Identity 101

A primer to Identity Management

This document provides an understanding of the meaning of Identity. An in depth explanation of the constituents of identity and how identity works in the real world is provided. This is used to establish how current identity systems are designed. Subsequently, their shortcomings are discussed. Finally, we look into Self Sovereign Identity Systems being developed using blockchains that present an exciting opportunity into a global identity management system. Along the lines of which we also look into Sovrin, the leading provider in this space, and present an analysis of their system.

Understanding Identity

What is Identity ?

In order understand how Identity Management Systems work, we need to gain an understanding of what exactly is Identity. Identity is a term that encompasses a wide variety of definitions and can never be limited to a defined set given its changing definition with its varied perception. However, more important than defining identity is defining the purpose that the identity carries. Any form of interaction and subsequent exchange of information between a given set of parties is based upon trust. Trust in the fact that a person or an entity is in reality whom they claim to be. This establishment of trust is built on what we call as an identity. Identity is that set of basic attributes that provides an individual, entity or an asset a legal representation. Identification is the prerequisite that entails access to any kind of service. A lack of identity can result in denial of access to basic services such as education, financial services, healthcare, social welfare benefits, economic development, and the right to vote. Yet, nearly 1.1 billion people around the world lack a formal proof of identity.

The burden of providing primary forms of identification rests on the shoulders of the government of the country to which the citizen belongs. The advancements made in the technology sector have brought a drastic change in how we create and process identity. Prior to this wave of digitisation , identity verification was limited to the authentication of physical documents presented by the owner. The introduction of the internet and subsequently increased digitisation has allowed for a gradual replacement of physical documents. Another significant change is the introduction of a person's online identity. The widespread use of the internet has led to the creation of multiple silos of user information owned and controlled by third parties. Given the importance of data and how it can be used for generating monetary benefits or even to harm people, a question ultimately arises, are the users really owners of their own identity?

At the most basic level of understanding, identity is a sum of attributes belonging to the entity in question. A comprehensive study by World Economic Forum concluded that an entity and its identity can be broadly categorized into three classes each, as shown in figure 1. An entity can be either be an individual, a legal entity or an asset and the attributes defining the entity can be broken down into inherent, accumulated and assigned attributes. We refer to a Third Party as being a separate entity, which can either be a single individual or a legal entity, for example, a government, organisation or an institute.

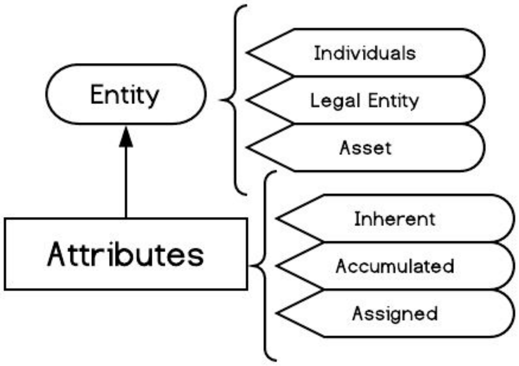

Fig. 1. Types of attributes and entities

Attributes of Identity

1. Inherent

The inherent set of attributes primarily consist of any set of information that is drawn directly from the entity in question. Some typical examples include name, date of birth, biometrics, gender or any other physical attributes. Inherent attributes being unique to the identity in question provide the strongest set of verification. As a result in most cases, they tend to serve as prerequisites to other sets of attributes. These inherent attributes cannot be duplicated by an entity to create pseudo-entities. Typically governments are engaged in the task of verifying and attesting these set of attributes for their citizens.

2. Assigned

These set of attributes are defined by third parties and are created with a specific purpose in mind. The most important point to note here is that these third parties are in control of the identities they create for the user as well as of the associated data. The primary purpose behind creating these is either to serve as a verification standard, for example, driver's license, smart cards or to provide access to the services offered by the third party, for example, creation of login ids on internet portals. Unlike inherent attributes, people may be allowed to create multiple sets of pseudo-identities based on assigned attributes depending on the use case.

3. Accumulated

These attributes are a function of time and hence change or modify with the progression of time. These basically include factors that are defined by the interaction of an entity with a third party, for example, a person’s credit history or education records. These accumulated attributes have always been a prime target

> for surveillance, and lately with monetisation driven by insights from data analysis, aggregating these particular set of attributes has become of prime importance.

Identity as a Claim

Analysing these attributes with respect to an entity and taking into consideration a number of use cases, we can conclude that,

> "An identity, as a sum of its attributes does not hold any value. The value of an Identity is a direct function of the perceived trust that the accepting entity has upon the issuer of an Identity."

Figure 2 provides a clear depiction of the Identity creation and distribution cycle. An attribute is always linked to a third party and derives its functioning from this relation. Even though some of the aspects of an Identity, such as the inherent attributes are derived from the entity itself, yet they provide no value by themselves. The only way an Identity is created and consumed in the real world is in the form of *claims. Claims* are a form of backing that an entity provides to a set of attributes of an entity. An identity can be composed of multiple claims. For example, the driver's license is a claim provided by the government, consisting of a mixture of inherent and assigned attributes and serving a direct purpose of verifying an individual's ability to drive.

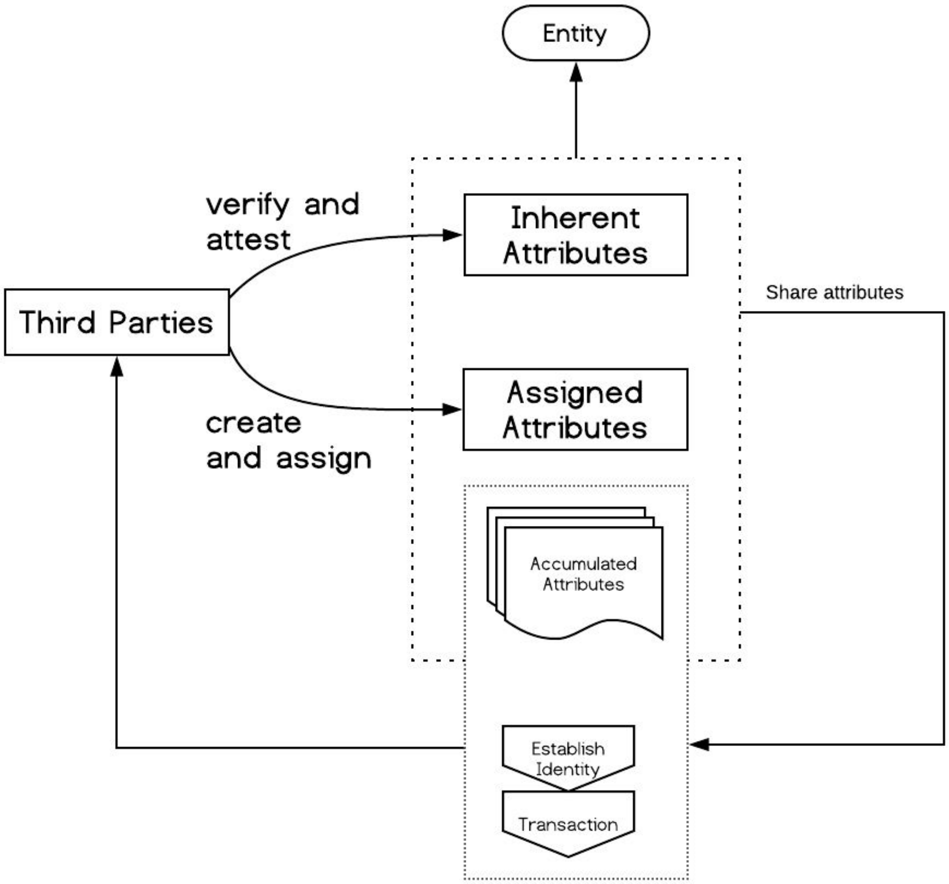

Fig.2. Representation of creation and distribution of attributes

To put it another way, an Identity is a form of trust, and this trust is provided by the third party that is attesting to the attributes of given the entity. A lack of trust in the attestor can virtually make the value of identity void. A secondary derivative of value is from the accumulated attributes of an identity with a third party, however, this is only a derivative and not a determinant of validity. The establishment of validity entails a derivation of value from accumulated attributes.

Digital Identity Systems

The massive scale of digitisation has allowed a majority of legacy identity systems to transition away from the physical methods of authentication towards digital ones. Although card-based systems are still prevalent around the world, they are either being complemented by other means of identification or themselves undergoing some forms of digital transformation. Instead of registering target populations manually and storing identity information in paper registers, electronic capture and storage of data provides a number of benefits, such as

1. *Unique Identifiers*: Digital Identity Systems provide significantly superior accuracy in capturing data. This allows for better deduplication procedures, hence ensuring removal of ghost enrollments and thus bringing down the associated losses due to waste and fraud.
2. *Increased Scalability*: Digital Identity Systems are easier to expand and accommodate to changing requirements. Moreover, they are much easier to integrate with other services and facilitate fast data processing and collection.
3. *Better Monitoring and Reporting*: The data records generated by such systems provide better accountability and monitoring of the services being dispatched by creating auditable transaction records. This prevents fraud and helps in aid development planning.

There are three components central to any digital identity system flow

1. *Enrollment* : enrollment introduces the user to an identity system. The enrollment procedure is dictated by requirements. Enrollment involves taking inputs in the form of inherent or assigned attributes, verifying them for their authenticity and producing unique identifiers with a namespace limited to the system. This is a one time procedure, though a user might be asked to regularly provide updates for their credentials depending on the level of enforcement for consistent information.
2. *Authentication and Authorisation* : once enrolled, a user is entitled to access of services provided by the third party. The service can range from an entitlement in the form of claim, like a drivers license, to something in exchange of monetary rewards, like a subscription service.

3. Accumulated Attributes : some services, specifically involving distribution of claims are limited to authentication and authorisation. The purpose of a driver's license is to just provide proof that the user is legally allowed to drive. Identity systems formed by most third parties allow them to store the data of the transactions and exchanges between them and users. These additional data points are termed as accumulated attributes. The accumulated attributed are specific to a relationship, which means that while inherent and assigned attributed can remain static between two or more third parties, accumulated attributes are unique to every relationship.

There are a vast number of identity systems being implemented presently. Each of these brings in a new set of unique features to the table, yet they can broadly be defined along a definite set of paradigms dictating their architectures. In this section, we categorize these systems based on their design philosophies and technological considerations. Any identity system is built with two criterias in consideration

1. *Purpose*

Purpose is one of the core ideologies that shapes an identity system, as it defines the why of building an identity system. Identity systems are always a means to an end. If the system is constructed to fulfill a singular specific purpose, then it is known as *Functional Oriented.* This narrow thinking makes the system highly limited in its functioning. Most of the third parties providing services to their users need a way to uniquely identify them. Hence the identity systems built by them are functional oriented, central to the services provided by them. Similarly, identity credentials issued for driving, voting, availing welfare schemes are all functional identities developed by the government with the aim of providing a specific service.

The information flow of a functional oriented system is represented in figure 3. The most important differentiating factor of a functional oriented system is that the whole digital identity system flow is managed by a single third party. This means that all databases, namely enrollment, authentication, authorisation and accumulation are owned and controlled by the same organisation. This architecture is flawed from the outset in a number of ways, yet the most widely used and adopted.

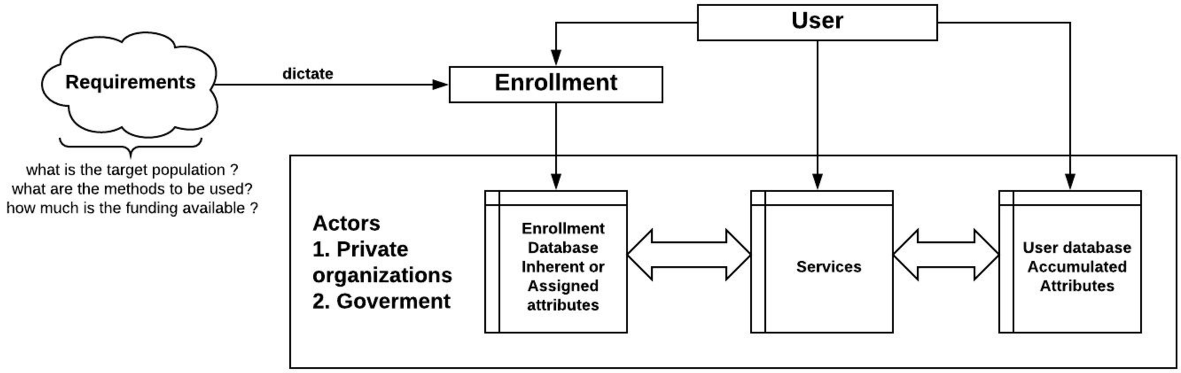
User
Requirements
dictate
Enrollment
what is the target population ?
what are the methods to be used?
how much is the funding available ?
Actors
1. Private organizations
2. Goverment
Enrollment Database Inherent or Assigned attributes
Services
User database Accumulated Attributes

Fig. 3. Representation of a Functional Identity System

On the other hand, systems that are designed with the primary objective of providing support to other systems, exhibiting the ability to expand and adapt, serving as the basis for a number of add-on services are known as *Foundational Purpose Systems*. The true purpose behind the construction of an identity systems should be to make them Foundational Purposed. This ensures that they are true to their task of acting as points for enrollment, authentication and authorisation. Any other functionality expected or asked for from an identity system should be considered as an add on service built on top of these Foundational Systems.

Figure 4 provides a representation of a Foundational Identity System. The key differentiator in the architectural designs of such systems in comparison to functional oriented systems is the distribution of tasks between different third parties. While some of them are engaged primarily in enrolling and maintaining user identities, others merely employ the generated identifiers for providing their specific services. However, the accumulated attributed are still owned and controlled by the organisation providing the service.

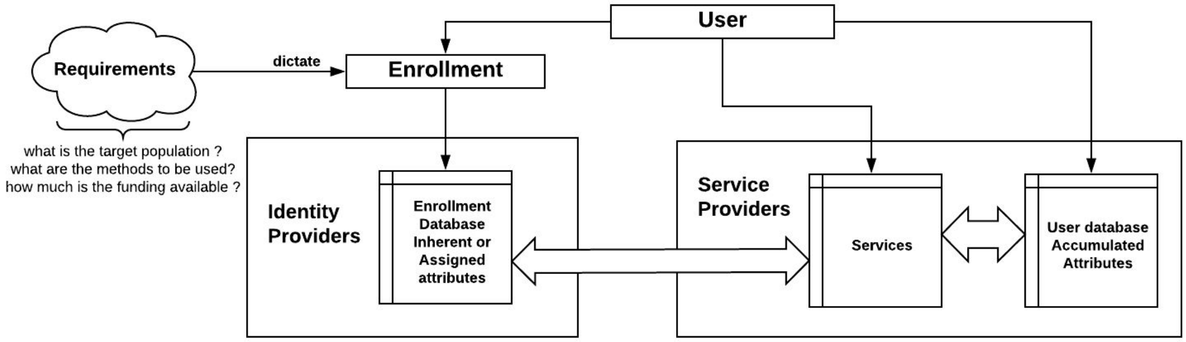
User
Requirements
dictate
Enrollment
what is the target population ?
what are the methods to be used?
how much is the funding available ?
Identity Providers
Enrollment Database Inherent or Assigned attributes
Service Providers
Services
User database Accumulated Attributes

Fig. 4. Representation of a Foundational Identity System

2. *Requirements*

Requirements state the properties and characteristics that the system must possess in order to fulfill the purpose it is being deployed for. Requirements are the key enablers in determining the selection of the technological aspects of the system. For example, whether the system should employ use of biometrics, what level of training will be required to operate the equipment, the associated costs required to deploy the system, what will be the target population etc.

Most of the systems built today function on a minimalistic basis, designed to fulfill their objectives and no more. Such systems are termed as *Instrumental Design Systems* and and are usually function oriented in their design philosophy. *Instrumental Design Systems* are extremely hard to repurpose for other purposes as use cases can differ vastly in their requirement. Technological constraints and lack of foresight of such systems makes them vulnerable to changes in expectations and severely undermines their longevity.

On the other hand, systems that are built keeping in mind future expansion and allowing for interoperability are termed as *Infrastructural Design Systems*. The design choices made for such systems are decided by factors affecting all participating parties and stakeholders. Considerations are kept in mind for the possible changes that might affect the system in the future. There is heavy emphasis on making the system extensible and allowing for maximum interoperability with the aim of allowing third party services to easily integrate with the system. The aim of the system is to increase longevity by providing long term support to third party services.

Key Issue with Identity Systems : Fragmentation

This is the key factor responsible for the operational inefficiencies, and in some cases for the early demise of current identity systems. Theoretically, an Identity is an identifier that belongs to a certain entity, hence all those attributes of the identity should be managed by the entity itself. However, in reality the legal entities engaged in identity management act as stewards of the information, therefore resulting in fragmentation of Identity. The reason why identity systems tend to result in fragmentation is because of flaws in their design philosophy.

Most of the Identity systems in use today are Functional Identity Systems. Functional systems having been designed with regards to fulfilling a particular objective. In addition, the way the system is structured dictates whether it is feasible to expand in the future. Systems designed on such philosophy result in broken identifiers separated by the barriers of the systems themselves. As each system is unique to its environment, it leads to a fragmentation of the identity of an individual which in principle should have been a unified identifier bounded to the entity itself. Functional Identity Systems are most widely used because they are easy to design, implement and control. Key problems arising due to fragmentation can be summarised as follows :

1. *Data Isolation:* identity data and associated accumulated attributes are centralized in repositories of different organisations. This, not only leads to duplication of data but also prevents linking and access of this siloed information which ultimately is tied to a single user.
2. *Recurring Efforts:* construction and maintenance of identity systems is a costly proposal, and has to be taken up whenever the situation demands. This is primarily because existing systems are hard to modify and adapt to changes, and have locked in functionality resulting in recurring efforts.
3. *Multiple Identities:* a customer is supposed to maintain multiple sets of identities, one for every set of services they avail. This means a typical person is a holder of multiple sets of Identities both online and offline. Fragmentation prevents single access points for management of these Identities and increases the workload of management of these identities on the user.

Functional Systems : A primary cause of Fragmentation

1. *Donor Organisations*

Humanitarian programs are driven by an agenda and attached funding. According to the previous secretary-general of the UN, Ban Ki-moon, corruption in 2011 prevented nearly 30% of all UN development assistance from fulfilling it's targeted purpose. Due to these reasons, organisations engaged in such programs are extremely particular about ensuring that records be established about where the funds are being spent. This involves

1. Identifying the target population
2. Tracking the distribution of aid

Given these reasons most of the donor organisations prefer to opt for their own systems, rather than repurposing previous ones, hence the choice of Functional Identity Systems. As programs build their databases, the cost of capturing the data, enrolling users and maintaining the system are recurring with each cycle and a major impediment to the functioning of aid distribution.

1. Malawi is currently being supported by various organisations such as Gavi, UNHCR, UNDP and UNICEF. However each organisation implements its own identity system, resulting in *Data Isolation* and *Recurring Efforts*. This makes it near impossible to coordinate and prevent overlap of distribution of services.
2. A nutrition program in South Asia being implemented by USAID made significant use of an identity system employing unique identifiers. However, a lack of coordination and design mistakes led to its early demise as it was not fit to be repurposed for other agenda's, leading to waste of resources and capital.

2. Private Organisations

Private organisations are concerned with fulfilling their own interests, and hence invest in building their own identity systems to provide services to users. Coordination in the industry is hard to come by due to competing interests, hence most organisations tend to prefer construction of their own identity system leading to *data isolation*. Each of these systems contributes to the fragmentation of identity in-spite of that the fact that most of them share a vast number of common attributes. Yet they repeat the same set of procedures for identity enrollment and authentication leading to *recurring efforts* both for the organisation and the user.

This results in an increased burden for the user, who has to repeatedly provide the documents for identification as well as be responsible for maintaining the associated identifiers. In addition, with the given boost in the field of data analytics, these companies have intensified their efforts of data aggregation in order to profile their customers and increase sales. In most cases, this aggregation is done without asking for the consent of the user.

3. National Identity Programs

Governments are typically the first movers in regards to technological advancements in identity management. Most of the recently launched government national identity programs have started to shift from functional to foundational identity systems. However, the deployment has been patchy and disperse due to lack of coordination and political inertia. A study by International Communication Union of the 48 identity systems being implemented by developing countries found that only nine countries had well integrated ID programs, wherein a single identification was being used to fulfill five or more functional categories. The employment of functional identity systems leads to a excessive wastage of human resources and capital. For example, while driver’s license and voters card fulfill two completely different functional aspects, yet they share a number of common inherent attributes. The repeated collection and storage of such identity information presents problems displaying all three aspects of fragmentation, that is *data isolation, recurring efforts* and *multiple identities*.

37 of those programs were found to involve some form of biometric identification highlighting the higher rates of adoption for digital identity. A key differentiator here is that 45 of those programs use some form of physical credential to authenticate an individual’s identity, while Yemen,India and BVN of Nigeria work without one. 28 of the programs had an electronic component on their physical credentials. Of all the systems under implementation, two of them stand out for their significant adoption of increased functional categories, namely *Aadhar* of India and *Nadra* of Pakistan. Both of them are being developed with a focus on providing identification services to third parties. Aadhar particularly is being used by the Indian government as a first step towards overcoming the fatigues of fragmentation in identity systems. The creation of Digital India Stack allows user credentials stored in aadhar to be used for authentication and authorisation by third parties through the API provided by aadhar. This also allows for ekyc, esign, payments services and digital signatures to be implemented. Moreover, aadhar does not carry with itself any physical identification, requiring an individual to verify using their biometrics. The current systems of Digital Identity Management are not very uniform in

their implementation and vary in their applications. In the following subsections we will gloss over some ways authentication takes place online.

Public Key Infrastructure System(PKIX)

X.509 Public Key Infrastructure System is the gold standard when it comes to verification of entities online. Communication on the internet is highly vulnerable to various types of security risks such as phishing, man in the middle attacks(MITM), snooping etc. Unlike physical ID systems there are a lot more factors at play on the internet. This is the reason why encrypted communication is extremely important for communications that take place over the internet. Encryption programs like PGP are the top of the line when it comes in providing complete security and privacy in the digital realm. However in order to establish encrypted communication online requires the authentication between entities regarding who they are and preventing masquerading attacks. This is where PKIX comes in.

The establishment of a secure connection requires the fulfillment of a prerequisite which is the identification of the server. This identification is done through intermediaries know as Certificate Authorities(CA). Any organisation that wishes to authenticate itself to users on the internet needs to sign up with a CA and obtain a digital certificate. CA’s are very thorough in their verification and background checks and require the fulfillment of certain requirements as well as fees from the client to sign up for a digital certificate. These certificates are revocable and have an expiration period. Therefore a server needs to produce its digital certificate to the client browser to authenticate itself. The verification of the signatures of the CA themselves is done through pre installed public keys stored in the browser by default. This vault of keys is secure and cannot be modified by any malicious user. The addition and removal of keys from this bootstrapped vault is done after rigorous checks and scrutiny. Currently there are more than 1000 CA engaged in the distribution of digital certificates. The exact mechanics of CA are further divided into three authorities as shown in figure x in order to facilitate ease of registration, access and make the system manageable and more secure.

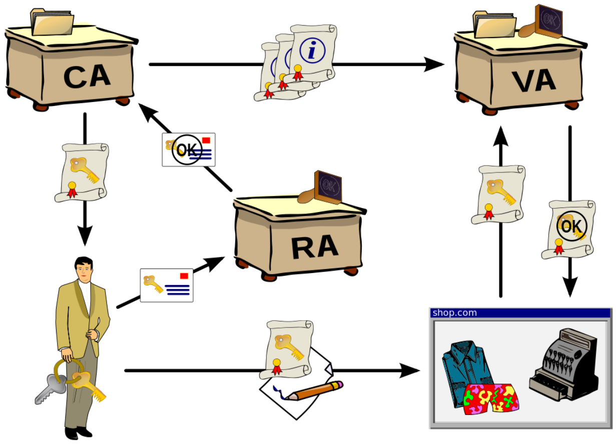

Fig 5. A typical workflow of a X.509 based PKI system

Despite the widespread acceptance, the systems is plagued by a large number of problems such as

1. **Being Overly Complex**

   This is no hidden secret that the current system is overly complex in its functioning and requires a lot of moving parts and coordination. The procedure of obtaining digital certificates, managing them, revoking them is not an easy task. This complexity is the very reason why the system fails to capture widespread adoption.

2. **Limited Use Case**

   The PKIX system was designed to authenticate websites online. To that end it fulfills its purpose. The problem is that extending it work for sharing Digital Credentials does not work. More importantly the system has a very high barrier of entry, which means that most of the websites actually work with unencrypted communication. The question of enrolling individual users on this network thus completely falls out the equation.

3. **Prone to attacks on CA**

   The final nail in the coffin is the presence of the CA themselves. These are for profit organisations. Any form of compromise to the server of a CA can undermine the integrity of digital certificates associated with it.

PKIX was never supposed to be used for sharing of Digital Credentials. There is no support for making and verifying claims. Moreover, its overly complex protocol and method of operation and extremely high barrier of entry restrict its use case to limited entities online, usually organisations that have an online presence. The natural question that arises is, how are users authentication online then ?

Username and Passwords

As we have discussed before, private organisations tend to construct their own functional identity systems thus leading to fragmentation. This problem becomes visibly pronounced in the modern digital world wherein the same user has to remember an ocean of usernames and passwords for authorising themselves to a given entity. The present solutions, rather than aiming to solve the problem from the root are coming with added alternatives to make life easier, for example the development of password management systems. Moreover, again none of this helps with the question of how will Credentials be shared online in a way that makes them both authentic as well as easy to interpret. There is no common standard to guide the progress.

Self Sovereign Identity

Identity systems are a necessary prerequisite for access to services in the modern world. However, as discussed in previous sections they suffer from challenges that limit their functionality. The key issue here is that most of the solutions are being developed to address the front end problems pertaining to enrollment and authentication, without any emphasis on how and where the data is stored, as well as functionality to bring in interoperability and consented control and aggregation of data. Any identity system that needs to tackle fragmentation will have to be built by the cooperation of all participating parties. More importantly, the task is to reduce the data of every user into a single mesh, such that it acts as a central point for every identity-related task and yet is immune to centralisations risks, cyber attacks or privacy problems. A solution is to build a self-sovereign identity system.

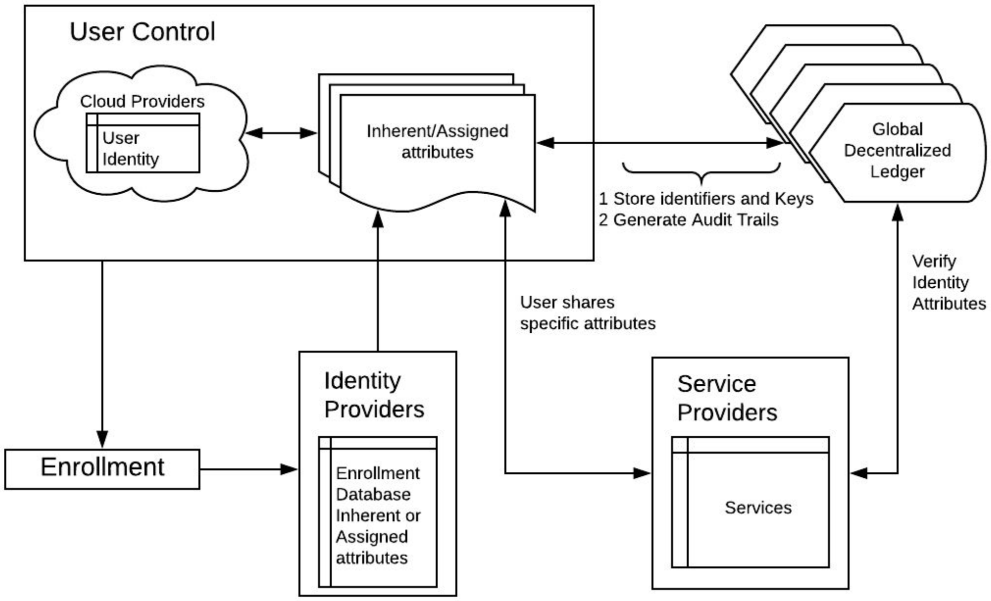
User Control
Cloud Providers
User Identity
Inherent/Assigned attributes
Global Decentralized Ledger
1 Store identifiers and Keys
2 Generate Audit Trails
Verify Identity Attributes
User shares specific attributes
Identity Providers
Enrollment Database Inherent or Assigned attributes
Enrollment
Service Providers
Services

Fig. 6. Self Sovereign Identity System

Architecture of a Self Sovereign Identity System

Figure 6 provides the blueprint for the functioning of a self-sovereign identity system. As opposed to the approach taken by the Foundational systems presented in figure 4, SSI systems are inherently designed around the user. Identity can be reduced to a set of claims. Whether a third party accepts a claim about an entity or not is completely dependent on the level of trust the third party has over the issuer. Taking this fact into

account, the backend of an self sovereign identity system can be made such that it fulfills generation and consumption of claims in a decentralised and secure manner.
The method of verifying and issuing will be kept completely separate from the system, that is the process an issuer entails in order to provide an entity a claim based on its reputation will in no way be related to the functioning of the self sovereign identity system. The main purpose of self sovereign identity system will be to ensure a

1. Standard for generating claims
2. Standard for verifying claims in a machine readable format

A typical flow of information involves issuing with regards to a particular user. A third party instead of producing a claim in the form of physical credential will generate a digitally verifiable form of the claim, similar to how digital certificates work. These certificates can then be produced by the entity to any third party and will be cryptographically verifiable. There are some points to be taken into consideration for this exchange to take place

1. Digital Certificate Verification : as per the present standards of modern cryptography, asymmetrically generated keys can be used to verify digital certificates. However, for this method to work, one needs to verify whether the keys themselves belong to an entity or not. In present implementations of Certificate Verification in use on the internet, a given set of Certificate Authorities are in charge of generating certificates certifying the keys. The root certificate of these authorities themselves, comes pre-installed in the browser hence preventing man in the middle attacks. However, if such a model were to be followed, it would fall victim to the issue of *centralisation*. Also given the stakes, it would be difficult to bring about cooperation among mutually distrusting parties.
2. Standard for Claims : in order to bring about interoperability and tackle all problems of fragmentation, the claims will have to generated in a manner such that they can be understood by any other third party.

A way to tackle this situation is to make use of a decentralised key management system. The solution is store the public keys and certificates in a decentralised ledger. This way any service provider can simply ask for a reference to the claim provided by a user on the ledger. This ensures the validity of the claim as all the data and keys on the ledger are considered to be the correct source of information. In order to generate unique identifiers and claims, Sovrin proposes an architecture involving use of Decentralized Identifiers. The proposal states use of a decentralised ledger in control of a number of organisations spanning governments and private organisations. The claims generated will be following the Decentralised Identifiers format, providing maximum flexibility for various types of claims.

Advantages of a Self Sovereign Identity system

1. ***Addressing Fragmentation***

Most of the current identity systems are siloed and hence result in fragmented identifiers. This arose primarily from the fact that the third parties themselves were the stewards of the identities of the user. SSI focuses on shifting this stewardship to the user themselves. Employing a global decentralised ledger allows for users to share their previous identity attributes to other third parties in a trustful and transparent way using claims. This means that instead of duplicating the same set of information, a user can merely provide specific sets from the same pool for authentication. The value of authentication here will be derived from the value the issuer has in the market. Moreover, most of the third parties can work on their services instead of building an identity database for each cycle.

2. Distributed Decentralised Ledger

The introduction of blockchain technology through bitcoin has opened the doors to a range of new features that were not possible before. The key takeaway is that a global distributed ledger maintained by governments and private organisations alike ensures that nobody is in control, yet everybody is a steward due to the consensus protocols making up the blockchain. These stewards are incentivised to perform honestly as identity theft can hurt their services. Moreover, generation of an illegitimate transaction would require a collusion by a majority of those stewards. In addition, a distributed ledger allows for the generation of audit trails of information, as every block is verified by all participating nodes, and the blockchain itself is an immutable ledger hence no changes or modifications are allowed once appended to the ledger. The advantages of using a using a decentralised ledger to store the identifiers can be listed as follows

1. Distributed decentralised ledger functions without a single authority in command. As a result of this, centralisation is prevented.
2. The ledger will be global in nature, public in access and permissioned in transactions.
3. A global and public ledger serves the purpose of maintaining the same state for every participant on the network, ensuring non repudiation and increasing trust.
4. Distribution among all participants provides resiliency against change of data by malicious actors.
5. Permissioned transactions by a select group of third parties ensure that only valid data is added to the ledger as per the rules set in stone by the source code of the ledger.
6. Privacy of the data is ensured by storing the information in encrypted format, ensuring that only the authorised parties can access the information.

3. Authenticating Identity

An issuer basically enrolls a user based on some pre-requisite checks and then provides the user with some form of attributes that substitute as authentication methods. However, every such attribute can be broken into two parts, the template and the data itself. While the template is constant, the data is dependant on the user. This is similar to providing trust certificates for websites, which allow you to differentiate between fake and real sites. Here the veracity of the certificate is provided by the trust a party has on the issuer. While the verification certificates and the templates are uploaded by the issuer on the ledger itself, the data still remains private and in control of the user. The concurrent maintenance of data on the global ledger by multiple parties along with timestamping ensures that the data was signed by the issuer, as well as that is current and up to date.

4. Secure Data Storage

The data generated in identity systems belongs to the user, but a user cannot manage all the data that they accumulate over their lifetime. With the advent of cloud services, immediate data retrieval at any location provides ease of access. Identity data storage should be like a rental service, where users can store all their personal information in encrypted form and these services providers can retrieve the data and connect to the global ledger in order to share it to a requesting party. The most important aspect is the need of multiple such providers to ensure that a user is not restricted to a single provider, thus preventing centralisation. In essence, the data of a user is secured in a locker, which can be shifted to any service provider, just like mobile number portability. Decentralised data storage provides a possible solution to this problem. Also, given the data identifiers and certificates are stored on the neutral global ledger, a user is never tied to a single provider.

5. Identity Monetisation

Currently, the data of users is aggregated and used to provide market insights. The introduction of a self-sovereign identity system means the user now plays a role in deciding who accesses their information. This also allows for the formation of a new type of data marketplace where users can sell their data to interested third parties, without revealing other sensitive information. For example, a user can provide a trail of their spending habits without revealing the other specifics of their transactions. Similarly, information exchange will become more accurate. For example, a market survey which requires users between certain age groups to fill a form can verify whether the person filling the form is actually of that age group or not, ensuring accuracy. This analysis is just a tip of the iceberg, SSI as a whole can result in fructification of a vast new set of use cases.

Sovrin and its Unique Problems

Sovrin, originally developed by Evernym presents the most exciting and developed solution in regards to Self Sovereign Identity. We shall not present how Sovrin works here. It is highly recommended to go through the official sovrin papers available on their website. However, in this section, we shall discuss our analysis of Sovrin.

Sovrin presents till date the most advanced and theoretically reliable approach in regards to a SSI system. It fulfills the crucial principles of design by privacy, provides every user the ability to decide what is shared and how it is shared. It also supports the ability to manage backups and claims in a truly decentralised manner without any form of coordination with a central entity. Sovrin, indeed is a perfect system, and therein lies its biggest fault. It is the perfect system for a perfect world. The principles and choices made by the system do not fall in line with the real world motives and scenarios. This makes it highly susceptible to failure due to lack of practicality. We highlight some of the reasons which make Sovrin impractical for real world application

1. **Everyone is Equal**

   Sovrin will support the construction and distribution of claims for any participant without any distinction or biasness. This presents a thorny problem for issuers. Even though we can reduce the formation of a claim to any entity, in the real world most of these entities are very limited in number. In other words, the requirement of trust severely limits the number of issuers on whom the verifiers rely. For any given citizen, the variety of issuers primarily ranges from government functionaries to educational, health and financial institutions. It is true, that many private organisations also engage in distribution of claims. Still the total number of registered businesses pales in comparison to the human population. More importantly, most of these business would have an extremely small footprint and do not engage in mass issuance of claims. Trust in issuers depends on a lot of factors, the most important of which is time and the backing/origination, which is why the number of issuers engaging in mass issuance of claims is limited.

   The crux of the matter is, the generation of most the claims is backed by a very limited number of issuers in comparison to the number of individual entities. Most of the claims and trust bootstrapping needs to revolves around these limited issuers. The failure by Sovrin to address this insight is its most significant shortcoming. This particular shortcoming in fact, rears its head further in the following two issues.

2. **Economics of Participation**

Issuers have their own particular reasons to operate with high efficiency and security when it comes to construction of claims, their verification and distribution. This is because they stand to lose the most if the system is compromised in any way. State governments engage in Identity Management because of necessity, whereas most private organisations work with Identity Management because of its prerequisite in business operations. For example, institutions like banks are very thorough in claims verifications and issuance due to their dealing with finances. Any error on their part can result in significant damage to both the bank as well as its customers. This is the reason why most of the issuers tend to form closed systems for Identity Management.

The functioning of an identity system is driven as a synergy between both the issuers and the users. If the issuers move en masse to a new system, the individuals would automatically also register. The vice versa however is very difficult to achieve, simply due to the law of incentives. Any entity would do something only if it has some incentive to take that particular course of action. Providing incentive to the limited number of issuers is far easier as compared to the vast number of individuals.

The next question which naturally arises is, why would these issuers move to a decentralised system ? The system indeed presents a lot of functionalities to the individual users, but does not make much sense for the issuers to move their systems. And if enough issuers fail to join the network, the lack of supported claims can make it nearly impossible to bootstrap the network.

3. **Scalability**

Assuming, even if most of these issuers would simply give up their interests and take the leap of faith and to move to a decentralised system, there is still the issue of scalability that needs to be addressed. As mentioned before, the whole system runs on the same blockchain. Sovrin explicitly states that none of the data is stored on the chain itself as well as private DID connections and claim exchange will take place off chain. Still, considering the fact that the numerous claim definitions, transactions involving changes in authorisation policies take place on the chain and that there is no way of pruning the chain, the ability to scale the chain becomes questionable. Moreover, there is no limit on how many DID's every entity can construct thus making it very hard to judge the abuse a malicious entity can do to the system by bombarding with junk requests leading to Denial Of Service Attacks.

4. **Centralisation, "A necessary Evil"**

> Even though Sovrin provides the perfect recipe for a possible SSI system. The truth is, that centralisation, despite its inherent flaws due to human exploitation from malicious internal actors, will continue to dominate and fledge in its application. The simple reason behind this that people have an inherent trust to rely and fall back upon known entities. It is the primary reason that corporations invest heavily on building brand value. People not only associate themselves with the organisations, but they also need someone to lay the blame on in the scenario where things do not go the way as intended. This does not mean that blockchain will never find its place, the question is that the use case will always be something that works in conjunction to a central reliable system. Bitcoin works on the backbone of real world currency manage by central institutions. Blockchain shines through where multiple parties need to transparency to function, or quite simply where there is a lack of a reliable mediator. Identity cannot be centralised because it presents to much risk to pool all the information at a single node, but handing the whole system in the hands of decentralised service builds even less trust amongst the participants.

Sovrin as a system provides numerous new innovations in the field of Digital Identity Management, especially the development of Decentralised Identifiers(DID) and Verifiable Credentials. It might just be able to succeed in doing where all the systems have failed before. However, its success will more or less likely be dictated by how accepting both the issuers as well as the people are to the network. And given the current analysis, it seems highly unlikely that will happen. Keeping in mind these shortcomings we present a solution that will likely addresses these shortcomings.

In a separate document we propose a solution which revolves around the synchronous functioning of two independent systems. One of the systems is centralised and is intended for well known identities. It is supposed to act as a central repository of identities and respective claims issued by well known identities. The other system is decentralised and supposed to be used by any remaining entities, which are primarily composed of individual users. The reasoning behind this approach and how these two systems are supposed to work will be explained in the following sections. For purposes of reference and uniformity, the central system will be referred as SkyNet and the decentralised system as OneId.