Towards Self-Sovereign Identity using Blockchain Technology

Rabobank

UNIVERSITY OF TWENTE.

Author:
Djuri Baars

Supervisors:
Hans Moonen
(University of Twente)
Marten van Sinderen
(University of Twente)
Roel Steenbergen
(Rabobank Nederland)

Executive Summary

With more than three billion internet users, each with multiple digital identities, the management of these identities is very important.

Surveys show that people often use the identity management systems they don't want to use. They don't have full control over their information, have no way to know what is shared with other parties and are dependent on trusted parties when logging in to websites.

Blockchain technology is used as basis for a secure and transparent distributed ledger for the Bitcoin cryptocurrency. Its decentralized, public and immutable properties solve the double spending problem and allow every participant of the network to read the transaction history, help in the validation process and pay and receive Bitcoin.

Cryptographically complex math ensures that everyone can do transactions with everyone without the need for a trusted third party. Next to financial transactions, this also holds for other claims. Entities can put claims on a decentralized ledger by digitally signing it, which allows any other entity to verify that these claims are made by that specific entity.

This allows authorities like governments to make claims about individuals, which can be combined with other claims to create a very strong claim about someone. Because both the claimant and the claimee can be verified, this allows entities like mortgage lenders to outsource their Customer Due Diligence (CDD) processes.

In this research we will explore the possibility of self-sovereign identity, where you are in control of your digital identity.

We started with a desk research on currently available identity management solutions. We concluded that in most systems, the end-user is not able to store their own data. Currently only one decentralized system is available, but has not gained wide adoption yet.

A case study has been performed on a solution which allows the exchange of KYC attributes, resulting from thorough Customer Due Diligence (CDD) as is often performed when opening a bank account. These attributes can be used by other entities, like insurance companies and mortgage lenders to make their on-boarding process easier for customers, since they don't need to supply copies of the same documentation all over again. Also, the companies themselves could outsource their Customer Due Diligence (CDD) this way to lower costs and make fewer errors. Although the idea is very interesting, the studied solution did not meet the expectations. At the time the company behind the solution was very small and the process to improve very complex. The solution was also proprietary, creating dependence on the vendor, which heightens the adoption barrier.

Because of the lessons learned from the case study, the results of the literature research and the desk research, we designed an architecture for a Decentralized Identity Management System (DIMS) using the concept of claim-based identity and blockchain technology.

To lower adoption barriers and create a self-sustaining ecosystem, it will be developed on a public blockchain and source code will be made open-source. The solution will be privacy-friendly by using privacy-enhancing techniques and storing only claims about one's identity. We also provide a solution to allow retrieval of more sensitive data, and made it as modular as possible to make integration within existing IT architecture easier.

The Decentralized Identity Management System (DIMS) can be useful in a wide range of use cases, like proving your age when buying liquor at the supermarket or applying for a health insurance where you get a student discount if you can show your are enrolled at a university.

This shows that our work resulted in a solid foundation for self-sovereign identity using blockchain technology.

List of tables and figures

List of Tables

Table 1 Design-Science Research Guidelines by Hevner [46] 8
Table 2 Consulted experts 11
Table 3 Secure idenTity acrOss boRders linKed (STORK) Quality Authentication Assurance (QAA) levels 16
Table 4 Overview running initiatives 18
Table 5 Top 5 cryptocurrency market capitalization (21-08-2016) . . 27
Table 6 Consensus algorithm comparison 29

List of Figures

Figure 1 Multiple digital identities 1
Figure 2 Sequence Diagram of centralized single-sign-on 2
Figure 3 Centralized vs. decentralized vs. distributed 3
Figure 4 Visualization of a hash function 4
Figure 5 Research Design 9
Figure 6 Visualisation of the funnel method by Hofstee (2006) 12
Figure 7 Identity Management at Facebook 16
Figure 8 Example of establishment during onboarding and re-use of credentials for logging in 17
Figure 9 Attribute Management in BuddyPress 18
Figure 10 Screenshot TrustTester 19
Figure 11 SURFconext screenshot 20
Figure 12 iDIN 20
Figure 13 Screenshots of Jumio Netverify application 22
Figure 14 Screenshot Idensys selection-page 23
Figure 15 Blockchain concepts mindmap 25
Figure 16 Simplified visualization of a blockchain 26
Figure 17 How the Interledger-protocol works 31
Figure 18 Architectural model blockchain 32
Figure 19 High-level architecture of solution 38
Figure 20 Hierarchical deterministic derived keys 46
Figure 21 High-level architecture overview 48
Figure 22 Archimate model issuance of claim 50
Figure 23 Sequence diagram of validation steps 51
Figure 24 Archimate model attribute disclosure 52
Figure 25 Sequence diagram of merchant claim validation 53
Figure 26 Screenshot consumer identity wallet 54
Figure 27 Simplified traditional Application Programming Interfaces (APIs) integration model 57
Figure 28 Simplified decentralized model 58

List of source codes

1 Example integration of GitHub OAuth using passport.js 55
2 Issuance of ”older than 18” claim in smart contract 55
3 Message format of acquirers request 56
4 Trust registry entry with IPNS reference 56
5 JSON access descriptor 57

Glossary

I Reveal My Attributes Project by Radboud University to selectively disclose your attributes, see section 2.3.7

Oracle an information provider or bridge to the blockchain

Taint see section 3.1.9

Wet ter voorkoming van witwassen en financieren van terrorisme Dutch Ani-Money Laundering and Counter-Terrorist Financing Act

Wet toezicht trustkantoren Dutch Act on the Supervision of Trust Offices

Wet op het financieel toezicht Dutch Act on Financial Supervision

Acronyms

ABC Attribute Based Credential

AML Anti Money Laundering

API Application Programming Interface

CDD Customer Due Diligence

CT Confidential Transactions

DIMS Decentralized Identity Management System

DLP Distributed Ledger Platform

DLT Distributed Ledger Technology

DNB De Nederlandsche Bank

FATF Financial Action Task Force

IAF Identity Assurance Framework

IAM Identity & Access Mangement

idP Identity Provider

IGF Identity Governance Framework

IPFS Interplanetary File System

IPNS Interplanetary Naming System

IRMA I Reveal My Attributes

KYC Know Your Customer

NIST National Institute of Standards and Technology

PDS Personal Data Store

PII Personal Identifiable Information

PKI public key infrastructure

PoA Proof of Authority

PoC Proof-of-Concept

PoI Proof of Identity

PoS Proof of Stake

PoW Proof of Work

QAA Quality Authentication Assurance

RBAC role-based access control

SAML Security Assertion Markup Language

SSO Single-Sign-On

STORK Secure idenTity acrOss boRders linKed

tps transactions per second

Wft Wet op het financieel toezicht

Wtt Wet toezicht trustkantoren

Wwft Wet ter voorkoming van witwassen en financieren van terrorisme

ZKP Zero Knowledge Proof

Reading Guide

Both digital identity and blockchain technology are comprehensive concepts. Chapter 1 aims to bring every reader up to speed with both concepts. In section 1.3 we also explain the research itself.

In chapter 2, we will go deeper into digital identity. We start with concepts in section 2.1. These concepts will be used to create a classification in section 2.2. The classification will be used to look at several identity management solutions in section 2.3.

Chapter 3 is about blockchain technology. We will continue explaining about concepts in section 3.1. In section 3.2 we discuss methods to improve privacy and confidentiality.

In chapter 4, we will be using the established conceptual framework to study an existing solution which should allow exchange of Know Your Customer (KYC) attributes. It did not meet all stakeholders expectations, but did provide some insights in how a decentralized architecture for identity management can look like.

In chapter 5, we will design a new solution for self-sovereign identity. It takes the insights from previous chapters to develop modular building blocks to make this possible.

Chapter 6 explains the roles within the Decentralized Identity Management System (DIMS) and shows there are new business opportunities. This should create incentive to participate in the solution and therefore hopefully reach widespread adoption.

In chapter 7, we will discuss the maturity of blockchain technology and the development of similar solutions.

Chapter 8 will conclude with the answers to the research questions and offers starting points for future work.

Disclaimer

Everything written in this research are solely the findings and opinions of the author. It does not represent the public opinion of Rabobank Group or any of the other involved companies nor its employees unless explicitly stated.

Contents

1 Introduction **1**
- 1.1 Digital identity . . . 1
- 1.2 Blockchain technology . . . 4
- 1.3 Research . . . 6

2 Digital Identity Management Systems **12**
- 2.1 Concepts . . . 12
- 2.2 Digital identity system classification . . . 16
- 2.3 Identity management systems . . . 18
- 2.4 Consumer expectations . . . 24

3 Blockchain Technology **25**
- 3.1 Concepts . . . 25
- 3.2 Privacy and confidentiality . . . 33

4 Case study: KYC on Blockchain **36**
- 4.1 Introduction . . . 36
- 4.2 System architecture . . . 37
- 4.3 How the solution works . . . 38
- 4.4 Lessons learned . . . 39
- 4.5 Considerations . . . 40
- 4.6 Validation . . . 42
- 4.7 Discussion . . . 42
- 4.8 Conclusion . . . 42

5 Solution Design **44**
- 5.1 Design motivation . . . 44
- 5.2 Features . . . 45
- 5.3 Benefits . . . 46
- 5.4 Design . . . 47
- 5.5 Result . . . 53
- 5.6 Accessing more sensitive data . . . 56
- 5.7 Comparison existing solutions . . . 57
- 5.8 Known issues . . . 59
- 5.9 Limitations . . . 60
- 5.10 Validation . . . 61

6 Business model **63**
- 6.1 Actors . . . 63
- 6.2 Use cases . . . 64

7 Discussion **67**

8 Conclusion **69**

9 Acknowledgements **73**

10 References **74**

1 Introduction

1.1 Digital identity

Electronic information associated with an individual in a particular identity system is called a digital identity. These identity systems can be used for authentication and authorization [15].

Authentication is the process of verifying a user's identity [44]. There are three methods of authenticating a person:

- something you know (password, pincode)
- something you have (smartcard, hardware token generator)
- something you are (biometric; like fingerprints)

Determining what an entity is allowed to do and enforcing this policy once they are authenticated is called authorization [44][15]

According to Internet World Stats there were more than three billion internet users at the end of 2015 [120]. Next to the digital identity at their internet service provider, they probably have a lot more digital identities e.g. at social networks and their bank as illustrated in figure 1.

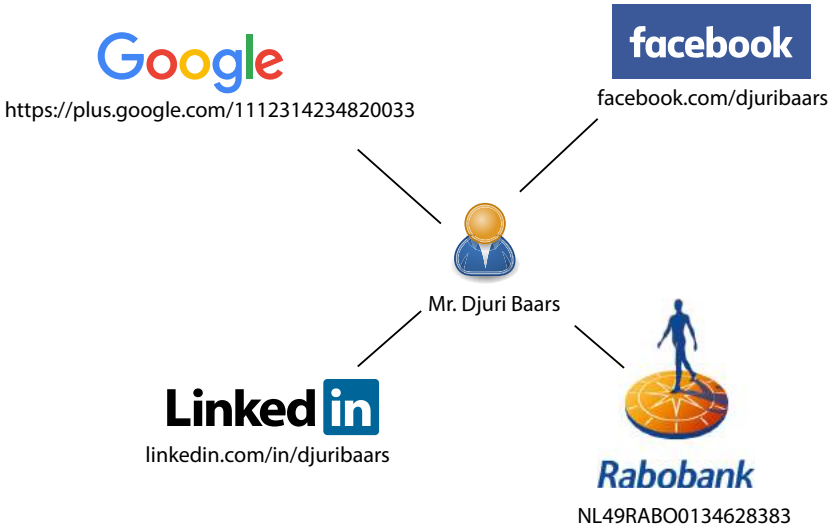


Figure 1: Multiple digital identities

The importance of managing identities has already been noticed because of the ever growing variety of applications and growth of the internet [43]. Although several initiatives like *OpenID connect* provide more convenience for individuals by providing Single-Sign-On (SSO) functionality [74][89], there does not yet exist a solution that allows consumers to manage and store their digital identity completely by themselves.

Organizations offering hosted identity management systems are able to register who does business with whom, which has some serious privacy consequences [5]. In particular, parties offering a wide range of services are able to link these across domains which allows targeted advertising and financial exploitation [73][114].

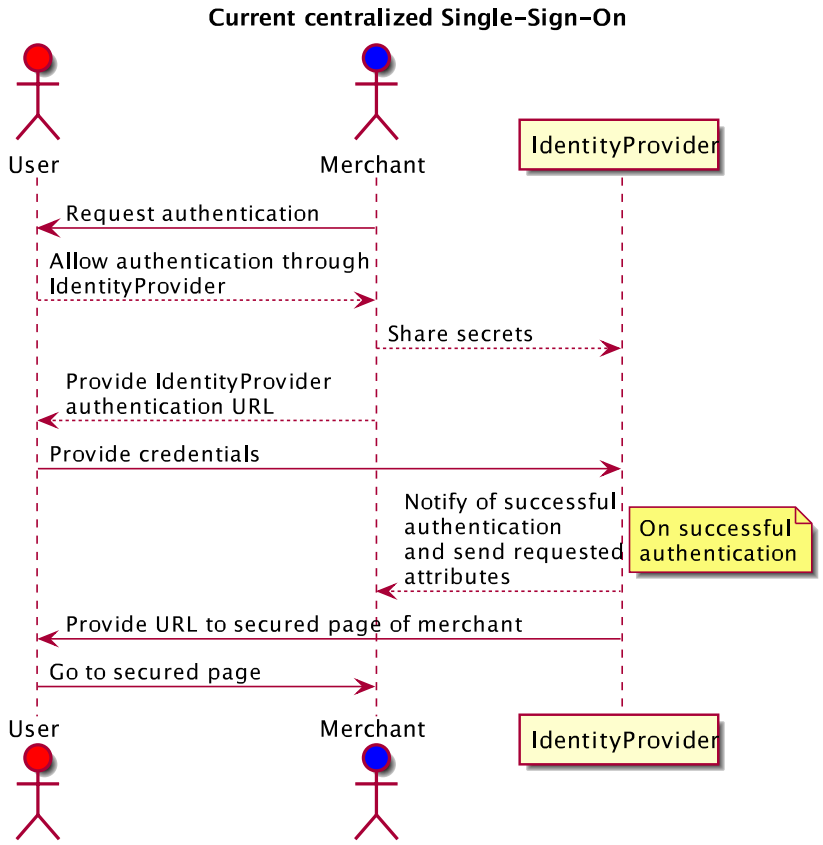


Figure 2: Sequence Diagram of centralized single-sign-on

Microsoft Passport was in 1997 the first initiative which allowed to use the same identity on multiple websites [65][2]. This used a solution that is referred to as *federated identity*.

Next to failing to remember user preferences and a bad user experience [95], it put Microsoft at the center which makes it just as centralized as normal identity systems. This dependency is visualized in the sequence diagram in figure 2 and put in comparison to alternatives in figure 3.

An organization formed in 2001 called *The Liberty Alliance*, established several standards, guidelines and best practices for federated identity as an alternative to initiatives like Microsoft Passport [102]. Their work contributed to the foundation for Security Assertion Markup Language (SAML), an open XML-based data-format for exchanging authentication and authorization between identity providers and service providers [102]. According to [57] SAML 2.0 gained wide acceptance in 2007.

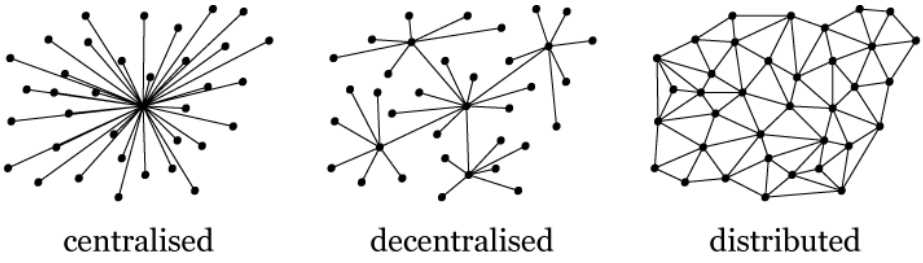


Figure 3: Centralized vs. decentralized vs. distributed

With federation there is a chance that, if the authenticating party is unavailable, the user can not access resources depending on that party.

A considerable amount of literature has been published on Attribute Based Credential (ABC) [56][6][43][4]. ABC is based on the idea of data-minimization and unlinkability of transactions, which makes it very privacy-friendly.

The Dutch IRMA project (short for: I Reveal My Attributes) uses strong cryptography and ABC to create a Decentralized Identity Management System [50]. Attributes like "I'm a student" can be digitally signed by your educational institution and loaded on a smartcard. This could be used to prove your enrollment when a store grants educational discounts on software [33] (see figure 3).

Because of legislation, compliance and accountability it might not be possible for parties like financial institutions to participate in Decentralized Identity Management System (DIMS) where the origin of such claims can not be traced back, since they are required to monitor their data exchanges and be able to validate the origin of claims made [77][48].

A well-designed DIMS is expected to be beneficial for organizations, removing the need to implement one-to-one proprietary integrations (silo's) between back-end systems and reducing dependencies on centralized systems.

Furthermore, it offers organizations with thorough and regulated identity establishment processes ("Know Your Customer") like banks, which are relatively expensive [64]; to create a business case out of sharing their verified attributes with entities like mortgage lenders and insurance companies. Because of the sensitivity of the data, this requires high requirements on privacy and confidentiality.

This could be solved by exchanging claims (answers to questions, like "Are you 18?") instead of sharing the raw data. The claim that you are older than 18 is a lot less sensitive than your birth date. Instead of a fully decentralized architecture this allows for a more distributed landscape. The information required to make a claim remains at the issuer, but the claim itself is available on the distributed ledger of claims (see figure 3).

Blockchain technology could function as the foundation of such system being a network for decentralized trust and exchange. Because everyone can participate as issuer or acquirer (and both), there are low adoption barriers and low costs. This allows new business opportunities for governments, banks and other authorities and more transparency and control for end-users.

1.2 Blockchain technology

Blockchain is best known as the underlying technology of the Bitcoin cryptocurrency [66]. It functions as a Distributed Ledger Platform (DLP) and contains the rules of the platform and the ledger of all transactions since the beginning.

The most characterizing property of blockchain is its immutability. Every block contains a hash of the preceding block. This creates a chain of blocks from the first (genesis)block to the current [22]. This makes it computationally impractical to modify information once it is in the chain because all subsequent blocks should also be regenerated [22], see figure 16.

Hashing is a one-way mathematical operation to compile a stream of data in a summarized form (a fixed-length binary sequence) called a digest. Because of this, there is no easy way to find out what the original stream of data was when you only have the digest. When hashes are smaller than the data, hash collisions can occur which makes it more difficult to find out the original stream of data (see figure 4).

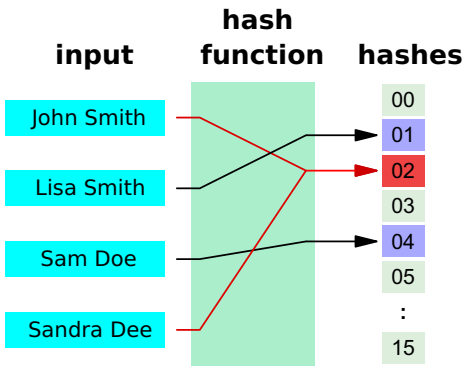


Figure 4: Visualization of a hash function

The Bitcoin blockchain is public and permission-less, transactions details are readable for anyone, anyone can send transactions if they are valid and anyone can participate in so-called mining. Mining is the process for determining validity of transactions and what blocks get added to the chain. To reach this consensus, several methods exists. For Bitcoin this is the ”hashcash” proof of work function [23].

To provide incentive for participation in this block generation process, a block also contains an answer to an extremely difficult mathematical puzzle, where the answer is unique for each block. When solving the block, bitcoins are rewarded to the solver which is also recorded in this block. The processing of transactions of others is also incentivized because of attached transaction fees [21].

Blockchains also exist in more restricted and access-controlled variants, which can be divided in consortium blockchains and fully private blockchains [69].

With consortium blockchains, validity is determined by a predefined set of validators. For example, a consortium existing of fifteen entities require at least ten of the participants to sign a block in order for the block to be valid. Reading the blockchain might still be public or limited to participants.

Fully private blockchains have centralized write permissions, which can be useful for internal auditing within a single organization [69].

Public and permission-less blockchains do not seem to fit privacy-sensitive use cases like managing ones digital identity at first glance. However, work by [100] shows how a metadata-field of the popular Bitcoin blockchain in combination with a commitment scheme can be used for non-financial transactions, like access control or storing consumer consent for sharing data between two trusted parties.

The second most popular cryptocurrency, Ethereum [32] is a platform to build decentralized applications [24]. The possibility of decentralized applications makes it a valuable contribution to research the potential of this technology, and enhance the ability of self-sovereign identity.

1.3 Research

1.3.1 Research Motivation

A survey conducted by Innovalor shows that (Dutch) people have the feeling they don't have any control over their personal data [51]. The value of certain platforms is deemed useful enough to accept the uncertainty about which information is stored about them and who this is shared with. Multiple publications confirm that this desire is held for people worldwide [36][93][96].

This survey also shows that there is a desire to have more control. Next to managing who has access to your personal data, people want more insight in who is using their personal data and modify and delete (parts of) this data.

MIT introduced the concept of Personal Data Stores (PDSs). They describe a personal metadata management framework [63] and developed a prototype called SafeAnswers. They also did a qualitative evaluation of the system, which show 81% of the individuals would use it in their personal life. Although the authors are convinced there is an amazing potential for PDSs, their work faced a number of challenges:

- (Semi)automatic validation
- Privacy preserving techniques
- Development and adaptation of privacy preserving data-mining algorithms
- Better user interfaces which help better understanding the risks and the monitoring and visualization of the large-scale metadata

In this research we will explore how blockchain technology could be used to create such a Personal Data Store and allow self-sovereign digital identity.

1.3.2 Research Question

To explore the potential of creating a self-sovereign identity solution with blockchain technology, the following main research question has been formulated:

> *How to design identity management architecture that is decentralized so that entities can exchange attributes and verify claims without being dependent on a single central authority?*

First we will need to learn about the current state of digital identity management systems and how these are perceived by consumers. Therefore the following sub research questions are formulated:

1. What are the properties of current digital identity management systems?
2. What do consumers expect from identity management systems?

This shows that there is a desire for a DIMS. We continue to design an architecture for this based on blockchain technology:

3. What does characterize blockchain technology?

4. Can blockchain technology be used as infrastructure for identity management?

5. What does an architecture for a DIMS look like?

A DIMS still depends on establishment of identities by entities with proper Customer Due Diligence (CDD) in place, which is very costly. For these entities providing claims resulting from those processes, there should be a business case for participating in a DIMS instead of their own solutions.

6. What is the business model for business participating in a DIMS?

1.3.3 Research Methodology

Because we believe that a successful identity management system can improve effectiveness and efficiency of an organization and the experience of consumers, we based the design of the research on the Information Systems Research Framework by Hevner as shown in table 1. For an activity diagram of this research, see figure 5.

Guideline	**Description**
1: Design as an Artifact	Design-science research must produce a viable artifact in the form of a construct, a model, a method, or an instantiation.
2: Problem Relevance	The objective of design-science research is to develop technology-based solutions to important and relevant business problems.
3: Design Evaluation	The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods.
4: Research Contributions	Effective design-science research must provide clear and verifiable contributions in the areas of the design artifact, design foundations, and/or design methodologies.
5: Research Rigor	Design-science research relies upon the application of rigorous methods in both the construction and evaluation of the design artifact.
6: Design as a Search Process	The search for an effective artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment.
7: Communication of Research	Design-science research must be presented effectively both to technology-oriented as well as management-oriented audiences.

Table 1: Design-Science Research Guidelines by Hevner [46]

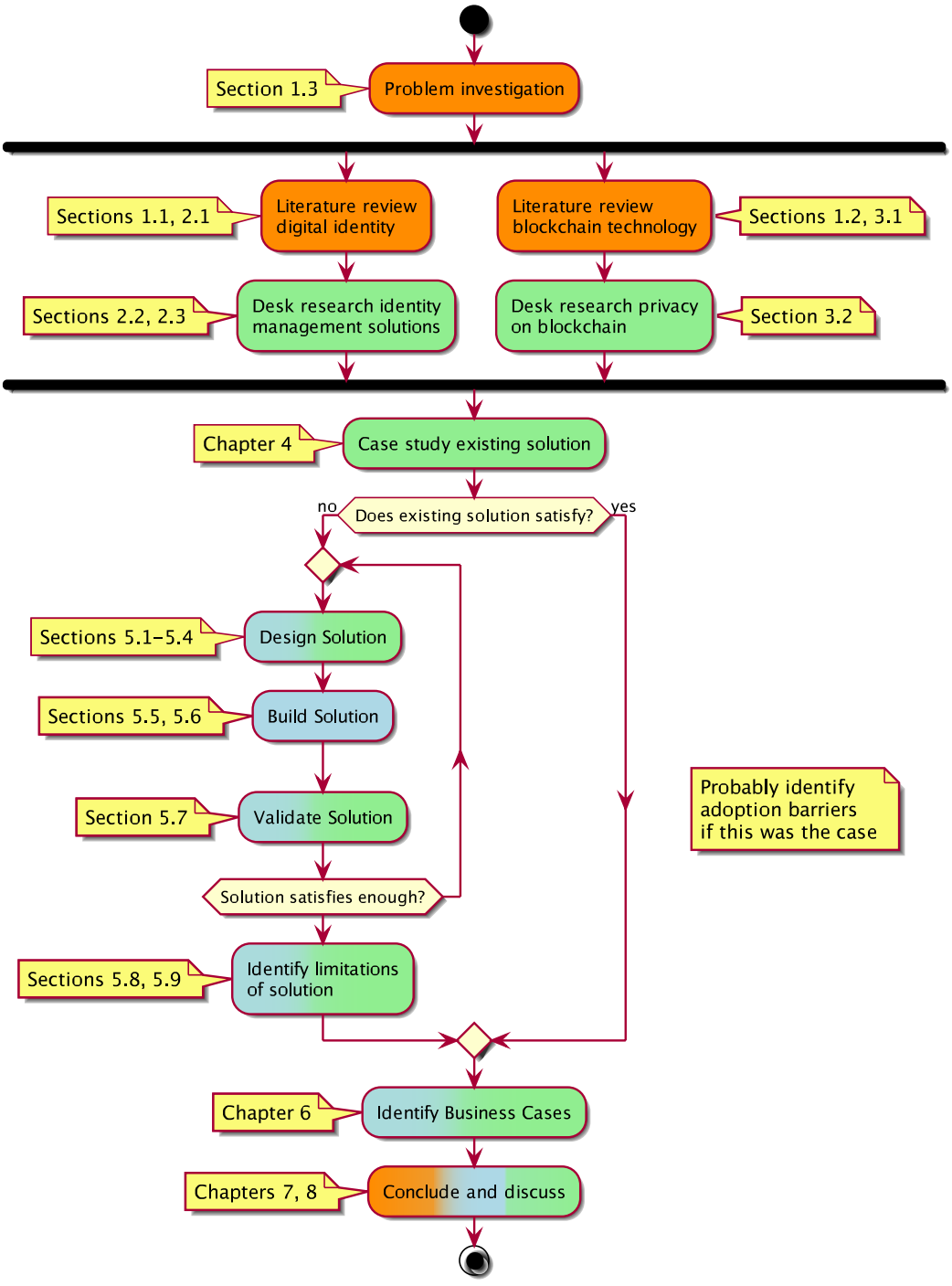


Figure 5: Research Design

Relevance cycle

During the relevance cycle we learn about the current state of technology, related work and learn about general opinion. This will be accomplished by performing literature research.

Tasks related to the relevance cycle are orange in figure 5 and are based on guideline 2 and 4 by [46] as shown in table 1.

Rigor Cycle

During the rigor cycle we develop a knowledge base with properties of identity management solutions and privacy-enhancing techniques for blockchain technology.

A big issue with current popular blockchain implementations is that everything is visible for everyone. Desk research will be performed on how privacy and confidentially can be preserved when using blockchain technology.

We will also look at existing identity management solutions to learn about their properties and to understand why self-sovereign identity has not yet been possible.

A case study of an existing identity management solution based on blockchain technology has been performed. Its properties did not satisfy the solution needs, but did provide usable insights for the design cycle.

Tasks related to the design cycle are green in figure 5 and are based on guideline 5 and 6 by [46] as shown in table 1.

Design Cycle

Based on the results from the relevance and rigor cycle, principles specific to self-sovereign identity are used to extract relevant techniques to build a Decentralized Identity Management System (DIMS).

Because of this we took the lessons learned to design and build a new solution. This solution will be validated by experts. An overview of consulted experts is given in table 2. This provided insight in both the potential and the limitations of the designed solution.

A decentralized solution could potentially render parties redundant, but could also create new business opportunities. We will describe what roles exist within the designed solution and how they could benefit from a DIMS based on blockchain technology.

Tasks related to the design cycle are blue in figure 5 and are based on guideline 1, 3 and 7 by [46] as shown in table 1.

#	**Name**	**Organization**	**Job Title**
1	Rob Guikers	Jibes, Rabobank	Technical Innovation Expert
2	Andrew Mooijman	Uniqom, Rabobank	Project Manager Identity
3	Perry Smit	Chamber of Commerce	Innovator
4	Henk van Cann	Blockchain Workspace	Blockchain & identity expert
5	Marlies Rikken	Innovalor	Advisor

Table 2: Consulted experts

2 Digital Identity Management Systems

To learn about current projects, related concepts and their characteristics, a literature research on concepts and a desk research on trends within digital identity management systems is conducted.

The domain of identity and access management is very comprehensive. We begin with explaining relevant concepts. Using these concepts we create a classification, which is used to look at running projects related to digital identity. This will contribute to the first research question:

> *1. What are the properties of current digital identity management systems?*

The last section will consider consumer expectations of identity management systems to answer the question:

> *2. What do consumers expect from identity management systems?*

Methodology

The method of literature research conducted for this section is based on the post-positivist model.

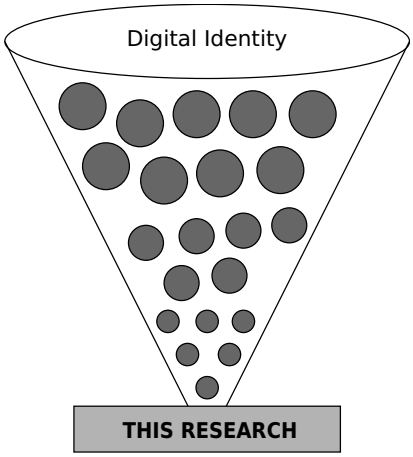


Figure 6: Visualisation of the funnel method by Hofstee (2006)

To investigate the causes and effects, we delved in the concept of digital identity. The funnel method by [47] (see figure 6) has been used to structure the concept in the domain of digital identity relevant for the context of the research question.

2.1 Concepts

We already introduced some concepts in chapter 1. Here we continue by explaining them in more depth.

2.1.1 Federated identity

Federated identity management systems can provide authentication and authorization capabilities across organizational and system boundaries. It requires agreements that an identity at one provider is recognized by other providers and contractual agreements on data ownership [11].

This makes the user and merchant (acquirer) very dependent on the availability of an identity provider (issuer). When the identity provider goes down or discontinues their service and the only offered authentication method is using federation the user can not log in anymore and the merchant might lose many customers.

2.1.2 Self-sovereign identity

Sovereignty is the principle that entities should be able to have control of their own digital identity. Christopher Allen shares a vision about self-sovereign identity and provides ten principles specific to it [2]:

- Existence: Entities must have an independent existence, it can never only exist digitally.
- Control: Entities must be able to control their identities, they should always be able to refer, update or hide it.
- Access: Entities should have direct access to their own identity and all related data. All data must be visible and accessible without gatekeepers.
- Transparency: The system and its logic must be transparent in how they function, how they are managed and how they are kept up to date.
- Persistence: Identities must be long-lived, at least for as long the user desires but it should not contradict a ”right to be forgotten”.
- Portability: All information about identities must be transportable. The identity must not be held by a singular third party.
- Interoperability: Identities should be as widely usable as possible.
- Consent: Entities must agree to the use of their identities and the sharing of all related data.
- Minimization: Disclosure of claims must be minimized.
- Protection: The right of entities must be protected, when there is a conflict between the needs of the network and the right of entities, the priority should be the latter.

Most solutions existing today fall short on access, transparency and portability principles, because they are facilitated by third parties which do not disclose the workings of the system. Because they all strive for the highest adoption themselves or not willing to compromise on security, there are only few solutions portable.

Although it is debatable a fully self-sovereign identity complying to all these principles will ever exist for all identity use cases, they can at least function as ideals to strive for when developing the next solution.

2.1.3 Claim-based identity

Claims are statements which can be made about subjects. They are issued by a provider, which can be the same subject the claim is about. These statements can be made about anything like names, identities or privileges. This way they can provide a powerful abstraction for identity, by decoupling authentication from authorization [10].

Claims can be used to implement role-based access control (RBAC), because they can contain information about role membership. When trusting the issuer, you can choose to receive claims from external providers. This is the case with federated identity [10].

A familiar use of claim-based identity is public key infrastructure, used with SSL digital certificates. Certificate authorities are the issuer of the claim which contain the information to verify authenticity of a domain name [81].

2.1.4 Attribute Based Credentials

An Attribute Based Credential (ABC) is a cryptographic container where attributes like your last name, date of birth or license number, are represented as integers [56][84].

Earlier work already researched the technical possibilities of ABC on smart cards [79]. There are several ideas and concepts presented in this work which will be taken into consideration when designing a new solution. Although technical feasibility is an important factor, the success of a specific implementation still depends on the adoption of both the consumers and suppliers.

Koning et al [56] provide legal and socio-technical exploration of ABC. They mention that users themselves are a serious security and privacy threat, but this is no different from other identity management solutions.

There seems to be a lack of sufficiently appealing business cases for ABC that compete with current data processing practices [56]. We expect to improve this for a DIMS using micro-transactions, which will be described in chapter 6.

In section 2.3.7, we will look at I Reveal My Attributes (IRMA), a system for Attribute Based Credentials (ABCs) by Radboud University in more detail.

2.1.5 Knowing Your Customer

Know Your Customer (KYC) is a regulation governing the activities related to verifying identity of clients of business (Customer Due Diligence). Its objective is to identify, understand and mitigate risks posed by customers, and is part of Anti Money Laundering (AML) initiatives [12].

In the Netherlands the Wet op het financieel toezicht (Wft) and Wet toezicht trustkantoren (Wtt) impose an obligation to operate an adequate Customer

Due Diligence (CDD) system for regulated institutions [34]. The CDD policies should also incorporate the ongoing monitoring of accounts and transactions. The cost of failure to comply can be punitive, as illustrated by PayPal having to pay $ 7,7 million for not having a real-time system to scan and block prohibited payments at that time [38].

It is reported by banks in the USA that the average cost of customer acquisition is $ 1.500 on average [13]. Although organizational changes aimed to increase Anti Money Laundering (AML) compliance efficiency, AML compliance budgets are still increasing [90]. It is assumed by [90] that this is because of fragmented and single use data sources and that a shared services model could lower costs and improve efficiency and responsiveness.

Digital-only banks like the German Fidor Bank and Dutch bunq make use of services like Jumio's Netverify (see section 2.3.9) to eliminate manual document handling and save time and money [54]. This way Fidor Bank manages to keep the total cost of customer on-boarding below €20 [110].

2.1.6 Reference Frameworks

To be able to communicate trustworthiness of authentication mechanisms, reference frameworks which define discrete levels of risk and trustworthiness exist [102].

In Europe the STORK QAA framework is commonly used to unambiguously describe the guarantees which can be expected when using a given authentication method. An overview of the STORK QAA levels is given in table 3. The Dutch eHerkenning, described in section 2.3, offers different levels of assurance which are also based on STORK.

Similar is the Identity Assurance Framework (IAF) by Liberty Alliance, based on guidelines by the American National Institute of Standards and Technology (NIST) [102].

The Liberty Alliance aimed to develop standards for federated identity and web services in relation to Identity & Access Mangement (IAM). Next to IAF they also developed Identity Governance Framework (IGF) which defined how information related to identity is stored and exchanged in a privacy-friendly way. Unfortunately there are no developments related to IGF and there are no known implementations of the framework [102]

Level	Description	Guarantees
1	No or minimal assurance	Minimal or no confidence in asserted identity. Identity credentials are accepted without any verification.
2	Low assurance	Real-world identities must be validated. Authentication should provide enough warranty that the legitimate user uses the identity credentials.
3	Substantial assurance	Registration of identities are processed with methods that unambiguously and with high level certainty identify the claimant. Authentication must be based on at least two factors. The identity provider is supervised or accredited by the government.
4	High assurance	Comparible to level 3. The registration requires at least once either the physical presence of the user or a physical meeting with the user. Furthermore, the identity provider can only use hardware tokens or smartcards which comply to specific requirements.

Table 3: STORK QAA levels

2.2 Digital identity system classification

Because digital identity covers of a broad spectrum of use cases, we will present a classification for digital identity systems. Then we will look at some projects currently in use or in development, relevant to designing a solution for the presented problem in section 1.3.1.

2.2.1 Identity Management

Identity Management will be defined as the process of managing your digital identity. It can be compared to managing who owns a paper copy of your physical identity documents (see figure 7).

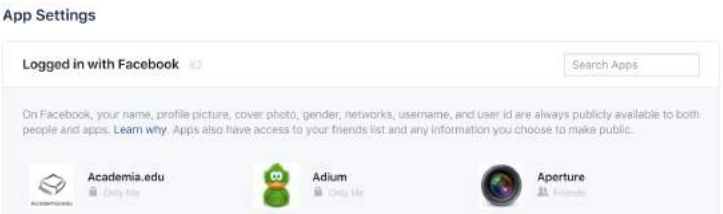


Figure 7: Identity Management at Facebook

2.2.2 Establishment

Establishment covers the tasks and processes related to establishing a digital representation of someone or somethings identity, also known as CDD. At this moment this is often still performed using human validation of similarities between a government issued identity document and physical presence. A visualization where establishment occurs within the process of on-boarding is shown in figure 8.

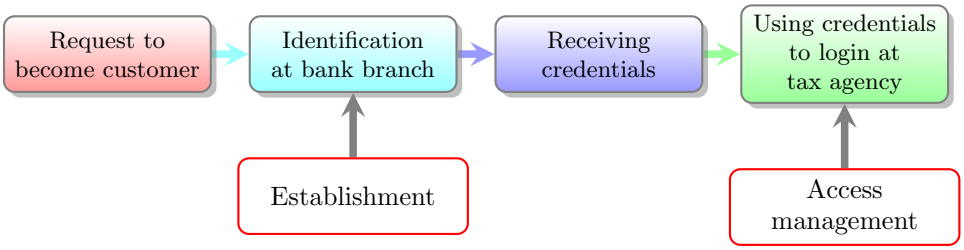


Figure 8: Example of establishment during onboarding and re-use of credentials for logging in

2.2.3 Access Management

Access management concerns authentication and authorization. Authentication is the process of verifying a user's identity [44].

Determining what an entity is allowed to do and enforcing this policy once they are authenticated is the called authorization [44][15].

2.2.4 Attribute Management

Data-minimization and privacy-driven solutions only share strictly required attributes.

Attribute management will be defined as the management of individual attributes, being a subset of an identity belonging to one entity. Examples are sharing only your age and not your date of birth when purchasing alcoholic beverages.

An example of attribute management is shown in figure 9.

Figure 9: Attribute Management in BuddyPress

2.3 Identity management systems

In this section we will look at the properties of several identity management systems, an overview of the systems and how they fit in the created classification is presented in table 4.

Name project	Identity Management	Establishment	Access Management	Attribute management	Use of blockchain	Access	Storage	Technology	Status
Onename.io		X			X	Decentralized	Distributed, Federation		Production
Qiy	X			X		Unknown	Decentralized		In development
iDIN			X			Centralized	Distributed, Federation	SAML	Pilot
eHerkenning			X			Centralized	Distributed, Federation	SAML	Pilot
IRMA				X		Decentralized	Decentralized		PoC successful
PKIoverheid			X			Centralized	Centralized		Production
Jumio		X				Centralized	N/A		Production
Tradle	X				X	Decentralized	Decentralized		Proof-of-Concept
Idensys			X			Centralized	Distributed, Federation		Pilot
uPort	X		X		X	Decentralized	Decentralized		Released September 2016

Table 4: Overview running initiatives

2.3.1 Onename.io

With Onename you can create an Blockchain ID which could function as your digital passport around the web. The verification of your identity is performed using multiple identity providers [71].

Although at the moment these blockchain IDs can only connected to social accounts, in the future they could also be linked to more concrete credentials like social security numbers and insurance information. This area is still far from mature although the first independent verifiers of physical address and phone numbers providing proof on the blockchain already exist [87][86].

2.3.2 Qiy/Digital Me

The Qiy Foundation claims to offer a ”human-centric solution to access, manage and share personal data”. They claim their mission is to ”give people control over

their data and facilitate them to do smart things with it", which are implemented in an open standard. The openness of this standard is questionable though, you are only invited to member events and participate in the Review Board if you pay an annual fee of at least €1.500 (as an individual).

Their scheme consist of rules, regulations and standards for the exchange of personal data. The standard includes considerations about security and privacy considerations which should contain methods to manage these. At the time of writing these were not available for the public [88].

2.3.3 TrustTester

TrustTester allows customers to prove their self disclosed attributes by trusted third parties of the TrustTester platform. After validation, the customer can chose to share the validation result with the merchant. The merchant will only see the attributes are validated by a trusted party but not which one [113].

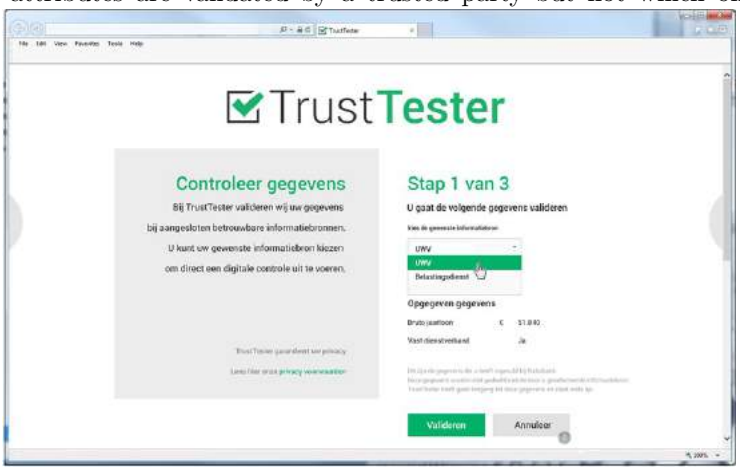


Figure 10: Screenshot TrustTester

2.3.4 SURFconext Federation

SURFconext federation is part of the SURFconext infrastructure. SURFconext offers educational organizations functionality to facilitate inter-organizational collaboration.

They offer a federated identity management service where you can authenticate with your credentials if you are student or employee of one of the almost 120 organizations. According to their website they currently have one million users which generate almost two million logins per month [105].

Among many others, it can be used at SURFspot, a webshop where students and employees of educational organizations get special discounts.

2.3.5 iDIN

iDIN, formerly known as "BankID", will allow customers of Dutch banks to use their trusted bank log-in methods to authenticate themselves. It is development

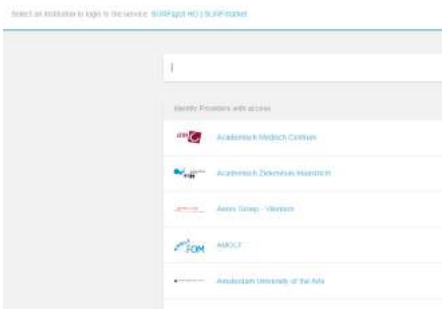
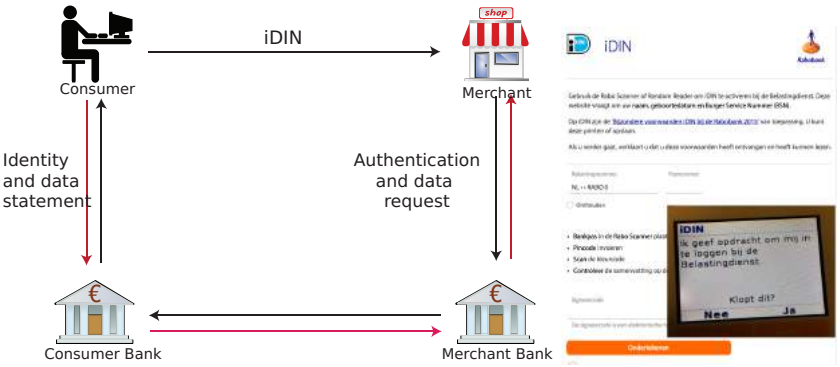


Figure 11: SURFconext screenshot

by the Dutch banks under supervision of the Dutch Payments Association.

It is very similar to the Dutch iDeal system for instant payments, and make use of the same infrastructure. However, during the pilot-phase the payment and authentication functionality are not (yet) combined.

iDIN is based on a bank centric four-corner model, similar to how interbank payments are working, see Fig 12a.



(a) Four corner model (b) authentication screen

Figure 12: iDIN

The four-corner model works as follows:

1. The consumer (debtor) sends an authentication instruction to her own bank
2. The debtors bank verifies the authentication and authorization of the consumer
3. The debtors bank sends the identity to the merchants (creditors) bank
4. The merchant gets informed of a successful authentication by the consumer.

The biggest advantage for consumers is that they don't have to remember a new username/password combination and don't require additional hardware. Also, they don't require registration and validation with another organization.

As of May 2016 they are in a pilot-phase in collaboration with De Belastingdienst. It is expected they go public in the third quarter of 2016.

2.3.6 eHerkenning (eRecognition)

eHerkenning is developed as the successor of the DigiD for organizations. It facilitates authentication and authorization for everyone who wants to use online services. The resources required for authentication differ per provider.

Dependent on the nature of the service, an certain level of assurance is required. eHerkenning supports five assurance levels based on the European STORK framework which allows participants to establish cross-border relations (see section 2.1.6).

The lowest level allows authentication using username and password, the highest level requires authentication using a PKIoverheid certificate (described below). The levels in between require the use of two-factor authentication methods like hardware-generated secure tokens or sms-codes. There are several suppliers where you can purchase the required resources, where higher assurance levels come with more rigorous validation and higher fees.

2.3.7 IRMA

IRMA is an acronym for "I Reveal My Attributes". It is a decentralized Attribute Based Credential (ABC) solution developed by Radboud University. The owner of the attributes is able to share a subset of all attributes, which makes it very privacy-friendly.

Because it is an academic project they do not intend to make a profit. They are however convinced of the desire for such a system and are looking for private parties which are willing to take over the further development of this product.

Their solution is based on the following requirements:

- **Non-transferability**: My younger sister should not get my "over 18" attribute
- **Issuer-unlinkability**: The university should not be able to track where I do my shopping
- **Multi-show unlinkability**: The liquor store should not be able to use my "over 18"-attribute to track my buying behavior
- **Revocation**: Stolen or lost tokens should be blockable

A smartphone or smartcard contains a secret key which is used to make credentials non-transferable. After proving ownership of that secret key, issuers like the government can issue address attributes, which can then be selectively disclosed in transaction.

2.3.8 PKIoverheid

PKIoverheid is the Dutch Public Key Infrastructure of the Dutch Government. On the technical level it is no different from any other PKI-solution. The only difference is that the highest authority is the Dutch Government instead of a private organization.

Although the highest authority is the Dutch government, it can only be purchased at privately held organizations which are under strong supervision of Logius [82].

2.3.9 Jumio

Jumio offers ID scanning and verification solutions for web and mobile. It is able to use webcams and cameras embedded in smartphones for scanning identity documents. This way they are able to help fill in and replace forms required for customer on-boarding for financial institutions or purchasing airplane tickets [54].

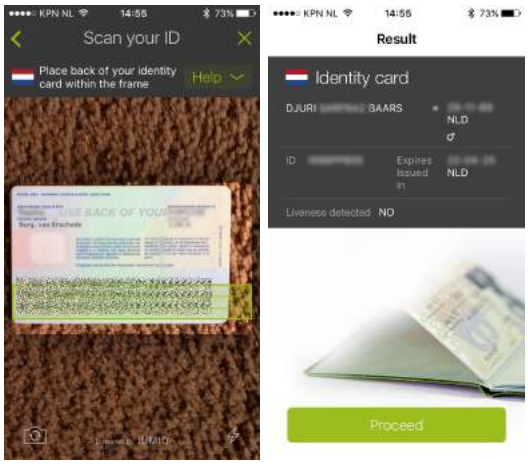


Figure 13: Screenshots of Jumio Netverify application

2.3.10 Tradle

Tradle is a platform for exchanging KYC attributes with the use of blockchain technology. It puts the customer in control of their own data, stored by multiple organizations. After giving explicit consent they are able to share attributes stored by one organization with another, lowering barriers for customer on-boarding and reducing KYC costs for e.g. mortgage lenders and insurance companies.

They offer a server application which is able to exchange KYC attributes which can be mapped to existing data-models. Thereby avoiding the need to replace complex and expensive back-end systems.

2.3.11 Idensys

Idensys, formerly known as "eID-stelsel" aims to be a portal which integrates multiple authentication methods in one portal. Currently the only authentication method supported is eHerkenning and because of that they are often confused.

It is developed by the Dutch government in collaboration with private parties, as part of the Generic Digital Infrastructure (GDI) of the Dutch Government [42]

To maintain interoperability with transboundary eID facilities, it is based on the European enforced requirements and those related to the eIDAS regulations [118].

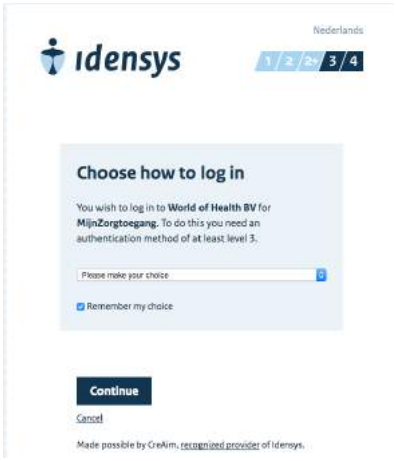


Figure 14: Screenshot Idensys selection-page

2.3.12 uPort (unreleased)

uPort claims to be a web-based wallet and identity system, based on blockchain technology. It has not been released yet, but their platform is said to be open-source. From the available information they seem to focus on personas and identity [116]. Their planning is to release the platform September 2016.

2.4 Consumer expectations

As mentioned in the research motivation (see section 1.3.1), people have the feeling they don't have any control over their personal data and that there is a desire to have more control [51]. People want more insight in who is using their data and modify and delete (parts of) this information. This is confirmed as general opinion by multiple publications [36][93][96].

The work by [63] which introduced the concept of PDSs and describes a personal metadata management framework. The qualitative evaluation of their SafeAnswers system showed that 81% of the individuals would use it in their personal life.

This is further endorsed in [18], where the author advocates a world where endorsements or entitlements can be decoupled from underlying identities to resolve the paradox of more security and privacy. This is consistent with the *Existence*, *Consent* and *Minimization* principles of self-sovereign identity (see section 2.1.2).

Blockchain technology, which is mentioned as potentially suitable platform for bottom-up identity by the same author [17] could contribute to *Access*, *Transparency*, *Consent*, *Portability* and *Interoperability* principles. We will look at blockchain technology in the next chapter.

3 Blockchain Technology

In the first subsection we will go deeper into the concept of blockchain technology and answer the following research question:

> *3. What does characterize blockchain technology?*

The biggest challenges with blockchain technology are privacy and confidentiality of transactions, which are very important when managing ones identity. A desk research has been conducted to learn about methods to improve privacy and confidentiality. This contributes to the research question:

> *4. Can blockchain technology be used as infrastructure for identity management?*

Methodology

Using thematic analysis and selective coding the concept of blockchain technology. For coding the 'open coding' technique by [104] is used. It allows building theory about new phenomena of interest and exploratory build a model to gain understanding of the phenomena. This resulted in concepts and key ideas, visualized in the mindmap shown in figure 15.

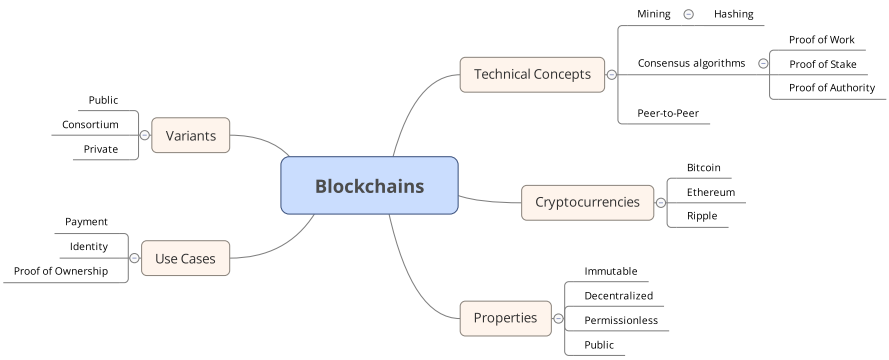


Figure 15: Blockchain concepts mindmap

A wide range of sources is used to collect and structure information about the phenomena. Using thematic analysis by [45] and selective coding by [16], the collected data was systematically and logically related to the concepts and key ideas identified using the initially obtained model.

3.1 Concepts

3.1.1 Relation with Distributed Ledger Technology (DLT)

Although blockchain technology and Distributed Ledger Technology (DLT) are closely related, there is a distinct difference:

- **Distributed ledger**: A ledger maintained by a group of peers, rather than a central agency [78]
- **Blockchain**: A chain of blocks, where each block contains unchangeable records [66]

In combination with consensus mechanisms, blockchain technology can be used as distributed ledger technology for cryptocurrencies and decentralized applications, which will be explained in the next sections.

3.1.2 Cryptocurrencies

A variety of definitions of cryptocurrencies exist, we will use the definition given by [101] defining it as:

> *A cryptocurrency is a digital medium of exchange that relies on a decentralized network, that facilitates a peer-to-peer exchange of transactions secured by public key cryptography.*

To keep track of the legitimate owners of such cryptocurrency, Satoshi Nakamoto presented the concept of time-stamping transactions by hashing them onto a chain of blocks, a blockchain [66]

Each block confirms the integrity of the previous block, making it effectively impossible to overwrite previous records [66]. This makes blockchain an ideal ledger for cryptocurrencies like Bitcoin. A simplified representation is given in figure 16.

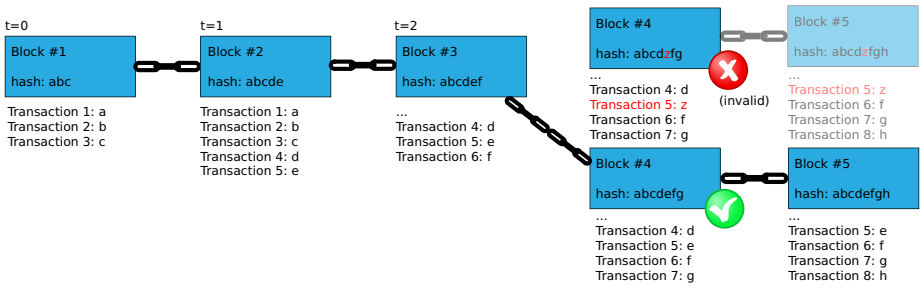


Figure 16: Simplified visualization of a blockchain

At the time of writing, there are 758 cryptocurrencies [32], an overview of the most popular ones are shown in table 5.

Bitcoin

Although many cryptocurrencies exist today, the in 2008 implemented Bitcoin [66] is still by far the most popular [32].

Bitcoin was designed to allow payments to be sent directly to another party without relying on any trusted third party like a bank [66]. Because issuance of currency is part of the ledger it effectively solves the double spending problem.

#	Name	Symbol	Market Capitalization	Price	Supply	Market Share
1	Bitcoin	BTC	$ 9,195,625,072	$ 581.09	15,824,868 BTC	80,13%
2	Ethereum	ETH	$ 924,264,525	$ 11.11	83,203,360 ETH	8,05%
3	Ripple	XRP	$ 214,658,722	$ 0.006036	35,562,073,617 XRP	1,87%
4	Steem	STEEM	$ 173,638,669	$ 1.45	120,147,708 STEEM	1,51%
5	Litecoin	LTC	$ 170,619,517	$ 3.61	47,209,929 LTC	1,49%
	Total (758 cryptocurrencies)		$ 11,476,000,884			

Table 5: Top 5 cryptocurrency market capitalization (21-08-2016)

In 2014 functionality was added to attach a user-defined sequence of up to 40 bytes to each transaction [72]. This allows arbitrary data to be added to the Bitcoin blockchain.

This arbitrary data is already used to store proof-of-ownership of digital art [8] and create two-way links to less-public blockchains [9]. The advantage of these so-called sidechains will be described in more detail in section 3.2.

Despite its popularity currently there are some issues with the protocol [85]. The blockchain gets bigger very fast while block confirmations required to ensure valid payments take longer.

At this moment there are two new versions of the Bitcoin protocol likely to be adopted. Besides choosing one of the two, there is also the possibility both versions will be used with the possible side effect of wallet owners being able to double their Bitcoin [35].

Ethereum

Vitalik Buterin created Ethereum, a next generation blockchain which functions as smart contract and decentralized application platform [24].

Ethereum aims to be an "ultimate abstract foundational layer". Their decentralized ledger technology has a built-in Turing complete programming language, which allows anyone to create programs called "smart contracts" with their own definition of ownership, messaging formats and state transition functions.

These decentralized applications can contain value and perform transactions with that value if certain conditions are met.

Every transaction in Ethereum is a state transition function which can contain data. Although the ledger is still comparable to blockchains like Bitcoin, the contents are optimized for small differences in state in a so-called "patricia tree", which allows shorter block times and thus faster confirmations of transactions [76].

3.1.3 Programmable transactions

The concept of "Smart Contracts" was first described by [107], the Bitcoin protocol does implement a weak version of this concept. It uses a scripting system Script, a simple stack-based language which can be used in transactions [99].

Script can be used for several use cases, like the requirement of two out of three private keys to validate a transaction ("multisig"), or lock funds for a certain amount of time [99].

While the name might suggest otherwise, smart contracts on a blockchain do not have any legal status and are not legally enforceable.

3.1.4 Public vs. private

Public blockchains are accessible for everyone. Participation is unconditional and free. Public blockchains achieve consensus without central authority and thus can be considered fully decentralized [80]. Consensus mechanisms will be addressed in the next paragraph.

When the consensus proces is controlled by a pre-selected set of nodes, the blockchain is only partially decentralized [69]. Reading the ledger can be public or restricted (permissioned). Blockchains can also be hierarchical which allows more complex access control and subcurrencies [59].

When write permissions are kept centralized to one party, you have a fully private centralized blockchain. Reading the ledger can still be public or also permissioned. Practical uses are limited as the only advantage over "normal" distributed database systems is cryptographic authentication [69].

3.1.5 Consensus mechanisms

The Bitcoin cryptocurrency and most other cryptocurrencies currently available make use of Proof of Work (PoW) to reach consensus [58][98].

PoW is very slow and requires an enormous amount of energy. The power consumption of the mining network is estimated to be equal to the power consumption of Ireland [60]. It can be compared to a competition where every participant (miner) tries to solve the same puzzle and validate the same transactions.

The miner who provides a perfect block with the correct solution to the proof of work and complies to other shared rules in the protocol gets rewarded. Subsequently, that block gets connected to the already existing blockchain. All other miners, both cheating and non-cheating waste their energy [7].

Alternative methods for consensus do exist. With Proof of Stake (PoS) participants who own the currency can put this at stake in return for the right to mine. It is assumed that the miner will be honest, because if they eventually prove be dishonest they will be punished by losing their stake [53].

The first cryptocurrency to use PoS is Peercoin [55], but other variants like NXT [67] do exist. There is some criticism on using PoS as single consensus method [83], so some cryptocoins implement a hybrid algoritm [49].

For permissioned blockchains, nodes can be given the right to validate transactions from whitelisted addresses. Because it is assumed only trusted addresses are whitelisted, participants should be able to rely on that fact and so that one confirmation should be enough for finality. This is also known as Proof of Authority (PoA) or Proof of Identity (PoI) (see table 6)

In permissioned blockchains used for cryptocurrency, the currency is issued by a centralized party. There is no need to incentivize mining, often there are no transaction costs. Also, "mining" is computationally cheap because only

validating nodes use energy and there is no need to make the computations more difficult.

An overview of consensus algorithms is given in table 6.

	Proof of Work	**Proof of Stake**	**Proof of Authority**
Speed	Slowest	Average	Fastest
Power consumption	Inefficient	Efficient	Efficient
Permission type	Permissionless	Permissionless	Permissioned
Finality	No finality	Finality (possible)	Finality
Maturity	Tested	Untested	Safe
Costs	Costly	Less costly	Free

Table 6: Consensus algorithm comparison

3.1.6 Finality

Many people claim that public blockchains can't be an acceptable settlement mechanism. Tim Swanson argues that public blockchains can't definitively guarantee settlement finality [106]. However, [70] explains that from a philosophical point there is no system in the world that truly offers 100% settlement finality.

The Proof of Work (PoW) consensus algorithm technically never allows transactions to be truly finalized, because of the probability that someone is always able to create a longer chain that starts one block before and does not include that block [70]. By waiting at least six block confirmations, a transaction is sufficiently close to being final for most entities.

The Proof of Stake (PoS) consensus algorithm offers very strong incentives to never cheat the system. If you cheat and have a block or state that is not present in any other blockchain you will lose your entire deposit required for having the right to validate (stake). Although this does not give the guarantee that transactions will never be reverted, it does give the guarantee that the transaction will never be reverted or a large group of validators will destroy their value at stake. [70]

3.1.7 Privacy considerations

With the Bitcoin blockchain being public, everyone is pseudonymous which poses some privacy issues [62]. Privacy on blockchains will be discussed in section 3.2.

3.1.8 Pseudonymity

Often people mention anonymity is a big advantage of using Bitcoin. There are varying degrees of anonymity and to some extent it is, when you spend bitcoin it is comparable to write under a pseudonym. Everything you spend using a wallet with one or more addresses, it linked to that wallet. If someone knows the addresses linked to your identity, then everything you transacted will also be linked to you.

3.1.9 Taint

Most public blockchains have a cryptocurrency which are mined. When you send Bitcoin to another address, either your own or someone else's, it is will be recorded on the blockchain where other parties are able to analyze how related two addresses are when they both held a particular bitcoin, this is called a taint analysis.

This can for example be used how many steps it takes for bitcoins from an address known for stolen coins, to the current address.[26]

3.1.10 Scalability

Scalability of blockchain-based cryptocurrencies are measured in transaction throughput and storage requirements.

3.1.10.1 Transaction throughput

Transaction throughput is influenced by the technical implementation of the concept of blocks and the consensus algorithm.

As of June 2016, the Bitcoin network has a throughput of 7 transactions per second (tps) because of the block size restriction of 1 megabyte [97].

The current Ethereum blockchain averages on less than 1 transactions per second (tps) [103], the new consensus algorithm CASPER should allow for faster block times and incentivize investment in transaction processing hardware [53].

In comparison, VISA averages around 2.000 tps with a peak capacity of 56.000 tps [119].

Because mining in permissioned blockchain is not required or even absent, they allow for less computationally expensive validation of transactions. Participants are usually known and only allowed if they are trusted. The Proof-of-Authority consensus algorithm and the Tendermint protocol allow for much higher performance. The Tendermint promises transaction speeds of over 10.000 tps [109].

3.1.10.2 Storage requirements

On June 27st, 2016 the complete Bitcoin blockchain is approx. 73 GB large [19].

For most nodes it is not necessary to store the entire chain from the beginning (genesis block) to the most recent one. The Satoshi whitepaper describes the concept of pruning [66], where all information about fully spent transactions is deemed unnecessary.

Although Ethereum is similar to the Bitcoin blockchain, Ethereum blocks only contain the transaction list and the state in that block [24]. When nothing changes, a pointer can be used to store once and use multiple times. This should provide space savings between 5 and 20 times when applying the same to Bitcoin.

3.1.11 Ledger interoperability

Because of the availability of a wide variety of blockchains [32], it would be desirable to have a protocol which allows transfer of data or value between implementations.

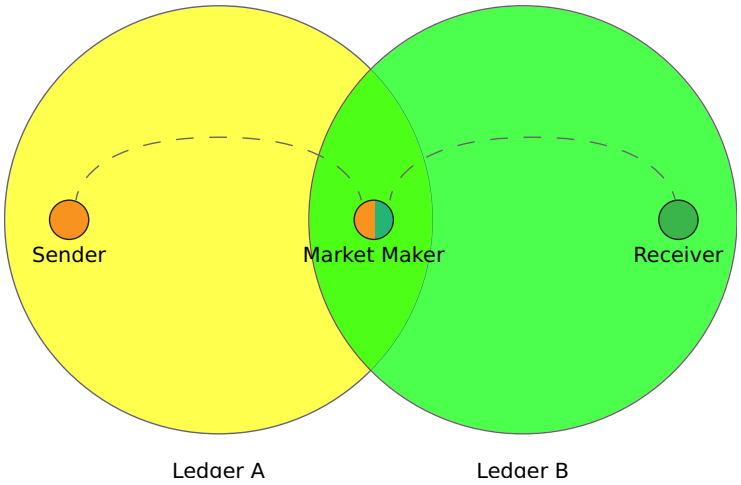


Figure 17: How the Interledger-protocol works

The interledger protocol enables secure transfers between ledgers, which are not limited to decentralized ledger protocols. It provides ledger-provided escrow based on cryptography, which should remove the need to trust a market maker [112], see figure 17.

The protocol as proposed is designed for interledger payments, but its specification does specify a data-field to allow for an arbitrary set of data related like state of decentralized apps [52].

It is similar to sidechains, which will be discussed in section 3.2. The biggest difference is that sidechains allow two-way pegs between blockchains whereas Interledger allows this across payment systems.

3.1.12 Comparison with traditional databases

Although the technical details differ greatly, blockchains can be seen as append-only distributed databases.

Traditional databases like the relational database MySQL do also support a distributed setup but require well-managed access control lists and coordination between nodes for proper replication.

The biggest advantages for blockchain technology at the moment is that there is no need for central coordination and the cryptographically secured immutability. When looking at confidentiality, robustness and scalability, blockchain technology has currently still a way to go compared to traditional databases.

Because each entry is timestamped and its chronological order is cryptographically secured, the use cases for blockchain technology are different and therefore not objectively comparable to traditional databases.

3.1.13 Architectural model of blockchains

We present an architectural model of current blockchain technology using the Archimate modeling language. We modeled one blockchain node, which has a persistent connection with other nodes. There is no dependency on any networks except for the internet, and communication paths are direct. Since all nodes connected to the blockchain are equal in rights (peers), there is no distinction in functions. Every node can mine, read and send transactions.

The only artifact that is produced by the blockchain are validated blocks, which are added after reaching consensus. With Ethereum Virtual Machine based blockchains, the blocks also make up the state of the decentralized applications.

A blockchain client, like a wallet can be used to become a node in a blockchain. The available blocks are synchronized from peers to verify integrity of the chain. It can be used to crytographically sign transactions and relay these into the network, which are validated by other participants using the consensus algorithm and confirmed when included in blocks. The resulting model is shown in figure 18.

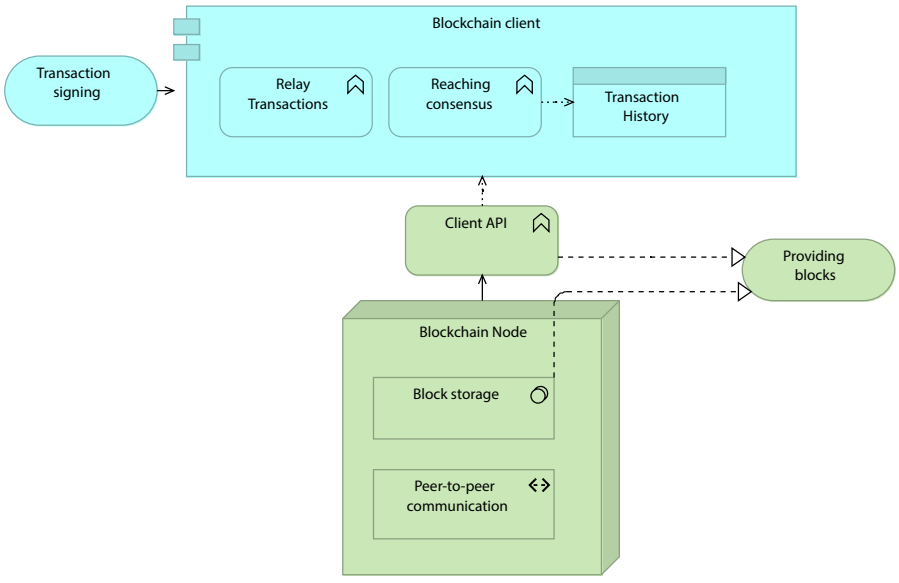


Figure 18: Architectural model blockchain

3.2 Privacy and confidentiality

Transactions on current public blockchain implementations are transparent for every participant of the network (node). This is a desirable property for most people on the Bitcoin blockchain acting for themselves. For financial organizations which are required to adhere to strict regulations and might manage wallets for organizations this is not desirable. In this section we will explore methods and concepts to improve privacy and confidentiality on blockchains.

3.2.1 Address Reuse

To avoid linking transactions to a common owner, it is recommended that users use a new address for each transaction.

When you received cryptocurrency on multiple addresses and then spend them together, the outgoing transaction includes multiple addresses as input, this way multiple addresses could be linked to one individual [1] [91].

Reuse of addresses does not only harm the privacy of an individual but also others. If you are a retailer accepting payments from your customers through a singular address, and put this on your website, it can be strongly linked to your corporate identity. The transaction history which is public on the Bitcoin blockchain can reveal economic activity between customers which harms their privacy [1][117].

3.2.2 Hierarchical deterministic wallets

When combining inputs someone can conclude with high certainty that the owner of multiple source destinations is one individual, which is required if the collection of inputs of one address is not large enough for the desired payment to a single destination address.

By using extended public keys, another party (e.g. consumer, webserver) is able to calculate multiple wallet addresses without revealing private keys. By sending multiple transactions without combined inputs and maybe even some seconds apart from each other allows for a much higher degree of privacy without requiring explicit coordination by the recipient [20].

Because wallets can derived in a deterministic manner, this can be used to detect relations between a master and child key which could function for selective claim disclosure, which will be described in chapter 5.

3.2.3 Coin Mixing

Mixing (also called ”shuffling”) collects the coins from several transactions, shuffles and gives the same amounts back. It is similar to money laundering [94].

Bitcoin (and similar cryptocurrencies) transactions consume one or more inputs and creates one or more outputs. There is no requirement that all inputs need to originate from the same scriptpubkey. It is possible to join transactions and get the effects of mixing without being dependent on a possibly malicious third party. This transaction style is called CoinJoin [28].

3.2.4 Confidential Transactions

Confidential Transactions (CT) is a privacy protocol to hide the value of transactions, it can be combined with other techniques to hide participants of this transaction.

This flexibility allows that only users with strong privacy needs have indeed strong privacy, but this could actually threaten their privacy.

The Confidential Transactions (CT) protocol relies on a variant of Schnorr signatures, which already enjoy widespread deployment as opposed to the Knowledge Of Exponent (KoE) variants which are the basis of zkSNARKs used in ZeroCash/Zcash [14].

The size of a CT transaction is variable, but can be very big compared to other solutions. A draft of a compact variant has been developed, but the cryptosystem for its range-proofs has been broken [31].

Sidechain Elements by Blockstream (alpha) is the best known existing implementation at this time of writing.

3.2.5 Off-chain data

The Bitcoin wiki describes off-chain transactions as the movement of value outside the blockchain [68]. Besides doing transactions off-chain it is also possible to store (meta)data related to the transaction off-chain and exchange this through another communication channel which provides more confidentiality. The owner of the (meta)data can then share this only with intended recipients.

This of course adds the cost of another system to be maintained next to the blockchain. Therefore, one should consider if it is not better to choose a solution which integrates the advantages of using a blockchain and allows to manage visibility of metadata for improved privacy and confidentiality.

3.2.6 Sidechains

A sidechain is a separate blockchain which can be linked with another blockchain using a two-way peg. This enables transfer of assets between both in a cost effective way.

This creates new opportunities like decentralized security, visibility and development of new concepts like smart contracts, digital identity and other research without the need for a new currency [9].

Sidechain Elements, a project by Blockstream (intended for research purposes) are already investigating some concepts described above, including confidential transactions [37].

3.2.7 Cryptographic methods

3.2.7.1 Zero Knowledge Proofs (ZKPs)

Zero Knowledge Proofs (ZKPs) are a concept where one can prove that a statement X is true (prover) to a verifier without revealing anything except the

statement is indeed true. In non-interactive variants there is no interaction necessary between provider and verifier.

zk-SNARK is a non-interactive zero knowledge proof of knowledge, where proofs are short and easy to verify (succinct) [14].

Any verifier can use the verification key to verify a statement X without having to interact with the prover. This is used for constructing a Decentralized Anonymous Payment Schema which hides transaction data. A concrete implementation is used in Zcash [121].

3.2.7.2 Homomorphic Encryption

Homomorphic encryption allows working on encrypted data, without the need to decrypt it first. This allows multiple providers to perform a chain of operations without exposing the data itself [30].

3.2.8 Ring Signatures

A ring signature is a form of digital signature similar to group signatures. Both allow multiple members of a group to sign a message on behalf of the group, without the need to reveal their identity. Ring signatures differ from group signatures that any group of users can be used as a group and there is no tracing authority which provides anonymity of signers unconditionally [61].

3.2.9 Transactional Privacy

The currently unfinished Hyperledger project describes the concept of transactional privacy. In their own implementation it is possible to put chaincodes (decentralized applications) with confidentiality requirements on the blockchain [40].

This can be used for confidentiality against users. Contents are encrypted with secret keys only known to originators, validators and authorized auditors.

4 Case study: KYC on Blockchain

We will apply the conceptual framework to a Proof-of-Concept (PoC) where Know Your Customer (KYC) attributes will be exchanged between two entities after explicit consent of the customer.

Giving and taking away consent is recorded on a blockchain, where the use of a commitment scheme allows the use of public blockchains to function as storage for these actions [41].

Exchange of the KYC attributes from the owner of validated data (issuer) to the acquirer will be off-chain. The exchange action itself and a signature of the exchanged information will be logged on the blockchain as well.

The goal is the to share validated attributes, resulting from CDD processes, between different organizations. This can be useful to on-board a new customer without requiring them to supply paper-based documentation.

This section will contribute to the research question:

5. *What does an architecture for a DIMS look like?*

Methodology

The author of this research observed the project from exploration until completion. To avoid influencing the project, no feedback was given until delivery of the solution.

Using the concepts and ideas from the literature research, the deliverables were critically analyzed. Findings were discussed with all relevant stakeholders and compared with the report of the project manager. The presented findings below are opinions of the author and confirmed with the relevant stakeholders.

Attempt has been made to generalize properties and results of the project and match them with concepts found in literature to avoid bias.

4.1 Introduction

Rabobank decided to do a proof-of-concept of a blockchain-based solution where KYC attributes can be shared with other entities after explicit consent of the customer.

Usually these are conducted within one organization. Since the Rabobank Group consists of organizations in different domains, this allowed for an intra-organizational project with different IT landscapes, which makes this proof-of-concept very interesting.

Its purpose was to gain experience with blockchain technology in general and the suitability of the technology for secure exchange of KYC attributes using this technology.

The use case was to on-board customers for the following products:

- Current account (Bank)

- Mortgage
- Health insurance

Currently each of the participating organizations performs their own identity establishment, which is expensive. As mentioned in section 2.1.5 the cost of customer acquisition is $ 1.500 on average [13].

When validated attributes resulting from CDD processes can be shared, new synergies and business opportunities can arise. Insurance and mortgage companies can depend on the thorough KYC processes from banks, making onboarding easier and faster for their customers while lowering the costs.

4.2 System architecture

The solution consists of the following components:

- **Blockchain**: Used for storage of consent and integrity checksums
- **Server application**: For each issuer and acquirer
- **Customer smartphone application**: For the end-user

4.2.1 Blockchain

The solution is said to be blockchain-agnostic. During the proof-of-concept the Bitcoin testnet was used.

Blockchain technology is used as storage for consent actions and for storing the required information to validate the integrity of received information.

4.2.2 Server application

The server application is required for both issuing and acquiring. It is connected to the back-end systems of the organization to retrieve and store information.

The client application connects to the server of the organization where its attributes are stored. For that reason a webserver is built-in to provide a secure websockets connection.

To check for consent actions given by the consumer and to store data integrity information, there also is a persistent connection to a blockchain. To ensure maximum security, it is recommended to run a full blockchain node in this case.

4.2.3 Customer smartphone application

The customer interface is a smartphone application. At the time of the proof-of-concept only a smartphone application for iOS was available.

The smartphone application is not communicating to the blockchain directly. It uses the server of the organization which contains the attributes to share

as gateway. When attributes are stored at multiple providers, multiple server connections have to be made.

It was not disclosed how a consent action is stored. It is assumed that this is done by sending the consent transaction to the server, which then sends it to the blockchain. This assumption is made because there is no direct connection between the customer application and the blockchain.

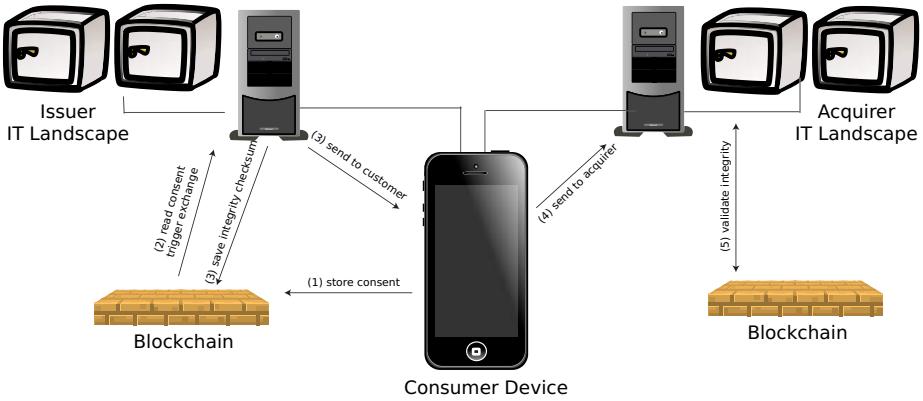


Figure 19: High-level architecture of solution

4.3 How the solution works

At the time of the proof-of-concept there only was a user interface for the customer. Therefore, only a very limited set of functionality could be studied. For confidentiality reasons, only the process is shown. An high-level overview of the architecture and the interactions are shown in figure 19.

4.3.1 First time on-boarding

1. When the customer starts the application, she is required to enter the URL of the server of the organization it wants to on-board.
2. A new identity represented by a key-pair is created on the device of the customer. The identity is sent to the server of the organization.
3. The server seals the customers' identity and sends a product list.
4. The customer chooses a product, choice is sent to server.
5. The server sends the first on-boarding form to the customer.
6. The customer fills in the form, adds attachments and sends it to the server.
7. The server seals the form on the blockchain and sends the next form if required.
8. After each seal a verification is sent by the server to the customer.

9. When the on-boarding is completed including all related processes for CDD, the customer is notified by the server.

In this process there is only communication between the customer application and the server of the organization the customer wants to on-board. Confirmation of the seal could be up to eight hours, because of block confirmation times of the bitcoin network.

4.3.2 Share validated attributes

This process only works when at least one set of attributes is available by another organization than the organization the customer wants to on-board in this case.

1. The customer is required to enter the URL of the server of the new organization it wants to on-board.
2. The existing identity is sent to the server of the organization.
3. The server seals the customers' identity and sends a product list.
4. The customer chooses a product, choice is sent to server
5. The server sends the first on-boarding form to the customer
6. If there are forms which contain the same attributes as another organization, the choice is given to share the already validated form. Otherwise, the customer fills in the form, adds attachments and sends it to the server
7. The server seals the form on the blockchain and sends the next form if required
8. After each seal a verification is sent by the server to the customer
9. When the on-boarding is completed including all related processes for CDD the customer is notified by the server.

In this process there are only communications between the customer application and the server of the organization the customer wants to on-board. The validated attributes need already be present on the customer device.

4.4 Lessons learned

4.4.1 Know the process you are trying to improve

One of the lessons which became clear after some time, is that the processes they tried to improve was very specific to The Netherlands, because of the legislation and regulations around mortgages.

The complexity of the process required simplification of the use cases to be able to validate the mechanism. This also resulted in spending a lot of time on the data-model before a decision was made.

4.4.2 Role of blockchain technology is very small

One of the most important observations of the studied solution, is that blockchain technology is only used for decentralized exchange of commitments. The KYC attributes itself are stored centralized at the organizations.

Consent to share the attributes by the entity (consumer) the attributes are about, is granted and revoked using transactions on the blockchain.

After this consent is given, the data is sent by the data-owning organization to the consumer, and after that from the consumer to the recipient. The action of sending this information is stored on the blockchain, together with cryptographic seals to validate that the data was not modified by the consumer.

4.4.3 Blockchain, a means to an end

The smartphone application showed that the focus was a technical proof-of-concept. It was enough to validate the mechanism, but not very user-friendly. For average customers, the PoC smartphone application was unclear and explanation of what to do was necessary.

4.4.4 Decentralization has big impact on business case

Especially for the vendor spending time and money creating the platform, it will be hard to create a business case. Because the infrastructure (blockchain) is open for everyone and maintains itself, there is no business case for managing this. Anyone could create the same components on the same infrastructure and sell it for a lower price or make it open-source. Making the software proprietary and requiring high license fees heightens adoption barriers. The most realistic business case is in offering consulting services to implement software at organizations and interface it with their existing IT architecture.

4.5 Considerations

4.5.1 Reversibility of cryptographic seals

The used solution is said to be blockchain-agnostic [64]. The blockchain is only used to store cryptographic seals (using hashes or encryption) of identity objects.

It is unclear if the algorithm is reversible, and if it is reversible to sensitive information. This should be taken into consideration when choosing a public blockchain (like Bitcoin) or a private blockchain.

4.5.2 Blockchain interoperability

There are many blockchain implementations available, which differ in functionality and openness. Especially with permissioned blockchains, it is likely that entities who want to exchange attributes are on different blockchains.

Although generic software to connect blockchains exist (like the Interledger protocol, described in section 3.1.11), it is unclear if this is supported with the used solution.

4.5.3 Block confirmation time

The proof-of-concept was used in combination with the Bitcoin testnet blockchain. When using the solution in production, it can be assumed that the Bitcoin production network will be utilized. Currently the block confirmation time of the Bitcoin network is 10-100 minutes. Because transactions should only be considered as valid after a certain number of blocks (exchanges often require 6+ confirmations), it can take up to 10 hours before data exchange is triggered.

4.5.4 Vendor lock-in

Although the solution is blockchain-agnostic, software components are specific to this solution and proprietary. Without the complete solution it is not possible to verify sent data using the solution, which requires other entities to invest in a license. This heightens the adoption barrier and lowers the chance this solution will gain high adoption in general.

4.5.5 Trust

When do other organizations trust CDD performed by other parties than themselves? In The Netherlands, banks are required to adhere to strict regulations, laws and protocols. Is a bank allowed to trust attributes originating from organizations with less strict requirements?

The answer to these questions are not specific to this solution, but need to be answered before a DIMS can be taken into production. A possible solution could be a CDD registry which contains public keys of all regulated banks, maintained by De Nederlandsche Bank (DNB).

4.5.6 Key management

In the current solution, private keys are stored on the customers' mobile phone. This gives the customers full control, but also full responsibility over all actions (transactions) signed with the key. Because of its architecture it is not possible to revoke key-pairs in case of loss or theft, as is possible with normal public key infrastructure (PKI).

4.5.7 Adoption barriers

Because the user interface is very complex and not very user-friendly, the application requires a lot of explanation before it can be used. This can however be improved over time, but will very likely also prevent organizations to use this solution.

Because it took a very long time until a simplified version of the in reality used data-model could be used, it took very long until a minimally working proof-of-concept specific for the participants was developed. The slow development speed suggests that the developing organization is too small and not able to scale up when required.

4.6 Validation

Project manager Andrew Mooijman and technical innovation expert Rob Guikers who were both directly involved with the proof-of-concept at Rabobank, confirmed their experience matches the findings presented in this chapter.

The tested solution is part of learning how KYC cost can be lowered and to provide end-users a privacy-friendly method to share Personal Identifiable Information (PII) between organizations.

Typical KYC processes consist of many parts. The solution itself did not contribute to lowering the costs since it did not have impact on existing KYC and CDD processes for Rabobank.

According to Andrew, Rabobank Group policies, at this moment, do not allow for external validation of Personal Identifiable Information (PII) so that these can be immediately used as part of the Rabobank CDD process. At the same time Rabobank does function as a provider for PII towards other organizations.

He expects that Rabobank Group Policies will be adapted in the future, to allow partial external validation for new customers and new services for existing customers.

Rob noted that it would help if they better understood the complexity of integrating the solution in the existing IT landscape, which they were unable to investigate.

Because there was not enough information about costs and the business model it was not possible to determine if it would be profitable enough to continue using the tested solution. This is also important to determine how likely the solution is to be adopted by other entities.

4.7 Discussion

It should be noted that the organization offering the implementation at that time was still a start-up, consisting of a low number of employees. Because they were unfamiliar with Dutch legislation and did not have much experience with corporate organizations yet, it is very likely that this learning process had a negative influence on the time they could spend on the development of the solution itself. We expect that the involved organization learned from this and is very likely to meet and manage expectations better in successive collaborations.

4.8 Conclusion

Because the used blockchain technology is public and blockchains in general are immutable, it can be understood that sensitive data is not stored on the blockchain itself.

This does raise the question of why separate, new proprietary software should be used together with blockchain technology, instead of modifying existing software to store commitments in the same way.

Because of the high probability of vendor lock-in and slow development it is currently not likely that this implementation will be adopted by large organizations.

We expect that releasing the core product as open-source and consultancy services on the side will help the organization to adapt to different requirements a lot faster than they do at this moment, while still be able to help develop larger organizations to apply it in more complex use cases.

The added value of blockchain technology for this use case did become clear during the case study. It is assumed that offering the same benefits can be achieved with lower dependencies on proprietary components, which will be explored in the next chapter.

5 Solution Design

In this chapter the design of a new solution is explained.

We start with the motivation for designing a new solution. Next we will describe the features which contribute to the goal of accomplishing self-sovereign identity and how it will benefit stakeholders.

Following is the design itself, where we take a top-to-bottom approach. We begin with a high-level overview of the architecture using the ArchiMate Modeling language, where each of the three layers is explained in more detail in subsequent sections.

We then present the result, which consists of four components that can be utilized by each actor to participate in the ecosystem.

In the subsequent section we compare the designed solution to existing solutions and the case study of previous chapter.

We continue discussing the known issues and limitations of the designed solution and conclude this chapter with a validation by experts by explaining the proof-of-concept and a subsequent survey.

This section contributes to the research question:

> *5. What does an architecture for a DIMS look like?*

5.1 Design motivation

The previous sections have shown that:

- People have the feeling they don't have any control over their personal data. They want more insight in who is using their personal data and modify and delete (parts of) this data (Section 1.3.1).
- There are a lot of identity management solutions (IMS) we -the consumer- only use because we have to but none of those put us in control of our information (Chapter 2).
- Blockchain technology has the properties to function as a foundation for Decentralized Identity Management System (DIMS), although the transparency of public blockchains raises some concerns for privacy and confidentiality (Chapter 3).
- A proprietary solution on a public blockchain heightens the adoption barrier, which lowers incentive to participate in such system (Chapter 4).

We will combine our findings, the principles of self-sovereign identity (see section 2.1.2) and the design motivations of the I Reveal My Attributes (IRMA) project (see section 2.3.7) to create design new solution of a DIMS based on blockchain technology.

The lessons learned from the case study in chapter 4 resulted in the following requirements to lower adoption barriers and incentivize the use of the designed solution:

- The system should not be dependent on a trusted third party
- The system should allow acquirers to determine the validity of a claim
- The system should allow issuers to connect existing systems
- The system should be made modular
- The system should be made open-source to create the possibility for each entity to connect themselves

5.2 Features

5.2.1 Decentralized exchange, centralized issuance

Reliability of an identity is only as good as the authority issuing that identity [27]. Although there are a lot of cases where community-based reputation systems can be useful, most business transactions are required to trace back a chain of responsibility in case things go wrong.

Here we will use decentralized exchange of claims using blockchain technology, where claims will be linked to a token by an issuer at the edge of the network.

The platform will be built on an Ethereum Virtual Machine-based blockchain, we leverage the decentralized apps functionality to create trust registries and claim storage. This way we will be independent from the systems of the issuer and allows availability of claims even when the issuer itself stops its services.

5.2.2 Privacy preserving techniques

No storage of sensitive information on blockchain

Because every two-way encryption method can be cracked over time, the aim is to never store any sort of sensitive information.

Although hashes are one-way (see section 1.2), they can be potentially sensitive since attacks have unlimited time to guess (*brute-force*) the input of the digest. For this reason there also will be no hashes stored on-chain.

Instead of storing raw data like your birth date, only answers to questions will be stored within a smart contract on a blockchain. It can be seen as a claim complying to a requirement.

These claims will be binary and generic. When you need to prove you are at least 18 years old, it is sufficient to show the claim that you are by an authorized party (like the government).

The recipient of the information (the acquirer) is able to validate the claim is issued by a trusted party. This way it is not required to share your birth date and make that conclusion themselves.

Proof of the validity of the claim regarding an individual will be stored off-chain, by the issuing party.

No address reuse

Because public keys are pseudonymous, each address should be used one time only (see section 3.2.1). To accomplish this, the proof-of-concept will make use of hierarchical deterministic wallets, where child key-pairs can be derived deterministically from a master key, see figure 20.

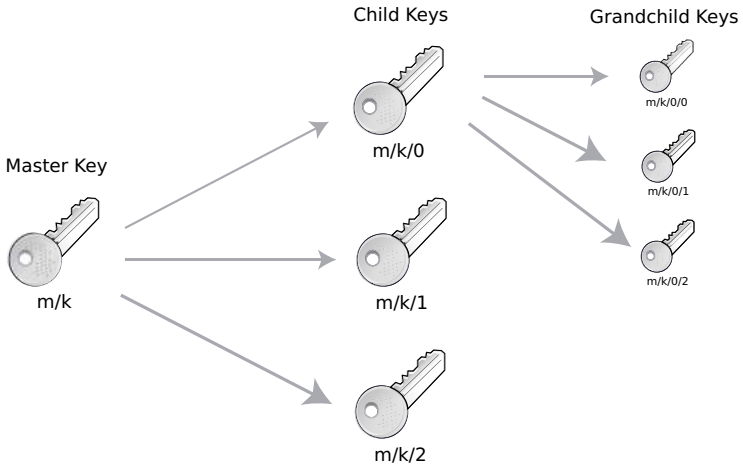


Figure 20: Hierarchical deterministic derived keys

By sharing the derivation path to the identity provider from master to child key off-chain, they can endorse a child key without the need to through the whole authentication and validation process again when the master key is already validated.

Identity verification of acquirers

When an acquirer desires claims, they can ask for it using a digitally signed request. This allows the consumer to verify that the endorsed claims are only shared with the correct entity using public key cryptography. Acquirers should register their public key at trust registries, curated by authorities like the Chamber of Commerce and National Central Banks. This way they can contribute to more privacy and security in both online and offline transactions.

5.3 Benefits

5.3.1 Issuers

Because most issuers already have authentication systems in place, it is desirable to integrate these existing systems in the designed solution. This will be accomplished by offering a modular authentication module, which is able to integrate the most used strategies for authenticating entities, like SAML. This lowers the cost and the adoption barriers to implement the solution.

5.3.2 Acquirers

Acquirers will be able to request virtually any claim they require using the same infrastructure. Claims and the information required to verify the validity of those claims can be encrypted to their public key, using the same infrastructure they are able to verify the claim is issued by a certain authority. It depends on the availability of issuers before this will become practical, however when an acquirer decides to integrate the solution; it will instantly give the possibility to acquire claims about consumers by all connected issuers.

5.3.3 Consumers

For the consumer a big benefit will be to own and manage claims themselves. The process to do this should be as friction-less as possible. The aim is to create a straightforward and easy to understand user experience.

Being able to be sure that only selectively disclosed claims will be shared with a verified acquirer, will put the consumer in full control of their own information and thereby creating self-sovereign identity.

5.4 Design

The solution makes use of existing web technologies to authenticate consumers at Identity Providers (idPs). After successful authentication at the idP and proving ownership of a certain private key by the consumer, the idP stores the validity of the claim in a smart contract on the blockchain.

To show the relation between business (yellow), application (blue) and infrastructure (green) the ArchiMate modeling language will be used. Interactions between the actors are visualized using sequence diagrams.

5.4.1 High-level architecture

A high-level overview of the architecture is shown in figure 21. The architecture will be described from bottom to top.

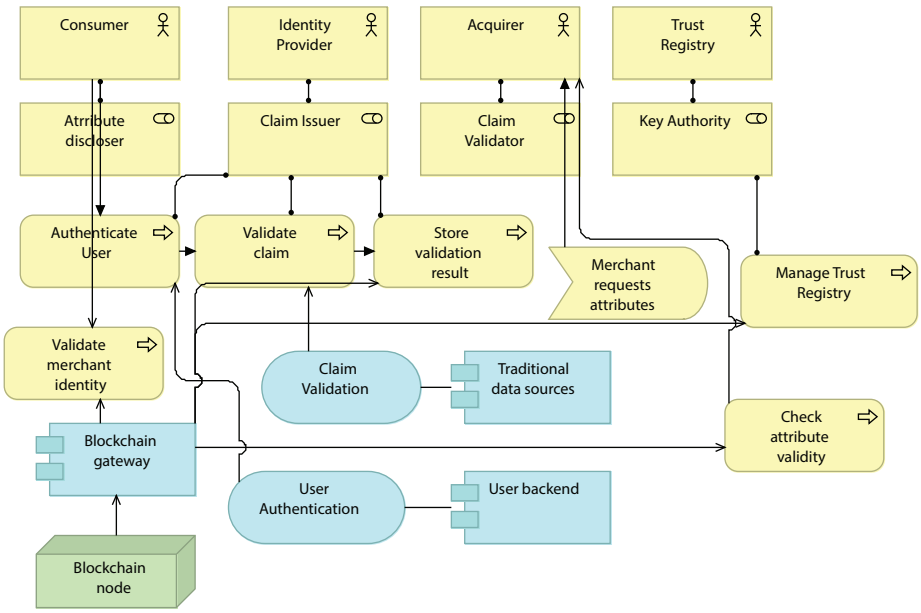


Figure 21: High-level architecture overview

Infrastructure layer

The Ethereum blockchain will function as decentralized storage platform of the claims. Issuance and revocation of claims will be stored in smart contracts which are also part of the infrastructure.

Although the Ethereum blockchain consists of many nodes, they function as one single shared ledger for all claims (transactions). The same holds for the Ethereum Virtual Machine which is responsible for executing the smart contract code; every node executes this the same way.

The infrastructure part is simplified in this model, a detailed model has been described in section 3.1.13.

Because the infrastructure supporting the traditional data sources and user back-end will differ per identity provider, they are left out of the model for readability.

Application layer

For each identity provider there is a smart contract on the blockchain. The application services can be generalized to user authentication and claim validation. They are free to implement these services as they desire. When there is sufficient information retrieved to be able to make the claim, the result will be written by the idP to a smart contract on the blockchain.

Business layer

On the highest level, the following business processes can be distinguished:

- Claim validation by consumer at identity provider
- Attribute disclosure by consumer to merchant
- Managing the trust registry by key authorities

In the next section we will focus on the business layer in more detail.

5.4.2 Business actors

- **Consumer**: The entity who wants to share claims about her identity
- **Issuer (Identity Provider)**: An entity who is able to verify claims about an identity
- **Trust Registry**: An entity which curates a registry where it is the authority on
- **Acquirer (Merchant)**: The entity that desires specific claims to be either true or false

5.4.3 Business processes

Claim validation by consumer at issuer

The consumer is able to generate its own key-pair on the Ethereum blockchain. This key can be used to ”collect” claims about herself using multiple identity providers.

First, an entity needs to authenticate at the identity provider (e.g. using username, password and SMS-code). Next, the identity provider can verify ownership of a certain private key by sharing a message the consumer needs to sign. After ownership of the key is proved, the consumer can chose to let the identity provider store claims about her identity, when the idP has sufficient proof that they can do so.

When desired, the consumer is able to derive keys from the master key in a hierarchical deterministic way. When providing the path from master to child key to an idP, it can relate master-child and endorse the same claim without requiring the consumer to go to the authentication process again.

The business interactions are shown in figure 22, the technical mechanisms are shown in 23.

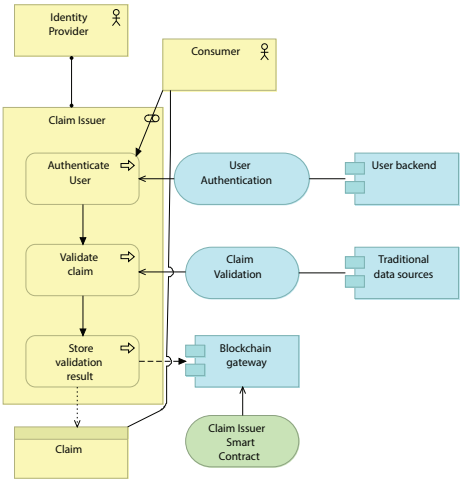


Figure 22: Archimate model issuance of claim

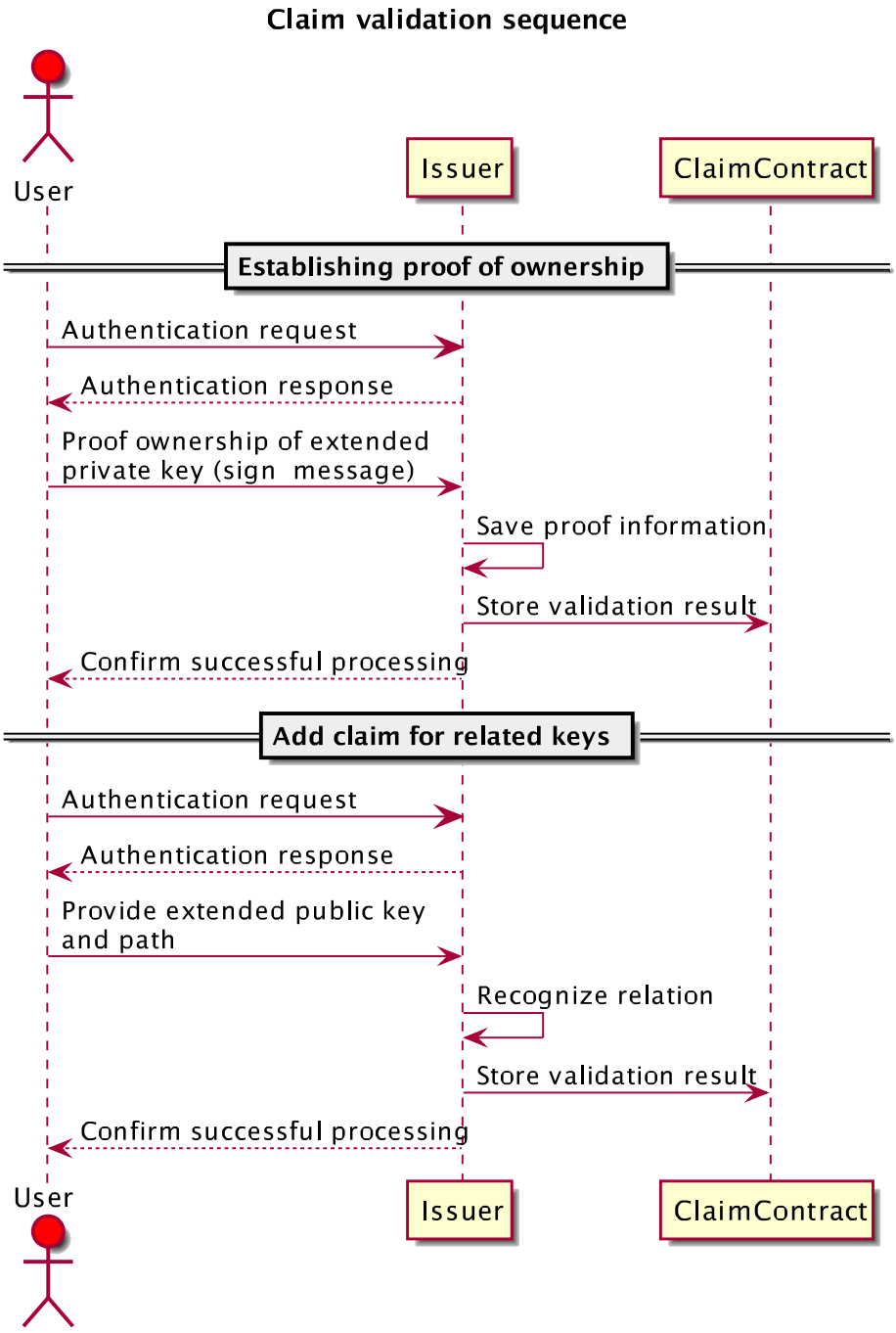


Figure 23: Sequence diagram of validation steps

Selective attribute disclosure by consumer to acquirer

In business transactions, a merchant (the acquirer) could require verification of certain attributes (e.g. a discount for students only). During this business process, the consumer can share a public key loaded with only those claims relevant to the transaction.

Before doing the transaction, the consumer can verify the identity of the merchant by looking up the public key in the trust registry. By encrypting information to this public key, the consumer can make sure it only shares the claims with the intended acquirer (recipient) only.

After verifying the merchants' identity, a derived key will be made using the hierarchical algorithm. The path from master key to derived key can be shared with each identity provider, which allows each identity provider to assign the same truth value to the derived key as was done with the master key.

Finally, after making claims about the derived key, this key and location of the smart contract with the claim value, can be shared with the merchant. The acquirer is able to verify that the claims are indeed made by trusted parties. The simplified business interactions are shown in figure 24 and the technical mechanism is shown in figure 25.

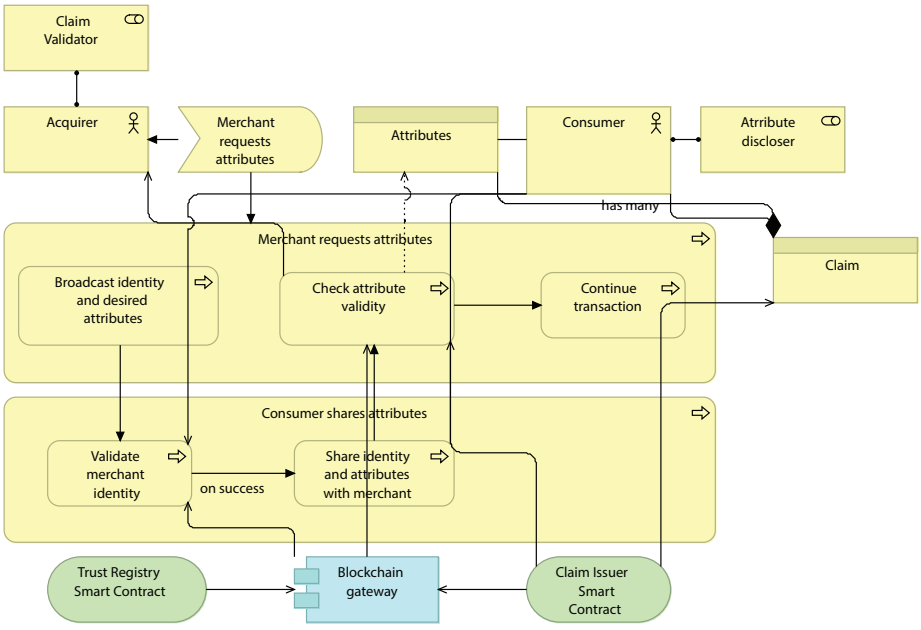


Figure 24: Archimate model attribute disclosure

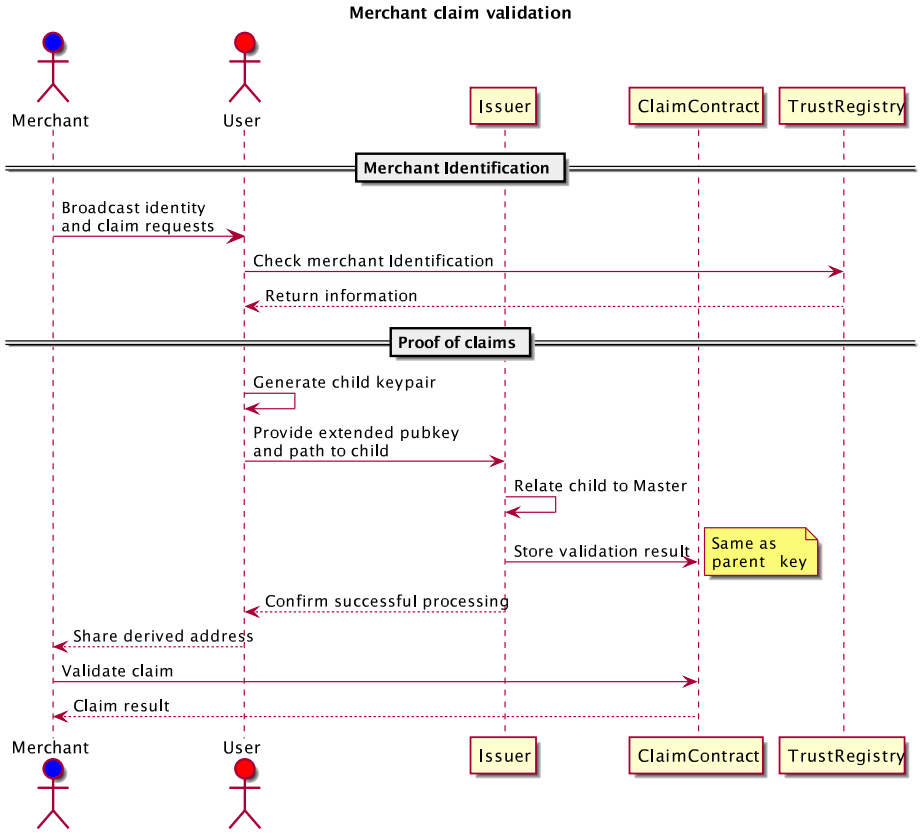


Figure 25: Sequence diagram of merchant claim validation

Managing the trust registry by key authorities

Managing the trust registry is similar to claim validation of consumers by identity providers. Here entities belonging to a specific type of organization (e.g. local banks) can proof to authorities (central banks) that they have ownership of a private key. The authority can then add this information to the trust registry it curates, which is also a smart contract.

5.5 Result

For each of the business actors and its roles a technical proof-of-concept has been developed. For convenience and demo purposes an user interface has also been created.

5.5.1 Consumer identity wallet

The identity wallet is a smartphone application where the consumer can manage the digital identity, represented by a key-pair. Identity providers make claims

about someone by making statements about one's public key, in a smart contract on the blockchain.

For optimal convenience, most steps described above should be automated. Using Bluetooth beacons for example, a merchant can broadcast their public key and request for claims, which can be shown in the app. A screenshot of the app is shown in figure 26.

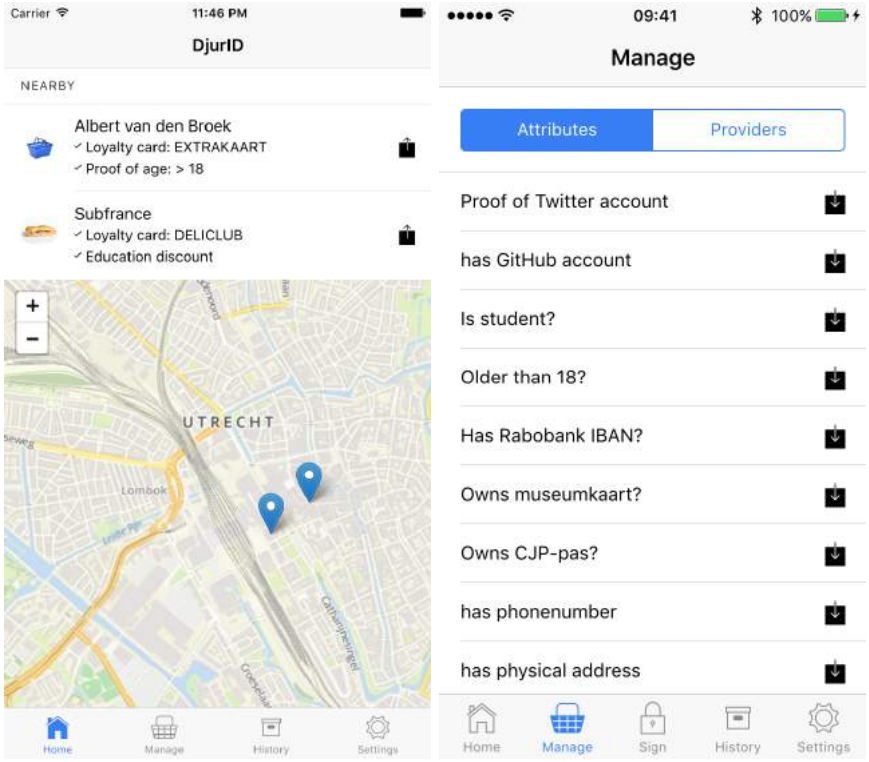


Figure 26: Screenshot consumer identity wallet

Using the broadcast information, the application can check the trust registry. If the public key indeed belongs to the merchant which the consumer wants to transact with, it can show the merchants' information.

If everything is in order, consent can be given to exchange the relevant claims. The application automatically generates a derived key and requests the relevant endorsements of the claims at the identity providers. When this process completes, the derived key can be shared using Bluetooth or using a QR-code.

5.5.2 Claim issuance by issuers

To enable Identity Providers to add claims to the blockchain, a modular oracle has been developed. An oracle is an information provider or bridge to the blockchain. It allows almost every authentication method to be plugged in using passport.js strategies, which supports more than 300 strategies like OAuth and SAML [75]. Integration can be as easy as adding a few lines of code, as shown in listing 1.

```
var GITHUB_CLIENT_ID = config.github_client_id;
var GITHUB_CLIENT_SECRET = config.github_client_secret;
passport.use(new GitHubStrategy({
    clientID: GITHUB_CLIENT_ID,
    clientSecret: GITHUB_CLIENT_SECRET,
    callbackURL: "http://127.0.0.1:3000/login/github/return"
  },
  function(accessToken, refreshToken, profile, done) {
    process.nextTick(function () {
      return done(null, profile);
    });
  }
));
```

Listing 1: Example integration of GitHub OAuth using passport.js

After successful authentication and receipt of proof of ownership of a private key, claims can be saved to the idP's smart contract. An example of the "older than 18" claim is shown in listing 2.

```
{
  "pubKey": "0x6395F09b3ED5E1FD1E482773a6784bC0a79529ed",
  "validatedOn": "2016-06-23T18:25:43.511Z",
  "pastAge": "18"
}
```

Listing 2: Issuance of "older than 18" claim in smart contract

In the case of claims which might differ in the future (like subscriptions), a `validUntil` property can be added.

To be known as an issuer for specific claims, a claimant can use the register-event on the blockchain to broadcast his service to interested parties.

5.5.3 Trust registry management

The management tool for authorities is a simple form where public keys of entities can be related to legal entity identifiers.

5.5.4 Acquirer point-of-transaction

For offline (physical) transactions between customer and acquirer, a simple beacon has been developed which broadcasts its public key, together with desire for certain attributes over bluetooth. An example request for being a student and older than 18 is given in listing 3.

The merchant should make sure its public key is registered at the trust registry so the consumer is able to verify the merchants' identity.

```
# <pubkey>|<claim1><?params>|<claim2><?params>|....
# 0|FIRST_NAME
# 1|LAST_NAME
# C|IS_STUDENT
# D|IS_OLDER|AGE
0x6395F09b3ED5E1FD1E482773a6784bC0a79529ed|C|D18
```

Listing 3: Message format of acquirers request

The receipt of the public key over bluetooth is not finished in the developed proof-of-concept. The merchant can scan the QR-code from the consumers' wallet and check the identity provider's smart contracts for the validity of the requested claims.

5.6 Accessing more sensitive data

The use cases of simple claims are limited. However, next to the statements a reference could be made to any other system, which contains the more sensitive data related to the claim. This could be referencing the same public blockchain, a private blockchain or non-blockchain storage. In the following subsection we will give a IPFS-based solution.

5.6.1 Example: IPFS-based pointer

Interplanetary File System (IPFS) can be seen as a decentralized file allocation table (FAT), also used with file systems on storage devices. It allows a small amount of mutability using Interplanetary Naming System (IPNS), allowing content to change under a persistent identifier.

This allows trust registries and claim storage to be supplemented with a reference which contains additional information about a statement. An example is given in listing 4.

```
{
  "pubKey": "b14ab53e38da1c172f877dbc6d65e4a1b0474c3c",
  "kvkNumber": "59581883",
  "validationDate": "20160622",
  "details": "/ipns/XLF2ipQ4jD3UdeX5xp1KBgeHRhemUtaA8Vm/59581883"
}
```

Listing 4: Trust registry entry with IPNS reference

This pointer could reference either the information itself or a standardized descriptor to access this information. The latter could be useful to perform authentication and authorization checks to access sensitive information. An example descriptor to access the detailed Dutch Chamber of Commerce information about an organization is given in listing 5.

```
{
  "pubKey": "b14ab53e38da1c172f877dbc6d65e4a1b0474c3c",
  "kvkNumber": "59581883",
  "validationDate": "20160622",
  "authType": "multiSignature",
  "returnType": {
    "contentType": "application/ld+json",
    "@context": "http://schema.org/",
    "@type": "Organization"
  },
  "accessType": "https",
  "accessUrl": "https://api.kvk.nl/api/v2/profile/companies"
}
```

Listing 5: JSON access descriptor

By using a shared vocabularies like schema.org, which are supported by the bigger search engines, acquirers can make use of the more sensitive data with a lot less effort than compared to proprietary API integration. [92].

5.7 Comparison existing solutions

5.7.1 API integrations

Similar functionality can be obtained by using API integrations. However, this requires a separate integration for each system you want to connect to, of which each needs to be maintained. Furthermore, APIs are often not standardized, which makes integration a very costly undertaking. A high-level model of traditional API integrations is given in figure 27.

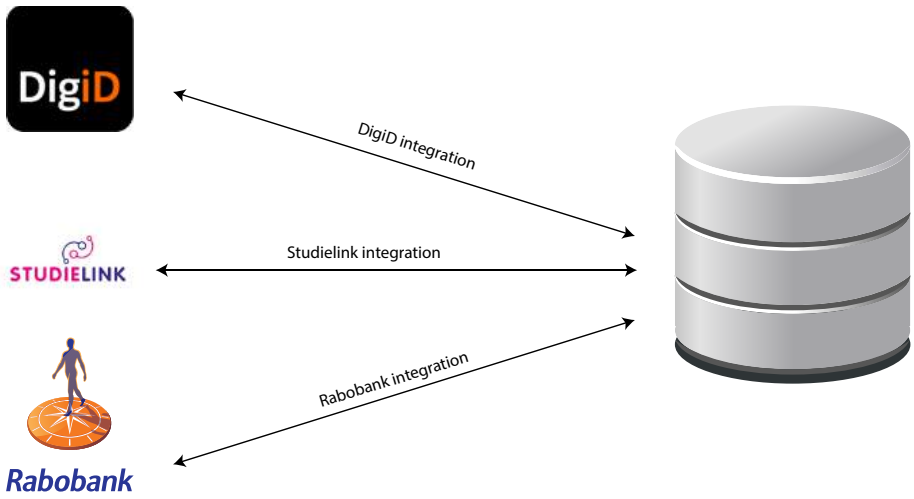


Figure 27: Simplified traditional APIs integration model

Because everything remains stored in the systems at each side of the integration, you are fully dependent on the availability of the service.

With the DIMS you only need to create an integration with the blockchain technology to get access to and integrate with all other entities connected to the system. Because simple claims are stored on the ledger itself, you are still able to make decisions based on that claim when the source systems are temporary or permanently unavailable. A high-level visualization of this decentralized model for comparison is given in figure 28.

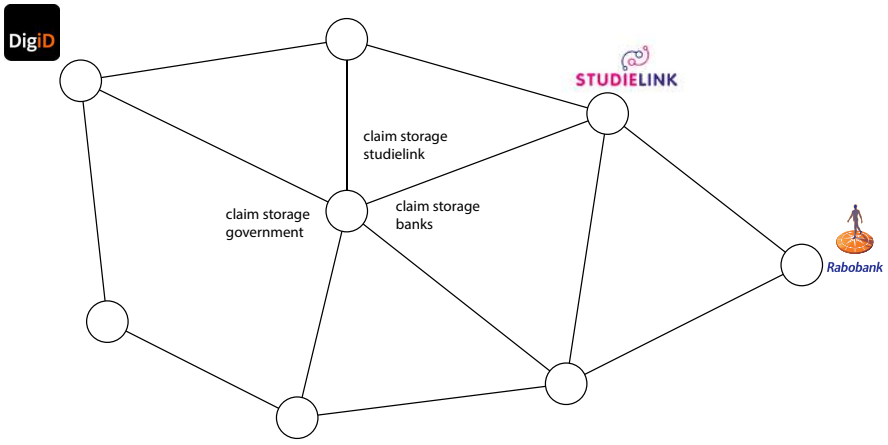


Figure 28: Simplified decentralized model

Because everyone is able to participate in this ecosystem, entry barriers for new entrants are much lower which levels the playing field. Although API responses can be digitally signed, this is not as mandatory as in our designed solution. This could be exploited using man-in-the-middle attacks.

5.7.2 IRMA

Because the design considerations are based on IRMA, it is very similar. IRMA is also decentralized and uses Attribute Based Credential (ABC). Also, it allows selective disclosure of attributes which makes it privacy-friendly.

Because of issuer-unlinkability acquirers seem not able to only allow specific claims from certain providers, like age claims from government only and not from banks.

The platform is only designed for disclosing attributes and does not (yet) offer functionality to add references to more sensitive data. Furthermore, it does not allow for (micro)payments which can be useful for incentivizing quality claim issuance and validation.

5.7.3 Federated authentication

Signed attributes could be used for claim validation but is still dependent on systems of the issuer and the certificate authority. Some federated authentication mechanism use both an open standard and a proprietary standard which makes it more complex and expensive to implement.

Similar to the comparison with API based solutions, there is no dependence on the availability of infrastructure of the issuer for the simple statements with DIMS and claims from different providers can be stacked to create a better quality claim about an entity.

5.7.4 Case study solution

The case study solution is proprietary and only uses the blockchain for explicitly storing consent and data validation using "seals". All participants (consumer, issuer and acquirer) are required to use proprietary software and in the case of issuer and acquirer also need to connect the proprietary solution to their existing IT infrastructure.

The designed solution will be released open-source and implemented as part of existing applications and integrated in existing IT infrastructure or existing API management tooling. This should also help faster adoption of the solution.

Furthermore, the case study solution is fully dependent on the availability of the issuing party and the consumer to relay the information to the acquirer. The designed solution stores simple claims in a privacy-friendly way on the blockchain, and allows the acquirer to retrieve the requested information directly at the issuer after gaining the required consent and proper authentication.

One of the challenges was to reach consensus on the data-model. In the case study solution there was no shared data-model or elaboration of simple data-models. In the developed solution the shared vocabulary at `schema.org` for structured data is suggested, presenting the more sensitive data in a more standard and machine-readable fashion.

5.8 Known issues

5.8.1 Master key linkability

In current proof-of-concept code all claims are linked to single master key, allowing identity providers to learn about other claims not made by themselves. Creating a child key for every attribute and multiple grandchildren to share with each transaction will solved this.

This results in multiple (derived) keys to be shared with the merchant, which sounds cumbersome and inconvenient at first. However, since all of these tasks are automated it does not have an impact on user experience.

5.8.2 Key management

The management of private keys is very important in this solution. If the master key is stored on only one device like a smartphone and are not protected in any way identity theft is very likely. Use of secure storage facilities on devices like Apple's secure enclave should be enforced.

Derived keys could also be given out by governments or banks, this has however some privacy issues because a derived key, no matter how many levels deep can be recognized from a master key.

In the proof-of-concept smart contracts each claim has an expiration date, which let claims expire. Furthermore, it is recommended to add a revocation property for emergencies.

5.8.3 Blockchain forks

On June 18th, an attacker managed to transfer more than 3,6 million ether out of a smart contract known as "The DAO" because of a vulnerability in the application code [115]

Although application programming errors on the platform should not concern the application code itself, the magnitude of the theft was big enough that it could put the option to switch to a proof-of-stake algorithm at stake itself.

The developers chose to perform a hard fork which did not roll back transactions in this case, but it does relocate the stolen funds to another contract to let the original owners withdraw them [111]

This fork shows that blockchains are not strictly immutable [25]. Although Ethereum is still in an early stage, it should be taken in consideration that developers and miners are able to influence the rules of the game and might roll back transactions or change behaviour of decentralized applications using forks.

5.8.4 Block confirmation times

For financial transactions on a blockchain, it is required to wait before the transaction is confirmed in a block, otherwise it could be the case that money is spent more than once. With claims about one's identity double claiming is less of an issue, but the claim should be available for all nodes which can take up to the block generation time. This is currently 15 seconds on average on the public Ethereum blockchain [39].

5.9 Limitations

The designed solution focuses on exchange of claims about ones identity. It does not replace or supplement existing validation processes like Customer Due Diligence. However, it does allow to share (a subset of) the outcome of these processes. To make sure origin of the shared attributes can be trusted, the system uses authorities like central banks and chambers of commerce which keep a curated registry. The resulting business opportunities will be described in the next section.

5.10 Validation

According to the oral explanation and the high-level architecture at the time (July 2016), Marlies Rikken confirms that the designed solution fits within the concept of personal data stores.

Technical Innovation expert Rob Guikers is also involved within API management project within Rabobank. Currently API management (or service integration) is the only manageable method to exchange information in a machine-readable way.

Because of differences between the methods of service integration, this is a time-consuming and work-intensive process. Next to technical issues, it also requires contractual agreement between the parties.

Rob agrees that it a decentralized solution could make this process more efficient if a well-documented API is also part of the solution. It requires support of multiple parties before such solutions will make a difference.

Building upon open standards and sufficiently describing metadata about the used standards, as is seen with the JSON-descriptor in section 5.6.1 will also contribute to lowering barriers for adoption.

Because Rob is also involved is also involved with other projects utilizing blockchain technology, he notes that privacy, scalability and block confirmation times are the biggest challenges.

Although he understands the methods to enhance privacy within the designed solution, he is not yet sure that the used methods will be sufficient. This requires more time to investigate.

He understands how the solution will benefit the consumer in terms of owning their data, but mentions that the consumer is still dependent on entities which need to be trusted, to be able make those claims. It requires adoption by governments and other trusted parties (like banks and supermarkets) before it will be usable in practice. It is however a good foundation for collecting and selectively disclosing attributes from and with multiple parties.

Rob mentioned the easier adoption as advantage over the IRMA solution, since the IRMA solution (at that time) required smartcards and corresponding transceivers.

Releasing the designed solution as open-source is the only way such systems will reach high adoption. Closed systems do not work when trust needs to be established. Next to releasing the solution open-source, it also requires blockchain technology needs to be more performant and scalable.

The Dutch Chamber of Commerce (Kamer van Koophandel) is also considering new technologies like blockchain to improve their service. This should also result in new products and services which fit the purpose of the chamber; the registration of legal entities and offering and providing legal certainty, of course within the framework of privacy laws.

Because digital technology is subject to change and the data-model of the current systems are not very flexible, the duration of projects to develop new products and services is long. An important factor is the complexity of the integration of the new solution within the existing IT landscape. A business model of

the proposed solution sounds very attractive, but should first be subjected to intensive validation to say whether such advantages actually could be achieved.

The Chamber of Commerce is open to further research into the integration of the building blocks in their IT service portfolio. That the building blocks are made available open-source is particularly appealing.

Henk van Cann, an expert in the field of identity and blockchain thinks the architecture of the designed solution sufficiently solves the challenges related to privacy. His presumption is that the scalability will be a minor issue in this solution, since the technical developments on this aspect will go faster than the adoption by identity providers.

Looking at the benefits of the consumer, Henk mentioned that the user experience might be more important than achieving self-sovereign identity. Furthermore, he questions if being fully responsible for your digital identity is something everyone really wants. He expects it could be some hassle in the long term to collect and manage all relevant claims.

6 Business model

This section concerns the research questions:

> *6. What is the business model for business participating in a DIMS?*

Blockchain technology should allow any number of entities to do transactions with each other, without the need to entirely trust each other. Permissioned blockchains create dependencies on trusted third parties in the infrastructure itself. This creates an unfair advantage and heightens entry barriers for new entrants.

Although permissioned blockchains can be linked as sidechains to public and permissioned blockchains to handle more sensitive data, the connecting chain should be public. This allows any entity to connect to the system without having to pay licensing fees.

It will be assumed that the ideal DIMS will be built on a public blockchain, where identities are established at the edge of the network and not in the network itself. This is the reason why we will focus on non-infrastructural roles and its business opportunities in this section.

The presented business models will be based on the developed proof-of-concept.

6.1 Actors

In section 5.4.2, we defined the following actors in the system:

- **Consumer**: The entity who wants to share claims about her identity.
- **Issuer (Identity Provider)**: Some entity who is able to verify claims about an identity
- **Trust Registry**: An entity which curates a registry where it is the authority on
- **Acquirer (Merchant)**: The entity that desires specific claims to be either true or false

6.1.1 Consumer

As seen with big data, information about the consumer is what is valuable to organizations [108]. With a DIMS the consumer should be fully in control about sharing information, she is in a very powerful position.

It allows consumers to ”sell” (parts of) their information for discounts or lower fees. For example, when insurers are able to verify that the consumer is taking good care of herself and her belongings, a discount can be given on insurance premiums.

6.1.2 Issuer

A lot of commercial entities already use authentication of entities to enhance their customers' experience. Depending on the domain the entity operates in, it can be rewarding to participate as identity provider within the DIMS.

It should be noted that because existing authentication mechanisms can be reused and the infrastructure is open, there are almost no fixed costs except for running a blockchain node. Depending on the quality of the claims itself and the volume of deeper checks and its pricing, this can become very rewarding without big investments.

6.1.3 Trust Registry/Authority

As authority, it is expected that traditional database are already used to keep information about the governed entities. Depending on the availability and requirement of authentication for these entities, a one time investment in offering trust registration services for commercial entities on a DIMS helps faster adoption. This heightens transaction volume of the services itself, which in turn could give high revenues from low fee transactions that can be used as investment for other activities which have high costs but low to no revenue.

6.1.4 Acquirer

For acquirers of the information, the most beneficial is being able to acquire attributes from multiple issuers and directly being able to validate it at the source.

It allows outsourcing costly due diligence processes which are similar for each entity. Because banks are required to adhere to strict KYC regulations, acquirers can reuse their validations. This can save a lot of money in on-boarding costs for organizations offering insurances and mortgages.

6.2 Use cases

To illustrate how this DIMS can be used in practice, two use cases will be presented.

6.2.1 Offline: Proving your age at the liquor store

Situation

When buying liquor, you have to prove you are at least 18 or 23 years old when checking out at the cash register. Currently you have to physically hand over a government issued proof of your identity to the cashier who has to calculate the age from your birth date.

Complication

The cashier is obligated to ask for government issued proof, but isn't able to verify if the offered proof is really government issued or forged. Also, it is possible that the proof is borrowed from someone who looks like the person.

When the process is automated with the same government issued proof, the liquor store might be able to track customer behaviour by other available attributes available on the proof, like their social security number which raises a lot of privacy concerns.

Answer

A DIMS could make the claim about ones' identity directly verifiable by the acquirer (liquor store). Multiple claims could be made about the identity by several issuers, but only the required attributes for the transaction can be shared by the buyer. The claim could be "Older than 18" or "Older than 23" and only accepted from the government, or also from banks which can give the same claim about their customers because of the strict CDD processes.

Because only the required attributes need to be disclosed and only the answers to questions (the answer to "older than 18" instead of birth date) and the result of the claim can be encrypted to the public key of the recipient, this is much more privacy-friendly than disclosing all attributes available in your passport.

6.2.2 Online: Collective discount for students on health insurance

Situation

Some health insurance companies offer a collective discount for students following an education programme at selected universities. To verify eligibility documentation to verify this has to be provided. This can be done by sending a copy of a college card or sending a notarized proof of enrollment. Furthermore, some kind of identification of the person to be insured is required.

Complication

Although the required documentation can be scanned and sent using e-mail or using upload forms, documentation is only given at the beginning of each academic year. When students finish their study or cancel their education programme, insurance companies are not able to notice this for at least the beginning of the next academic year, unless they request proof every month.

Because both a scan of a card or a document is machine-readable, this process requires human checks and validation which is error-prone and more expansive than using semantic sources.

Answer

In this case, documentation needs to be provided from several sources. First about the identity of the person to be insured and second if the person to be insured is really a student at a given university.

Like in the offline case, the government can issue claims about the identity of the person to be insured on the DIMS. Since health insurance in some countries is linked to your social security number, this should not be stored on the blockchain itself but on a protected location referenced on the blockchain.

The claim about being enrolled at a specific university or one of the eligible universities could be made available. The validation process can then be fully automated which saves a lot of time and money.

Since the claim of enrollment can be given with a expiration date of one month and only renewed if the person is still studying, insurance companies can use this to continuously validate eligibleness of the discount.

7 Discussion

First of all it should be noted that blockchain technology not the only technology to approach this problem.

Blockchain technology is moving in a fast pace and this should be interpreted as one of the many possible solutions at the time of writing. Furthermore, blockchain and related concepts are very mathematical and technically complex. We tried to explain the basic idea of relevant concepts without going in technical detail.

One of the most interesting implications remains the legal aspect, especially when it comes to finality of transactions and forks (which one will be considered the truth?).

Maturity

The concept of decentralized applications on a decentralized infrastructure is very interesting, especially since it's supposed to be tamper-free. Although the technology is around for several years now, it is still in its early stage.

Exploitation of a decentralized application functioning as a decentralized autonomous organization (see section 5.8.3) shows that the "transparency" of deployed smart contract code doesn't necessarily mean that it is free of errors. Special care should be taken before participating in such projects.

Trust

Blockchain is a platform for decentralized trust, which allows for community-based reputation systems. It is not likely that organizations are willing to accept attributes or claims from pseudonymous sources. Even if they are willing, legislation about customer data, data retention and privacy will -at least for now- not permit this.

It is important to understand that this network of decentralized trust works best for solving the double spending problem, since it is proven that certain currency has been "made". For claims this is different. The network does not know why a claim has been made by a certain entity. The only thing that is non-reputable is that a certain claim is made by a certain entity on a certain moment in time.

Developments

While finalizing this research, the IRMA project made some advancements which were not taken in consideration. This includes the release of source code at http://credentials.github.io/. We tried to get in touch with the people behind IRMA using their website, but unfortunately they did not respond.

The uPort project was released recently which shows similarities to the designed solution, it also uses the Ethereum blockchain, is based on ABC and offers self-sovereign identity. Furthermore, it offers interesting methods to recover your identity when the medium gets lost or stolen [29].

Since IRMA, uPort and the designed solution are all open-source, investigating the possibility to merge (parts of) the projects is expected to further improve possibility of self-sovereign identity.

The developments in the world of blockchain are hard to keep up. State channels is one of the many developments to watch, which is supposed to contribute to challenges related to privacy and scalability [3].

8 Conclusion

Answers to research questions

We will begin with giving answer to the sub research questions.

> *1. What are the properties of current digital identity management systems?*

The domain of identity and access management is very comprehensive. We started with explaining relevant concepts in section 2.1.

Using these concepts we create a classification, which is used to look at running projects related to digital identity, presented in section 2.2.

We then investigated several currently running initiatives to learn about the characteristics which are described in section 2.3.

Looking at the properties of current identity management solutions in production, almost all of them are centralized. The consumer has no full ownership and control over their own attributes.

> *2. What do consumers expect from identity management systems?*

The survey conducted by Innovalor shows that (Dutch) people have the feeling they don't have any control over their personal data, but have the desire to do so. Next to managing who has access to your personal data, people want more insight in who is using their personal data and modify and delete (parts of) this data. This is confirmed as general opinion by multiple publications [36][93][96].

> *3. What does characterize blockchain technology?*

Blockchain is best known as the underlying technology of the Bitcoin cryptocurrency. Although it still is in an early stage, it does allow for the development of decentralized applications. We explained blockchain technology in chapter 3.

The most characterizing property of blockchain is its immutability. Every block contains a hash of the preceding block. This creates a chain of blocks from the first (genesis)block to the current. This makes it computationally impractical to modify information once it is in the chain, because all subsequent blocks should also be regenerated.

The current challenges with blockchain are scalability, privacy and the energy-inefficient Proof of Work (PoW) consensus algorithm.

We investigated where to place blockchain in the ArchiMate modeling language. It became clear that the blockchain technology itself should be put in the infrastructure layer.

Next-generation blockchains like Ethereum offer a turing complete programming language which allows running code that changes the state of the system, this should be put in the application layer.

4. Can blockchain technology be used as infrastructure for identity management?

With the characteristics of blockchain technology in mind, in chapter 4 we first looked at a solution which facilitated the exchange of attributes which were established during KYC-validation at a bank. An exchange could occur between for example banks and mortgage companies.

The company responsible for the proof-of-concept was at the time of the case study relatively small. Their business model heightened adoption barriers by creating a dependence of their own proprietary software next to the blockchain. It showed that the responsibility of blockchain was actually very small, compared to the overall solution. Blockchain technology is only used as storage for the attribute checksums and access control list for data to be exchanged.

Another observation during the proof-of-concept was that all involved parties were using different backend-systems with different data-models, which was not related to the offered solution itself, but gave the insight that other things need to be sorted out before such a system could be useful.

Although the proof-of-concept did not meet the expectations, it showed the potential of blockchain technology as a basis for self-sovereign identity.

The desk research on privacy and confidentiality on blockchain (section 3.2) also shows methods which can be combined with the concept of claims and Attribute Based Credentials (ABCs) to design an architecture for a DIMS.

5. What does an architecture for a DIMS look like?

Based on the answers on the prior research questions, we designed our own Decentralized Identity Management System (DIMS). It should allow decentralized exchange of attributes by the consumer itself, without the need for proprietary software and based on open vocabularies. The main components should be and will be made available open-source which should help to gain a high adoption rate.

It should be run on a permission-less blockchain with enough nodes available to avoid dependence on any entity for the network itself. It will be kept available by all peers. This should also make sure that there is no value in the decentralized network itself but at the edge of the network.

Using a modular architecture for the so-called oracles, it should be easy for most identity providers to create an environment where their users can authenticate as they usually do, present proof of ownership of their private key and then store the relevant claims to a smart contract automatically.

By using claims instead of the raw data, no Personal Identifiable Information (PII) is stored on the blockchain itself. Because it makes use of hierarchical deterministic keys, a child key can be endorsed by the identity provider without the need to authenticate again. By creating a unique derived key for each transaction, attributes can be selectively disclosed. This allows only required data to be exchanged, which makes the solution very privacy-friendly.

Since claims will not always be sufficient for every transaction, we also presented an example how to access more sensitive data, accessible through traditional methods using JSON metadata descriptors on a decentralized file system.

Because blockchain is an infrastructural innovation, end-users should not care whether blockchain is used or more traditional storage systems. The adoption by consumers is an important factor, that is why we decided to create a demo smartphone application that shows that no matter how complex the underlying technology might be, the user interface can still be straightforward and easy to use.

6. What is the business model for business participating in a DIMS?

The connecting chain should be public to allow any entity to connect to the system. This avoids creating unfair advantages for system administrators and keeps the entry barrier for adoption low.

We identified the following business models for the actors in the designed solution:

- Consumer: Their information is what is valuable to organizations [107]. With a DIMS the consumer should be fully in control about sharing information. It allows consumers to "sell" (parts of) their information for discounts or lower fees.
- Issuer: A lot of commercial entities already use authentication of entities to enhance their customers experience. Existing authentication mechanisms can be reused, there are almost no fixed costs except for running a blockchain node. Depending on the quality of the claims itself and the volume of deeper checks and its pricing, this can become very rewarding.
- Trust Registry: As authority it is expected that traditional database are already used to keep information about the governed entities. Because they can offer more certainty with their validations, selling them for low enough prices could generate high revenues from low fee transactions.
- Acquirer: The most beneficial is being able to acquire attributes from multiple issuers and directly being able to validate it at the source. It also allows outsourcing costly due diligence processes which are similar for each entity.

Answer to main research question

In section 1.3.2 the following research question was defined:

> *How to design identity management architecture that is decentralized so that entities can exchange attributes and verify claims without being dependent on a single central authority?*

The answers to the sub research question lead to a proof-of-concept of a Decentralized Identity Management System (DIMS) based on blockchain technology.

The design is dependent on a storage layer where everyone can be sure that if claims are made, it is non-reputable that the claim has been made and by what entity. In our solution, blockchain technology provides this functionality

by replacing the need for a central authority using cryptographically secured consensus algorithms.

To become useful, all relevant actors need to adopt the system. Our case study showed that proprietary solutions are undesirable if trust is important.

We designed modular building blocks for all relevant actors to participate in the ecosystem, and will make it open-source to allow everyone to implement and customize it to their own needs. Our hope is that this will lower adoption barriers and work together to a self-sovereign identity.

Code of the proof-of-concept will be published at `https://github.com/djurid` and `http://djurid.me/`.

Future work

Because of the possibility of new interactions between all kinds of entities, a lot of legislation questions arose when discussing especially the exchange of more sensitive data attributes. An attempt was made to get an idea about this, but appears to be a too much exhaustive and out of scope of this research.

The scalability problem seems to be the biggest challenge with public blockchains, which might be partially solved with sidechains. Furthermore, there seems to be some developments in centrally managed blockchains, where the main reason for centralizing some parts of the technology means more security for the end-users. When still allowing all users to "read" the blockchain, it still offers more transparency and more efficient ways of doing business.

A lot of interesting developements with regard to Ethereum were presented on DevCon 2. This included "Raiden Network", a solution which could improve performance for cryptocurrency payments. It could be useful to explore how this technique could be applied to the designed solution.

9 Acknowledgements

I would like to thank Innovalor for providing their questionnaire information which showed a clear desire of the Dutch people for more control over their digital identity and introduction to the concept of personal data stores. Also, I want to thank everyone at Rabobank Group for their critical feedback and helpfulness.

Furthermore, I am very grateful to the organizers of the meet-ups and conferences about blockchain in Amsterdam and Utrecht, where I met a lot of great people and learned a lot about blockchain and its many use cases.

Thank you Andrew Mooijman, Henk van Cann, Marlies Rikken, Perry Smit and Rob Guikers for your contribution as experts.

Last but not least I would like to thank my supervisors Hans Moonen, Marten van Sinderen and Roel Steenbergen for their guidance, feedback and patience.

10 References

[1] *Address reuse - Bitcoin Wiki*. URL: https://en.bitcoin.it/wiki/Address%7B%5C_%7Dreuse (visited on 04/08/2016).

[2] Christopher Allen. *The Path to Self-Sovereign Identity*. 2016. URL: http://www.coindesk.com/path-self-sovereign-identity/ (visited on 07/05/2016).

[3] Ian Allison. *Ethereum's Vitalik Buterin explains how state channels solve privacy and scalability*. 2016. URL: http://www.ibtimes.co.uk/ethereums-vitalik-buterin-explains-how-state-channels-address-privacy-scalability-1566068 (visited on 10/23/2016).

[4] Gergely Alpár and Jaap–Henk Hoepman. "A secure channel for attribute-based credentials". In: *Proceedings of the 2013 ACM workshop on Digital identity management*. ACM. 2013, pp. 13–18.

[5] Gergely Alpár, Jaap-Henk Hoepman, and Johanneke Siljee. "The Identity Crisis. Security, Privacy and Usability Issues in Identity Management". In: *arXiv preprint arXiv:1101.0427* (2011), p. 15. arXiv: 1101.0427. URL: http://arxiv.org/abs/1101.0427.

[6] Gergely Alpár and Bart Jacobs. "Credential design in attribute-based identity management". In: *Bridging distances in technology and regulation, 3rd TILTing Perspectives Conference*. 2013, pp. 189–204.

[7] Andreas M Antonopoulos. *Mastering Bitcoin: unlocking digital cryptocurrencies*. " O'Reilly Media, Inc.", 2014, pp. 175–213.

[8] *Artists & Creators | ascribe*. URL: https://www.ascribe.io/ (visited on 08/21/2016).

[9] Adam Back et al. "Enabling blockchain innovations with pegged sidechains". In: (2014).

[10] Dominick Baier et al. *A Guide to Claims-Based Identity and Access Control: Authentication and Authorization for Services and the Web*. Second Edi. Microsoft Press, 2011, p. 374. ISBN: 9780735640597.

[11] Sriram Balasubramaniam et al. "Identity management and its impact on federation in a system-of-systems context". In: *Systems conference, 2009 3rd annual IEEE*. IEEE. 2009, pp. 179–182.

[12] Bank for International Settlements. "Basel Committee Publications - Customer due diligence for banks (Consultative Document, Issued for comment by 31 March 2001) - Jan 2001". In: (2001). URL: http://www.bis.org/publ/bcbs85.htm.

[13] *BBVA Bagged a Bargain with Simple | anthemis*. 2014. URL: http://www.anthemis.com/bbva-bagged-a-bargain-with-simple/ (visited on 08/13/2016).

[14] Eli Ben Sasson et al. "Zerocash: Decentralized anonymous payments from bitcoin". In: *Security and Privacy (SP), 2014 IEEE Symposium on*. IEEE. 2014, pp. 459–474.

[15] Abhilasha Bhargav–Spantzel, Anna C. Squicciarini, and Elisa Bertino. *Establishing and protecting digital identity in federation systems*. Jan. 2006. DOI: 10.3233/JCS-2006-14303. URL: http://content.iospress.com/articles/journal-of-computer-security/jcs261.

[16] Anol Bhattacherjee. *Social Science Research: principles, methods, and practices*. Vol. 9. 2012, p. 144. ISBN: 9781475146127. DOI: 10.1186/1478-4505-9-2.

[17] David Birch. *Blockchain revolution*. Amsterdam, 2016. URL: http://www.slideshare.net/15Mb/blockchain-revolution (visited on 10/21/2016).

[18] David Birch. *Identity is the New Money*. Ed. by Ed Conway. London Publishing Partnership, 2014. ISBN: 9781907994128.

[19] *Bitcoin Blockchain Size*. URL: https://blockchain.info/charts/blocks-size (visited on 06/27/2016).

[20] *Bitcoinism: Reclaiming Financial Privacy With {HD} Wallets*. URL: http://bitcoinism.blogspot.nl/2013/07/reclaiming-financial-privacy-with-hd.html (visited on 02/16/2016).

[21] *Block – Bitcoin Wiki*. URL: https://en.bitcoin.it/wiki/Block (visited on 04/04/2016).

[22] *Block chain – Bitcoin Wiki*. URL: https://en.bitcoin.it/wiki/Block%7B%5C_%7Dchain (visited on 04/04/2016).

[23] *Block hashing algorithm – Bitcoin Wiki*. URL: https://en.bitcoin.it/wiki/Block%7B%5C_%7Dhashing%7B%5C_%7Dalgorithm (visited on 04/04/2016).

[24] Vitalik Buterin. "A next-generation smart contract and decentralized application platform". In: *White Paper* (2014).

[25] Vitalik Buterin. *Hard Fork Completed - Ethereum Blog*. 2016. URL: https://blog.ethereum.org/2016/07/20/hard-fork-completed/ (visited on 07/24/2016).

[26] Coindesk. *Bitcoin glossary - {CoinDesk}*. URL: http://www.coindesk.com/information/bitcoin-glossary (visited on 02/16/2016).

[27] *CoinDesk London Workshop Explores Blockchain Identity in Finance - CoinDesk*. URL: http://www.coindesk.com/benefits-challenges-blockchain-identity/ (visited on 05/31/2016).

[28] *{CoinJoin}: Bitcoin privacy for the real world*. URL: https://bitcointalk.org/index.php?topic=279249 (visited on 02/16/2016).

[29] ConsenSys. *uPort*. 2016. URL: https://uport.me/%7B%5C#%7Dhome (visited on 10/01/2016).

[30] *Craig Stuntz's Weblog : What is Homomorphic Encryption, and Why Should I Care?* URL: http://blogs.teamb.com/craigstuntz/2010/03/18/38566/ (visited on 04/08/2016).

[31] *[Crypto] Compact Confidential Transactions for Bitcoin*. URL: https://bitcointalk.org/index.php?topic=1085436.msg13765368%7B%5C#%7Dmsg13765368%7B%5C%7D (visited on 02/16/2016).

[32] *Crypto-Currency Market Capitalizations*. URL: https://coinmarketcap.com/ (visited on 08/21/2016).

[33] Daphne Riksen. "'In de afweging tussen functionaliteit en privacy legt privacy het altijd af'". In: *Surf Magazine* september (Sept. 2013), pp. 4–6.

[34] *DNB Guidance on the Anti-Money Laundering and Counter-Terrorist Financing Act and the Sanctions Act Preventing the misuse of the financial system for money laundering and terrorist financing purposes and controlling integrity risks*. Tech. rep. Version 3.0. De Nederlandsche Bank, 2015.

[35] *Double Your Money? Looming 'Hard Fork' Uncovers Fatal Bitcoin Flaw - Forbes*. URL: http://www.forbes.com/sites/jasonbloomberg/2016/02/07/double-your-money-looming-hard-fork-uncovers-fatal-bitcoin-flaw/%7B%5C#%7D516b06f3cef6 (visited on 04/08/2016).

[36] Chair Digital Economy. *Digital Identity 3.0 The Platform for People*. Tech. rep.

[37] ElementsProject. *ElementsProject/elementsproject.github.io*. URL: %7Bhttps://github.com/ElementsProject/elementsproject.github.io%7B%5C%7D%7D (visited on 02/16/2016).

[38] Rachel Louise Ensign. *PayPal to Pay $7.7 Million to U.S. Over Alleged Sanctions Violations - WSJ*. Mar. 2015. URL: http://www.wsj.com/articles/paypal-to-pay-7-7-million-to-u-s-over-alleged-sanctions-violations-1427312161 (visited on 08/16/2016).

[39] *ethstats.net*. URL: https://ethstats.net/ (visited on 09/08/2016).

[40] *fabric/protocol-spec.md at master ů hyperledger/fabric*. URL: https://github.com/hyperledger/fabric/blob/master/docs/protocol-spec.md (visited on 06/24/2016).

[41] *Factom: A Data Layer for the Blockchain*. URL: https://www.factom.com/factom-a-data-layer-for-the-blockchain/ (visited on 04/01/2016).

[42] *GDI*. URL: https://www.digitaleoverheid.nl/digitaal-2017/digitalisering-aanbod/gdi (visited on 05/20/2016).

[43] Uwe Glässer and Mona Vajihollahi. "Security Informatics". In: ed. by C. Christopher Yang et al. Boston, MA: Springer US, 2010. Chap. Identity Management Architecture, pp. 97–116. ISBN: 978-1-4419-1325-8. DOI: 10.1007/978-1-4419-1325-8_6. URL: http://dx.doi.org/10.1007/978-1-4419-1325-8_6.

[44] Michele D. Guel. "A Framework for Choosing Your Next Generation Authentication/Authorization System". In: *Information Security Technical Report* 7.1 (Mar. 2002), pp. 63–78. ISSN: 13634127. DOI: 10.1016/S1363-4127(02)00107-3. URL: http://www.sciencedirect.com/science/article/pii/S1363412702001073.

[45] Greg Guest, Kathleen M MacQueen, and Emily E Namey. *Applied thematic analysis*. Sage, 2011.

[46] AR Hevner et al. "Design science in information systems research". In: *MIS quarterly* 28.1 (2004), pp. 75–105. URL: http://dl.acm.org/citation.cfm?id=2017217.

[47] Erik Hofstee. "Constructing a good dissertation". In: *Johannesburg: EPE* (2006).

[48] Xinyi Huang et al. “Preserving Transparency and Accountability in Optimistic Fair Exchange of Digital Signatures”. In: *IEEE Transactions on Information Forensics and Security* 6.2 (June 2011), pp. 498–512. ISSN: 1556-6013. DOI: 10.1109/TIFS.2011.2109952. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5706363.

[49] *Hybrid Proof-of-Work Proof-of-Stake – Decred Wiki*. URL: https://wiki.decred.org/Introduction%7B%5C#%7DHybrid%7B%5C_%7DProof-of-Work%7B%5C_%7DProof-of-Stake (visited on 04/08/2016).

[50] *I Reveal My Attributes | IRMA (I Reveal My Attributes) project*. URL: https://www.irmacard.org/irma/%7B%5C#%7D05 (visited on 04/08/2016).

[51] InnoValor. *Persoonlijke data, onder controle? - InnoValor*. 2016. URL: https://innovalor.nl/personal-data-store/ (visited on 05/23/2016).

[52] Interledger. *interledger/rfcs: A set of draft specifications for enabling interledger payments*. URL: https://github.com/interledger/rfcs (visited on 08/28/2016).

[53] *Introducing Casper ”the Friendly Ghost” - Ethereum Blog*. URL: https://blog.ethereum.org/2015/08/01/introducing-casper-friendly-ghost/ (visited on 04/08/2016).

[54] *Jumio Turns Webcams Into Credit Card Readers, on Desktop and Mobile - NYTimes.com*. URL: http://www.nytimes.com/external/readwriteweb/2011/07/26/26readwriteweb-jumio-turns-webcams-into-credit-card-reader-15531.html (visited on 05/20/2016).

[55] Sunny King and Scott Nadal. “Ppcoin: Peer-to-peer crypto-currency with proof-of-stake”. In: *self-published paper, August* 19 (2012).

[56] Merel Koning, Paulan Korenhof, and Gergely Alpár. “The ABC of ABC – An Analysis of Attribute-Based Credentials in the Light of Data Protection, Privacy and Identity”. In: (2014).

[57] Gregg Kreizman, John Pescatore, and Ray Wagner. *The U.S. Government’s Adoption of SAML 2.0 Shows Wide Acceptance*. Tech. rep. Gartner, Inc., 2007.

[58] Jae Kwon. “TenderMint: Consensus without Mining”. In: (2014).

[59] *Ledger structure – Openchain 0.5.0 documentation*. URL: https://docs.openchain.org/en/latest/api/ledger.html (visited on 04/08/2016).

[60] D. Malone and K.J. ODwyer. “Bitcoin Mining and its Energy Footprint”. In: *25th IET Irish Signals & Systems Conference 2014 and 2014 China-Ireland International Conference on Information and Communities Technologies (ISSC 2014/CIICT 2014)*. Institution of Engineering and Technology, 2014, pp. 280–285. ISBN: 978-1-84919-924-7. DOI: 10.1049/cp.2014.0699. URL: http://digital-library.theiet.org/content/conferences/10.1049/cp.2014.0699.

[61] Sarah Meiklejohn. “An Exploration of Group and Ring Signatures”. In: (2011).

[62] *MIT Madars Virza: Bitcoin Privacy Issues and How Zerocash Can Help.* URL: http://insidebitcoins.com/news/mits-madars-virza-bitcoin-privacy-issues-and-how-zerocash-can-help/30751 (visited on 04/08/2016).

[63] Yves-Alexandre de Montjoye et al. "openPDS: protecting the privacy of metadata through SafeAnswers." In: *PloS one* 9.7 (Jan. 2014), e98790. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0098790. URL: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0098790.

[64] Andrew Mooijman et al. *PoC KYC on blockchain with Tradle.* Tech. rep. Utrecht: Rabobank Nederland, 2016.

[65] *MSN Historical Timeline: A brief history of milestone events in the life of MSN from the past ten years.* 2005. URL: http://web.archive.org/web/20050618082125/http://www.microsoft.com/presspass/press/2002/nov02/11-08MSN8GlobalTimeLine.mspx (visited on 07/09/2016).

[66] Satoshi Nakamoto. *Bitcoin: A peer-to-peer electronic cash system.* 2008.

[67] NXT Community. "Nxt Whitepaper". In: (2014), pp. 1–28. URL: https://wiki.nxtcrypto.org/wiki/Whitepaper:Nxt.

[68] *Off-Chain Transactions - Bitcoin Wiki.* URL: https://en.bitcoin.it/wiki/Off-Chain%7B%5C_%7DTransactions (visited on 02/16/2016).

[69] *On Public and Private Blockchains - Ethereum Blog.* URL: https://blog.ethereum.org/2015/08/07/on-public-and-private-blockchains/ (visited on 04/08/2016).

[70] *On Settlement Finality - Ethereum Blog.* URL: https://blog.ethereum.org/2016/05/09/on-settlement-finality/ (visited on 06/27/2016).

[71] *OneName: The Bridge Between Physical & DigitalIdentity | Blockchain for the Billions on WordPress.com.* URL: https://rywalk.wordpress.com/2015/02/13/onename-the-bridge-between-physical-digital-identity/ (visited on 05/20/2016).

[72] *OP_RETURN and the Future of Bitcoin - Bitzuma.* URL: http://bitzuma.com/posts/op-return-and-the-future-of-bitcoin/ (visited on 08/21/2016).

[73] Richard Paap. "Modelleren van Scanner Panel Data". In: *Medium Econometrische Toepassingen: MET/uitg. door het Econometrisch Dispuut van de Economische Faculteitsvereniging Rotterdam* (2001), pp. 12–15.

[74] Andreas Pashalidis and Chris J Mitchell. "Information Security and Privacy: 8th Australasian Conference, ACISP 2003 Wollongong, Australia, July 9–11, 2003 Proceedings". In: ed. by Rei Safavi-Naini and Jennifer Seberry. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. Chap. A Taxonomy of Single Sign-On Systems, pp. 249–264. ISBN: 978-3-540-45067-2. DOI: 10.1007/3-540-45067-X{_}22. URL: http://dx.doi.org/10.1007/3-540-45067-X%7B%5C_%7D22.

[75] *Passport.* URL: http://passportjs.org/ (visited on 10/21/2016).

[76] *Patricia Tree - ethereum/wiki wiki.* URL: https://github.com/ethereum/wiki/wiki/Patricia-Tree (visited on 06/07/2016).

[77] Siani Pearson and Andrew Charlesworth. "Accountability as a way forward for privacy protection in the cloud". In: *Cloud computing*. Springer, 2009, pp. 131–144.

[78] Peter Evans-Greenwood, Robert Hillard, Ian Harper, Peter Williams. *Bitcoin, Blockchain & distributed ledgers Caught between promise and reality*. Tech. rep. 2016. URL: https://www2.deloitte.com/content/dam/Deloitte/au/Images/infographics/au-deloitte-technology-bitcoin-blockchain-distributed-ledgers-180416.pdf.

[79] Antonio Piedra, Jaap-Henk Hoepman, and Pim Vullers. "Cryptology and Network Security: 13th International Conference, CANS 2014, Heraklion, Crete, Greece, October 22-24, 2014. Proceedings". In: ed. by Dimitris Gritzalis, Aggelos Kiayias, and Ioannis Askoxylakis. Cham: Springer International Publishing, 2014. Chap. Towards a Full-Featured Implementation of Attribute Based Credentials on Smart Cards, pp. 270–289. ISBN: 978-3-319-12280-9. DOI: 10.1007/978-3-319-12280-9{_}18. URL: http://dx.doi.org/10.1007/978-3-319-12280-9%7B%5C_%7D18.

[80] Marc Pilkington. "Blockchain Technology: Principles and Applications". In: (Sept. 2015). URL: http://papers.ssrn.com/abstract=2662660.

[81] *PKI - Public Key Infrastructure - What is it? | Comodo*. URL: https://www.comodo.com/resources/small-business/digital-certificates1.php (visited on 07/11/2016).

[82] *PKIoverheid - Logius*. URL: https://www.logius.nl/diensten/pkioverheid/ (visited on 05/20/2016).

[83] Andrew Poelstra et al. *Distributed Consensus from Proof of Stake is Impossible*. 2014.

[84] Franz Stefan Preiss. "Minimizing Information Disclosure in Authentication Transactions with Attribute-Based Credentials (Minimalisatie van vrijgegeven informatie in authenticatietransacties met behulp van attribuutgebaseerde credentials)". In: (2012).

[85] *Problemen in Bitcoin–land: Wat is er aan de hand? | PCM*. URL: http://www.pcmweb.nl/nieuws/problemen-bitcoin-land-wat-er-aan-de-hand.html (visited on 04/08/2016).

[86] *Proof of Phone*. URL: https://www.proofofphone.com/ (visited on 05/20/2016).

[87] *Proof of Physical Address*. URL: https://proofofphysicaladdress.com/ (visited on 05/20/2016).

[88] *Qiy Foundation | Technology*. URL: https://www.qiyfoundation.org/qiy-scheme/what-is-a-scheme/technology/ (visited on 05/20/2016).

[89] David Recordon and Drummond Reed. "OpenID 2.0: A Platform for User-centric Identity Management". In: *Proceedings of the Second ACM Workshop on Digital Identity Management*. DIM '06. Alexandria, Virginia, USA: ACM, 2006, pp. 11–16. ISBN: 1-59593-547-9. DOI: 10.1145/1179529.1179532. URL: http://doi.acm.org/10.1145/1179529.1179532.

[90] *Reducing the Cost of Anti-Money Laundering Compliance*. Tech. rep. accenture, 2015.

[91] Fergal Reid and Martin Harrigan. "An Analysis of Anonymity in the Bitcoin System". In: (July 2011), p. 28. arXiv: 1107.4524. URL: http://arxiv.org/abs/1107.4524.

[92] Jason Ronallo. "HTML5 Microdata and Schema. org". In: *Code4Lib Journal* 16 (2012).

[93] John Rose, Olaf Rehse, and Björn Röber. "The value of our digital identity". In: *Boston Cons. Gr* (2012).

[94] Tim Ruffing, Pedro Moreno-Sanchez, and Aniket Kate. "CoinShuffle: Practical decentralized coin mixing for Bitcoin". In: *Computer Security-ESORICS 2014*. Springer, 2014, pp. 345–364.

[95] Patrick Salyer. *A Step Back in Time: The History and Evolution of Digital Identity*. 2015. URL: http://www.iotevolutionworld.com/iot/articles/410328-step-back-time-history-evolution-digital-identity.htm (visited on 06/09/2016).

[96] Christine Satchell et al. "Identity crisis: user perspectives on multiplicity and control in federated identity management". In: *Behaviour & Information Technology* 30.1 (2011), pp. 51–62.

[97] *Scalability - Bitcoin Wiki*. URL: https://en.bitcoin.it/wiki/Scalability (visited on 06/27/2016).

[98] David Schwartz, Noah Youngs, and Arthur Britto. "The Ripple protocol consensus algorithm". In: *Ripple Labs Inc White Paper* (2014), p. 5.

[99] *Script - Bitcoin Wiki*. URL: https://en.bitcoin.it/wiki/Script (visited on 06/07/2016).

[100] Lu Jack Johnston David Kirby Peter Snow Paul Deery Brian. *Business Processes Secured by Immutable Audit Trails on the Blockchain*. Tech. rep. 2014.

[101] H.F. Spenkelink. *The Adoption Process of Cryptocurrencies - Identifying factors that influence the adoption of cryptocurrencies from a multiple stakeholder perspective*. Aug. 2014. URL: http://essay.utwente.nl/65677/1/Spenkelink%7B%5C_%7DMA%7B%5C_%7DMG.pdf.

[102] Rob van der Staaij. *Handboek identity & access management*. Academic Service, 2014. ISBN: 9789462450882.

[103] *Statistics - etherchain.org - The ethereum blockchain explorer*. URL: https://etherchain.org/statistics/basic (visited on 06/27/2016).

[104] Anselm Strauss, Juliet Corbin, et al. *Basics of qualitative research*. Vol. 15. Newbury Park, CA: Sage, 1990.

[105] *SURF | Op SURFconext aangesloten diensten*. URL: https://www.surf.nl/diensten-en-producten/surfconext/op-surfconext-aangesloten-diensten/index.html (visited on 05/04/2016).

[106] Tim Swanson. *Settlement Risks Involving Public Blockchains*. Mar. 2016. URL: http://tabbforum.com/opinions/settlement-risks-involving-public-blockchains (visited on 07/05/2016).

[107] Nick Szabo. "The idea of smart contracts". In: *Nick Szabo's Papers and Concise Tutorials* (1997).

[108] Curtis R Taylor. "Consumer privacy and the market for customer information". In: *RAND Journal of Economics* (2004), pp. 631–650.

[109] *Tendermint*. URL: http://tendermint.com/ (visited on 06/27/2016).

[110] *The Fidor Bank Story, Frank Schwab, Fidor TecS AG*. 2014. URL: http://www.slideshare.net/ashridge/the-fidor-bank-story-frank-schwab-fidor-tecs-ag (visited on 08/28/2016).

[111] *The Hard Fork: What's About to Happen to Ethereum and The DAO - CoinDesk*. URL: http://www.coindesk.com/hard-fork-ethereum-dao/ (visited on 09/08/2016).

[112] Stefan Thomas and Evan Schwartz. "A protocol for interledger payments". In: *URL https://interledger. org/interledger. pdf* (2015).

[113] *TrustTester: Secure validation of personal data | TNO*. URL: https://www.tno.nl/en/focus-area/industry/networked-information/information-creation-from-data-to-information/trusttester-secure-validation-of-personal-data/ (visited on 05/20/2016).

[114] Joseph Turow, Lauren Feldman, and Kimberly Meltzer. "Open to Exploitation: America's Shoppers Online and Offline". In: (2005).

[115] *Understanding The DAO Attack - CoinDesk*. URL: http://www.coindesk.com/understanding-dao-hack-journalists/ (visited on 09/08/2016).

[116] *uPort The Wallet is the New Browser - Medium*. URL: https://medium.com/@ConsenSys/uport-the-wallet-is-the-new-browser-b133a83fe73%7B%5C#%7D.1l0vsfq2p (visited on 05/20/2016).

[117] Max Van Kleek et al. "Social Personal Data Stores: the Nuclei of Decentralised Social Machines". In: *Proceedings of the 24th International Conference on World Wide Web Companion*. International World Wide Web Conferences Steering Committee. 2015, pp. 1155–1160.

[118] *Verordening (EU) nr. 910/2014 van het Europees Parlement en de Raad van 23 juli 2014 betreffende elektronische identificatie en vertrouwensdiensten voor elektronische transacties in de interne markt en tot intrekking van Richtlijn 1999/93/EG*. URL: http://eur-lex.europa.eu/legal-content/NL/TXT/?uri=uriserv:OJ.L%7B%5C_%7D.2014.257.01.0073.01.NLD (visited on 05/20/2016).

[119] "Visa Inc. at a Glance". In: (). (Visited on 06/27/2016).

[120] *World Internet Users Statistics and 2015 World Population Stats*. URL: http://www.internetworldstats.com/stats.htm (visited on 03/21/2016).

[121] *Zcash, an Untraceable Bitcoin Alternative, Launches in Alpha | WIRED*. URL: http://www.wired.com/2016/01/zcash-an-untraceable-bitcoin-alternative-launches-in-alpha/ (visited on 04/08/2016).