Whitepaper

Self-Sovereign Identity

Whitepaper about the Concept of Self-Sovereign Identity including its Potential

Version 1.0, 13.10.2017

Andreas Abraham (andreas.abraham@egiz.gv.at)

Abstract: This document provides three main contributions. First, it details the Self-Sovereign Identity concept including its underlying blockchain technology. Second, related technologies are identified; evaluation criteria are defined and used to evaluate these technologies. Finally, the SSI potential is identified and described.

E-Government Innovationszentrum
Inffeldgasse 16/a, A-8010 Graz
Tel. +43 316 873 5514
Fax. +43 316 873 5520
E-Mail post@egiz.gv.at
www.egiz.gv.at

Das E-Government Innovationszentrum ist eine gemeinsame Einrichtung des Bundeskanzleramtes und der TU-Graz

BUNDESKANZLERAMT ÖSTERREICH

TU Graz
Graz University of Technology

Table of Contents

Abbreviation

[CA]	Central Authority
[DDNS]	Decentralized Domain Name Service
[DDO]	DID Description Object
[DID]	Decentralized Identifiers
[DLT]	Decentralized Ledger Technology
[DPKI]	Decentralized Public Key Infrastructure
[eIDAS]	Electronic Identification, Authentication and Trust Services
[EU]	European Union
[GDPR]	General Data Protection Regulation
[IdP]	Identity Provider
[LoA]	Level of Assurance
[PET]	Privacy Enhancing Technologies
[PIMS]	Personal Identity Data Management
[RA]	Registration Authorities
[SP]	Service Provider
[SSI]	Self-Sovereign Identity
[SSO]	Single-Sign On
[ZKP]	Zero-Knowledge-Proof

1 Introduction

Not long time ago, Bitcoin was introduced, a crypto currency with a promising underlying technology called blockchain [1]. This new technology, which builds the basis of Bitcoin, can be used in different other use cases because of its promising features. It can be seen as distributed storage that ensures integrity of the data and solves an integral trust issue. A Blockchain consists of a cryptographically linked list of blocks, which allows only to append blocks. New blocks are added and linked after they have been verified from Blockchain nodes in the network. This process ensures the integrity because blocks in the Blockchain cannot be changed.

Every day we are using more and more online services. This is enabled by the fast developing IT technology. To be able to access these services we are using digital identities, which are simplified digital representations of ourselves. This digital identity consists of a set of attributes related to an identity.

Identity management manages those digital identities and their corresponding data. It was further developed through different stages. One of the promising identity management models is the user centric model. In this identity model, the user's identity data are stored in the user's domain. The identity models come with certain issues such as there is always trust to a central authority required. Transparency cannot be fully provided, since there is a trusted authority involved.

These issues can play an important role in certain use cases, which leads to the conclusion that a new identity model for these use cases has to be developed. Identity management built on the promising blockchain technology would enable an identity model, which reduces the previous issues for certain use cases. This work introduces Self-Sovereign identity (SSI), a new identity management model. This identity model tries to remove the trust issue that comes with identity management. Moreover, SSI also tries to give the user fully control over his/her own data.

In this work, we introduce the term Self-Sovereign identity and derive requirements from it. The identified requirements are used to define criteria, which are used to evaluate related Blockchain technologies. Finally, we recommend the most promising technology to realize a SSI system.

This whitepaper is structured as follows. The concept of Self-Sovereign identity is definded in Section 2 including the blockchain technology. Section 3 gives a technology overview followed by the definition of evaluation criteria, which are afterwards being used for the technology evaluation. Section 4 identifies and describes the possible potential of the SSI technology. Finally, Section 5 concludes this work.

2 Self-Sovereign Identity (SSI)

The increasing everyday usage of different online services requires an efficient digital identity management approach. These identities often contain sensitive personal data especially when used in the eGovernment context. Many people are concerned about how these sensitive data are being managed in terms of where these data are stored and who can access it.

Thus, identity management becomes more and more important. Different identity models evolved during time triggered by the increasing demand of online services as well as the further development of those services. The four main identity models are as follows.

Isolated Identity Model

The evolution of identity models started with the isolated identity model, which is still the most common model. Its main concept expresses the combination of the service provider (SP) and the identity provider (IdP), which means that the SP manages the user's identity data as well as their credentials. In this case, the user authenticates herself directly at the SP.

Central Identity Model

In contrast to the isolated model, the central identity model separates the IdP from the SP. This separation is the main difference and advancement because the identity data are stored at the IdP. When a user wants to access an online service, she has to first authenticate herself at the IdP and afterwards the identity data are transferred to the SP. In this model, the user does not have any control over her own identity data. An example for this scenario is using Facebook because the user does not have control over her own data stored at Facebook.

User-Centric Identity Model

The user-centric model differs from the central model in storing the user's identity data in the user's domain. This domain could be a secure token such as a smart card. Sharing identity data of a user requires explicit user consent. An example for a use case scenario where this model is used is the Austrian Citizen Card.

Federated Identity Model

The federated identity model differs from the previous defined models by distributing identity data across multiple IdPs instead of storing it in one central place. In this model, multiple IdPs provide the required identity data to access a service. These IdPs are working together in a federation, which requires a trust relationship between the IdPs. Federated IdPs share a user's common identifier. This model can be used to realize Single-Sign on (SSO). SSO would be the authentication subset of federated Identity Management.

Self-sovereign identity as next identity model

The next further stage of identity models is the Self-Sovereign identity (SSI) model. In this model, the user fully owns and controls her own data. A SSI system creates new requirements – detailed in Section 2.2 - on the technology that is used to create such a system. The blockchain technology fulfills most of these requirements. The blockchain technology is described in section 2.1 and the full description of Self-Sovereign identity concept is described in section 2.2.

2.1 Blockchain / Distributed Ledger

The blockchain was introduced by Nakamoto [1] as part of the peer-to-peer crypto currency Bitcoin. If a user wants to use Bitcoin, she has to install a wallet on her device. The user manages his account using this wallet. Additionally, the user can make transactions such as buying goods and paying with Bitcoin using her wallet. To perform such a transaction, the user has to transfer the right amount of bitcoins to the seller.

Bitcoin utilizes the blockchain technology as transaction register to keep record of all bitcoin transactions. Figure 1. Blockchain and its Blocks Architecture [2] shows the blockchain architecture. A blockchain is a cryptographically linked list of blocks where each block consists of transactions, the hash value of the previous block and a nonce. A transaction consists of input and output amount of bitcoins as well as the address of the sender and receiver. Bitcoin uses public key cryptography where the public key represents the address of a user.

Bitcoin is a public-permissionless blockchain, which means that anybody can host a copy of the Bitcoin blockchain. Special self-appointed entities of the Bitcoin peer-to-peer network, so-called miner, collecting transactions and try to solve a crypto graphical problem. Solving this problem serves the proof-of-work that ensures the validity of the transactions in the block.

The cryptographic problem of the proof-of-work consists of finding the right hash value of a block that starts with the predefined prefix. Miners are using special hardware to

calculate the hash values. A nonce is added to the calculation and increased each time another hash results. To calculate the proof-of-work, a few quintillion hash values or even more have to be calculated [2].

After the proof-of-work was successfully calculated, the miner broadcasts the block to the p2p network. Every miner in the network verifies the calculation and if it is correct, the block is added to the blockchain. This approach increases the transaction security because the whole block is verified. Additionally, the blockchain solves a general trust issue. Different independent miners are verifying a new block, which solves a general trust issue. There is no trust relationship between the different miners required.

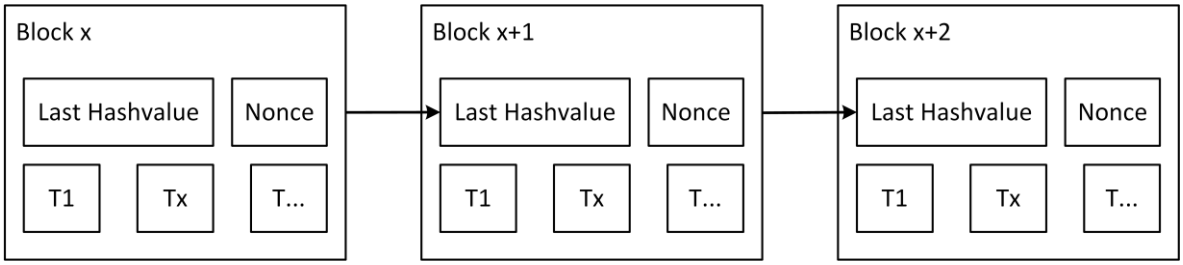

Figure 1. Blockchain and its Blocks Architecture [2]

Miners are motivated to do mining by receiving a commission from the Bitcoin network. This commission should increasingly motivate miners to buy better mining hardware. Only the first one who finished calculating the proof-of-work for a block receives the related commission. This commission is created by sum up the transaction fees. A transaction fee is charged indirectly when processing a transaction. It appears indirectly because the bitcoin transaction output is slightly lower than the input and the difference is the fee for the miner. This system motivates thousands of miners in the world to calculate the proof-of-work. Because each miner has a copy of the blockchain and verifies each new broadcasted block, no trusted central authority is required that observes the blockchain and its activity. The commission for calculating the proof-of-work for one block is now 12.5 bitcoins and changes with the active number of miners. The cryptographic problem becomes more complex after a certain time, which should motivate miners to buy stronger hardware. The provided commission motivates thousands of miners doing their job and invest into computing power.

A known problem with the proof-of-work is the 51% attack. This attack describes the case if there is one miner who holds 51% of the computation power of the whole Bitcoin network. This miner could double spend bitcoins as well as invalidate other transactions and potentially prevent people from sending bitcoins.

This attack is not a theoretical problem, it can become real when the big mining pools start incorporating with each other and become the dominant mining power. Chinese bitcoin mining pools are together already the most dominant in the world [3]. Between May and June 2016, China's mining pools mined about 70 percent of the new bitcoins.

The Blockchain Technology

Bitcoin's underlying blockchain technology offers huge potential in many directions. Many researchers and developers recognized the power of this novel technology and created their own blockchain architectures and implementations. A blockchain is an append-only, ordered and replicated log of transactions. Some of the blockchain implementations are similar to bitcoin – used as crypto currencies – whereas others are used in a total different manner such as for e-Voting or as decentralized domain name service (DDNS).

Various blockchain forks or own implementations try to add features to fulfill several requirements related to different use cases. During the further development of the blockchain technology, a new term was introduced namely distributed ledger. This synonym is used for some implementations to differ from the blockchain.

This work describes the usage of the blockchain technology to store digital identities and its corresponding attributes. Using the blockchain technology, we are going to develop a system, which enables Self-Sovereign identity. In order to design such a system we identified criteria for such a system and evaluate different blockchain implementations such as Sovrin, Blockstack, Multichain, uPort, Ethereum, Shocard, Kyc-Chain and Idcubed to conclude their advantages and disadvantages when using them for building a Self-Sovereign identity system. The concept of Self-Sovereign identity is described in Section 2.2. The different blockchain implementations are described in Section 3 together with the description of the evaluation criteria as well as the evaluation of the different technologies.

2.2 SSI Concept

Sovereignty is, per definition, a supreme power or authority, which governs itself without any outside influences. Sovereignty for identity management means that the user's identity data are fully owned and controlled by herself.

The concept of Self-Sovereign identity can be seen as the next stage of evolution in identity management. The blockchain technology provides a good basis to create a SSI system. The requirements of such as system are detailed as follows.

Each individual has to have the full control over her data

Each user must have full control over her own identity data. This includes not only what identity data are being stored but also who has access to these data. The user should be able to add or import identity attributes as well as delete or revoke them at her leisure. Also, all access of identity data of a user should be logged for later verification.

Ensure security and privacy of user's identity data

All identity data have to be stored and processed in a highly secure manner. Additionally, the user's privacy has to be preserved. For instance, unlinkability between the user wallet and her identity data increases the user's privacy.

Fully portability of the data

This requirement describes that the user should be able to use her identity data wherever they want. For instance, a SSI system can be used as identity provider when the user tries to access an online service.

No trust in a central authority is required

The underlying blockchain technology solves the required trust related to a central authority.

Ensure data integrity

The integrity of identity data can be ensured by utilizing the blockchain. This is one of the main advantages using the blockchain technology.

Transparency of the identity data is maintained

The blockchain technology provides data transparency of all in the blockchain stored data. All changes to the data in the blockchain are fully transparent so that no one can alter of delete data without someone else noticing it.

2.2.1 Benefits for Stakeholders

The stakeholder of a SSI system are citizens, public administration and businesses.

From a citizen perspective, a SSI system would give the citizens power to fully own and control their own identity data. Furthermore, trust in a single authority is not necessary anymore. Security and privacy will be maintained. The citizen only decides what kind of data are going to be stored in the ledger and who is going to access it. For example, if a SSI system is deployed in the whole European Union, it can be utilized for cross-border authentication and for cross border services. A SSI system provides the citizens a platform to use their electronic identity all the time and everywhere.

Such a SSI system is also beneficial for the government. It could help decrease costs related to identity management. Additionally, a SSI system offers full transparency in the identity management, which can help to increase the citizen's trust in the government. In addition, cross border government processes and services can easier be realized.

Businesses can benefit from a SSI system when providing services and a qualified identity provider is required. It can save time and costs for the companies.

2.2.2 Challenges

When applying a SSI model, various challenges arise such as access permission level of the blockchain, proof of work calculation, missing technical understanding and data storage issues.

The access permission level of the blockchain depends on the chosen technology. This can vary between public (permissionless) to private (permissioned) access permission level. Using a public blockchain requires the proof-of-work calculation to ensure secure and tamper resistant consensus. This proof-of-work requires huge amount computation power. In contrast, using a private blockchain requires trust in the parties that are responsible for writing and reading data from or to the ledger.

Another obstacle when using SSI model can be the missing technical understanding of this novel technology. This lack of knowledge can lead to problems such as integration issues when trying to integrate a SSI model into existing infrastructure.

Utilizing SSI emerges data storage issues especially when storing sensitive data. Some blockchain implementations are using additional external storage. When the sensitive data are stored encrypted in the blockchain key distribution problems arising. The limited storage capacity of blockchains describe another storage issue.

2.2.3 SSI Architecture

The architecture of a Self-Sovereign identity system based on the blockchain technology is described as follows.

The blockchain offers the possibility to realize a system without semi-trusted parties such as central certificate authorities (CAs) or registration authorities (RAs). Nevertheless, the degree of required trust depends on the specific blockchain implementation. Only public or permissionless blockchains provide a fully semi-trust less environment. If a blockchain is private or permissioned, only authorized parties have access to the ledger, which requires at least some kind of trust relationship to these entities. Even though authorized parties are independent of each other, trust in the chosen parties or in the selection process of becoming a trusted party is still necessary.

Independent authorities should host a copy of the ledger that helps reducing the required trust. These authorities can be for example different countries of the European Union. The citizen of a country A could then import their identity data using the existing eIDAS infrastructure. eIDAS nodes are used to import qualified identity data into the user's ledger. A citizen should be able to use her smart phone to access or share her identity data wherever they want where each identity's data access is logged.

The blockchain ledger is the central part of such a SSI system. Nevertheless, there are additional parts required to fully support all features and use cases which are described in Section 4. Therefore, the distributed ledger has to be extended by at least two additional parts namely the off-ledger storage and the data import part.

Off-Ledger Storage

The first additional part is the off-ledger storage. Storing sensitive data in the blockchain might not be a good idea even though if these data are encrypted. The problem that occurs by storing data in the blockchain is that these data cannot be modified or deleted anymore afterwards. This might be an important feature in some cases but not when dealing with sensitive data.

In the SSI system, person related data are stored off-ledger only a unique identifier is stored in the blockchain. This special identifier is cryptographically linked to the off-ledger data storage. Different storage services such as cloud storages can be used as off-ledger storage.

Data Import

The second part is the data import extension. A SSI system deals with identity data. These data can have different levels of quality. For instance, a national authority can issue qualified identity data of a person. In contrast, a person enter person related data on their own. The authenticity of these self-entered data cannot be guaranteed. Therefore, it can be beneficial for a user that the SSI system supports the import of qualified data.

The qualified data import is not straightforward because a special transformation of the data during the import process has to be performed. This transformation process converts the data format from the received format to the SSI system supported format. The received data format can vary depending on its source. This is a necessary step in order to provide selective disclosure and attribute attestation at a later point.

3 Technology Evaluation

This consists of three parts. First, the evaluated technologies are introduced. Second, the evaluation criteria are identified and detailed. Finally, the evaluation results are described.

3.1 Technologies Overview

This section gives a short overview of the technologies, which are relevant for a SSI system. During our technology research, we also looked at other blockchain technologies such as slock.it, Namecoin, Certipeople, Respectnetwork, Qiy Foundation, Openreputation, Openpds, BYU Domain and SpidChain but they have not been included in our evaluation because either their focus was on something different or the project already ended.

Sovrin

NAME	SOVRIN
URL	https://www.sovrin.org/
OPENSOURCE	Yes (https://github.com/sovrin-foundation/sovrin)
LICENSE	Apache2 License

Sovrin is an open source project that focus on Self-Sovereign identity. The goal of this project is to provide identity data for everybody all the time and everywhere. It is a very promising technology for a Self-Sovereign identity system also because it is based on the distributed ledger technology and utilizes Plenum[1], a crypto library, which offers an advanced distributed consensus algorithm. This algorithm is utilized instead the proof-of-work from the Bitcoin system. Sovrin allows the import of certified identity data and offers attribute redaction. Each attribute has to be signed by the issuing authority, which enables the selective disclosure. Sovrin offers functionality to increase the privacy of users when for instance a service provider wants to verify if a user is over 18. It only returns the required information that the user is or is not over 18 and hides the actual birthdate. This is achieved by applying Zero-Knowledge-Proof (ZKP) crypto primitives as privacy enhancing technology (PET).

It not only uses PETs to preserve user's privacy, but also attaches one or more attestations to identity attributes to prove their authenticity. The attributes are signed either by the user itself of by relevant third parties. (It's one thing for you to say you have a degree; it's

[1] https://github.com/evernym/plenum Plenum is an advanced distributed consensus algorithm that supports elliptic-curve cryptography, digitally signing and message encryption operations.

quite another when the university says so.) Sovrin was designed for one single purpose: globally trustable Self-Sovereign identity.

Sovrin is a permissioned ledger where only permissioned nodes can read or write data to the ledger. The Sovrin foundation decides who is going to be one of these nodes. Two kinds of nodes have to be distinguished depict in Figure 2. Sovirn Architecture Overview **[4]**. First, the validator nodes – placed in the validator pool - that are responsible for the write operations to the Sovrin ledger. 60 to 120 of these nodes are worldwide expected. Second, the observer nodes, which handle the read operations as well as keeping their state synchronized with the validator nodes. A few thousands of these nodes are expected worldwide.

Figure 2. Sovirn Architecture Overview [4] shows the Architecture of Sovrin where each circle identifies a different purpose. The Sovrin agents can be seen as middle layer between the clients and the inner ledger [4].

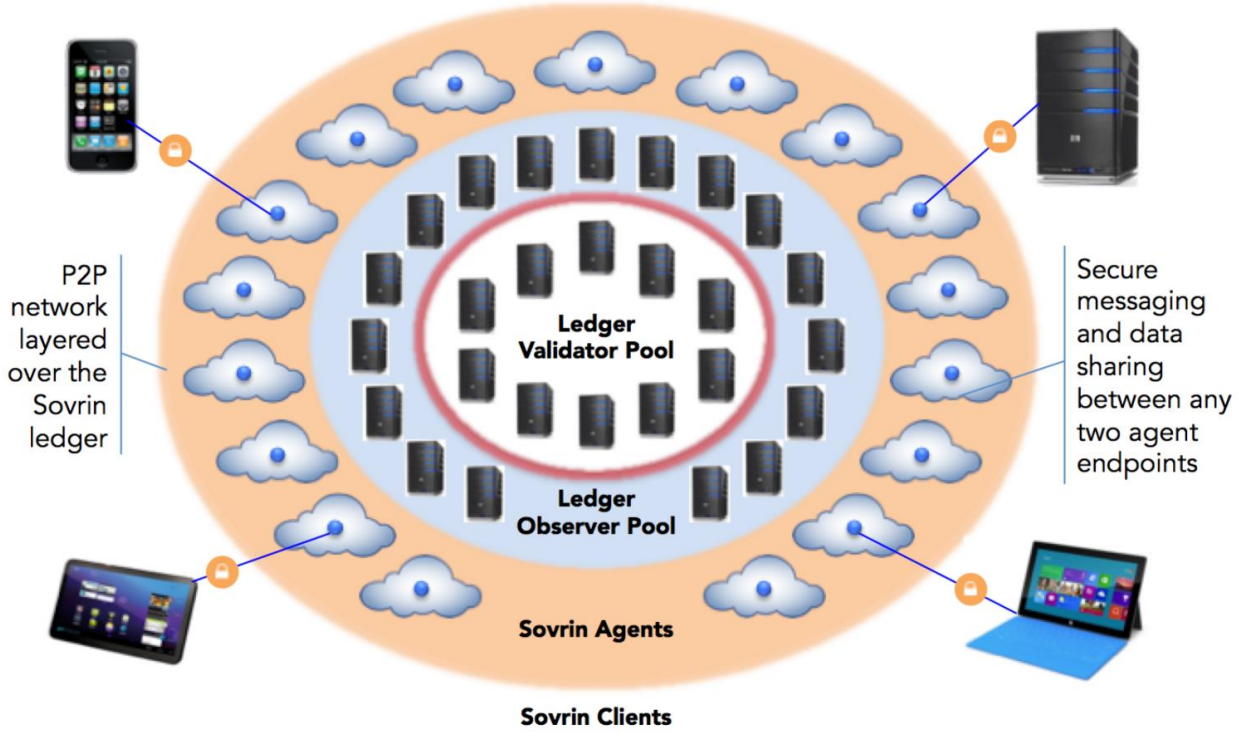

Figure 2. Sovirn Architecture Overview [4]

Blockstack

NAME	BLOCKSTACK
URL	https://blockstack.org/
OPENSOURCE	Yes (https://github.com/blockstack)
LICENSE	GPL v3 License

Blockstack implements a decentralized domain name service (DDNS) combined with a public key infrastructure [5]. It separates the control and data plane. This is realized by the Blockstack's four-tier system consisting of four different layers such as the storage layer, the routing layer, the virtual layer and the blockchain layer shown in Figure 3. Blockstack Architecture [6] The blockchain layer and the virtualchain layer form the control plane and the routing and storage layer form the data plane. The different layer shown in Figure 3. Blockstack Architecture [6] are detailed as follows.

Blockchain Layer

The Blockchain layer stores sequences of Blockstack operations. In addition, it provides consensus of the written operations.

Virtualchain Layer

The virtualchain layer is utilized to define the Blockstack operations. These operations are encoded in the virtualchain layer to validate blockchain transactions as additional metadata.

Routing Layer

Blockstack separates routing requests and actual storage of data. Routing information are stored in zone files where and the virtualchain layer binds names to the zone files.

Storage Layer

The storage layer uses already existing storage solution to store data. Possible storage solutions are Dropbox, Google Drive etc. The stored data consist of name value pairs.

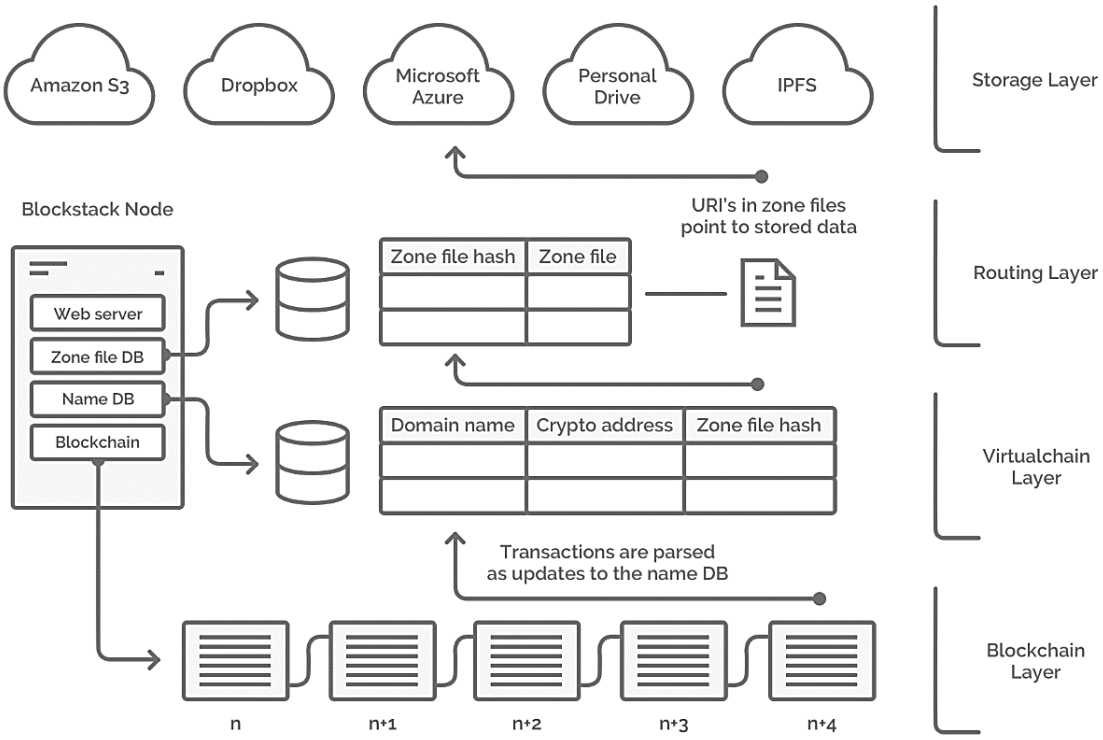

Figure 3. Blockstack Architecture [6]

Multichain

NAME	MULTICHAIN
URL	http://www.multichain.com/
OPENSOURCE	Yes (https://github.com/MultiChain/multichain)
LICENSE	GPLv3 License

Multichain is a fork of the Bitcoin blockchain, which differs from Bitcoin. Multichain provides a platform for creating private blockchains with integrated user permission management [7]. These private blockchains can be used within or between organizations and aims three main goals. First, Multichain ensures that the blockchain and its activity is only visible to chosen participants. Second, it introduces controls over which transactions are permitted. Finally, it enables secure mining without the need and the related costs of calculating the proof-of-work. The approach used by Multichain is called permission-based mining. Only authorized participants are permitted to mine.

In contrast to Bitcoin or other blockchain technologies, Multichain supports not only a single blockchain, instead it supports different blockchains at the same time [7]. Running more than one blockchain simultaneously can offer institutions advantages such as using

one blockchain to monitor incoming funds, which then triggers a transaction on another blockchain.

Ethereum

NAME	ETHEREUM
URL	https://www.ethereum.org/
OPENSOURCE	Yes (https://github.com/ethereum/)
LICENSE	GPLv3 License

Ethereum is a crypto currency similar to Bitcoin as well as a platform for decentralized applications. It was the first blockchain technology, which introduced smart contracts. With the smart contract feature, Ethereum can be seen as the blockchain 2.0 [8]. Smart contracts are applications where its state can be stored in the blockchain. They contain code and can interact with other smart contracts. These smart contracts can be implemented in various touring complete scripting languages.

Ethereum does not directly focus on SSI but it could be used as platform to realize such as system. Summarizing, Ethereum is a crypto currency with support for smart contracts, which makes Ethereum to a blockchain platform that supports decentralized applications.

uPort

NAME	UPORT
URL	https://www.uport.me/
OPENSOURCE	Yes (https://github.com/consensys/uport-lib) -> (https://github.com/uport-project/uport-connect)
LICENSE	Apache License Version 2.0

uPort is built on Ethereum and focuses on sovereign identity for people, businesses, organizations and devices. uPort consists of three components namely a mobile app, smart contracts and the developer libraries.

1. The mobile app is used to create a Self-Sovereign identity and the user's keys are stored on the mobile phone as well.
2. Smart contracts provided by Ethereum form the core of the identity. Additionally, the core contains the logic that is required to recover the user's identity if the mobile phone is lost or broken.
3. The developer libraries offer third party application developers an integration platform when developing an application, which should support uPort.

uPort uses already existing cloud or storage solutions as data storage such as IPFS[2], Azure, AWS or Dropbox. uPort uses this solutions to store an attributed data blob that contains the user's data.

Figure 4. uPort's High Level Architecture depicts uPort's high-level architecture including the different contracts and the standard flow is described as follows.

a) The uPort mobile app communicates with the controller contract, which contains the main access control logic.
b) The controller contract forwards this transaction to the proxy contract, which is a layer between the user's private key, stored on the mobile device, and the application contract.
c) The proxy contract is replacing the private key with a persistent identifier.
d) The application contract is the actual application running on the uPort.

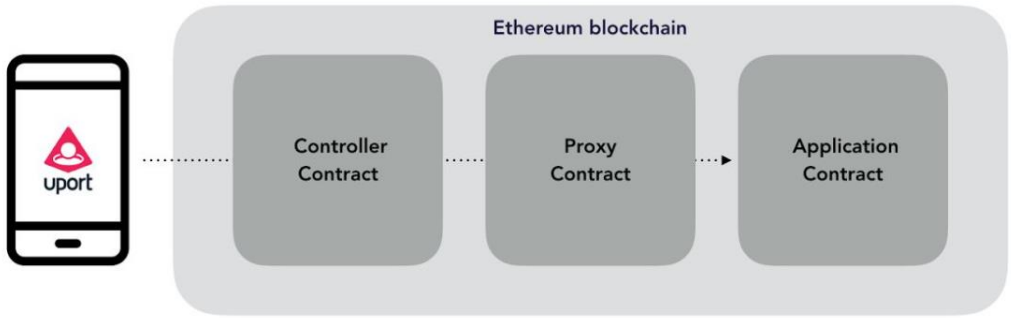

Figure 4. uPort's High Level Architecture [9]

Idcubed

NAME	IDCUBED
URL	https://idcubed.org/
OPENSOURCE	Yes (https://github.com/IDCubed/)
LICENSE	Proprietary

Idcubed is an open source project developed by the ID3 group. It utilizes the Open Mustard Seed[3] (OMS) framework, which provides a powerful new self-deploying and self-administrating infrastructure layer for the internet. OMS uses a combination of different

[2] https://ipfs.io/
[3] https://idcubed.org/open-platform/platform/

technologies such as blockchain 2.0, trusted execution environments, machine learning, mobile security and cloud based computing. Idcubed aims sovereign identity, pseudonyms and verifiable attributes. It provides an openPDS[4] integration for cloud storage, sharing and secure computing using OIDC. OpenPDS is a software that allows user to store, collect and grants fine-grained access to their data while still preserving the user's privacy.

Nevertheless, Idcubed is at a very premature stage where it is difficult to evaluate this project. In addition, it is difficult to gather information about the implementation or the specification; therefore, this project is not included in the evaluation.

Shocard

NAME	SHOCARD
URL	https://shocard.com/
OPENSOURCE	No
LICENSE	Commercial

Shocard is a commercial mobile identity management solution using BlockCypher's blockchain infrastructure. The problem here is that there are no technical details available because it is not open source nor freely available. Because of this, Shocard is not relevant for our evaluation.

Kyc-Chain

NAME	KYC-CHAIN
URL	http://kyc-chain.com/
OPENSOURCE	No
LICENSE	Commercial

The Key-Chain is a commercial project based on distributed ledger technology and allows the users to manage their own digital identities in a secure manner. No technical information could be gathered and only very little information about the project itself are available. Key-Chain is not relevant for our evaluation because it is neither open source nor freely available.

[4] http://openpds.media.mit.edu/

3.2 Criteria

In this section, we identify criteria and describe them. These criteria are derived from the requirements of a SSI system and being used to evaluate the blockchain technologies. The criteria are listed and detailed as follows.

Permissionless (Public) / Permissioned (Private)

This criterion describes if the Blockchain is public or private. The difference here is that either everybody can or cannot access the Blockchain with or without permission. Both permissioned and permissionless blockchains have advantages and disadvantages.

On the one hand, permissionless (public) blockchains prove the consensus of the data with the proof-of-work. This proof usually consists of solving a cryptographic problem. Special entities (miner) try to solve this problem because they receive a commission after successfully solving the problem. The arising problem is to motivate the miner to solve this problem, which is the received commission. Many independent miners are trying to receive the commission by solving the proof-of-work. Therefore, no trust relationship between the parties is required.

On the other hand, the advantage of using a permissioned blockchain is that the proof-of-work does not have to be a difficult to solve cryptographic problem, i.e. transactions can be confirmed within a short timeframe. The need for motivating the miner with commissions is not required. As disadvantage can be seen that the permissioned entities have to be semi trusted or at least the authority, which decides if an entity becomes an authorized entity with write-access to the ledger.

Proof-Of-Work (Mining)

The proof-of-work ensures the consensus of the blockchain to proof that the new blocks are valid before they are added to the chain. Different approaches of the proof-of-work are available depending if the blockchain is private or public. For this proof are different approaches available also depending on if the ledger is permissioned or permissionless. This criterion describes the proof-of-work approach.

Key management system

Key management is an integral part in identity management especially when dealing with sensitive data. Keys are used for several operations such as encryption or signature operations. Theft of key material should be as difficult as possible.

Identity Data Import / Gathering

This criterion describes how the evaluated technology imports or gathers the user's identity attributes. When a recognizable authority signs these attributes, the authenticity of these data can be verified. Whereas, self-signed identity data do not have such a strong authenticity. For instance, it differs when someone says that she has a master's degree or when the university says that this person has one.

Selective Attribute Disclosure Support

This criterion describes if the evaluated technology supports selective disclosure. Selective disclosure is increasing the user's privacy by only disclosing the necessary information to the requester. Different approaches are available described using this criterion.

Data storage

Some blockchain technologies store all data in the blockchain whereas others use different storage approaches such as already existing cloud storage. This criterion describes which data and where these data are stored, which affects security and privacy of the data.

Trust Required

This criterion describes and evaluates the required trust of a blockchain technology. The trust situation can differ depending on if the blockchain is public/private or permissioned/permissionless.

Identifiers

Each technology uses identifier assigned to identities or data. Identifiers, in particular if they are unique, can produce privacy leakage or even cause security issues. This criterion evaluates the used identifiers of a blockchain technology.

Smart Contracts

Smart contracts are an extension of the blockchain technology proposed by Nick Szabo in 1996 [10]. This extension elevates the distributed ledger technology to a higher level by adding functionality to it. Smart contracts are small programs that are executed on the blockchain nodes triggered by various actions such as a predefined date. The idea behind smart contracts was to give developers the power to write their own scripts that are going to be executed in the blockchain. For instance, possible use case could be a regularly transfer money to a company such as an insurance or the property owner. One of the main features of smart contracts is that they are written in a programming language which is Turing-complete.

In the SSI system, smart contracts could be used to not only increase the functionality of the system, but also to make it more flexible for extensions and further developments.

3.3 Evaluation

This section describes the evaluation of the corresponding technologies and details the results.

3.3.1 Sovrin

Permissionless (Public) / Permissioned (Private)

Sovrin is a permissioned distributed ledger. The Sovrin foundation defines nodes, which are both authorized and responsible for reading and writing from and to the distributed ledger. Sovrin defines two types of nodes that are responsible and authorized for read and write operations to the ledger. The observer nodes are responsible for the read operations from the ledger and the validator nodes handle the write operations. Additionally, the observer keep the ledgers synchronized between them and the observer nodes.

Proof-Of-Work (Mining)

This technology uses a distributed consensus protocol named Plenum Byzantine Fault Tolerant Protocol[5]. This open source protocol – developed by Evernym[6] - is optimized for

[5] https://github.com/evernym/plenum
[6] https://www.evernym.com/

security and scalability. Because the ledger is permissioned, a difficult computation to calculate the proof-of-work is not required.

Key management system

The keys are stored in a decentralized public key infrastructure (DPKI). An identity owner uses keys, which are stored in a keychain. The keychain itself is stored in a DPKI. Approaches for key discovery, rotation and revocation and recovery are defined.

Identity Data Import / Gathering

The imported identity data have to fulfill the Sovrin data format. Additionally, to be able to apply selective disclosure, each identity attribute has to be individually signed.

Selective Attribute Disclosure Support

Sovrin supports selective disclosure of identity attributes. The imported attributes have to fulfill the Sovrin format. If authenticity is required each attribute has to be individually signed.

Data storage

Sovrin offers two different storage options: the on- or off-ledger storage. Identity data, keys, transaction proofs an pointers are stored on-ledger. Other digital data can be stored off-ledger. Nevertheless, multiple factors influence the decision what kind of data are stored on- or off-ledger.

Trust Required

The board of trustees governs Sovrin as well as decides and approves who is going to be a steward. A steward is a trusted entity, which is performing operations on the ledger. Stewards take the role of observer and validator node. Therefore, trust in the Sovrin foundation and their decisions is required.

Identifiers

Sovrin uses key-value pairs to identify each identity owner on the ledger [11]. Decentralized identifiers (DIDs) are the key and DID description objects (DDOs) are its associated value. These key-value pairs are called DID record. The design of DID records removes the dependency on a central authority.

Smart Contracts

Smart contracts are not supported by Sovrin.

3.3.2 Blockstack

Permissionless (Public) / Permissioned (Private)

Blockstack is a permissionless (public) blockchain built on top of the Bitcoin blockchain[7]. Everybody can download and run the Blockstack core, which will then be a node in the Blockstack network.

Proof-Of-Work (Mining)

The mining in Blockstack is performed indirectly. Indirectly means that the mining is performed by the underlying Bitcoin blockchain.

Key management system

Blockstack uses a distributed key management where the public keys are stored in the virtualchain layer and its corresponding private keys are stored in a secure element of the client side.

Identity Data Import / Gathering

Blockstack user can enter their identity data on their own.

Selective Attribute Disclosure Support

Blockstack supports selective disclosure.

Data storage

Blockstack uses the storage layer as data storage. This layer utilizes already existing storage infrastructure such as Dropbox, Microsoft Azure, Personal Drive or BitTorrent to store data.

[7] https://github.com/blockstack/blockstack

Trust Required

No trust is required because Blockstack is built on Bitcoin and uses a public blockchain where everybody can host his/her own Blockstack node.

Identifiers

Blockstack identifiers are stored in zone files in the routing layer. The binding of zone file to name is stored in the virtualchain layer.

Smart Contracts

Blockstack does not support smart contracts.

3.3.3 Multichain

Permissionless (Public) / Permissioned (Private)

Multichain is a platform that offers the possibility to run private (permissioned) Blockchains. It is a Bitcoin fork built on the Bitcoin core.

Proof-Of-Work (Mining)

The mining approach in Multichain is a lightweight proof-of-work compared with Bitcoin. Only restricted entities are allowed to mine, which removes the need of solving a difficult cryptographic problem. Multichain implemented the mining diversity parameter to prevent the arising issue that one miner could monopolize the mining process. The mining diversity describes a constraint on the number of blocks, which may be created by the same miner ($0 \leq mining\ diversity \leq 1$) [7]. Summarizing, Multichain enables secure mining without the costs of calculating the proof-of-work.

Key management system

The Multichain built-in wallet stores – as default storage option – the private keys of the user [12]. The wallet can be encrypted to increase security. Multichain offers the possibility to store private keys outside the wallet in an external hardware security module (HSM) or external computer.

Identity Data Import / Gathering

Identity data are both self-entered as well as self-signed.

Selective Attribute Disclosure Support

The stream feature[8] of Multichain supports selective disclosure.

Data storage

Multichain uses the blockchain itself as data storage to store any kind of data up to 64MB per transaction.

Trust Required

There is no trust in any party required and the definition of a transaction includes a proof of authorization and a proof of validity.

Identifiers

Identifier are together with quantities encoded and stored within each transaction output. Every Multichain node is responsible for verifying and checking the quantity of assets in transaction, which is similar to the process of the native currency of the blockchain.

Smart Contracts

Multichain does not support smart contracts.

3.3.4 Ethereum

Permissionless (Public) / Permissioned (Private)

Ethereum is a public (permissionless) blockchain used as distributed computing platform, which introduced the smart contract feature.

Proof-Of-Work (Mining)

[8] http://www.multichain.com/developers/stream-confidentiality/

As mining algorithm, the Ethash[9] algorithm is used. It was previously known as the Dagger-Hashimoto algorithm. This algorithm is designed to hash fast within a slow CPU environment and additional it provides speed-ups for the mining process[10].

Key management system

Keyfiles store key pairs related to an account. These keyfiles are JSON text files with by default encrypted private keys. The keystore stores the keyfiles of each own Ethereum node's data directory. The accounts are indexed by an address that is derived from the public key[11].

Identity Data Import / Gathering

Ethereum is mainly used as platform for distributed applications and the identity data are self-entered only. Applications built on Ethereum could handle this differently.

Selective Attribute Disclosure Support

Ethereum does not support selective disclosure of identity attributes out of the box. Smart contracts could be used to implement this feature.

Data storage

Ethereum considers two different approaches as data storage. The former, the blockchain directly stores smart contracts including a minimal set of data that must be available at any time. The latter, larger objects are stored in distributed object storage or distributed file system such as Swarm[12], Storj[13] or IPFS[14].

Trust Required

Ethereum is a public-permissionless blockchain that also utilizes a proof-of-work algorithm; therefore, there is no trust required.

Identifiers

---

[9] https://github.com/ethereum/wiki/wiki/Ethash
[10] http://ethdocs.org/en/latest/mining.html
[11] http://ethdocs.org/en/latest/account-management.html
[12] Swarm is a decentralized storage and content distribution platform as part of Ethereum.
[13] storj.io is a distributed cloud storage.
[14] https://ipfs.io/ Inter Planetary File System

During the identity creation, a smart contract creates for each identity a stable identifier that cannot be changed afterwards anymore.

Smart Contracts

Ethereum supports smarts contracts.

3.3.5 uPort

Permissionless (Public) / Permissioned (Private)

uPort is public-permissionless built on Ethereum blockchain.

Proof-Of-Work (Mining)

The proof-of-work is performed by the underlying Ethereum blockchain. More details can be found in Section 3.3.4.

Key management system

The client application stores the private keys on the mobile phone. Whereas, the public keys are stored on distributed storage such as IPFS in order to support a decentralized public key infrastructure (DPKI).

Identity Data Import / Gathering

uPort supports a method to enter identity attributes including a signature to prove the attestation. These attributes can be signed from a qualified authority or simply just be self-signed.

Selective Attribute Disclosure Support

The selective disclosure feature is supported by uPort. In the whitepaper says that uPort allows end-users to securely and selectively disclose their data to counterparties [9].

Data storage

uPort uses off-chain data stores such as IPFS, Microsoft Azure, AWS[15], Dropbox, etc. The uPort identities are cryptographically linked to the off-chain data stores.

Trust Required

There is no trust required because uPort is built on the public Ethereum blockchain.

Identifiers

The core of the uPort identity is the uPort identifier, which is a globally unique and persistent identifier. It is the address of an Ethereum smart contract known as proxy contract. The identity is using this proxy contract to interact with other smart contracts on the blockchain.

Smart Contracts

Smart contracts are supported by uPort, especially because also the underlying Ethereum blockchain supports it.

[15] AWS (Amazon Web Services) https://aws.amazon.com/

3.4 Evaluation Results

This section summarizes the evaluation results shown in Table 1. Technology Evaluation Results Sovrin is the most promising technology because of its key management, no expensive proof-of-work has to be calculated and the support of identity data import even though some trust is required. All other technologies require the proof-of-work calculation, which is a huge disadvantage. Besides Sovrin is uPort the only other technology that supports identity data import.

	Sovrin	Blockstack	Multichain	Ethereum	uPort
Permissioned/ Permissionless	Permissioned	Permissionless	Permissioned	Permissionless	Permissionless
Mining	No	Yes[16]	Yes[17]	Yes	Yes[18]
Key Management	DPKI[19]	DKM[20]	Wallet or External	Ethereum Node	User Device and DPKI
Identity Data Import	Yes	Self-Entered	Self-Entered	Self-Entered	Yes
Selective Disclosure	Yes	Yes	Yes	No	Yes
Data Storage	On- and Off-Ledger Storage	Existing Storage Infrastructure[21]	Multichain Blockchain	Ethereum Blockchain and DOS/DFS[22]	Off-Chain Data Store[23]
Trust Required	Yes	No	No	No	No
Smart Contracts	No	No	No	Yes	Yes

Table 1. Technology Evaluation Results

[16] Indirectly, by using the underlying Bitcoin blockchain.
[17] Only a lightweight approach is being used.
[18] Mining is performed by the underlying Ethereum blockchain.
[19] DPKI (Decentralized Public Key Infrastructure)
[20] DKM (Decentralized Key Management)
[21] Such as Dropbox, Microsoft Azure, Personal Drive and more.
[22] DOS/DFS (Distributed Object Storage/Distributed File System)
[23] Such as IPFS, Microsoft Azure, AWS, Dropbox and more.

4 SSI Potential

This section identifies and details the technical potential of the SSI technology and points out possible use cases.

4.1 Technical Potential

The SSI technology offers new technical innovation potential when applying it as identity management system.

4.1.1 Extending Trust Models

A SSI system decreases the required trust, compared to traditional identity management systems, by using the distributed ledger technology in Section 2.1. Nevertheless, the trust in such a system can be elevated to an even higher level by combining the SSI system with current trust schemes such as the EU TSL (trust service list) or the Web-of-Trust.

4.1.2 Decentralized Public Key Infrastructure

A SSI system can be seen as the implementation of an open decentralized identity layer that includes a decentralized public key infrastructure (DPKI) based on the decentralized ledger technology (DLT). The DPKI differs from the well-known public key infrastructure (PKI) by not depending on central certificate authorities, such as authorities for issuing certificates (CA) or registration authorities (RA). The dependence from these authorities can be eliminated by changing the root of trust from the authorities to the identity owner.

Central authorities have the characteristic that their failure can cause critical consequences for the user.

1. **No single point of failure.** Without the need of a central CA or other registration authorities, there is no single point of failure anymore, which could cause severe problems for citizens.
2. **Interoperability.** The SSI system uses various methods, like not relying on proprietary software, to increase the interoperability significantly.
3. **Resilience.** Combining the decentralized architecture with cryptographically verifiable data increases the resilience of such a system.
4. **Key recovery.** DPKI offers the opportunity to build robust key recovery systems by using a combination of key escrow services and social recovery of keys shared across trusted DPKI connections.

4.1.3 GDPR Compliance

On 25th May 2018, the General Data Protection Regulation (GDPR) [13] will enter into force in all European Union member states. The GDPR aims to increase data protection and data privacy of each individual citizen within the EU.

A SSI system can support the GDPR compliance when dealing with eIDs. The following paragraphs detail how the SSI system can fulfill certain GDPR articles.

Consent

The GDPR mentions that a user must give explicit consent in order to process and collect the user's data. A SSI system can be extended by a, for the user and its experience optimized, graphical user interface a so-called personal identity data management system (PIMS). This PIMS will provide fine granular control over what data are being shared together with revocation mechanisms. Some Blockchains support smart contracts, which could be used to enforce the user's consent decisions.

Pseudonymization

The GDPR describes pseudonymization as process, where citizens' personal data are being transformed in such a way that the resulting data cannot be linked to a specific person without providing additional information.

In the SSI system, only identifiers are stored in the Blockchain. Those identifiers are cryptographically generated and cannot be linked to a specific person. An idea is to extend the SSI system and introduce qualified Anonymity by generating service provider and sector specific identifiers associated to a citizen. This is achieved by pairwise creating identifiers where a specific identifier is associated to each single counterpart. Different techniques can be used for the creation process, depending on data protection considerations. This way, a SSI system fulfills the pseudonymization requirements of the GDPR.

Right to Erasure (Right to be Forgotten)

The GDPR's right to erasure describes the right of a EU citizen to request the deletion of personal data. In a SSI system, the user fully owns the identity data; therefore, the user can simply delete the whole identity and its related data. The erasure is realized by not storing any private data in a public accessible place such as the decentralized ledger itself. Instead, only identifiers, linked to identity data, are stored in the ledger. The access

relies on the user's consent and is enforced by SSI consent mechanisms. The consent mechanism also enables the revocation of previously granted data access.

Records of Processing Activities

Maintain a record of processing activities that includes variety of information such as the processing purposes, the involved categories as well as envisaged time limits. The design of the SSI system provides opportunities to realize this.

Data Portability

In the GDPR, the right of data portability describes the right that a person within the EU is able to transfer personal data from one place to another. The SSI system supports this right by providing an open identity layer for the Internet, which offers the possibility to access and use it worldwide. This possibility is enabled because the SSI system combines different technologies that enable the data portability such as the distributed ledger with standardized data exchange formats such as XDI[24].

Data Protection by Design and by Default

Data protection by default means that the data protection mechanisms are already a part of the system's design, which is the case in the SSI system. This includes that, by default, appropriate technical and organizational measures should ensure the protection of the processed personal data related to a citizen within the EU.

The SSI system implements state-of-the-art techniques for both to preserve the user's privacy as well as to protect the processed data. One of these state-of-the-art techniques are Zero-Knowledge Proofs (ZKP), which allow an identity owner to prove the correctness of identity specific elements without revealing any unnecessary additional information. For instance, by proving the possession of a driving license without disclosing the complete driving license.

4.1.4 Identity Derivation from existing eID Infrastructure (eIDAS)

Another big possibility is offered by the SSI system when it comes to identity derivation. With extending the SSI system, it should be possible to derive identity data from existing

[24] XDI (eXtensible Data Interchange) is a standard format and protocol for data interchange developed by OASIS. https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xdi

eID infrastructure such as eIDAS network. This is realized by transforming the identity assertion into the SSI format. The big advantage when communicating directly with eIDAS is to have only one transformation method instead of deriving identity data from different member states of the EU, where each state would require its own identity translation. This way, the eIDAS infrastructure is elevated to a global scale by generating a global Self-Sovereign identity.

4.1.5 Qualified Self Sovereign Identity

The eIDAS regulation defines the requirements for a high Level-of-Assurance (LoA), which is required for authentication to the eIDAS node in order to receive qualified identity data. With the elevation of SSIs to qualified SSIs, the system should be able to provide qualified identity attributes. Thus, the system requires a high LoA according to eIDAS for authentication.

4.1.6 Verifiable Claims

A new concept that is introduces by the SSI system are verifiable claims. A claim is an attribute related to a specific person. Verifiable claims are non-reputable sets of statements made by an entity about another entity. These claims are cryptographically generated. The W3C group has a working group on verifiable claims[25].

For instance, a verifiable claim could be issued by a University, which affirms that a related person is holding a degree of this University or a healthcare provider that provides medical attestations. Electronic health data are in particular even more sensible data, which are specially mentioned in the GDPR.

4.2 Possible Use Cases

Utilizing a SSI system helps to realize many different use cases. SSI was intended to build the missing identity layer on the Internet. This way, SSI enables various use cases especially when dealing with eID in a cross-border context. A small excerpt of the possible use cases together with a short description is shown as follows.

[25] https://www.w3.org/2017/vc/charter.html

4.2.1 Creating a Student Bank Account

A student, named Alice, wants to create a bank account at bank B. Bank B offers a special student account without any annual fees and other benefits for students. To be able to apply for this type of account, the student must prove that she is registered as student at University C.

To do so, the Bank B requests claims from Alice such as the name, birthdate and address. Alice owns a Self-Sovereign identity (SSI), which can be used to provide these claims. B requires an additional claim, which proves that Alice is registered as student as C. C issues this verifiable claim for student Alice. Next, Alice gives her consent to provide the requested claims to the bank. After the bank has successfully verified the claims, the bank account for Alice can be created.

4.2.2 Applying for a Job

A person, named Bob, wants to apply for a job at company XYZ. The job requires that the applicant holds a master's degree of some related study. Therefore, the employee can request the claim that includes this information from Bob. Bob has a verifiable claim, issued by the University C that attests Bob's degree at University C.

Bob has to give his consent in order for XYZ to receive the verifiable claims. XYC can then verify these claims.

During the application process, XYZ figured out that they need additional information namely the certificate of the master program including the grades from their applicants. If Bob gives his consent, the University can issue these verifiable claims and the company XYZ will receive those afterwards.

4.2.3 Privacy Preserving Claim Attestation

This subsection describes in two short use cases how privacy preserving claim attestation is working. The first use case describes the request of special services related to physical disabilities and the second describes proving the age of majority.

Use case requesting special services:

A student named John with physical disabilities requires special services from his University. The University is located in a different member state of the EU. When John

requests these special services from the University, he provides medical attestations provided by his health operator. These attestations can contain different information such as that John might have reduced mobility capabilities but without revealing additional sensitive information such as his disease.

Use case prove age of majority:

When John moved to the city where the University is located, he must prove his age of majority at a public authority. John provides a verifiable claim that attests his age is over 18 without revealing the actual birth date.

The above-described use cases are only two out of a variety of scenarios where privacy persevering claim attestation would be important.

5 Conclusion

In this work, we have described the concept of Self-Sovereign identity, performed a technology evaluation based on criteria related to an SSI system as well as identified the potential of such a SSI system when applying it as identity management system.

The criteria were derived from SSI system related requirements and used to evaluate related blockchain technologies such as Sovrin, Blockstack, Multichain, Ethereum and uPort. Other technologies were considered for the evaluation but did not fulfill the minimum requirements to be evaluated.

Out of all the evaluated technologies, Sovrin checks most of the boxes when developing a Self-Sovereign identity system. Sovrin got the best result especially because its design is made to realize a SSI system. Additionally, the documentation provided by Sovrin made the evaluation much easier.

In contrast, Blockstack, Multichain, Ethereum and uPort did not match as much criteria compare to Sovrin. Ethereum offers a powerful platform to develop distributed application using smart contracts. However, the platform itself would only provide the basis.

uPort is an application built on Ethereum, which focus on Self-Sovereign identity. Nevertheless, uPort is in a very premature stage where there is only an alpha version available.

Blockstack focus on decentralized domain name service by combining different layer and parts of already existing infrastructure such as Bitcoin blockchain for the blockchain layer or Dropbox for the storage layer. The usage for SSI might be more difficult than using Sovrin.

Multichain provides a platform that allows running several different private blockchains parallel. It also provides a built-in permission management system. Nevertheless, the focus here is not directly on SSI and the usage for such a system might be easier using Sovrin.

Finally, technical potential as well as possible use cases of a SSI system have been identified and discussed.

Summarizing, the SSI technology brings various opportunities and large potential when applying it as identity management system but the challenges have to be considered as well.

References

[1] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," *Www.Bitcoin.Org*, p. 9, 2008.

[2] A. Marsalek and B. Prünster, "TECHNOLOGIEÜBERBLICK BLOCKCHAIN," pp. 1–27, 2016.

[3] N. Popper, "How China Took Center Stage in Bitcoin's Civil War - The New York Times," 2016. [Online]. Available: https://www.nytimes.com/2016/07/03/business/dealbook/bitcoin-china.html?_r=0. [Accessed: 05-May-2017].

[4] D. Reed, J. Law, and D. Hardman, "The Technical Foundations of Sovrin A White Paper from the Sovrin Foundation," no. September, 2016.

[5] M. Ali, J. Nelson, R. Shea, B. Labs, and M. J. Freedman, "Blockstack: A Global Naming and Storage System Secured by Blockchains."

[6] M. Ali, J. Nelson, R. Shea, and M. J. Freedman, "Bootstrapping Trust in Distributed Systems with Blockchains," vol. 3, no. 41, 2016.

[7] G. Greenspan, "MultiChain Private Blockchain - White Paper," pp. 1–17, 2013.

[8] B. S. Contracts and K. Bharadwaj, "The Origins of Smart Contracts," pp. 1–7, 2016.

[9] C. Lundkvist, R. Heck, J. Torstensson, Z. Mitton, and M. Sena, "UPORT: A PLATFORM FOR SELF-SOVEREIGN IDENTITY," 2017.

[10] "Nick Szabo - Smart Contracts: Building Blocks for Digital Markets." [Online]. Available: http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart_contracts_2.html. [Accessed: 12-Oct-2017].

[11] B. Drummond Reed *et al.*, "DID (Decentralized Identifier) Data Model and Generic Syntax 1.0 Implementer's Draft 01 A Draft from Rebooting the Web of Trust III Design Workshop," no. November, 2016.

[12] Multichain, "External key management | MultiChain," 2017. [Online]. Available: http://www.multichain.com/developers/external-key-management/. [Accessed: 25-Apr-2017].

[13] The European Parliament and The European Council, "General Data Protection Regulation," *Off. J. Eur. Union*, vol. 2014, no. October 1995, pp. 20–30, 2016.