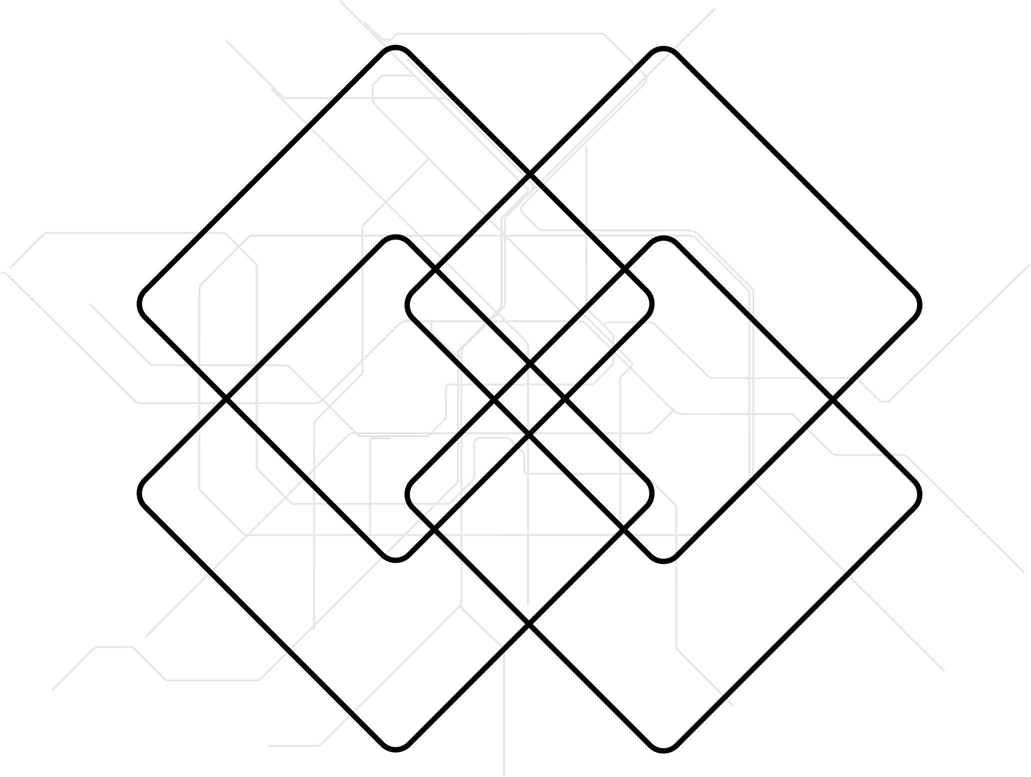


Decentralized Identity Management for Public Transporation

Authors	Georgios Kokosioulis (121735) & Lukas Stockburger (121695)
Programme	MSc Business Administration and E-Business
Course	Master's Thesis
Place, Date	Copenhagen, 16.03.2020
Supervisor	Raghava Rao Mukkamala
Pages/Characters	98/192.244

CBS COPENHAGEN BUSINESS SCHOOL

Table of Contents

Table of Figures

Table of Tables

Table of Abbreviations

Abbreviation	**Term**
ABT	Account Based Ticketing
BVG	Berliner Verkehrsbetriebe
CBS	Copenhagen Business School
DID	Decentralized Identifier
DIF	Decentralized Identity Foundation
DIdMS	Decentralized Identity Management System
DLT	Distributed Ledger Technology
DPKI	Decentralized Public Key Infrastructure
DSB	Danish State Railways
EMC	Euro Mobility Card
GDPR	General Data Protection Regulation
KYC	Know Your Customer
PTA	Public Transportation Authority
PTT	Public Transportation Ticketing
PT	Public Transport
IdM	Identity Management
IdMS	Identity Management System
SSI	Self-sovereign Identity
sEMC	Student Euro Mobility Card

Abstract

Identity Management on the Internet has been a challenging topic over the last few years. Personal information theft and identity data breaches are not uncommon and are often a result of current Identity Management practices and system architectures. Identity management is largely centralized, harming the privacy of users. With the rise of blockchain technology, a new concept of Identity Management is evolving called Self-sovereign Identity (SSI). This new paradigm has been ignited by the evolution of digital Identity Management Systems and distributed ledgers. In this research, blockchain technology is examined to determine how a Decentralized Identity Management System (DIdMS) could be used for public transportation. In order to exploit its full potential, it is essential to review attributes of a Self-sovereign Identity as well as analyze existing DIdMS. The paper describes how the system utilizes blockchain technology to provide a high-level of security, trust and transparency for all the involved parties. The aim is to develop a system that serves as a proof-of-concept and provides a Self-sovereign Identity to the users. Following the Design Science Research methodology, the proposed system has been analyzed based on existing ticketing solutions, and principles of SSI. From these requirements a prototype was developed to showcase how passengers can utilize a standardized travel credential that's valid across different transportation networks in Europe. This removes the barriers of having multiple travel cards for each transportation provider. Moreover, it empowers individuals to have full control over their identities while creating an interoperable ticketing system across Europe. This aligns with the goal of the European Union to create a Single Transportation Market by 2050 while focusing on privacy and data integrity between different Public Transportation Providers.

Keywords: Identity Management, Self-sovereign Identity, Blockchain, Distributed Ledger Technology, Public transportation

Chapter 1

1. Introduction

Over the years technology has become a major influence on our society. Old systems and paradigms have been shifted and disrupted by the introduction of novel technologies like the Internet. The Internet and its underlying protocol layer introduced the first global interconnected system that could be used by anyone. Since today it is the backbone of many applications and is one of the drivers of globalization and interconnectivity. What first started as a data transfer medium is now disrupting whole industries like the mobility sector. The rise of *Intelligent Transport System (ITS)* allows for improved transportation and traffic management systems in and around cities; lowering traffic congestion, CO2 emission and improved transportation systems. This has also been recognized by the United Nations as one of the 17 Sustainable Development Goals (United Nations, 2015). Thus, Goal 11 recognizes the need for more sustainable cities including the need for improved public transportation options. This leads to new and more sustainable transportation concepts like Mobility as a Service (Holmberg et al., 2016). Here different types of mobility modes are integrated into one single service. This includes, for example, ridesharing offers, public transportation, or ride-hailing applications like “Uber”. However, all those systems seem to work in isolation without being interoperable. Especially the public transportation systems in cities are usually highly isolated. Ticketing systems are dependent on single transportation systems which are linked to the city or country they are operating in. However, the European Union aims to create a smart ticketing solution until 2050, whose objective is to create an interoperable solution between public transportation authorities around Europe (Urban ITS Expert Group, 2013). This will allow users to have a single point of entry into any public transportation system within Europe.

The current public transportation landscape is highly scattered around different solutions, pricing models, languages, etc. Some systems are more advanced and offer integrated solutions around different transportation modes within a country, whereas other systems only work within a specific network of a city. Here ticketing solutions range from Paper Tickets and Smart Cards over to Account based solutions. For example, the public ticketing system OV Chipkaart (Smart Card) offers one solution for all public transport modes in the Netherlands, whereas the public transportation system in Berlin (BVG) only issues tickets that are linked to its own services and not across Germany. The scattered landscape comes along with many challenges. One of those challenges is the management of user accounts across independent public transportation. Issues like data integrity, data privacy as well as data ownership of such systems are important considerations when implementing a pan European or even global transportation system. Moreover, the internet not only offered an easy way of data exchange and connectivity but also gave rise to many data scandals like Cambridge Analytica in early 2018 (Kozlowska, 2018). Here centralized platforms misused private data of their users for their own benefit or faulty security measures lead to data breaches by external attackers. However, the Internet has evolved over time and improved its security standards by introducing more secure protocols like HTTPS. Yet, it still suffers from scandals around centralized applications that are built around it.

Over time the idea of the internet as a global protocol has been manifested in many other follow up technologies. One of those technologies is Blockchain Technology, which allows for high data integrity and data ownership between all stakeholders. Blockchain systems are known for high transparency which is important when dealing with many different stakeholders like different transportation authorities. However, the use of any service usually comes with some kind of account management that allows users to access and manage their data. In public transportation, for example, accounts can be used to manage account balances, ride histories or user settings. Thus, public transportation providers are highly dependent on the ability to allow users to manage their accounts in order to create the best possible user experience. Usually, *Identity Management Systems (IdMS)* are hosted in centralized databases that are controlled and managed by the service providing authority. However, organizations like the *World Wide Web Consortium (W3C)* have lately been working on new concepts and standards for decentralized identity solutions. Here *Decentralized Identifiers (DIDs)* are introduced. Those identifiers allow linking users with their associated data without relying on third parties. Thus, users not only gain more control over their own data but also gain overall sovereignty from existing systems. This leads exposing them less to risks of data misuse by more centralized systems. The ability to own and control your own data is one of the core principles of Self-sovereign Identity. Those fundamentals are also rooted in the concept of Blockchain Technology. Moreover, the introduction of new data protection legislation like the GDPR in Europe illustrates the need for a paradigm shift when it comes to privacy, user data and control. A unified IdMS for public transportation in Europe will involve many stakeholders from different countries. This system generates sensitive user data through ride histories, GPS locations, account balances and other account-related personal information, etc. Thus, exposing users to many new data risks potentially handled by a variety of publicly as well as privately held transportation companies. This raises the overall question of how such an Identity Management System for the public transport sector could look like.

1.1 Motivation

The amount of services online is increasing on a steady basis. This also means that more user data is aggregated and new accounts are created on a daily basis. However, many of those accounts are still living in isolation and users need to create new identities on any new platform they are using. This can also be seen in the public mobility sector where many users have multiple accounts for different public transportation providers. For example, passengers who have visited Copenhagen and Berlin are having two different public transportation accounts that are not connected to each other. However, all of these services provide one functionality; providing mobility to their users. Having multiple accounts for different services is also referred to as balkanization (Bakre & Patil, 2017). Over the last years, the so-called *Identity Providers* have managed to provide a more universal login experience. Identity Providers allow users to authorize and authenticate with just one account. A primary example for an Identity Provider is Facebook Connect, which lets users sign in with their Facebook account to any service that integrates with Facebook's identity solution. However, those Identity Providers are centralized entities where only a few major providers exist. This creates an oligopoly market structure which is prone to many risks for users (Wagner et al., 2018). Facebook, for example, has a market share of about 66% of the total market, followed by Google with 20% as can be seen in the figure below (D’Onfro, 2015). This exposes users to a great threat of identity theft, data leaks or blocks by private companies. Although centralized systems are widely popular, the introduction of Blockchain Technology brings up a new opportunity to implement decentralized systems that can disrupt the existing Identity Provider landscape. This opens up the ecosystem to new types of applications and use cases that focus on privacy, trust and interoperability.

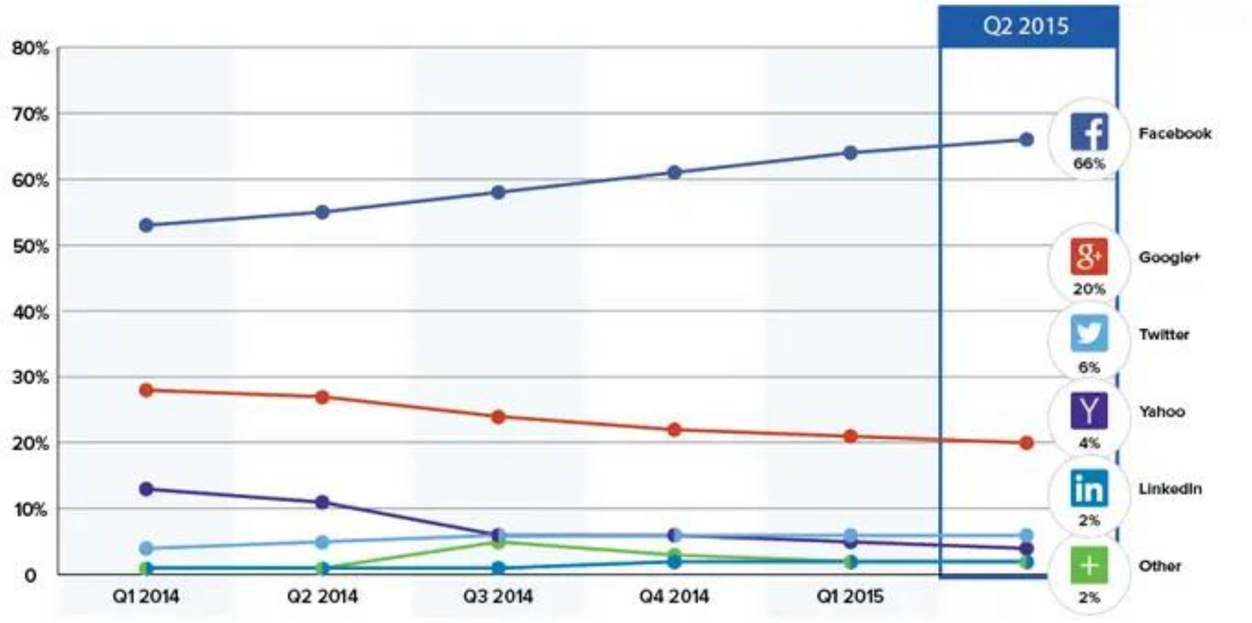


Figure 1. Identity Provider Market Share. Adopted from (D'Onfro, 2015).

In order to achieve the long-term goal of a Single European Transport Area, adequate identity solutions are needed for the mobility sector in Europe (Urban ITS Expert Group, 2013). This becomes apparent when looking at the broad landscape of available public transportation solutions. In Europe, with its 27 member states, there is a great variety of different systems in place with varying pricing schemes, policies, languages, etc. In some countries even each city has its own transportation provider which acts independently (AECOM, 2011). In order to create a Single European Transport area users need to be able to use a dedicated account to authenticate with transportation providers. This would not only improve usability but also increase the adoption of using public transportation. However, designing such a system brings up the question of the underlying system and architecture that is needed to connect all systems with each other (Calypso, 2017). Since these systems vary to a great extent high interoperability of those systems is required to create a combined solution. Thus, this paper aims at analysing a possible solution that on the one hand is taking the power away from big centralized Identity Providers and on the other hand keeps the independence of single transportation providers across Europe.

As identified in the pertinent literature, there has been academic research conducted on the theoretical concept of Decentralized Identity Management. However, a practical application utilising a *Decentralized Identity Management System* (DIdMS) in the public transportation sector is missing. The existing research gap in this field derives the motivation to conduct research in this particular field and outline a possible solution. Therefore, the research addresses the paradigm shift from centralized to decentralized Identity Management and applies it to the public transportation sector.

1.2 Current State of Public Transportation

Public transportation is an important mode of transport all over the world. However, it usually lives in isolation and differs largely from city to city. Looking at Table 1 shows that the interoperability of different public transportation systems in Europe is low. This leads to many different standards and integrated ticketing schemes. For example, examining ten major European cities and their *Public Transportation Ticketing (PTT)* systems illustrates the variety of ticketing schemes available. In most cases, cities have implemented their own ticketing schemes which are available to passengers through pre-paid tickets. This leads to different pricing, time-validity, and transportation schemes throughout Europe. Thus, complicating travel from country to country. The European Union has realized the need for a combined ticketing system across countries in order to improve usability and allow for seamless travel across the European Union. In order to facilitate the implementation of such a system, the EU has financed, for example, the ETC project. The project tries to develop a combined standard for public transportation across Europe. First test cases have been deployed in cross border travel from Aachen (Germany) to Maastricht (Netherlands). Here about 600 passengers were able to travel within the areas of those cities with one single ticket (European Travellers Club, 2019). A single ticket creates a one size fits all solution and therefore brings more usability to the overall market. Improved mobility inside cities has also been in the focus of the United Nations. It has been defined as an important step towards a more sustainable future in cities and is part of the UN Sustainability Goals (United Nations, 2015).

In Europe, the average occupancy rate for a passenger car is around 1.4 passengers (Fiorelloa et al., 2016). This low rate shows the necessity to incentivize people to switch from single travel to shared travel options. This can, for example, be achieved through improved usability in the overall transportation system (Steg, 2003). The public transportation sector has been in the focus of new directives in order to improve the overall landscape and support the achievement of the UN Sustainability Goals by 2050. Thus, the idea of a Single European Transport Area has been developed by the European Commission. This also includes the improvement of the current public transportation system in order to facilitate travel in Europe. The idea of one overall system has been introduced which includes the implementation of so-called smart wallets. These wallets should enable users to manage their travel experience in one solution without the need to switch the application when traveling between countries (Urban ITS Expert Group, 2013).

City	Ticketing
Berlin	Pre-paid ticket
Madrid	Smart Card
Rome	Pre-paid ticket
Paris	Pre-paid ticket
Amsterdam	Smart Card
Vienna	Pre-paid ticket
Warsaw	Pre-paid ticket
Budapest	Pre-paid ticket
Prague	Pre-paid ticket
Copenhagen	Smart Card

Table 1. Ticketing systems in the ten major cities in Europe.

1.3 Current State of Digital Identity Management

Most digital identities are nowadays managed on the Internet. Here users are usually required to create new accounts for every single service they are using. For example, using a travel planning app like omio.com requires the user to create an account in order to plan trips or view their travel history. Accounts like this live in isolation and cannot be used on any other platform that is not affiliated with omio.com. Here account credentials are stored in a centralized database and account data is retrieved by providing those credentials. This leads to many users having different credentials for different services which are not interoperable across applications. However, over the last decade, many services implemented so-called *Identity Provider* solutions that allow users to use their existing credentials from social login services like Facebook, Twitter or Google to log into new services (Wagner et al., 2018). This improves user experience by allowing them to use their existing social media accounts to login to different applications. Thus, users do not have to create new accounts every time they want to use another service.

A popular solution for many website owners to offer social logins on their website is oAuth which is an open standard for authorizing users with existing applications like Facebook (Wagner et al., 2018). This type of Identity Management is part of a federated identity that allows for technical interoperability and authentication across the internet. Federated identity systems have been a popular solution over the last decade. However, those systems are highly prone to attacks since they manage user information in a centralized way. Thus, exposing those systems to a high risk of data breaches. For example, in 2018 over 50 million Facebook accounts were breached allowing hackers to gain oAuth access tokens. This allowed them to access all other applications and data sets of breached users that have been using Facebook's SSO for authentication (Wagner et al., 2018). Identity providers like Facebook allow third-party applications to easily implement login mechanisms into their services while providing them with trust and validity over user credentials. Although those systems offer increased usability they are highly intransparent when it comes to sharing meta-data of users. Often companies like Facebook or Google do not share information on how they reuse information they gain through their identity providing solutions.

Also, when it comes to marketing efforts or data analytics, many services are not transparent of data usages. Many scandals about account hacks and intransparent data handling of Identity Providers have led to a rise in gaining back control of users over their own identity and data. This idea is manifested in another type of identity, the so-called Self-sovereign Identity. Here users do not rely on a centralized identity service in order to verify their claims. The establishment of centralized Identity Providers is becoming a threat to many users. Through the centralization of accounts that are controlled by third parties, users can easily lose their rights and overall sovereignty of their identities. Identity providing services can control the access and issuing of identities. Thus, a high dependency on such services are a major threat to the freedom of any user (A. Reed & Drummond, 2017). In order to give back users the control of their Identity Christopher Allen (2016) has defined a set of principles to follow when creating self-sovereign identities.

1.4 Blockchain Technology in Identity Management

Blockchain Technology has been widely publicly known through the rising popularity of the blockchain-based currency Bitcoin. For a few years blockchain technology has been the testing ground for many new applications. Bitcoin is only one of many solutions blockchain is used for. The technology is mainly characterized through its transparency, immutability and decentralized nature (Nakamoto, 2008). Thus, many technologies that use blockchain technology are built around trust-related issues. For example, the transfer of assets which requires high levels of trust in order to confirm the value exchange between multiple parties. Through its Distributed Ledger Technology

(DLT), blockchain can ensure that assets cannot be duplicated or double-spent even if the parties do not trust each other. In general, the underlying network holds this information and confirms transactions on the chain. Also, different Identity Management solutions on top of blockchain have evolved over time (Nakamoto, 2008). Here, for example, applications like Jolo.com and Civic are trying to improve overall usability of managing decentralized identities on the blockchain. Decentralized networks like the Sovrin Network are facilitating blockchains like Hyperledger Indy to allow the creation and management of decentralized identities on the internet. Identity management networks utilize blockchain to remove any intermediary as an Identity Provider. Sovrin facilitates a decentralized network to provide authentication services to identity holders (A. Reed & Drummond, 2017). Compared to a centralized federated identity system the network can’t be shut down, use data without consent, or block users from using their identities.

The idea of controlling your own data through cryptographic keys enforces the idea of controlling your own identity. Thus, when it comes to Self-sovereign Identity blockchain can be considered as an important technology to give users back the control over their identities. Identities are linked to so-called *Decentralized Identifiers* which are created and stored on blockchains. Those identifiers can be linked to certain documents and credentials which users are able to control without the need of a third-party provider (Bakre & Patil, 2017). Moreover, interoperability between systems can be ensured since users are not locked into one specific Identity Provider which is unwilling to integrate into services outside their own defined scope. This leads to an independent system that can be integrated by any service without any restrictions on content type, location or government (Fiorelloa et al., 2016).

Compared to a centralized IdMS, a distributed system relies on a shared ledger which is validated and stored by several network nodes. The stored information belongs to different users, which in a decentralized identity system can be split into Identity Owners, Service Providers, and Identity Providers. However, user sensitive information is stored off-chain and is not accessible to anyone else than the controlling entity. Therefore, storage can be off-chain and on-chain depending on the use case. On-chain storage is responsible for verification and revocation of claims and identities, off-chain storage is used for static data like private data (Tobin, 2017). The following diagram describes the main components of a DLT-based Self-sovereign Identity system.

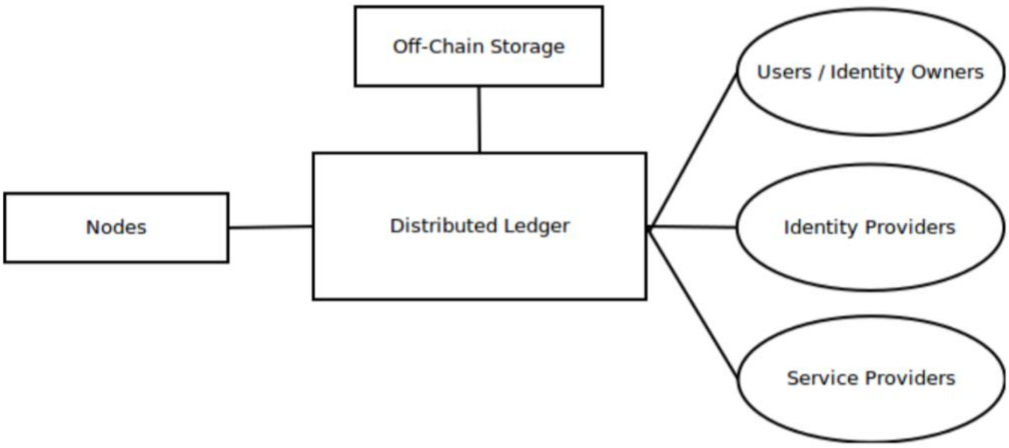


Figure 2. Components of a Blockchain-Based Self-Sovereign Identity System.

1.5 Problem Formulation

The ITS Directive (2010/40/EU) of the European Commission outlines the goal of establishing a European wide public transportation system that allows for seamless door-to-door mobility within and across member states. It provides the foundation of the deployment of an interoperable public transportation system until 2050 (Urban ITS Expert Group, 2013). The idea of combining different public transportation systems while still giving the member states the freedom to decide which system to invest in comes with many technical as well as legal challenges. Thus, to promote the development of a Single European Transport Area a possible technical solution needs to be examined. According to the Urban ITS Expert Group (2013), one of the key issues to investigate is the concept of smart ticketing solutions which provide a seamless ticketing experience for end-users. Smart ticketing solutions allow for the interoperability of fares and ticketing systems between different transportation providers. However, when it comes to the implementation of smart ticketing solutions there are a number of implementation choices to be made. One of the possible cases to consider is smart ticketing based on secure identity and back office processing (Urban ITS Expert Group, 2013). This raises the question of how user data is processed and how users are authenticated to use the system. This emphasizes the need for trustworthy and secure data processing across different transportation systems. Moreover, the various different types of implementations of public transportation networks in Europe lead to many challenges when it comes to cross-system data handling, like fare management. This in particular requires a proper identity solution that allows for cross-platform identity management.

Since a cross-platform solution relies on many different stakeholders a system that is backed by trust and transparency needs to be established. This opens up the sector to a Blockchain-based solution which enables trust between each stakeholder and aims at creating an interoperable system with transparent accounting mechanism for each entity. Thus, the resulting solution needs to provide data privacy and trust for all different entities of the Blockchain network.

Moreover, the need for high data privacy and trust between those systems needs to be taken into account to minimize fraud. Data privacy of such a system is a crucial part since there are many different stakeholders involved which increases the risk of data fraud. Moreover, many sensitive user data like ride history and personal information will be aggregated. Thus, users should be in full control of the usage of their own data. This relates to the idea of *Self-sovereign Identity (SSI)* which is gaining popularity within the identity space. Looking at current system implementations and the goal of a Single European Transport Area leads to the research gap of designing a feasible identity solution that aligns with emerging SSI standards in Identity Management and Blockchain Technology.

> ***How can a Decentralized Identity Management System that utilizes Blockchain Technology and core principles of Self-sovereign Identity for the public transportation sector be designed?***

The aforementioned research question can be divided into smaller sub-questions as presented below. Each of these sub-questions identifies key areas that build to the general problem formulation. Insights into the challenges and problems that need to be addressed will be given by answering the following questions:

1. *How can users be enabled to have full control over the management of their identities?*

   One of the core parts of any Identity Management solution is the actual structure, storage and the overall lifecycle of the data processed. Answering this question will define the requirements that have to be met in order to structure, store and provide the users' identity data. This will be achieved through an identity solution that aligns with SSI principles and complies with European data protection regulations.

2. *How can Blockchain Technology provide the infrastructure for users to use their credentials for different public transportation providers across Europe?*

   In this regard it is critical to investigate how data can be exchanged seamlessly between stakeholders. This answer will particularly shed light on how to form and validate identity transactions between the different parties. This will also give insight into the users' data attributes in regards to their identity information that needs to be provided to various Identity Provider entities.

3. *How can the ecosystem of transportation Service Providers benefit from a* DIdMS*?*

   The Identity Management solution is primarily designed to benefit the end-users without excluding the transportation Service Providers from benefiting as well. The aim of this question is to investigate how a digital identity solution can be built in order to benefit all the involved stakeholders of the ecosystem.

1.6 Scope

Since the primary topic of this paper is the use of digital identity on Blockchain, this research investigates the way it is used for the public transport sector. It is conducted within the perspective of information technology and Self-sovereign Identity principles such as the components and attributes it needs. Hence this section defines more specific objectives and adds delimitations.

1.6.1 Aim

The main aim of this study is to propose an IdMS that interacts with public transport providers and can be used by individuals to provide identities in accordance with SSI principles. The current ticketing systems in public transport will be mapped out and be dissected into their components. Additionally, Blockchain-based Identity Management Systems will be analyzed along the SSI principles. Using this information, requirements for designing a DIdMS in the public transport sector can be specified. In order to answer the aforementioned research question, user scenarios and a proof-of-concept of the system will be produced. The outcome of the study will be a contribution to the field of Blockchain and SSI and provide a foundation for designing a decentralized identity system in the public transport sector.

1.6.2 Delimitations

A few delimitations have been acknowledged during the research. The paper will not cover a full-scale implementation of the system as this would be impossible in the given time frame. This means that some of the system objectives discussed in the paper will only be reviewed but not implemented. The research will focus on developing a system that is providing access rights to passengers for public transportation service. Another delimitation of this research is that stakeholders involvement is omitted in the development process of the prototype due to the abstraction of the system.

Chapter 2

2. Literature Review

The following chapter will provide insights and a description of the different fundamental concepts that serve as the theoretical foundation that will result in developing an identity solution for the public transport sector. This chapter is divided into two parts; the first part introduces Identity Management and the work in the field of *Self-sovereign Identity (SSI)*. The second part explores Blockchain Technology along with relevant Identity Management standards that have been produced to facilitate the SSI principles.

2.1 Digital Identity Management

According to Windley (2005), digital Identity Management is the concept of managing records of different identities. This can, for example, include creating, managing, using and destroying records linked to a specific identity. A record can be, for example, the real name of the digital identity, that it is representing. However, a digital identity is not only related to a person interacting with a digital system. Digital Identities are the representations of external agents which can also be devices, organizations or applications. Thus, digital Identity Management is the overall layer that handles permissions and authorizations to execute certain tasks within a system. Therefore, whoever controls a digital identity has access to certain actions within a closed system that are defined by rules and permissions encoded into a digital identity. Thus, security and access management are crucial tasks of any Identity Management System (Windley, 2005). Identity holders usually get access to their digital identities through credentials which they can use to authorize. Moreover, digital identity systems are starting to become more complex since they need to provide access to an increasingly heterogeneous technology environment. Thus, digital identity systems are moving from centralized systems to more federated or even decentralized solutions. According to Christopher Allen (2016), decentralized identity systems offer the benefit of increased portability and user control across different applications.

2.1.1 Types of Identities

Christopher Allen (2016) separates online identities into four different types namely; centralised, federated, user-centric and self-sovereign. Those types can be categorized along the axis of user control and portability. Whereas, user control refers to how much control a user has over her own identity. For example, low control would mean access to an identity can be withdrawn by a centralized authority like a database. Portability describes how easy an identity can be reused across different systems or applications (Allen, 2016).

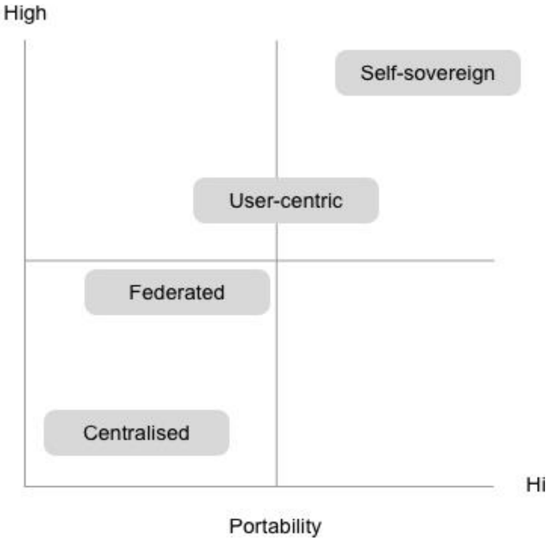


Figure 3. Types of Identity.

2.1.1.1 Centralised Identities

According to Allen (2016), centralised identities are issued by a centralised authority. Here the underlying authority controls the access to the identity. This can be, for example, an online service like Amazon. The service can easily deny access to the identity by revoking the users credentials. Moreover, if it is centralised there is only one single source of truth. This can result in fake identities which have been only confirmed by the centralized authority. This in general gives more power to the issuing authority than to the users that actually own the identity. Centralised identity systems also lead to high balkanisation of identities. Many websites and online services force users into creating separate identities; leading to data silos, less control for users and more power to the website. Those services could easily disappear or block users from using their own data. However, this is not in the best interest of users since most of current online identities are issued through centralised systems (Laurent et al., 2015).

2.1.1.2 Federated Identities

Federated identities are those that can be used within a collaboration of systems. Here users are able to login with the same credentials into different services that form a federation. For example, Google offers its users to log into multiple applications that are affiliated with each other. Users can, for example, use the same credentials for their Google mail account as for their YouTube account. Thus, after a user logged into one application, she can also use other applications within the collaboration. This is possible because they are using a federated identity which is shared across services. Most literature defines this type of login mechanism as *Single-Sign-On* (SSO) which is a subset of federated identity management. However, federated Identity Management is usually referred to as a collaboration of trust where Identity Providers never share user credentials with external Service Providers (Laurent et al., 2015). Thus, it does not rely on a collaboration of single systems but more on a specific Identity Provider. This leads to a more user-centric approach to managing identities.

2.1.1.3 User-centric Identities

In a user-centric approach, users are able to control their identities outside a specified collaboration of systems. Here the federated systems refer more to a collaboration of trust between an Identity Provider and any type of external application. Thus, creating a trust relationship between each other. This trust relationship opens up the possibility of using Service Providers by only exchanging credentials through the Identity Provider (Wagner et al., 2018). In such a system a user only grants permission to the Service Provider to use the identity provided by the Identity Provider to manage the service. For example, the Facebook feature "Login with Facebook" allows users to use their Facebook account to use any application that integrates it without exposing their identity credentials to them on authentication. However, user-centric identities improve the portability of an identity they don't give full control over an identity. Thus, if users for example get banned by the Identity Provider, they lose access to any application they have been using. According to Allen (2016), truly user-centric identities are those that allow the user to fully control their identity in an infinite amount of systems. This is also referred to as Self-sovereign Identity.

2.1.1.4 Self-Sovereign Identity

Self-sovereign identity aims at putting the user into the center of control of their own identity. This means that a user can fully decide over her identity. Thus, a Self-sovereign Identity creates full autonomy of uses over any type of identity system. In previous systems, users were relying on a centralized Identity Provider to be authorized and give access to identity information to third parties. However, in Self-sovereign Identity systems identities are decoupled from any centralized source that could block, alter, or delete their identity. Identity claims are stored within the identity itself that is controlled by the user (Wagner et al., 2018). Christopher Allen (2016) outlines ten guiding principles for a Self-sovereign Identity. According to him, a Self-sovereign Identity consists of the following principles:

Principle	**Meaning**
Existence	An identity must be linked to a real person outside the digital world. Thus, having an independent existence.
Control	The user has full control and authority over the usage of her identity and its claims.
Access	A user needs to be able to access all claims and data related to her identity.
Transparency	The system that operates and manages identities needs to be fully transparent.
Persistence	An identity needs to be long-lived. If data changes the identity keeps the same.
Portability	Identities cannot be linked to one specific party. They need to be portable to any other type of system.
Interoperability	Identities should be possible to be used in any type of system.
Consent	Sharing of data can only happen when the user consents.
Minimization	Data should be exposed only to verify claims.
Protection	The freedom and rights of individual users are most important to the network.

Table 2. Christopher Allen (2016) Principles of SSI.

However, these are only guiding principles and there is no clear consensus yet of how a Self-sovereign Identity is truly defined. Thus, a Self-sovereign Identity can be an identity that fulfils only a few of those principles. Nonetheless, it can be assumed that identity becomes self-sovereign if the user grants full control over it and it does not rely on a centralized system. Thus, moving towards a decentralized Identity Management solution where no central institution holds control over it would pave the way to a Self-sovereign Identity system.

2.1.2 Identity Management Ecosystem

Identity management is a crucial concept when it comes to managing access rights and authentication of services. There are three important roles when it comes to modern online Identity Management Systems. The following will explain the concept of Identity Owner, Identity Providers and services providers which is used throughout this research.

2.1.2.1 Identity Owners

Identity Owners are those who receive credentials by different services. The wallet may also contain further personal information about the Identity Owner, the so-called self-attested claims. The Identity Owner could present entire credentials, parts of them or even combinations of multiple credentials in the form of proofs to Service Providers. The credentials can be entirely or selectively disclosed, therefore the Identity Owners have full control over their data on how data is used and what is shared.

2.1.2.2 Identity Providers

An Identity Provider (IdP) is a trusted system that manages identities on behalf of an entity. The IdP is providing authentication and authorization for external Service Providers when requested. Thus, an IdP is storing credentials of an identity issuer and generating claims for a relying party. Thus, Identity Providers act as third parties that are responsible for a seamless exchange of credentials in order to authenticate users with services that are integrated within the ecosystem of IdPs. Through the integration with IdP users are able to use the same credentials across systems. This reduces the balkanization of user accounts across the Internet (Wagner et al., 2018).

2.1.2.3 Service Providers

Service providers are those who consume and verify identities in order to provide a specific service. The Service Provider needs an identifier in order to recognize existing users and associate them with their data. Many Service Providers are at the same time also issuers. This is because many services use their own IdMS and databases to authenticate and onboard new users (Windley, 2005). Thus, the identity-consuming entity is in most cases also the service providing entity. Here identity credentials are exchanged for services in order to provide a customized user experience.

2.1.3 Know Your Customer

Know your Customer (KYC) processes describe due diligence processes often used in the banking sector that facilitates the onboarding of new customers. KYC is initiated when a customer intends to work with a financial institution. It includes the exchange of documents between parties as well as the collection of basic identity information of the beneficiary. Other processes that are concerned with the customer onboarding consist of risk management, monitoring of transactions and other specific policies that would give useful insights on the customer. This process is repeated every time, for example, the customer wants to open a bank account. The current KYC processes are outdated and require significant costs. Due to the growth of technology and regulations, the domain of KYC is undergoing a lot of changes by utilizing Distributed Ledger Technology. This provides more cost-efficient and faster identity verification processes when onboarding new customers (Parra-Moyano & Ross, 2017).

2.2 Blockchain Technology

A detailed explanation of how a Distributed Ledger Technology (DLT) technically operates and works is out of the scope of this study. However, it is important to understand why DLT and Blockchain are related to SSI and what their use entails for Identity Management software that interacts with a system utilizing them.

Blockchain was introduced in 2008 with the purpose of creating an electronic cash system that did not rely on a trusted third party (Nakamoto, 2008). It was implemented by creating a network where all transactions are visible and broadcasted to everyone and when validated, are added into a block. The block containing these approved transactions is created and chained to the previous block of approved transactions, thus creating a chain of blocks, therefore the name Blockchain.

Distributed Ledger Technology is the underlying technology of Blockchain (UK Government Chief Scientific Adviser, 2016). Due to the emergence of the DLT, it is now possible to achieve a fully decentralized identity. A distributed ledger is a database that is geographically dispersed and spanned across multiple places. The data stored in a distributed ledger is consensus agreed by the majority of participants of the network and all of them can own their copy of the ledger. If the data is altered, the changes are applied to all the copies of the ledger to reflect the new state. However, in order to prevent situations where anyone can alter the data as they wish, distributed ledgers are divided into two main categories: permissioned and permissionless. Blockchain is a permissionless distributed ledger which means that anyone can change the state of the ledger by adding new blocks of transactions as long as these have been validated by consensus algorithms like Proof of Work or Proof of Stake, etc. A permissioned ledger, as the one used for this project, requires some sort of a governance model before new transactions are added to the ledger. The accuracy and security are established cryptographically with the use of keys and digital signatures, without the need to rely on a central authority to be in control of the changes made to the ledger.

2.2.1 Types of Blockchains

As presented so far in this study, distributed ledgers can be utilized to facilitate a Self-sovereign Identity resulting in Blockchain to become almost synonymous with SSI. However, it is important to distinguish them as separate concepts that do not necessarily intertwine. Since every Blockchain is a distributed ledger, Blockchain systems can be categorized into three types. As can be seen in the table below, the columns represent data ownership, while the rows represent the read, write, or commit permissions granted to the participants:

	Permissionless	**Permissioned**
Public	Public-Permissionless (Public)	Public-Permissioned (Consortium)
Private		Private-Permissioned (Private)

Table 3. Types of Blockchains.

In a public Blockchain, any transacted party is allowed to read the ledger and can take part in consensus. The ability to write the ledger differs based on the granted read and write permissions.

2.2.1.1 Public

In a *public permissionless* Blockchain, everyone in the network has equal authority. Some notable Blockchain networks are Bitcoin and Ethereum. The data privacy in a public Blockchain can not be guaranteed due to its public nature, thus it is argued that personal data should not be stored in this type of Blockchain. Some use cases for a public permissioned Blockchain include cryptocurrency, supply chain management, SSI and a lot more.

2.2.1.2 Consortium

In a *public permissioned (consortium)* Blockchain, everyone participating in the network has the ability to view the transactions. Writing access is given to only a few nodes participating in the network and allowed to take part in the consensus. Therefore, an elected consortium is responsible for the governance of this type of Blockchain. They decide who is granted the ability to write into the Blockchain. In order for a public permissioned Blockchain to work, a set of case-specific agreements needs to be established between the participating parties. Also, SSI can potentially be one use case that utilizes this type of Blockchain.

2.2.1.3 Private

A private Blockchain is a *permissioned* Blockchain. In a private Blockchain, access is granted only to the entity or entities that participate in a transaction and have knowledge about it, whereas the other will not be able to access it.

2.2.2 Data Privacy

Data privacy has always been a concern for users. Recent scandals of data breaches have increased awareness on this issue resulting in new legislations on data privacy. Some of the actions that have been taken toward this approach is the human-centric model proposed by the Ministry of Transport and Communications of Finland. Additionally, the European Union taking a similar direction to address the increasing concerns of data privacy, they imposed a legal framework, the so-called *GDPR (Tankard, 2016)*.

2.2.2.1 MyData Model

A concept explored in literature in this paper is called MyData and refers to a paradigm shift from the current organization-centric focus to a human-centric focus in personal data management (Poikola et al., 2015). The main idea behind the approach of MyData is that individuals should have a better overview of how their data is used, where it is stored, who can have access to it. Moreover, users should also be involved in the process of deciding who can use their data and for what reason. Therefore, it is a concept that aims to give full control over data back to the users. This is achieved by placing the users in the center of the personal data ecosystem. MyData proposes to change the way personal data is managed at the infrastructure level which will result in interoperability and data portability. Additionally, data management in MyData is consent-based which means that the user can control the flow of the data without having to store all his or her data on centralized repositories. Finally, MyData is a progressive approach to personal data management that facilitates data sharing across sectors with the aim of benefiting users, organizations and society at large. MyData focuses on giving certain rights to the users over their data and more concretely MyData outlines: a) the right to know what personal information exists, b) the right to see the actual content of personal information, c) the right to rectify false personal information, d) the right to audit who accesses and processes personal information and why, e) the right to obtain personal information and use it freely,

f) the right to share or sell personal information to third parties, g) the right to remove or delete personal information (Poikola et al., 2015).

2.2.2.2 GDPR Requirements and Blockchain

When developing an Identity Management system, it is important to consider regulations regarding the processing of personal attributes and data concerning users like the *General Data Protection Regulation* (GDPR) that came into effect in May 2018. This means that Service Providers will need to meet certain restrictions and requirements imposed by GDPR, when it comes to handling personal data. In regards to the research question and more concretely to the first sub question as presented in section 1.5, this section will analyze how GDPR requirements can map into what Blockchain-based Identity Management solution may offer.

First GDPR requirement is to ensure *availability*. This holds true with Blockchain through the distribution of nodes that store the same copy of the ledger and even if a certain number of these nodes becomes unavailable, the others will still hold the same copy of the data which means that everyone can always have access to the information (Piekarska et al., 2018).

The second rule of GDPR is *completeness,* meaning that every event and data have to be recorded. This happens by design in Blockchain and DLT, as each new block is connected to the chain of blocks through cryptographic hashes of previous blocks ensuring that any change will be immediately reflected to the Blockchain and it is up to the participants of the network to agree on the content of it in order for a new block to be published and stored forever on the Blockchain (Piekarska et al., 2018).

Another GDPR requirement is *correctness* there is assurance of data accuracy. Blockchain and DLT have been designed in a way that data needs to be verified before it is amended to the Blockchain and participants of the network must achieve a consensus over the correctness of the audited data before changes apply (Piekarska et al., 2018).

Furthermore, *integrity* of data is protected in a Blockchain. The design of the Blockchain is aligned with GDPR statement that data stored should be protected from unintentional or malicious changes. If new changes are introduced this means that new inputs to the blocks will be introduced without overriding any data (Piekarska et al., 2018).

Another rule of GDPR is *immutability* which holds true for Blockchain by combining cryptography and distribution and makes it very difficult to alter or delete any information stored (Piekarska et al., 2018).

Confidentiality is another rule from GDPR which means that only the involved parties in an exchange of data can be able to view the details of the transaction. Here any event of processing data has to be traceable and linkable to previous data processing events. Blockchain can offer a permissioned environment and confidential transactions where only a group of participants knows the content of the exchanged information, while the other parties of the network acknowledge such a transaction taking place (Piekarska et al., 2018).

Transparency is implemented in a Blockchain by tracking the transactions using their hashes to their exact position in the chain (Piekarska et al., 2018).

2.2.2.3 Blockchain and Data Erasure

In Art 17 of GDPR it is stated that any data subject needs to have the right to have personal data erased. This is also referred to as the "*Right to be forgotten"*. This right is an important consideration when it comes to Blockchains which by nature are immutable. Thus, GDPR and Blockchain Technology might become mutually exclusive in their fundamental design choices. However, Art 4 does not clearly define personal data making it up subject to interpretation. Therefore, it can be argued whether or not encrypted data can be considered as personal data. This gives Blockchain Technology, even with its immutability, the possibility to act within the requirements of GDPR. However this depends on its specific implementation and further definitions of what exactly constitutes as personal data (Berberich & Steiner, 2016).

2.2.3 Data Storage in Blockchains

Data is critical in any software application and over the past years databases have evolved a lot to be more efficient and scalable based on the volume and complexity of the stored data. In the case of Blockchains, there are two ways that data can be stored and this includes data stored in the Blockchain itself or data stored off-chain in a third-party database.

2.2.3.1 On-chain Storage

The simplest form of storing information in a Blockchain is to simply store the data in the chain itself. Binary data can be stored as part of the transaction and will then be distributed to the network along with all the other transactions. This leads to the question of what data should be stored on-chain that is GDPR compliant since, for example, data related to natural persons imposes an infringement to the GDPR. Due to the fact that the encryption algorithm can still be compromised, the encrypted data stored on-chain is not recommended even after the key that decrypts the data is shown to have been destroyed (Wagner et al., 2018). In order to increase security, Blockchains can use hash functions to encrypt data. Therefore, when storing personal data on-chain, the hashes generated from the data are stored instead.

2.2.3.2 Off-chain Storage

Although storing encrypted data directly on Blockchains works well, it can suffer from scalability. This strengthens the fact that on-chain storage is not technically and financially practical. However, the advantages that Blockchain provides can be applied to off-chain storage methods. Thus, another solution is to store data in a separate storage location and provide a link between this location and the hash in the chain. The corresponding cost of storage also becomes lower due to the relatively small size of a hash. A good technique to achieve integrity is to add a timestamp to the hash value of the data when applying changes to the data stored off-chain, diminishing the effect of the reference stored on-chain (Wagner et al., 2018).

2.2.4 Trust Infrastructure

Trust plays a vital role in the Self-Sovereign Identity ecosystem since it has a decentralized infrastructure that is not established by a central institution. It is essential, therefore, to achieve accountability and have reliable and tamper-proof information. This ensures trust in the SSI ecosystem which at the same time introduces one of the biggest challenges to solve. While in a centralized system a user has to trust the central institution, in a decentralized system the community needs to agree on certain technologies, formats, and standards. A number of challenges arise when establishing trust in a decentralized manner. In decentralized systems, the verification processes become slow and expensive, especially when data has to be verified by multiple parties. Also, the users have no registry to look up if another party can be trusted. In addition to that, the users have to trust that their information in the system is protected against any form of attack. The solutions to these problems is to establish trust with technology, achieved by standardized, open-source and transparent processes. In an SSI case this takes away the management of identities from central institutions and gives it back to the users.

2.2.5 Foundations and Entities

In this section, the relevant foundations, entities, workshops and working groups working in the fields of identity and data privacy are introduced. The activities of these entities are relevant in order to develop an SSI based Identity Management system for the public transport sector.

2.2.5.1 W3C

Word Wide Web Consortium (W3C) is an organization that consists of different actors (*W3C*, 1994). Their task is to organize and develop web standards that can be applied by the web community. Some of their latest work involves the development of standards around decentralized Identity Management. Technical concepts like a new type of identifier to provide verifiable, decentralized digital identity are introduced.

2.2.5.2 Decentralized Identity Foundation (DIF)

The Decentralized Identity Foundation (DIF) is another foundation working in the respective area of decentralized identity. It consists of several Working Groups that their objective is to shape the future of decentralized identity technology and standards. One of the most notable projects is called “Identifiers and Discovery” which aims to develop protocols and implementations that would allow to create, resolve and discover Decentralized Identifiers and names across a decentralized system. Another important working group is “Storage and Compute” that focuses on the storage and management of a user’s identity data whose maximum control is retained by the Identity Owner. They are also working on a project called “Universal Resolver” that facilitates the creation and registration of self-sovereign identifiers by their own owners. Another project that the organization is working on is “Identity Hubs” which are off-chain data storages for private data of the identity subject. Microsoft and Sovrin are members of the DIF and have announced their work on Hub implementations as well (*Decentralized Identity Foundation*, 2018).

2.2.5.3 Rebooting Web of Trust

Rebooting Web of Trust is a workshop where people with various backgrounds from the areas of cryptography, computer science, anthropology, and others, are gathered with a common objective to exchange knowledge and help create the next generation of decentralized web-of-trust based identity systems (*Rebooting Web-of-Trust*, 2018). The relevant projects they are working on are the following:

- Decentralized Identifiers (DIDs)
- Decentralized Public Key Infrastructure (DPKI)
- JSON-LD

These projects enable trust in decentralized networks and were developed along the principles of SSI. The next section will present in a more thorough way some of the aforementioned standards and technologies.

2.2.6 Technologies and Standards

This section will introduce the most important technologies, standards, and specifications produced by foundations and entities working in the area of identity and data privacy. For the Self-Sovereign Identity ecosystem to achieve interoperability, the relying parties have to agree on standardized formats and processes. Standardization is a crucial pillar in the goal of creating a decentralized ecosystem for SSI.

2.2.6.1 Decentralized Identifier (DID)

Usernames and passwords are the most common identifiers currently on the internet. Those can be described as conventional identifiers and do not belong to the users since they are valid only within the Service Provider’s ecosystem and are borrowed from them. This ultimately results in lack of control for users over their identifiers. Moreover, having their personal identity information stored on Service Providers’ storage can pose a security risk since attackers may easily compromise access to valuable data.

As a result, a decentralized system builds upon Decentralized Identifiers (DIDs). The DID specification was created and published by the W3C Credentials Community Group under the W3C Community Final Specification Agreement (FSA) (D. Reed et al., 2017). A DID is a new type of globally unique identifier for digital identity. DIDs are recorded on a distributed ledger and their publishing, reading, updating and revoking is done via a Decentralized Public Key Infrastructure (DPKI). Thus, a DID is a self-generated identification number that derives from the user’s public key. Hence, the DID does not require a centralized registration authority for its management. A DID represents a portable digital identity generated and addressed by a public key on a ledger, while the DID holder is in possession of the corresponding private key, hence enabling full control of the DID for the DID holder (D. Reed et al., 2017). This makes it a truly self-sovereign digital identifier (D. Reed et al., 2017). The generation of a DID is based on cryptography and randomness which makes it almost impossible to create the same DID pair twice by different users. This makes any DID a unique id in the overall ecosystem. Apart from people represented by a DID, also other entities can have their own DIDs such as an organization or even a machine. DIDs are a key concept in SSI solutions and utilized by emerging frameworks like Sovrin, uPort, and Jolocom.

DIDs' syntax consists of three parts. The first part contains a scheme referring to Uniform Resource Name (URN) followed by a namespace as a Universally Unique Identifier (UUID). The last part contains a namespace-specific identifier that uniquely identifies the entity inside the namespace. The URN and UUID must be lowercase. The syntax of a DID can be seen in the below figure:

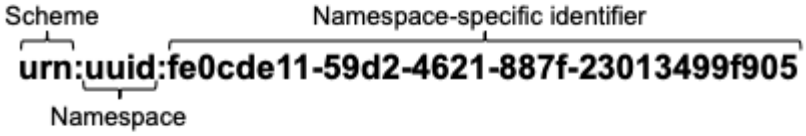

Scheme
Namespace-specific identifier
urn:uuid:fe0cde11-59d2-4621-887f-23013499f905
Namespace

Figure 4. DID Syntax.

A real and valid DID generated by using Sovrin's method can be demonstrated in the below figure:

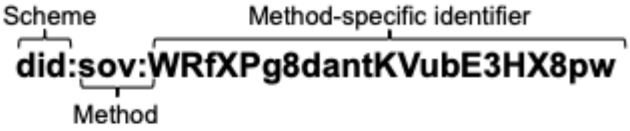


Figure 5. Sovrin DID Example.

As it is apparent, a DID differs from a traditional human-readable identifier like the social security number. Sovrin is using "pairwise pseudonymous identifiers" that reduce the correlation of the identity and the identifier meaning that the identity data is separated from the identifier and that for every new relationship a new identifier can be created (D. Reed et al., 2016).

It should be noted that DIDs are universal identifiers and are not tied to a specific ledger and can be used in all the ledgers that support them. DIDs can be generated in different ways depending on the Blockchain or DLT that provides the functionality to create a public and private key pair. This process is defined in DID methods and there are different implementations of the CRUD operations regarding the DID. The following attributes must be able to be defined by a method specification (D. Reed et al., 2017):

- The DID method name.
- The structure of the method-specific identifier.
- How the method-specific identifier is generated or derived.
- How the CRUD operations are performed on a DID and DID document.
    - **C**reate or Register a DID on the ledger

- ○ **R**eading or Resolving a DID document
- ○ **U**pdating a DID document
- ○ **D**eleting or Revoking a DID

There are currently some DID methods officially registered in the DID Method Registry of W3C that can satisfy the aforementioned requirements of a DID method specification. Some of the most popular are listed in the table below:

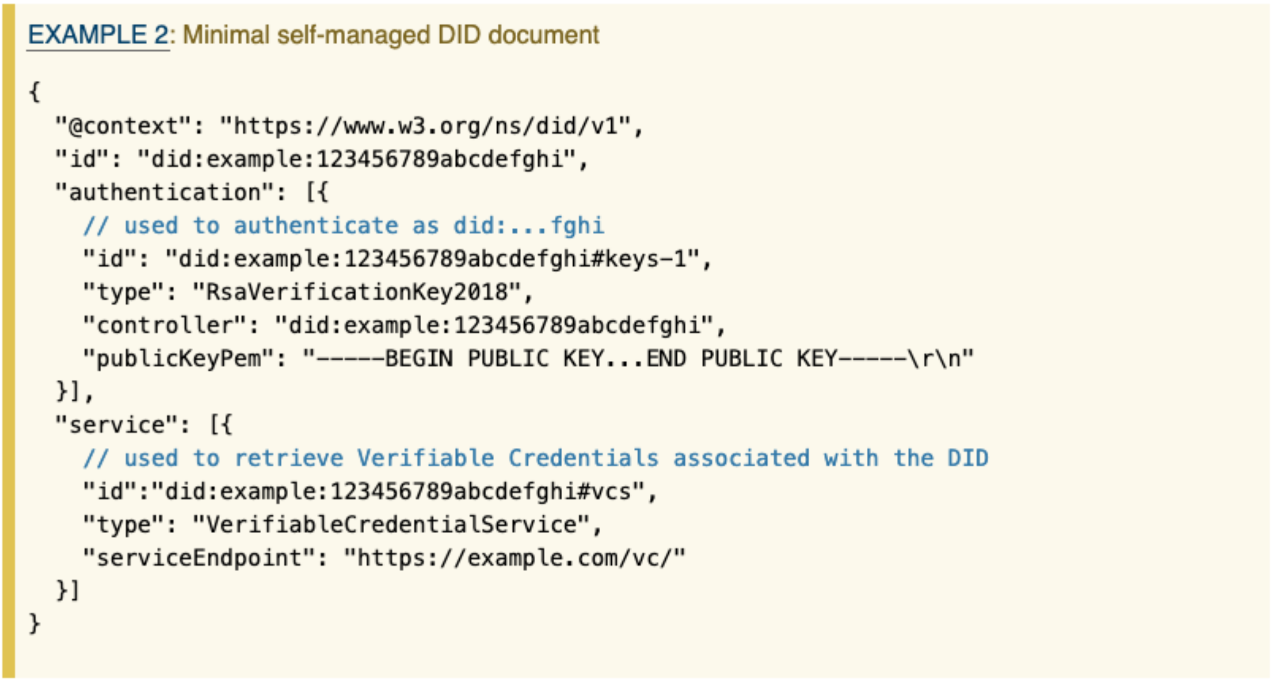
Method	DID prefix
Sovrin	did:sov:
Bitcoin Reference	did:btcr:
Ethereum uPort	did:uport:
Veres One	did:v1:
Jolocom	did:jolo:
IPFS	did:ipid:

Table 4. The Most Popular DID Methods.

2.2.6.2 DID Document (DDO)

While a DID is a key-value entry of an identifier inside the system, a DID document (DDO) contains meta-data about the DID (D. Reed et al., 2017). In other words, DID documents are the counterpart of a DID meaning that DIDs are simply the indexes, while the DID documents provide the information. This information can be associated with public keys and service endpoints allowing the DID to interact within the SSI ecosystem in a verifiable and safe way.

DID Documents are created as a JSON-LD (linked data) file (Sporny, Longley, Kellogg, et al., 2018). This schema allows data to be exchanged in a standardized structure between two different systems that both understand each other. A DDO is stored on a distributed ledger along with the DIDs. The document contains service endpoints to reference off-chain data like *Verifiable Credentials (VC)* that are connected to a specific DID. In the figure below, an example of a basic schema of a DID document is presented:



EXAMPLE 2: Minimal self-managed DID document

```
{
  "@context": "https://www.w3.org/ns/did/v1",
  "id": "did:example:123456789abcdefghi",
  "authentication": [{
    // used to authenticate as did:...fghi
    "id": "did:example:123456789abcdefghi#keys-1",
    "type": "RsaVerificationKey2018",
    "controller": "did:example:123456789abcdefghi",
    "publicKeyPem": "-----BEGIN PUBLIC KEY...END PUBLIC KEY-----\r\n"
  }],
  "service": [{
    // used to retrieve Verifiable Credentials associated with the DID
    "id":"did:example:123456789abcdefghi#vcs",
    "type": "VerifiableCredentialService",
    "serviceEndpoint": "https://example.com/vc/"
  }]
}
```

Figure 6. Example of Basic Schema of a DID Document. Adopted from (D. Reed et al., 2017).

Usually, as it is shown above, these key-value pairs would be written on a Blockchain or a distributed ledger and be used by other parties. The DDO contains a mandatory “@context” attribute which defines the version of the DID document format and must be the same for every DID in the environment. The second attribute is the “id” which describes the subject of the DID to which this document is associated with. The next key-value pair contains at least one public key which is important for authentication and CRUD operations. Attributes of this include id, type, owner and encoded public key properties. If there is no public key in a DDO document, this means that the DID has been revoked and it is not valid. Last but not least, the “service” attribute contains information about provided service endpoints where a requesting party can interact with the DID owner. A service endpoint could be an API for communication, a link or even a QR code.

2.2.6.3 Universal Resolver

The Decentralized Identity Foundation (DIF) has been working on defining the standard for the Universal Resolver which aims to develop a software unit that takes a valid DID as input and resolves it to a DID document (Sabadello, 2017). The universal resolver is established to provide an interoperable system that is able to resolve any kind of DID over a generic API that targets the underlying ledgers of a DID method. Hence, it makes it easier to resolve DID methods within the SSI ecosystem. The model of the DIF’s universal resolver is demonstrated in the figure below:

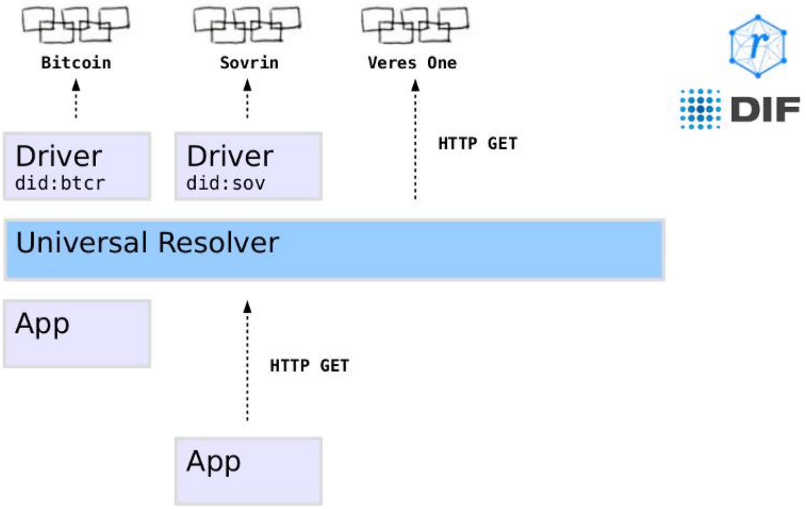


Figure 7. An illustration of DIF's universal resolver model. Adopted from (Sabadello, 2017).

As it can be seen above, the universal resolver can consist of different drivers for each supported identifier type and use them to achieve operations requesting information from a DID document. This resolver can also be used as a service by other clients. They could connect to the universal resolver by using an HTTP API or a native linked library enabling the clients to use decentralized networks (Sabadello, 2017). For example, ledgers like Veres One support the DID resolution to a DID document natively, while other ledgers like Bitcoin need a specific driver implementation that creates a DID document based on a specified DID.

It should be noted that the universal resolver enables data formats like Verifiable Credentials and protocols like DIF's Hub protocol to be built on top of the identifier layer without the need to know which Blockchain or ledger has been used to register the identifier (Sabadello, 2017).

2.2.6.4 DIF Identity Hub

Today a vast amount of user data is stored on the cloud. There are cloud services like Google Drive or Dropbox that let users store data with the confidence that it will be secure, highly available and accessible anytime, anywhere. However, there is a need for a similar service that gives users full control over their data. The Identity Hub developed by the Decentralized Identity Foundation (DIF) is trying to solve this problem by allowing an easy management of digital identity credentials through an interoperable protocol (Buchner, 2016).

The main concept of hubs is that they act as data managers of a decentralized identity (Buchner, 2019). Expanding on that, hubs can be seen as the central tools that allow users to store their data on them in a secure way and give them control over it. DIF defines each hub as a data store that is owned and signed by an identity and accessible via a globally recognized API (Buchner, 2016). The Identity Hub is still work in progress and some requirements are subject to change. The most notable requirements stated by the DIF are presented below:

- **One DID to many Hub Instances** - According to this requirement, a single Identity Owner may have one or multiple instances of an Identity Hub and all of their instances must be addressable via a URI routing mechanism that is linked to the entity's identifier. In order to ensure that the owner can have access to them anywhere, the instances have to sync state changes (Buchner, 2016).
- **Syncing data between Hubs** - Hub instances owned by the same Identity Owner must be able to sync data in a seamless way. Although there is still no selection of a protocol that can reproduce Hub states across different hubs, DIF claims that it should be fairly easy to have this capability implemented on top of any NoSQL datastore (Buchner, 2016).
- **Hub data serialization and export** - Here data contained in the hubs must be exported in a serialized state. Thus, the users retain full control over their data portability (Buchner, 2016).

DIF also proposes the following Hub URI scheme "hub://did:foo:123abc/" that allows to achieve different links to an identity's owner data without being dependent on a specific hub instance. Services that understand this scheme will be able to leverage the Universal Resolver to look up the different instances of an entity's hubs via its DID and then access the hubs via the specific Service Endpoints contained in the identifier itself, as presented in the section of DDO (Buchner, 2016).

Regarding the authentication, DIF's specifications suggest that authenticating requests to Identity Hubs will follow the DIF/W3C DID Authentication scheme. Furthermore, DIF defines some specifications regarding the Identity Hub's API in order to achieve a high level of security and privacy. The response and request format of the request URLs will not follow a traditional REST-based API schema due to the sensitive nature of the data being transmitted to Identity Hubs (Buchner, 2016). In conclusion, Identity Hub is a concept of a decentralized personal data store that puts control over personal data in the hands of users and enables them to navigate their digital identity in a way closer to a true Self-sovereign Identity.

2.2.6.5 Public Key Infrastructure and Decentralized PKI

Rebooting Web of Trust proposed in a white paper the concept of Decentralized Public Key Infrastructure or DPKI. As it is explicitly stated, "identities belong to the entities they represent". This decentralized approach aims to give the identity control back to its principal owner and not to a trusted third-party as it occurs in the traditional PKI (Allen et al., 2015).

Before diving into DPKI, the concept of PKI needs to be described first. Today, the dominating way for managing authentication and distribution of public keys is PKI. In its essence, it is a centralized database that stores pairs of identities and public keys. In order to prove ownership of the identity, typically a password (the private key) that corresponds to a public key is required. Although an identity can prove its ownership with the private key, there is a lacking mechanism to trust and proof of identity. This means that the organization which is in control of the database is also the actor that needs to be trusted in the first place. PKI involves the use of Certificate Authorities (CAs) and trusted third parties to solve this type of problem. One security challenge that PKI is prone to, is the creation of a single point of failure. For instance, if a centralized repository containing public keys is compromised, this may result in jeopardizing the integrity of the whole database. Another security problem associated with the traditional approach of PKI is that if a CA is compromised this would allow a MITM (Man-in-the-Middle) attack which is extremely difficult to detect (Allen et al., 2015).

On the other hand, DPKI is an approach that has been proposed as an effort to remove or at least reduce the trust that has been placed on centralized and trusted third-party systems. It aims to improve the way keys are managed on the Internet. According to Allen et. al (2015), the answer to solve the aforementioned problems is not to abandon PKI but rather decentralize it (Allen et al., 2015). The decentralization of elements of trust across all the participating entities can result in no single third-party being able to compromise the integrity and security of the system. DPKI is based on decentralized key-value datastores like Blockchain and other DLTs. The place of a third-party is taken by the validators (or miners) which have the responsibility of ensuring the integrity and security of the Blockchain or the decentralized ledger through a consensus protocol (Allen et al., 2015). Hence, with this approach Identity Owners can have complete control over their identity data instead of central actors.

2.2.6.6 Claims and Verifiable Credentials

The SSI ecosystem was created not only to have full control and autonomy about an identifier but also to enable actors to create attestations about themselves or others. The Verifiable Credentials Working Group (VCWG) of the W3C has created a data model that covers claims and Verifiable Credentials in an effort to standardize a format on how claims and credentials should be used or issued (Sporny, Longley, & Chadwick, 2018).

The following figure illustrates the ecosystem in which Verifiable Credentials are expected to be useful for its core actors as well as their roles and relationships between them. This section will give an overview of claims and credentials, their structure and the difference between them.

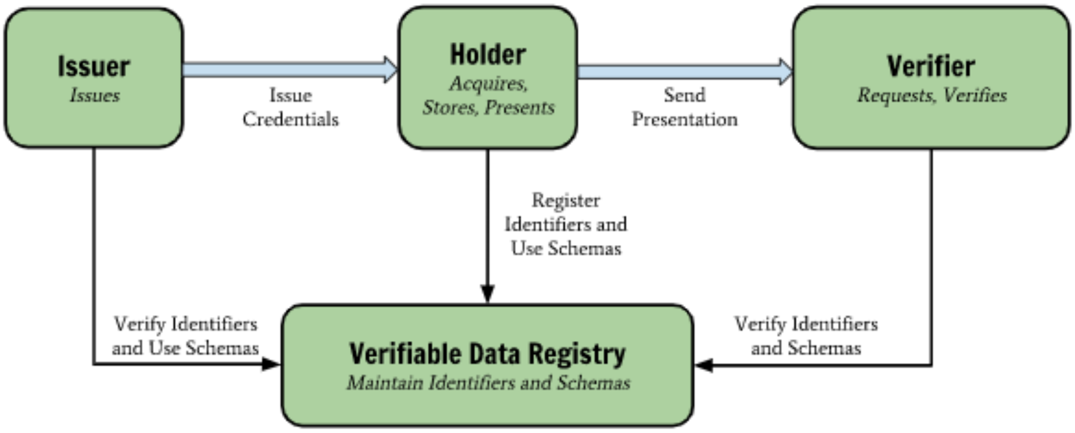

Issuer
Issues
Issue Credentials
Holder
Acquires, Stores, Presents
Send Presentation
Verifier
Requests, Verifies
Register Identifiers and Use Schemas
Verify Identifiers and Use Schemas
Verifiable Data Registry
Maintain Identifiers and Schemas
Verify Identifiers and Schemas

Figure 8. An Example Ecosystem That Illustrates the Roles and Information Flows Forming the Specification Basis. Adopted from (Sporny, Longley, & Chadwick, 2018).

Claims

According to Sporny et al. (2019) from the VCWG, a claim can be used to make an assertion about a subject and can be described as a subject-property-value relationship. For instance, a claim can be about someone being a student or graduate of a certain university as seen below. It can be viewed as a key-value pair that states something about a subject (holder).

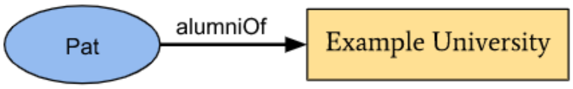


Figure 9. An Example of a Basic Claim Illustrated as a Directed Graph. Adopted from (Sporny, Longley, & Chadwick, 2018).

It should be noted that claims can be made by anyone about anybody and anything. This aligns with the informal nature of claims in contrast to credentials. Expanding on the aforementioned example, a student could claim that she has the highest-grade point average in her course without explicitly knowing about it, thus making the claim untrusted. In order to assert a higher level of trust in this claim, it shall be verified by trusted third parties. In this case, either the university or a professor of the course can verify the claim as they know whether what student claims is valid or not. This is a concept of making claims verifiable and it depends on the third party whether the verifier is to be trusted or not.

Verifiable Credentials

As opposed to claims, a Verifiable Credential is a more formal certification since it is represented in a similar form of physical documentation. It can contain more than one claim as it typically does and comes with a signature when and by whom the credential is issued. In the SSI context, a digital credential is required in order to manage online identities in a simple manner.

Physical credentials, most of the time, have to be digitized and stored. This means that often complex and unnecessary processes are employed to create a digital format of a physical credential which in the end can be lost, faked and may disclose more than is needed. The utilization of Verifiable Credentials has many benefits over physical credentials. Integrating digital claims, cryptographic signatures or minimal disclosures can lead to a more secure, scalable and trusted processing information of an Identity Owner. As can be seen in the following figure, a Verifiable Credential can consist of three components (Sporny, Longley, & Chadwick, 2018).

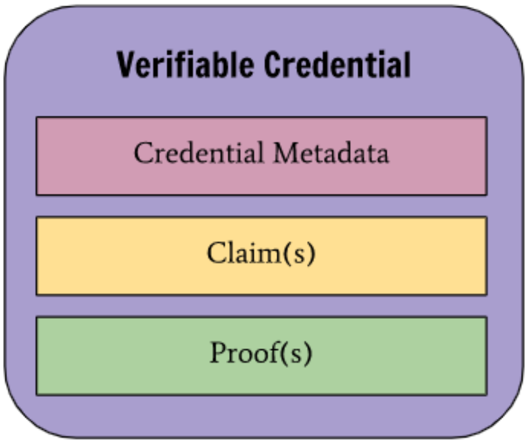


Figure 10. An Illustration of Verifiable Credential Basic Components. Adopted from (Sporny, Longley, & Chadwick, 2018).

The first component refers to credential metadata that describes the properties of the credential related to a public key of the issuer, a credential identifier or the issuance date. The issuer may sign these metadata. The second component expresses whether a Verifiable Credential contains one or more claims. It is possible to have a credential containing multiple claims and often, it is the case that each attribute of the credential is a single claim. Thus, only the requested claims have to be disclosed leaving the other claims untouched. The last component is concerned with the digital proof, which is usually a digital signature. It is a highly important component as it creates a higher level of privacy. Once the third-party signs the Verifiable Credential, a cryptographic proof is generated to demonstrate who issued it. The proof contains the type, the nonce, a timestamp of date and time created, the signature value and the public key of the issuer. By having this information, only then third parties can verify the presented data.

After receiving the credential, the holder stores it inside a digital wallet on a device which in that case can be a smartphone. The wallet serves as a means to form connections with other parties, in which credentials can be used for authentication or to prove authorization to use a service. Generally, the information of Verifiable Credentials should not be written on the ledger because of its immutability, but this depends on the DID method chosen.

Verifiable Presentations

When a holder needs to provide information from one or more Verifiable Credentials to a verifier, she can combine them into a structure called Verifiable Presentation (Sporny, Longley, & Chadwick, 2018). The Verifiable Credentials are either presented directly or derive data formats from Verifiable Credentials that are cryptographically verifiable. The verifiable presentation shares a similar data model to a Verifiable Credential as it is illustrated below.

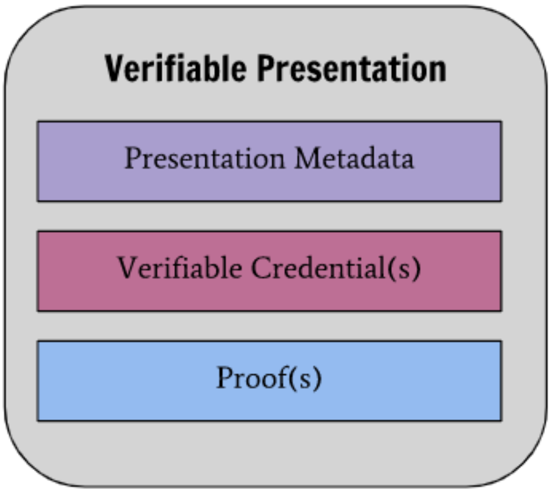


Figure 11. An Illustration of Verifiable Presentation Basic Components. Adopted from (VCWG, 2019).

More specifically, a verifiable presentation can contain metadata that describes information about the presentation including the type and terms of use. It also consists of Verifiable Credentials that contain metadata, relevant claims, and proofs. On the same lines with a Verifiable Credential, the verifiable presentation is signed and thus a digital proof is generated which by itself contains the type, the nonce, a timestamp of date and time created, the signature value and the public key of the presenter. The proof can be created, among other ways, using the cryptographic method of Zero-Knowledge Proofs (ZKPs).

2.2.7 Zero-Knowledge Proofs

One key tool for verifiable presentations is ZKPs mechanisms which enable the presentation of information in a privacy-enhancing manner. This implies that the holder can prove the validity of the issuer’s signature without disclosing the values that were signed. Thus, only selectively disclosing the requested values (Sporny, Longley, & Chadwick, 2018). For example, if a verifier wants to ensure that the user is above 18, only then the proof that the user is indeed above 18 is required and not the actual age. In this way, the privacy of the user is respected and additionally, the verifier received the requested information in a trusted way. In SSI, ZKPs are commonly used and mainly utilized for authentication and proving statements digitally.

Chapter 3

3. Design Science Methodology

After outlining the problem, the research question and the theoretical foundations with their core concepts, the next step is to shed light on the research methods for this research.

3.1 Design Science Research

Existing business research methods mainly focus on positivist and interpretive research paradigms and methods. They typically feature quantitative and qualitative empirical research techniques for research to identify, define, explain and evaluate existing business practices. However, when research is conducted in applied disciplines such as those in business and Information Systems (IS), there is a need to address other goals apart from explanation or evaluation of existing phenomena. This means that such goals require more adequate research paradigms to be addressed besides or in conjunction with positivism and interpretivism (Venable, 2011).

One important goal of business research constitutes the invention and development of new business practices besides simply investigating existing ones. This type of research is better supported by the Design Science Research (DSR) paradigm (Hevner et al., 2004). The DSR paradigm places heavy if not exclusive emphasis on the invention, design, and development of new technologies, techniques, and methods without undermining research rigor. Hence, DSR is used to solve problems in the real world while contributing to the scientific community. Due to the lack of a standardized process in DSR, researchers started to map DSR into a structured process. This group of researchers as well as other DS researchers do not limit the DSR process due to its iterative and exploratory nature. For this reason the process is expected to differ from project to project (Peffers et al., 2007). The DSR process adjusted for this research can be seen in the figure 11.

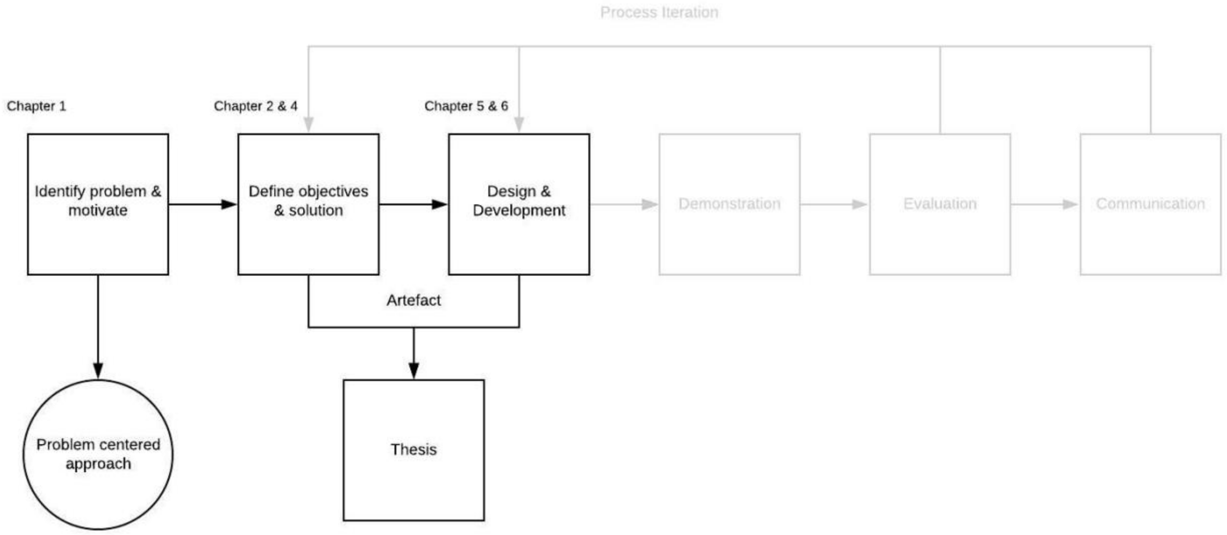


Figure 12. Design Science Research Methodology (DSRM) Process Model.

Peffers et al. (2007), proposed a process model for doing DS research. The research design of this project generally aligns with the DSR process model by following the predefined steps. Hence, firstly identifying the specific research problem and justifying the value of the solution. In a second step, defining the objectives of the solution and knowledge of what is possible and feasible. The next step is concerned with designing and developing artefacts and lastly by evaluating how well the artefact supports the proposed solution to the problem. The solution that has been developed and evaluated including its utility and effectiveness will be communicated and made available with the scope of this study. As can be seen in the figure 11, the DSR process is adjusted to reflect this research including the relevant constraints taken by the authors. More specifically, the steps of demonstration, evaluation, and communication are omitted for this study.

The DSR process is based on three inherent research cycles. The Relevance Cycle which connects the requirements from the contextual environment of the research project with design science activities (Hevner, 2007). The Rigor Cycle links the design science activities to grounding theories and methods. It incorporates new knowledge into the knowledge base of scientific foundations, domain experience and expertise generated by the research (Hevner, 2007). The shifting between theory and data and vice versa places the research between inductive and deductive approaches. The Design Cycle which loops over the core activities of developing and evaluating the produced artefacts and processes of the research (Hevner, 2007).

Sonnenberg and vom Brocke (2012) suggest that each DSR activity including identification, design, construction, and use, shall be followed by an evaluation activity. In a DSR process, an evaluation activity can occur ex-ante i.e. before an artefact is constructed and ex-post, i.e. after an artefact is constructed (Sonnenberg & vom Brocke, 2012). However, because of certain limitations including time constraints in this research, the focus is placed only on the development of artefact which is based on the Relevance Cycle and Rigor Cycle. The evaluation and iterative process which is part of the Design Cycle is not included. Because of technical immaturity of the project and unfamiliarity of stakeholders with the concept of decentralized systems, no feedback has been collected by external stakeholders like industry experts or potential end-users. Their feedback could have potentially been incorporated into the final artefact. This would have initiated the first Design Cycle by incorporating feedback from interviews and user testing into an improved solution. Thus, figure 12 represents the adjusted DSR Framework that excludes the Design Cycle. This opens up new research possibilities based on the produced artefact that outlines an initial version of a decentralized system for public transportation.

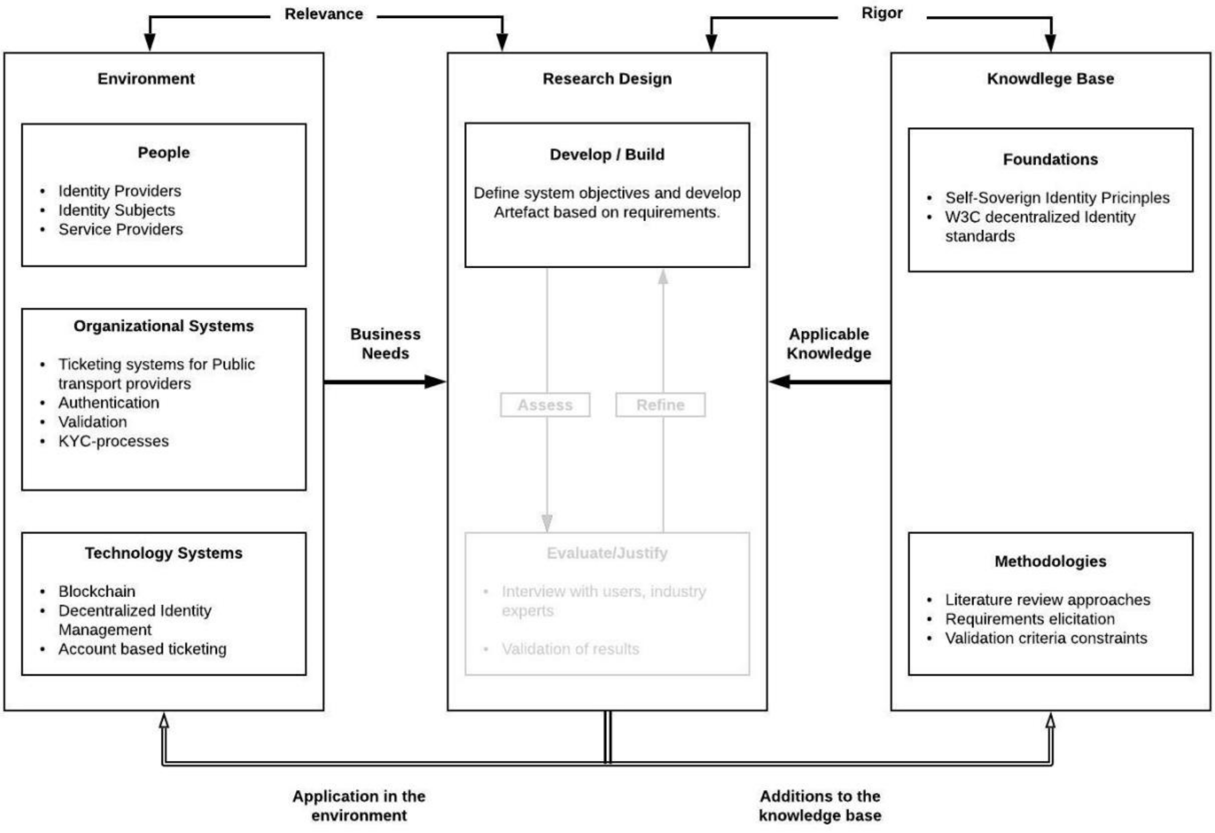

Relevance
Rigor
Environment
People
Identity Providers
Identity Subjects
Service Providers
Organizational Systems
Ticketing systems for Public transport providers
Authentication
Validation
KYC-processes
Technology Systems
Blockchain
Decentralized Identity Management
Account based ticketing
Business Needs
Research Design
Develop / Build
Define system objectives and develop Artefact based on requirements.
Assess
Refine
Evaluate/Justify
Interview with users, industry experts
Validation of results
Applicable Knowledge
Knowdlege Base
Foundations
Self-Soverign Identity Pricinples
W3C decentralized Identity standards
Methodologies
Literature review approaches
Requirements elicitation
Validation criteria constraints
Application in the environment
Additions to the knowledge base

Figure 13. Design Science Research Framework for this project.

3.2 Research Approach and Design

Gleasure (2015) notes that a DSR approach is, among others, suitable for research topics that their prescriptive aspect is below the level of maturity of the analytical aspect. This means that although there are topics for which traditional research has already contributed to the scientific knowledge base, there is a lack of actual application of that knowledge, making the DSR a good fit as a research approach. In chapter 2, extensive research on the fundamentals of Self-Sovereign Identity principles and core technology standards were presented in light of digital Identity Management. However, the actual application of a digital IdMS on Blockchain for the public transport sector is lacking. This project is tapping into this field and since prescriptive knowledge is missing in academic literature, DSR qualifies as the most suitable approach (Gregor et al., 2013).

DSR projects typically are characterized by either more product-centric approaches or more process-oriented approaches (Gleasure, 2015). For this project, a product-centric approach is followed since the aim is to develop and test a solution that facilitates digital Identity Management for public transport providers. Furthermore, this paper places a greater focus on the usefulness of the situated artefact over the theoretical characteristics of it, since the latter has already been addressed in pertinent studies. In the given context, inductive and abductive DSR is preferred which is based on secondary data rooted in pertinent literature leaving primary data out of the scope. Moreover, deductive or “theory-first” DSR is another approach to ensure that the solution fits with the given relying entities’ needs and requirements (Gleasure, 2015). However, because of the lack of academic research in this field a deductive approach has not been followed. Finally, this research develops a better solution in the form of a more efficient service, thus contributing knowledge to an already known problem (Gregor et al., 2013).

3.3 Research Philosophy

According to Saunders (2009) and colleagues, research philosophy is associated with the development of knowledge and its nature. It consists of important assumptions which have been adopted through the way the world is viewed (Saunders et al., 2009). In the world of research, there are two branches of philosophy; ontology and epistemology. Ontology, from a philosophical point of view, deals with the nature of reality and the different entities within reality (Saunders et al., 2009). Epistemology, on the other hand, is concerned with the theory of knowledge and what constitutes acceptable knowledge in a field of study (Saunders et al., 2009).

In Design Science Research, the community of researchers has yet to produce clear definitions of world-views. According to Purao (2013), most researchers within DSR avoid referring to their ontological and epistemological assumptions. They usually choose to view their research under pragmatism, while acknowledging the difference of DSR to traditional research and their respective world-views (Purao, 2013). Along the lines, this research acknowledges that traditional world-views – ontology and epistemology – might not suit design science research.

This research is argued to follow a more positivist world-view characterized by a more realist ontology and objectivist epistemology. This is reasonable in the sense that the literature review is more technology-driven. At the same time, the conducting research is not based on subjective views gathered by a specific group, but instead the solution is meant to be designed for different target groups. In terms of epistemology, the research follows a "knowing-via-making" approach over a "knowing-via-participating" approach from the subjective world-view (Purao, 2013). Thus, the research operationalizes the idea contained in this approach by developing a meaningful artefact for the identified problem.

Besides world-views, the validity of this research needs to be assessed based on how well the results of the underlying theoretical models can be generalized and used for future research (Bryman & Bell, 2011). In the case of this project, changes are accounted for through the interaction of the users with the solution. This might imply that the outcomes cannot be generalized and are unique to the project. The study is conducted on a solid theoretical foundation and follows a reliable research process. Thus, it can be argued that this study can be used as a foundation for future research in its specific field.

According to Goldkuhl (2012), the basis of DSR constitutes a problematic situation that calls for improvement. The improved state knowledge is developed through inquiry processes starting by evaluating the situation as-is and it is advanced through building and evaluating relevant artefacts until a desirable situation "to-be" is realized (Goldkuhl, 2012; Hevner, 2007). This entails that a major concern in the DSR community is the utility and useful contribution to practice through "means-to-ends" descriptions and development of artefacts that solve the identified problem (Goldkuhl, 2012; Hevner, 2007).

From an epistemological standpoint, as a result, DSR knowledge is grouped into three blocks that are sequentially related. The first block consists of evaluative knowledge of a problematic as-is situation. The second block is related to prospective and normative knowledge about the desirable situation. The last block describes the to-be artefact up until prescriptive knowledge is reached. This helps clarify the use values of the underlying solution (Goldkuhl, 2012). As Goldkuhl states in his paper, the purpose of knowledge in DSR is to improve a given situation through action.

For this research, the sources of knowledge need to derive from various sources such as the knowledge base of known theories, the as-is situation and its analysis as well as the authors experience and reasoning. Due to the purpose of this research which is to explore and propose a DIdMS for the public transport sector, the chosen methods need to reflect upon this by contributing to the scientific knowledge base by proposing an effective solution.

In this research, the knowledge inquiry is built through studying pertinent literature and analyzing a phenomenon. It is not characterized by an iterative cycle of knowledge which includes gathering observations from stakeholders or building and evaluating artefacts based on their input. The outcome is solely explained by previous academic research and is not grounded in empirical data. Instead input is taken from previous studies in the field of digital Identity Management and Blockchain Technology, the given phenomenon as well as the authors experience to draw inferences for the given research question.

Chapter 4

4. Analysis of Relevant Systems

In order to develop a DIdMS for the Public Transportation Sector a deeper analysis of existing technologies needs to be considered. Thus, current system implementations are dissected into different components in order to develop a framework for assessing and evaluating the developed artefact. Moreover, Blockchain-based IdMS are analysed in accordance with SSI principles. This will produce a matrix to identify the best suited underlying system that can be used as a foundation for developing a DIdMS that aligns with Christopher Allen’s (2016) SSI Principles.

4.1 Ticketing in Public Transportation

Ticketing systems for public transportation in Europe are implemented in all kinds of variations. Some systems are dependent on Paper Tickets whereas other systems make use of Smart Cards or even account-based solutions (Calypso, 2017). Since each transportation agency decides on its own what kind of system they are implementing, there is no consistent solution on a pan European level. Most systems run independently and cannot interact or communicate with other systems outside their Ecosystem (*Roadmap to a Single European Transport Area – Towards a competitive and resource efficient*, 2011). The following describes the most common ticketing solutions and their system architecture.

4.1.1 Pre-paid Tickets

Comparing European cities based on their ticketing systems shows that pre-paid tickets are one of the most popular ticketing systems. Pre-paid tickets are tickets that have been paid before the start of a trip. They can be bought on vending machines or apps that are provided by the transportation service. Tickets are usually valid for a specific timeframe and are linked to specific transportation modes and zones. The validity of those tickets is mostly proven through timestamps or magnetic stripes that are added onto the ticket (AECOM, 2011). Thus, after the time has been exceeded a ticket becomes invalid and needs to be purchased again. Paper tickets live independently from each other and cannot be reused. The ticket is only valid for the passenger that is in possession of it. This means the value of the ticket cannot be restored when lost.

4.1.2 Smart Cards

Smart cards are physical cards which are issued by transportation authorities. Usually, It can be purchased on-site at a vending machine or at the ticket office. Smart cards are not reliant on any back-office processing since all of the information is saved on the card itself. Travel rights and account balances are registered on the card and get processed by validator terminals that control access to the underlying transportation systems. Through visual and acoustic signals users and other involved parties can check if a transaction was valid or not (Calypso, 2017). For example, when entering an underground station, a user can enter the property only after successfully validating her Smart Card on the gate. Since all credentials are stored on the card itself users either need to have preloaded balances or pay-as-you-go terminals available in order to access public transportation. Card-based systems rely on fare calculation and applicative software that is stored on memory in equipment like validators, sales machines or inspection readers. This is also referred to as front office processing (Calypso, 2017). Thus, these applications always need to be synced with up to date fares and software in order to function correctly. Logic and business aspects are encoded into the front office and are not stored in a centralized database. Most Smart Cards rely on specific system implementation and cannot be used outside the system; they are also described as closed-loop payments. Cards, terminal validators are unique to each system and therefore these systems have high implementation and maintenance costs. In card-based ticketing systems, the card itself gives authentication to travel and there is no centralized Identity Management in place that stores data into a database. Thus, when losing a card, balances cannot be recovered and existing balances are lost. Smart cards have evolved over time but one of the biggest bottlenecks of these systems is that every information needs to be stored on the card itself. Thus, adding new balances in a card-centric system needs to be done through systems that are implemented inside the ecosystem of a card-centric transportation system (Calypso, 2017). Figure 13 describes the high-level architecture of a card-centric system. Here user data gets updated on the card itself when interacting with any validator machine inside the system. Each validator usually has its application logic like fare prices directly saved on the machine itself.

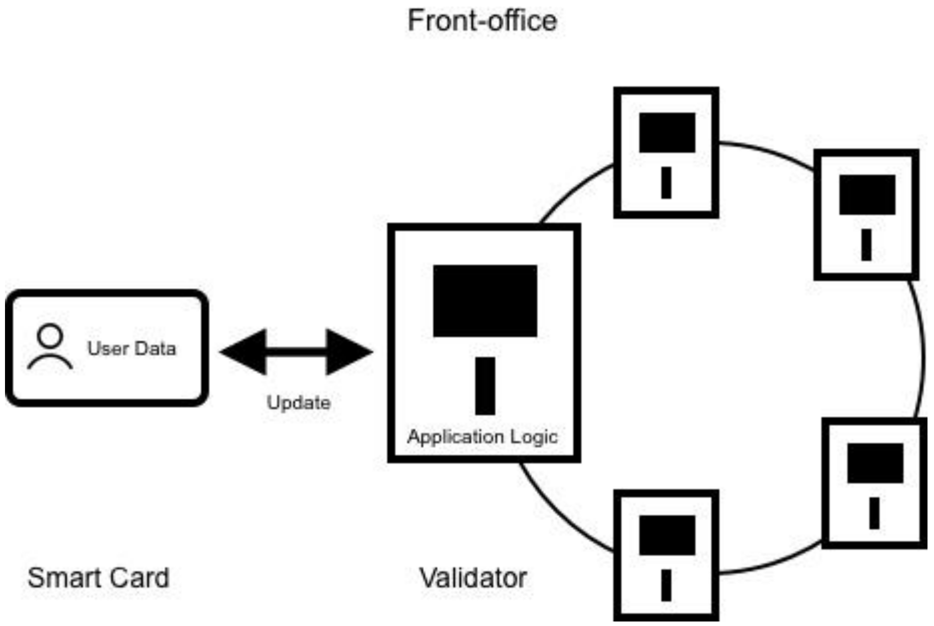


Figure 14. High-level System Design of a Card Centric System.

4.1.3 Account-Based Ticketing (ABT)

The improved network availability is facilitating the shift away from Smart Cards to account-based systems where user data is stored and accessed from centralized databases. Thus, terminals, validators and other equipment that are needed to provide access management in public transportation do not need to store information in memory anymore. Smart cards are replaced through identifiers which link users to their associated account and validate transactions on the network. Account-based ticketing is increasing in popularity and slowly replacing Smart Cards that are issued by each transportation agency. In account-based ticketing, EMV technology is often used to facilitate transactions and use existing media to control access and account management. EMV payment technology is already implemented in most credit cards. Thus, many account-based systems like the TF London Underground system started to accept EMV to access their public transportation services (Giesecke & Devrient, 2011). ABT systems allow the usage of different cards and are considered as open-loop systems. ABT moves all the fare price calculations and application logic from the front office to the back office (or on the cloud). Thus, on-site terminals do not have a ticketing function anymore and portable objects like Smart Cards are only a means of authentication. Handling all processes in centralized places allows for shorter update cycles, applying algorithms to an entire dataset and also decreases maintenance cost of equipment. ABT accounts are accessible online, and changes are immediately in effect. Thus, systems are starting to move from a card-centric approach to a more system-centric approach (Calypso, 2017). Figure 14 illustrates the high-level architecture of an ABT system. Here different types of media are used as an identifier. The validator machine passes an identifier like a user id to the centralized back-office. This in return validates and updates the user account based on a system centric application logic. The validator then rejects or accepts a passenger based on the returned values.

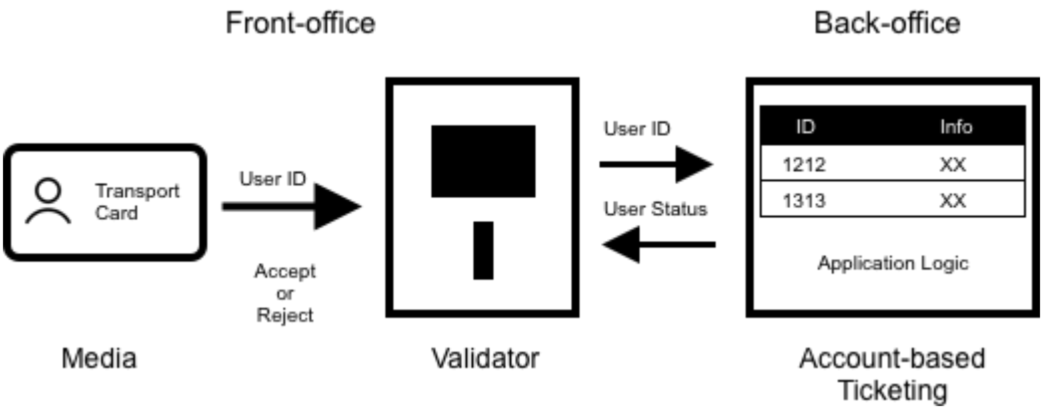


Figure 15. High-level Architecture of ABT.

4.2 Comparing Ticketing Systems

Looking at three different public transport ticketing systems highlights the advantages of an account-based solution over Paper Tickets and Smart Cards. Thus, the following attributes from an ABT system can be defined in order to develop a framework to evaluate public ticketing solutions.

- **System centric -** Users have a specified account in a centralized database. Users are controlling a custom account that holds their account information like account balances.
- **Separation of concerns** - The ticketing system separates between front and back office processes. Each part of the system is focusing on a different set of functionalities. The front office takes care of access control while the back-office controls the application logic. This lowers maintenance and updates costs for transportation providers.
- **Reusable -** The user is able to reuse one identifier multiple times. There is no need to create a new ticket on every purchase.
- **Interoperable -** The system is able to integrate with other systems outside their ecosystem. Passengers can reuse their accounts independently of the system.
- **Restorable** - The value inside the system is not bound to a specific ticket. The value is transferable inside the system. This attributes value to an identity and not a physical object.

	Paper Tickets	Smart Cards
System-centric	✗	✗
Separation of concerns	✗	✗
Reusable	✗	✓
Interoperable	✗	✗
Restorable	✗	✗

Table 5. Functionality Comparison of Public Transport Ticketing systems.

The table above compares traditional Paper Tickets and Smart Cards based on the high-level attributes of account-based solutions. The table highlights that Smart Cards are different when it comes to reusability of identifiers in the system. The reusability of Smart Cards is defined by the possibility of adding additional balances to an existing card. Paper tickets usually have to be reissued when expired. However, compared to ABT Smart Cards lack many features that can be linked to better usability for the end user and cost savings for public transportation providers. Thus, the defined ABT attributes will be used to evaluate the outcome of this research.

4.3 Analysis of Blockchain-based SSI Management Systems

This section will present three solutions that leverage Blockchain Technology to develop Self-sovereign Identity systems. These systems have been deployed at the application layer where a Blockchain-based system resides underneath. A few notable examples of such systems are Sovrin, uPort, and Civic. Next, the paper will explore the functionalities of these systems and investigate whether or not they satisfy the principles of Self-sovereign Identity. In order to provide a better overview in the analysis, each related SSI principle is highlighted along its evaluation criteria.

4.3.1 Sovrin

Sovrin Foundation is a private non-profit entity whose goal is to standardize and create an infrastructure for self-sovereign identities by utilizing its own Blockchain called Sovrin ledger which is based on Hyperledger’s Indy. Sovrin is leveraging a consensus algorithm called Plenum which is responsible for validating new transactions (D. Reed et al., 2016). The SSI model of Sovrin is independent of any available distributed ledger but it has the flexibility to work with any Blockchain that satisfies the fundamental principles. Sovrin utilizes a public permissioned Blockchain using nodes also known as Stewards to achieve a global consensus. Sovrin is managed by the Sovrin Foundation, the governance model which decentralizes the power to multiple parties and approves new Stewards to participate in the network in an open and transparent manner. The system of Sovrin also provides functionality for issuing and managing credentials under a privacy-preserving way by generating Zero-Knowledge Proofs. Those are based on the users' credentials allowing them to exercise control. Here they can choose selectively which data they want to share with someone else. Users can utilize third-party mobile or web apps that act as a Sovrin client to interact with the ledger in order to perform, create, read and manage their identity data.

4.3.1.1 Analysis

The public permissioned ledger utilized by Sovrin requires a governing body to approve the participating nodes in a trusted way depending on the reputation of the Stewards (*Protection*). Although the Sovrin Foundation has power over the ledger, attributes are not shared with Stewards and administrators without the consent of the Identity Owners (*Consent*). All their private information is encrypted by the Identity Owners' own keys. Sovrin's codebase is open source which means that the code that is used for different operations, validations and access to the ledger is publicly available to everyone (*Transparency*). It can also be deployed on any distributed ledger that meets the requirements making it interoperable with other ledgers (*Interoperability*). Selective disclosure of verifiable claims based on Zero-Knowledge Proofs is provided by Sovrin Decentralized Identifiers and public keys (*Minimization*). Users' private data is stored on their device or a selected Agent and does not reside in any of the system's database (*Access*). Also, a Sovrin Agent can enable secure messaging between the clients and maintain an encrypted backup of private storage by utilizing Local Containers (*Persistence*). The only way to unlock the identity of the Identity Owners is by using their key pair which enables everything else in the system (*Control*). Portability of data is ensured by utilizing system-independent data formats like JSON-LD (*Portability*).

4.3.2 uPort

uPort is a decentralized identity system built on top of the Ethereum platform supporting the Self-sovereign Identity model (Lundkvist et al., 2016). It comprises a mobile app and several Ethereum contracts including a public registry of uPort identity. The uPort mobile app generates a key pair that allows them to create, update and share identity information with other users. In the backend, three smart contracts are utilized to control the users' data. More specifically, a Proxy Contract is deployed as the user's unique identifier, a Controller Contract to provide access and a Recovery Quorum

Contract to assist with recovery of a user's identity in case they lose access to it. The bulk of identity data is stored on a distributed file system (IPFS) while the corresponding private key of a uPort identity is stored on the mobile app. The public registry is used to create a correlation between IPFS data and a uPort identifier. It is also notable to mention Jolocom which is another SSI system whose functionality is quite similar to uPort. Jolocom is also developed on top of Ethereum and utilizes several Ethereum smart contracts. The main difference of a Jolocom identity from a uPort one is the way identity data is structured and represented (Fei et al., 2018).

4.3.2.1 Analysis

uPort allows users to create, update and control their identity (*Control*) and also share personal information with third parties at their own discretion (*Consent*). The core identity is stored on the Ethereum Blockchain. It is also replicated and stored on other distributed computers across the globe (*Persistence*). The private information is stored on the users' devices as well as off-chain with IPFS which makes it always accessible by the users (*Access*). uPort is partly decentralized by having only a few centralized elements (*Protection*). Those include the messaging server which is responsible for transferring attributes, a push notification center and an application manager that allows developers to create and manage identities for their own applications (*Portability*). uPort is developed with open source libraries and open standards (*Transparency*). Also, uPort gives the possibility for developers to create their own compatible applications with the system (*Interoperability*). uPort generates JSON Web Tokens as verification on claims and provides the ability for "Selective Disclose Request" for private data. The profile of the user is stored in the public registry of uPort which inevitably makes it prone to information leakage regarding specific attributes of the user and compromises the privacy of users.

4.3.3 Civic

Civic is building an ecosystem that is designed to facilitate low-cost access to Identity Verification (IDV) and Know Your Customer (KYC) processes. Civic utilizes the Ethereum Blockchain and has created an ERC20 token called CVC that is stored in an Ethereum wallet. The token is used to reward and pay for services in the ecosystem and it has a fixed supply. Identity information is stored on the users' device and Civic receives only hashes of the data which are stored on the Blockchain (Civic Technologies, 2017). The user provides identity information to a validator for validation. The user registers to a service by providing a requested proof to the Service Provider. Civic also stores attested personal identity information in a Merkle tree and records it in the Blockchain.

4.3.3.1 Analysis

The identity information is stored on the user's device (*Control*) which makes it always accessible to the user (*Access*). Civic is based on the Ethereum Blockchain and therefore it has no proprietary software or infrastructure (*Transparency*). Although the network is likely to be available in the foreseeable future, the actual data storage lifespan depends entirely on the user (*Persistence*). Applications connected to the Civic ecosystem can use the identity information (*Protection*), although the information cannot be ported to any other devices (*Portability*). Civic among others can provide claimed and verified identity attributes for all types of services and also passwordless login (*Interoperability*). Since all the information is stored on the user's device, the Identity Owner selects what information to share and with who (*Consent*). The user selects what information to reveal which is stored with hashes in a Merkle tree (*Minimization*).

4.4 Comparison of Available SSI Management Systems

As presented in section 2.1.1, Cristopher Allen has suggested ten principles that define a Self-sovereign Identity. They represent an ideal system for building an SSI based Identity Management System. In this section, these principles will be used as criteria in the comparison for Blockchain-based systems for SSI. A more detailed comparison of the three selected systems is presented in table 6.

All three SSI systems provide the Identity Owner with full control over their identity and the ability to disclose claims and attributes selectively. To some extent, they all also provide portability and persistence by utilizing established data formats and self-hosting of data. More specifically, uPort utilizes IPFS (InterPlanetary File System, a distributed file system) for storage, while storage and backup are facilitated by Sovrin with trusted networks Agents. In the case of uPort and Civic, both of the systems store the identity information on the user device by using Blockchains for verified and signed hashes of the data. This provides the user with full control over the information. Sovrin, on the other hand, is storing most of the identity records on-chain. This offers better accessibility and utilization.

uPort and Civic are based on the Ethereum Blockchain which means that the computational power of the network is inherited and thus both systems are more resilient to third party influence. In the case of Sovrin, the system is designed to be agnostic of the Blockchain utilized. This makes it more appealing to a bigger market since it can cover a broader range of identity needs as opposed to the Ethereum based systems.

Although all the presented systems leverage some decentralized techniques, none of them is entirely decentralized. More concretely, Sovrin is running on a Hyperledger Blockchain and its consensus algorithm is based on approved nodes also known as Stewards that are run by organizations interested in maintaining the network health. In addition, Civic utilizes the verification providers, Validators, that can verify identity information that their role is centralized in the ecosystem.

As it is depicted below in the tabular comparison of self-sovereign Identity Management Systems, the (✓) describes if the criterion is met by the system, while the (✗) symbol means the opposite effect which is that the respective criterion is not met.

Criterion	Sovrin	uPort	Civic
1. Existence	✓	✓	✓
2. Control	✓	✓	✓
3. Access	✓	✓	✗
4. Transparency	✓	✓	✓
5. Persistence	✓	✓	✓
6. Portability	✓	✓	✗
7. Interoperability	✓	✓	✓
8. Consent	✓	✓	✓
9. Minimization	✓	✓	✓
10. Protection	✗	✗	✗

Table 6. Comparison of Blockchain-based SSI Systems.

Chapter 5

5. System Objectives & Requirements

This chapter will shed light on identifying the system objectives in order to develop a functional prototype. It sets the requirements that need to be followed in order to implement a feasible solution for the public transport sector. Those requirements are being translated into use cases and scenarios. Those can be used to describe a first DIdMS implementation based on Blockchain and SSI principles.

5.1 System Objectives

In order to develop and design a DIdMS for public transportation certain system objectives have to be determined first. Those objectives are extracted from current implementations of DIdMS as well as requirements that derive from specific interactions and relationships between stakeholders in public transportation. Defining the interactions and relationships between stakeholders are an important step in order to design technical specifications of an identity system. Here decentralized identity concepts like DIDs, schemas, credentials, and the overall identity verification process are modelled along the lines of public transportation. Those objectives will lead to more specific technical requirements that can be translated into an overall system design.

5.1.1 Stakeholder Relationships

The idea of a pan European transportation ticket that is implemented through a DIdMS requires the accounting for certain types of relationships between stakeholders. Those relationships will be the foundation for trust within the system and ensure a seamless travel experience between different public transportation systems across Europe. In order to account for those relationships, it's important to define the different types of stakeholders and their responsibilities in the system. Stakeholders include Public Transportation Authorities, Passengers, Institutions, Identity Owners and Identity Providers.

5.1.1.1 Public Transportation Authorities / Service Provider

PTAs provide access to certain types of transportation modes and are responsible for issuing and verifying tickets. They are considered as the service providing authority that is constantly requesting access rights by users in order to provide transportation services. Moreover, PTAs act as trusted partners for any other PTA and can issue certain types of Verifiable Credentials to users inside the system. Those credentials can include, for example, student discount rights. This would allow verified users to travel on a discount in the entire system.

5.1.1.2 Passengers / Identity Owners

Passengers are the service seeking party that uses credentials to authorize themselves to any *Public Transportation Authority (PTA)* in Europe. They are able to selectively disclose information that is requested by Service Providers. Thus, they are reliant on identity proof of PTAs and other involving institutions. In the current system passengers are acting as Identity Owners.

5.1.1.3 Institutions / Identity Providers

Institutions are all stakeholders that are able to issue Verifiable Credentials that attest certain user attributes like age, student status. Institutions, thus, act as Identity Providers. For example, universities are responsible for providing attestations of student enrolment that can be used to claim discount rights to PTAs on Public Transportation. Moreover, it includes institutions like governments that issue national identities as well as banks that can attest payment liquidity for users in order to pay for services.

5.1.2 Domain-specific Trust Framework

In order to establish trust between verifiers there needs to be a domain-specific trust framework established. For example, a PTA needs to be able to know all other PTAs in the system in order to function on a European level. Here the European Union Transportation Authority could act as a framework that pools together all domain-specific trust partners. The same accounts for any other institutional stakeholder in the system. Thus, a system objective will be that the initial trust of all institutions in the system is granted by default.

5.1.3 Stakeholder Identification

In order to ensure trust between stakeholders it is important that each of them is uniquely identifiable in the system. Therefore, the system needs to be able to register DIDs for each stakeholder. DIDs can then be used to create connections between stakeholders in order to issue and verify credentials between them. Moreover, users need to be able to get easy access to manage their DIDs and thus allow them to claim ownership over credentials.

5.1.4 GDPR Considerations

Data privacy through GDPR is an important consideration in the system objectives. Thus, the system needs to account for any type of privacy concern and act along the lines of European data regulations. Therefore, the underlying technology needs to be examined on whether it satisfies the GDPR requirements and to what extent. In other words, several requirements will be derived from GDPR and arguments will be made whether Blockchain Technology satisfies them.

According to GDPR, any personal data of a subject shall be accessible by other entities with informed consent of the subject. Data recorded on a Blockchain is pseudonymous which means that it does not contain explicit personal data but only unique references to it. However, on some Blockchains the identifiers can be traced back to an IP address which inevitably classifies as personal data. Verifiable claims that are recorded off the Blockchain are shared only with informed consent of the subject with other entities. In addition to this requirement, another GDPR consideration is to provide means for the subject to control consent. This allows them to decide who they share personal data with and whether consent can be provided and revoked. Consequently, a Blockchain-based system can provide suitable methods to record and revoke consent by recording on the ledger an immutable and granular proof of consent. In a similar manner, a consent revocation can be published by the subject and can be recorded on the ledger.

Article 17 of GDPR which states the right to erasure or the so-called "right to be forgotten" constitutes a major consideration when designing a system. According to the aforementioned article, any records of personal data owned by other entities shall be erased when requested by the subject without undue delay. Therefore, the system shall not record any private claims or proofs on the Blockchain but only a proof of issuance or revocation can be published on the Blockchain. However, if an entity publishes a public statement on the Blockchain, it cannot be erased. In order to fully comply with GDPR, the entity also has to erase copies of personal data on its servers, which is out of the scope of the SSI system.

Furthermore, two more considerations arise from GDPR that shall be addressed by the system. The first is concerned with the portability of personal data in an automated fashion. On a Blockchain-based system, personal data is stored in a machine-readable format in the subject's claims repository which makes it Blockchain-independent and, therefore, can be ported to other systems. The second is to ensure that personal data is managed in a secure and private manner by design, through the underlying technology of the system. Data stored on a Blockchain is cryptographically secure by default. The system shall not be utilized to record private data since it may pose risks to a lack of privacy due to the transparent nature of Blockchains. Thus, only non-private data should be stored on the ledger.

5.1.5 Read and Write on the Blockchain

As it was covered in section 2.2.1, there are several different types of Blockchain systems characterized along two dimensions: access and validation. The developed system needs to be publicly accessible in order to verify credentials by different stakeholders. However, when it comes to writing permissions the system can either be permissionless or permissioned. One the one hand, the permissioned model means that anyone can access it (public) but it is limited to who can participate in the validation process (permissioned). Specifically, the system is designed to operate in a way that everyone can see the contents of the Blockchain, but only approved participants are permitted to validate transactions. On the other hand, in a permissionless system anyone could participate in the validation process. Thus, the system can be either public-permissionless or public--permissioned in order to act as an appropriate solution for Identity Management in the public transport sector.

Access	**Validation**		
		Permissionless	**Permissioned**
	Public	X	X
	Private		

Table 7. Read and Write Permissions of the System.

The publicly designed system is an important constraint on the type of data that can be stored on the Blockchain. More specifically, only public data can be stored on the Blockchain and not any other data, even if it is encrypted. It is a good design choice not to store any private data on-chain. Although this data can be cryptographically stored and accessed on a public Blockchain by those with the decryption key. This occurs in order to ensure that encryption algorithms used today are not broken in the foreseeable future. Thus, the data can be exposed to a set of vulnerabilities. Therefore, data could be eventually decrypted and be publicly accessible by anyone.

Taking a closer look at what data can be stored on the ledger, there are four types of transactions:

- **DIDs** - the public Decentralized Identifiers created by each participating entity in the system (Identity Owners, Service Providers, etc.) and associated DID documents with verification keys and endpoints.
- **Schemas** - entries that define the claims which contain the data attributes that will be part of a Verifiable Credential.
- **Credential Definitions** - entries created by a credential issuer using a specific schema and optionally a revocation registry.
- **Revocations Registries** - entries that give an issuer a way to revoke issued credentials in a non-correlatable way.

Most significantly, there are neither credentials nor private data of any kind on the ledger. The only data stored on the ledger is public data that constitutes public information, and anyone can access it. All private data is exchanged between identities directly and stored in their secure wallets. The data on the ledger is utilized to prove claims about the Identity Owner's data that can be trusted.

5.1.6 Off-chain Data Storage

In order to ensure a secure peer-to-peer data exchange, the overall system's objective is to integrate with off-chain data storage solutions. The user must be able to decide what kind of data storage he or she wants to choose. This can range from self-hosted solutions like device storage to cloud-based storage providers like Amazon Web Services or Microsoft Azure. Cloud-based solutions are referred to as Identity Hubs. Thus, the system needs to be able to authenticate with those Identity Hubs to retrieve credentials. Identity Hubs can bring the advantage that the Identity Owner does not need to be online to provide proofs to validators. Moreover, it would be device agnostic and credentials could be managed on multiple devices. This allows the connection via service endpoints to Identity Hubs in a seamless way. Moreover, this ensures that data can be recovered and is not lost, for example, when losing the specific device that functions as an offline storage solution.

5.1.7 Schema Registration

Schemas are an important part of any decentralized identity system. It defines the overall structure of credentials and allows issuing and proof requesting credentials inside the system. Thus, the system needs to be able to register new schemas as well as retrieve existing ones. The system has to differentiate between so-called *Schema Definitions* and *Credential Definitions.* In a broader context Schema Definitions should be defined by an overarching authority that is able to define a common structure. This can be used by anyone in the industry in order to issue specific Credentials. For this reason, a fully functional system needs to be able to write Schema definitions on the ledger. This makes it possible for any other stakeholder to discover and retrieve commonly used Schema Definitions. For example, in the underlying use-case, it can be assumed that some Schemas are registered by a governing body like a specific European Agency. Thus, industry standards across different countries in Europe can be achieved. The following table describes the different roles and purposes of a specific Schema.

Authority / Schema Designer	**Credential Issuer**	**Purpose**
European Transportation Authority	Public Transportation Provider (PTP)	Providing consistent proof to travel across different Public Transportation Systems in Europe.
European Banking Authority	Banking Provider (BP)	Providing consistent proof of liquidity across European Banks in order to pay for rides on public transport.
European Education Authority	University Office (UO)	Providing consistent proof for student status across European Universities.
National Government	National Office (NO)	Providing proof of national identity inside a national jurisdiction.

Table 8. Schema Registration by Different Entities.

As table 8 describes there are three European Agencies that are responsible for registering a Schema in order to achieve consistency across the system. These are namely the banking, education, and transportation industry that are currently overseen on a European level. In regards to national identity, the system acknowledges the sovereignty of each state. Thus, each government is responsible for its own schema design which can be used by national institutions to issue their own credential definitions on the ledger.

5.2 System Functional and Non-Functional Requirements

In order to create a usable system, there are certain types of use-cases taken into account that the system needs to handle. Those use-cases are derived from existing public transportation solutions as well as from other Identity Management solutions that have been analysed in chapter 4. An important use-case is the overall management of user accounts. Thus, the user needs to be able to access and store data as well as manage cryptographic keys for authentication in a user-friendly manner. Therefore, the system will provide a user interface that helps to store and manage credentials. Managing credentials will include the retrieval and storage of documents on Identity Hubs as well as self-storage solutions like the local storage of mobile phones. Moreover, the system needs to handle the following use-cases in order to function as a decentralised identity system. Each step will describe its rationale as well as functional and non-functional requirements.

In this section, preliminary functional and non-functional requirements are defined. This list of requirements is not exhausted however aligns with the requirements for a minimum viable system. In future implementations those requirements can be extended for new use-cases. According to technical specifications of functional and non-functional requirements, the keywords, several words are used to signify the requirements. In this documentation, the keywords that have been used are “MUST” and “SHOULD” and follow the guidelines. More specifically, “MUST” is used to indicate that the requirement is an absolute requirement and must exist in the developed system. On the other hand, “SHOULD" is used to describe that a requirement may not exist in the full implementation of the system, but the implications of choosing not to implement it must be understood.

The core functionalities of the system have been identified and are viewed as a **MUST**. Moreover, the requirements have been separated into functional requirements which are denoted by “FR” and non-functional requirements which are denoted by “NFR” (Bander, 1997).

5.2.1 Establishing a Digital Identity

The identity subjects need to be able to register an identifier in order to participate in the decentralized system. Thus, Identity Subject or Holder needs to be in control of the identifier. This ensures that data can be associated with a specific identifier.

No.	**Requirement**
FR1	The system **MUST** allow any natural person to create a decentralized identifier.
FR2	The system **MUST** allow the subject to create many Decentralized Identifiers as needed.
FR3	The user **MUST** be able to control his or her associated data with the corresponding identifier.
NFR1	The system **MUST** provide a public-private key pair to the user.
NFR2	The system **MUST** operate on a decentralized information system to issue identifiers.

Table 9. Requirements for the Use-case of Establishing a Digital Identity.

5.2.2 Establishing Relationships

In order to manage different relationships and interactions between stakeholders there needs to be a way to establish connections between the actors. Those relationships need to be accepted by all involved parties. Thus, the system needs to provide a way to discover users based on their DIDs in order to allow the establishment of a connection. This ensures a secure data exchange without leaking data to any outside relationship.

No.	**Requirement**
FR4	The user **MUST** be able to connect with other DIDs.
FR5	The user **SHOULD** be able to accept and reject connections.
NFR3	The system **MUST** support secure data exchanges.
NFR4	The system **SHOULD** display relationships in a human-readable way.

Table 10. Requirements for the Use-case of Establishing Relationships.

5.2.3 Issuing Credentials

In order to prove claims to other stakeholders there needs to be the possibility to issue Verifiable Credentials or self-issued credentials. Those credentials can either be credentials that are self-claimed by the identity subject or by an institution that is able to issue Verifiable Credentials about the subject. This can be, for example, an official document like an ID card issued by a government. Thus, the system needs to be able to add values to any attribute defined in a specific Credential Schema. This can be either done through manual inputs or through more sophisticated direct database integrations to databases of the credential issuer. Issuing credentials can also include the acceptance by the subject in order to verify the correctness of the proven data.

No.	**Requirement**
FR6	The system **MUST** allow entities to act as Identity Providers to issue verifiable claims to a subject.
FR7	The user **SHOULD** be able to create self-issued credentials.
FR8	The user **SHOULD** be able to approve and reject Verifiable Credentials.
FR9	The user **MUST** be able to register Schema definitions in a non-programmable way.
NFR5	The system **MUST** enable claims to be issued in a human-readable format with standardized semantics**.**
NFR6	The system **SHOULD** be able to integrate with external databases.

Table 11. Requirements for the Use-case of Issuing Credentials.

5.2.4 Credential Management

Identity subjects need to have access to their own credentials in order to manage those. The system needs to be able to connect to user-defined Identity Hubs in order to access all issued credentials. The management of those credentials needs to be done in a secure way. Thus, data needs to be stored encrypted and on request be decrypted by the identity subject. Moreover, issuing parties need to be able to revoke credentials as soon as those are not valid anymore.

No.	**Requirement**
FR10	The user **MUST** be able to manage and store credentials in a secure way.
FR11	The user **SHOULD** be able to decide where to store credentials.
FR12	A user **SHOULD** be able to revoke claims or attestations to claims that have been issued.
NFR7	The system **MUST** enable data encryption.
NFR8	The system **MUST** expose authentication methods with external Identity Hubs and Registries.

Table 12. Requirements for the Use-case of Credential Management.

5.2.5 Proving Claims/Assert Claim

In order to prove claims from certain credentials, the system needs to be able to ask the identity subject to give consent in order to share specific claims. This can ensure data minimization on presenting claims for any type of proof request. Thus, the user always needs to be informed on what type of claims are requested by outside parties and what data is presented to connected stakeholders. This increases the transparency of the system and aligns it with the overall principles of Self-sovereign Identity.

No.	Requirement
FR11	The user **MUST** be able to prove claims.
FR12	The user **MUST** be able to give consent to request a proof of claim.
FR13	The user **MUST** be able to request Verifiable Credentials.
FR14	The user **MUST** be able to selectively disclose data requested.
NFR9	The system **MUST** handle claims in a transparent way.
NFR10	The system **MUST** enable data minimization through Zero-Knowledge proofs.

Table 13. Requirements for the Use-case of Proving and Asserting Claims.

5.3 User Scenarios and Use Cases

In order to better understand the different functionalities of the system different use-case scenarios are showcased. The following section describes three scenarios a user can be found in while engaging with the system. The scenarios assume that the user who is the Identity Holder is already in control of a decentralized identity. This identity, for example, could have been registered through Apps like Jolocom or any other Identity Management solution that is built upon a decentralized infrastructure. The different use cases describe the process from requesting base credentials to requesting and using a student specific Travel Credential.

Use cases are an Unified Modelling Language (UML) tool that allows to present how core functionalities could be facilitated in the domain of the system (Ashbacher, 2004). This helps to get a better understanding of the functionality of the system, therefore use case diagrams have been constructed.

5.3.1 Scenario 1: Requesting Base Credentials

This scenario is built upon the requirements that a student needs to fulfil in order to become eligible for a discounted Travel Credential. The requirements for a student include student status, student name and proof of bank account. Thus, in order to get a discounted Travel Credential issued the student needs to engage with different organisations that are able to issue the necessary base credentials. Overall the process of issuing one set of credentials will follow similar actions. For this reason, the described scenario can be applied to any other base credentials issuing scenario. In order for the student to receive a student status credential, the student needs to directly engage with her university. Firstly, she would directly visit the university to ask for a student transcript. After showing up at the university she would prove with her student card or any other identification document that she is an official student. The university office would then confirm the authenticity of the shown document. If the university confirms that she is an active student, she would receive a digital student transcript by the university. Now the student is able to confirm the correctness of the issued document and save a digital representation of the credential into her Identity Hub or wallet.

5.3.1.1 Use Case 1

From a system perspective, the Identity Owner is establishing a relationship with the university first. The trust within the relationship is formed through an authentication process between both parties. This ensures that both parties engage in a trusted relationship where both identities are validated. After establishing the relationship, the students request through the system his or her student status at the university. The university checks in a back process whether or not the student is an active student. If this holds true, the university prepares the requested credential. This can be done via a database request to retrieve specific student information from the university backend. After receiving the information, the credential is passed on to the student. This is done through issuing a Verifiable Credential that structure has been defined inside the system. After the student has received the student credential, she stores it on her own device or an Identity Hub.

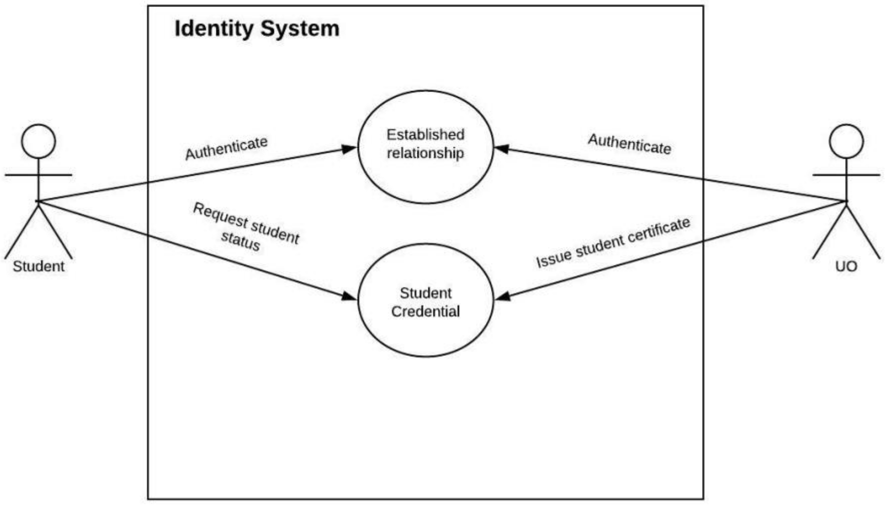


Figure 16. Use-case Diagram of Requesting Base Credentials.

5.3.2 Scenario 2: Requesting a Discounted Travel Credential

After the student has gone through the processes of gathering all necessary base credentials for applying for a Travel Credential, she has to approach her local Transportation Authority. The Transportation Authority asks the student to show all the proof in order to assess whether or not she is eligible for a discounted Travel Credential. Here the transportation office asks the student to present a bank account statement, general identity information like name, and the current student transcript. For each request, the student gives consent to the transportation authority to access the needed information. While presenting the documents the transportation authority validates the authenticity of the documents. If all documents are valid the student receives the Travel Credential which she can use all over Europe to travel via public transportation at a discounted price.

5.3.2.1 Use Case 2

The diagram for scenario 2 depicts the issuing process of a Travel Credential. Here the student and the Identity Provider in this case the transportation authority establish a trusted relationship. This is similar to the process described in use case diagram 1. After establishing the relationship and requesting the Travel Credential a presentational request is sent to the student. The student presents a set of claims that are needed for successfully issuing the Travel Credential. The Identity Provider then validates the correctness of the provided claims through cryptographic proofs. After successfully validating that the transportation authority is able to issue a new credential, the Travel Credential which has like any other Verifiable Credential a predefined Schema that allows for a standardized data structure. After receiving and validating the issued credential the student stores the credential for future proofs in her digital wallet or Identity Hub.

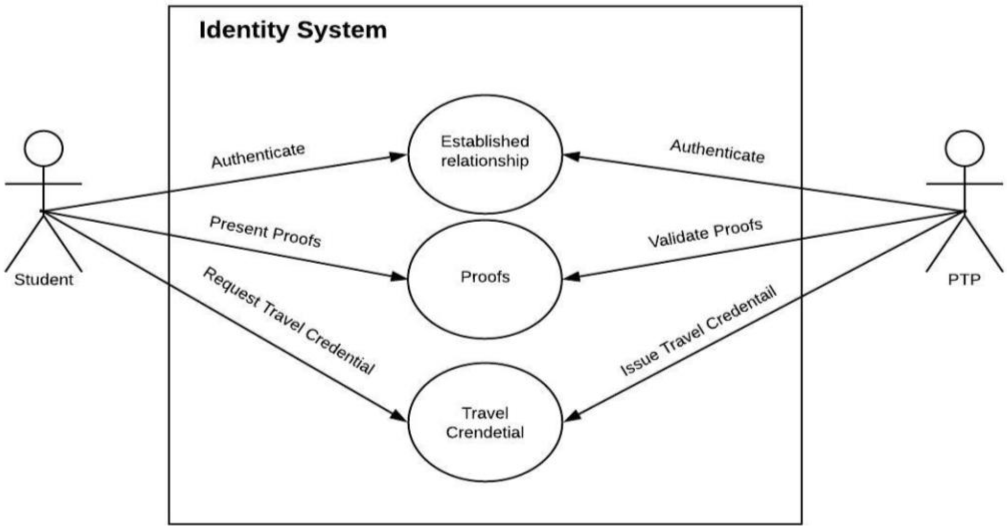


Figure 17. Use-case Diagram of requesting a discounted Travel Credential.

5.3.3 Scenario 3: Using the Travel Credential across Europe

After successfully issuing a Travel Credential the student wants to use the card for traveling on public transport in Europe. Here, for example, the student goes from one European country to another and uses the card to travel from A to B via the metro. When entering the underground system, the student passes some validator machines that give access to the underground. She uses her digital wallet and presents the Travel Credential to the validator machine. The machine checks the shown credential for validity. Following the verification, the validator calculates the student fees and passes those on to the digital wallet. This can be treated as a Transaction Credential like a receipt for a specific journey which entails price, location and time after the check-in and check-out. While this is happening, a payment request gets issued to the bank account which is referenced in the Travel Credential. After doing this the gates open and the student gets access to the underground.

5.3.3.1 Use Case 3

Use case diagram 3 depicts the process that is happening when the student uses her card. Again, like in previous use cases a trusted relationship needs to be established first. After establishing this relationship, the proof needs to be presented by the student through the system. After validating the proof by the Service Provider, in this case, another transportation authority gives access to the system. While giving access to the service the system treats the service offer as a credential offer. Here the passenger receives a Transaction Credential. This credential can be a receipt for a journey that can be used to prove to anyone else that a certain check-in or check-out has happened. Moreover, the system can engage with any off-chain solution to trigger additional events like a charging request. This can be settled because referenced proves in the overall Travel Credential.

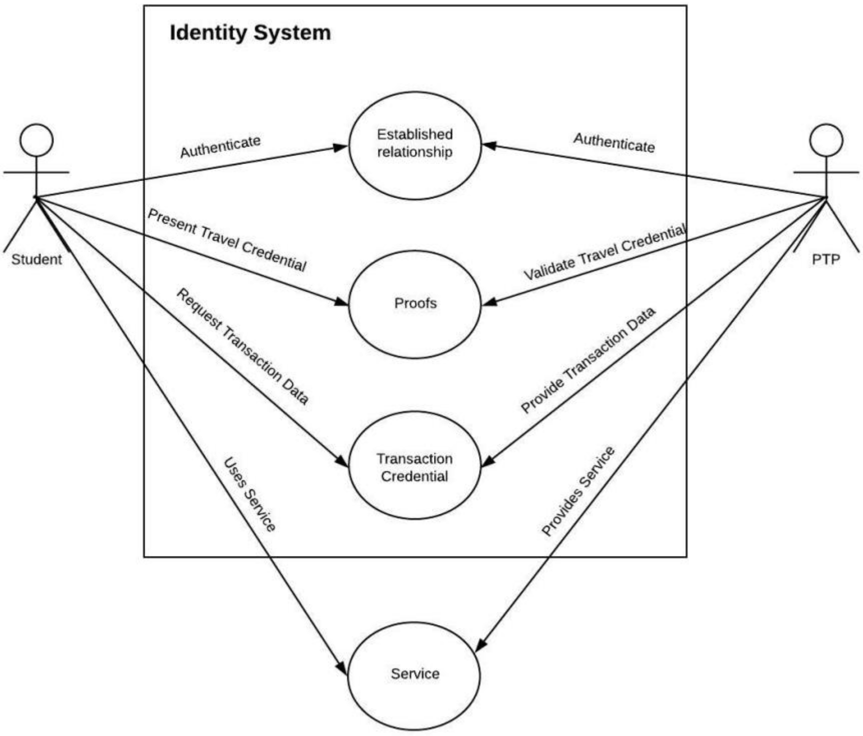


Figure 18. Use-case Diagram of using the Travel Credential across Europe.

5.4 Sequence Diagrams

After defining the most important use cases in the system a sequence diagram can be designed. A sequence diagram helps to showcase the different interaction flows between the different stakeholders in the system. It takes the defined use cases and puts it into the context of the overall system. Stakeholders and their specific interaction flows are mapped along the time they are happening through the entire process. The sequence diagram is time focused and visualises the different steps that are involved in order to build a usable prototype. The interactions between stakeholders can be split up between actions that are happening on a national and European wide level. For illustrational purposes, the diagram involves stakeholders from a domestic country system that are required to apply for a Travel Credential. Here a University Office, a Banking Provider and a Public Transport Provider are used to model a domestic system. In order to showcase the interoperability of the system, a transportation service is used to demonstrate the usage of the Travel Credential in two European countries. Thus, showcasing the overall functionality of the system in a European wide context. Any instance in the system such as a University Office could be replaced through a stakeholder within the same institutional context.

5.4.1 Sequence Diagram 1: Base Credentials

The sequence diagram below describes the process the user needs to go through to acquire all the necessary base credentials in order to satisfy the requirements of applying for a Travel Credential. The system assumes that the user first acquires a government ID. This happens through the attestation by the government which is requested by the user. Following a successful attestation by the government, the user receives a government credential. The same process is used to request attestation for a university transcript at a University Office as well as a bank account statement by a Banking Provider. Every time a new credential gets issued the user stores the credential in a personal wallet. Those credentials are referred to as the base credentials.

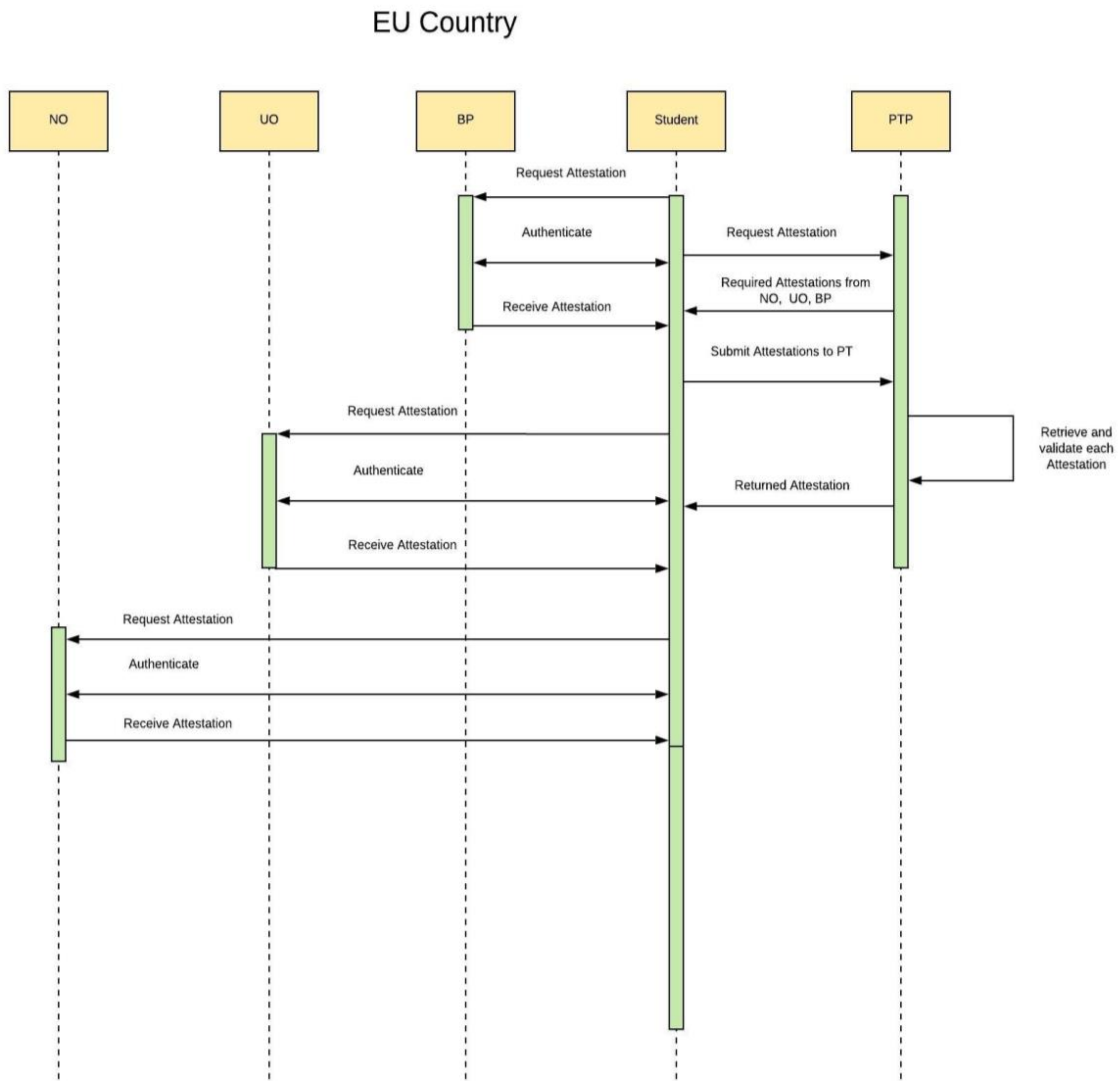


Figure 19. Sequence Diagram of Base Credentials.

5.4.2 Sequence Diagram 2: Using the Travel Credential to Travel

After receiving all the necessary credentials for applying for a Travel Credential the users continue with the process of applying for it at the local transportation authority. Here the Public Transport Provider requests the user to present the required attestations in order to issue a Travel Credential. After requesting the attestation, the user presents all the three acquired attestations. Those attestations are processed by the Public Transport Provider and checked for their validity through cryptographic proofs which are encoded inside the credential. After the proofs have been validated and confirmed, a Travel Credential is issued. It gets saved by the user inside their personal wallet. In the last step the user uses the Travel Credential to travel in a European country. Here the user presents the Travel Credential to the domestic countries Public Transport Provider which validates the Travel Credential and allows the user to use the service. This is achieved through cryptographic proofs and the general assumption that transportation Service Providers in Europe trust each other. Therefore, the validity and usage of the Travel Credential across borders can be ensured.

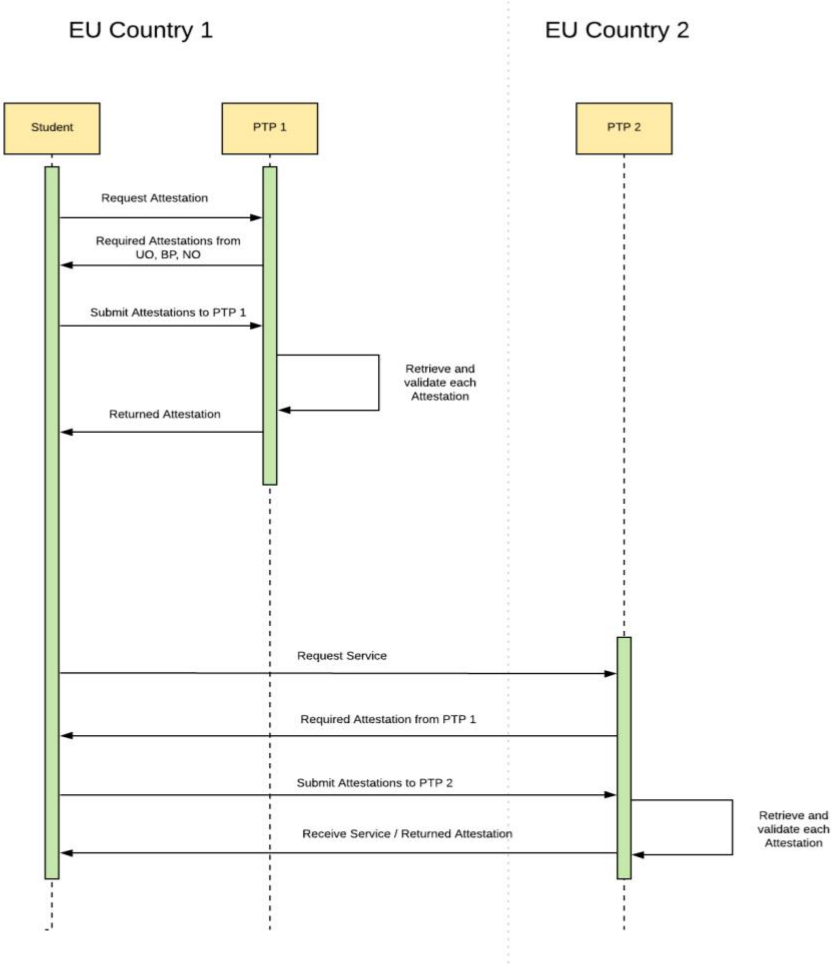

EU Country 1
EU Country 2
Student
PTP 1
PTP 2
Request Attestation
Required Attestations from UO, BP, NO
Submit Attestations to PTP 1
Retrieve and validate each Attestation
Returned Attestation
Request Service
Required Attestation from PTP 1
Submit Attestations to PTP 2
Retrieve and validate each Attestation
Receive Service / Returned Attestation

Figure 20. Sequence Diagram of Using the Travel Card to travel.

Chapter 6

6. System Implementation of the Artefact

After outlining the overall system design of a decentralized identity application, it is now possible to go into more detail about a possible identity solution for public transportation in Europe. Here a specially designed travel credential for the public transportation sector is defined and issued to passengers. The *Euro Mobility Card (EMC)* allows passengers to travel across different public transportation systems in Europe. The card replicates a Verifiable Credential that can be issued by any PTA in Europe. The passenger can prove by controlling this card that he or she is eligible for certain discounts or has sufficient bank liquidity to pay for transportation on the go. The EMC is a digital card that users can access from their mobile device and show on any credential presentation request. Those can be requested programmatically by validator machines on system entry or by manual request through ticket inspectors while traveling. However, the specific requirements for validation in the real-world are out of the scope of this research thus will only be described for illustrational purposes.

6.1 Underlying Technology of the Artefact

According to the analysis of Blockchain-based systems, Sovrin aligns the most with the fundamental principles of SSI. This makes it the perfect candidate for the preferred system choice. The underlying system has been forked and modified from an official learning material repository of Hyperledger (Hyperledger, 2019). Thus, the developed artefact will leverage Hyperledger Indy, that Sovrin ulitizes for its instance, as the underlying system to develop a prototype. Indy is classified as a public permissioned Blockchain and on that note, the proposed system will share the same access and validation permissions. Indy is a special purpose Blockchain implemented specifically for decentralized identity that enables certainty about who is communicating with whom in a digital transaction. Therefore, Indy allows you to create digital identities that would be rooted in the Blockchain or any other form of the distributed ledger.

6.1.1 Key Features of Hyperledger Indy

- **Self-Sovereignty** - Indy software embodies the concepts of SSI by creating artefacts that include cryptographic proof of existence, public keys and many more. These are special artefacts that the platform of Indy will store for the users who are the only ones that can change or remove their identity from it.

- **Privacy** - Indy is designed in a way to preserve users' privacy. According to documentation of Indy, it adheres to the concept of "privacy by design" that is also now a standard by privacy experts worldwide and outlines how Indy preserves those privacy settings.

- **Verifiable Credentials** - Indy supports the emerging W3C standard of Verifiable Credentials that enables a trusted way to provide identity attributes about users. VCs take the model of issuing paper credentials in the real world such as passports, university degrees and put it online in a trusted manner.

- **Purpose-built** - Indy is built for just one purpose and it is all about identity on the internet. Indy's platform can be used in any use case about decentralized identities and allows users to prove to others who they are while being certain who others are.

- **Correlation-resistant** - Indy resolves the problem of correlation which is made possible due to the common identifiers such as emails, used daily on so many different websites. Identity is completely correlation-resistant which means that there is no way that Indy's platform will connect two identities or have two similar identifications on the ledger.

- **Decentralized Identifiers (DIDs)** - Indy supports the emerging W3C standard for DIDs and according to documentation, DIDs are globally resolvable identifiers, unique and independent of any central authority.

- **Peer-to-peer Connection** - Indy allows for a peer-to-peer connection and uses pairwise DIDs to establish connections between two identities in order to communicate in a secure way.

- **Zero-Knowledge Proofs (ZKPs)** - Indy provides advanced features in the proving of claims and utilizes the cryptographic method of zero-knowledge proofs that allows to selectively disclose some data elements from credentials to be provided in a proof.

6.1.2 Typical Components of Hyperledger Indy

- **Cryptography** - Any identity solution that leverages Indy's platform, will eventually use Zero-Knowledge Proof, a cryptographic protocol that ensures users can prove to any other users that they know certain information without revealing the underlying data. ZKP support in Indy is one of the most useful components which is utilized for proof requests of credentials.

- **Nodes** - As with all Blockchains, Indy uses a consensus algorithm called Plenum to make a decision on the contents of the next block to be added to the chain. Plenum is an implementation of the Redundant Byzantine Fault Tolerance (RBFT) algorithm and is quite efficient as opposed to other algorithms. Indy, also, implements a novel deployment of "Stewards", trusted nodes within the network that have been granted permission to participate in the validation process. In general, the validation process requires enough Stewards to be robust in case of a fault, but not too many as this may compromise performance in reaching consensus.

- **Agents and Wallets** - According to Hyperledger Indy documentation, an "Agent" refers to the software that interacts with other entities, e.g. via DIDs, and a "wallet" to store data for DIDs and all the related information like private keys. Indy Agents can come in many varieties, for instance, a user might have installed a mobile Agent app on a smartphone, while an organization might have installed an enterprise Agent running on the cloud. All of the Agents have a secure Wallet for identity data storage.

- **Ledger** - At the heart of Indy is an immutable public ledger in which Nodes, the permissioned instances, accept and process read and write requests from Indy clients or the so-called Agents. The ledger is backed up by Merkle trees and all the nodes hold a copy of the ledger.

- **State and storage** - The state of the Indy network is maintained by nodes for each ledger using a Merkle Patricia Trie which provides a cryptographically authenticated data structure to store all key-value bindings. On the other hand, the storage in the system is implemented with RocksDB as a key-value database where the key represents the sequence number and the value represents the transaction.

6.2 Development of the Artefact

The following table will describe the issuing and usage of the EMC in more detail and explain the different steps from requesting the card over to using it in public transportation. For a better understanding, each stakeholder is mapped to a real-world entity in the Danish system. Thus table 14 describes the stakeholders and roles from the identity space that are used throughout the system.

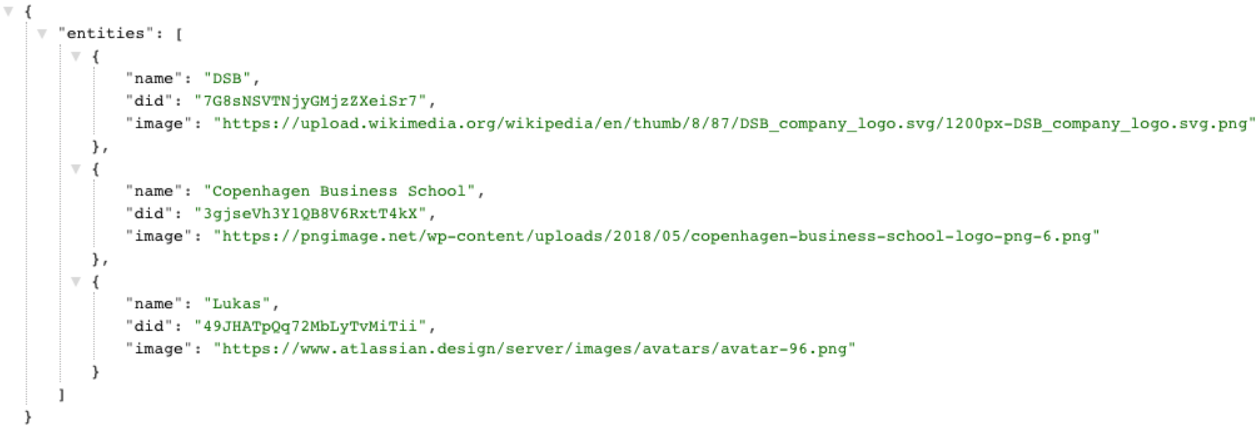
Denmark	**Overall role**	**Identity Management Role**
Danske Bank	Bank	Identity Provider / Issuer
Danske Statsbaner	Transportation Authority	Service Provider / Identity Provider / Issuer
Copenhagen Business School	University	Identity Provider / Issuer
Danish Government / Nem ID	Government	Identity Provider / Issuer
Lukas	Passenger	Identity Holder

Table 14. Role Definitions of Actors in the System.

In a decentralized system, the roles of certain stakeholders can overlap compared to any traditional Identity Management System. In the underlying solutions, the role of an Identity Provider is overlapping with the roles of a Service Provider as well as an Issuer. This is due to the fact that credentials are stored by nature with the user and any third-party is excluded as the identity providing authority. However, since there is a one-to-one validation process happening between Verifier and Issuer the Issuer gets a similar role to an Identity Provider in a centralized system. Without the cryptographic proof encoded into the credential by the Issuer, the validation process would not be possible without an intermediary. Thus, the stakeholders within the system roles are not mutually exclusive allowing certain stakeholders to act within multiple roles.

6.2.1 Registration of DID

Each stakeholder in the system receives an endpoint DID which can be used to identify a stakeholder within the system. The DIDs are registered on the ledger and can be read by anyone. The prototype uses an external JSON API to connect users to their DIDs. After requesting an Agent in the browser, the DID and additional meta information are being written to an hosted JSON file. Here a new object with username, avatar, and DID is getting stored. This functionality has been implemented in order to allow for a simplified user experience in the prototype. Thus, users can open a modal to see all possible relationships within the system. Simply by clicking a user avatar and name the user can send a relationship request. In a more advanced system, QR codes or other discovery methods via HTTPS could be sent in order to avoid connecting through non-human readable identifiers like DIDs.

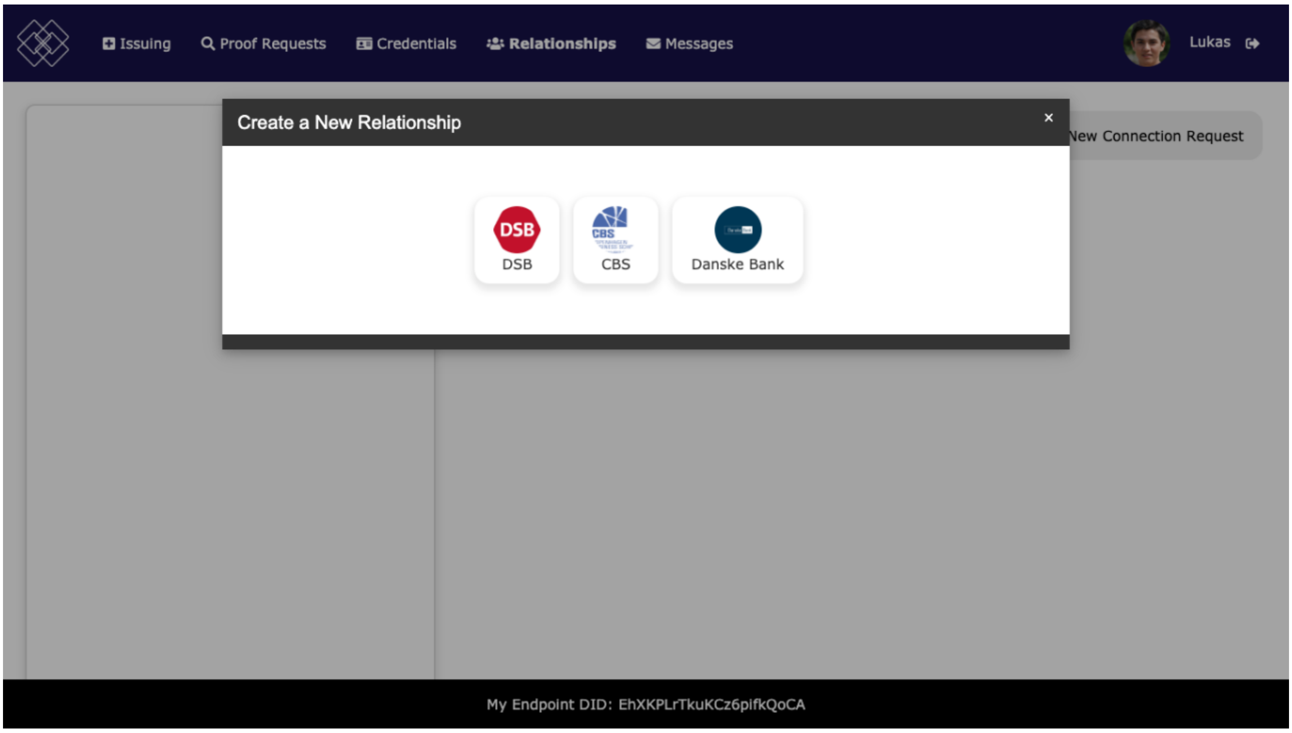


```
{
  "entities": [
    {
      "name": "DSB",
      "did": "7G8sNSVTNjyGMjzZXeiSr7",
      "image": "https://upload.wikimedia.org/wikipedia/en/thumb/8/87/DSB_company_logo.svg/1200px-DSB_company_logo.svg.png"
    },
    {
      "name": "Copenhagen Business School",
      "did": "3gjseVh3Y1QB8V6RxtT4kX",
      "image": "https://pngimage.net/wp-content/uploads/2018/05/copenhagen-business-school-logo-png-6.png"
    },
    {
      "name": "Lukas",
      "did": "49JHATpQq72MbLyTvMiTii",
      "image": "https://www.atlassian.design/server/images/avatars/avatar-96.png"
    }
  ]
}
```

Figure 21. External JSON for Storing Metadata of Agents.

6.2.2 Relationships between Stakeholders

In order to create a secure communication channel between the two stakeholders, a relationship needs to be formed first. For this reason, each party creates and exchanges a unique pairwise DID with its counterpart. The request is demonstrated through an invitation request. After the invitation has been accepted both pairwise DIDs and verkeys are getting exchanged. The created pairwise DIDs are unique to the relationship and cannot be reused across other relationships. Depending on the design choice the DIDs are registered either on the ledger or stored inside the wallet. In the existing prototype, the pairwise DIDs are registered on the ledger. Thus, Lukas forms relationships with all parties he wants to exchange information in a secure way.



My Endpoint DID: EhXKPLrTkuKCz6pifkQoCA

Figure 22. Sending Relationship Requests.

In the prototype, Lukas establishes relationships with the Government, Danske Bank, CBS, and DSB. After creating all necessary relationships Lukas has four pairwise DIDs which are unique to each relationship. The creation of isolated DIDs for each relationship helps to defeat the correlation between Lukas and each stakeholder. Thus, this increases the level of privacy within the system and helps to create a multidimensional identity that is able to control multiple DIDs.

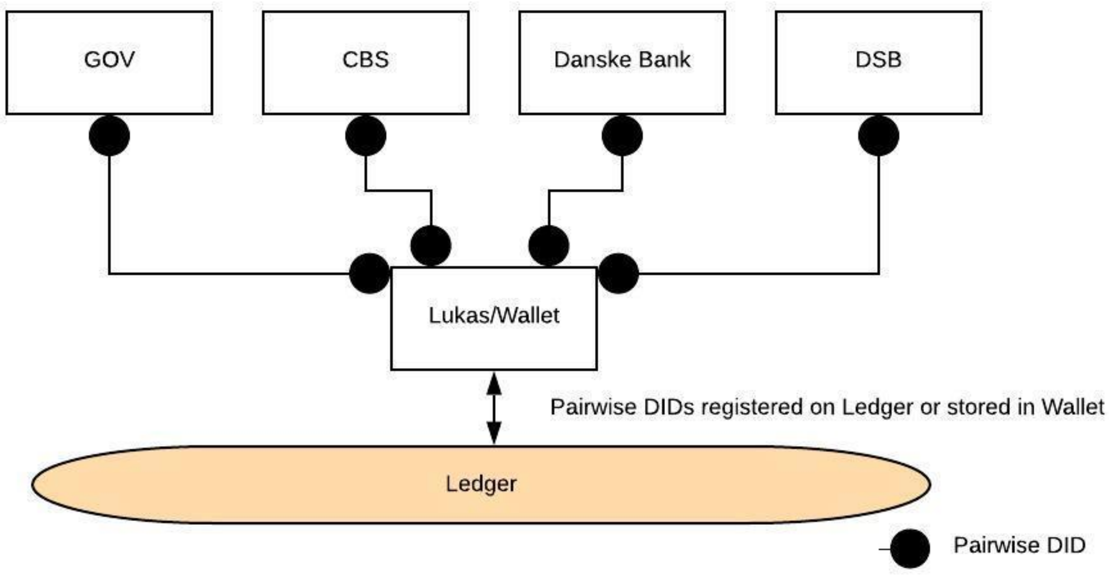
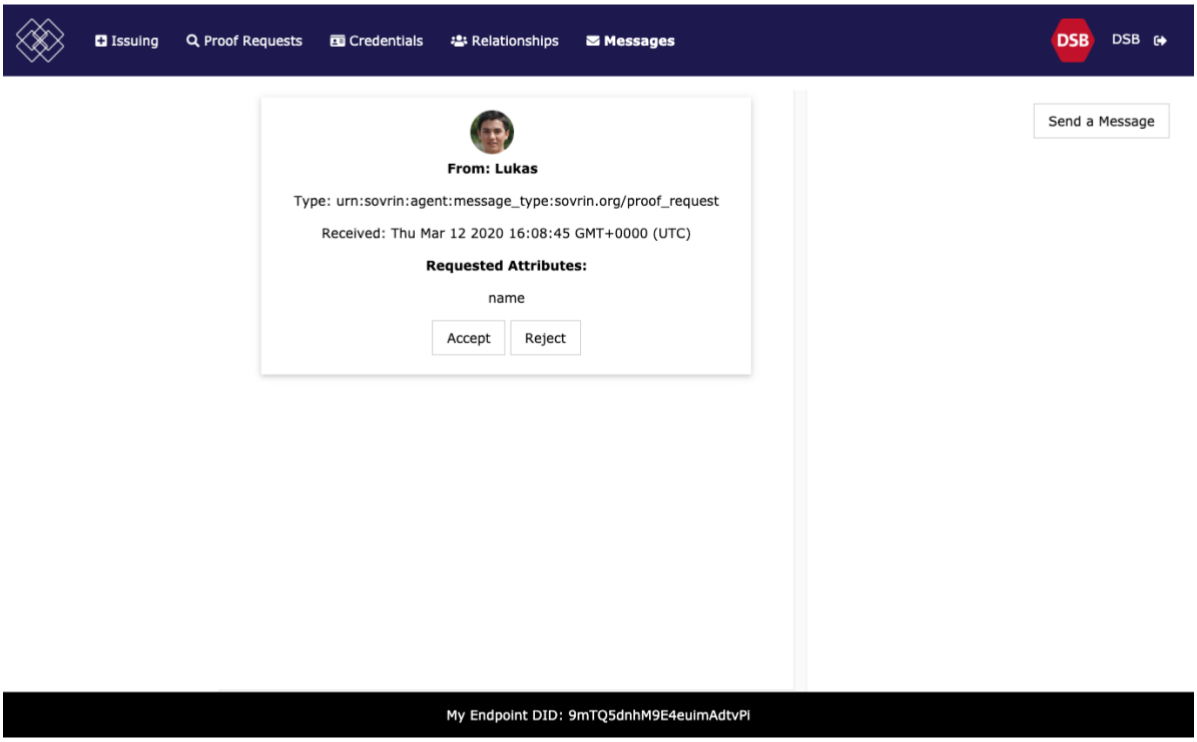


Figure 23. Forming Relationships and Creating Pairwise DIDs.

In the prototype, relationships can be formed through opening the relationship managing modal. Here all possible relationships are displayed, and the user can simply click on an avatar to request a relationship. The avatars are just an abstraction of the DIDs of each stakeholder. In order to discover all possible Agents in the network the JSON file from the registration process is retrieved. Here all registered stakeholders are listed and mapped out as possible users for engagement. After selecting a user by the avatar and name the prototype system uses the specific stakeholders DID to send a relationship request in the background. After sending an invitation both parties are requested to accept the invite in order to establish a secure connection.



From: Lukas

Type: urn:sovrin:agent:message_type:sovrin.org/proof_request

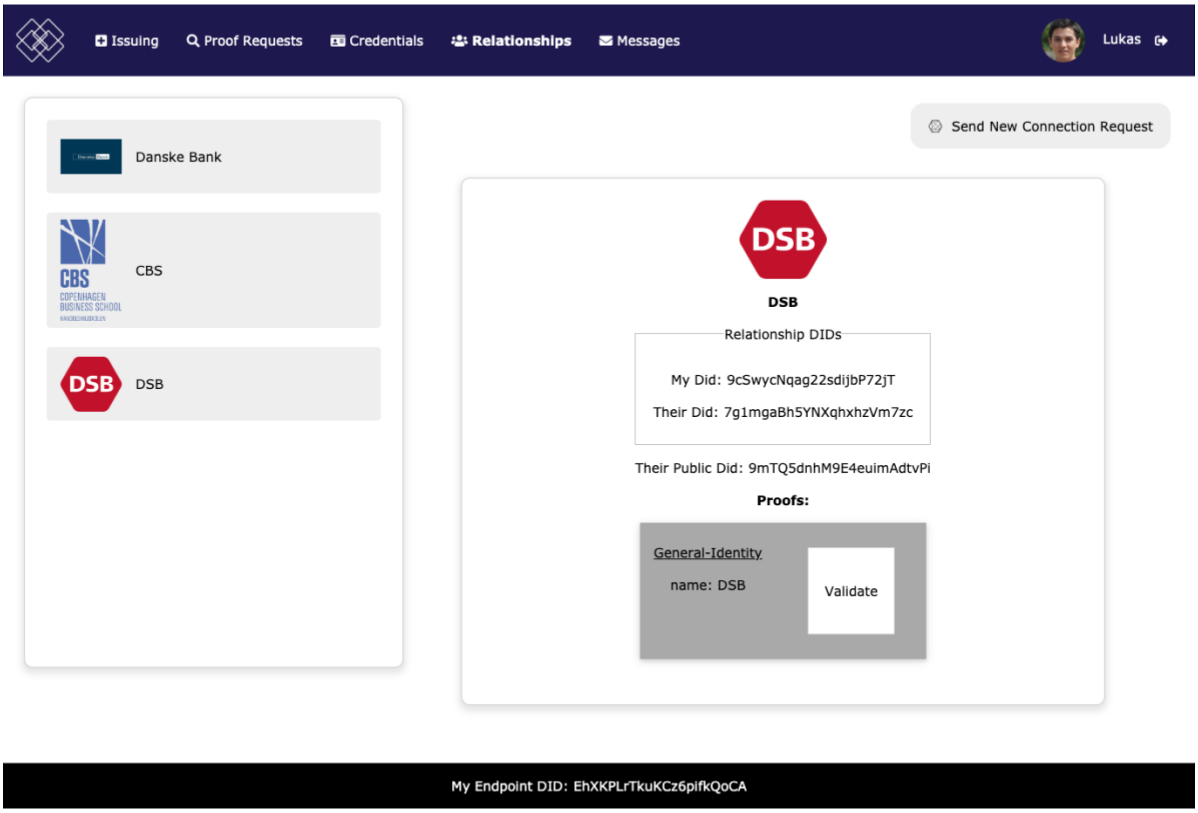
Received: Thu Mar 12 2020 16:08:45 GMT+0000 (UTC)

Requested Attributes:

name

My Endpoint DID: 9mTQ5dnhM9E4euimAdtvPi

Figure 24. Accepting a Relationship Request.



DSB

Relationship DIDs

My Did: 9cSwycNqag22sdijbP72jT

Their Did: 7g1mgaBh5YNXqhxhzVm7zc

Their Public Did: 9mTQ5dnhM9E4euimAdtvPi

Proofs:

General-Identity

name: DSB

My Endpoint DID: EhXKPLrTkuKCz6pifkQoCA

Figure 25. Overview of Accepted Relationships.

6.2.3 Schema Designs & Definition

Schema design is an important aspect of the prototype. The current prototype does not offer the possibility to discover different schemas on the ledger. For this reason, schemas are either hardcoded into the system or are getting saved on creation in an external JSON file. This file can be fetched by any Agent in order to discover schemas. This allows for a system-wide schema discovery and enables the different stakeholders to create their own schema definitions based on existing schema designs. This is particularly useful when it comes to the issuing of EMC from different transportation authorities in Europe. The prototype assumes that there is a European transportation committee that has defined the overall structure of the EMC schema. Thus, any national transportation company can create their own schema definition based on the EMC schema. In an ideal system, Schemas could be discovered on the ledger by anyone. Moreover, each industry would need to agree on a specific Schema design like a bank account schema. Thus, the prototype assumes that each industry is represented by a European Agency which is overseen by the EU. The industry-specific agencies are responsible for defining schemas that can be used by agents to establish credential definitions. The following describes the Schemas and their related institution in the prototype.

Agent	**Defining institution**	**Schema Name**
Danske Bank	European Banking Authority	Banking Schema
Danske Statsbaner (DSB)	European Transportation Authority	Euro Mobility Card Schema
Copenhagen Business School	European Education Authority	University Status Schema
Danish Government	Government	Identity Card Schema

Table 15. Schemas and Related Institutions.

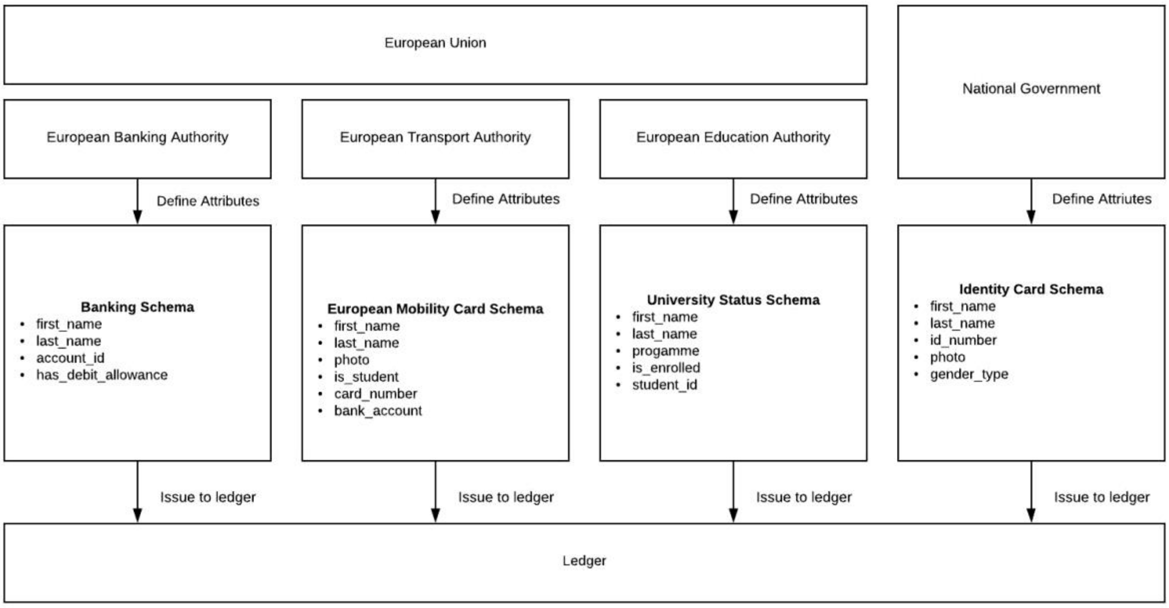


Figure 26. Schema Definition.

Figure 25 describes each schema and the attributes that are defined in the prototype. Each schema has its overarching authority that agrees on attributes that are defined in order to create a standard. The standard is then specified in the schema design and issued to the ledger. Thus, anyone can discover the schemas and use them for creating their own credential definitions. The prototype assumes commonly used attributes from each industry. This allows to better visualize the use case scenario of issuing an sEMC card to a Danish user. Moreover, the prototype assumes that a national government is responsible for its own identity card schema which can vary on a European level. However, it assumes that banking, transportation, and education can be overseen on a European level.

In order for an authority to issue a schema definition to the ledger, the prototype allows adding attributes through an input form. Here the user can define a schema name and different attributes via a JSON array. After submitting the form, a Schema Definition is issued on the ledger with each element in the array representing an attribute in the Schema. For example, the EMC Schema could be added by adding the following array to the input [“first_name”, “last_name”, “photo”, “is_student”, “card_number”, “bank_account”] and submitting the form. At the current stage the prototype issues schema definitions on the ledger. However, across server instances, the discovery of Schemas on the ledger is not implemented. Thus, to demonstrate the functionality and have consistent Schema designs across agents the prototype uses hardcoded values that are set by default for each instance. This allows each agent to retrieve all possible Schema Definitions across the system.

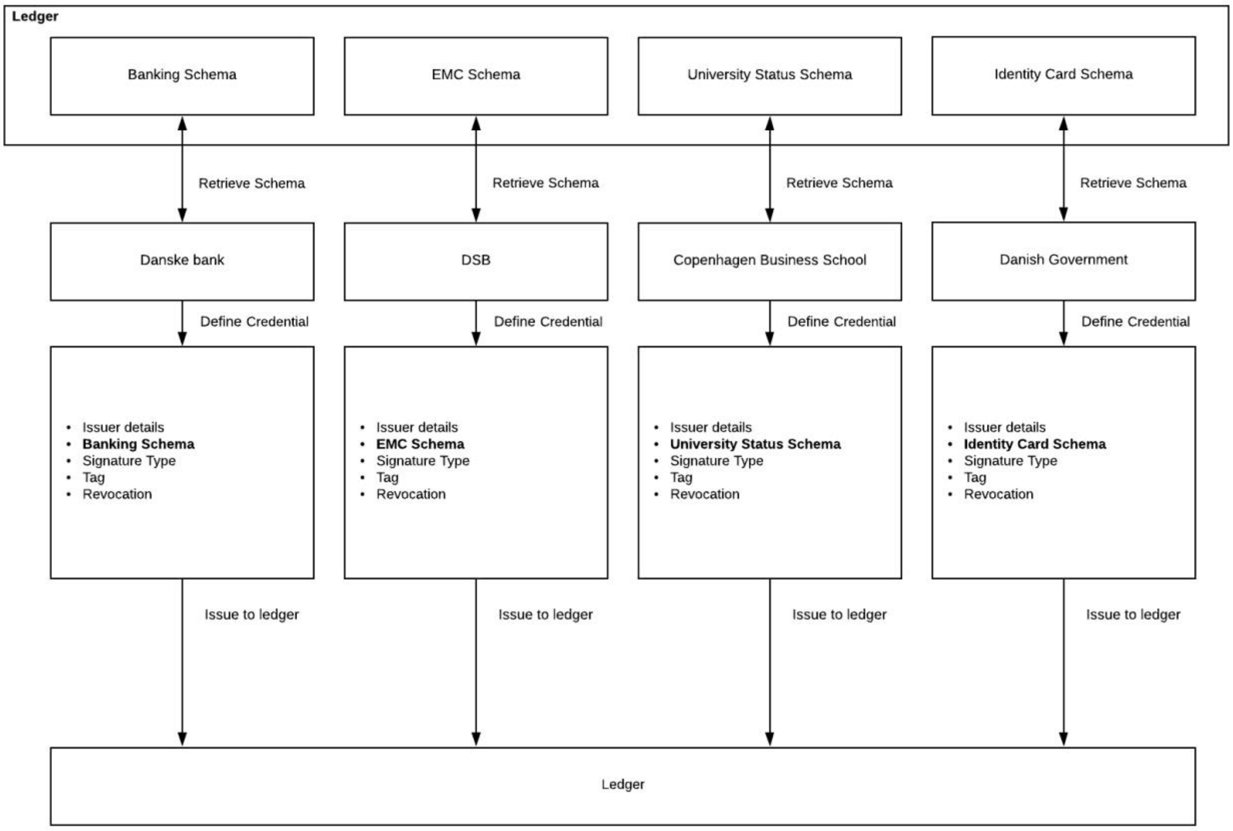
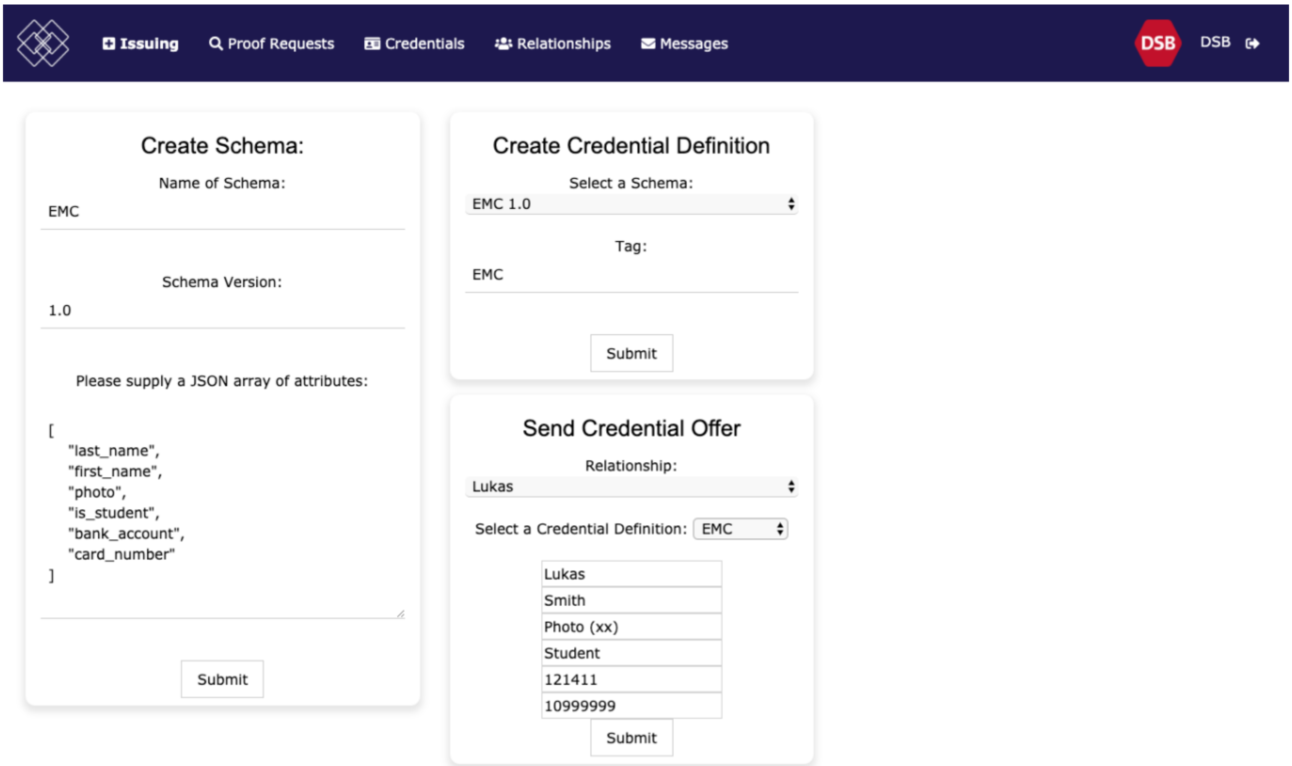


Figure 27. Credential Definition.

Figure 26 describes the process of an Issuer to create their own Credential Definitions. Here each Issuer is able to retrieve a specific Schema in order to create their own Credential Definition. As previously mentioned, the different Schema Definitions are hardcoded values which in a fully functional system would be retrieved from the ledger instead of the local system. After retrieving the Schema Definition, the Credential Issuer (Danske Bank, DSB, CBS, and Danish Government) is able to create their own issuer-specific Credential Definition. The Credential Definition is an instance of the selected Schemas data structure and the issuer-specific attributes like Issuer details, Signature Type, Tag, and Revocation. The definition is dynamically generated by the system after selecting the Schema Definition by the Issuer and adding a tag. The Credential Definition is then registered on the ledger and can be reused by the Issuer. Moreover, the Credential Definition is used by Verifiers to validate the integrity of the received proofs.

6.2.4 Issuing Verifiable Credentials

After an Issuer has created their issuer-specific Credential Definition the Issuer is able to create attestations for users in the prototype. In order to do this, the issuer needs to select their credential definition for the credential they would like to issue. The system then converts each attribute defined in the credential definition into an input field. This input field can then be used by the issuer to manually enter credentials to each attribute. For example, Danske Bank would add into the "account_id" input field the bank account number of the user that has requested an attestation of their bank account. In a more advanced prototype, the system could autocomplete these values by requesting their internal databases. Thus, the issuer would not need to manually populate the input fields with data. As soon as all the required information has been provided by the issuing party and a secure relationship has been established between the issuer and the identity holder the credential offer can be initiated. After sending the invitation to accept the credential the identity holder needs to verify the key-value pairs for each attribute. If everything is correct the identity holder accepts the credential and stores it locally in the prototype. The identity holder can now manage their credentials in their wallet. This means any other verifier can request proof of the credential or specific claims that can be made from it. For example, the person has a bank account. In the specific case of requesting an sEMC. The identity holder would need to request credentials from three different sources, the Danske Bank, the Government, and the CBS. This would add all three necessary credentials to their wallet.



Create Schema:

Name of Schema: EMC

Schema Version: 1.0

Please supply a JSON array of attributes:

```
[
    "last_name",
    "first_name",
    "photo",
    "is_student",
    "bank_account",
    "card_number"
]
```

Create Credential Definition

Select a Schema: EMC 1.0

Tag: EMC

Send Credential Offer

Relationship: Lukas

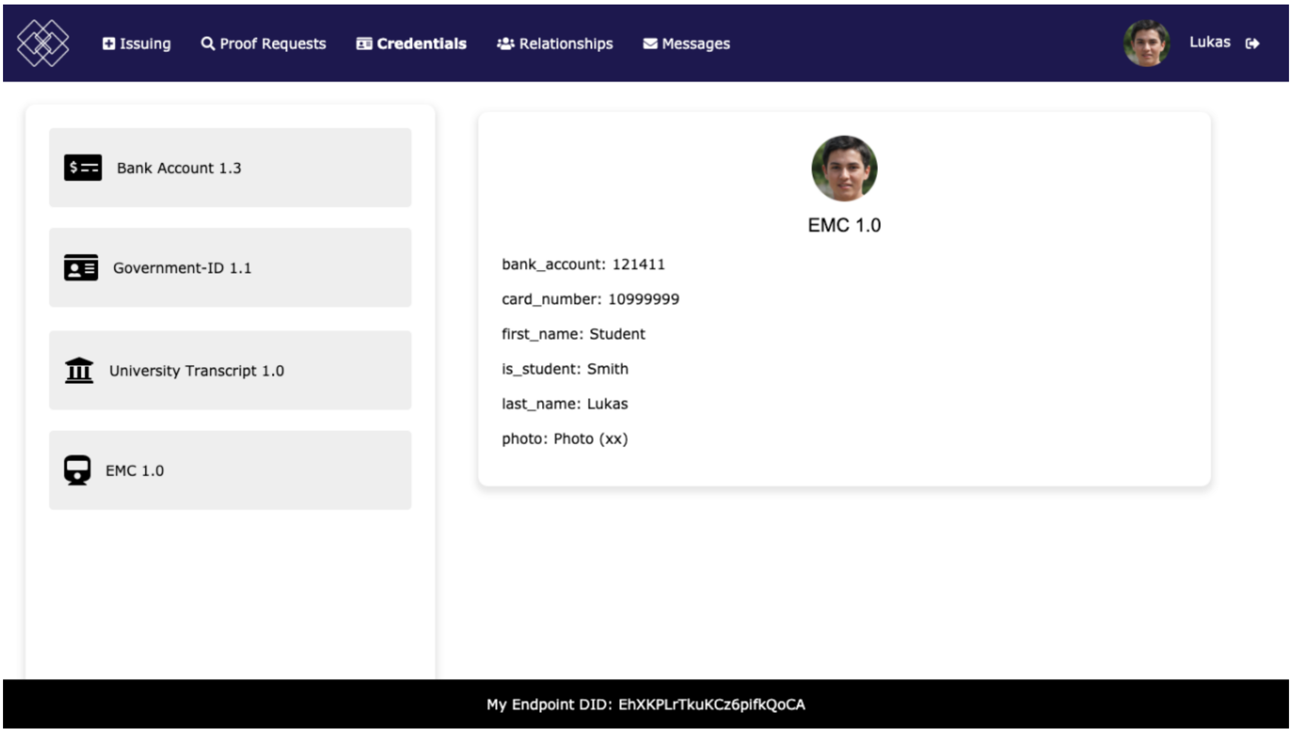
Select a Credential Definition: EMC

Lukas
Smith
Photo (xx)
Student
121411
10999999

Figure 28. Issuing EMC Based on Predefined Schema.

6.2.5 Proof Requests of Credentials

The prototype offers the ability to request proofs from an identity holder. For this, the identity holder and the verifier need to establish a secure relationship first. This is used to exchange credentials and issue new credentials that require the validation of underlying credentials. In the case of the prototype, DSB would be allowed to issue a Euro Mobility Card based on the requirements that have been set by the European Transportation Authority. The prototype assumes that the sEMC can only be issued when a user can present credentials about their bank account, their university status, and their general identity. Thus, the user presents the acquired attestations from the previous process. For this, the user uses the credential definitions from each issuer and posts them via a form to the verifying party. This retrieves the claims from the identity holders' wallet and enables the verifier to verify the origin and integrity of the provided data. This is possible through advanced cryptography and cross-checking of verification keys that are registered on the ledger. In order to expose claims to the verifier the user first needs to accept the attributes that are requested by the verifier. After all the verification and data exchange has been handled the verifier becomes an issuer and issues the EMC card. In the prototyped use case, the identity holder receives an sEMC card. This is due to the availability of the university status credentials which satisfies the requirement for receiving a student discount on travel.



EMC 1.0

bank_account: 121411

card_number: 10999999

first_name: Student

is_student: Smith

last_name: Lukas

photo: Photo (xx)

My Endpoint DID: EhXKPLrTkuKCz6pifkQoCA

Figure 29. EMC Credential for a Passenger.

Chapter 7

7. Discussion

This chapter presents some topics of discussion in areas that have not been investigated thoroughly in the paper due to time and scope constraints. These areas expand from the need for decentralization of user identity in public transportation and what this entails as well as the feasibility pillars for the proposed system.

7.1 Why Decentralization of User Identity in Public Transportation

The current public transportation system across Europe is highly scattered and needs to be further developed into a coherent transportation solution. In order to achieve the goal of a Single European Transport Market by 2050 the way user accounts are managed needs to adhere to a common system. This means that systems that are currently implemented have to settle on a common implementation in order to create an interoperable solution that can be used by any existing or upcoming service. The implementation of a decentralized Identity Management solution can be the turning point for a new paradigm in public transportation. The new system is focused on user control, privacy, and interoperability. The current system is characterized by many different standards and a high dependency on different vendors. This leads to vendor lock-in where private software solutions lock services into exclusive software architectures and proprietary data formats. This causes high switching costs within the industry and isolation between those systems. In order to arrive at a combined public transportation solution data standards and system architectures need to be commonly developed and applied.

However, a one size fits all identity solution could cause the centralization of power and lead to a monopolized access management landscape in the public transportation sector. Public transportation is a public good and the underlying identity architecture should not be managed similarly to structures that have been seen by social media services and other data harvesters. Moreover, an SSI based solution does not only enforce privacy by design and user control but can also improve the efficiency of the overall system. The reuse of existing information leads to faster service integrations, higher efficiency, better interoperability and better user experience of the entire public transport sector. Moreover, the focus on privacy improves the overall security as well as reduces costs for public transportation agencies when it comes to storing and managing private data. The improved user experience that adheres to an ever-growing need for privacy will incentivize more and more users over time to shift their mobility habits. In order to profit from the benefits of decentralized Identity Management in transportation, the European Union has to not only come together as a political and fiscal union but also as a transport union. Thus, trust and a legal framework have to be developed that involves all managing and operational aspects of a decentralized identity network as well as engages all public transport stakeholders.

7.1.1 SSI Framework Alignment

After proposing a first prototype that involves the issuing of a Euro Mobility Card the solution can be assessed according to two core SSI principles namely user control and portability. The EMC allows for a high level of user control. This is because of the general nature of it being a Verifiable Credential that is in full control of the user and stored in their personal wallet. This means that the EMC is part of their identity that will consist of many different credentials like an identity card. Thus, issuing a card that is only dependent on its direct verification by the issuer and not reliant on the goodwill of any Identity Provider. This will give the user full control over this part of their identity.

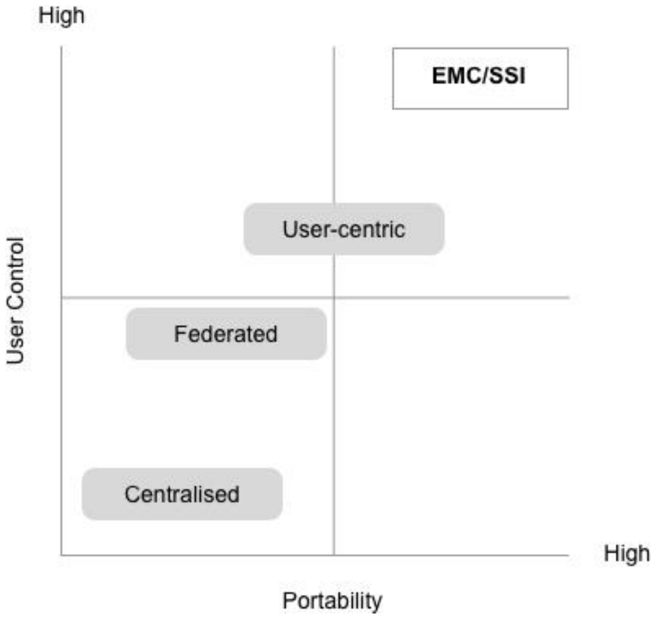


Figure 30. Types of identity and EMC.

In regards to portability the prototype also scores high. This is especially because of the underlying technology. The prototype follows the guidelines of the W3C working group on decentralized identity. It adheres to the sets of standards that have been defined by the group. This allows an application-agnostic utilization of credentials. Any similar application can be used to retrieve credentials. The user can easily restore any of the acquired credentials through the use of their private key. Simply providing the recovery seed of their wallet gives them access to their credentials. Thus, they are free to choose the application they would like to use. However, when it comes to certain types of features some wallets might be superior. Nevertheless, the passenger can reuse any of their DIDs in any different SSI application that follows the W3C standards.

7.2 Positioning in the Current Ticketing Systems

Section 4.2 defined a set of attributes which can be used to compare the developed prototype against existing Account Based Ticketing solutions. The defined framework opens up the discussion of how the SSI based prototype aligns with other evolving technologies in this sector. Looking at table 16 Part 1, it highlights that the developed prototype aligns with most of the attributes that modern Account Based Ticketing solutions in public transportation fulfil. The EMC removes any logic like fare price calculating to a backed process that is controlled by each PTA. Thus, allowing for a separation of concerns of different system activities. Moreover, it is highly reusable since it acts as an identifier in an overall ecosystem. The EMC is not bound to any specific system; it simply acts as an identifier that can interact with any other system that follows W3C standards. Thus, it is interoperable with other systems outside the ecosystem of one transportation company. This compares to the EMV-chip standard that has been used in London's ABT solution which has been previously mentioned in section 4.1.3. However, looking at the comparison of the two systems based on the system-centric attribute highlights an important difference. The EMC is not built around a specific system with a central database to manage user accounts. It is based on a decentralized system where every user is in full control of their own data and credentials. Thus, there is no particular need besides business interest to manage user accounts by the transportation company. This clearly acts in favor of higher user control, portability and privacy for the end user. Furthermore, it can add a new dimension in offering the potential for monetisation. This possibility, from both the user’s and business perspective, acts as an incentive for them to join the network since they could benefit from sharing or buying data which can be used for producing intelligent insights for marketing purposes (Faber et al., 2019). Moreover, it might even suggest introducing new high-level attributes to the framework for evaluating ticketing systems namely; user-control, portability and privacy (Table 16, Part 2).

	Euro Mobility Card	**Account Based Ticketing**
Part 1		
System-centric	✗	✓
Separation of concerns	✓	✓
Reusable	✓	✓
Interoperable	✓	✓
Restorable	✓	✓
Part 2		
Privacy by design	✓	✗
User Control	✓	✗
Portability	✓	✗

Table 16. EMC and its Extension of Account Based Ticketing.

7.3 Possible Feasibility Assessment

In order to examine the validity of the proposed system, a feasibility study needs to be conducted to give the authors and other stakeholders a clear picture of the proposed system. This could help uncover new ideas that potentially could change the scope of the project. It is essential to conduct this before diving into the actual implementation of a full-fledged system.

To assess the usability of the system, a high-fidelity prototyping method could be employed in order to offer an advanced user interface that looks similar to the final solution. This could be developed in conjunction with the user scenarios and allows stakeholders to understand the concept of the proposed system. While stakeholders interact with the prototype, other methods can be employed to gather data that is useful to determine what needs improvement. The think-aloud protocol is a qualitative method used to capture the verbalized thoughts of the participants that emerge when a certain task is being completed. This aims to elicit information about the performed tasks that should reflect the thought process of the participants that will give insight on how the prototype is used.

To assess the practicalities of the proposed system an assessment on the process of validating the technical capabilities could be conducted. This would shed light on the technological challenges that the integration of the system into existing ticketing systems brings in order to be fully operational. The developed prototype constitutes a small subset solution of the identified problem which entails a lot of technical efforts on the side of identity and Service Providers. In the context of public transportation, this would require advancements in upgrading the current ticketing system and terminals to support the final solution. Specifically, the identities provided to passengers such as the Euro Mobility Card need to be presented in a form that is recognizable by the ticketing terminals in order to check-in and check-out a journey in a seamless manner. Additionally, the infrastructure needs to be upgraded in such a way that the inspection devices can successfully read the presented mobility card. Therefore, a more exhaustive technical feasibility is required in order to assess the full implementation of an Identity Management System that facilitates cross-border travel at a European level.

Another pillar of feasibility for this project is the operational feasibility. This involves the legal aspects of the system as well as the organizational conflicts and policies that the system entails. Considering the fact that the underlying technology of the proposed system does not rely on a central authority this means that a governance trust framework needs to be in place in order to support the use case of Identity Management for the public transport sector. In accordance to how Sovrin operates with their Stewards, a governmental body should be formed that is accountable for governing and maintaining the network. This would help to mitigate any concerns regarding the credibility of the system as well as enable individuals and Service Providers to join the network. Since relying parties cannot trust anyone to assert claims, the governing bodies should be present and have their own identity on the network through which they implement the functionality of issuing Verifiable Credentials. This means that a number of bodies needs to be defined via a consortium by European Commission such as one for transportation, one for banking, another one for educational institutions, etc. These bodies can act as representatives of their respective entities which are able to assert claims. Consequently, relying parties need to know and trust only one issuer. Another benefit of establishing a more regulated and governed ledger is that in this way the legal aspects of data privacy can be addressed by the involved entities which can ensure that the data processing within the system conforms with legislations such as GDPR.

7.4 Limitations

This research has been carried out over a four-month period that inevitably poses certain time limitations at the expense of exhaustiveness of the study conducted. It is important to note that it is an exploratory study since the concept of Self-sovereign Identity is fairly new and the underlying technology standards are bound to undergo significant design changes. Both Self-sovereign Identity and Blockchain Technology are new research domains, thus there is a lack of prior extensive academic research. In order to tackle this limitation, academic research on traditional Identity Management and current institutional research on the technical foundations of Blockchain Technology has been used as a theoretical framework.

The artefact has been developed by deriving the system requirements from the analysis of various SSI management Blockchain-based systems and existing public transport ticketing systems. A limitation of this study is that the system requirements could not be validated by experts from the public transportation industry since the prototype is limited in functionalities as well as the complexity of the concept poses a significant limitation. In addition to that, the set of requirements as presented in this thesis are not exhaustive. A Self-sovereign Identity system will possibly need to satisfy additional requirements to be fully functional and ready for operational use. Another limitation constitutes the fact that no evaluation was conducted at the end of the design and development phase of the design science research process which can be justified due to the time limits and the lack of unfamiliarity of the involved parties with such concepts.

In terms of the developed artefact, a few weaknesses can be pointed out. The artefact could not incorporate all the proposed system requirements; thus, a minimum viable product was developed. It can be utilized as a blueprint for future implementation of a full-fledged and scalable system. Furthermore, since the authors had no prior knowledge with Hyperledger Indy and the Blockchain Technology, aspects such as architecture design choices might be lacking better clarity and justification. The user-facing components of the developed artefact could be further improved and also tested with end-users from all relying entities, for example, with a high-fidelity prototype.

Also, the problem that this thesis addresses is not a vital problem for people but rather is an emerging concern on data privacy and interoperability of the current system. Especially with recent European legislations on data privacy and regulations such as the GDPR, users are more concerned and conscious about who has full control over their data and how it is processed. However, the legal framework around the utilization of Blockchain Technology is yet to be clearly defined as there is uncertainty about accountability and jurisdiction issues.

7.5 Future Outlook

The current prototype showcases the potential of SSI technology in the public transport sector. However, it is just the starting point for further development and research that needs to be undertaken in the future to arrive at a feasible solution that can be used in a real-life scenario. The standards and underlying system requirements have not been fully developed yet and are constantly changing. Thus, the idea of an SSI based system in public transport needs to be constantly updated to future development work that has been done in this space. However, the SSI landscape and its ecosystem of applications are constantly growing. Compared to other industry sectors the SSI space does not rely on competition but rather a collaboration between different system applications. The nature of SSI allows to grow the ecosystem as a whole and shift more people into creating their own decentralized identities while offering interoperable identity solutions. In regards to public transportation it is important to take a closer look into payment integrations to facilitate the value exchange between Service Providers and users in a seamless way.

Moreover, more features should be taken into account to keep up with existing transportation platforms. Here features like travel planners, travel history and other additional features should be taken into account in order to develop an application that can substitute existing ones. The future of public transport in combination with Self-sovereign Identity looks promising. It opens up the industry to the integration with any other type of application through respecting privacy, data protection regulations and interoperability. This would, for example, allow the public sector to integrate with private sector applications without compromising data protection laws. Through extended integration of applications, the entire ecosystem could grow in its usability. For example, private transportation offerings like car-sharing or scooter sharing could integrate with public transportation offerings and offer a combined solution.

Here KYC processes of registering with different companies could be combined and help to reduce costs. A particular use case would be the verification of driver licenses. The license could be stored as a Verifiable Credential and reused by anyone that requires it to sign up. However, the possibilities of integration are not just limited to transportation offerings. Here, for example, other services could be integrated. For example, tax reimbursements for business travel through the integration of credible user-centric travel history that can be shared with authorities by the user.

Furthermore, due to legal uncertainty, several opportunities for future research could arise. One of them might be concerned with developing a legal framework for the legislative implications of a Blockchain-based IdMS. This could lead to a bunch of new research questions to be answered. Investigating, for instance, who are the data controllers and processors in a system like the one proposed by this thesis. Also, how publicly visible data is controlled on such a system, or which laws apply and in what jurisdiction could potentially qualify as future research questions.

Besides the legal field, further research may be conducted on the societal implications of the proposed system. The paradigm shift of removing centralized authorities that today act as trusted third parties may face resistance from end-users. This mainly lies in the fact that multiple different stakeholders need to be present on the network. Thus, users and Service Providers need to participate in the network as well as issuers who facilitate assertions of claims. Without any users, there would be no incentives for Service Providers to join. Moreover, without any Service Providers, no users would likely join. This is the classic “chicken and egg” problem and in order to overcome it, the adoption of such networks should be accompanied by incentives for Service Providers.

Chapter 8

8. Conclusion

In most existing Identity Management solutions users delegate the control of their identity to Identity Providers. Those providers facilitate the communication between users and different Service Providers to exchange identity specific information. This leads to the answer of the first question posted in this research. How users are enabled to get full control over the management of their identities.

In order to give users full control over their own identity a direct identity layer between a verifier, an issuer and an identity holder needs to be established. This removes any intermediary that is acting as an Identity Provider and guarding identity information of users. The need for a trusted and reliable system layer highlights the fundamentals of a Blockchain-based system. Here trust is delegated to a network instead of a single party. Thus, cryptographic principles like DPKI enables users to directly control their data through a public private key pair. Any person or organization is publicly verifiable through a decentralized identifier (DID) on the ledger. This allows the direct issuing and verifying of credentials between stakeholders. Through cryptographic proofs that are encoded into the credential users are able to directly verify each other's information. Any identity information can be associated with its associated identity holder and issuer without the need to engage with a third-party provider. Thus, a trusted, reliable, transparent and immutable network as well as cryptographic proofs are the fundamental layers in order to grant full control to users. Moreover, to decentralize the full control, the network has to be publicly available. For this reason, a public permissionless or at maximum a public permissioned Blockchain has to be implemented in order to build a feasible SSI solution.

Following the description of the underlying technology different infrastructural components for an interoperable Identity Management solution for public transport can be highlighted. Those act as the building blocks for the entire system and provide the answers to the second sub-question of this research. This in particular tries to answer the question of how Blockchain Technology can enable users to use their identities in the public transportation sector. The final artefact in this research introduces a prototype that utilizes a Verifiable Credential (Euro Mobility Card) that acts as an identifier throughout the public transportation network in Europe. The design and requirements of the EMC are built upon a set of globally defined technical standards by the W3C. Moreover, domain-specific trust networks like the European Transportation Authority define credential specific attributes and store them on a decentralized ledger. This creates a consistent solution for issuing credentials and authorizing passengers independent of the used transportation network. Thus, domain-specific schema designs for credentials, organization dependent schema definitions as well as Blockchain enabled validation processes are providing the first set of infrastructural components for a decentralized Identity Management solution for public transportation.

Furthermore, this sheds light on sub question three which is concerned about the benefits for Service Providers and end-users of such a system. Current data protection regulations like GDPR and a growing demand for higher privacy standards by users lead to additional operational expenses for companies. Thus, implementing a privacy centric system that focuses on data minimization reduces storage costs for Service Providers. Those costs are usually associated with storing private data of users for validation purposes. This moves identity to the edges of the network and allows Service Providers to focus on key features of their service without the need to oversee and secure big amounts of private data. Moreover, through standardized issuing and validation processes the overall market would benefit from reduced Know Your Customer costs. Those can be reduced through the overall trust of the network which enables users to enter the system at multiple entry points without sacrificing data integrity. This also improves the overall usability for end users and opens up the system to a broader audience. The benefits can also be seen when looking at the overall data quality. The system enables Service Providers through the implementation of claims repositories to consistently validate the accuracy of existing user data. This creates a central spot for data updates and reduces data redundancies. Last but not least a decentralized system would improve the overall landscape of the market and break free from vendor lock-in by third parties. Through SSI embedded principles users could simply port their data to the next best solution without leaving the underlying system architecture. An SSI based identity solution for public transport is therefore seamlessly compatible with any other type of service that builds on W3C identity standards. Thus, this also encourages a common identity solution that is decoupled from any previous application centric solution.

The described technology, components and benefits illustrate how a Decentralized Identity Management System based on SSI could look like. The proposed system design acts as a first stepping stone towards the implementation of a DIdMS in public transport. However, the mentioned methods and principles that define the current system design are still evolving and need to be constantly updated for future research in this field. Moreover, the proposed artefact is an abstraction of the existing transportation network, which needs to take further real-world consideration into account. Thus, to move from a more abstract solution to a real-world implementation many more aspects, which are out of the scope of this study, have to be considered. Those would for example include the implementation of physical validation machines and devices that validate in real time

Verifiable Credentials and give users access to premises. Moreover, the question of what standards to adapt for a decentralized system have to be assessed. The progress of the W3C has to be further studied. However, the W3C standards should not be the only standard to be considered in future work. Relying on one standard by one authority would lead to a re-centralization of the entire system.

This paper presented a Decentralized Identity Management System in the public transport sector that utilizes the building blocks of SSI and Blockchain Technology. Thus, it can be concluded that this research provides a theoretical answer to the problem statement as it puts all the relevant components into practice. The conducting research contributes to developing a foundation for designing a DIdMS by facilitating the SSI principles in the public transport sector.

Bibliography

AECOM. (2011). *Study on Public Transport Smartcards*. European Commission.

Allen, C. (2016, April 25). *The Path to Self-Sovereign Identity*. Life With Alacrity. http://www.lifewithalacrity.com/2016/04/the-path-to-self-sovereign-identity.html

Allen, C., Brock, A., Buterin, V., Callas, J., Dorje, D., Lundkvist, C., Kravchenko, P., Nelson, J., Reed, D., Sabadello, M., Slepak, G., Thorp, N., & T. Wood, H. (2015). *Decentralized Public Key Infrastructure*. Web of Trust. https://danubetech.com/download/dpki.pdf

Ashbacher, C. (2004). The Unified Modeling Language Reference Manual, Second Edition, by James Rumbaugh. In *The Journal of Object Technology* (Vol. 3, Issue 10, p. 193). https://doi.org/10.5381/jot.2004.3.10.r1

Bakre, A., & Patil, N. (2017). Implementing Decentralized Digital Identity using Blockchain. *IJETSR*, *4*(10).

Bander, S. (1997). *Key words for use in RFCs to Indicate Requirement Levels*. http://www.rfc-editor.org/rfc/rfc2119.txt

Berberich, M., & Steiner, M. (2016). Practitioner's Corner · Blockchain Technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers? In *European Data Protection Law Review* (Vol. 2, Issue 3, pp. 422–426). https://doi.org/10.21552/edpl/2016/3/21

Bryman, A., & Bell, E. (2011). *Business Research Methods 3e*. Oxford University Press.

Buchner, D. (2016). *DIF Identity Hubs*. https://github.com/decentralized-identity/identity-hub/blob/master/explainer.md

Buchner, D. (2019). *Rhythm and Melody: How Hubs and Agents Rock Together*. https://medium.com/decentralized-identity/rhythm-and-melody-how-hubs-and-agents-rock-together-ac2dd6bf8cf4

Calypso. (2017). *White paper Account Based Ticketing with Calypso*. Calypso. https://www.calypsonet-asso.org/sites/default/files/170529-

CalypsoWhitePaperABT_%20v2.3.pdf

Civic Technologies. (2017). *Civic - White Paper*. Civic. https://tokensale.civic.com/CivicTokenSaleWhitePaper.pdf

Decentralized Identity Foundation. (2018). Decentralized Identity Foundation. https://identity.foundation/

D'Onfro, J. (2015, August 1). *Facebook is dominating Google, Yahoo, and Amazon in this one niche area*. Business Insider; Business Insider. https://www.businessinsider.com/facebook-third-party-identity-provider-2015-7

European Travellers Club. (2019). *European Travellers Club - For interoperable Public Transport in Europe*. European Travellers Club. http://www.europeantravellersclub.eu/.cm4all/mediadb/ETC%20brochure%2C%20version%201.1.pdf

Faber, B., Michelet, G., Weidmann, N., Mukkamala, R. R., Vatrapu, R., Centre for Business Data Analytics, Copenhagen Business School, Denmark, & Department of Technology, Kristiania University College, Oslo, Norway. (2019). *BPDIMS:A Blockchain-based Personal Data and Identity Management System*. 52nd International Conference on System Sciences, Hawai.

Fei, C., Lohkamp, J., Rusu, E., Szawan, K., Wagner, K., & Wittenberg, N. (2018). *JOLOCOM: Own your digital self*. Jolocom. https://jolocom.io/wp-content/uploads/2018/07/Jolocom-Technical-WP-_-Self-Sovereign-and-Decentralised-Identity-By-Design-2018-03-09.pdf

Fiorelloa, D., Martino, A., Zani, L., Christidis, P., & Navajas-Cawood, E. (2016). Mobility data across the EU 28 member states: results from an extensive CAWI survey. *Transportation Research Procedia*.

Giesecke & Devrient. (2011). *The Ticket to Global Mobility*.

Gleasure, R. (2015). When is a problem a design science problem? *Systems, Signs & Actions*, *9*(1), 9–25.

Goldkuhl, G. (2012). Design Research in Search for a Paradigm: Pragmatism Is the Answer. In *Communications in Computer and Information Science* (pp. 84–95).

https://doi.org/10.1007/978-3-642-33681-2_8

Gregor, S., The Australian National University, Hevner, A. R., & University of South Florida. (2013). Positioning and Presenting Design Science Research for Maximum Impact. In *MIS Quarterly* (Vol. 37, Issue 2, pp. 337–355). https://doi.org/10.25300/misq/2013/37.2.01

Hevner, A. R. (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, *19*(2), 87–92.

Hevner, Hevner, March, Park, & Ram. (2004). Design Science in Information Systems Research. In *MIS Quarterly* (Vol. 28, Issue 1, p. 75). https://doi.org/10.2307/25148625

Holmberg, P.-E., Collado, M., Sarasini, S., & Williander, M. (2016). *MOBILITY AS A SERVICE-MAAS: Describing the framework*.

Hyperledger. (2019). *Hyperledger training material* [Nodejs]. https://github.com/hyperledger-archives/education/tree/master/LFS171x/indy-material/nodejs

Kozlowska, I. (2018, April 30). *Facebook and Data Privacy in the Age of Cambridge Analytica - The Henry M. Jackson School of International Studies*. The Henry M. Jackson School of International Studies. https://jsis.washington.edu/news/facebook-data-privacy-age-cambridge-analytica/

Laurent, M., Denouël, J., Levallois-Barth, C., & Waelbroeck, P. (2015). Digital Identity. In *Digital Identity Management* (pp. 1–45). https://doi.org/10.1016/b978-1-78548-004-1.50001-8

Lundkvist, C., Heck, R., Torstensson, J., Mitton, Z., & Sena, M. (2016). *UPORT: A PLATFORM FOR SELF-SOVEREIGN IDENTITY*. uPort. https://blockchainlab.com/pdf/uPort_whitepaper_DRAFT20161020.pdf

Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*.

Parra-Moyano, J., & Ross, O. (2017). KYC Optimization Using Distributed Ledger Technology. In *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2897788

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. In *Journal of Management Information Systems* (Vol. 24, Issue 3, pp. 45–77). https://doi.org/10.2753/mis0742-

1222240302

Piekarska, M., Lodder, M., Larson, Z., & Young, K. (2018). *When GDPR becomes real*. Web of Trust. https://github.com/WebOfTrustInfo/rwot5-boston/blob/master/final-documents/gdpr.pdf

Poikola, A., Kuikkaniemi, K., & Honko, H. (2015). *MyData - a Nordic Model for human-centered personal data management and processing*. Finnish Ministry of Transport and Communications. http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/78439/MyData-nordic-model.pdf

Purao, S. (2013). Truth or Dare: The Ontology Question in Design Science Research. *Journal of Database Management*, *24*(3), 51–66.

Rebooting Web-of-Trust. (2018). https://www.weboftrust.info/

Reed, A., & Drummond, T. (2017). *The Inevitable Rise of Self-Sovereign Identity*. Sovrin Foundation.

Reed, D., Law, J., & Hardman, D. (2016). *The Technical Foundations of Sovrin*. Sovrin. https://www.evernym.com/wp-content/uploads/2017/07/The-Technical-Foundations-of-Sovrin.pdf

Reed, D., Sporny, M., Longley, D., Allen, C., Grant, R., & Sabadello, M. (2017). *Decentralized Identifiers (DIDs) v1.0*. https://www.w3.org/TR/did-core/

Roadmap to a Single European Transport Area – Towards a competitive and resource efficient. (2011). European Comission.

Sabadello, M. (2017). *A Universal Resolver for self-sovereign identifiers*. https://medium.com/decentralized-identity/a-universal-resolver-for-self-sovereign-identifiers-48e6b4a5cc3c

Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research Methods for Business Students*. Pearson Education.

Sonnenberg, C., & vom Brocke, J. (2012). Evaluation Patterns for Design Science Research Artefacts. In *Communications in Computer and Information Science* (pp. 71–83). https://doi.org/10.1007/978-3-642-33681-2_7

Sporny, M., Longley, D., & Chadwick, D. (2018). *Verifiable Credentials Data Model 1.0*. https://www.w3.org/TR/vc-data-model/

Sporny, M., Longley, D., Kellogg, G., Lanthaler, M., Champin, P.-A., & Lindström, N. (2018). *JSON-LD 1.1*. https://www.w3.org/TR/json-ld/

Steg, L. (2003). CAN PUBLIC TRANSPORT COMPETE WITH THE PRIVATE CAR? In *IATSS Research* (Vol. 27, Issue 2, pp. 27–35). https://doi.org/10.1016/s0386-1112(14)60141-2

Tankard, C. (2016). What the GDPR means for businesses. In *Network Security* (Vol. 2016, Issue 6, pp. 5–8). https://doi.org/10.1016/s1353-4858(16)30056-3

Tobin, A. (2017). *Sovrin: What Goes on the Ledger?* Sovrin. https://sovrin.org/wp-content/uploads/2018/10/What-Goes-On-The-Ledger.pdf

UK Government Chief Scientific Adviser. (2016). *Distributed Ledger Technology: beyond blockchain*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/492972/gs-16-1-distributed-ledger-technology.pdf

United Nations. (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*. https://sustainabledevelopment.un.org/post2015/

Urban ITS Expert Group. (2013). *Guidelines for ITS deployment in urban areas* (No. 1). URBAN ITS EXPERT GROUP. http://www.smart-ticketing.org/downloads/reports/2013-urban-its-expert_group-guidelines-on-smart-ticketing.pdf

Venable, J. (2011). Incorporating Design Science Research and Critical Research Into an Introductory Business Research Methods Course. *The Electronic Journal of Business Research Method*, *9*(2), 119–129.

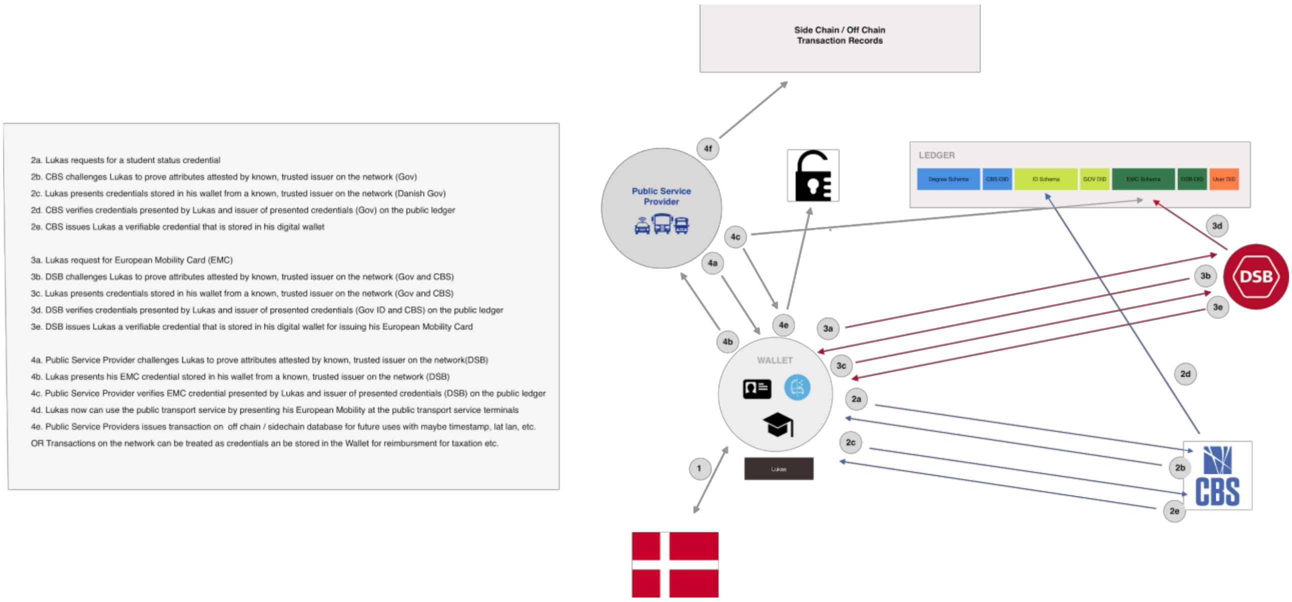
W3C. (1994). World Wide Web Consortium (W3C). https://www.w3.org/

Wagner, K., Némethi, B., Renieris, E., Lang, P., Brunet, E., & Eric, H. (2018). *Self-Sovereign Identity - A position paper on blockchain enabled identity and the road ahead* (No. 1). Blockchain Bundesverband E.V.

Windley, P. J. (2005). *Digital Identity* (Allison Randal And (Ed.)). O'Reilly.

Appendix

EMC Flow



2a. Lukas requests for a student status credential

2b. CBS challenges Lukas to prove attributes attested by known, trusted issuer on the network (Gov)

2c. Lukas presents credentials stored in his wallet from a known, trusted issuer on the network (Danish Gov)

2d. CBS verifies credentials presented by Lukas and issuer of presented credentials (Gov) on the public ledger

2e. CBS issues Lukas a verifiable credential that is stored in his digital wallet

3a. Lukas request for European Mobility Card (EMC)

3b. DSB challenges Lukas to prove attributes attested by known, trusted issuer on the network (Gov and CBS)

3c. Lukas presents credentials stored in his wallet from a known, trusted issuer on the network (Gov and CBS)

3d. DSB verifies credentials presented by Lukas and issuer of presented credentials (Gov ID and CBS) on the public ledger

3e. DSB issues Lukas a verifiable credential that is stored in his digital wallet for issuing his European Mobility Card

4a. Public Service Provider challenges Lukas to prove attributes attested by known, trusted issuer on the network(DSB)

4b. Lukas presents his EMC credential stored in his wallet from a known, trusted issuer on the network (DSB)

4c. Public Service Provider verifies EMC credential presented by Lukas and issuer of presented credentials (DSB) on the public ledger

4d. Lukas now can use the public transport service by presenting his European Mobility at the public transport service terminals

4e. Public Service Providers issues transaction on off chain / sidechain database for future uses with maybe timestamp, lat lan, etc.

OR Transactions on the network can be treated as credentials an be stored in the Wallet for reimbursment for taxation etc.

Prototype

If you want to run the prototype please follow the instructions in this repository:

- https://gitlab.com/zbkrt/european-mobility-card/-/tree/master/LFS171x%2Findy-material%2Fnodejs

Register DID

```
{
  "auditPath":[
    "UqByQMaMyMQ9zUL8qKm3EgbMepdVLgK3QiiwVpQiX8E",
    "J1TQyqB6zvirTU6bBGWxReZQvvVLXkHkrYW9NB95YqvQ",
    "4JKM7N6RVZbgNcSt8oLY4bmtJ8ua3CbPSYkduRnftNys",
    "DpAWbvEGkPstyskc4mCtb7E9si2Sm1UVGx8T8nJMnXTw",
    "FWuXMCH7Bc5aBMnUfGCBm7kpquDxW2yCAMF42jWYSZRL",
    "3hu8GLowfe2q98wXrXXqc6sN6PhCKTqvxqm1RXUPrKed"
  ],
  "reqSignature":{
    "type":"ED25519",
    "values":[
      {
        "from":"Th7MpTaRZVRYnPiabds81Y",
"value":"3RHeTGWtTioPAkQChytG4przZoSiFyNU94HvzVuvZ7nGL3diiPK6PDS4puMUZdLWGwRDZ8rdftskt7rMwSpFpsct"
      }
    ]
  },
  "rootHash":"37LJpcFN6v8MEgob88MEZAMHExSHYoyUvsB8MDfaG7CJ",
  "txn":{
    "data":{
      "dest":"3qmTXyyFRn8EttFkNUQXU8",
      "role":"101",
      "verkey":"2YobTiTFaxXHHE9Bw2CZZUhBxeEp1RZdqn7rHSfuQbut"
    },
    "metadata":{
      "digest":"a9af402ef68849fb2bcbbc93904bd9ed3bf85c1e50d39853e6d420975d3a19a1",
      "from":"Th7MpTaRZVRYnPiabds81Y",
      "reqId":1584025815138533000
    },
    "protocolVersion":2,
    "type":"1"
  },
  "txnMetadata":{
    "seqNo":4806,
    "txnId":"4e70c82f34751cd8b3fd36606d7827dcba029d5849f72a064b88498d09db328c",
    "txnTime":1584025822
  },
  "ver":"1"
}
```

EMC Credential Definition

```
"Euro-Mobility-Card-Data":{
  "name":"Euro Mobility Card-Data",
  "version":"0.1",
  "requested_attributes":{
    "attr0_referent":{
      "name":"card_number",
      "restrictions":[
        {
          "cred_def_id":"5LYiBBHCgZvcPTqaamVrG8:3:CL:4444:Euro Mobility Card"
        }
      ]
    },
    "attr1_referent":{
      "name":"is_student",
      "restrictions":[
        {
          "cred_def_id":"5LYiBBHCgZvcPTqaamVrG8:3:CL:4444:Euro Mobility Card"
        }
      ]
    },
    "attr2_referent":{
      "name":"last_name",
      "restrictions":[
        {
          "cred_def_id":"5LYiBBHCgZvcPTqaamVrG8:3:CL:4444:Euro Mobility Card"
        }
      ]
    },
    "attr3_referent":{
      "name":"first_name",
      "restrictions":[
        {
          "cred_def_id":"5LYiBBHCgZvcPTqaamVrG8:3:CL:4444:Euro Mobility Card"
        }
      ]
    },
    "attr4_referent":{
      "name":"photo",
      "restrictions":[
        {
          "cred_def_id":"5LYiBBHCgZvcPTqaamVrG8:3:CL:4444:Euro Mobility Card"
        }
      ]
    }
  },
  "requested_predicates":{

  }
},
```

EMC Credential Issuing

```
{
  "auditPath":[
    "4pts5t3ujsGE7EYchNtWeQ631sAoEbbG8krfsLtja4S1",
    "6vTjkrays4gcgmoeXCwAi9atmiTkVycMKy3jkmU5APaq",
    "ikkwwdZryqVwVGXCGQ7pSJenryivVZGYSLjBw1wZ75K",
    "9KN1UGdhKhiGMRFxr3srd4edgSb7RJK6wcxTnZMA4FaP",
    "2rGUCVk1NpsV38sGYtRE4EuG8nfhV5Xy3eYgn8sdjnKH",
    "FWuXMCH7Bc5aBMnUfGCBm7kpquDxW2yCAMF42jWYSZRL",
    "3hu8GLowfe2q98wXrXXqc6sN6PhCKTqvxqm1RXUPrKed"
  ],
  "reqSignature":{
    "type":"ED25519",
    "values":[
      {
        "from":"MseGJssFMUkmNutSSBSNT5",

"value":"37FN6CD8N9JCAh4JiW3ZFpNJXfLj1dKzWd95Lcep2TkBZQwTsaT2h9Z6BNWDwQ6MXwgioMV8rX9anNvP
Hi3dhADG"
      }
    ]
```

```
   },
   "rootHash":"4MeCnTezFPyQSCUVtsTstJY12uaDSRaMppqUbBFiiN9f",
   "txn":{
      "data":{
         "data":{
            "primary":{

"n":"799492209924368651909165643934422419053805922957851777639076501785139659102808230611357938
4721072186545104927161707519281136311182020342300322097218885706218780493827888824741778647432
0385307510677929764749671216731468133923227501858436530493566955575521290941917044297697893102
2471135689326660731860672113984792676593327538755515751835399227030048612242490114720046347159
2203593211450808639104159988637028299339824466889666577622853044279291863654339982945708036805
3833980711096558913040257327262842880828858460988612448685341305918258173134445669023344281122
378701979357119465855283049043979929992430420945706711261",
                "r":{

"bank_account":"750110300097988567044868007969219304567760681544103145975508109075534405031994044
6559648134233997796178286291025496267357747032077902604368029957991347309687335427989405413517
2673873603354568926986844488476509611149029142268581153785265721372652858006747222086863824511
0297199677663773371984854282941528424557090971940339064214686312910556068165030940443008974760
8404317270246205790270627402105598942624144744580490095705909818069204322005101717641153620883
8061723125691543772481328406303351016190017943302503646912334599409172950002409427684383767378
2780190978412931799160160121109210921607311150834619518333627981699",

"card_number":"350126316295418137104662151419706386845801973085779669206470303270844483786843563
7045535180552830793390670228043387492986164330494911382410206831313513989354728026173905641756
2149275571269254006427821808670161518610744122180148769216235452440481752800079258790130692899
3493350846373861984508632592713810727549718473321334140544123302413191604983896118445001222531
6054554198952529562945031640418815347197327494448396205021774888027151248715488866379950079870
1841500789427805042335461551843076874687710984030772692918908961961655021145819165512461703542
212639964372486214066646432443790373800928906020466440767909563409",

"first_name":"14716269944216369836654295222583487071053183985492397980180577465938188439690991953
0833539643641360207859497896793090115379689510962708129107260891133071019752119003652441862483
0787600844055853613718069096043456328721369129700066580745930084315760336700081509927361514620
1691679681782508753543687489061631781555882839127805260870537143201538320896220704845616088973
8086552615545977072167004944213200919298022799220977183763140807889384868615008473934273284497
7884724260672856107362539693952362458968937426417005432521541125808042298440559383848628183611
9955974252978258497502273748958664986165939795028621860541372566",

"is_student":"22327723296362527794438836752730294394944975445261781216648200312830970046380671376
5740769219978324010306987107493552339359122520852326051498209618911946435358433163106990457727
3170464679010207911535576278511840155479766263249281706413214735061160090946458601731578391487
2755052015764350514244845953451554224821964220494235033686812346756887710762008276196881787901
2326344880833861517333657562961512431363557341331376932066103523984436128856287432955958971263
5456775694279483335055791266691664924118851409867239873467765464034284970035501859084219152031
6937630154095930722828327952641640449266442859337835145754914415",

"last_name":"44853248792978648673615941908382363168224305636007279994407003091038942379833860280
6624622823301534087559719447592105043428387186681539986038599777920529453638143293731080727186
2078798405186697693082927136639430273213894666758994835808578301771713965885797725283144759194
0935048575637180477109904911640021415139267848632340370619845222546658225229912983913879963988
7866316544089157667792390738603075094063266521204537054045481687855853257246947473885835098075
6851594971950118431727291157344057445816740919800365144794246887358115322079103383614831838668
1063558475996049185428710872274427055079127974233697379727977273",
      },

"master_secret":"6743028705269822241974379337534822729201527906446912252294409840500109141306422164187779912745404075895438993978262327869043590801882921067473617432157237519906315295411080422204019340605988247481270529832216318202387941101067870461570473787285986316364524005959571529539200732716754653539542612753774794337219476257326922830465947743011224482702558880679578378303694412004994986686116571934042282479964156169860058425608158105183146877973354480617273211823220471038196860583963121520407073178156490883466758998597523345227229163398864221956393016985577137629671403194363066086033411767814501838078866889831773079544 3",

"photo":"6269174509664995498048390455125394768204056474436630191275615940529910405881279295377115332135905250009215195791922218244712180668872106610229597407972424699678161916000898277189992333912407284857723329014689632962766991264411130214706322138546875082219003348056524888369026521248216736717082729374868982478159357043871154557144851788785605399243787231216844877431288993422132212136063694740696260323493952203941795959434370998581355421708067556858405510587550908727730840810683849910757430176317011559162497947825985363908831417400234743164019695919159617342622591196416068554763601362852745149168805621182298112283 2"
          },

"rctxt":"3558486113536544999180474060351357564164047783338879856081146480667721171191850731225966441615310323319918452366765889376241737403121138925250246938216162204703526752550281562255479723399498836322123156326949835793972487052437734460258956266673414553736387186230902930654517418667513752908373115838841376366687470253546527624166675997003845764366584059755158389309312611172838685914995996569200951414119039048513595132058056782267489620848020636698620824420877707654822506045768120912075106619179651618198187274224212976527413748208565899683676390111742702572603687164308722460767741019353404770409873073904849121689 5",

"s":"2976191075289163829669498661302922071716561414394439296132311311654148817445181619470257528268380213177902741757455298045380304094154634951543533820460812208977772029052073200223060464761376892294199004349408078979909038051076684006120090244510009096014462178119798793783296936537247264904763972603021611751600146943852053782561607657270732214677704690222810704179401002592475927718705177581801915253056095631649550210672659180760035921166899308239926001047921893877916966559170833637360236285443249175462180439177283928221145997694277480633954082678350906999353268689862396070155110558807528525361993854005140632092 2",

"z":"3106732419695289706121069165300742675423231163426381266610371173445501526208513949412927287934818331426030246833097982723202594275008662181914449473780106190198637684665794889957974387717484308449625891889847511066531144187221707259831427909339170817866358375651862522773796922901207578243655872861207081964399674122620086628363915274586481026763265163311998959997896425480285088590314032782251260890036171877185597212762623410063661625493028184604539427418252417803451371104085455121107300818495839280658743681833881704858123143600923688308695932596122294754536503328219553641218213302506130123548078110446081015576 9"
        }
      },
      "ref":4915,
      "signature_type":"CL",
      "tag":"EMC"
    },
    "metadata":{
      "digest":"234a53637ee6c2ef63a37eeaf145ac8e859e773bdf32c5f38d238b522f6bebf5",
      "from":"MseGJssFMUkmNutSSBSNT5",
      "reqId":1584270231243599600
    },
    "protocolVersion":2,
    "type":"102"
  },
   "txnMetadata":{
      "seqNo":4916,
      "txnId":"MseGJssFMUkmNutSSBSNT5:3:CL:4915:EMC",
      "txnTime":1584270231
   },
   "ver":"1"
}
```