ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

UNIVERSITY OF PIRAEUS

University of Piraeus

School of Information and Communication Technologies

Department of Digital Systems

Postgraduate Program of Studies

MSc Digital Systems Security

Self-Sovereign Identity using the Hyperledger Indy framework

Master Thesis

Supervisor Professor: Christos Xenakis

Course:Network Security

Abstract

Blockchain is a peer-to-peer network that ,unlike traditional widely used networks, does not require a central authority in order to function. It operates in a trustless way. This means that the nodes in such a network do not have to be specifically permitted or validated by a central authority to participate in it. The network is assumed to be trustworthy because each transaction is validated by the whole network. This ensures trust since an adversary would have to seize control of the majority of the network nodes in order to enforce his authority in it. This thesis leverages the capabilities offered by the Hyperledger project to create a Self Sovereign Identity management solution. Such a management solution could provide users the ability to choose which of the her/his identity attributes will share with specific service providers in order to access their services.

Contents

1 Introduction 1

1.1 Research Structure 1

1.2 Introduction 1

1.3 Problem specification 2

1.4 Aims and Objectives 3

2 Background and literature review 4

2.1 Problems with current trust system 4

2.2 Privacy Protection in traditional system 6

2.3 Privacy Preserving Computation(PPC) 9

2.4 Blockchain Technology 10

2.5 Self Sovereign Identity 13

2.6 Hyperledger Indy 15

2.7 Related Work 16

3 Security and privacy consideration on blockhain 21

4 Application demonstration 25

4.1 Requirements 25

4.2 Trinsic demo 26

4.3 Application Review 35

5 Conclusion 37

List of Figures

1 Figure 1, PKI overview[16] 4

2 SMPC model,[29] 10

3 Merkle Tree,[34] 12

4 Example DID[42] 14

5 Four layer SSI[46] 16
6 Privacy preserving techniques for blockchian[38] 22
7 Signing of attributes with DID with extra proof[55] 23
8 High level overview 25
9 Faber credential template 27
10 Faber credential schema ledger transactions 28
11 Faber credential template part of json call 28
12 Verification Template 29
13 Issued Credential 30
14 Available connections 31
15 Job application template 32
16 Job Certificate 33
17 Job Application 34

List of Tables

1 k-2 anonymous 6
2 k-1 anonymous 6
3 Pre anonymized data [26] 7
4 k-4 anonymous[26] 7
5 k-4 anonymous[26] 8
6 0167-closeness wrt Salary and 0278-closeness wrt Disease.[27] 9
7 Blockchain based comparison. Satisifed(1) and Non Satisfied(0) criteria[53]. 20
8 Non-Blockchain based comparison. Satisifed(1) and Non Satisfied(0) criteria[53]. 20

1 Introduction

1.1 Research Structure

This research comprises of 5 chapters. First chapter serves as an introduction. A brief mention of the blockchain history and how blockchain technology can help solve some security and privacy issues current technologies suffer from. The second chapter includes in depth information about current trust systems ,models,blockchain technology and the Linux Foundation Hyperledger framework[1] features. In the third chapter security considerations regarding blockchain technology are discussed. Fourth chapter includes a demonstration of the Self Sign Sovereign Identity(SSI) application using Hyperledger Framework. The last chapter acts as the conclusion to the research.

1.2 Introduction

Blockchain was first used in 2008 by an individual or group of people under the name Satoshi Nakamoto to develop Bitcoin.[2]. The first description of bitcoin is found in 1998 on the cypherpunks mailing list by Wei Dai[3]. Cypherpunks is a group of people that formed in 1992 by Eric Hughes, Timoty C May and John Gilmore.[4] and aimed to promote privacy and security. Some of the individuals who joined this group include Philip Zimmermann[5],creator of PGP 1.0, Brunce Schneier[6], Julian Assange, founder of WikiLeaks, and more.[4]. Adam Black , also a cypherpunk member, invented Hashcash[7] which was used as a countermeasure for denial-of-service attacks and email spam. Nowadays a variance of Hashcash is used as the mining algorithm for Bitcoin[7]. It employed a proof of work algorithm which will be discussed later in this research.

As of 2021 the total value of existing bitcoins is over 800 billion dollars.[8]. There are even search engines that allow users to search businesses that accept bitcoin as currency[9]. Universities like the University of Nicosia(UNIC) accept bitcoin as payment for tuition fees[10]. The increase in popularity attracted the attention of bigger cooperations and technology entities not only to bitcoin as a currency but to the underlying technology, blockchain, as well. The Linux Foundation maintains the Hyperledger project[1] with the aim of developing frameworks, tools and libraries to facilitate enterprise leverage blockchain technologies. One example of a global cooperation actively using and developing applications with blockchain technology is IBM who developed Food Trust. It is used by companies such as Nestle[11] and Walmart Inc[12] to enhance their food supply chain services. The innerworkings of blockchain technology will be discussed in chapter 2.

1.3 Problem specification

In today's world online trust is depended on third parties. Apart from SSI there is no way for an individual to prove its online identity the same way he/she do for his/her offline identity. In order for users to authenticate themselves to different services they have to register to them using usernames,passwords or other more modern and user friendly authentication mechanisms like FIDO usb keys[13]. One of the biggest problems with this implementation is that individuals' details are scattered across multiple databases across multiple organizations.

In order for digital credentials to be verified two steps have to be performed. Their format need to be standardized following the same thought process behind passports. Making it standardize facilitates the process of adopting the technology globally.Secondly the mechanism through which such digital credentials are verified should also be standardized. The current answer to this problem is the Public Key Infrastructure(PKI) which is not ideal, with its current implementation,regarding blockchain technology because it is centralized. Blockchains rely on three main pillars. Each transaction is digitally signed, linked to a previous transaction all validated transactions are replicated in all entities which use an agreed upon consensus algorithm. This way the information becomes immutable since it can at any time be revalidated across many machines. This way there is not a single point of failure in any system leveraging blockchain technology making it more redundant and secure.[14]. Blockchain has applications in a wide variety of use cases, from SSI to privacy preserving threat sharing. This report focuses mainly on the SSI applications but notes from other use cases will be mentioned to further solidify the potential advantages of blockchain.

1.4 Aims and Objectives

The research it aims to answer the following questions:

- What is a blockchain and how does it work?
- What are the advantages and disadvantages of blockchain over current used technologies?
- What are the security considerations with blockchain?

Simultaneously it aims to demonstrate how using the Hyperledger Indy framework, which will give a user the ability to choose which of his/her identity attributes (e.g. age) will share with specific service providers in order to access certain online services and resources.

2 Background and literature review

2.1 Problems with current trust system

Network trust nowadays relies on some selected trusted authorities to verify different entities. The centralized nature of it can potentially pose a great threat. The core trust mechanism of the internet is the Public Key Infrastructure(PKI). PKI certification mechanism is described in RFC5280[15]. The Certification Authority(CA), responsible for revoking and validating the certificate status, the registration authority(RA), optional component that when used is responsible to register new certificates on the request of entities and transfer them to a CA. The repository which stores and distributes the certificates. Also there is the Certification Revocation List(RCL) issuer which is the CA in most cases. Certificates are based on the x509 format[15]. It is apparent that since CAs are the central trust providers that everybody else relies on , for verifying if a service or node is trustworthy, it means that if a malignant entity gains control over them, it can directly control the trust in the part of the network a CA is responsible for.

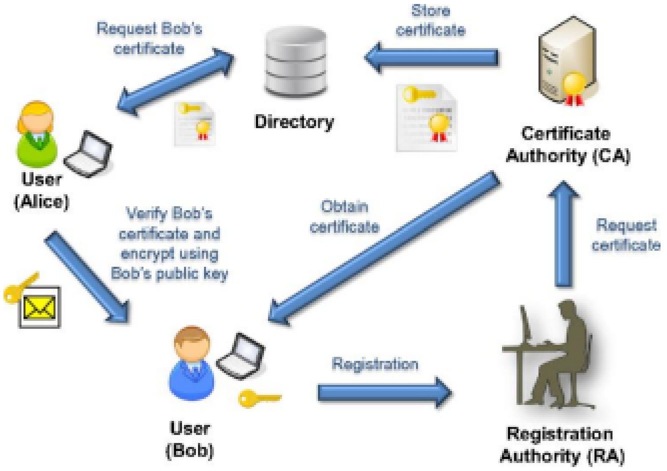

Request Bob's certificate
Store certificate
Directory
Certificate Authority (CA)
User (Alice)
Verify Bob's certificate and encrypt using Bob's public key
Obtain certificate
Request certificate
Registration
User (Bob)
Registration Authority (RA)

Figure 1: Figure 1, PKI overview[16]

An incident that proved that a centralized system like PKI can case major problems occurred in 2011[17] [18]. Diginotar was a CA based in Netherlands. According to the Fox-it report[18]. The attacker managed to gain administrative privileges on hte servers hosting the certificates and was able to produce forged certificates, drastically disturbing the service and being able to perform Man in the middle attacks abusing the certification mechanism. Reportedly the threat actor had access to over 300.000 Gmail accounts through this attack. Enisa points out that sine Diginotar had no records of the rogue certificates the ony viable solution was the removal of Diginotar root certificate from all browsers [17]. The incident resulted in the bankrupsy and dissolve of Diginotar. This incident was a clear indication that technology constantly needs to be challenged and evolved. The idea challenged in this case, apart from the usual security requirements when setting up a service, is the acceptance of a single entity having full control of the trust in a network or part of it.

According to Enisa[19] adversaries are steering their attention towards supply chain to indirectly compromise systems, which is what happened in the Solarwinds' attack. which challenged the idea behind a central authority. There is still limited information regarding how the initial breach occured[20] [21]. The Advanced Persistent Threat(APT) behind the attack was able to stay dormant, undetected for 14 days, scan the infected network for sandboxes and selected domains to avoid running on and the proceeded to attack selected networks, infecting Orion update software and spread to Solarwinds' clients. Solarwinds' clients include but not limited to Microsoft, Intel, Cisco, Nvidia and more. These vendors also provide software and/or hardware solutions to many organizations which can be described as critical infrastructures under the Network and Information Systems(NIS) directive. Concidering the level of intrusion, the APT could spread to the critical infrastructures in Europe and globally. It could potentially cause great disturbance of service and buisiness continuity in several critical sectors such as Health, Power, Transport and more.Although it was not directly connected to the PKI , like the Diginotar incident, this scenario proves once again that difficulties in sharing information between entities working towards a common goal, in this case cybersecurity firms, Solarwinds and its customers, hinder the ability of collaborative ability of professionals to combat such attacks. Vasu Jakkal, Microsoft's corporate vice president of security, compliance and identity told to ZDNet that the industry needs to act together against the evolving threats. The industry has to collaborate and find ways to share information in privacy respecting way across private and public sectors[22]. The events described in this section are very strong indications that blind trust towards a centralized authority and being unable to communicate sensitive yet critical information fast thought a secure, privacy preserving process is crucial to fortify services and develop a safer environment. Blockchain technology could allow the industry overcome these obstacles. Information is valuable and ways must be found to use it and process it to everyone advantage with respect to security, privacy, ethics and regulations.

2.2 Privacy Protection in traditional system

Another factor that needs to be considered when dealing with sensitive information, in particular Private Identifiable Information(PII) is the General Data Protection Regulation(GDPR). According to Gartner by 2023 over 25% of GDPR driven proof of consent implementations will involve blockchain technology up from less than 2% in 2018 [23]. Blockchain technology is very promising for data privacy since it allows for immutability, transactions are cryptographically signed, decentralized information, no single node in the network can dictate what information the rest of the network has access to. In conjuction with blockchain other data privacy protection mechanisms can be used. Commas and Ferrer defined privacy models as ”privacy models specify conditions that the data must satisfy to keep disclosure risk under control”[24]. K-anonymity[25] is a privacy model indicating that a data set is k-anonymous if each record cannot be distinguished from at least k-1 other records regarding the quasi identifiers. Quasi-identifiers is a collection of attributes which when combined with external information can lead to identification of respodners. In simpler words if in a dataset for each row there is at least 1 more identical row then it is k-anonymous.

name	X	Y	Z	W
age	18	18	20	20

Table 1: k-2 anonymous

name	X	X	Z	W
age	18	18	20	20

Table 2: k-1 anonymous

As pointed out by Gehrke et al[26] k-anonymity is not sufficient for some cases since it does not provide attribute disclosure prevention. The relation and frequencies between values are preserved so correlations can be made if context is known. Gehre et al[26] aimed to improve upon k-anonymit with l-diversity. A dataset is said to have l-diversity of X when there are at least X " well represented values for each sensitive attribute". For example even if all sensitive information is removed from a dataset to satisfy k-anonymity it might still be possible to identify an individual if the leftover information in the groups is not diverse enough. If there is a group with X amount of people in specific location and through a k-anonymized dataset it is derived that all people in said location are patients with a particular disease then it is trivial to link the dataset information to real people, since if someone knows a person from this specific area he wil know with 100% certainty that he suffers from this disease.

	Non-Sensitive		Sensitive
	Zip Code	Age	Condition
1	13053	28	Heart Disease
2	13068	29	Heart Disease
3	13065	21	Viral Infection
4	13053	23	Viral Infection
5	14853	55	Cancer
6	14853	47	Heart Disease
7	14850	49	Viral Infection
8	14850	49	Virtal Infection
9	13053	31	Cancer
10	13053	37	Cancer
11	13068	36	Cancer
12	13068	35	Cancer

Table 3: Pre anonymized data [26]

	Non-Sensitive		Sensitive
	Zip Code	Age	Condition
1	13***	<30	Heart Disease
2	13***	<30	Heart Disease
3	13***	<30	Viral Infection
4	13***	<30	Viral Infection
5	14***	>40	Cancer
6	14***	>40	Heart Disease
7	14***	>40	Viral Infection
8	14***	>40	Virtal Infection
9	13***	3*	Cancer
10	13***	3*	Cancer
11	13***	3*	Cancer
12	13***	3*	Cancer

Table 4: k-4 anonymous[26]

	Non-Sensitive		Sensitive
	Zip Code	Age	Condition
1	1305*	≤40	Heart Disease
4	1305*	≤40	Viral Infection
9	1305*	≤40	Cancer
10	1305*	≤40	Cancer
5	1485*	>40	Cancer
6	14853	>40	Heart Disease
7	14850	>40	Viral Infection
8	14850	>40	Virtal Infection
2	13053	≤40	Heart Disease
3	13053	≤40	Viral Infection
11	13068	≤40	Cancer
12	13068	≤40	Cancer

Table 5: k-4 anonymous[26]

Another option is t-closeness[27] which employs mathematical formulas. In contrast to l-diversity it also takes into consideration the distribution of an attribute in a data set. t-closeness essentially calculates how ”close” is the anonymized dataset to the pre anonymized one.

$$E(p,q) = \frac{1}{m-1}\sum_{i=1}^{m} |\sum_{j=1}^{u}(pj - qj)|$$

	Zipcode	Age	Salary	Disease
1	476**	2*	3K	gastric ulcer
2	476**	2*	4K	gastritis
3	476**	2*	5K	stomach cancer
4	4790*	>40	6K	gastritis
5	4790*	>40	11K	flu
6	4790*	>40	8K	bronchitis
7	476**	3*	7K	bronchitis
8	476**	3*	9K	pneumonia
9	476**	3*	10K	stomach cancer

Table 6: 0167-closeness wrt Salary and 0278-closeness wrt Disease.[27]

The issue with the above mentioned models is that they are difficult to implement when different entities need to collaborate in computing a result. The solution to this problem is Privacy-Preserving-Computation(PPC).

2.3 Privacy Preserving Computation(PPC)

It is possible to process data without compromising privacy. PPC includes techniques like Secure Multi-Party Computation(SMPC).

SMPC enables different entities collectively perform computations while keeping their inputs and outputs private. It was first introduced by Yao[28]. It can be explained with the millionaire's problem. The problem describes 2 millionaires who want to know which one is richer without revealing their wealth to each other. The same logic applies to if for two given numbers one must determine if they are equal or not without revealing their actual values. An issue with SMPC is that the different parties need to provide their data on a trusted third party to avoid other participants extracting sensitive information. The trust is based on the parties action being in accordance to the protocol used. If a party doesn't comply it is deemed not trustworthy. Zhong et al[29] proposed a conceptual model to utilize SMPC via blockchain.

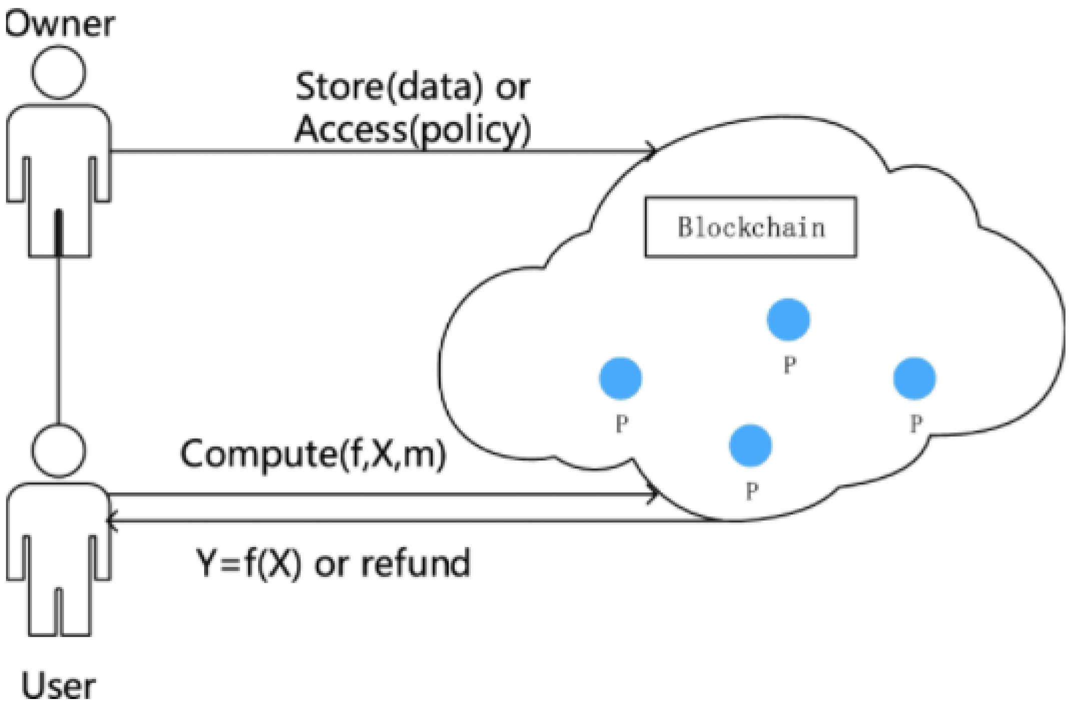


Figure 2: SMPC model,[29]

The data must be encrypted by the data owner before uploading it to the blockchain who also has to set access policy for said data. User wants to perform computations on the data. The user must give a deposit first to perform a computation. When the computation is completed successfully he gets the deposit back and the data. If it is interrupted he gets the deposit back. If a party is dishonest it receives penalties. The computation parties, represented as P in Figure 2, provide computing and storage resources. Other notable reports include the Zysking et al Enigma[30][31], a blockchain based computation platform and Choudouri et al[32] who used a bulletin board system based on Bitcoin to enhance fairness in SMPC.

There are other PPC enabling techniques like Homomorphic encryption but in depth analysis of such techniques are out of scope for this research.

2.4 Blockchain Technology

Previous section mentioned widely used models and techniques to preserve privacy in several cases. Technologies based on blockchain can help overcome obstacles present in current implementations. In this section the components of a blockchain will be analyzed.

Blockchain is based on transparency. At its core any participant can read all the contents, this can change if needed in certain use cases since the technology is flexible. The network operates on a trustless bases with consensus algorithms. The core components of a blockchain are the distributed ledger, a record of all transactions, which is the database essentially. A node application which is run on network nodes, for example the wallet containing a participant's funds is a node application. No single individual can tamper data since the blockchain infrastructure is based on digital signatures, distributed among network participants and transactions are timestamped. The consensus algorithm defines how all the nodes come to an agreement about transaction validity, rendering a CA redundant. When new transactions validated a new block is created containing the information. Each block consists of the block header, containing metadata like timestamp, hashed version of block data, previous block header hash and maybe a nonce[33].

The data is encoded using Merkle tree[34]. Merkle tree is efficient to use for verification purposes. For example in order to prove that transaction 1, shows as Data 1 in Figure 3, is included in the block only 1 item per depth is needed. In this case A and F blocks only need to be used. If transaction 1 was present then A=Hash(1) would get hashed with B=Hash(2) producing E=HASH(A,B). Then E and F will get hashed produced the Merkle root. If this ”new” Merkle root is the same as the one with transaction 1 in it then it means the transaction 1 was included.

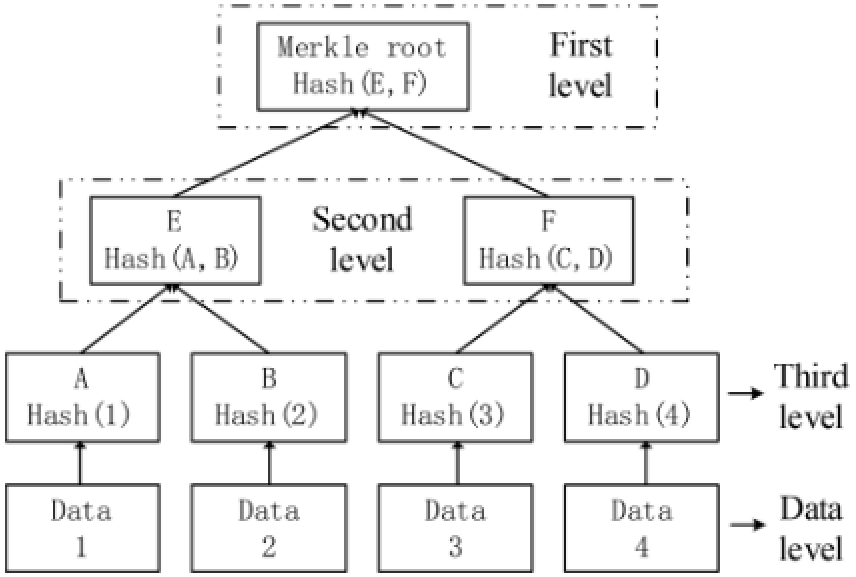


Figure 3: Merkle Tree,[34]

Consensus algorithms include but not limited to Proof of Work(PoW), Proof of Stake(PoS), Proof of Elapsed Time(PoET), Simplified Byzantine Fault Tolerant(SBFT) and Proof of Authority(PoA). Yao et al[35] performed a survey on Blockchain consensus mechanisms. PoW, used in bitcoin, demands the users to find the nonce in order to create a new block. The fact that significant processing power is required to calculate the correct value it means the blockchain is safe from malevolent entities since as long as they don't posses enough computing power ,to be concidered the majority in the network, they cannot arbitrarily add transactions to the block chain[36]. Hyperledger indy which will be used to implement the SSI application in chapter 4 is a consensus algorithm similar to Redundant Byzantine Fault Tolerance(RBFT)[37] called Plenum[38]. The basic idea behind RBFT is that when a message is sent to N amount of nodes, then each node has to propagate the message to other nodes in the network. As long as $f = \frac{N-1}{3} + 1$, >33%, of nodes are trustworthy then their consensus towards a decision can be used safety.In Plenum client requests are executed in batches. Each REQUEST message contains the requested operation, request identifier, client id all signed with the client's private key. A MAC authenticator is used on each node for each message. If the message is authenticated then it verifies the signature. Then the message is propagated with a PROPAGATE message. The receiving nodes verify the siganture and if valid propagate the message again to other nodes. In RBFT each node, they are refered to as replica in the paper[37], run an instance of the BFT algorithm acting as the Master protocol instance. Performance indicators dictate if the primary replica is malicious so a new view state is initiated to change it. After that each protocol instance replica performs a three phase commit before answering back to the client. A PRE-PREPARE message is sent from the primary to all other replicas, which in turn store the message. After verifying the MAC the replica sends a PREPARE message to all other replicas if the node received f+1 same request copies. This allows for better performace since clients maliciously targeting 1 node are unable to abuse it. After receiving 2f PREPARE messages from differente replicas, the node sends a COMMIT message. After receiving 2f+1 COMMIT messages from individual replicas, the node returns the ordered request. This is how requests are ordered. Finally the nodes execute the reqeust and reply to the client with a REPLY message. When a client receives f+1 valid replies from distinct nodes, it accepts the result[35].

2.5 Self Sovereign Identity

Self Sovereign Identity(SSI) aims to allow individuals and organizations be in complete control of their identity. As described earlier the current trust system is based on centralized entities. Users put a lot of trust in service providers to keep their sensitive data safe. The information stored is not always used solely for the reasons users agreed up initially. More often than not the information is used by the company, holding the data, in collaboration with marketing firms, other organizations with political motives to abuse the data and enforce their agenda. Famous incidents involving personal data abuse include the Cambridge Analytica incident,[39], the Yahoo incident[40] in 2013, the Facebook one in April 2019[41] and more. The concentration of sensitive information in select databases has proved to be a desirable target of adversaries. Another problem is that most of the times a user can't reuse his/her identity with other service providers. The fact that one user is verified on X company doesn't necessary allow him to re use this trust to other services. This in turn forces users to maintain multiple accounts which come with its own problems, for example regarding username and password memorization etc.

SSI proposes a different approach to digital identities. In SSI the identifiers are exchanged directly between the parties without the need for an central authority. The most fundamental part of the implementation are the Decentralized Identifiers(DIDS). World Wide Web Consortium(W3c) standardizes DIDs[42] to facilitate their adaptation. Standardization always was and always will be a major step towards any wide adoption of any technology. A user generates a DID then stores it in the ledgers. DIDs , in Figure 4 example did:example:123, *did* is the scheme, *example* is the method and *123*is the identifier. The DID subject is the entity identified by the DID. DID controler has the capability, defined by DID method, to make changes to the DID document. DID documents can have more than 1 controllers.[42]. DID documents are the resolve result of DIDs. DID documents information about the DID such as the id, public keys, services etc. The verifiable data registries are where the DIDs are stored. These could be distributed ledgers, decentralized file systems, databases etc[42].

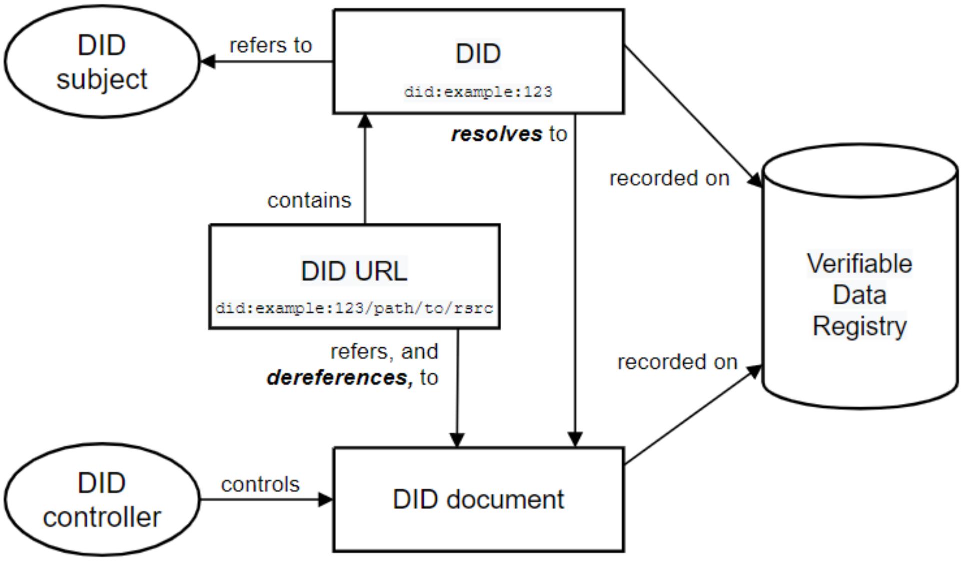


Figure 4: Example DID[42]

An example peer communication between peers follows[43]:

1. Alice has her private key and a DID document from a service with an endpoint and a public key.

2. The service has a private key and a DID document from Alice with her public key.

3. Alice encrypts her message, signs it and sends it to the endpoint.

4. Service receives the message and verifies signature with Alice's public key.

5. If the signature is verified the message can be read.

Another significant component to build a reliable SSI system are the verifiable credentials[44]. The roles in verifiable credentials are the issuer, holder, verifier and the verifiable data registry. Issuers can assert claims regarding subjects by creating a verifiable credential from these claims. A holder is essentially the subject or organization the credential is regarding. A verifier, as the name suggests verifies the authenticity of a credential. The verifiable data registry can be considered to be the ledger. In order for the credential to be valid there are 5 steps. First the issuer's DID must be resolved to a DID document on the ledger, containing the public key to check the integrity. The holder generates a zero-knowledge proof for his/her claim. The issuer needs to have the authority to release DID for the specific credential, an economic Ministry does not have the authority to issue a DID about vaccination certificate for example. The credential must not be present in the revocation registry. Lastly it is verified that the credentials provided meet the authorization requirements. The communication is done with DIDComm standard[45].As an example the government issues a driver's licence to Spiros. Spiros is the holder of the licence. During a police check the officer becomes a verifier for the licence Spiros presents.

2.6 Hyperledger Indy

Hyperledger Indy is a distributed ledger built for decentralized identities[43]. Evernym[46] is one of the original founders of the Sovrin Network, a public permissioned blockchain. It means everyone can use the blockchain but only permitted entities,called Stewards, can run the validator nodes. There are also private, meaning only selected entities can participate like the IBM Food Trust implementation, and permissionless ledgers, like bitcoin where everyone can act as the miner-validator.They also made the first contributions to the

Hypeldger Indy probject which spawned Hypeledger Ursa and Hyperledger Aries. Hyperledger Ursa is the shared cryptographic library for all Hyperledger projects implementing cryptographic protocols such as CL-RSA signatures[47]. Hyperledger Aries is the protocol for peer to peer connections, wallet, messaging and key management. Hyperledger [46].

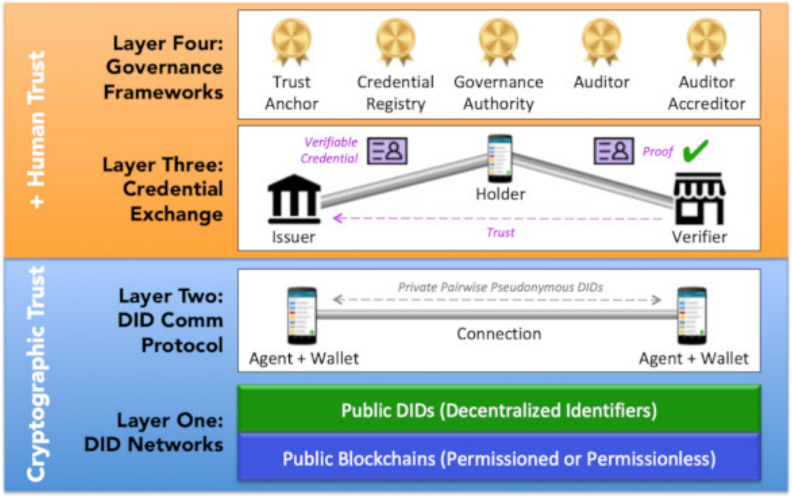


Figure 5: Four layer SSI[46]

Initially Indy covered the first three stack layers all the way from wallets(L1), DIDS communication(L2) to the Zero knowledge Proof(ZKP) credentials(L3). The problem was that it was confusing since it gave the impression to new developers that L2,L3 were tied to the underlying Indy blockchain code. More clearly dividing the 4 layers clarifies the fact that each of these layers is interoperable between SSI ledgers, verifiable credentials etc. So now Hyperledger Indy is responsible for L1 and Hyperledger Aries is responsible for L2,L3. Aries was the first implementation of open source wallets using Decentralized Key Management System(DKMS) architecture.

2.7 Related Work

Papadopoulos et al[43] combined blockchain technology with machine learning(ML) to create a proof-of-concept(PoC) that allows participants use Federated Learning(FL) to train models and share them with researchers without exposing sensitive information. They used the ledger developed by the British Columbia Verifiable Organization Network(VON)[48].

Their research main purpose was to prove that FL was able to run over Hypeledger Aries agents using DIDcomm protocol through a trusted network. This implementation can potentialy some problems with FL. Namely malignant entities can provide the network with corrupt data or the entity providing the models supplying malicious models to extract sensitive data. There are still potential threat such as the compromisation of in-house computers which in turn allows adversaries to create legitimate VCs and even perform DDOS attacks. The implementation proved resilient against simulated rogue agents attack scenario while also remaining relatively perfomant.

Kondova and Erbguth[49] research if blockchain can potentialy help organizations towards GDPR compliance Although some aspects of blockchain might indicate otherwise, data immutability for example. GDPR applies only to PII. According to Recital 26[50] a dataset must be examined for its ability to identify a person taking into account all the reasonable possible means available. In relation to blockchain technology credentials and revocations pools must be examined closely since they can contain such information. DIDs are created by data subjects with selected attributes in mind. Unless a DID specifically discloses the data subject's identity it is not possible to prove his/identity when looking at a DID in isolation. Konda and Erbguth tho mention that when used multiple times, in conjuction with metadata they carry like creation timestmap, can be used for data corellation. Commission nationale de l'informatique et des libertés(CNIL) does not concider the node operators as data controllers, the ones responsible for the data[50], but accept the fact there could be no data controller in that case[49]. On the other hand on permissions blockchains the entity determining permissions could be considered the data controller. An entity signing transactions with his/her private key can be assumed to be the data controller in the transaction phase[49]. CNIL concluded that there is no universal blanket statement for every use case thus they should be considered individually. Justification for data processing as described in Artical 6[50] can be covered possibly by consent Article 6.1, by law Article 6.2 or other articles but each use case must be examined ad hoc[49]. In regards to the right to be forgotten[50], Article 17, off-chain data can be deleted, data not present in the immutable ledger, so it depends on the implementation if it can comply or not.

Cameron[51] defined the 7 laws of identity regarding digital identity systems.

1. User Control and Consent: The systems must be designed with usability and simplicity in mind. The users must be in control of their identity to trust the system so the development of it should incorporate such features. The user's information must be protected and any access to that information should be clear to them.

2. Minimal Disclosure for a Constrained Use: The information stored in a system should be the minimum required for its operation to ensure minimum damage during a breach. Apart from the quantity of information, special care must be taken for the identifiability of it. For example if a individual's age is needed, instead of the full birth date it should store only the year of birth. Uniquely identifiable numbers, id numbers,licence numbers etc, that provide context are more valuable, therefore more dangerous to be breach targets, that non-context providing information. A unique id number provides more context than a randomly generated number to be used as an employee number.

3. Justifiable Parties: Disclosure of information to third parties should be justified and limited to only the information required for them to perform their strictly defined duties. Strict policies must govern the information relationship.

4. Directed Identity: Identity systems must support both omni-directional,public, and unidirectional,private, identifiers. A public website with a public certificate is a good example of a omni-directional identifier. No issue can occur if a user examines the Uniform Resource Locator(URL) or the certificate. On the other hand passport readers should be omni-directional since they do not inherently store any sensitive information while passports themselves should only allow be read by verified machines since they hold sensitive information.

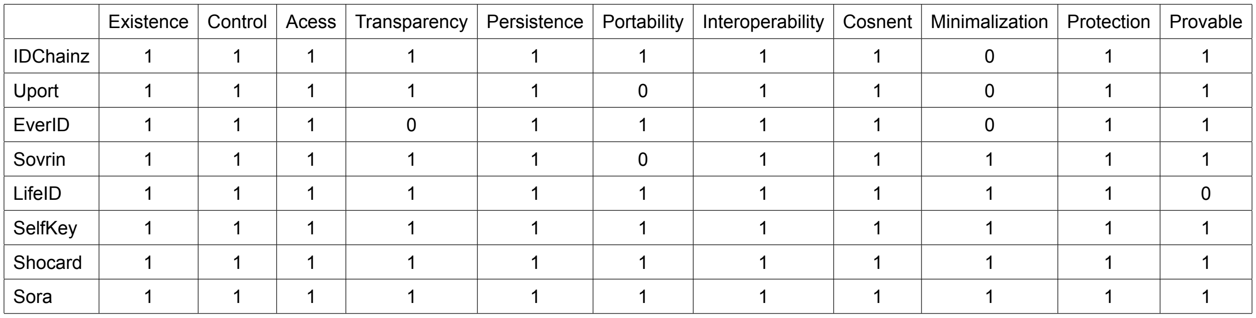
5. Plurarism of Operators and Technologies: There should not be a single identity system that provides identities for everything. As Cameron points out most people would not want to use a government issued certificate, like their id, to log in to their workplace systems. The characteristics and intended purpose of one certificate does not

apply to every use case. The only common denominator of all such systems should be technology protocols and standards all would work and agree on.

6. Human Integration: Humans should be seen as integral part of identity systems. There have been great strides made in the communication part between the browser and the server, via cryptography, but the most vulnerable part still remains, the part where a human interfaces with the browser. The communication procedure should be streamlined to the point the user knows exacly what to expect from the application when exchanging information. Cameron brings up the example of a pilot and air traffic control tower communicating. It is based on a well defined protocol with each party knowing exacly what to expect from the other.

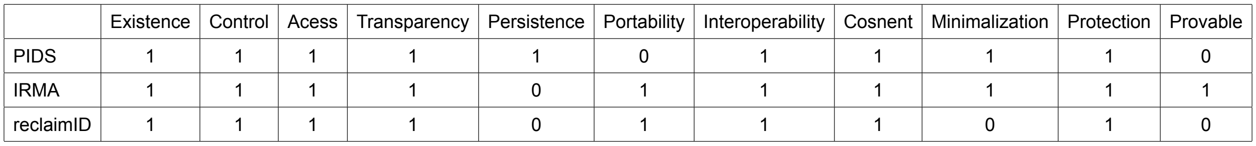
7. Consistent Experience Across Contexts: The user experience should remain consistent and simple across all contexts. As an example recall how computers become more usable after introducing modern User Interfaces(UI) and Icons. The user should be able to easily select which identity to use in a specific context. When trying to browse the internet for example the user should be able to select a less identifiable identity than the one she/he would use to perform bank transactions.

Allen[52] expanded on Camerons work[51] and provided a different perspective to the laws proposed by Cameron.Bokkem et al[53] analyzed several SSI solutions, blockchain based and non blockchain based, aiming to find out if blockchain is a necessity in building SSI solutions. They defined criteria,based on Allen's work[52], and measured implementations according to them. Existence is concidered satisified if each user is allowed to create her/his own identity account. All of the mentioned implementations are based on SSI so they satisfy this requirement. Control implies that users have full control of their identity during its existence with proper authentication. Access allows to trace the authenticity of users and the origin of data. It does not mean the data is public it just demands that it is possible to know where the data came from. In order for a system to satisfy Transparency it has to incorporate algorithms that are open source and well documented. Persistence requires a system to preserve identities while respecting the right to be forgotten. Portability demands that identidy data is not in possesion of a single third party. The reason Uport, although being blockchain based, does not satisfy this criteria is because Uport identities are only accessible to other Uport identities.In terms of Interoperability identities must be interoperable between different identity systems to allow the user use her/his identity as she/he sees fit. The Consent criteria is met when systems are build with respect to the users ability to know what exacly she/he consents to and to what extend her/his data is used, which she/he must first agree on. Minimalization is statisfied if the disclosure of claims are minimized to the absolute minimum when used. Protecton refers to the fact that users protection should always be top priority and favorited against other needs in the network. Lastly Provable criteria is met if it is possible to see what claims the trusted parties have verified.



	Existence	Control	Acess	Transparency	Persistence	Portability	Interoperability	Cosnent	Minimalization	Protection	Provable
IDChainz	1	1	1	1	1	1	1	1	0	1	1
Uport	1	1	1	1	1	0	1	1	0	1	1
EverID	1	1	1	0	1	1	1	1	0	1	1
Sovrin	1	1	1	1	1	0	1	1	1	1	1
LifeID	1	1	1	1	1	1	1	1	1	1	0
SelfKey	1	1	1	1	1	1	1	1	1	1	1
Shocard	1	1	1	1	1	1	1	1	1	1	1
Sora	1	1	1	1	1	1	1	1	1	1	1

Table 7: Blockchain based comparison. Satisifed(1) and Non Satisfied(0) criteria[53].



	Existence	Control	Acess	Transparency	Persistence	Portability	Interoperability	Cosnent	Minimalization	Protection	Provable
PIDS	1	1	1	1	1	0	1	1	1	1	0
IRMA	1	1	1	1	0	1	1	1	1	1	1
reclaimID	1	1	1	1	0	1	1	1	0	1	0

Table 8: Non-Blockchain based comparison. Satisifed(1) and Non Satisfied(0) criteria[53].

3 Security and privacy consideration on blockhain

Bernabe et al[38] performed a survey about privacy techniques in public and permisionles blockchains as well as permissioned and private blockhains. They identified several privacy challanges in blockchain scenarios as well as solutions to them. Multi-entry transcations require the user to have different addresses, with the intent being to obfuscate the users identity even more. Adversaries could potentially relate transactions with wallets resulting in discovering about balances, destinations and more. A solution to this could be the use of one-time addresses for every transaction blocking the traceability of them. Since they are one time use, in place of the actualy addresses, the formed are protected. During a transactions, in Bitcoin for example, it is possible for a transaction to have change which is charged back to the original address. As a solution a user could use a different address to charge the changes. Browser cookies can be used to link transactions and identities. Mixing services lke CoiJoin cannot be defend against this threat since it is cookie based. The mixing services themeselves are also liable because the service provides must be trustworthy. Mixing allows for transactions to mix together to enchance users privacy. The P2P communication of blockchain nodes can prove a threat as corellations can be made between nodes. To preserve privacy in blockchain mechanisms exist such as ZKPs, SMPC, HE all mentioned earlier in the report. Ring signatures is another privacy preserving mechanism. The idea behind ring signatures is that a function exists such as after computing a value with only the public keys then it is possible to sign a message with it using only 1 private key of the used public. All participating public keys are needed but only one private key. So it is possible verify a message from a certain group but it is impossible to know from which member it originated. For data anonymization in particular there are also tools available such as Mixing through protocols like CoinJoin. Diferential Privacy[54] is based on introducing random noise in a dataset but without interfering with statistical analysis performed on the entirety of the data in a significant degree.The full taxonomy follows in Figure 6.

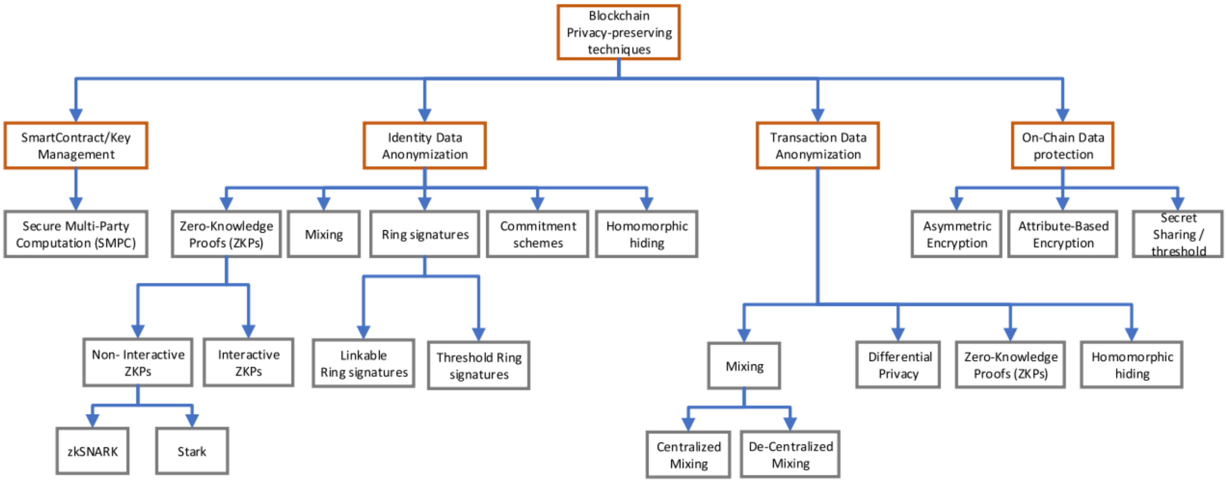

Blockchain Privacy-preserving techniques
SmartContract/Key Management
Identity Data Anonymization
Transaction Data Anonymization
On-Chain Data protection
Secure Multi-Party Computation (SMPC)
Zero-Knowledge Proofs (ZKPs)
Mixing
Ring signatures
Commitment schemes
Homomorphic hiding
Asymmetric Encryption
Attribute-Based Encryption
Secret Sharing / threshold
Non- Interactive ZKPs
Interactive ZKPs
Linkable Ring signatures
Threshold Ring signatures
Mixing
Differential Privacy
Zero-Knowledge Proofs (ZKPs)
Homomorphic hiding
zkSNARK
Stark
Centralized Mixing
De-Centralized Mixing

Figure 6: Privacy preserving techniques for blockchian[38]

Zavarsky et al[55] explored ways to enhance security and privacy of SSI specifically on Hyperledger Indy Blockchain in regards to MITM attacks. A MITM attack, during initial communication setup is possible, meaning the adversary can have access to both the request and the response,the proof, which means verifiable credentials fail. To mitigate that risk Zavarsky et al propose an extra proof is needed. Claims should be signed using the signing key of the DID original key-pair to verify that the proof was actually sent by the correct party. The addition of the extra proof affected performance but not in a significant way. They tested different credential numbers per test ranging from 50 to 500, increasing by 50 in each interval. The median value of all tests was 0.725 seconds per credential, without extra signing, and 0,78 seconds per credential, with extra signing.Figure 7 shows the DID exchange of two nodes.

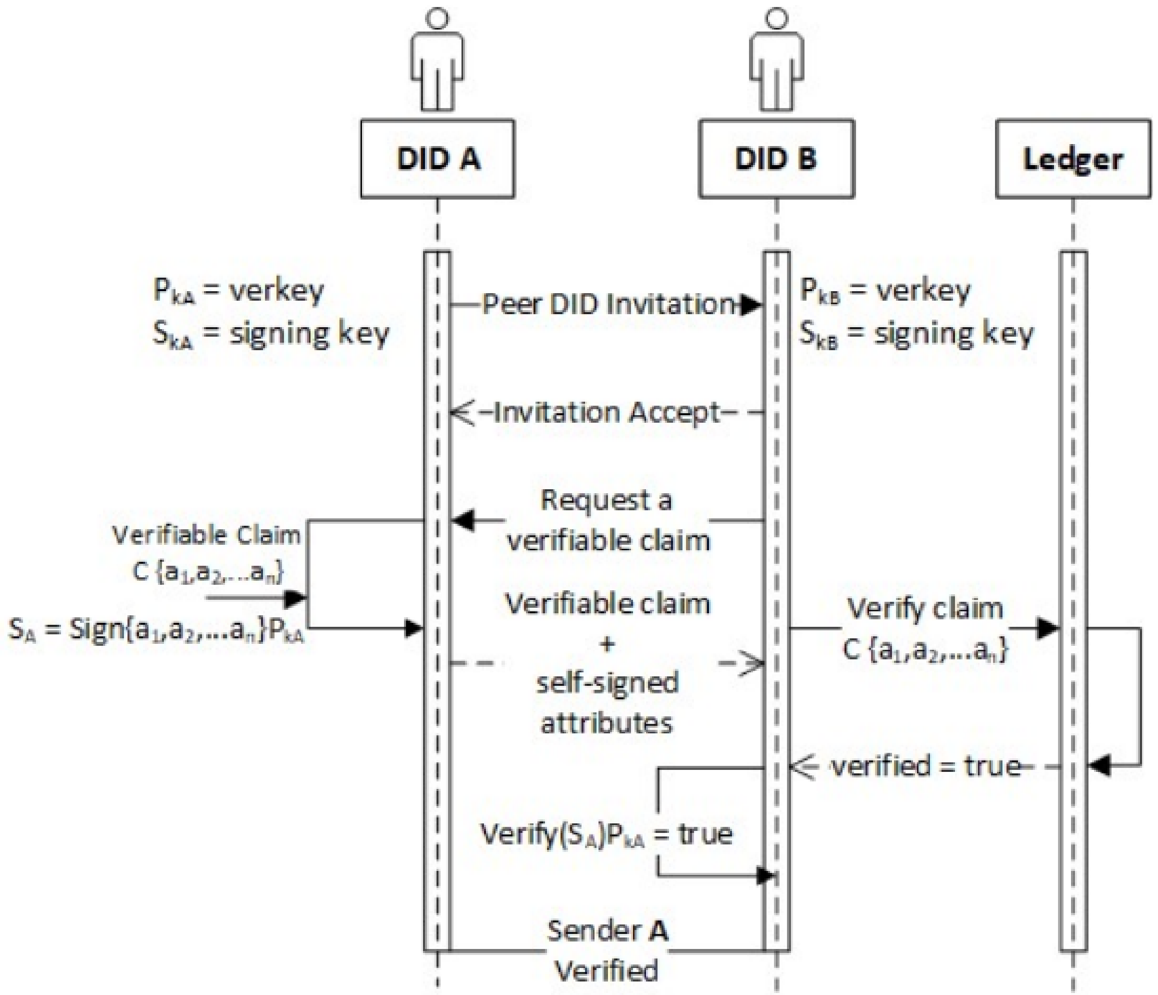


Figure 7: Signing of attributes with DID with extra proof[55]

They created a sample attribute categorization and gave each attribute a score to later determine their sensitivity. In Hyperledger Indy there are tree main entities, issuer, verifier and identity holder. To establish a DID connection, first two peers without prior knowledge of each other must exchange verifiable credentials. Zavarsky et al's[55] proposed a peer scoring system, based on Gruner et al's work[56], which can be helpfull in forming new relations or breaking existing ones if the confidence level drops. That proposal sets the maximum Trust level value as 1. Newly created identities are assigned a initial reputation score based aspects like compliance, certifications and more. Sovrin Foundation, operators of the Sovrin network and initial contributors to the Hyperledger Indy codebase, have an established governance framework that can be used as reference for these aspects[57].

Reputation ranges from 1 to 100 to better demonstrate fluctuations while remaining proportional to the Trust value. Reputation is depended from the initial reputation score and the average number of credentials issued in a set amount of time[56]. The confidence level is dependend on the trust and reputation levels. Gruner et al[56] defined 5 trust levels. Peers with in the No Trust bracket, $9 \leq t \leq 0.2$, up to Superior trust incrementing bracker borders by 0.2. A entity in the No Trust zone should only provide non-critical services such as approving comments from already existing users in a forum. The trust scales up to superior trust which means the entity is trused to run critical applications like opening a bank account. Their test results showed their model allows peers Trust to only increase marginally during high certificate volume days and decrease marginally during low volume days making it hard to exploit.

4 Application demonstration

In this section the Trinsic demo will be presented as well as a high level overview of the underlying infrastructure.

4.1 Requirements

Regarding demo perquisites, only a mobile phone is needed with the Trinsic wallet application installed and a Trinsic account. Trinsic website allows to quickly develop demos without requiring anything to be installed localy on a computer. All actions can be done via the user interface. The exact steps are described in section 4.2. Trinsic leverages the Hyperledger components including Indy, Aries and Ursa. In addition it uses Sovrin staging network as the blockchain. All these components were introduced in section 2.6. An overview follows in figure 8:

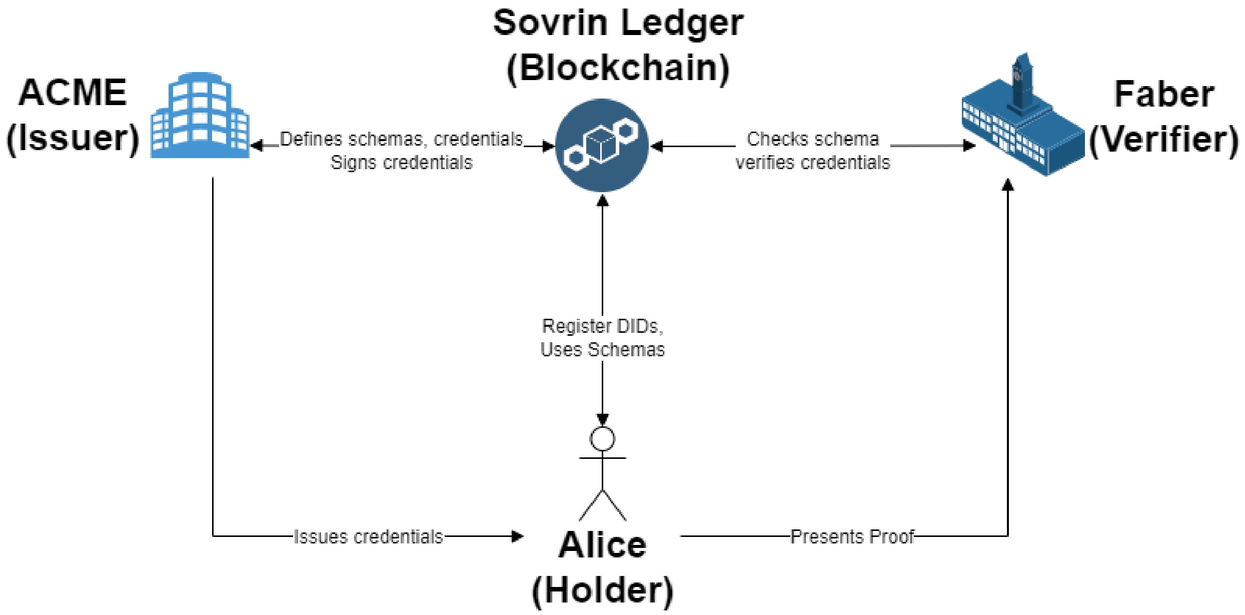


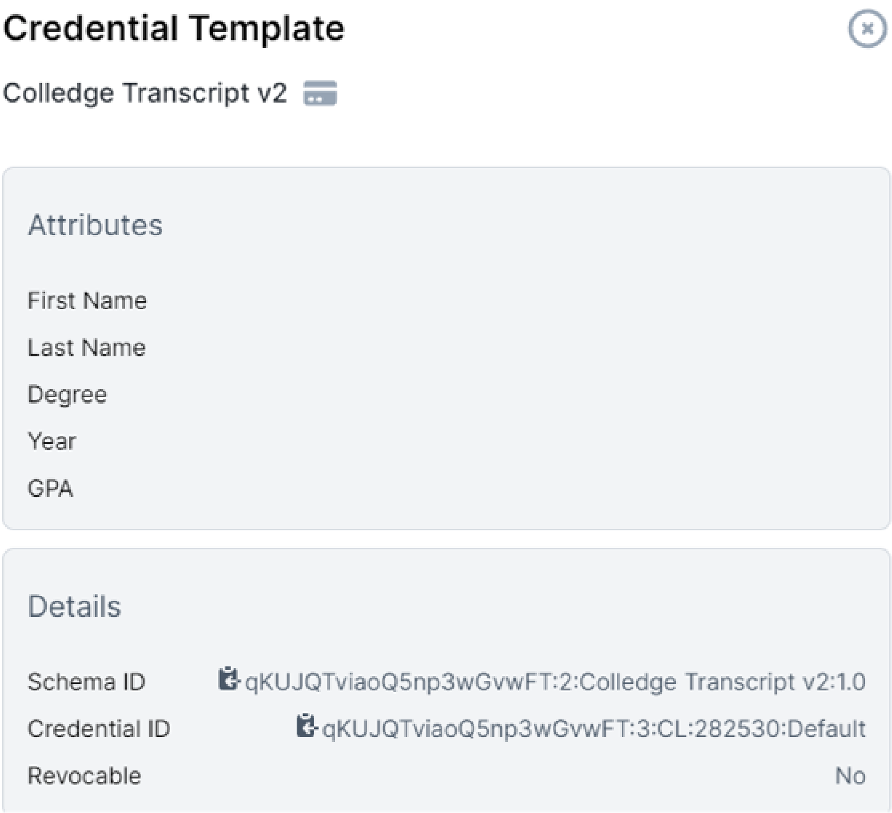
Figure 8: High level overview

Trinsic also provides SDKs in various languages including Python, Ruby, Javascript and .Net. Applications can be made for web, desktop and mobile. In this particular example a mobile phone application use case is demonstrated. Since the platform is API based every language can work assuming the correct calls to the API endpoints are implemented.The calls are done json format. Examples of the json calls will be given throughout the demo to better demonstrate its functions. The Sovrin Staging Network ledger can be explored at `https://indyscan.io`.Trinsic allows to use the platform for free for up to 50 credentials exchanged per month. Each time an issuer sends a credential to someone's wall it is considered an exchange. Same goes for when a service verifies a credential from a user's wallet.

4.2 Trinsic demo

The demonstration will be based on the Faber, Acme using Trinsic wallet. The concept of the demo is that Spiros graduated from Faber college, so he needs a certificate to prove that which he will use to apply for a job at a company called Acme.

First create the 2 organizations Faber, the college, and Acme, the company. Then click on Faber icon, Credentials and Create template. For this demo purpose the credential will need 5 attributes. First name, Last name, Degree, Year and GPA. Trinsic will generate the template and provide the schema and credential id. There is also an option to be able to be able to revoke it but there is need in this demo's scope.Some code snippets will be include in each step to demonstrate how each step would look code wise. In this thesis everything is done on the Trinsic website and the certificates are store on a mobile phone. The snippets are included as an example of what it would like if someone wanted to implement it locally.The json call can be found by searching indyscan website for the schema id qKUJQTviaoQ5np3wGvwFT:2:Colledge Transcript v2:1.0. Part of the json call to write on the ledger is available on Figure 11. It shows the DID of the schema issuer, its id, name, time of creation etc.



Credential Template

Colledge Transcript v2

Attributes

First Name

Last Name

Degree

Year

GPA

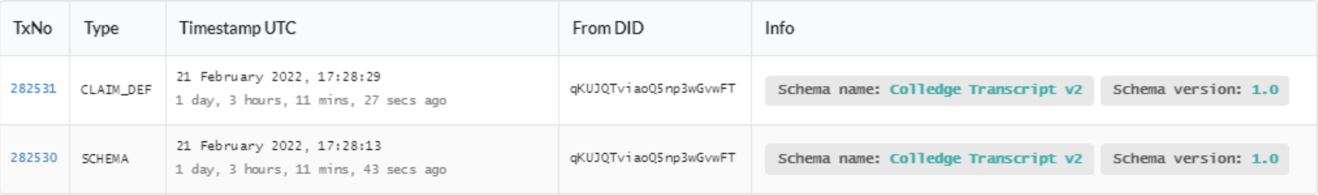
Details

Schema ID	qKUJQTviaoQ5np3wGvwFT:2:Colledge Transcript v2:1.0
Credential ID	qKUJQTviaoQ5np3wGvwFT:3:CL:282530:Default
Revocable	No

Figure 9: Faber credential template

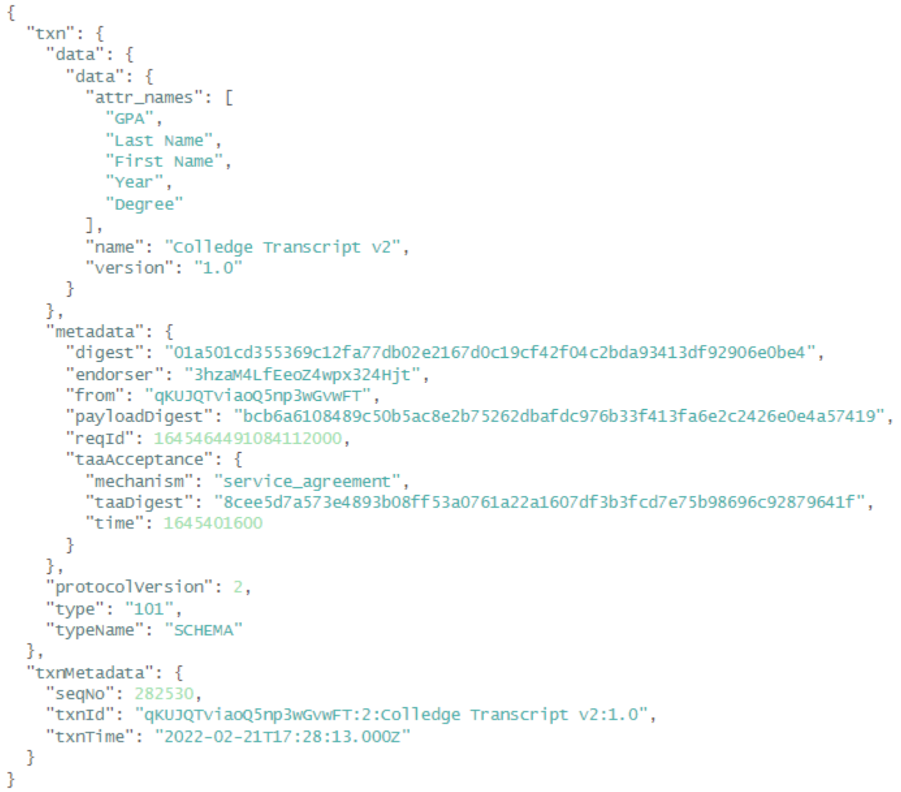
```
let transcriptCredential = await faberClient.createCredentialDefinition({
  name: "College Transcript",
  version: "1.0",
  attributes: ["First Name", "Last Name", "Degree", "GPA", "Year"],
  supportRevocation: false,
  tag: "default"
});
```

Listing 1: Faber credential template in Javascript



TxNo	Type	Timestamp UTC	From DID	Info
282531	CLAIM_DEF	21 February 2022, 17:28:29 1 day, 3 hours, 11 mins, 27 secs ago	qKUJQTviaoQ5np3wGvwFT	Schema name: Colledge Transcript v2 Schema version: 1.0
282530	SCHEMA	21 February 2022, 17:28:13 1 day, 3 hours, 11 mins, 43 secs ago	qKUJQTviaoQ5np3wGvwFT	Schema name: Colledge Transcript v2 Schema version: 1.0

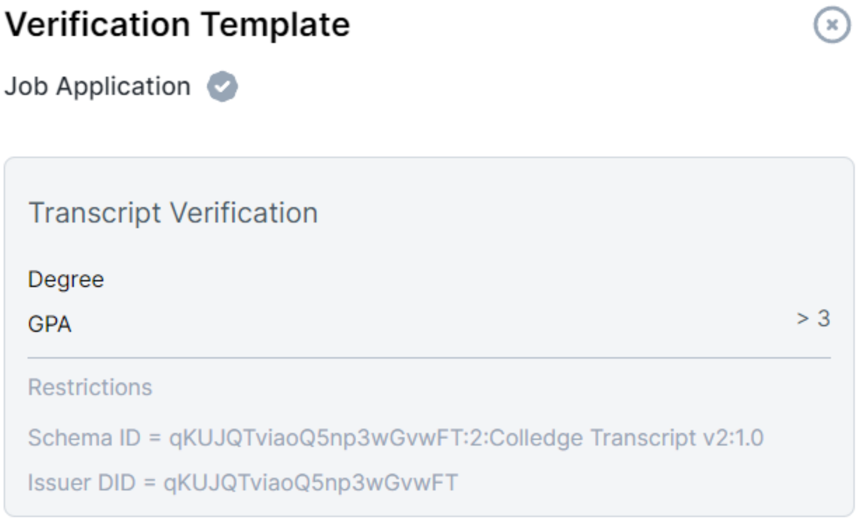
Figure 10: Faber credential schema ledger transactions



```
{
  "txn": {
    "data": {
      "data": {
        "attr_names": [
          "GPA",
          "Last Name",
          "First Name",
          "Year",
          "Degree"
        ],
        "name": "Colledge Transcript v2",
        "version": "1.0"
      }
    },
    "metadata": {
      "digest": "01a501cd355369c12fa77db02e2167d0c19cf42f04c2bda93413df92906e0be4",
      "endorser": "3hzaM4LfEeoZ4wpx324Hjt",
      "from": "qKUJQTviaoQ5np3wGvwFT",
      "payloadDigest": "bcb6a6108489c50b5ac8e2b75262dbafdc976b33f413fa6e2c2426e0e4a57419",
      "reqId": 1645464491084112000,
      "taaAcceptance": {
        "mechanism": "service_agreement",
        "taaDigest": "8cee5d7a573e4893b08ff53a0761a22a1607df3b3fcd7e75b98696c92879641f",
        "time": 1645401600
      }
    },
    "protocolVersion": 2,
    "type": "101",
    "typeName": "SCHEMA"
  },
  "txnMetadata": {
    "seqNo": 282530,
    "txnId": "qKUJQTviaoQ5np3wGvwFT:2:Colledge Transcript v2:1.0",
    "txnTime": "2022-02-21T17:28:13.000Z"
  }
}
```

Figure 11: Faber credential template part of json call

The two transactions show in Figure 10 are the schema declaration at the bottom and the credential offering at the top. These 2 mean that first the Faber college, this is derived from the DID column, created the credential schema and then created a credential with it. Next step is for Spiros to receive his transcript. Faber allows students to get their credentials via QR codes. For this functionality Trinsic wallet will be downloaded on a phone. Select Credentials tabs again, locate the template and click offer. In a real world scenario the info would be populated from a database and each student would get their credentials using their verifiable wallets. In this case define the credentials attribute values to be issued to Spiros. Scan the QR code with a phone to be presented with the credential.



Verification Template

Job Application

Transcript Verification

Degree

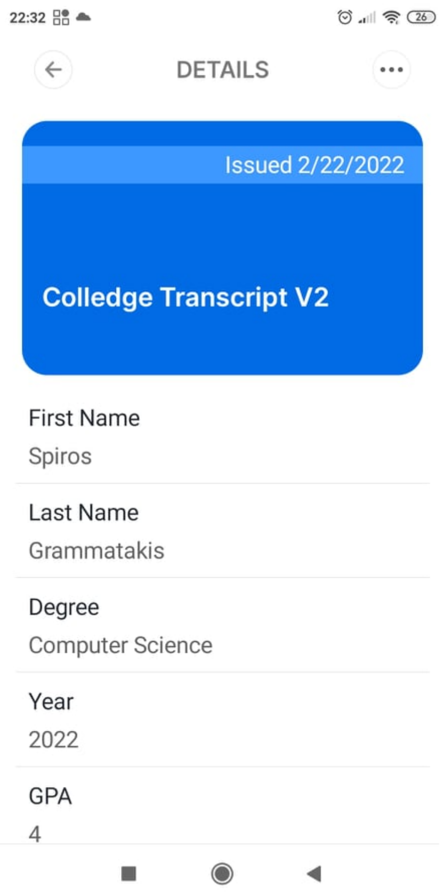
GPA > 3

Restrictions

Schema ID = qKUJQTviaoQ5np3wGvwFT:2:Colledge Transcript v2:1.0

Issuer DID = qKUJQTviaoQ5np3wGvwFT

Figure 12: Verification Template



DETAILS

Issued 2/22/2022

Colledge Transcript V2

First Name
Spiros

Last Name
Grammatakis

Degree
Computer Science

Year
2022

GPA
4

Figure 13: Issued Credential

```javascript
let transcriptCredential = await faberClient.createCredential({
  definitionId: transcriptCredentialId,
  connectionId: faberConnectionId,
  automaticIssuance: true,
  credentialValues: {
    "First Name": "Spiros",
    "Last Name": "Grammatakis",
    "Degree": "Computer Science",
    "GPA": "4.0",
    "Year": "2022"
  }
});
```

Listing 2: Faber credential offer in Javascript

Now it is time for Spiros to connect with ACME and apply for the job. During this part ACME and Spiros will connect with a persistent pairwise key to exchange messages. This is an optional step. Depending on the use case it can be ignore. One use case where one should not create such a connectio with another entity is for example when buying a train ticket. The transaction will only be one time therefore a persistent connection is not required. Having too many persistent connections can make the wallet harder to navigate. Click on ACME and select Invite Connection. The QR code is generated and can be scanned with a phone. Figure 14 shows the available connections.

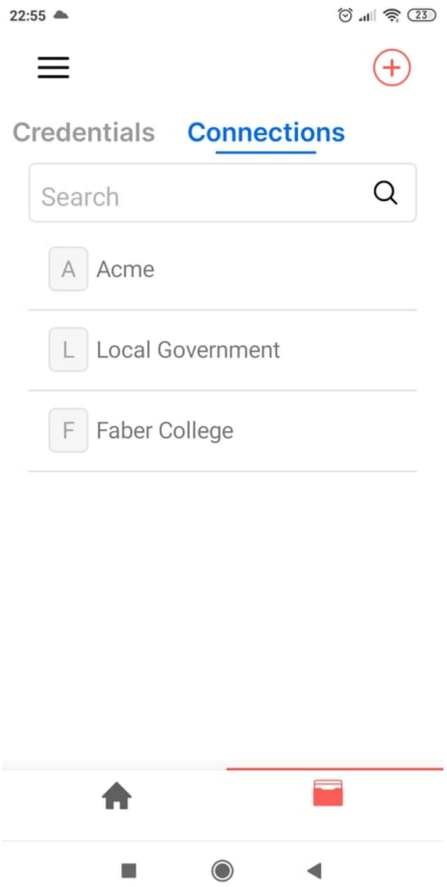
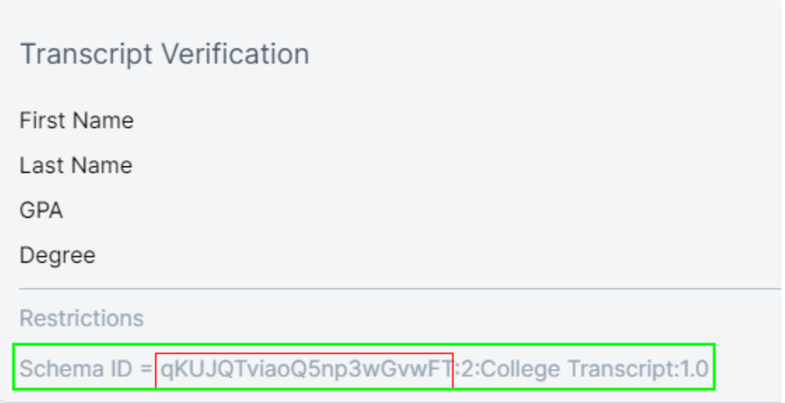


Figure 14: Available connections

ACME must now decide on its requirements for an applicant. Select ACME then Verification Templates. The attributes names have to match exactly the ones from the credential template Faber issued earlier.



Transcript Verification

First Name

Last Name

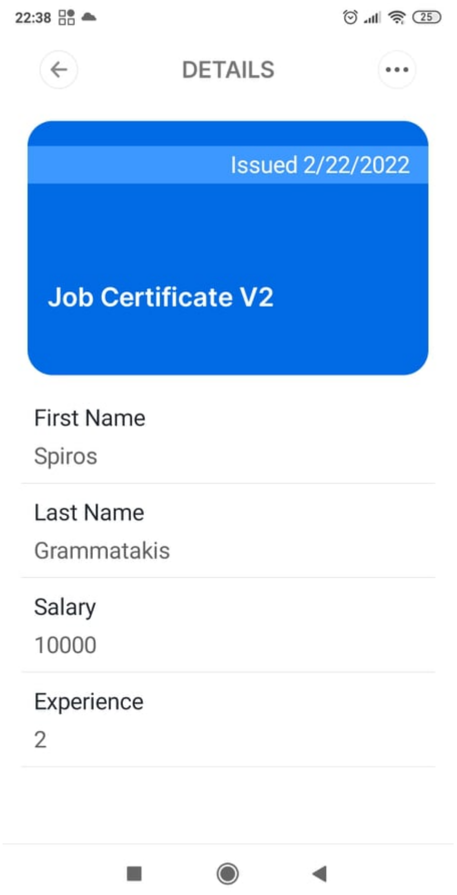
GPA

Degree

Restrictions

Schema ID = qKUJQTviaoQ5np3wGvwFT:2:College Transcript:1.0

Figure 15: Job application template



Issued 2/22/2022

Job Certificate V2

First Name
Spiros

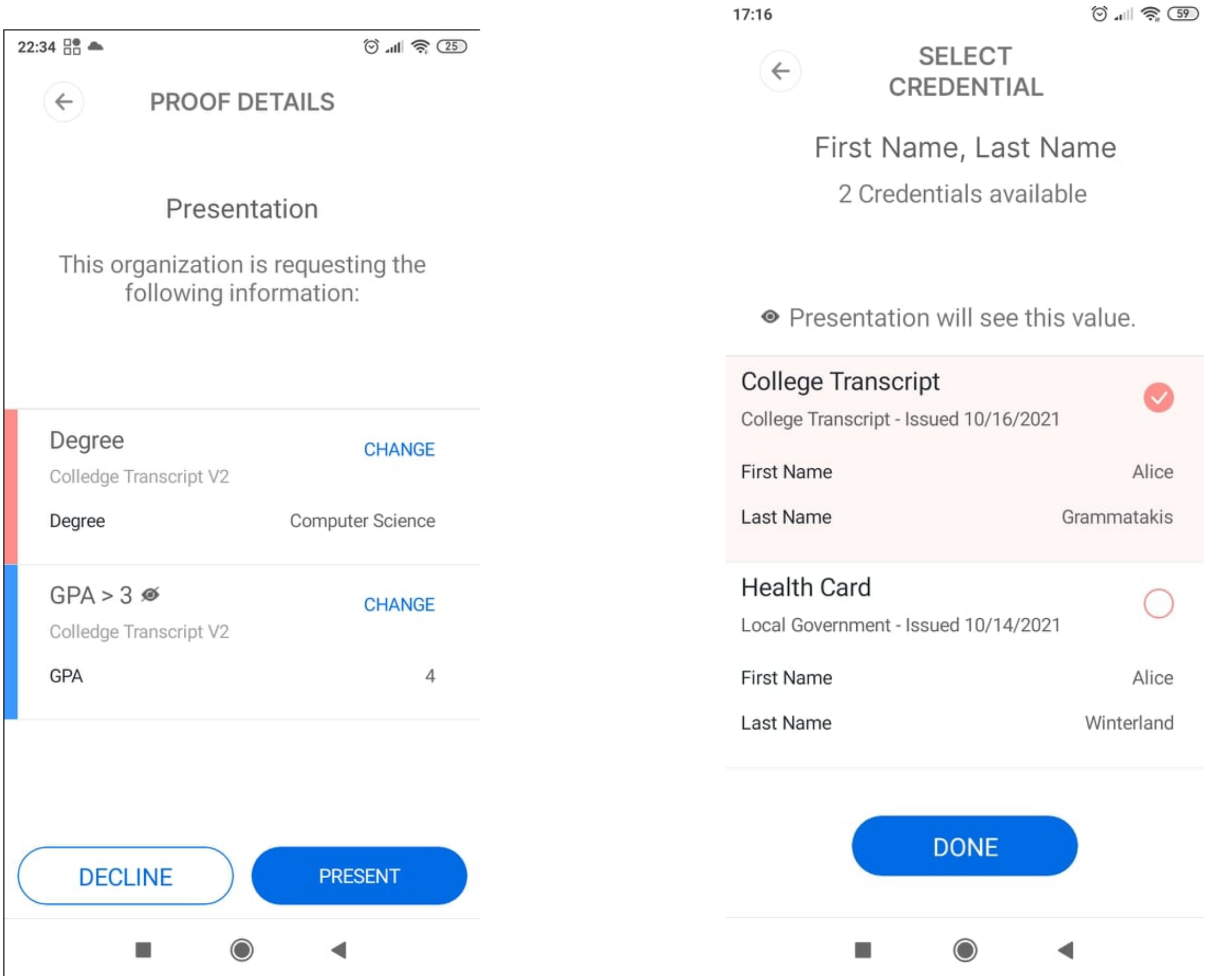
Last Name
Grammatakis

Salary
10000

Experience
2

Figure 16: Job Certificate

ACME now can use the secure connection between it and Spiros to request to prove the mentioned requirements. Acme can specifically define what Schema ID it accepts. Figure 12 shows the Verification template. The text inside green border is the Schema ID issued by Faber college earlier. ACME can also choose to just define the Faber DID, shown in red border in Figure 12, to accept any credential from Faber college containing the required fields. Using the before mentioned Connection tab the ACME can ask Spiros to present proof based on templates. Figure 15.a shows the phone notification for the job application. If ACME does not define a specific Schema ID then Spiros can choose whichever credential satisfies ACME's needs as shown in Figure 17.b.



PROOF DETAILS

Presentation

This organization is requesting the following information:

Degree CHANGE
Colledge Transcript V2

Degree	Computer Science

GPA > 3 CHANGE
Colledge Transcript V2

GPA	4

(a) Defined Schema ID

SELECT CREDENTIAL

First Name, Last Name

2 Credentials available

Presentation will see this value.

College Transcript ✓
College Transcript - Issued 10/16/2021

First Name	Alice
Last Name	Grammatakis

Health Card ○
Local Government - Issued 10/14/2021

First Name	Alice
Last Name	Winterland

(b) Non-Defined Schema ID

Figure 17: Job Application

One problem that could potentially arise if no specific Schema ID or issuer DID is used is that if a forged credential is made, as the one shown in figure 17.b , it is possible to provide false attributes to a Verifier. It is better to specifically define which issuer a company trusts in addition to the Schema ID to be sure the proof is valid. By using specific IDs it is also easier to revoke credential templates if a problem is found and they need to be revoked from the ledger. Lastly ACME can in turn issue a credential to Spiros proving he works for them following the same steps Faber took.

```
let employeeCertificate = await acmeClient.createCredential({
  definitionId: employeeCredentialId,
  connectionId: acmeConnectionId,
  automaticIssuance: true,
  credentialValues: {
    "First Name": "Spiros",
    "Last Name": "Grammatakis",
    "Salary": "1000",
    "Experience": "2 years"
  }
});
```

Listing 3: Employee Certificate in Javascript

4.3 Application Review

In this subsection closing thoughts will be included about the application user experience as well as blockchain in general regarding user experience.

It is important to consider the user's experience when introducing new technology. Familiarity with using mobile applications will certainly help with blockchain mobile wallet applications. Regarding the demo a user would just need to follow on screen prompts. The most important part prior to introducing people to such applications is to educate them on the importance of SSI.Education will help combat the inherent fear as with every new technology that gets introduced.An potential disadvantage is that assuming a user has many concurrent connections, these can come from her/his college, work, telecommunication provider, power provider etc, it can prove to be confusing to manage all of them. To avoid confusion it is vital for a governance framework to be established which will explicitly define which organizations can justify keeping concurrent connections with a user and what information is required. Private and public sector organizations and unions should work together to establish a solid framework.

As a first approach universities degrees could be grouped under 1 schema but that would not be optimal. Let's assume 2 scenarios. Computer science graduates and electrical engineering graduates. If a diploma from a computer science University curriculum certifies a graduate to teach in schools she/he should be able to prove it with just 1 certificate. A simple flag named ”Certified to teach” with true or false as values is enough. A case where special care is required is for example electrical engineers diplomas. In some countries depending on the institution, which comes with different levels of expertise, the ability of a graduate to sign validating a generator's functionality for example depends on the generator's power level. There is no one size fits all approach so it needs extensive work where all parties must participate, government, universities, legal and other interested parties to figure out the details.

5 Conclusion

This research aimed to answer questions regarding blockchain, challenges that can arise with its use and its advantages and disadvantages over current technologies. Blockchain technology is researched in chapter 2, considering both current technology shortcomings and advantages, disadvantages present in blockchain. Security considerations are mentioned in chapter 3. Blockchain is the natural evolution of the internet. The technology companies of the world are all moving towards decentralized infrastructures, cloud computing, to store,process and analyze their data and perform their workflow. The next step in the internet evolution is to allow everyday users enjoy the advantages of decentralized identities. No single point of failure either in trust or computing context means the infrastructure as a whole is more rigid, scalable and friendlier to users. People will have control over their identity without relying on third parties to do so. The immutability of the ledger in addition to the consensus algorithms ensure trust and privacy are preserved to a higher degree than what is offered now. Apart from technology also governance has to catch up and provide the necessary framework for technology companies, both willingly and unwillingly, comply with it and provide quality services to users while earning their trust, respecting their privacy and simultaneously being able to operate on a universally agreed upon framework creating a more equal playing field.

References

[1] Hyperledger. URL: https://www.hyperledger.org/.

[2] Satoshi Nakamoto. *Bitcoin: A Peer-to-Peer Electronic Cash System*. 2008. URL: https://bitcoin.org/bitcoin.pdf.

[3] Bitcoin.org. *Frequently Asked Questions*. URL: https://bitcoin.org/en/faq.

[4] Jameson Lopp. *Bitcoin and the Rise of the Cypherpunks*. 2016. URL: https://www.coindesk.com/markets/2016/04/09/bitcoin-and-the-rise-of-the-cypherpunks/.

[5] Philip Zimmerman. *Creator of PGP*. URL: https://philzimmermann.com/EN/background/index.html.

[6] Bruce Schneier. *Schneier on Security*. URL: https://www.schneier.com/.

[7] Adam Black. *Hashcash*. 1997. URL: http://www.hashcash.org.

[8] Bitcoincharts.com. *Bitcoin Network*. URL: https://bitcoincharts.com/bitcoin/.

[9] Bitcoinwide.com. *Bitcoinwide*. URL: https://bitcoinwide.com/.

[10] University of Nicosia. *UNIC*. URL: https://www.unic.ac.cy/blockchain/.

[11] Nestle. *Nestle expands blockchain to Zoegas coffee brand*. 2020. URL: https://www.nestle.com/media/news/nestle-blockchain-zoegas-coffee-brand.

[12] Hyperledger. *Case Study: How Walmart brought unprecedented transparency to the food supply chain with Hyperledger Fabric*. URL: https://www.hyperledger.org/learn/publications/walmart-case-study.

[13] Fast Identity Online(FIDO) Alliance. *Simpler, Stronger Authentication*. URL: https://fidoalliance.org/.

[14] Sovring foundation. *Sovrin: A Protocol and Token for Self-Sovereign Identity and Decentralized Trust*. 2018. URL: https://sovrin.org/wp-content/uploads/2018/03/Sovrin-Protocol-and-Token-White-Paper.pdf.

[15] D. Cooper et al. *Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List(CRL) Profile*. 2008.

[16] Syh-Yuan Tan, Wei-Chuen Yau, and Boon-Hock Lim. “An implementation of enhanced public key infrastructure”. In: *Multimedia Tools and Applications* 74 (2014), pp. 6481–6495.

[17] Enisa. *Operation Black Tulip: Certificate authorites lose authority*. URL: https://www.enisa.europa.eu/media/news-items/operation-black-tulip/.

[18] Fox-it. *Diginotar Certificate Authority Breach*. September 5 2011. URL: https://web.archive.org/web/20190406122114/https:/www.rijksoverheid.nl/documenten/rapporten/2011/09/05/diginotar-public-report-version-1.

[19] Enisa. *Understanding the Increase in Supply Chain Security Atttacks*. July 29 2021. URL: https://www.enisa.europa.eu/news/enisa-news/understanding-the-increase-in-supply-chain-security-attacks.

[20] Solawinds. *Solarwind Security Advisory*. April 6 2021. URL: https://www.solarwinds.com/sa-overview/securityadvisory.

[21] Cyber Security Infrastructure Security Agency(CISA). *Advanced Persistent Threat Compromise of Government Agencies, Critical Infrastructure, and Private Sector Organizations*. December 17 2020. URL: https://us-cert.cisa.gov/ncas/alerts/aa20-352a.

[22] Liam Tung. *SolarWinds attack is not an outlier, but a moment of reckoning for security industry, says Microsoft exec*. 2021. URL: https://www.zdnet.com/article/solarwinds-attack-is-not-an-outlier-but-a-moment-of-reckoning-for-security-industry-says-microsoft-exec/.

[23] Gloria Omale. *Gartner Predicts for the Future of Privacy 2019*. January 14 2019. URL: https://www.gartner.com/smarterwithgartner/gartner-predicts-2019-for-the-future-of-privacy/.

[24] Jordi Soria-Comas and Josep Domingo-Ferrer. “Big Data Privacy: Challenges to Privacy Principles and Models”. In: (15 September 2015). URL: https://link.springer.com/article/10.1007/s41019-015-0001-x.

[25] Clemens Samarati and Latanya Sweeney. *Protecting privacy when discolsing information: k-anonymity and its enforcment through generalization and suppresion.* August 1998.

[26] A. Machanavajjhala et al. *l-diversity:privacy beyond k-anonymity*. ACM Trans. Knowl. Discov. Data. March 2007. DOI: `10.11-2345/1217299.1217302`.

[27] N.Li, T. Li, and S.Venkatasubramanian. *t-Closeness: Privacy Beyondk-Anonymity and l-Diversity*. 2007. DOI: `10.11-2309/ICDCS.2015.40`.

[28] Andrew C. Yao. “Protocols for secure computations”. In: *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982)*. 1982, pp. 160–164. DOI: `10.1109/SFCS.1982.38`.

[29] Hanrui Zhong et al. *Secure Multi-Party Computation on Blockchain: An Overview*. 26 January 2020. URL: `https://link.springer.com/chapter/10.1007/978-981-15-2767-8_40`.

[30] Guy Zyskind, Oz Nathan, and Alex Pentland. *Enigma: Decentralized Computation Platform with Guaranteed Privacy*. 2015. arXiv: `1506.03471 [cs.CR]`.

[31] Zyskind Guy. *Efficient secure computation enabled by blockchain technology*. 2016. URL: `https://dspace.mit.edu/handle/1721.1/105933`.

[32] Arka Rai Choudhuri et al. “Fairness in an Unfair World: Fair Multiparty Computation from Public Bulletin Boards”. In: *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*. New York, NY, USA: Association for Computing Machinery, 2017, pp. 719–728. ISBN: 9781450349468. URL: `https://doi.org/10.1145/3133956.3134092`.

[33] Di Pierro. *What is blockchain?* 2017.

[34] Jin Wang et al. “An optimized transaction verification method for trustworthy blockchain-enabled IIoT”. In: *Ad Hoc Networks* 119 (2021), p. 102526. ISSN: 1570-8705. DOI: `https://doi.org/10.1016/j.adhoc.2021.102526`. URL: `https://www.sciencedirect.com/science/article/pii/S1570870521000780`.

[35] Wei Yao, Renita Murimi, and Guiling wang. *A Survey on Consortium Blockchain Consensus Mechanisms*. 23 August 2021. URL: https://arxiv.org/pdf/2102.12058.pdf.

[36] L. M. Bach, B. Mihaljevic, and M. Zagar. “Comparative analysis of blockchain consensus algorithms”. In: *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. 2018, pp. 1545–1550. DOI: 10.23919/MIPRO.2018.8400278.

[37] Pierre-Louis Aublin, Sonia Ben Mokhtar, and Vivien Quema. *RBFT: Redundant Byzantine Fault Tolerance*. 2013. URL: http://pakupaku.me/plaublin/rbft/5000a297.pdf.

[38] Jorge Bernal Bernabe et al. “Privacy-Preserving Solutions for Blockchain: Review and Challenges”. In: *IEEE Access* 7 (2019), pp. 164908–164940. DOI: 10.1109/ACCESS.2019.2950872.

[39] Nicholas Confessore. *Cambridge Analytica and Facebook: The Scandal and the Fallout So Far*. April 4 2018. URL: https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html.

[40] Jonathan Stempel and Jim Finkle. *https://www.reuters.com/article/us-yahoo-cyber-idUSKCN1C82O1*. 2013. URL: https://www.reuters.com/article/us-yahoo-cyber-idUSKCN1C82O1.

[41] Upguad. *S3 Security is flawed by design*. 3 April 2019. URL: https://www.upguard.com/breaches/facebook-user-data-leak.

[42] Drummond Reed et al. *Decentralized Identifiers (DIDs) v1.0*. W3C Proposed Reccommendation. W3C, August 2021. URL: https://www.w3.org/TR/2021/PR-did-core-20210803/.

[43] Pavlos Papadopoulos et al. “Privacy and Trust Redefined in Federated Machine Learning”. In: *Machine Learning and Knowledge Extraction* 3.2 (2021), pp. 333–356. ISSN: 2504-4990. DOI: 10.3390/make3020017. URL: https://www.mdpi.com/2504-4990/3/2/17.

[44] Daniel Burnett et al. *Verifiable Credentials Data Model 1.0*. W3C Recommendation. https://www.w3.org/TR/2019/REC-vc-data-model-20191119/. W3C, Nov. 2019.

[45] Daniel Hardman. *DIDComm Messaging*. URL: https://identity.foundation/didcomm-messaging/spec/.

[46] Drummon Reed. *Hyperledger Aries: The Next Major Step Towards Interoperable SSI*. 30 May 2019. URL: https://www.evernym.com/blog/hyperledger-aries/.

[47] Jan Camenisch and Anna Lysyanskaya. “A Signature Scheme with Efficient Protocols”. In: *Security in Communication Networks*. Ed. by Stelvio Cimato, Giuseppe Persiano, and Clemente Galdi. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 268–289. ISBN: 978-3-540-36413-9.

[48] British Columbia. *British Columbia’s Verifiable Organizations*. 2018.

[49] Galia Kondova and Jörn Erbguth. “Self-Sovereign Identity on Public Blockchains and the GDPR”. In: *Proceedings of the 35th Annual ACM Symposium on Applied Computing*. SAC ’20. Brno, Czech Republic: Association for Computing Machinery, 2020, pp. 342–345. ISBN: 9781450368667. DOI: 10.1145/3341105.3374066. URL: https://doi.org/10.1145/3341105.3374066.

[50] Paul Voigt and Axel von dem Bussche. *The EU General Data Protection Regulation (GDPR)*. Springer International Publishing, 2017.

[51] K. Cameron. *The laws of identity*. November 2005. URL: https://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf.

[52] Christopher Allen. *The Path to Sovereign Identity*. April 25 2016. URL: https://github.com/WebOfTrustInfo/self-sovereign-identity/blob/master/ThePathToSelf-SovereignIdentity.md.

[53] Dirk van Bokkem et al. *Self-Sovereign Identity Solutions: The Necessity of Blockchain Technology*. 29 April 2019. URL: https://arxiv.org/abs/1904.12816.

[54] C. Dwork and A.Roth. *The algorithmic foundations of differential privacy*. 2014.

[55] Manas Pratim Bhattacharya, Pavol Zavarsky, and Sergey Butakov. “Enhancing the Security and Privacy of Self-Sovereign Identities on Hyperledger Indy Blockchain”. In: *2020 International Symposium on Networks, Computers and Communications (ISNCC)*. 2020, pp. 1–7. DOI: `10.1109/ISNCC49221.2020.9297357`.

[56] Andreas Grüner et al. “A Quantifiable Trust Model for Blockchain-Based Identity Management”. In: *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*. 2018, pp. 1475–1482. DOI: `10.1109/Cybermatics_2018.2018.00250`.

[57] Sovrin. *Sovring Governance Framework*. 28 Jue 2017. URL: `https://sovrin.org/library/sovrin-governance-framework/`.