Technische Hochschule
Brandenburg
University of
Applied Sciences

Self-sovereign Identity: Development of an Implementation-based Evaluation Framework for Verifiable Credential SDKs

Brandenburg University of Applied Sciences
Department of Economics

Master's Thesis

submitted by
Philipp Bolte
July 28, 2021

First supervisor: Prof. Dr. rer. nat. Vera G. Meister
Second supervisor: Jonas Jetschni, M.Sc.

Statutory Declaration

I hereby attest that I have written this thesis independently without any outside help and that I have used only the sources cited.

Brandenburg, July 28, 2021

Philipp Bolte

Abstract

Diese LaTeX-Vorlage ist für Berichte, Bachelor- sowie Masterarbeiten gedacht. Natürlich ist sie nicht perfekt und jede Art der Verbesserung wird dankend angenommen. Bei Fragen zur Verwendung oder Anregungen zur Verbesserung können Sie mir diese gern an markus.brandt1992@gmail.com senden.

An entsprechender Stelle werden Beispiele für die Verwendung von Abkürzungen, Zitaten, Abbildungen, Tabellen und die Einbettung von Code gegeben.

Contents

List of Figures

List of Tables

List of Abbreviations

SDK Software Development Kit
SSI Self-sovereign Identity
VC Verifiable Credential

1 Introduction

The Internet has become a cornerstone of coexistence in today's world. With over 4.66 billion Internet users worldwide [Jo21], it determines how we communicate, think, inform ourselves, and interact with one another. As a result, huge networks of people are being created in which different cultures are coming closer together and knowledge is being shared like never before. A central enabler for the functioning of such a digitalized society are digital identities [Li20].

Over the course of our lives, we generate a large amount of digital identities from a wide variety of services, including Facebook, Twitter, WhatsApp, GitHub, LinkedIn, and many more. They represent us in this digital realm, are part of our personality, and allow us to identify ourselves online. Because of the way we manage digital identities in the current era, users mostly own separate identities for each service or go through centralized, federated identity providers like Google or Facebook. As a result of these key developments, silos of identity data emerged, which are problematic concerning efficiency, security, and privacy. This creates a dependency towards the services that have full control over the identity data. This makes it difficult for users to control how services exploit this power for their own interests. In addition, various data leaks and hacks in which sensitive user data became public show that the current approaches are not suitable for the problems of these modern times [Sw21]. [Eh21, pp. 2-3]

In contrast, the Self-sovereign Identity (SSI) paradigm takes a new approach by giving users full control of their digital identities through various novel approaches [FCA19, p. 103059]. This work examines this new approach from a developer's point of view to test its practical applicability. In the next sections, the scope, related work and the research approach will be discussed.

1.1 Scope of Work

For a successful realization of Self-sovereign Identity (SSI) concepts, the existence of good solutions for developers is critical. This ensures that the barriers to a successful adoption of SSI are kept to a minimum, simplifying and speeding up the entire process. A good toolset and developer experience is thus a key enabler for SSI.

With this in mind, an overview of the most important solutions[1] in the SSI space is established throughout the thesis. To scope the work accordingly, this work looks at the solutions in terms of how closely they can map the lifecycle of a Verifiable

[1] synonymous to Software Development Kits (SDKs), libraries, frameworks, and platforms

Credential (VC). This decision was made due to VCs being a key artefact in SSI as they hold the actual verifiable data, e.g. vaccination status or birthdate, of a subject [MDC19]. The overview is intended to serve as an entry point for developers to get a general view of the capabilities of existing solutions and to give starting points for further research.

Furthermore, a use case agnostic reference implementation is presented that implements four of the presented solutions based on the lifecycle. It can serve developers as a basis for their own work, but above all enables practical validation and the gathering of experience during its development. This way, the knowledge gained flows directly into a new evaluation framework, which, in addition to other software selection frameworks, can provide concrete help in selecting the most suitable solution from the developer's point of view. In addition, it can reveal shortcomings in current solutions that need to be addressed for successful adoption of SSI in practical use cases. So the objective of this work, besides the scientific contributions, is to generate added value for the whole ecosystem.

1.2 Related Work

At the current time, there does not appear to be any comparable work that addresses the topic in a manner corresponding to section 1.1. The most similar is [NJ20] who have developed a mobile wallet based on uPort that covers login, VC issuance as well as verification. Based on the experience gained, an evaluation of uPort has been made as well. However, uPort is currently no longer being developed, and the assessment is also based on only a fraction of the VC lifecycle and basic principles for SSI.

Another paper by [Ku20] defines a comprehensive evaluation framework from an enterprise perspective that, compared to other papers, also covers aspects such as user experience, technology and compliance. It is characterized by a wide range of questions that are used for the evaluation of 43 solutions. However, the list of solutions considered is outdated and missing important players (see e.g. MATTR and Trinsic). Furthermore, the assessment does not provide any practical guidance for developers. A clear analysis of the SSI-relevant features, e.g. with regard to the VC lifecycle, does not exist.

Otherwise, many papers seem to focus on theoretical foundations or evaluation of existing solution based on two things: (i) architecture [GMM18] concerning privacy [Be19], performance [Bo20], use case [Ku20], various variations [Al16, Re21, AL20, Bo20, FCA19, Ca05] of SSI principles [Bo19, Bo20, DT20, DP18, FCA19, SNE20], and (ii) the interoperability between those systems [Ho20, Jo20]. This clearly shows that there is a deficit in terms of works that look at existing solutions based on their practical features and applicability from a developer's point of view. This thesis addresses some of these gaps and thus clearly contributes to the field of research.

1.3 Methodology

The process for achieving the objectives from section 1.1 can be divided into four basic steps: gain theoretical foundation, create solutions overview, develop reference implementation, and define the evaluation framework (see figure 1.1).

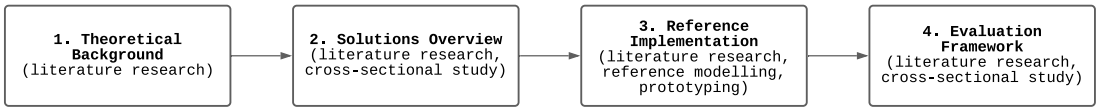

Figure 1.1: Research Approach

For this purpose, various methods of business and information systems engineering according to [WH07] are applied. Through a literature research, an overview of existing papers and books is created, which on the one hand serves for building a theoretical foundation, but also represents the basis for all other steps.

To find the literature, keywords such as "Self-sovereign Identity" and the following complex query based on [Bo19] were used at Google Scholar:

```
("Self-sovereign identity" OR "Self sovereign identity ")
OR (
    ("block-chain" OR "blockchain") AND ("identity management")
    AND ("solution" OR "implementation" OR "review" OR "survey")
    AND ("verifiable credentials" OR "decentralized identifiers")
)
```

Furthermore, the method of cross-sectional studies, mainly in the form of expert questionnaires, is used. A clearly defined set of questions is used to validate the solutions overview, but also to identify evaluation criteria for the evaluation framework. More details on the selection of experts and the questions are described in section X.X. For the creation of the reference implementation, the methods of reference modelling and prototyping are used. In combination, they allow the development of a software prototype that represents a particular problem in a simplified way and whose analysis can contribute to the discovery of new knowledge. This is especially interesting for the development of the evaluation framework. Moreover, it should also be noted that this thesis follows the general idea of generating real-world artefacts defined by [He07] as part of design science research.

To conclude the methodology, a combination of research approach and applied methods defined previously is reflected in the following research questions:

1. What libraries, platforms, or SDKs are available for implementing Verifiable Credentials?
2. Which SDKs do experts from the field recommend using?
3. Which criteria for evaluating Verifiable Credential SDKs can be derived after developing a reference implementation?

2 Self-sovereign Identity

Health care, social security, education, access to financial services — this is just a small list of requirements that are essential for a decent life and are usually taken for granted by people in the Western world. Yet there are more than 1.1 billion people worldwide who cannot provide identification and thus cannot access the most basic services. Digital identities could make a significant contribution towards solving this problem and giving people the chance to participate in society on a more equal playing field. [Wo17]

As already mentioned in chapter 1, however, current implementations of such digital identities are insufficient for the problems of our modern times. [SNA21] divided these problems into four categories:

1. Data Ownership and Governance
2. Password-Based Authentication
3. Fragmented Identity Data
4. Data Breaches and Identity Fraud

The former describes the fact that users have no ownership over their digital identities and thus cannot exercise any control over them. Service providers often take advantage of this and use collected data to create comprehensive profiles of their users and thus sell tailored advertising space on corresponding marketplaces for high figures. The lack of control also means that service providers can temporarily or permanently deny users access to their digital identity at any time. At the beginning of 2021, this led to much discussion as the account of former U.S. President Donald Trump was permanently banned from Twitter. One of the central concerns was whether service providers have too much power over users' liberties [No21]. In addition, given the frequent and often repeated use of weak passwords, the heavy reliance on password-based authentication is a security risk that may lead to identity theft. If users want to protect themselves, they need to use different and complex passwords for each of their accounts, which quickly becomes a complicated undertaking without a password manager. A study by the password manager LastPass, for example, found that a business customer manages an average of 191 passwords [St17]. While the use of such tools greatly simplifies the management of passwords, they too can pose a major security risk and do not completely protect the user [OR20, Or21, To21]. Alternatives such as single sign-on, where users authenticate to other service providers using for example their Google account, can solve this problem but lead to even greater dependency and centralization. The third issue involves identity data being spread across a large set of service providers, making it difficult to maintain. As a result,

duplicates, errors, and outdated data sets are common. The lack of open standards also complicates interoperability between providers, which could theoretically be used to retrieve, move, or delete data. [SNA21, pp. 2-3]

One of the biggest problems, however, are data breaches. In June 2021 alone, there were 235 breaches with 1.16 billion stolen records, with a total of 18.9 billion records stolen in 1,785 breaches in the first half of 2021 [Ri21]. Looking at the past, there have been quite a few major hacks [Sw21], including:

- Yahoo (2013): 3 billion accounts
- Marriott (2018): 500 million customer records
- Alibaba (2019): 1.1 billion entries
- LinkedIn (2021): 700 million accounts

A survey of 413 people by [Ma21] found that 73 % of participants had been affected by at least one but an average of 5.3 data breaches. In addition, the majority blamed themselves for the breaches, with only 14 % aware that service providers were responsible.

These are decades-old problems that were already critically discussed by Kim Cameron in 2005. Cameron, who last worked as Chief Architect of Identity at Microsoft from 1999 to 2019, wrote the following on a blog article [Ca05]:

> *"The Internet was built without a way to know who and what you are connecting to. This limits what we can do with it and exposes us to growing dangers. If we do nothing, we will face rapidly proliferating episodes of theft and deception that will cumulatively erode public trust in the Internet."*

Cameron attributes these problems to the lack of an identity layer on the Internet, which has resulted in a number of services having to find their own solutions. He calls this a *patchwork of identity one-offs*, which fundamentally still exists today and are difficult to resolve. The reason for this, he says, is a lack of consensus and an unwillingness to give up too much control over identity data. A solution for this is, according to him, an *identity metasystem* that abstracts away deeper complexities similar to hardware drivers or TCP/IP and only loosely couples digital identities to the systems. Such an open identity layer could only be successful if it fulfilled the ten laws of identity defined by Cameron. These include criteria such as user control, consent, pluralism and minimal disclosure. [Ca05]

Over the years, these ideas, among others like [De12, Ca05, id14], gave rise to the concept of Self-sovereign Identity (SSI). It is intended to eliminate the shortcomings of today's established concepts by placing the user in the center and giving them back complete control over their identity data. The user can decide what, to whom and how much data is shared without being dependent on a central authority. The emergence of blockchain technology and various new standards in recent years gave a new boost to implement SSI in reality. [St21, pp. 6-7; To17, pp. 8-9]

SSI is a completely new approach to digital identities on the Internet and is seen as

a paradigm shift that deeply affects the infrastructure and power distribution of the Internet [Re21, p. 3]. For a deeper look at the topic, this chapter takes a closer look at Self-sovereign Identity. For this purpose, the concept of identity and the different types of identities will be discussed first. This is followed by a historical look at the different stages of digital identities, taking a closer look at the previous concepts of SSI. After a basic foundation has been built, standards that have been established in recent years and are intended to make SSI feasible in reality are described. Finally, the SSI architecture with its components and roles will be looked at.

2.1 Identity

What defines a human being? One would probably get various answers to this question, such as its name, gender, place of residence, profession, hobbies, religion, charitable activities, party affiliation or even a combination of all these characteristics. [CK01, p. 206] describes in his work that a person's identity is not just a single, fixed construct, but consists of several partial identities. Thus, depending on the context in which a person finds itself, it takes on one of its various partial identities, which represents it as a human being more or less. For example, a partial identity for health care consists of its medical history, while the partial identity towards work contains received certificates. Nevertheless, these different parts of the identity are not necessarily considered separately, as they can also overlap in certain aspects of information. It is important to mention that a person decides which information to share at which time towards which entity. In 2.1 the concept of partial identities is illustrated exemplarily by a person Alice.

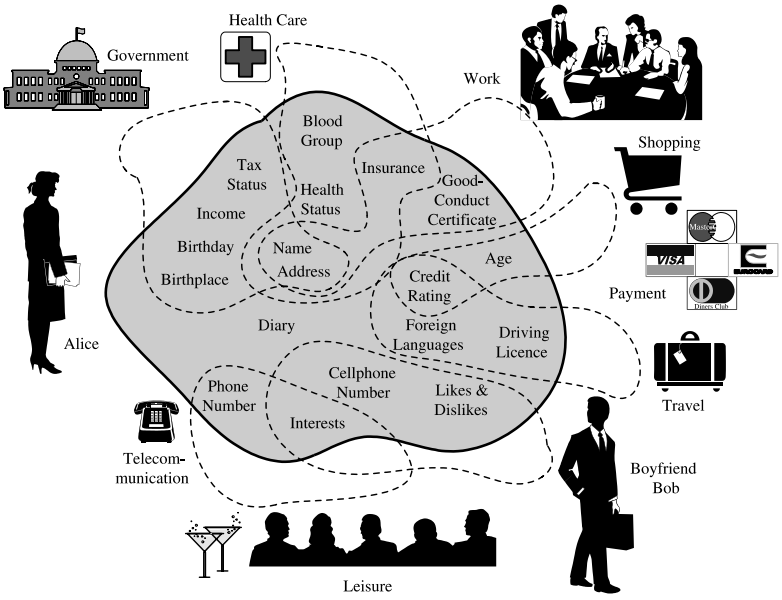

Figure 2.1: Partial Identities of Alice extracted from [CK01]

An important balancing act is to disclose the right amount of data in order to maintain anonymity, but also to provide the other person with the necessary information.

For the purchase of a water, the kiosk vendor should not ask for any personal data, whereas verification of age when buying alcohol is a valid reason for information disclosure. In reality, official documents, such as state identification documents, or sometimes unofficial documents, such as customer cards, are usually used for such proof of identity. Here, users have full control over their documents as they are under their control, and only they can decide self-sovereignly whom and when to show them. Official identification documents are also produced and standardized to ensure the highest possible level of security and interoperability. Other countries can verify such documents without explicitly contacting authorities, simply by looking at the document. Confidence in the validity of the data arises from the fact that the verifying party trusts the authority issuing the document. [St21, p. 6]

As a result of the increasing digitalization of various branches of life, many processes are shifting to the digital world. Digital identities, which are similar to analogue identities in terms of their basic idea, are now being used for interacting with digital services. They allow entities, such as people or objects, to authenticate themselves online through certain attributes and thus prove their identity [MG20, p. 103; Bu20]. A more precise definition is given by Cameron [Ca05], who defines digital identity as *"A set of assertions that a digital subject makes about itself or another digital subject"*. In this context, a digital subject is *"A person or thing represented or existing in the digital realm which is being described or dealt with."* and the attributes mentioned can be represented in the form of claims, which are defined as *"An assertion of the truth of something, typically one which is disputed or doubted"*. The problem is that analogue identities and their documents usually have no or not widely accepted digital representations in the form of digital identities. From this emerged the patchwork of identity one-offs described in chapter 2, resulting in a divergence of digital identities from their original counterparts concerning their characteristics. To better understand this development, the next section describes the different stages of digital identities in more detail. [St21, p. 10; Eh21, p. 2]

2.2 Stages

2.3 Standards

2.3.1 Overview

2.3.2 Decentralized Identifier

2.3.3 Verifiable Credentials

2.3.4 DIDComm

2.3.5 Zero Knowledge Proofs

2.4 Architecture

2.4.1 Roles

2.4.2 Components

3 Expert Questionnaire

3.1 Expert Selection

3.2 Content

3.2.1 Solutions Overview Draft

3.2.2 Questions

3.3 Results

4 Reference Implementation

4.1 Requirements

4.2 Base Implementation

4.3 Architecture

4.4 Solution Integration

4.4.1 Solution Selection

4.4.2 MATTR

4.4.3 Trinsic

4.4.4 Veramo

4.4.5 Azure AD

4.5 Results

5 Evaluation Framework

5.1 Requirements

5.2 Criteria & Questions

5.3 Results

6 Conclusion

Bibliography

[Al16] Allen, Christopher: , The Path to Self-Sovereign Identity, April 2016.

[AL20] Allende López, Marcos: Self-Sovereign Identity: The Future of Identity: Self-Sovereignity, Digital Wallets, and Blockchain. Inter-American Development Bank, September 2020.

[Be19] Bernabe, J. Bernal; Canovas, J. L.; Hernandez-Ramos, J. L.; Moreno, R. Torres; Skarmeta, A.: Privacy-Preserving Solutions for Blockchain: Review and Challenges. IEEE Access, 7:164908–164940, 2019. Conference Name: IEEE Access.

[Bo19] van Bokkem, Dirk; Hageman, Rico; Koning, Gijs; Nguyen, Luat; Zarin, Naqib: Self-Sovereign Identity Solutions: The Necessity of Blockchain Technology. arXiv:1904.12816 [cs], April 2019. arXiv: 1904.12816.

[Bo20] Bouras, Mohammed Amine; Lu, Qinghua; Zhang, Fan; Wan, Yueliang; Zhang, Tao; Ning, Huansheng: Distributed Ledger Technology for eHealth Identity Privacy: State of The Art and Future Perspective. Sensors, 20(2):483, January 2020.

[Bu20] Bundesdruckerei: , So funktionieren digitale Identitäten, March 2020.

[Ca05] Cameron, Kim: , The Laws of Identity, May 2005.

[CK01] Clauß, Sebastian; Köhntopp, Marit: Identity management and its support of multilateral security. Computer Networks, 37(2):205–219, October 2001.

[De12] Devon, Lofretto: , What is "Sovereign Source Authority"?, February 2012.

[DP18] Dunphy, Paul; Petitcolas, Fabien A.P.: A First Look at Identity Management Schemes on the Blockchain. IEEE Security & Privacy, 16(4):20–29, July 2018.

[DT20] Dib, Omar; Toumi, Khalifa: Decentralized Identity Systems: Architecture, Challenges, Solutions and Future Directions. Annals of Emerging Technologies in Computing, 4(5):19–40, December 2020.

[Eh21] Ehrlich, Tobias; Richter, Daniel; Meisel, Michael; Anke, Jürgen: Self-Sovereign Identity als Grundlage für universell einsetzbare digitale Identitäten. HMD Praxis der Wirtschaftsinformatik, February 2021.

[FCA19] Ferdous, Md Sadek; Chowdhury, Farida; Alassafi, Madini O.: In Search of Self-Sovereign Identity Leveraging Blockchain Technology. IEEE Access, 7:103059–103079, 2019.

[GMM18] Grüner, Andreas; Mühle, Alexander; Meinel, Christoph: On the Relevance of Blockchain in Identity Management. arXiv:1807.08136 [cs], July 2018. arXiv: 1807.08136.

[He07] Hevner, Alan R: A Three Cycle View of Design Science Research. 19:7, 2007.

[Ho20] Homeland Security: , Preventing Forgery & Counterfeiting of Certificates and Licenses – Phase 1 Interoperability Plug Fest Test Plan, May 2020.

[id14] idcubed.org: , ID3 - idcubed.org - The Windhover Principles for Digital Identity, Trust, and Data, November 2014.

[Jo20] John, Anil: , DHS SVIP Blockchain/DLT/SSI Cohort - Multi-Product Phase 1 Interop Artifacts/ Scaffolding / Information, December 2020.

[Jo21] Johnson, Joseph: , Internet users in the world 2021, July 2021.

[Ku20] Kuperberg, Michael: Blockchain-Based Identity Management: A Survey From the Enterprise and Ecosystem Perspective. IEEE Transactions on Engineering Management, 67(4):1008–1027, November 2020.

[Li20] Liu, Yang; He, Debiao; Obaidat, Mohammad S.; Kumar, Neeraj; Khan, Muhammad Khurram; Raymond Choo, Kim-Kwang: Blockchain-based identity management systems: A review. Journal of Network and Computer Applications, 166, September 2020.

[Ma21] Mayer, Peter; Zou, Yixin; Schaub, Florian; Aviv, Adam J: "Now I'm a bit angry:" Individuals' Awareness, Perception, and Responses to Data Breaches that Affected Them. USENIX Security Symposium, 30:18, 2021.

[MDC19] Manu Sporny; Dave Longley; Chadwick, David: , Verifiable Credentials Data Model 1.0, November 2019.

[MG20] Meinel, Christoph; Gayvoronskaya, Tatiana: Blockchain: Hype oder Innovation. Springer Berlin Heidelberg, Berlin, Heidelberg, 2020.

[NJ20] Naik, Nitin; Jenkins, Paul: uPort Open-Source Identity Management System: An Assessment of Self-Sovereign Identity and User-Centric Data Platform Built on Blockchain. In: 2020 IEEE International Symposium on Systems Engineering (ISSE). IEEE, Vienna, Austria, pp. 1–7, October 2020.

[No21] Noor, Poppy: , Should we celebrate Trump's Twitter ban? Five free speech experts weigh in, January 2021. Section: US news.

[OR20] Oesch, Sean; Ruoti, Scott: That Was Then, This Is Now: A Security Evaluation of Password Generation, Storage, and Autofill in Browser-Based Password Managers. In: USENIX Security Symposium. pp. 2165–2182, 2020.

[Or21] Ormandy, Tavis: , Password Managers., June 2021.

[Re21] Reed, Drummond; Sporny, Manu; Longley, Dave; Allen, Christopher; Grant, Ryan; Sabadello, Markus: , Decentralized Identifiers (DIDs) v1.0, March 2021.

[Ri21] Risk Based Security: , Data Breach QuickView - June 2021, July 2021. Section: Featured.

[SNA21] Soltani, Reza; Nguyen, Uyen Trang; An, Aijun: A Survey of Self-Sovereign Identity Ecosystem. Security and Communication Networks, 2021:1–26, July 2021.

[SNE20] Satybaldy, Abylay; Nowostawski, Mariusz; Ellingsen, Jørgen: Self-Sovereign Identity Systems: Evaluation Framework. In (Friedewald, Michael; Önen, Melek; Lievens, Eva; Krenn, Stephan; Fricker, Samuel, eds): Privacy and Identity Management. Data for Better Living: AI and Privacy, volume 576, pp. 447–461. Springer International Publishing, Cham, 2020. Series Title: IFIP Advances in Information and Communication Technology.

[St17] Steel, Amber: , LastPass Reveals 8 Truths about Passwords in the New Password Exposé, November 2017.

[St21] Strüker, Dr Jens; Urbach, Dr Nils; Guggenberger, Tobias; Lautenschlager, Jonathan; Ruhland, Nicolas; Sedlmeir, Johannes; Stoetzer, Jens-Christian; Völter, Fabiane: Grundlagen, Anwendungen und Potenziale portabler digitaler Identitäten. p. 52, June 2021.

[Sw21] Swinhoe, Dan: , The 15 biggest data breaches of the 21st century, January 2021.

[To17] Tobin, Andrew; Reed, Drummond; Windley, Foreword Phillip J; Foundation, Sovrin: The Inevitable Rise of Self-Sovereign Identity. p. 24, 2017.

[To21] Toth, Marek: , You should turn off autofill in your password manager | Marek Tóth, July 2021.

[WH07] Wilde, Thomas; Hess, Thomas: Forschungsmethoden der Wirtschaftsinformatik. p. 8, 2007.

[Wo17] World Bank: , 1.1 Billion 'Invisible' People without ID are Priority for new High Level Advisory Council on Identification for Development, 2017.

A Appendix