The second key, called the public key or verification key—is used to verify the signature and ensure the document has not been tampered with, and it does not need to be kept secret. For universal adoption of digital credentials, we need a standard way to verify the public key of the issuer, which would then prove the authenticity of the credential.
The W3C Verifiable Claims Working Group was formed in April 2017.

Its mission is summarized in the [charter](https://verifiable-claims-wg.w3.org/):

*It is currently difficult to express banking account information, education qualifications, healthcare data, and other sorts of machine-readable personal information that has been verified by a 3rd party on the Web. These sorts of data are often referred to as verifiable claims. The mission of the Verifiable Claims Working Group is to make expressing, exchanging, and verifying claims easier and more secure on the Web.*

Verifiable claims are a standard way of defining, exchanging, and verifying digital credentials. The strength of the claim depends on the degree of trust the verifier has in the issuer. For example, if a bank issues a claim saying that you have a certain credit card number, a merchant can rely on the claim if the merchant has a high degree of trust in the bank.

**A worldwide standard format for digital credentials should have a far-ranging impact.**

Standardized network packets enabled the Internet. Standardized hypertext pages enabled the Web. Likewise, standardized digital credentials can enable a worldwide ecosystem of credential issuers, owners, and verifiers all exchanging interoperable verifiable claims as shown in this diagram.
But this leaves the second problem: standardizing how to verify the digital signatures of credential issuers

The usual answer has been public key infrastructure (PKI).

The premise of public key cryptography is that anyone can verify a digital signature from anyone else as long as you have access to their public key. The two keys are cryptographically linked so that every private key has only one public key and vice versa.

The core challenge is verifying that you have the correct public key for the issuer. For the past several decades the answer has been PKI.

PKI is what powers the green padlock in your browser.

The PKI used in modern browsers relies on a small number (a few hundred) certificate authorities (CAs) to be the roots of trust. The number is small so that your browser can easily manage them. The owner of a private key, such as a website, gives their public key to a CA who signs it with their own private key and issues a public key certificate. That’s what your browser is checking for each time you connect to a website that offers an encrypted HTTPS connection. This is how you know you’re dealing with the site you think you are.

The fundamental problem with PKI is that it is cumbersome, costly, and centralized.

Certificates from reputable CAs take real time and effort to obtain. Being a CA has been described as having a license to print money, because these centralized roots of trust are built into browsers and other software. This is why most digital certificates are purchased by companies, not individuals. They are just too hard for most people to deal with.

What’s worse, inserting a middleman into our digital trust infrastructure is a vulnerability. If a CA makes a mistake on a digital certificate, or if their service goes down or has a security lapse, or if they raise their prices, or if they go out of business—the whole system falls apart. It is centralization of this type that can lead to censorship and single points of failure.
PART TWO

The Solution
A public blockchain² is a decentralized root of trust that nobody owns, but everyone can use.

Blockchain technology turns the centralized root of trust model on its head. Rather than relying on CAs, consortia, or governments to be a cryptographic root of trust, it uses a consensus algorithm operating over many different machines and replicated by many different entities in a decentralized network. The Bitcoin network has proven this model by operating for eight years without a breach.³

Blockchains replace trust in humans with trust in mathematics.

Regardless of their specific design, all blockchains represent a cryptographic triple play:

1. Each transaction⁴ in the blockchain is digitally signed by the originator.
2. Each transaction—singly or in blocks—is chained to the prior via a digital hash.⁵
3. Validated transactions are replicated across all machines using a consensus algorithm.

The result is a cryptographic ledger of immutable records that makes it very difficult, if not almost impossible to change past transactions or maliciously control future ones.

So a blockchain is ideal to serve as a decentralized self-service registry for public keys.

Since every transaction in a blockchain has a digital signature that requires a private key, it is an obvious choice to use the blockchain itself for the storage of the associated public key—or any other cryptographic key over which the key owner needs to prove ownership. This is the core idea behind moving from centralized PKI to decentralized PKI (DPKI).⁶

² In this paper, the term “public blockchain” means a blockchain network available to anyone to use, just like the Internet is a public network available to anyone to use.
³ While the Bitcoin blockchain network itself has not been hacked, individual Bitcoin exchanges and multi-sig wallets have been breached. This is the cryptocurrency equivalent of a bank being robbed.
⁴ In blockchain technology, a “transaction” is any action that successfully writes a new record to the distributed blockchain database.
⁵ A digital hash is an electronic fingerprint of data that is globally unique and extremely difficult to forge.
⁶ Sovrin Foundation Trustee and Trust Framework Working Group Chair Drummond Reed is a co-author of the DPKI paper.
In fact, with blockchains, every public key can now have its own address

This address is called a decentralized identifier (DID)—another standard coming from the W3C.

DIDs provide a standard way for individuals and organizations to create permanent, globally unique, cryptographically verifiable identifiers entirely under the identity owner’s control. Unlike a domain name, IP address, or phone number, a DID is not rented from any service provider, and no one can take it away from whomever owns or controls the associated private key.

DIDs are the first globally unique verifiable identifiers that require no registration authority.

A DID is stored on a blockchain along with a DID document containing the public key for the DID, any other public credentials the identity owner wishes to disclose, and the network addresses for interaction. The identity owner controls the DID document by controlling the associated private key.⁷

Because DIDs are an open standard, any blockchain can create a DID method defining how DIDs can be registered (written) and resolved (read) on that blockchain. And because control over a DID is asserted entirely using cryptography—by digitally signing the transaction with the blockchain where the DID is registered—no central authority is needed to register the DID. Nor is any central authority needed to track or manage DIDs.

DIDs enable true self-sovereign identity—lifetime portable digital identity for any person, organization, or thing that can never be taken away.

DIDs are a sea change in digital identity. For the first time in history, an identity owner is no longer dependent on an external provider to gain the power of a permanent unique identifier that can be looked up on the Internet. Furthermore, given the right blockchain economics, DIDs can be cheap, so people can generate as many as they need to protect their privacy. Lastly, and most importantly, every person and organization with access to the Internet can have the means to prove their ownership of a public key, thereby enabling their claims to be verified.

⁷ Evernym, the original developer of Sovrin, is developing an open standard for managing these private keys called Decentralized Key Management System (DKMS).
With a public blockchain for DIDs, anyone can issue a digitally-signed credential, and anyone else can verify it.

We can finally use the same process to verify identity online that we have been using offline for centuries. Because every DID has an associated public-private key pair, anyone with a DID should be able to digitally issue and sign verifiable claims and other documents. So long as the verifier has the DID of the issuer (which in most cases is contained in the credential itself), it is a simple matter to look up the issuer’s public key on the blockchain and verify the signature on the claims. This is so straightforward that it should become the default behavior of any software that uses digital credentials.

Issuers and verifiers of digital credentials should no longer need to form identity federations. The DID specification ensures that issuers and verifiers everywhere can look up the necessary public keys on a public blockchain regardless of whether they belong to the same organization or identity federation. This progression—from disconnected “islands of identity” each with its own PKI to a global identity network based a decentralized PKI (DPKI)—is the same progression that occurred in moving from “islands of networking” (local area networks) into the global Internet.

Finally, we can move from reliance on centralized CAs to a more resilient, decentralized web of trust model.
PART THREE

Identity for All
To be truly universal, a blockchain for self-sovereign identity (SSI) must operate as a global public utility

If the system is going to work for all identity owners, issuers, and verifiers, it must operate like the Internet, the Web, or the Domain Name System (DNS).

The Internet, Web, and DNS are based on open protocols and open standards, run on open source software, and have open governance. As a result, no one owns them, everyone can use them, and anyone can improve them. An identity system based on a public blockchain needs to function the same way, so it can truly provide identity for all.

It cannot be based on proprietary technology, or be under the control of a single company or federation.

The irony of some proposed blockchain identity solutions is that they plan to use the unique properties of a blockchain—the ability of multiple competing parties to interact with the same universal source of truth—to provide a solution that will be under the control of a single company or consortium.

This fails the first test of true self-sovereign identity—that an identity is and will always be under the control of the identity owner, and that it can never be taken away. In fact the fundamental feature of all federated identity systems—that your digital identity is issued by an external “identity provider”—no longer applies.

The starting point must be a public blockchain.

With a public key on a public blockchain, every identity owner can now be their own self-sovereign identity provider. As Constellation Research analyst Stephen Wilson points out, this doesn’t mean that an identity owner controls every aspect of their identity. Rather, as Sovrin Foundation Chair Phil Windley has written, self-sovereign identity is powerful because it clearly delineates the boundaries within which the identity owner has complete control, and outside of which the owner must form relationships with others. For example, identity owners will need to work with other issuers of verifiable claims if they want to build trust in their identity.
Every public blockchain was designed for a specific purpose

**Bitcoin’s purpose is a global cryptocurrency.**

The blockchain industry exists because of the 2008 paper and code that the pseudonymous Satoshi Nakamoto published for a decentralized cryptocurrency network with a unique proof-of-work incentive model. The result speaks for itself: one of the largest distributed computing projects in the world, eight years of operation without a breach of the core blockchain, a community of thousands of developers and startups, and a cryptocurrency that is generating headlines nearly every day.

**Ethereum’s purpose is a global computer for smart contracts and decentralized apps (dapps).**

Ethereum has enjoyed similar success. Originally proposed in late 2013 by Vitalik Buterin to address Bitcoin’s lack of a scripting language, it has grown into a community of more than 30,000 developers,⁸ and spawned the Enterprise Ethereum Alliance, a global consortium creating a private permissioned version of the public Ethereum network.

**Now we need a public blockchain whose purpose is identity for all.**

As powerful as the Bitcoin and Ethereum networks are, providing identity for all was never their core purpose. They were not engineered from the ground up for the unique requirements of a global public utility exclusively for decentralized identity. But just as the Bitcoin and Ethereum networks are now frequently used with each other, both could interoperate with a new public blockchain designed for this purpose.

---

The Sovrin blockchain has been designed ONLY for identity

Sovrin is the first global public utility exclusively for self-sovereign identity and verifiable claims.

In 2015, a startup named Evernym recognized the potential for blockchain technology to solve the root-of-trust problem for self-sovereign identity. Evernym began designing a new blockchain called Sovrin to meet this need. However the deeper its developers delved into the problem, the more they recognized the solution would need to operate as universally as DNS.

They also recognized that they could not establish Sovrin alone. It had to be a community effort, just like every other core piece of Internet infrastructure. Evernym began engaging digital identity, security, and privacy experts from around the world to help with design and governance.

The Sovrin protocol is based entirely on open standards and open source—the Hyperledger Indy Project.

On September 29, 2016, the Sovrin Foundation was announced in London. It is now an international non-profit foundation with a board of twelve trustees plus a Technical Governance Board. In early 2017 the Sovrin Foundation transferred the open source code base—originally contributed by Evernym—to the Linux Foundation to become the Hyperledger Indy project. After a year of sandbox and alpha testing, the Sovrin Network was formally launched on July 31, 2017, with a genesis transaction between the first 10 participating organizations known as “stewards”.

Every facet of Sovrin architecture is designed to address the four major requirements of SSI: governance, scalability, accessibility, and privacy.

As the design of Sovrin progressed, Evernym and the other founders of the Sovrin Foundation realized that there were four overarching requirements to building a successful SSI system:

1. **Governance**: how the network can be trusted by all stakeholders.
2. **Performance**: how the network can provide self-sovereign identity at Internet scale.
3. **Accessibility**: how the network can ensure that identity is available to all.
4. **Privacy**: how the network can meet the strongest privacy standards in the world.
Sovrin currently uses a hybrid architecture that provides public access in a permissioned ledger. The overlying identity system doesn’t require permissioning, only the need for transactions to be cheap and fast. As ledger technology changes, these architectural choices could change as well.

The principles of self-sovereign identity transcend any particular type of blockchain or distributed ledger. There are a range of blockchain governance models from public permissionless Bitcoin, Ethereum, and IOTA, to private permissioned ledgers like R3 Corda and CU Ledger, and hybrid public permissioned blockchains like Sovrin. But SSI is not dependent on a particular type of blockchain distributed ledger technology (DLT)—it can work with any blockchain or DLT capable of meeting fundamental principles.

The first job of the Sovrin Foundation was to capture these principles in the Sovrin Trust Framework.

The Sovrin Trust Framework Working Group—composed of volunteer experts in digital identity, privacy, and policy from around the world—worked for eight months to develop the first trust framework that would provide the legal and policy foundation for a global public utility for SSI. Ratified by the Sovrin Foundation Board of Trustees on June 28, 2017, the first version of the Sovrin Trust Framework establishes 13 core principles of SSI and defines a first generation of business, legal, and technical policies for implementing them.³

Over 20 stewards have now signed on to operate under the Sovrin Trust Framework.

Any organization that wishes to run a node on the Sovrin public blockchain can qualify to become a steward by following the rules defined in the trust framework. The first 24 stewards span 11 countries and include:

- 11 Countries
- 8 Financial Institutions
- 8 NGOs
- 2 Law Firms
- 10 Self-Sovereign Identity Startups and Personal Data Networks
- 1 University
- 8 Credit Unions, Banks & Credit Card Networks

³ Sovrin currently uses a hybrid architecture that provides public access in a permissioned ledger. The overlying identity system doesn’t require permissioning, only the need for transactions to be cheap and fast. As ledger technology changes, these architectural choices could change as well.
The network must have the performance and scalability of DNS

A global public utility for self-sovereign identity should be used even more than DNS.

Although the DNS is capable of much more, the vast majority of DNS requests are simply to look up the IP address for a domain name. With over 1 billion host computers now available on the Internet, the DNS is serving over 100 billion lookups per day.

The same should be true of a global public utility for DIDs—only instead of looking up IP addresses from domain names, it will be looking up public keys from DIDs. If you imagine every person, organization, or thing needs a collection of DIDs—one for every relationship they have—then it is easy to imagine that there could be trillions of DIDs in a global decentralized identity system.

Sovrin is explicitly designed to achieve this scale.

DNS is simpler to scale (yet more susceptible to attack) because it does not use a consensus protocol to create an immutable blockchain. Yet all consensus protocols can only scale to a limited number of validator nodes. To overcome this hurdle, the Sovrin Network is designed to use two rings of nodes: a ring of validator nodes to accept write transactions, and a much larger ring of observer nodes running read-only copies of the blockchain to process read requests.

In addition, the Sovrin blockchain is engineered to be able to return a state proof with any response. This is a very lightweight cryptographic proof—capable of being processed on a smartphone—that the response is valid according to the current state of the ledger, which should prevent man-in-the-middle attacks on Sovrin queries.
The economics of the network must enable universal accessibility

According to the World Bank, one seventh of the world’s population has no legal identity today.

From the World Bank’s 2017 ID4D (Identity for Development) Counting the Uncounted census:

Imagine trying to open your first bank account, prove your eligibility for health insurance, or apply for university without an ID; quality of life and opportunities become severely restricted. An officially-recognized form of ID is the key enabler – critical not only for exercising a wide range of rights but also for accessing healthcare, education, finance, and other essential services. According to the World Bank Group’s latest estimates, this is problematic for an estimated 1.1 billion people around the globe.

A global public utility for self-sovereign identity must meet the identity needs of everyone.

As valuable as identity credentials are in the developed world—where they are so embedded in our everyday life that we often take them for granted—in many developing countries they can be the difference between life and death. Child kidnapping, sex slavery, forced labor, and other forms of human trafficking all depend on the ability to prevent a person from being accurately identified. Refugee systems around the world require identity solutions that must work in the harshest conditions—and of course they must work beyond the boundaries of a single nation.

Just like with the Internet, cost should not be a barrier to access.

The basic infrastructure of the Internet is not free, but costs have become low enough to make universal access feasible. A global public utility for SSI must have the same goal. Universal access can be achieved in several ways: by running a public permissioned ledger at cost; by designing a very low-cost permissionless ledger; or by subsidizing the cost of SSI on any ledger. One way or another, the framers of the Sovrin Trust Framework all agreed that the goal must be identity for all. The Sovrin Foundation has formed the Identity for All Council to help ensure that Sovrin as a global public utility is serving the needs of those who do not yet have a means of proving their identity.
PART FOUR

Privacy for All
Above all, a global public utility for SSI must meet the highest privacy standards in the world, including GDPR

Privacy is the third rail of identity—if you don’t build it into the very core of a global identity system, it could lead to great harm.

By definition, a global solution for digital identity must enable every person and organization to verify and safely share highly private information—banking records, tax records, health records. Protecting the privacy of such records is crucial—in some cases even a matter of life and death. So a global public utility for identity must be able to meet the most stringent privacy standards in the world, starting with the EU [General Data Protection Regulation](https://www.gdpr.eu) (GDPR).

The Sovrin Network implements Privacy by Design on a scale that has never been possible before.

This challenge is also an opportunity—the opportunity to apply the principles of Privacy by Design in an identity system that protects not just the citizens of one country, or the customers of one company, or the members of one social network, but every person and organization in the world who opts to use Sovrin as a global public utility.

This approach also has one major advantage: because these new privacy protections are not confined to only one country, company, or website, they have the potential to achieve a network effect at a scale not previously possible.

At the heart of Sovrin architecture are three fundamental examples of “privacy as the default setting.”

1. **Pseudonymity by default.** Sovrin supports pairwise-unique DIDs and public keys.
2. **Private agents by default.** To prevent correlation, no private data is stored on the ledger, even in encrypted form.
3. **Selective disclosure by default.** Sovrin verifiable claims use cryptographic zero-knowledge proofs so they can automatically support data minimization.

The following pages will examine each of these three innovations in more detail.
Every universal identifier is a major correlation risk—even your mobile phone number.

Government ID numbers, credit card numbers, and phone numbers are all examples of universal identifiers. Universal identifiers allow your activity to be tracked everywhere you use them. With today’s computer technology we don’t have to rely on this type of highly-correlatable identifier. They are a 20th century tool that is no longer appropriate for the digital age. It is time we stopped using these outdated methods and started using technology that is designed to protect privacy without sacrificing functionality.

The solution is pairwise-pseudonymous identifiers—a separate DID for every relationship.

Imagine that when you open a new account with an online merchant, instead of giving them a credit card number or phone number, you gave them a DID created just for them. They could still use this DID to contact you about your order, or to charge you a monthly subscription, but not for anything else. If the merchant suffered a breach and your DID were compromised in any way, you would just cancel it and give them a new one—without affecting any other relationship.

The extraordinary consequence of this shift is that a pairwise-pseudonymous DID is not worth stealing. Not only can the criminal not use it anywhere else, but the moment either you or the merchant detects a problem, you simply can change the DID. The giant data breaches we are experiencing today, like Equifax and Yahoo, would become a relic of the past.

A global public utility for SSI must support the scale and economics of pairwise-pseudonymous DIDs.

To make sure it can be the default for all relationships, a public blockchain for identity must be able to scale to the trillions of DIDs that will be needed (see page 17), and the cost of each pairwise pseudonymous DID should be as close to zero as possible while still ensuring the health and sustainability of the network (see page 18 and Part Six).
An immutable public blockchain never forgets.

So how can a global public utility for identity implement the EU GDPR’s famous right to be forgotten? The answer is to never put any private data on the ledger itself. Instead, put only pseudonymous identifiers (DIDs), pseudonymous public keys, and agent addresses (see below). This enables the exchange of any private data to happen entirely off-ledger.

And encryption has a limited lifetime.

If encrypted data is stored on a public ledger, eventually the encryption will be broken (for example, with quantum computing). However the more immediate risk is that the private keys for the encrypted data—or the raw inputs to hashed data—are stolen or accidentally compromised. Then this indelible, globally-replicated data record is revealed to all, forever.

The solution is a P2P network of distributed private agents working in parallel with the distributed ledger.

In Sovrin architecture, each DID has a corresponding private agent—with its own pseudonymous network address—from which the identity owner can exchange verifiable claims and any other data with another identity owner over an encrypted private channel.

As shown here, private agents can operate on edge devices (mobile phones, tablets, laptops, etc.), in the cloud, or both.
Selective disclosure lets identity owners control how much data is shared in a particular context.

The classic example is date of birth. When you show your driving license at a bar to prove you are old enough to drink, the bartender is able to see your entire date of birth. Not only is this more information than the he actually needs to know—it is also information frequently used in identity theft.

Selective disclosure uses a cryptographic technique known as a zero-knowledge proof (ZKP).

Suppose you had a digital copy of your driver’s license in the form of a verifiable claim. You could then use a mobile app to present a zero-knowledge proof that you are old enough to drink. The bartender could verify the proof using the public key of the issuer (similar to verifying a digital signature). But the bartender never learns (i.e., has “zero knowledge of”) your actual birth date.

A public blockchain for SSI makes it feasible to deploy ZKP as the default for all Sovrin verifiable claims.

Although IBM and Microsoft have been working on ZKP technology for over a decade, it has never been feasible to deploy at scale because the required infrastructure did not exist. It is like trying to travel by plane before there were any airports. Now, with the advent of Sovrin infrastructure, zero-knowledge proofs can become the standard for all interactions between all Sovrin identity owners.

---

Dr. Jan Camenisch, Principal Researcher at IBM’s Zurich Research Lab and a member of the Sovrin Technical Governance Board, is a pioneer in the application of cryptographic zero-knowledge proofs to digital credentials. He is a lead author of a seminal book in the space, Attribute-Based Credentials for Trust (ABC for Trust): Identity In the Information Society (Springer, 2015).
PART FIVE

The Impact
The hidden costs of our dysfunctional Internet identity infrastructure are staggering.

- The 2017 Hiscox Cyber Readiness Report estimates that cybercrime and data breaches currently cost the global economy US $450 billion per year.
- The U.S. Public Interest Research Group estimates consumers will have to directly shell out a collective US $4.1 billion to freeze their credit reports and prevent fraudsters from using personal information possibly exposed in the massive data breach at Equifax.¹¹
- IDG estimates that theft of trade secrets costs every nation from 1 to 3 percent of their gross domestic product (GDP), for a total ranging from US $749 billion to $2.2 trillion annually.

A global public utility for self-sovereign identity could mitigate these problems and reduce cybercrime.

As described in this paper, the shift to self-sovereign identity—to pseudonymous DIDs backed by private keys and verifiable claims—addresses the root problems with identity authentication and data protection that make cybercrime so prevalent today.

Sovrin could directly transform four major industries:

<table>
<thead>
<tr>
<th>Industry</th>
<th>Current Estimated Size (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identity and Access Management (IAM)</td>
<td>$8 billion (Orbis Research)</td>
</tr>
<tr>
<td>Cybersecurity</td>
<td>$86 billion (Gartner)</td>
</tr>
<tr>
<td>RegTech</td>
<td>$70 billion (Let’s Talk Payments)</td>
</tr>
<tr>
<td>Data Integration</td>
<td>$6 billion (Markets and Markets)</td>
</tr>
</tbody>
</table>

¹¹ To add insult to injury, U.S. PIRG spokesperson Mike Litt adds, “We’re not customers of the credit-reporting firms—we don’t get to choose them collecting and selling our information—and, in the case of Equifax, losing it—and now we have to pay a fee to protect it?”
The faster our digital economy grows, the faster enterprises need to move “beyond the password.”

An April article 2017 in the Wall Street Journal put it this way:

A recent survey of U.S. companies found that each employee loses an average of $420 annually grappling with passwords. With 37 percent of those surveyed resetting their passwords more than 50 times per year, the losses in productivity can be large. Factor in the cost of the support staff and help desks required, and the financial burden becomes even bigger.

Sovrin goes one step further: it not only eliminates usernames and passwords in favor of cryptographic authentication, but it adds the ability to exchange verifiable digital credentials for stronger, more flexible, and more resilient identity verification and access control.

Consumer IAM is also growing fast, spurred by the widespread adoption of mobile and Internet of Things.

Allied Market Research reports that this is driven by smartphone adoption:

Growing popularity of mobile devices and flexible functionalities of consumer IAM solutions to tackle increased network traffic, burgeoning demand, and peak usage requirements of consumers for different applications are expected to provide numerous opportunities for the growth and development of the global consumer identity and access management market.

The biggest impact on IAM will be the shift from proprietary and federated identity solutions to SSI.

The best analogy for this shift is the transformation of network infrastructure in the 1990s—from proprietary networks to the Internet—and of content publishing infrastructure in the 2000s—from proprietary platforms to the Web. Sovrin is the Internet for identity—it can standardize and automate how the vast majority of identity management functions are handled not just for consumers but within the enterprise as well. Just as the Internet standardized global communications, Sovrin’s goal is to standardize global identity.
Cybersecurity is arguably the largest single issue on the Internet today.

“No locale, no industry or organization is bulletproof when it comes to the compromise of data,” according to Verizon’s 2016 Data Breach Investigations Report. IBM President and CEO Ginni Rometty described cybercrime as “the greatest threat to every profession, every industry, every company in the world.”

Too much of our defenses are consumed trying to make up for the Internet’s lack of a secure foundation.

As Kim Cameron, Chief Architect of Identity at Microsoft, famously said in the preface to his Seven Laws of Identity, “The Internet was built without an identity layer.” In other words, we are backfilling for identity and security on a network where in the early days “everybody knew everybody” and trust was so implicit that routing tables for IP addresses were coded by hand. By contrast the ten current Internet Engineering Task Force (IETF) standards for DNS security alone run to over 500 pages.

Centralized PKI has not solved the problem—only a decentralized PKI (DPKI) can give us the strong cybersecurity foundation we need.

In the 2015 paper describing DPKI, whose co-authors include Vitalik Buterin (Ethereum), Christopher Allen (SSL 1.0), and Jon Callas (PGP and Silent Circle), the fundamental advance of blockchain-based public key management was summarized this way:

Control over the identifier is returned to the principal [owner]. No longer is it trivial for any one entity to undermine the security of the entire system or to compromise an identifier that is not theirs. This is how DPKI is able to address both the security and usability problems that plague DNS and PKI.
The second largest industry to be impacted is RegTech

RegTech includes is products and services that help companies comply with regulations of all kinds, including identity and security.

As outlined in a Let’s Talk Payments report, companies spend a lot of money on regulatory, compliance and governance software. Close to 55% of this spending is in consulting and business services. A major driver is the new EU GDPR requirements, where the penalty for non-compliance can be 4% of an enterprise’s worldwide pre-tax revenues.

Know Your Customer (KYC) and Anti-Money Laundering (AML) compliance alone costs every financial institution an average of $60M annually.

A 2016 Thomson Reuters survey reports that, “While financial firms’ average costs to meet their obligations are $60 million, some are spending up to $500 million on compliance with KYC.” This explains why the strongest demand for pilots of SSI technology are from banks and credit unions seeking ways to control their KYC and AML costs—and why these organizations are among the first stewards of the Sovrin Network.

Sovrin can be a double-win for RegTech: lower compliance costs for businesses and substantially better experience for customers.

As an example, for KYC and AML, a financial institution will simply need to ask the customer for the set of verifiable claims it needs to comply with regulations. With a few clicks the customer can both provide the requested credentials and give consent, and within a few seconds the financial institution can both verify the credentials and write a private audit stamp to the blockchain.

It’s an enormous win-win for the entire industry: customers are onboarded in seconds instead of days, and financial institutions and regulators have a cryptographically-verifiable and tamperproof KYC/AML audit trail.
Finally, Sovrin could change the course of the Data Integration industry

Data Integration is how computer systems are plumbed together to connect business processes.

Analysts at Markets and Markets expect the industry to double in size by 2022. According to their report, “The major factor driving this market is the high demand for tools that can combine numerous heterogeneous data sources, enabling users to get a consolidated view of data and extract valuable business insights. The rise in adoption of cloud computing...is another key driver fueling the growth of the data integration market.”

One of the hardest problems in data integration is how to establish trusted connections between different systems.

The challenge is not how to “hook the pipes” together—it’s how to authorize data to flow between the different systems, and how to reliably and securely encrypt and decrypt the data. For some integration projects, these problems represent the majority of the costs and time involved—or worse, the barriers that prevent the project from succeeding altogether.

For these problems, Sovrin verifiable claims could be a game changer.

Many systems integrate with APIs in order to exchange the same kind of information that is inside verifiable claims. With verifiable claims, data is conveyed by the person or organization presenting the claim rather than through a pre-programmed API integration. Your business can use the information in a verifiable claim without any API integration, without a contract, or even a without pre-existing business relationship. This can reduce the overhead required for business integration and make data sharing much more flexible—or even ad hoc.
PART SIX

The Token
Realizing these economic benefits requires a new means of exchanging value for digital credentials

Every exchange of verifiable claims represents a reduction of risk in a digital trust relationship.

Each time you present physical credentials from your wallet to a verifier to prove your identity, you are enabling the verifier to lower their risk in a transaction with you—whether boarding a plane, renting a car, occupying a hotel room, or borrowing a book. This reduction of risk has real, measurable value to the verifier, because it lowers their cost of doing business.

It also has real value to you, the identity owner, because it removes friction and can even lower the price of a product (e.g., “Good Driver discount”). So both the presenter of the credential and the verifier accepting the credential are receiving real value from the issuer of the credential.

The greater the risk being taken by the verifier, the greater the value of the verifiable claims.

At the low end, even a claim that just helps a website verify that a login is not from a bot has value. At the high end, a credit report or a background check may mitigate thousands of dollars of risk. In short, all verifiable claims transfer some level of value to owners and verifiers.
This antiquated approach has three major drawbacks. First, it restricts the market to only the highest-value credentials.

If every verifier has to set up a payment account with every issuer from which they need to purchase credentials, the inherent friction in this process means only credentials critical to the operation of a high-risk business—credit reports, background checks, insurance verifications—will have a market. We miss the vast potential of a long-tail market for lower value credentials, such as address verifications, social network memberships, peer endorsements, etc.

Second, it favors the largest providers who are the biggest targets for data breaches (e.g., Equifax).

The more friction in a market, the more economies of scale benefit large providers. In digital identity verification, this has led to a small oligarchy of companies with enormous market power—and equally enormous honeypots of consumer data. This is precisely why we can have catastrophic data breaches such as Equifax losing the personal details of 148 million people or Yahoo admitting 3 billion of its accounts were hacked.

Third, it prevents the use of powerful new privacy-preservation technologies like selective disclosure that could further protect personal data.

If value exchange for digital credentials is confined to conventional payment networks, there is a much deeper hidden cost: it means all the data being shared is flowing directly between third parties with strong correlation. Not only is the data subject (you) not “in the loop,” but there is no way to apply selective disclosure (page 23) to enable such exchanges to be more protective of personal privacy.

By turning you into an identity owner and putting you (and your Sovrin agent) in the loop, you not only regain this privacy control, but the data and claims being shared actually increase in value because they are: a) authoritative, b) permissioned, and c) current.
The Sovrin token addresses all three problems by providing a built-in incentive for the privacy-preserving value exchange of digital credentials. By enabling digital value transfer to take place directly in-line with the exchange of verifiable claims—and by incorporating the same privacy-preserving zero-knowledge proof technology (page 23)—the Sovrin token is designed to turn the Sovrin protocol into a digital marketplace for trust.

First, the marketplace can expand to encompass credentials at all levels of value.

With a digital token and protocol that can efficiently transfer any amount of value—down to fractions of a cent—suddenly there is a market for credentials of all kinds, even weak reputation signals (account age, community ratings, patronage). Anything that can become a measure of trust can now be exchanged for a token of value.

Second, credential issuers of all types and sizes can enter the marketplace.

Just as credit cards enabled millions of small merchants to go into business offering telephonic and online ordering, the Sovrin token is designed to enable thousands of new “merchants of trust” to begin offering digital credentials to meet any market need for trust verification: microcredit, tenancy, employment qualification, online recommendations, news verification—the list is endless.

Third, this new marketplace can support GDPR-compliant privacy preservation.

As described in Part Four, many digital options for trust verification run afoul of the third rail of privacy. Sharing credentials without the identity owner’s consent violates legal regulations and the owner’s trust. But with a Sovrin token linked to verifiable claims exchange, realizing the value of shared credentials can be done safely and with the owner’s consent. Everybody wins.
The Internet has flattened and globalized many markets. But not digital credentials...yet.

Retailing, Auctions. Classifieds. Software. All of these have been remade by the Internet and the Web, removing middlemen and expanding reach. However, the payment and privacy hurdles described above have prevented that from happening with digital credentials. Until now.

A global public utility for SSI can unlock this market and create a virtuous cycle of issuers competing on credential quality and cost.

With the Sovrin token and the Sovrin protocol, the value of verifiable claims can now flow either from verifiers to issuers—or indirectly from verifiers to owners to issuers—as often as every single time a claim is exchanged.

The first case uses the Sovrin zero-knowledge payment protocol so the issuer gains no knowledge of either who is using a credential nor where it is being used—only that the issuer is being paid the asking price in Sovrin tokens. In the second case verifiers can pay for credentials directly from owners, and owners can do the same with issuers.

The result is a marketplace where any source of trust—from a government to a family—now has an incentive to realize value from helping build trust with others. And those who have earned that trust—the identity owners—can now benefit from the ability to transfer that trust to other relationships. Verifiers, for their part, can now take advantage of a vastly expanded marketplace for trust information—a marketplace in which issuers are constantly competing to fill any “trust gaps.”
With any modern smartphone, your mobile carrier could produce a verifiable claim about your current location.

Smartphones include GPS capabilities, so your phone could easily share a verifiable claim about your location. However verifiers are unlikely to accept that claim because the issuer is your phone. They have no idea if they can trust it.

But your mobile carrier also knows the location of your phone—in fact they need that information to deliver your mobile service. So they can issue a verifiable claim about your location, and it would be backed by their full infrastructure and reputation.

This information, shared privately and only with your permission, can have real value to verifiers.

Your location is sensitive personal data whose privacy is in many cases protected by law. However, with self-sovereign identity, you would have the power to privately share a verifiable claim of your location from your mobile carrier with any verifier of your choosing. And unlike the GPS location data available to apps, which is easy to fake, this claim would be quite credible.

How valuable is proof to your bank that you are in your home or office when requesting a high-value transaction? How valuable is proof to an online merchant that you are ordering a new computer from your verified home or office address? How valuable is proof to a parent that a teenager out for the evening can verify their current location?

You get convenience, the verifier gets assurance, the carrier gets new revenue—everybody wins.

Even if the value is only a few cents per claim, one-click location verification can add up to a new revenue source for carriers while providing a great new service for subscribers.

To prove the location of the phone is also the location of the owner, the claim can include a verification of one or more of the owner’s biometrics.
The Sovrin token can also enable a market for digital credential insurance

Digital credentials are a tool for managing risk. In any such market, another standard tool is insurance.

The risk of misidentification in an identity transaction is never zero. This isn’t a problem of technology, it’s a problem of people. People make mistakes. People use fake IDs. People mix up records. In a seminal talk on this topic, original Ethereum team member Vinay Gupta said:

*Systems with very low defect rates, but with real consequences, need insurance. Identity risk can become an insurable risk...When you get closer and closer and closer to perfect systems, the need for insurance rises rather than falls, because nobody ever plans for them to go wrong. The Internet’s identity system is going to be an insurance system.*

Digital credentials paid for with the Sovrin token could also be insured with the Sovrin token.

Issuers could purchase insurance directly from insurers and build it into the price of their claims. Or verifiers could independently purchase insurance on claims from issuers that an insurer is willing to underwrite. In either case, since insurers would themselves be issuers and verifiers of claims, the Sovrin token is designed to establish a common, low-friction medium of exchange for all participants in the SSI marketplace. This low friction should attract more insurers and make insurance affordable to a much wider spectrum of participants.

Insurers have the potential to become de facto reputation providers for credential issuers, resulting in an efficient marketplace.

Insurers are very good at risk assessment because it is their business—they have a strong financial incentive to be accurate. Their assessments about issuers, reflected in their pricing, can become a reputation signal other verifiers can rely on, further incenting issuers to do a good job.
In 2011 it was estimated that “degree mills” were generating over USD $300M annually in fake diplomas.

In their book *Degree Mills: The Billion-Dollar Industry That Has Sold Over a Million Fake Diplomas*, Alan Ezell and John Bear estimate that more fake Ph.D.s (50,000) are purchased in the United States every year than real Ph.D.s are given out (45,000). A 2015 investigation by Motherboard magazine found that “LinkedIn was filled with fake degrees” from these scam operations. Every fake degree erodes the value of real degrees from real universities.

Since universities are bastions of trust; for a real one, credential insurance should be very inexpensive.

Universities go through extensive accreditation processes, so it should be easy for an insurer to verify the authenticity of a real one. And a real university has a strong incentive to ensure the authenticity of the academic credentials it issues—it is core to its reputation as an educational institution. As a result, the cost of digital credential insurance should be very reasonable and the value very high.

A small number of insurers can provide coverage—and reputation—for universities around the world.

This is a perfect example of a virtuous trust cycle: as insurance companies enter the business of providing credential insurance for universities, verifiers—such as employers around the world—will no longer need to verify the authenticity of every educational institution. They can trust the insurer, who in turn is paid in Sovrin tokens by the universities for this service. Now each verifiable claim of a diploma could be accompanied by the insurer’s verifiable claim. Employers get the assurance they need; alumni get the verification they want; universities protect their reputations; and insurers earn revenue.
In addition, the Sovrin token can unlock an *ethical* market for customer data

Today, our leaked personal data is aggregated and sold by third-party data brokers without our consent.

Without any direct permission from consumers, the big three credit reporting agencies (Equifax, Experian, and TransUnion) **collect 4.5 billion pieces of data each month** to feed into their credit reports, then turn that into **over US $3B in revenue in 2016.**

With the Sovrin token, companies can offer customers a direct incentive to share data with consent.

This disintermediation of data brokers has three major advantages:

1. Since the data comes directly from the customer, it is fully permissioned.
2. The data is fresher and more valuable than third-party data or inferences.
3. The reward for the data, in the form of Sovrin tokens, would go directly to the customer instead of a third party, building good will, trust, and loyalty with the party who really matters.

Direct sharing can simultaneously increase efficiency, build trust, and eliminate middlemen.

Empowering customers to do direct, permissioned, verifiable data sharing with the businesses they trust is turning into a worldwide movement. In the words of the **MyData Declaration**:

*By letting individuals control what happens to their data, we intend to create a truly free flow of data – freely decided by individuals, free from global choke points – and to create balance, fairness, diversity and competition in the digital economy.*
For example, Sovrin verifiable claims can finally solve the multi-billion dollar change-of-address problem

After nearly 30 years, there is still no simple way to do an automated change-of-address over the Internet.

Given all the magnificent Internet innovations over the past three decades—online auctions, one-click ordering, real-time driving directions—it is hard to believe we don’t have a simple way to send a change-of-address (COA) to everyone who needs to be notified. Just in the U.S. alone, over 40 million people move every year. At an average cost of US $6.00, for operational processing costs alone, change-of-address notifications cost businesses billions of dollars per year.

A fully permissioned verifiable claim from a trusted issuer can finally meet all the necessary requirements.

Of all the reasons it is hard for a business to accept a COA over the Internet, the most difficult one is trust. COA notifications are a classic attack vector for identity theft. So businesses need verified COAs from issuers they trust, like national postal services. Self-sovereign identity and verifiable claims infrastructure can finally provide this solution.

A verifiable digital change-of-address saves so much cost that a verifier could afford to pay BOTH the issuer and owner.

In addition to the US $6.00 average processing cost, a business can also save on shipping errors and fraud prevention. This means the business could pay issuers for the verifiable COA claim, plus it can incent identity owners to send this new type of COA notification by rewarding them in Sovrin tokens as well. Payment to the COA issuer would be blinded so the issuer does not know about the owner’s relationships, while payment to the identity owner can be directly in Sovrin tokens—which also helps build loyalty and repeat business.
In conclusion, the ultimate value of a “token for trust” is an Internet we can all believe in.

The Bitcoin and Ethereum public blockchains have proven that decentralization works.

Together they have already had the biggest impact on the Internet since the evolution of the Web. They have demonstrated that a global cryptocurrency can begin to change the nature of money, and that a global computer for smart contracts can begin to change the nature of government, business, finance, and law.

Now Sovrin can prove that decentralization works for self-sovereign identity and verifiable claims.

The open standards on which Sovrin is based—DIDs and verifiable claims—represent a third way for us to tap the unique cryptographic trust properties of a public blockchain. This time the goal is to give us interoperable digital credentials that can do for trust between any two peers on the Internet what packets did for communications between any two peers on the Internet.

Like the Bitcoin and Ethereum tokens, the Sovrin token is an intrinsic component of this new network.

Designed to enable value exchange with the same privacy-preserving properties of the Sovrin protocol for verifiable claims exchange, the Sovrin token aligns the incentives of issuers, identity owners, and verifiers everywhere to build value as they build trust. It is an essential element of a decentralized identity network and all the security and privacy benefits that spring from it.

Most importantly, this new infrastructure can help restore our trust in the network we all rely on to power our global economy and connect our global society.

We invite you to join us in making it a reality.
Next Steps

To become involved directly with the Sovrin Foundation, please consider joining the Sovrin Alliance, the worldwide group of individuals and organizations supporting the development of the Sovrin network. For other information about the Sovrin Foundation, including how your organization can become a Sovrin Steward or Trust Anchor, please visit us at https://sovrin.org. The following additional documents are also available in the Sovrin Library.

- The Inevitable Rise of Self-Sovereign Identity for more about background of the SSI industry.
- The Technical Foundations of Sovrin for more about the technology underlying the Sovrin ledger and Sovrin distributed agent architecture.
- The Sovrin Glossary for a complete guide to the terminology of self-sovereign identity and the Sovrin network.

Acknowledgements

The self-sovereign identity community stands on the shoulders of giants. The authors are grateful to numerous pioneers in digital identity for their contributions to the ideas in this paper. We apologize in advance to any we have forgotten on this list—rest assured the slight was not intentional.

Some of the core ideas behind SSI are over 20 years old. The January/February 1997 issue of Harvard Business Review included an article by John Hagel III and Jeffrey F. Rayport entitled The Coming Battle for Customer Information whose discussion of infomediaries presages the concepts of Sovrin distributed agents and agency providers:

Infomediaries operate on the assumption that personal information is the property of the individual described, not necessarily the property of the one who gathers it. The infomediary business model recognizes that there is value in this personal data and the infomediary seeks to act as a trusted agent, providing the opportunity and means for clients to monetize and profit from their own information profiles.

Andre Durand, Phil Becker, and Eric Norlin started and ran the Digital ID World conference and magazine where many of the earliest ideas about individual online identity were discussed and many of its later players came to know one another. Steve Gillmor hosted an “Identity Gang” podcast where the group was expanded and the impetus for action was born.

Internet Identity Workshop (IIW) grew out of a mailing list started after that conversation. Founded in 2005 by Doc Searls, Kaliya Young (@IdentityWoman), and Phil Windley, and held twice yearly with a unique open space ethos created and curated by Kaliya and Heidi Nobantu Saul, IIW has nurtured numerous identity standards and protocols including OpenID, OAuth, UMA, and OpenID Connect. It continues to be a center of innovation around self-sovereign identity.
Doc Searls’ **Vendor Relationship Management (VRM)** mailing list fostered numerous conversations about individual control and ownership of personal data and identity. It was the first place we saw the term “self-sovereign identity” thanks to Devon Loffretto.

Kim Cameron’s **Seven Laws of Identity** proposed ideas that were radical at the time and still point to the need for innovative thinking about digital identity.

Numerous groups have been formed over the past decade to promote user-centric identity ideals including Project Higgins, OpenID Foundation, Information Card Foundation, and Open Identity Exchange, Personal Data Ecosystem Consortium, and People Centered Internet. These organizations have pushed the envelope in what is possible when we put people at the center of their own digital identity and data flows.

Christopher Allen authored an **early paper on self-sovereign identity** that laid out important principles that distinguish SSI from earlier ideas about user-centric identity. Christopher’s **Rebooting Web of Trust** workshop series has helped foster important developments in the space such the **Decentralized Identifier** specification and **Decentralized Public Key Infrastructure**.

This work has been supported by the Sovrin Foundation, its **Board of Trustees and Technical Governance Board**, and its **Founding Stewards**. It has also benefited from participation in the **Linux Foundation Hyperledger** initiative. Evernym, Inc. has provided substantial support for the specifications and code behind this work.

For the white paper itself, we would like to acknowledge the contributions of numerous individuals including Aron van Ammers, Judd Bagley, Sean Bohan, Jamie Burke, Joe Cutler, Eden Dhaliwal, Misty Bledsoe, Jacob Farber, Brenda Francis, Steve Fulling, Nathan George, Kendra Haar, Dax Hansen, Charlyn Ho, Steve Havas, Nick Johnson, Jason Law, Daniel Hardman, Lawrence Lundy, James Monaghan, Greg Murphy, Antti Jogi Poikola, Alex Preukschat, Timothy Ruff, Joyce Searls, and Andy Tobin.

A special thanks to Monique Heileson for her fine work on graphic design, layout, and illustration.

Many more helped with proofreading and finding mistakes but are in no way responsible for the remaining errors.

—**Phil Windley and Drummond Reed, Managing Editors**

©2018 Sovrin Foundation. This is a living public document published by the Sovrin Foundation under a **Creative Commons Attribution 4.0 International License** at the following link: **https://sovrin.org/library/sovrin-protocol-and-token-white-paper**.