Identity Architecture for Smart Societies

A specification

Peter Waher

Trust Anchor Group AB
peter.waher@trustanchorgroup.com

Abstract. This paper describes the basic principles required for a digital identity architecture that fulfills modern requirements on digital identities, digitalization in smart society and corresponding use cases, as well as privacy protection and cybersecurity requirements for its citizens, companies, organizations and public authorities. The document is suitable as the basis for public procurement documents and requests for tenders, as it discusses generic principles, and is agnostic to specific implementations.

Keywords: Digital Identity, Smart Contract, Tokens, Distributed Ledger Technology, DLT, NFT.

1 Introduction

There is a huge incentive for societies to digitalize their bureaucracies in order to optimize processes, reduce costs, provide more and better services to their citizens, get better oversight and information in the inner workings of the society and desist illicit behavior and fraud. At the same time as this digitalization process is implemented, care must be taken to avoid pitfalls concerning scalability, interoperability, privacy, energy consumption, cybersecurity and a host of other concerns. As technology choices also affect future development in the society, it is also vitally important to not build in architectural limitations from the onset, maximizing extensibility, for future inventions and improvements. It is important that the emerging digital society mimics the goals of our analog society: We want a digital society, that is *open* and *free*, and at the same time is *secure*, and that it protects the basic rights of each citizen. The goal of this paper is to help writers of procurement documents for public tenders for the implementation of digital identities in their societies, to avoid such pitfalls, to assure their future digital societies evolve in accordance with their stated goals and objectives, and that they digitally mirror what we expect from an open and free society.

The document is divided into chapters, that list requirements for building a digital smart society, from a holistic overview, listing design principles, to the concrete implementation principles and use cases necessary to be supported.

2 Use Cases

Following is a short non-exhaustive list of simple use cases, in which a digital society can make great improvements to quality of life and service compared to analog societies:

- **Exchange of company information** can be greatly improved if it was digital, encoded in machine- and human-readable formats, and available using linked-data. Smart Contracts, as described in this document can greatly improve the way in which company information can be automatically managed, processed and referenced in all manner of bureaucratic and commercial interaction.
- **Legal Agreements**. The vast amount of interaction between companies and companies, companies and authorities, or between authorities, are based on legal agreements. Manual creation, signing, transmission, validation and auditing is both time-consuming and costly. But the vast majority of agreements are simple applications of templates that are modified using common sets of parameters. All such agreements can be easily digitalized and automated using the recommendations outlined in this document, with vast improvements in time of service, as well as cost of operation.
- **Forms and Applications**. All kinds of interaction with authorities, require filling out, processing, registration and archiving of standard forms and applications. As with legal agreements, this process can be digitalized, automated and optimized, to a great extent, freeing up public resources for better use, providing better services to the citizens of the society.
- **Taxes**. All taxes and company reports can be standardized and digitalized in a transparent and auditable manner using smart contracts and the principles in this document. This provides increased oversight by authorities, at the same time, as it facilitates administration and interaction with authorities, from a company point-of-view.
- **Book-keeping**. All book-keeping is based on ledgers. Most companies today use proprietary book-keeping software to maintain private book-keeping. The principles described in the following sections provide increased transparency, while at the same time protect the confidentiality of the company, and provide resilience of the information, through the use of next generation Distributed Ledger technology.
- **Company reporting**. All kinds of company reporting, whether it be commercial reporting to shareholders, internal reporting in groups of companies, or reporting to authorities, banks and financial oversight bodies, can today be digitalized and automated using the principles and technologies outlined in this document.

3 Design Principles

Following is a list of design principles that must be adhered to, in the design of a digital society, together with a short description why such principles are required. On the contrary, trust in the digital society will evaporate and parallel digital societies will be created over time by different operators, reducing the effectiveness and purpose of the construction of the digital society in the first place.

- **Open**. It is vitally important that the digital society is open. This means that it cannot refuse access to the digital infrastructure to any member of society. It must be possible for each citizen, except if restricted by a court of law, to be able to participate in the digital society, without the risk of harassment or coercion against their will.
- **Neutral**. Access to the network must be provided without concern to the person and must allow a person to change and evolve over time. Service charges must therefore be neutral, and not be based on cross-domain profiling or social credit scores of personal information, that last over time. Digital identities can therefore not be used as a means to collect such information from users, unwillingly or unknowingly, for purposes other than expressly stated to and consented by the users in a transparent manner, or required by law.
- **Free**. The interfaces and specifications that define the infrastructure of the digital society must be public and free, in public free repositories, without access restrictions or license requirements. This assures that public authorities do not become beholden to patents, copyright holders and commercial enterprises, and that the societies remain free to pursue their interests and development for the interests of the public good. Commercial enterprises are free to compete to provide digital services in the society based on service, quality and price, but not on basis of unfair ownership of the underlying architecture, interfaces and design of the smart society.
- **Interoperable**. A single technical operator cannot be given the responsibility to operate the entire digital society for eternity. Instead, multiple operators may be selected, or replaced, over time, at will, by the public authorities, and they must always interoperate. For technical systems to interoperate, ALL interfaces must be based on *public standards*, as far as this is technically possible, and the interfaces that are not yet in public standards MUST be made available for standardization in global standards bodies, suitable for these tasks, and without limitation to copyright or patents.
- **Extensible**. All interfaces must be made using a *loosely coupled* architecture, to assure they are extensible, and that services can adapt as new services are provided in the society and existing services evolve. Tight coupling must be avoided.
- **Scalable**. Technology must scale, both up and down, be able to be used in small societies, as well as mega-cities, regions, countries and at a supranational or global scale.

- **Local Governance**. The infrastructure must support local governance, where each participant and actor is permitted to govern their own local environment. This applies to small actors, such as persons, as well as medium actors, such as companies and organizations, and large actors, such as municipalities, regional, national or supranational authorities.
- **Secure**. The digital society must protect the privacy, security, cybersecurity and confidentiality of all digital information processed in the infrastructure, by design and by default. It must assure information can be accessed and be processed only by authorized parties.
- **Ownership**. Free and open analog societies are based on the principles of ownership and free enterprise. A digital smart society requires a digital equivalent, a mechanism to define ownership of information and digital assets, and the free trade of such digital assets.
- **Transparent**. A free and open society requires transparent interaction with authorities, and between parties making agreements. This requires a mechanism to assure all identities of participants in such transactions, dealings and information exchange. This makes anonymous or pseudonymous participation in such interactions practically not possible. Furthermore, privacy legislation require identity information only be shared between the parties in such interaction, and required brokers, unless explicitly stated otherwise, or required by law.
- **Autonomous**. A digital society must permit all its operations to run autonomously, once properly configured and necessary consents and approvals have been signed, to maximize efficiency, time consumption, and performance, and to provide the maximum amount of benefits to the society, based on the resources available.
- **Trust**. A digital representation of trust in the network is required. Human interaction and endeavors are based on social interaction along trust relationships. Similar relationships are necessary in digital interaction. For this to be possible, the digital infrastructure must support the concept of "trust-worthy" identities for all participants, and a mechanism whereby you can estimate the level of trust in such identities.

4 Architectural Principles

Once the overall principles for the design of a smart society are determined, the architecture must be defined. Defining a technical architecture for the digital society is vitally important to the infrastructure, as it is difficult to change, once it is in place. Implementations of technical services, features and functions can easily be changed, enhanced or replaced over time. But the underlying architecture is much more rigid, as it affects all participants, and therefore difficult to change. It is therefore necessary to early define the architecture, and assert it complies with the stated design goals. Following are architectural principles that permit the implementation of a digital society that complies with the stated design principles.

- **Decentralization**. Decentralization, as opposed to centralization, permits processing to take place at different nodes in the network. The more the network is decentralized, the less is processed on each node, or, the more information can be processed, in the network as a whole. A decentralized network can potentially grow organically with the number of nodes participating. All successful Internet Standards are decentralized. Decentralization is one of the cornerstones of the Internet architecture, and if designed correctly, permits the creation of scalable and resilient networks, where risk is reduced[1]. A digital smart society must also be decentralized, with no centralized processing required. This assures the infrastructure can grow in scale, both in content, in users, and geographically.
- **Federation**. Federation is the mechanism whereby something complex can be divided down into smaller parts and managed by cooperating decentralized nodes. Federation is based on the interoperability of the decentralized nodes and requires standards specifying the interfaces used between nodes when they interoperate. In a federated network, all nodes are peers; there are no master nodes. A federated network works as a single network, even if different domains process different sets of data. Federation allows multiple actors to participate in operations in the network, regardless of which domain they belong to individually, while at the same time permitting each operator the rights of local governance in their domains.[2]
- **Distribution**. All sensitive and private information should be stored as close to the source as possible, and only be transmitted, when absolutely necessary. Instead of collecting all information and processing it on central locations or replicating it on every node where it is to be processed, it is possible to distribute the actual processing algorithm to where the local information is stored, and let it be processed locally, returning only the results, or the aggregate information, back to the party requesting the information. Interoperation should be based on real-time communication between entities instead of the transport of all information to central data lakes for processing. Distributing algorithms, limiting the transport of information protects privacy, confidentiality and ownership of information. It also gives all parties the control necessary to protect their information and choose to whom to give it. By using local sources of information, in-

[1] The web is an example of a decentralized technology, where content is provided by different servers in a decentralized manner. The servers do not necessarily interoperate, but the web has the ability to grow without practical limits, simply by adding more and more servers to the network. It is also practically impossible to take down the web.

[2] E-Mail is a very successful example of a federated technology. Any e-mail user can send an e-mail to any other e-mail user, regardless of which e-mail server is used by the users, as long as the servers are configured correctly, and the operators of each server so permit. Even though e-mail is more than half a century old, it has shown itself to be very resilient, and has shown no apparent limits to scale and use, even if the Internet and contents transmitted has seen exponential growth during this time.

formation across the network is also easier to maintain synchronized, as information can be corrected or updated locally, with global effect.

- **Bidirectional Full-duplex ad-hoc Communication**. To avoid limitation, and permit interoperable communication between parties in the network, in accordance with the above principles, underlying communication infrastructure must be based on a single communication protocol that is (1) standardized by the Internet Engineering Task Force (IETF), appropriate standards body for Internet technologies, and (2) support the construction of federated networks where participants of different domains can interact without the required approval of a third party, and (3) support full-duplex bidirectional ad-hoc communication between the participants, without creating security vulnerabilities or back-doors.[3] This allows the participants to interact, as if they were peers (i.e. forming a peer-to-peer communication layer). It must also be (4) extensible, such that different participants can define their own messages, without risk of interfering with, or colliding with messages defined by others, and (5) apart from asynchronous messages support additional communication patterns such as (a) Publish/Subscribe, (b) Multi-casting, (c) Request/Response and (d) Event-subscription.
- **Network Identity**. All participants must use a well-defined global network identity, identifying the participant in the federated network, for communication purposes, in accordance with the communication protocol used in the Interoperable communication. The IETF-standardized protocol must include protection against spoofing, as well as to inform the recipient of messages who the sender are. Each participant must authenticate itself with SASL-compliant secure authentication mechanism over a state-of-the-art encrypted transport connection, to gain access to the network, using its associated network identity.
- **Network-based Consent**. The IETF-standardized protocol must include a mechanism to record and retract consent where each user is able to control who can communicate with it efficiently, or not.
- **Conceptual Identity**. All participants with a network identity, should be able to create one or more *conceptual identities*. Conceptual identities are human- or machine-understandable identities of humans, services or things, restricted in time, and consist of collections of meta-data items, in the form of standardizable key-value pairs.[4]
- **Self-Sovereign Identity**. All conceptual identities can be signed by its creators, together with the network identity and a duration, and attach-

[3] This means communication can flow between peers in both directions freely, without regards to specific restrictions in communication patterns, or roles, such as request/response between master and slaves, or client and servers (in HTTP) or publish/subscribe between publisher and subscriber (in MQTT for example), suitable only for certain use cases.

[4] For humans actors, first name, last name, personal number, address, country, etc., are examples of meta-data about a person, that collectively form a conceptual identity about the person.

ments, such as photos or other form of electronic documents, using a *cryptographic key*. Such cryptographic keys must use a public-key state-of-the-art cryptographic algorithm, where only the client itself maintains the corresponding *private key*. The private key must never leave the client device. Once such a conceptual identity has been created it becomes *immutable*, and anyone with access to the public key can also validate the contents of the underlying conceptual identity. The signed cryptographic identity sometimes also goes under the name of self-sovereign identity, as it provides the client with the sole capability to sign using the identity.

- **Trust Provider**. A trusted party can act as a Trust Provider in a network, by validating cryptographically signed self-sovereign identities, and once approved, provide a secondary signature of approval on the identity. Trust Providers act as *electronic notaries* in the network, facilitating interoperation across the network, between different actors in different domains, as long as both domains trust the Trust Provider.[5]
- **Legal Identities.** Once a cryptographically signed self-sovereign identity has been approved by a Trust Provider that takes legal responsibility as an electronic notary, it can be used as a *legal identity* in the network, for the basis of signing legally binding smart contracts, and interact in operations with a legal context. Such legal identities are the basis for *cross-domain* interoperation, where the knowledge of the legal identity of each participant is required, and often a legal approval necessary.[6]
- **Personal Identities** are approved legal identities with a minimum of predefined meta-data information defined, such as complete first and last names, address, city, postal code, personal number, country, etc.
- **Corporate Identities** are personal identities with additional meta-data encoded into it, identifying the company (name, organization number, address, country, etc.), as well as department and position information. Authorized corporate identities can act as signatories for the company, in the digital infrastructure.
- **Revoking Identities**. The Trust Provider must incorporate methods to revoke, obsolete or mark as compromised, any legal identity issued by itself, and provide an interoperable mechanism to allow users in the network to validate the current state of objects the trust provider has issued and approved.
- **Know Your Customer (KYC)**. One of the principal responsibilities of a Trust Provider, is to approve identity applications. This process is often named “Know Your Customer”, and is an integral part of operations of banks and financial institutions. The Digital Infrastructure must support

[5] Certificate Authorities is an example of a Trust Provider (in X.509). Large trusted companies, financial institutions, notaries and public authorities, who are accustomed at validating the identities of people in their domain, can all act as trust providers.

[6] Each network identity can, over time, have multiple legal identities, as the latter are limited in time.

modern methods to perform KYC tasks. This includes support for manual KYC, where the Trust Provider manually reviews information and perhaps interviews application, zero-configuration methods, such as peer-review approval of applications, and automatic forms of KYC using biometrics, and third party services that can be integrated into the infrastructure.

- **Legally binding Smart Contracts**. The digital infrastructure requires a mechanism to create legally binding smart contracts. They must be readable by humans, must be localizable and also be able to be automatically readable by machines. There must also exist a method to ensure the legal integrity of the contract, and that machine-readable and human-readable parts match. A manual process is permitted, when creating smart contract *templates*. In such a process, an application is made to a Trust Provider, that reviews the contents of the contract, ensures its legality and compliance with rules and regulations, and its integrity, and approves it. Once approved, it can be used as the template for the automatic creation of approved smart contracts, where only parameters in approved ranges, and signing parties are allowed to be modified.
- **Legal Consent**. Legal consent, as well as contractual obligations, can be modelled as legally binding smart contracts, as described above, with the difference that legal consent smart contracts must be revokable, while contractual obligations are legally binding for the duration of the contract.
- **Identifiers**. All objects in the digital infrastructure must be addressable with a simple *identifier* in the form of a URI, where the URI schema indicates the type of object the identifier points to. This to comply with *linked data* requirements[7]. Objects that must have such identifiers include network identities, legal identities, smart contracts and tokens. Other types of identifiers (such as personal numbers, etc.,) are typically encoded as metadata in legally binding identities and smart contracts.
- **Tokens** are cryptographically protected digital representations of some form of a digital object or claims that can be easily transmitted and processed in decentralized networks. For the purposes of a digital infrastructure for smart societies, such tokens can be used to protect the ownership of information and digital assets. They must be transparent, ownership clearly defined, and be easily validated by recipients, as well as be able to be audited. To assure legality of the token constructs, all token operations such as the creation of tokens, change of ownership and destruction of tokens are performed using legally binding smart contracts. The creation of a token is done using a creation smart contract, defining the token. Transfer of ownership of a token is done using token sale smart contracts. Destruction of tokens is done using token destruction smart contracts. A token always has exactly one *owner*. The infrastructure needs to allow the user to prove it owns the token. In all token transfers, it must always be

[7] https://www.w3.org/standards/semanticweb/data

clear to all parties involved who the other parties in the transaction are. From the creation contract, it is clear what the token represents, and what parties are responsible for the creation of the token, and what roles they have. Apart from an initial owner, there may also be creator, certifiers, valuators and witnesses, all to increase trust in the digital asset the token represents. Ownership can later be validated through a chain of ownership contracts. The Trust Provider responsible for the creation of the token also maintains a ledger with information about the current owner, and the history of the token.

5 Implementation Principles

The physical embodiments of the principles outlined above consist of the following elements:

- **Digital Identity App** allowing participants to create self-sovereign identities and interact in the digital smart society. The App should exist in at least three forms. A simple access-only app, that permits the user to create a legally identity and use it to access digital resources. A more advanced app can allow the user to use the legal identity to sign, manage and propose contracts, as well as create contracts from templates. A third version also includes a digital wallet containing the digital assets, in the form of tokens, that the user owns. All apps must be open, freely available in public repositories, free to fork and modify, providing trust and allowing a natural evolution of technology, for the benefit of users.
- **Trust Provider Broker** acts as a broker in the federated network and permits the operator to become an electronic notary in the infrastructure. The Trust Provider can receive and approve identity and smart contract applications. These can be used insofar the Trust Provider is trusted in the network. Organizations and authorities with wide societal trust can act as Trust Provider in a wider sense, than a small private organization can. All Trust Providers can act as facilitators in digital services and automation, as far as their digital trust allows.
- **Privacy-Protecting Distributed Ledger** is used to record all legally binding interactions in a way that can be audited by authorized individuals. The Ledger is the principal resource for auditing, while the broker is the principal source for real-time communication and validation. Access to the broker does not mean access to the ledger. The ledger must only be made available to authorized parties. Furthermore, the ledger must be decentralized and distributed, permitting different trust providers to manage their own ledgers with different collections of information. The federated network must facilitate the interoperation of blocks in the distributed ledgers, organizing them per domain. Information in the distributed ledger must be signed, to protect the integrity of the information, and make it immutable. At the same time, all information in the ledger must be pro-

vided a life-time individually, and must be automatically archived, and finally deleted, at the end of that life-time. Private and personal information must not be stored indefinitely, but only for as long as purposes require.

- **Smart Contract Designer** is a free, and public software, permitting users to design, create and publish smart contract templates in accordance with the interoperability interfaces defined in the infrastructure. It can also be used to publish and propose contracts between users.
- **Libraries** must be available publicly and free, allowing anyone to build their own client applications, services and automatic agents in the digital infrastructure. Libraries must be based on open and public source code available in public repositories on the Internet, and must cover all aspects of the interfaces underlying the digital infrastructure of the smart society.

6 Technology choices

Following is a list of standardized technologies suitable for use in the implementation of a digital smart society, in accordance with the principles outlines so far:

- **XML**, or Extensible Markup Language[8], is a standard created and maintained by the Word-Wide-Web Consortium, or W3C for short. It defines an extensible language for machine-readable information. It allows any domain, in control of a domain name, to define machine-readable semantics, that does not collide with any other definition defined by other domains with their domain names. XML does this through the use of *namespaces*. XML is also transparent and can be validated and audited through the use of *XML Schema* Definition Files, or XSD. XML with XSD makes content well-defined, auditable, possible to validate and deterministic, which makes it perfect for use as the machine-readable definition of legally binding smart contracts in an extensible and loosely coupled architecture. This implies that infrastructure components do not need to be updated, because new definitions and semantics is invented and evolved in the network.
- **XMPP**, or Extensible Messaging and Presence Protocol[9], is a federated secure communication protocol, standardized by the Internet Engineering Task Force, or IETF, and maintained by its own standardization body, the XMPP Standards Foundation. It supports a secure, full duplex, bidirectional ad-hoc communication between peers in the federated network. It contains a consent-based authorization mechanism, that does not require manual operator intervention. It also contains standardized extensions for various communication patters, such as multi-casting and publish/subscribe, and other patterns. XMPP is based on XML, which makes communication of information in XML format seamless across the net-

[8] https://www.w3.org/XML/
[9] https://xmpp.org/

work. XMPP is for these reasons, suitable as the basis for a digital infrastructure for a smart society.

- **XBRL**, or Extensible Business Reporting Language[10], is an internation standard for business reporting. It is standardized and maintained by XBRL International, a global not for profit organization. It is used in many countries to automate business reporting and administration. As it is based on XML, it is very suitable to use together with XMPP and smart contracts, as the basis for reporting information by and about companies, and in communication with authorities in a digital smart society.
- **LEI**, or Legal Entity Identifier[11], is an identifier for organizations and companies, standardized by International Organization for Standardization, or ISO. As a short and well-recognized identifier based on a global standard, it is easy to encode as meta-data in legally binding corporate identities, in a loosely coupled manner.
- **XMPP Interfaces for IoT Harmonization**, is a working group within IEEE with the goal of standardizing interfaces for smart societies[12]. This includes interfaces legally binding digital identities and smart contracts, as outlined in this document. All interfaces are open, publicly available and free to use. They are also based on XML and XMPP, and are therefore suitable for use when constructing the infrastructure of a digital smart society.
- **Neuro-Foundation** is a public foundation, not for profit, membership sponsored, which is under construction. The goal of the Neuro-Foundation is to make public interfaces and free software available, complying with the above-mentioned principles and technologies. The goal of Neuro-Foundation is to assure the longevity, support and quality of the technology underlying the construction of digital smart societies, as well as forming the basis of a community supporting and maintaining the effort.

[10] https://www.xbrl.org/
[11] https://www.gleif.org/en/about-lei/introducing-the-legal-entity-identifier-lei
[12] https://gitlab.com/IEEE-SA/XMPPI/IoT