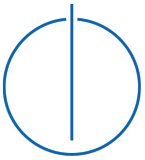
SCHOOL OF COMPUTATION, INFORMATION AND TECHNOLOGY – INFORMATICS

TECHNICAL UNIVERSITY OF MUNICH

Bachelor's Thesis in Information Systems

How to Wisely Identify Natural Subjects in Verifiable Credentials

Evan Christopher



SCHOOL OF COMPUTATION, INFORMATION AND TECHNOLOGY – INFORMATICS

TECHNICAL UNIVERSITY OF MUNICH

Bachelor's Thesis in Information Systems

How to Wisely Identify Natural Subjects in Verifiable Credentials

Vorausschauende Wahl des Mechanismus zur Identifikation von Natural Subjects in Verifiable Credentials

Author:	Evan Christopher
Supervisor:	Felix Hoops
Advisor:	Prof. Dr. Florian Matthes
Submission Date:	October 15th, 2023

I confirm that this bachelor’s thesis in information systems is my own work and I have documented all sources and material used.

Munich, October 15th, 2023 Evan Christopher

Acknowledgments

First and foremost, I would like to thank my advisor, Felix Hoops, who has offered me insightful guidance and invaluable support throughout this thesis. Thank you for taking a chance on me. I would also like to thank Prof. Dr. Florian Matthes for supervising this thesis.

Words fail to express how profoundly thankful I am to my parents for their unwavering belief in me all throughout my life. Your endless love and support have been my constant motivation in striving to be a better person with each and every passing day. Thank you for being there when no one else was, and for making me the person I am. I am, and will be, forever in your debt.

I would also like to extend my thanks to all the wonderful friends I've made along the way. I am grateful to have crossed paths with you.

Finally, thank you all once more for being a part of my life, as I conclude this chapter of my life and start with a new one.

Abstract

Self-Sovereign Identity (SSI) is an emerging concept that is gaining traction in various domains, including e-governance, healthcare, and IoT. It is the next step in the evolution of identity management, a core building block of a lot of applications and software, enabling users to prove claims made about themselves or by others without having to rely on a centralized third party. The goal of SSI is ultimately to give control over one's identity back to the hands of the users, who are natural persons or subjects.

As the adoption of SSI increases and becomes a standard among identity management practices, credentials in ecosystems of different domains will contain more sensitive information and authoritative claims. Thus, the importance of making sure that the credential subject is, indeed, who they say they are, as well as the accuracy of the subject's claims, becomes more prevalent than ever, especially considering recent trends in artificial intelligence. However, the lack of prevailing standards and solutions in the SSI space for this identification challenge poses significant obstacles to the adoption of the concept.

This thesis presents a comparative survey of existing standards and solutions for addressing the identification challenge in SSI, drawing from both white literature and grey literature. We construct a taxonomy of such approaches and examine how identity validation is accomplished by these approaches, providing a qualitative discussion and recommendations based on the newest standards and regulations. Subsequently, we propose a design of an approach based on our findings on the limitations of previous approaches, considered in the context of the GX Credentials project, which is part of the broader Gaia-X initiative. Gaia-X aims to create the next generation of data infrastructure for Europe, its states, its companies, and its citizens with special regard for data sovereignty.

This thesis aims to contribute to the ongoing development of SSI by offering insights into existing approaches and their applicability in real-world scenarios, as well as identifying the gaps and challenges associated with updating information in SSI ecosystems. The findings of this research have implications for enhancing the usability and practicality of SSI systems from various aspects, ultimately contributing to the advancement of digital identity management in the context of SSI.

Keywords: Self-sovereign identity, Verifiable Credentials, Taxonomy, Identity Validation, Natural subjects

Contents

1. Introduction

1.1. Problem Statement and Motivation

Credentials hold a pivotal role in our daily lives and are to be found everywhere – ID cards, driver's licenses, and even concert tickets to name a few. Ironically in our ever-digitalized and connected world, the majority of credentials are predominantly paper-based. The current prevalence of such credentials bears inherent limitations that warrant careful consideration. Susceptibility to theft, loss, and inefficiencies in credential verification are among the many challenges to be addressed by the introduction of digital identity. The centralization opportunities for identity providers such as Google, Meta, and many others are enticing, allowing for the consolidation of user information as well as customized and streamlined verification processes.

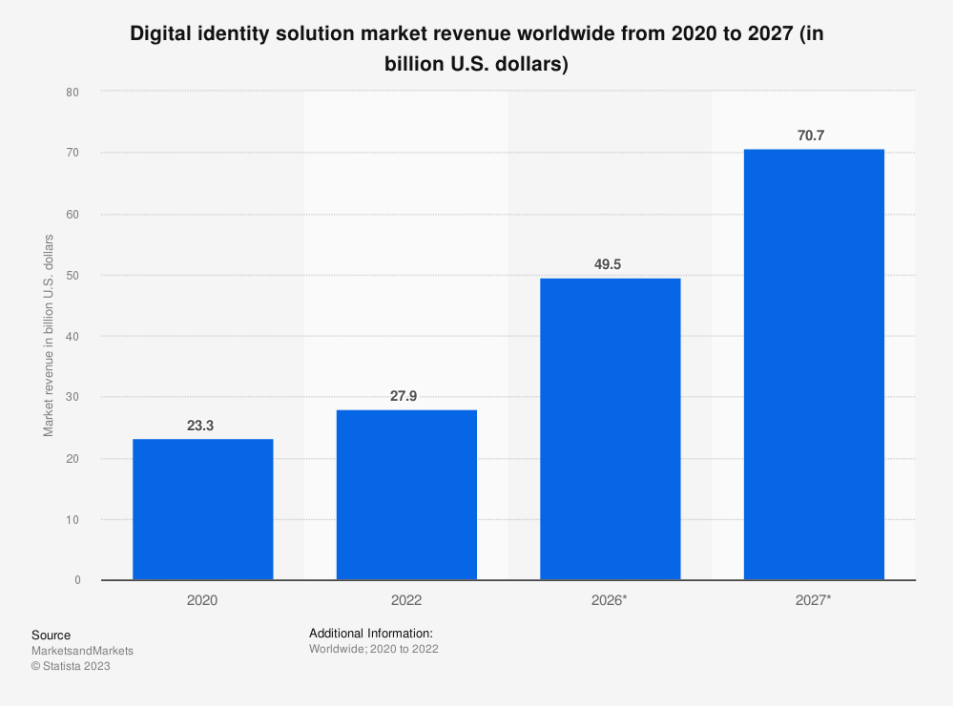


Figure 1.1.: Projected digital identity solution market revenue worldwide by Statista

The shift towards digital identity is reflected in the growth of the global digital identity solution market from nearly 28 billion U.S. dollars to an estimated 71 billion by 2027, driven by increasing instances of identity fraud and data breaches and new government

regulations [1][2]. However, it is evident that the inclination towards centralized solutions has proven to be a source of hacks and breaches, such as the Equifax data breach [3] and the Cambridge Analytica scandal [4], resulting in tremendous costs for all parties involved [5] [6].

As such, decentralized identity management solutions emerge as a compelling solution, with the global decentralized identity market size expected to reach $8.9 billion by 2028 [7]. SSI stands at the forefront of this decentralized paradigm, offering a promising avenue for individuals to regain control over their personal data, utilizing cryptographic keys to verify claims and distributed ledger technologies serving as the source of truth.

The primary target demographic of SSI users consists largely of individuals – natural persons and subjects, which we will occasionally denote as "end users" within the context of our thesis. Consequently, many use cases revolve around generating claims and attestations related to these end users, which entails processing their Personally Identifiable Information (PII) alongside other sensitive data.

With this in mind, the accuracy of such claims about the end user holds equal importance to making the claims more easily verifiable. This includes fundamental assertions such as the end user's name, birth date, and national ID number. Given the swift pace of development and the constantly evolving standards, there appears to be minimal attention directed towards instilling assurance in a presented identity Verifiable Credential (VC), a challenge highlighted by Čučko et al. [8]: *" [...] methods to guarantee a presented identity and, thus, respective VCs belonging to the claimed entity are still missing. [...] The solution here should aim for a strong guarantee of binding the identity to its owner "*.

1.2. Research Questions

Bearing the previously outlined motivation in view, we have formulated three research questions that will establish the cornerstones of this thesis.

For our first research question, we aim to comprehend the status quo of how the end user's identifying information is included in Verifiable Credentials. We will examine the particulars of the included information and the methods employed for its inclusion. We will then engage in qualitative discussions and comparisons of the identified methods.

RQ1. What are the existing solutions and proposals for including identifying information in Verifiable Credentials?

- How are identifying information included in VCs from these existing solutions?
- How do the existing approaches compare?

Building upon the insights from the previous research question, we will consider another mostly overlooked aspect of Verifiable Credentials, namely credential updates. We will explore the status quo, determining the current methodologies for updating data in VCs based on existing standards.

RQ2. How can updates to identifying information be handled in VCs?

As our last research question, results from previous research questions will serve as the foundation for evaluating the current state of the GX Credentials project, namely how identifying information are included in GX Employee Credentials and how they are issued.

RQ3. How can we engineer effective identity credentials within the GX-Credentials project?

- Should the existing schema be revised, and which identity attributes need to be included?
- How can the GX Credentials project be extended to facilitate the revised identity credentials?

Ultimately we aim to contribute to the overall SSI research by examining the status quo on identity binding of Verifiable Credentials to natural persons, which we believe will legitimize the concept and increase its adoption in the long run.

1.3. Thesis Outline

In the subsequent Chapter 2, we will discuss background concepts enabling Self-Sovereign Identity, encapsulating the history of identity management and the blockchain. We then proceed to Chapter 3, where related works to the thesis are briefly discussed. Subsequently, in Chapter 4 we will lay out our methodology for our research. In the following Chapter 5, we will detail the process of creating and presenting a taxonomy derived from the research objects collected in the previous chapter. Then, in Chapter 6, we explore mechanisms designed to facilitate updates of information in VC-based SSI approaches. Keeping the results from the preceding two chapters, in Chapter 7 we aim to evaluate the identity credentials from the GX Credentials project and discuss what needs to be changed to design an effective identity credential. We finally conclude the thesis in Chapter 8.

2. Background

2.1. Identity

Prior to delving into the specifics of the thesis, it is crucial to establish a foundational comprehension of the fundamental concept of identity and trace the trajectory that has brought the concept to its current status quo.

2.1.1. The Concept of Identity

Identity is a multi-faceted concept that can be understood and defined through various lenses within different domains. Depending on the context and discipline, it takes on distinct meanings and interpretations which highlights its complexity. The APA Dictionary of Psychology [9] defines identity as the individual's perceptions of themselves, shaped by their distinct physical, psychological, and interpersonal attributes that differentiate them from others. The feeling of continuation, i.e. an individual remains the same person despite physical changes, is a key aspect. In sociology, the scope of the definition expands to the individual's identity in social groups, including aspects such as ethnicity, race, social class, and roles. These roles are especially important for this paradigm of identity, as they tend to dictate how an individual acts within the given context or role. This phenomenon is comparable to one assuming a role identity, resulting in the merging of the role with the person [10].

When we think of identity, we often think about a person's identity, and, adopting the sociological definition of identity, the set of characteristics and qualities that define the individual and distinguish them from others. This set of characteristics and qualities, commonly summarised as attributes, is not limited to uniquely identifying individuals and their actions, but they also imply granting individuals access to services or resources that others aren't intended to have access to. An example of which can be observed in the context of an individual's age and permission granted for consuming alcoholic beverages. It can be argued that the attributes an individual wields hold power, and to properly manage this power, a way of proving that one owns a certain (set of) attribute(s) is needed. How would one go about doing so?

For this, we will observe the case of physical credentials and back to our previous example of proving one's age before purchasing alcohol. The seller of alcoholic beverages would validate the alcohol buyer's age by checking the buyer's credentials, e.g. a driver's license or an ID card. If the information stated on the card reflects the legal drinking age, the buyer would then be allowed to purchase them. This scenario illustrates proof of ownership, specifically ownership of a credential issued by a trusted party such as the

government in physical form. There are however plenty of downfalls to this approach, such as theft and fraudulent credentials. Additionally, in the context of the internet, presenting physical credentials poses significant challenges such as inconsistency in format and again, counterfeiting of credentials.

With the rise of social media platforms and online services in the Web 2.0 phase, a new paradigm has emerged: proof of knowledge. The usage of usernames and passwords to log in to these platforms is relatively manageable by the service providers and rather convenient for the user. However, when this method is extrapolated to more identity-critical services such as online banking, it fails to provide the necessary level of identity assurance, as it only confirms the intended individual knowing the username and password, but does not assure the identity behind the combination. It depends on the individual – and only that individual alone – on having the necessary knowledge to authenticate the user to the service. Once that piece of knowledge is not exclusively known by the individual, i.e. obtained through unintended means such as data breaches, problems naturally arise.

It has been pointed out by Kim Cameron as the main problem statement of his work “The Laws of Identity” [11], that the internet was built without a way to know who and what you are connecting to, stressing the need for a unifying identity metasystem. To this avail, Cameron proposed seven essential laws of identity which explain the successes and failures of digital identity systems.

Cameron’s first law relates to user-centricity, emphasizing the pivotal role of users in the success of the aforementioned identity metasystem. He states that the system must appeal to the user for its convenience and simplicity, as well as putting the user at the heart of the system, in control of the usage of their identity on the internet, only revealing the user’s identifying information after having given their consent. The second law advocates for minimal disclosure of identifying information, mitigating the risk of breaches by acquiring "least identifying information" on a "need to know" basis. The third law states the necessity for justifiable parties, implying that users disclose their identifying information solely to parties that can justify requiring such information and are recognizable to the user. The fourth law establishes "Directed Identity", achievable through the usage of identifiers with distinct relationship properties when used by public or private entities. This differentiation enables discovery while safeguarding against unwarranted exposure of correlation handles that can be used to assemble profile activities, enabling service providers to collude together to build global profiles of the user [12]. "Pluralism of Operators and Technologies" is mentioned as the fifth law, emphasizing that the universal identity metasystem should comprise multiple, interoperable identity technologies by different identity providers instead of the common monolithic structure of identity systems. As the sixth law, Cameron defines the importance of "Human Integration" in the metasystem. This implies changing the user’s experience to enable users better informed in making decisions related to their digital identity by fleshing out interactions in an unambiguous way. Cameron’s seventh law relates to creating a consistent experience across the user’s different contextual identities, to "thingify" identities and allow users to choose which persona is deemed suitable for a given context or needs of the relying party. Cameron implores all parties working on or

with identity systems to obey these seven laws, and failure to do so would be akin to "if civil engineers were to flout the law of gravity".

In the following section 2.2, we will see how Self-Sovereign identity aligns with these presented laws, serving as the unified identity metasystem he called out for.

2.1.2. The Scope of Digital Identity within the Thesis

Within the scope of this thesis, we will focus on the legal identity of an individual, i.e. natural persons. This identity, although assigned and not entirely inherent, enables individuals access to governmental and private services which are beneficial and pivotal for their well-being and livelihood. Referring back to digital identity, we adopt its definition from [11], defined as a set of claims made about a digital subject by itself or by another digital subject, whereby a digital subject is defined as “a person or thing represented or existing in the digital realm which is being described or dealt with”. In the context of the thesis, we define the intersection of both definitions as natural subjects, meaning a natural person represented or existing in the digital realm.

2.2. A Brief History of Online Identity

Self-Sovereign Identity results from an evolution of online identity paradigms over the years. In this section, we will be traveling toward the path of Self-Sovereign Identity in accordance with Christopher Allen’s work in “The Path to Self-Sovereign Identity” [13].

2.2.1. Centralized Identity

During the early days of the Internet, it emerged as initially as ARPANET (Advanced Research Projects Agency Network), a limited network primarily used by military and academic institutions contracted with the Defense Department to exchange information. This rudimentary form laid the foundation for the modern-day Internet. However, due to its military background, the Internet relied on centralized authorities as issuers and authenticators of digital identity, not to mention the not-easily-scalable nature. As a consequence, the TCP/IP protocol was standardized, facilitating reliable transmission of data in the form of "data packets" across the interconnected network and involving IP addresses. The organization called IANA was established in 1988 as a result of this, to determine the validity of these IP addresses. The need for a better way to keep records of assigned IP addresses to devices became quickly evident. Consequently, the Domain Name System (DNS) was proposed, with ICANN tasked with managing domain names. In the few years to follow, the Internet would evolve further through the development of the HyperText Transfer Protocol (HTTP) and the first ever web browser called WorldWideWeb by Sir Tim Berners-Lee in 1989. Many investors and businessmen became interested in this novelty, sparking tremendous growth in the number of startups. As commerce grew, Certificate Authorities (CA) were needed to help commercial sites establish their true identities on the network in the mid-1990s.

This centralized, hierarchical approach remains problematic, as the very existence of our identities depends on the centralized authority, which has the power to deny a true identity and confirm a false identity, depending on what best suits the purposes of the authority, in this case, the CAs.

Figure 2.1 indicates the relationship between the user and the service provider, also serving as the identity provider.

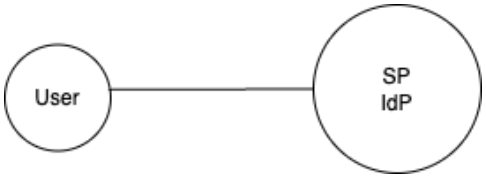


Figure 2.1.: An Illustration of the Centralized Identity Model

2.2.2. Federated Identity

The next evolution in the digital identity management landscape resulted in federated identity. Allen pointed out this phase of identity as administrative control by multiple federated authorities, and an attempt to debalkanize online identity, i.e. an effort to reunify the balkanization of identity due to centralized identities. This phase involves the cooperation of multiple organizations and entities, to allow users to access different services on the Internet with the same set of credentials, usually a combination in the form of a user ID and password. This gives users single sign-on (SSO) capabilities along with various benefits to service providers which include the delegation of managing user attributes to trusted third parties, scalability, and establishment of close relationships with end users [12].

One of the first attempts at establishing the feasibility of the concept was Microsoft's Passport, launched in 1999, a predecessor to Windows Live ID. It aimed to provide all the benefits of federated identity to the whole landscape of web commerce. As critiqued by Allen [13], this puts Microsoft at the center of the federation, making it almost as centralized as traditional authorities. Kim Cameron [11] also mentioned how Microsoft's attempt had failed and extracted lessons from which he formed, among others, his third law of Identity.

The anticipated benefits of federation were not realized, leading instead to the emergence of several dominant entities, resembling an oligarchy. This situation prompted the necessity for a new paradigm for identity.

2.2.3. User-Centric Identity

The subsequent progression takes shape as User-Centric Identity, a movement started before social media networks gained popularity. It is a term the Internet Identity Workshop (IIW) worked on, with the group founded back in 2005. The group is still actively participating in the advancement of identity on the Internet, pushing the idea of decentralized

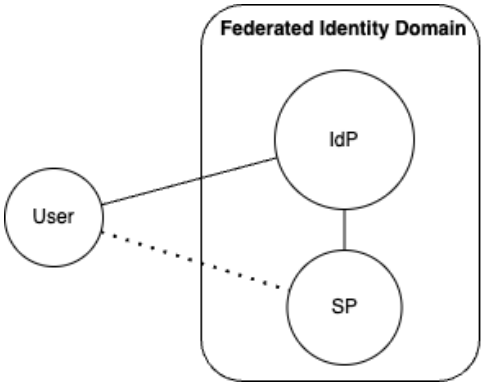


Figure 2.2.: An Illustration of the Federated Identity Model. The IdP acts as a mediator in establishing indirect trust between the user and the SP, illustrated here by the dotted lines. Both the IdP and service provider share the same identity domain, called the federated identity domain, which is created once a notion of trust is established among the IdP and corresponding SP(s) [14].

identity in recent years. The IIW mainly focused on putting users at the center of online identity development, in the middle of interactions with the service provider and the relying party with an emphasis on interoperability and user consent.

IIW's works have become the cornerstone for methods in creating digital identity, with OpenID's Connect (OIDC) Protocol still one of the most used standards in today's identity ecosystem. The OIDC Protocol involves three main actors: (1) the OpenID Provider, an entity or service provider that has implemented the OIDC and OAuth 2.0 protocols. (2) The user, which is a person using a client to access resources. (3) The Relying Party, an application or website that outsources its user authentication function to an identity provider [15]. The protocol cleverly reuses the user's login credentials after registering to an OpenID Provider, providing the user the option of using the same set of credentials to log in to other relying parties. The protocol relies however on SPs being registered with the desired IdPs to function with the already identified and authenticated users. This led to several large silos of valuable sensitive identity information [16].

The FIDO Alliance is another noteworthy initiative, cooperating with companies, governments, and experts in developing technical specifications that describe how a user is authenticated when accessing online services, mainly exploring non-password authentication methods [17]. The FIDO protocol utilizes standard public key cryptography techniques, creating a new key pair during the registration with an online service. The client's private keys are stored locally and the corresponding public key is sent to the online service, to be reused for future logins from the specific user [18].

Since then, a new model for approaching online identity emerged: one that relied on cryptography rather than a centralized trusted party. The seeds of Self-Sovereign Identity were sown and started to take root.

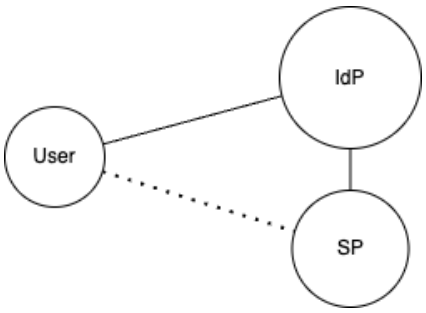


Figure 2.3.: User-Centric Identity Model. Also called the Open-trust model as a single IdP can be shared with multiple SPs. The absence of the federated domain eliminates the need to establish a notion of trust among the entities [14].

2.3. Self-Sovereign Identity

The concept of Self-Sovereign Identity started to gain traction in the mid-2010s, especially in the field of academia [19] following the publication of Allen's paper. In his paper, Allen defined SSI as the next step beyond user-centric identity, emphasizing the significance of user autonomy concerning their digital identity, allowing them to make self-asserted claims as well as claims asserted by other persons about the user. Additionally, he outlined ten principles of SSI, becoming the basis for plenty of research and implementations, and will be elaborated upon in subsection 2.3.2.

2.3.1. A Concise Definition of SSI

Simply defined, SSI is a framework for managing digital identities. It allows users to create, own, and manage their digital identity data without reliance on a third party. Identity data are secured and verifiable with the utilization of public key cryptography, allowing users to selectively share verifiable information across different services. This verifiable information consists of claims or attestations and is either made by the user themselves or by other entities about the user. A credential is simply a collection of these claims or attestations, similar to personal identifying information on a driver's license, with a data field representing a claim, e.g. name, date of birth, and the driving license category.

In the SSI ecosystem, three primary roles are present: issuer, holder, and verifier. The issuer issues credentials about a given entity, referred to as the subject. The holder, on the other hand, stores and manages these issued claims in their digital wallet, akin to storing ID cards in a physical wallet. The subject and holder may refer to the same entity, given that the issued claims were about the holder. This is typically the case, although exceptions occur, e.g. in cases involving guardianship. The verifier, often called the relying party, can be considered as a service provider that requires verification for their services. To this avail, the verifier requests the necessary claims from the holder. The holder then has the option to share the corresponding claims or not, based on their preference. These three roles establish what is called the SSI Trust Triangle, illustrated in figure 2.4.

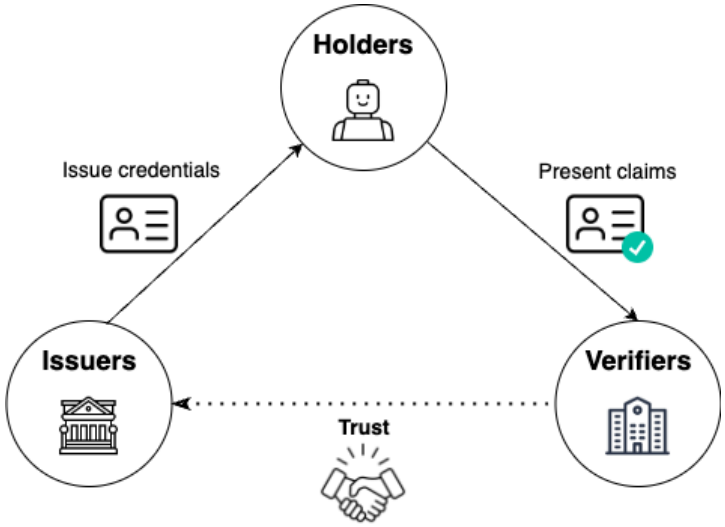


Figure 2.4.: The SSI trust triangle, adapted from [20].

2.3.2. SSI Principles

The following principles were derived from Allen's work [13].

1. **Existence.** Every self-sovereign digital identity presupposes the existence of an associated natural person aligning to that identity, as the self-sovereign identity works by sharing a subset of the person's whole identity. It is impossible for the self-sovereign identity to exist solely virtually as a consequence of this [16].

2. **Control.** Ultimate control over the identity should rest with the user, with the authority to manage their identity however they see fit, including actions such as publication, updating, concealment, and even making claims about other users. It is noteworthy here that while other users can make claims about them, the claim asserting user should remain peripheral to the core identity and not be central to it, unlike in previous identity models.

3. **Access.** Users should be made aware of their own data and have access to it, being able to easily present claims regarding their own identity without being hindered by unnecessary obstacles or gatekeepers.

4. **Transparency.** Algorithms and systems involved in the SSI ecosystem are to be made open-source and independent to increase trust and overall betterment, with the hopes of the flaws quickly being identified and resolved, leveraging the transparency of the underlying parts at play.

5. **Persistence.** Identities should ideally be persisted forever or at least until the user decides against the existence of their digital identity, meaning that SSI should be compliant with the "right to erasure" mandated by articles 17 and 19 of GDPR regulation [21].

6. **Portability.** As one of the consequences of the principle of "Control", the user's identity must be portable, unattached to a third party whose future existence or alignment with the user's interests might be uncertain. This highlights the importance of interoperability.

7. **Interoperability.** Identities should not be limited to single services and should be as widely usable as possible. The ultimate goal is the creation of a global identity, not confined within national borders while remaining under the user's jurisdiction, usable in a multitude of contexts.

8. **Consent.** User consent should be the basis of interactions. The sharing of user claims and other identity-related data occurs exclusively following the grant of consent, granted in a deliberate manner by the user.

9. **Minimalization.** Only the minimum amount of data in claims should be disclosed, e.g. when confirming the user's nationality, the user's national ID card number should ideally not need to be disclosed. The use of selective disclosure, range proofs, and zero-knowledge techniques can be employed for this purpose, with the highest level of minimalization being the Zero-Knowledge Proof.

10. **Protection.** The rights of the user should always be upheld and prioritized, should conflict between the needs of the users and the identity network arise. Therefore, it is essential to implement independent, censorship-resistant, and force-resilient algorithms within the systems, operating in a decentralized manner.

The principles are not set in stone and are meant to serve as a guideline for future standards and implementations. For instance, Naik et al. proposed a revised and extended set of specifications for SSI, splitting up existing principles and including additional ones. Notably, new principles such as Storage-Control, Cost-Free, and Sovereignty were introduced in their extended framework [22]. The proposed specifications were then used to evaluate two SSI solutions in uPort and Sovrin. Čučko et al. conducted a comprehensive analysis of SSI properties or specifications based on numerous studies of SSI principles, focusing on the opinions of experts in the field of decentralized and self-sovereign identity management from different domains. In total, 18 properties were identified and subsequently categorized. The perceived level of importance of each property was established, and additionally validated by experts [23].

2.3.3. On Sovereignty

Furthermore, it's important to provide a specific explanation for the term "sovereign" as used in "Self-Sovereign Identity." Naik et al. [22] outlined three distinct contexts of sovereignty: sovereignty in the real world, in cyberspace, and in digital identity. These contexts will be explained in the following text.

Sovereignty in the real world encompasses two main facets: state sovereignty and citizen sovereignty. State sovereignty involves territorial authority and jurisdiction, both of which

empower the state to exercise supreme control over all entities within its borders. Citizen sovereignty entails inherent rights under state laws, including juridical equality, social freedom, and political autonomy.

In contrast to the former, Sovereignty in cyberspace is much more complex, as it is not limited to only one state, a consequence of the internet having no fixed territorial boundaries. However, cyberspace is built upon numerous physical infrastructures spread over the globe. Hence, physical infrastructure located within a state will automatically be governed by it, as well as any cyber events that occur. Therefore, the question of who holds access or control over any network is irrelevant in this context. Nevertheless, it remains a challenging problem due to the fragmented nature of states, with no universally accepted sovereignty principle throughout the world.

The final context is the sovereignty of digital identity. Digital identities are issued to real-world entities. As such, their sovereignty is dependent on various factors, among others state laws and regulatory schemes within the SSI network itself.

Consequently, sovereignty could not be treated as a binary characteristic, implying that an identity is either fully sovereign or not at all. This implies different levels of sovereignty for the identity, depending on the state regulations for identity-related disputes and safeguards.

2.4. The Fundamental Components of SSI

We briefly discussed the SSI trust triangle in section 2.3.1, the workflow that is central to the SSI concept for establishing trust between issuers and verifiers in an indirect manner. Enabling this workflow are the three main components of SSI, which are Verifiable Credentials, Decentralized Identifiers, and the so-called Verifiable Data Registry. VCs and DIDs are standardized and officially recommended by the World Wide Web Consortium (W3C), a consortium developing standards and guidelines for the web, based on the principles of accessibility, internationalization, privacy, and security [24]. The Verifiable Data Registry is also crucial to this workflow, acting as a platform to manage DIDs. In the following subsections, we will dive deeper into each component and discuss them on a technical level, as well as explain their role in the overarching system.

2.4.1. Verifiable Credentials

As mentioned in section 2.3.1, credentials are commonly present in the physical world and are made up of claims. Preukschat and Reed defined credentials as “[...] any (tamper-resistant) set of information that some authority claims to be true about the subject of the credential—and which in turn enables the subject to convince others (who trust that authority) of these truths” [25]. This includes documents such as birth certificates, diplomas, and passports.

Naturally, they need to be verifiable in some way. In physical credentials, this is done by proof of authenticity directly present in the credential itself. This would then be

checked by the verifier either internally, through outsourcing to other document verification companies, or by simply contacting the issuer of the credential itself. This process takes time and goes through several intermediaries, leading to bottlenecks and long verification times. Furthermore, the digitization of physical credentials that are often needed in such workflows, when presenting or handing a copy of the authentic credential is not possible, causing not only a cumbersome workflow for both the holder and verifier, but also privacy concerns as the entirety of the credential needs to be presented. It is not possible to only disclose the necessary information as it wouldn't easily be verifiable otherwise.

The introduction of Verifiable Credentials is intended to address these issues. VCs are digital credentials that can represent all of the same information that a physical credential represents in a digitally verifiable and tamper-evident manner based on cryptography. A Verifiable Credential consists of three main parts: *Credential Metadata*, *Credential Subject/Claim(s)*, and *Proof(s)*. The *Credential Metadata* describes the properties of the credential and identifies it. Alongside the metadata is the payload of the credential, namely the *Credential Subject*. It is based on a certain schema and contains a set of claims describing the subject of the credential. Lastly, the *Proof* contains information regarding how the credential is digitally signed by the issuer and is used by the relying party to validate the credential.

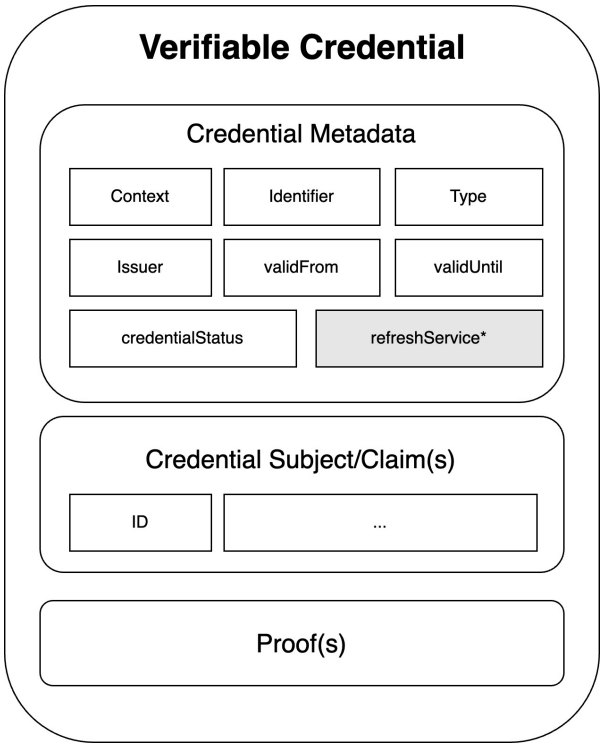


Figure 2.5.: An Overview of Verifiable Credential Components, adapted from [26, 27].

The following list further explains the properties of VCs according to [27]:

Context Since a VC is a JSON-LD file, the `@context` property is an array and must be defined, as it is needed to map the attributes present in the current VC to the correct format, specified via a base context URL `https://www.w3.org/ns/credentials/v2`, the first value of the array. Other context URLs can be added.

Identifier A globally unique identifier used for specifically identifying the credential in the form of a URL that may be de-referenced.

Type The `type` attribute in the form of an array is used by the verifier to help determine whether or not the credential can be processed or not, i.e. if the credential has the correct attributes that the verifier expects to find. The first value must always include the type `VerifiableCredential`.

Issuer The `issuer` property expresses the issuer of the VC. It must either be a URL or an object containing an `id` property. Additional attributes alongside the `id` about the issuer can also be included when expressed via an object.

validFrom The `validFrom` property helps issuers to express when a credential becomes valid.

validUntil In contrast to `validFrom`, the `validUntil` property helps issuers to express when a credential expires or ceases to be valid.

credentialStatus If the optional the `credentialStatus` attribute is present, it specifies whether the credential is suspended or revoked. More information for status schemes is further defined in this specification [28].

refreshService This optional attribute enables an issuer to include a link to a refresh service once the credential has expired and update it. This could be done either by a URL or an object containing an `id` attribute. Additionally, the type must be specified as `ManualRefreshService2018`. As of writing the thesis, this attribute is at risk of being removed from future iterations of the specification, hence the gray background and asterisk in figure 2.5. However, it would still be included as a reserved extension point in the specification, to accommodate possible future implementations. The relevancy of this attribute will be discussed in further detail in Chapter 6.

Credential Subject The `credentialSubject` property must be present in a VC and is an object containing claims made about one or more subjects. Each subject must have an `id` attribute to identify the subject. The claims must correspond to the structure defined in the context URL [26].

Proof The mandatory `proof` property is used to prove the integrity of

the information in a VC, making it tamper-evident. Cryptographic proofs include digital signatures and zero-knowledge proofs and are expressed via an object that may contain information related to the cryptographic suite, such as a digital signature and metadata related to it, e.g. metadata regarding the public key associated with the signature.

When presenting credentials to the relying party for verification, e.g. in the case of applying for a job, one would typically need to present multiple credentials from different issuers. Instead of presenting the credentials as is, the holder has the option to aggregate them and present them in the form of a ***Verifiable Presentation (VP)***. A VP can express data from multiple VCs and contain additional data, all encoded as JSON-LD. It is also possible to create VPs derived from a claim on a VC, thus creating indirect proof without revealing all claims in that VC. An important thing to note is that VPs are meant to be short-lived, bound to a challenge provided by a verifier. In general, VPs cannot be assumed to be correlated with the presented VCs it contains [27].

The VCs contained in a presentation are typically about the same subject and issued by multiple distinct issuers. Figure 2.6 provides an overview of how VPs are generally structured.

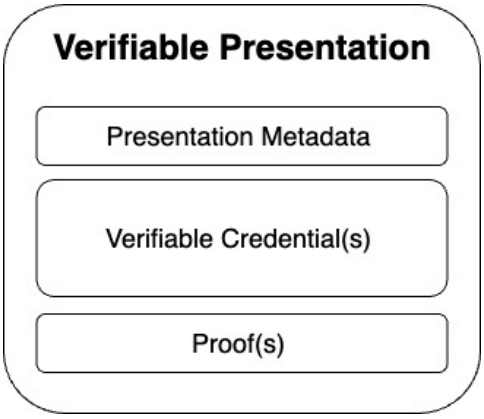


Figure 2.6.: An Overview of Verifiable Presentation Components, adopted from [27].

Moreover, as laid out by [29], VPs include a form of proof of ownership of the credentials that they contain. Some examples include JSON Web Signature (JWS) when using JSON Web Tokens (JWTs), Linked Data Signatures for Linked-Data Credentials, or Camenisch-Lysyanskaya ZKPs in the case of Anonymous Credentials.

2.4.2. Decentralized Identifiers

Decentralized Identifiers (DIDs) constitute another foundational pillar of SSI. In contrast to conventional identifiers like usernames and email addresses, which are issued and ultimately governed by a centralized or federated entity, DIDs are designed to be managed

in a decentralized manner, decoupled from centralized registries and federated identity providers. DIDs enable verifiers to help discover information related to a DID but are constructed in a way that enables the controller of the DID, i.e. the subject that has ownership over the DID, to prove control over it without having to rely on another party [30]. The DID subject meanwhile is the entity defined by the DID. It is important to note that the subject and controller might be the same entity. However, this is not always the case, as it is possible for a DID to have multiple controllers and subjects.

From a more technical perspective, DIDs are a new type of Uniform Resource Identifiers (URIs) and possess the functionality of both URLs and URNs in that it is globally resolvable and globally unique.

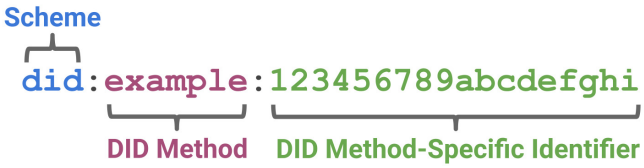


Figure 2.7.: The general scheme of a DID, adopted from [30].

The formal syntax of a DID is illustrated in figure 2.7. It is composed of three parts:

1. ***Scheme*** the `did` URI scheme identifier
2. ***Method*** the `did` a globally unique identifier specifying a DID method, e.g. *ebsi*, *sov*, or *key*.
3. ***Method-Specific Identifier*** a method-specific identifier for the DID that is unique within a DID method. Additional URL arguments akin to that of web URLs can be added here to retrieve resources or a specific part of these resources. These arguments include the DID path, query, fragment, and specific parameters [30].

A DID resolves and associates a DID subject to a JSON-LD document containing metadata regarding that DID, called the *DID Document*. The metadata it contains are typically verification methods, i.e. information regarding cryptographic suites or key types including the corresponding public keys, and services, i.e. means of interacting with the DID subject via service endpoints, e.g. file storage, social networking, and VC repository services.

Furthermore, possible operations related to the manipulation and creation of DID documents and DIDs are specified via *DID methods*. These operations comprise *DID Resolution*, which encompasses create and read operations, as well as *DID Revocation*, which encompasses update and delete operations [31].

A DID method is defined by implementers and is often associated with a specific type of verifiable data registry. Implementers must provide a corresponding specification of their DID method implementation, enabling interoperability between various implementations of the same DID method [30]. Over the years there have been plenty of such implementations,

with Hoops et al. [32] identifying around 160 distinct DID methods. The proliferation of DID methods provides encouraging signs of the growth of SSI and might be attributable to the fact that W3C's DID v1.0 specification [30] does not force the usage of specific technologies or cryptographic methods upon implementers of a DID method, nor does it forbid the creation of DIDs based on existing federated or centralized IdMs, hence creating a bridge between different identity management paradigms.

Naturally, with all of their inherent properties, DIDs are associated with VCs for the purpose of identifying the VC holder in SSI ecosystems and thus are linked to identities. The combination of both provides a solid foundation for enabling SSI and its goals. Nonetheless, it is essential to consider the significance of the verifiable data registry.

2.4.3. Verifiable Data Registry

As defined in [27, 30], a *Verifiable Data Registry (VDR)* is a system in which operations related to DIDs and VCs are facilitated. This includes the creation and verification of identifiers, keys, and other relevant information such as VC schemas, issuer public keys, and revocation registries. Examples of VDRs include distributed ledgers, decentralized file systems, databases, peer-to-peer networks, and other forms of trusted data storage [30].

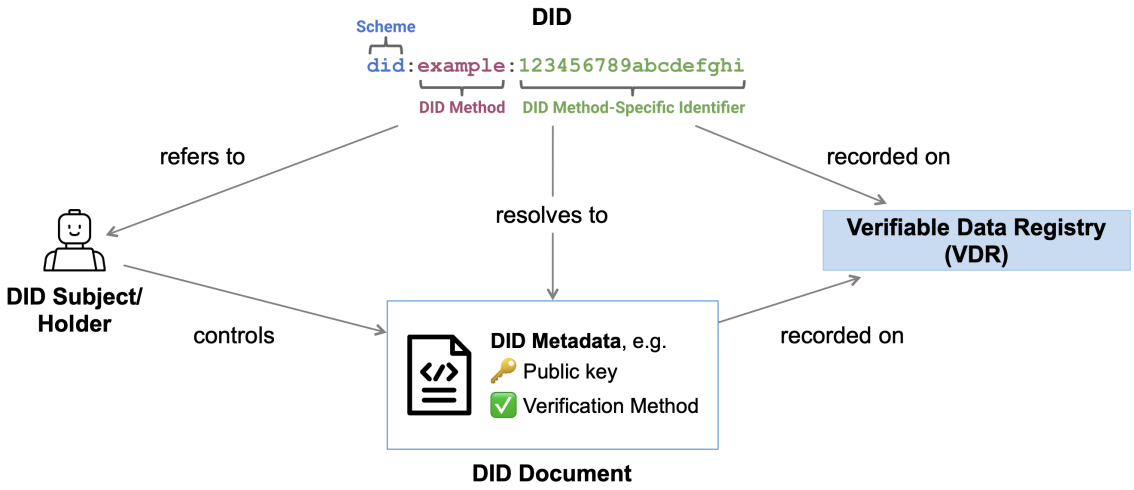


Figure 2.8.: DID architecture overview and its relation to the VDR, adapted from [30].

Most importantly, the VDR serves the purpose of an anchor to DIDs. The public keys associated with a DID are stored in its corresponding DID document, alongside which other relevant information is stored on the VDR, publicly available for verifiers to use when verifying credentials presented by the specific DID holder. This relationship is depicted in figure 2.8, and extends to the lifecycle of VCs, used by issuers for registering their public DIDs as well as verifiers for verifying the issuer of presented credentials, illustrated in figure 2.9.

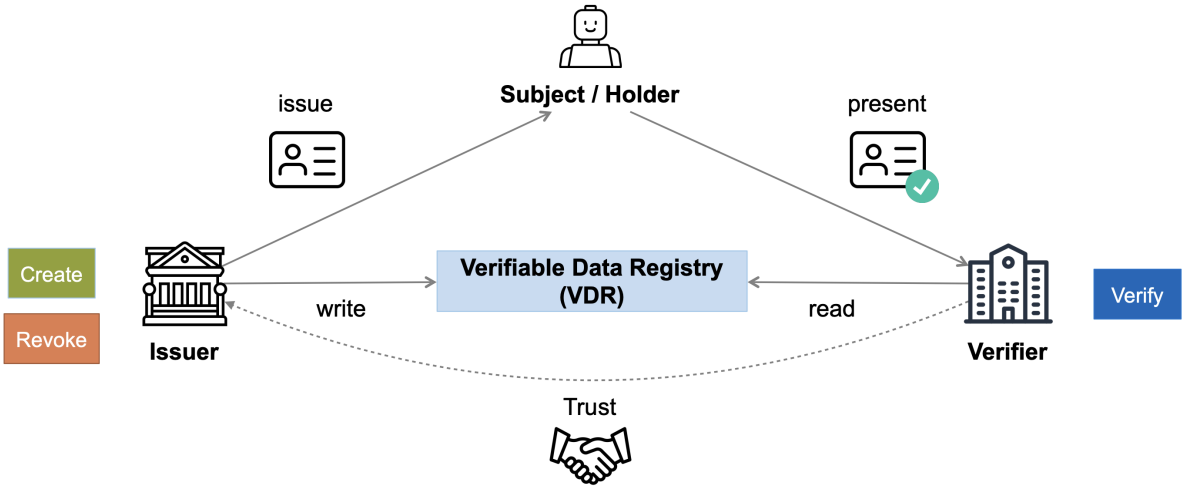


Figure 2.9.: An illustration of the VC Lifecycle and the role of the VDR

According to Mühle et al. [19], there are two competing models for the VDR, the first of which is the *Identifier Registry Model*. In this model, the blockchain or VDR acts as a replacement for the registration authority found in typical IdMs, and maintains the pairing of identification and authentication, the mechanism of which operates as we explained earlier in this section. The second model is an extension to the previous model and is described as the *Claim Registry Model*. Here, the VDR not only serves the purpose of storing DIDs of identities, but it is also for storing all the associated claims of the aforementioned identities. These claims are stored in the form of cryptographic fingerprints or hashes, along with definitions of VC schemes.

As such, the VDR must be trusted by all roles in the SSI ecosystem and it is crucial that it remains secure and tamper-resistant. For this reason, VDRs are commonly blockchains or distributed ledgers. Having stated this, it's important to note that blockchains may not always be essential, as non-blockchain-based solutions still fulfill the majority of the SSI properties laid out by e.g. C. Allen [13] and other researchers. It is noteworthy, however, that the examined blockchain-based solutions definitely still meet more SSI properties on average compared to their counterparts [33, 34].

3. Related Work

In the following sections, we will present and briefly discuss related work to this thesis, as outlined in the previous chapters, which explores how the identities of natural subjects are anchored in various SSI approaches. This encompasses both academically proposed system designs and frameworks, as well as implementations or commercially available solutions. Hence, related work in the rather general field of decentralized identity management will be considered, primarily work in the form of surveys, augmenting the literature review and compilation of SSI approaches. These surveys are presented in section 3.1. Within the literature review, we have compiled a range of non-SSI approaches, along with SSI approaches featuring identity proofing and other binding-related processes that may not exclusively rely on Verifiable Credentials. A selection of these approaches is outlined in section 3.2. Furthermore, considering that one of the main deliverables of the thesis is a taxonomy, section 3.3 provides an overview of previously constructed taxonomies for SSI, including noteworthy comparison matrices and overviews. In sections 3.4 and 3.5, we examine some regulations, standards, and specifications pertinent or related to SSI and the identification of natural subjects in credentials. Concluding the chapter, section 3.6 showcases relevant SSI initiatives relevant to the thesis.

3.1. Conducted Surveys on SSI Approaches

Bai et al. [35] summarize the advantages and disadvantages of the models from the four stages of the evolution of digital identity in the form of a comparison matrix to outline the importance of decentralized identity. Five comparative research objects are briefly examined and compared. Their research identifies four challenges in constructing an effective SSI architecture: good user experience, regulation, the right to be forgotten, and the commonplace conflict between user data privacy and enterprise data realization.

Kaneriya et al. [36] provide a comparative analysis of six blockchain-based SSI systems, mainly focusing on their architectural components. Several use cases for SSI are presented, including a use case where VCs are used to attest to claims in official identification such as an Aadhar card, a unique identification document issued by the Indian government for its citizens, including biometric and demographic data.

Badirova et al. [37] evaluate SSI concepts and applications including eIDAS, GAIA-X, Trust Over IP, and ESSIF, constructing a comparison matrix to this avail. Their work mainly focuses on the broader outlook of identity and access management concerns in a cloud setting, discussing access control models in the cloud as well as the application of artificial intelligence (AI) and machine learning (ML) in access control. They conclude that

decentralized identity is the future but has some issues such as the lack of integration with conventional IAM systems and problematic identity recovery.

Kuperberg [38] conducted a systematic criteria-driven survey of 43 market offerings and concepts of blockchain-based IAM solutions, based on a comprehensive set of requirements amounting to 75 evaluation criteria. He concluded that none of the offerings satisfies all of the mandatory criteria, most notably the lack of certification and standard-compliant interfaces such as OAuth or SAML for service providers to integrate. In his work, he noted that he failed to come across quantitative statements regarding overhead or performance in general, adding that it's too early to compare the established solutions with the relatively new blockchain-based solutions. Furthermore, he states that the verification of identities can not be seen as a unique proposition value to blockchain-based identities, as conventional identities can be extended to "trusted" ones via PostIdent and WebID.

Soltani et al. explore the origin of identity, defining digital identity and the evolution of SSI, examining relevant research initiatives, platforms, projects, regulatory framework, SSI components, and privacy engineering protocols. Additionally, their research evaluates a total of ten SSI platforms and lists various challenges in SSI research, such as finding the right balance of centralization and decentralization within the context of SSI to properly support the vision and requirements of its predecessor, the user-centric identity model.

Ahmed et al. [39] present an extensive literature review of state-of-the-art SSI approaches in academic literature as well as market offerings concerning the applicability of blockchain-based SSI solutions. According to their findings inspired by [40], IdMs based on DLT fall mainly into the two following categories, SSI and decentralized trusted identity (DTI). They explain that the main difference between both lies in how identity verification is handled, with DTI relying on mandatory identity proofing executed by centralized services to check government-issued IDs and keep cryptographic proof of validated data on a digital ledger for other parties to verify. In contrast, SSI does not necessarily rely on existing documents, offering users self-attestation of their identity information or forming their identities by gathering credentials from different issuers. However, distinguishing between these two categories is somewhat vague, as we have encountered several solutions that offer similar identity-proofing processes while still being marketed as self-sovereign identities. This validates the observation presented by Sedlmeier et al. [41], that "SSI" is a term given to very different projects. They highlighted that companies whose solutions may not necessarily involve SSI components adopt the term in their marketing, indicating SSI's ambiguity and lack of commonality. In the context of this thesis, we will use the definition stated in the preceding chapters, specifically section 2.3.

3.2. Identity Proofing and Binding-Related Processes

Toth et al. [42] evaluate research that outlines SSI properties and highlight the lack of properties that cover situations where the identity owner loses control of their digital identities, caused among others by weak identity verification and assurance. To this avail, several new SSI properties are added, including "Identity Assurance", according to which

relying parties must be provided assurances that digital identities truthfully characterize their owners rather than imposters. This is similar to how third-party identity proofing works in the physical world, wherein digital identity owners can submit their information to an issuer for verification and receive an attestation, consequently establishing remote identity assurances by linking the issuer's attestation to their digital identity. For this purpose, Toth et al. propose the use of a proof-of-existence identity registry that stores hashes of digital identities, allowing the verification of digital identities and their associated public keys, much like the "Identifier registry model" presented by Mühle et al. [19].

SelfKey [43] offers such a binding, aiming to "offer individuals a secure means of verifying their identity through AI-Powered proof of individuality methods" [44] by the establishment of a so-called *SelfKey iD*. To this avail, a mandatory KYC check will be conducted, typically involving selfie checks and the provision of personal information with government-issued identification documents, and is paid for by the user [45]. Following a successful check, a *Soulbound Token (SBT)*, i.e. a non-transferable NFT (Non-Fungible Token), is minted and serves as proof of identity in various contexts. The SBT might also be used in tandem with other identity credentials such as KYC credentials to provide a higher level of assurance for the user's identity. The aforementioned KYC credentials may be issued as VCs to the user by SelfKey network issuers, however, this feature is still under development as of writing this thesis. Alternatively, users can stake SelfKey's *KEY*, the main utility token within the SelfKey ecosystem, against their credentials, establishing so-called "MetaProofs" that can be used in place of KYC credentials and for minting reputation signaling tokens [46].

Zichichi et al. [47] seek to balance real user identification and minimalization of identity data within a use case in the Metaverse, i.e. Decentraland, specifically where a user's age needs to be verified before accessing an age-restricted movie screening in a decentralized cinema. They aim to design a system that enables this verification in compliance with eIDAS 2.0, implying the use of VCs, DIDs, and DLTs and emphasizing the importance of selective disclosure. Their implementation is based on a smart contract that enables on-chain verification of credentials based on ZKPs. They conclude that blockchain-based Metaverse platforms must integrate with other legally recognized instruments of online identification, as DLTs only guarantee data integrity and do not necessarily ensure the identifiability of blockchain users. Additionally, they acknowledge the high transaction cost of their smart contract implementation when verifying ZKPs being a limitation that should be optimized with further research.

Related to the work of Zichichi et al., Cali et al. [48] propose a decentralized SSI-based IdM system for decentralized virtual worlds, i.e. Metaverse, emphasizing legal and legislative aspects to the proposal as well.

Tahlil et al. [49] propose *AlgoCert*, a credential verification solution incorporating blockchain, SSI, and DIDs to tackle credential fraud, such as fraudulent certificates. In this system, institutions, i.e. issuers, leverage stateful smart contracts on the Algorand blockchain to store certificate documents in the InterPlanetary File System (IPFS) and use the unique content ID (CID) from IPFS, signed by both the prospective holder and issuer, to create an NFT certificate. This enables the credential holder to simply present a link to the NFT certificate when required by relying parties. As a conclusion to their research, they highlight that AlgoCert addresses concerns regarding unwanted transfer of ownership for important credentials, but acknowledge room for future research in scenarios where the holder's private key is lost.

Abubakar et al. [50] leverage VCs, DIDs, and smart contracts to enable on-chain verification and validation of COVID-19 vaccination certificates. Additionally, IPFS is used to store an encrypted copy of the vaccination certificate that was issued as a VC to avoid expensive writes to the blockchain. The resulting IPFS hash is kept in the smart contract and mapped to a public key pair, i.e. the citizen's digital identity. They conclude that their solution enhances security and user privacy but emphasize the significant drawbacks of high transaction costs associated with the smart contract deployment that warrant further consideration.

Worldcoin [51] was founded to create a globally inclusive financial network and identity, with proof of personhood being the main idea behind their digital identity network, *World ID*. World ID enables users to verify their humanness online by registering their biometric data instead of using other forms of verification, emphasizing the global issue of over 4.4 billion people lacking a digitally verifiable ID, restricting their access to financial and social services [52]. The user's biometric data is scanned with a custom device called *the Orb*, a device which scans the user's face and iris and computes what they call *The Iris Code*, a numerical representation of the texture of a person's iris. The Iris Code is then evaluated by a *uniqueness service* for deduplication, upon the success of which a new identity commitment is added to the on-chain Merkle tree of a smart contract. Worldcoin initially refrained from using biometrics but eventually opted for iris scanning, due to its superior accuracy compared to other biometrics, with false match rates beyond 2.5×10^{-14} [53]. Furthermore, they emphasize that the iris scan is not used to identify users but solely to confirm the user's uniqueness, adding that images collected by the Orb are promptly deleted unless specifically requested otherwise [54]. Biometric data, therefore, are only used for sign-up, with ZKPs and the *World ID Protocol* playing significant roles in every other operation. Essentially, the protocol is based on a list of public keys stored on-chain, each of which corresponds to an identity commitment, i.e. a private key generated by *World App*, the crypto wallet exclusive to World ID, in the sign-up process. Using these public key pairs to represent an identity is similar to the approach used in SSI. As of writing this thesis, Worldcoin has garnered worldwide attention, amassing over 2.25 million sign-ups

across 120 countries. Additionally, it is anticipated that World ID will incorporate support for VCs and DIDs in the future [55].

3.3. Taxonomies

Frederico Schardong and Ricardo Custódio [16] conduct a rigorous, reproducible systematic review and mapping of SSI, examining both conceptual and practical advances in SSI from over 80 sources of literature. They do not review and map standalone SSI solutions but rather work that aims to solve pragmatic issues about any aspect of the SSI ecosystem. Their survey culminates in a proposed taxonomy for SSI with two facets: conceptual and practical. Consequently, 69 SSI approaches are evaluated under the practical facet consisting of three multi-layered dimensions, encompassing a total of 21 characteristics. *Identity Verification* and *Identity Assurance* are among the characteristics present in the conceptual facet of their proposed taxonomy, with only two out of seventeen conceptually inclined works being assigned to owning either one of these characteristics, highlighting the lack of research in this area of SSI. This thesis focuses on these aspects of the SSI ecosystem, analyzing standalone SSI solutions in the process.

Schmidt et al. [56] presents a taxonomy for SSI based on a systematic grey literature review. Their goal is to clearly elaborate the members of an SSI ecosystem, classifying 147 SSI ecosystem members with a taxonomy consisting of four dimensions to derive patterns, resulting in eight SSI archetypes. An example would be the *Standard Setter* archetype with W3C as a representative member.

Bochnia et al. [57] examines a different aspect of SSI, focusing on credentials within the context of organizations and the consequent mapping of VCs to physical credentials. For constructing their taxonomy of credentials in organizations, the *Extended Taxonomy Design Process (ETDP)* [58] is employed. This method is an extension of the methodology for taxonomy building proposed by Nickerson et al. [59] that is used in this thesis. Subsequently, the resulting taxonomy dimensions are mapped to VCs, with dimensions such as *Modifiability* claimed to not be supported by the official W3C VC specification. Moreover, the implications of the mapping are discussed, mentioning the importance of PII handling in VCs. Concluding their research, they assert that VCs already possess the majority of characteristics found in physical credentials, with varying implementation and support for certain features depending on the vendor, and underscores the importance of further standardization for VCs within the context.

Kölbel et al. [60] aim to identify business model characteristics in order to distinguish enterprises leveraging SSI ecosystems. To this avail, they present a taxonomy of business enabled by SSI comprising 12 dimensions, nine sub-dimensions, and 51 characteristics. The taxonomy is constructed with the methodology proposed by Nickerson et al. [59] and is based on a final set of 18 active *Businesses Enabled by SSI (BESSI)* selected from CrunchBase's new venture database. The taxonomy encompasses dimensions such as *Regulatory Compliance* and *Customer charge*, which indicate how a consumer pays for a BESSI, including *cost-per-transaction* and *subscription models* as characteristics of the

dimension.

One of the white papers generated from the 11th annual Rebooting Web of Trust (RWOT) workshop by Kudra et al. [61] introduces a comparison matrix for various credential formats, encompassing VCs, AnonCreds, and the ISO-standard Mobile Driving License (mDL). Additionally, a variety of *credential profiles* were evaluated, with a credential profile being a configuration of the credential format, signing algorithm, revocation algorithm, and key management. These profiles were compiled by a group of domain experts during the duration of the workshop. The white paper explains the reasoning behind the properties in the matrix and serves as an application guide for making technical and non-technical decisions. The comparison matrix itself is an active document in the form of a spreadsheet in Google Sheets [62], accessible by anyone and maintained by technical experts. Ultimately, it is concluded that the creation of the comparison matrix fulfills all functions and provides value to the selected stakeholders.

TNO, an independent not-for-profit research organization in the Netherlands [63], presents an overview of SSI wallets and their characteristics, offering insight into what features are specifically offered by the wallets, as well as their interoperability with each other [64]. The list of wallets present is aggregated from the inputs of wallet vendors and TNO's contribution in the aforementioned RWOT11 on credential profiles. SSI wallets are characterized by multiple properties in the overview, including *credential format*, *revocation mechanism*, and *selective disclosure* [65]. In another GitHub repository [66], TNO created an SSI Standards Overview graphical overview page [67] accompanied by documentation [68] and structured in accordance with ToIP's Technology Stack [69]. These documents are, however, a first draft and should be considered a work in progress.

3.4. Regulations

3.4.1. GDPR

The General Data Protection Regulation (GDPR) [70] is a European Union (EU) regulation that governs data protection and privacy for individuals within the EU, defining strict guidelines for how organizations handle and process personal data. Its ultimate goal is to grant individuals, i.e. natural persons, greater control over their data. Consequently, organizations are required to ensure data privacy and security, with non-compliance being punished with substantial fines imposed by the regulation.

Decentralized digital identity is directly encouraged in Recital 7 of the GDPR, in which it is highlighted that "natural persons should have control of their own personal data". SSI, as an instance of this paradigm of identity management, is theoretically GDPR-conformant due to its architecture and its promise of data sovereignty. Identities in an SSI ecosystem are effectively represented by public key pairs, DIDs, and their corresponding DID documents. Consequently, DIDs that are used to manage data that refer to natural persons fall under the category of online identifiers associated with natural persons elaborated in Recital 30. Furthermore, the SSI architecture may lead discussions about whether data subjects

(credential subjects) can be considered data controllers (credential holders) when it comes to their own data within an SSI ecosystem [71].

3.4.2. eIDAS

The EIDAS (electronic Identification, Authentication and Trust Services) [72] regulation establishes a framework for electronic identification of natural and legal persons in the internet and trust services for electronic transactions in the EU. The goal of eIDAS is to promote eID interoperability across all 28 member states, simplifying identification for cross-border administrative services in the EU in the public and private sector. To this avail, the regulation has two primary concerns. Firstly, electronic identification or eID, meaning identity proofing for users looking to gain access to a service. A legal basis is therefore defined for mutual recognition of distinct eID implementations among all member states. All member states are mandated to “notify” their national eID scheme since September 2018. With eIDAS, the eID is categorized into three levels: "low," "substantial," and "high," depending on the required level of assurance. Secondly, eIDAS specifies various trust service solutions in the private sector, with the aim of improving trust and efficiency in administrative business processes in general. The specified solutions include such as eSignatures, eTimestamps, Qualified Web Authentication Certificates, eSeals, and Electronic Registered Delivery Services (ERDS).

With SSI gaining traction, questions regarding the malleability of the regulations are put forward. Dr. Ignacio Alamillo Domingo published an SSI eIDAS legal report [73] concerning the status quo of the matter, analyzing the compatibility of SSI and eIDAS trust framework. His assessment of the legal aspect of scenarios related to VCs is noteworthy, with him proposing the use of notified eIDAS and qualified certificates to issue VCs. In another scenario termed *eIDAS Bridge*, he proposed enchanting the legal certainty of any type of VCs by incorporating the issuer’s advanced or qualified electronic signature. He concluded that it is possible for SSI to be eIDAS compliant.

As part of the European Commission’s effort to realize set milestones in their 2030 Digital Compass [74], a new version of the regulation, i.e. eIDAS 2.0, was proposed in 2021 and introduced major changes, most noteworthy of which is the European Digital Identity Wallet (EUDIW). Every member state is encouraged to create a digital wallet in accordance with the regulation to enable all their citizens, and ultimately all EU citizens, to have identifiable digital identities. The proposal’s relevance is gaining traction, with the EU Commission releasing a provisional political agreement on the key elements of EUDIW and investing €46 million in various pilot projects [75, 76]. It remains to be seen how the new regulations handle privacy and security challenges, such as traceability and profiling concerns due to the provision of a single, permanent identification for users [37].

3.5. Standards and Specifications

3.5.1. NIST SP 800-63

The National Institute of Standards and Technology (NIST) is a U.S. government agency responsible for developing and publishing standards and guidelines, primarily for federal agencies and organizations operating in the United States. Relevant to this thesis is the *NIST SP 800-63 Digital Identity Guidelines* [77], which presents processes and technical requirements necessary for meeting digital identity assurance levels. The current draft consists of four volumes:

SP 800-63 Digital Identity Guidelines Provides the risk assessment methodology and an overview of general identity frameworks [78].

SP 800-63A Enrollment and identity proofing Provides a set of guidelines and requirements for the enrollment and verification of an identity in the use case of digital authentication, central to which is the identity proofing process by the Credential Service Provider (CSP). Additionally, technical requirements for each of the three identity assurance levels resulting from the proofing process are defined [79].

SP 800-63B Authentication and lifecycle management Provides recommendations on types of authentication flows and processes for all identity assurance levels, including authenticator lifecycles and invalidation in case of loss or theft [80].

SP 800-63C Federation and assertions Provides technical requirements on the usage of federated identity architecture and the corresponding assertions for implementing identity federations. Moreover, privacy-enhancing techniques such as selective attribute disclosure are fleshed out [81].

It must be emphasized that these guidelines are still under development and are currently in their fourth revision. As such, it is still evolving with noteworthy interest in feedback and suggestions for numerous topics. This includes methods for integrating digital evidence (e.g., Mobile Driver's Licenses and Verifiable Credentials) into identity proofing at various identity assurance levels, as well as whether or not emerging authentication models such as VCs are sufficiently addressed and accommodated by the guidelines. Although these guidelines are primarily meant for U.S. federal agencies implementing digital identity services, they could still be relevant in the international space and other use cases.

The Digital Identity Standards report [82] by Alamillo et al. for the European Union Agency for Cybersecurity (ENISA) outlines the most important standardization organizations and standards in the digital identity domain. The report divides groups of standards and specifications into two main categories: general and specific. General groups encompass standards and specifications used in identity management and trust services. Specific groups are differentiated by whether or not they provide authentication capabilities. We will focus on a selection of standards and specifications from both groups and briefly discuss each in the following subsections.

3.5.2. ISO/IEC 18013-5

ISO/IEC 18013-5 [83] is an international standard that specifies technical guidelines for the design format, data content, and implementation of ISO-compliant mobile driver's license (mDL) systems, acting as a base for international use and mutual recognition of the license, without interfering the autonomy of individual states or countries in enforcing their privacy rules. An mDL fulfills the same functionality as an ISO-compliant driving license (IDL) in a digital format called *mdoc* instead of the usual paper-based physical format. The mdoc is stored on a mobile device as a document or application, verifiable by entities other than the issuing authority. Besides the usual set of mandatory data, optional data in a mdoc include e.g. biometric templates. The Digital Identity Standards report [82] also mentions that *Concise Data Definition Language (CDDL)* [84] is used to express mdoc and may be encoded using *Concise Binary Object Representation (CBOR)* [85] or JSON Data Interchange Format [86].

3.5.3. ISO/IEC 23220

The ISO/IEC 23220 [87] series of standards is an evolution of the ISO/IEC 18013-5 and specifies the building blocks of general identity management through mobile devices, specifically for mobile Electronic Identification (eID) system infrastructure. This series of standards inherits and improves upon the functionality adopted by its predecessor, simultaneously ensuring backward compatibility. Furthermore, aspects such as generic data formats, trust models, as well as protocols and services for both the issuance and operational phases are specified. The generic nature and mobile-device-oriented approach of ISO/IEC 23220 compared to its predecessor make it the first real targeted attempt to facilitate a digital identity wallet, which has gained traction ever since it was proposed by the second rendition of the eIDAS regulations. As such, this series is identified as a standard to fulfill the needs of the EUDI Wallet [82].

3.5.4. OIDC with SIOPv2

The OpenID Connect Core 1.0 technical specification [88] defines the core functionality of OpenID Connect, including authentication and the usage of claims to specify information about the end user. OIDC is based on the OAuth 2.0 protocols and extends its functionality, as *Access Tokens* defined by OAuth 2.0 only allow access to resources but do not define standard methods to assert claims about the user's identity. Without profiling, OAuth 2.0 is incapable of providing information about the authentication of an end user. The end user's identity information is communicated by the *OpenID Provider (OP)* to the site the user is trying to log in to, i.e. a Relying Party (RP) or OAuth 2.0 Client, via a JWT called an *ID token* that can be reused for login into other OAuth 2.0 Clients for the specified amount of time. The OIDC Protocol and its actors are also briefly elaborated in this subsection 2.2.3.

The Self-Issued OpenID Provider v2 [89] further extends OpenID Connect with the concept of a *Self-Issued OpenID Provider (SIOP)*. It provides end users an option to bring their own identity by providing them OP to control on their own, called a SIOP. The end

user could then use this SIOP to issue identity information to RPs, enabling them to authenticate themselves with self-issued ID Tokens signed by keys under the control of the end users [82]. Hence, the identity would be “self-issued” [90]. SIOPs can also present cryptographically verifiable claims issued by the third parties trusted by the RPs, allowing end users to interact with RPs without the mediation of claims issuers or IdPs as specified in the OIDC core flow. SIOPv2 supports DIDs and is extended to cover Verifiable Credentials and Presentations.

3.5.5. OpenID4VC

The OpenID4VC specification consists of the following specifications [91]:

OID4VCI *OpenID for Verifiable Credential Issuance* defines an API that is used to issue Verifiable Credentials, supporting the VC data format as well as other credential formats, e.g. ISO 18013-5 [92].

OID4VP *OpenID for Verifiable Presentations* defines an extension of OIDC on top of OAuth 2.0 to allow the presentation of claims in the form of VCs as part of the protocol flow [93].

SIOPv2 *Self-Issued OpenID Provider v2* enables end users to use self-controlled OPs called SIOPs, as explained in the preceding subsection.

OpenID for Verifiable Presentations over BLE Defines the usage of Bluetooth Low Energy (BLE) to request VPs using the request and response syntax defined in OID4VP [94].

OpenID Connect UserInfo Verifiable Credentials Defines a new Verifiable Credential type called `UserInfoCredential` to provide OIDC UserInfo endpoints for the provision of user attributes to OpenID Clients in the form of a VC. Additionally, profiles for the OID4VCI protocol and for the StatusList2021 credential revocation mechanism are defined [95].

As of writing this thesis, these specifications are still in their early stages, with the last two being in draft state. Nevertheless, they have already been adopted in various other initiatives and used in the development of projects [96], due to their inherent flexibility in allowing implementers to make their own decisions for components of the VC technical stack [97].

3.6. SSI Initiatives

3.6.1. EBSI

The European Blockchain Services Infrastructure (EBSI) was born in 2018 as a joint initiative by the European Commission and member states through the European Blockchain Partnership (EBP) aimed at creating a blockchain-based digital infrastructure for the pan-European public sector. EBSI's vision is to "[...] build a secure, trusted and resilient infrastructure that enables public services to operate more efficiently, transparently and cost-effectively" [98, 99].

On a more technical level, EBSI consists of a peer-to-peer network of interconnected nodes running on a public permissioned blockchain-based infrastructure, with each member of the EBP running at least one node. Currently, EBSI is live and scaling up, supporting multiple environments for piloting, pre-production, and production with a comprehensive set of APIs for developing pilot projects. Compared to centralized and federated trust models with rigid, hierarchical, and various roles, the use of DID in EBSI enables greater flexibility and only requires two roles [100]:

1. **Trusted Accreditation Organization (TAO)** Verifies, accredits, and manages legal entities such as *Trusted Issuers* to extend the trust chain, where the so-called *root TAO* is the root of trust. A root TAO can accredit itself or other legal entities. In contrast, a TAO can only accredit other legal entities.

2. **Trusted Issuer (TI)** The leaf level of trust, is responsible for issuing VCs and the management of DID documents including their signing keys that are used to sign VCs.

EBSI also differentiates Legal Entities (LEs) and Natural Persons (NPs), introducing two distinct DID methods for each. The main difference is that LE DIDs are considered public identifiers and should therefore be unique in the DID Registry, requiring an additional *Verifiable Authorization* issued by the *EBSI Support* or a *Trusted Accreditation Issuer*. Conversely, NP DIDs are considered pseudonymous identifiers and are never registered in the DID Registry, exclusively created and stored within the holder's wallet [101]. Both LEs and NPs are additionally expressed through a special type of VC called *Verifiable ID*s [102, 103]. Verifiable IDs build upon existing standards and recommendations, namely eIDAS [72] and the W3C Verifiable Credentials Data Model [27].

3.6.2. ESSIF

ESSIF or the European Self-Sovereign Identity Framework [104, 105], is an EU-wide project that aims to build an identity layer for natural and legal persons based on the EBSI and rooted in the principles of SSI while being GDPR-compliant and in alignment with eIDAS. Its corresponding ESSIF-lab project gathered and funded small and medium-sized enterprises (SMEs), supporting these businesses to provide scalable and interoperable open-source SSI components. The project ended in December 2022, having funded up

to €5.6 million among 56 selected projects including a project to realize the SSI eIDAS Bridge scenario [106, 107] mentioned in Dr. Ignacio Alamillo Domingo's legal report [73].

On a more technical level, ESSIF is based upon a set of trusted registries defined by EBSI: a DID registry for issuers (acting as a self-controlled cryptographic trust anchor), a trusted issuer's registry (acting as a data trust anchor similar to a trust list), and a trusted schemas registry for the storage of VC schemas [82].

3.6.3. Gaia-X

Gaia-X [108] is a European initiative that strives to create a federated and secure data infrastructure for the collection and exchange of data across organizations, promoting interoperability and portability by connecting previously disjointed data and infrastructure ecosystems. With this objective, Gaia-X adopts concepts from SSI for its *Federation Services (GXFS)* and Trust Framework [109]. Furthermore, three conceptual pillars of the initiative are defined [110]:

1. **Gaia-X Compliance** Establishment of a common digital governance based on European values through decentralized services.
2. **Data Spaces/Federations** Interoperability and portability of data sets and services spanning multiple sectors.
3. **Data Exchange** Means to perform data exchange and enforce anchored contract rules for access and data usage within the infrastructure.

In practical terms, each pillar will have three main deliverables, consisting of Functional and Technical specifications, as well as software. Functional specifications describe the high-level functionality of Gaia-X, while technical specifications describe the technical requirements of Gaia-X, respective to the conceptual pillar. We will look into the most relevant pillar, namely the GXFS in greater detail in Chapter 7.

4. Literature Review Methodology

In recent years, plenty of academic publications and peer-reviewed articles have been published, researching numerous aspects of the new identity paradigm and its applicability in various domains. While academic publications or white literature (WL) undoubtedly provide valuable insights into the field, we recognize that the landscape of SSI is evolving rapidly, marked by the introduction of new standards and updated regulations, as well as the emergence of projects and initiatives. Therefore, to best answer the thesis's first research question we decided to include grey literature (GL) alongside academic publications, aiming to discover novel concepts and implementations for identifying natural subjects in Verifiable Credentials. This enables us to stay current with the latest developments, standards, and real-world implementations and better understand the current status quo.

Since implementations and their corresponding technical documentation are closely linked with the field of software engineering, we conducted an adapted version of the Multivocal literature review (MLR) method proposed by Garousi et al. [111], which presents a form of systematic literature review (SLR) that includes grey literature in addition to formal literature in software engineering. While the method is relatively new, the presented guidelines were formulated from a survey of 24 MLR guidelines and experience papers from various fields, following a rigorous process. This led to the establishment of 14 guidelines or recommendations across all phases of the MLR.

The MLR consists of three main phases: planning, conducting, and reporting the review. Each phase is further detailed by substeps, as seen in Figure 4.1. Each phase along with its corresponding steps will be elaborated in the following sections.

4.1. Planning

When planning for the MLR, Garousi et al. highlight these two steps:

1. **Establishing the need for an MLR** In the first step, it should be determined whether conducting a systematic review is necessary or not. Additionally, an assessment of which type of literature review (LR) to perform is needed. For this purpose, they present a checklist to aid the decision-making process. We intend to use the proposed criteria to justify our decision to include GL in our review presented in Table 4.1. For our thesis, the sum of "Yes" answers is five. According to Garousi et al., one or more "yes" responses suggest the inclusion of GL.

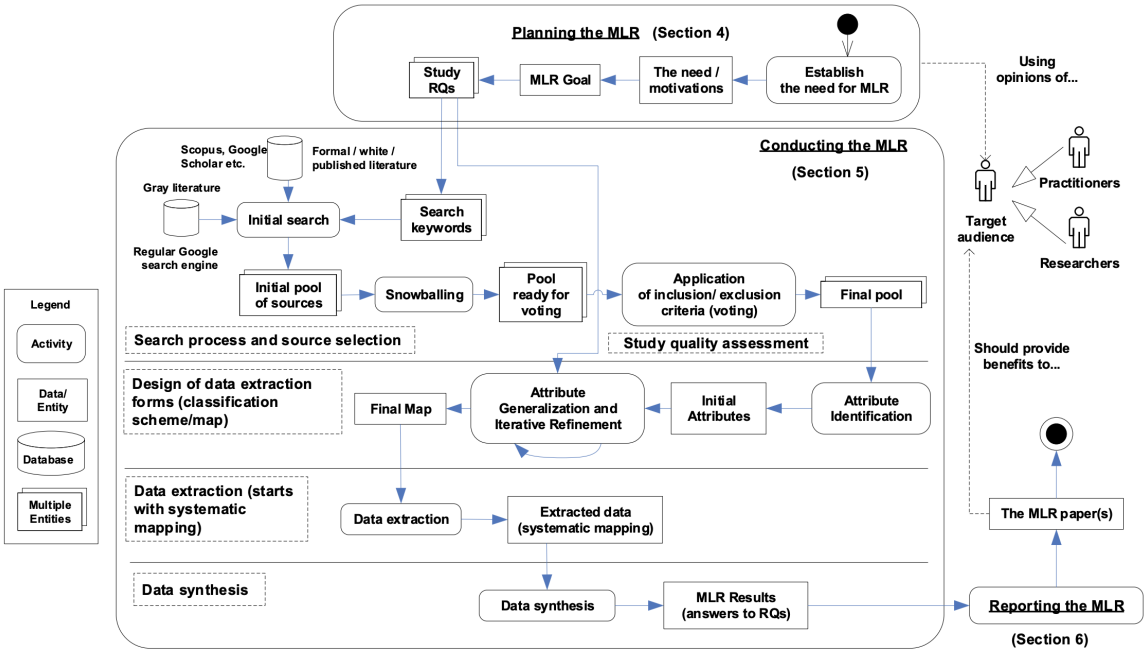


Figure 4.1.: An Overview of the Multivocal Literature Review process, adopted from [111].

2. **Setting the MLR Goal** Researchers should identify and review existing work in the field and determine the usefulness of an MLR for its intended audience before setting the appropriate research questions. This is a crucial part of an MLR, as it drives the review in aspects such as the search and data extraction process. RQs should also be constructed to match the needs of the target audience. For this purpose, a classification scheme for RQs is presented that differentiates RQs into five distinct categories, with subcategories for each. When we evaluate the RQs presented in chapter 1, our first two RQs are *exploratory* questions, assignable to the descriptive-classification subcategory as we seek to find existing approaches for the inclusion of identifying information in VCs as well as existing methods to update this information. The third RQ is assigned to the *design* category, as we seek to implement a solution for identifying natural subjects in VCs within the context of the Gaia-X project.

To summarize, after following the first phase of the method presented by Garousi et al., the inclusion of GL in our MLR is justified based on systematic reasoning. The specified research questions are intended for researchers, with two of the three categorized as exploratory questions. The inclusion of GL helps us to gather information found in technical documents and reports, enabling the thesis to paint a more accurate picture of the current landscape.

#	Question	Answer	Note
1	Is the subject "complex" and not solvable by considering only the formal literature?	Yes	The subject is the identification of natural subjects in VCs. To determine more technical details, technical documentation needs to be considered, which are categorized as grey literature.
2	Is there a lack of consensus on outcome measurement in the formal literature?	Yes	The subject is not the main topic of most academic works and is usually only addressed briefly without much consideration, with the exception of regulations.
3	Is the contextual information important to the subject under study?	Yes	Different approaches across various projects implemented by parties with different contexts are involved.
4	Is it the goal to validate scientific outcomes with practical experiences?	Yes	The subject arose after realizing that not much concern has been directed toward handling PII in VCs based on experience from working on Gaia-X.
5	Is it the goal to challenge assumptions or falsify results from practice using academic research or vice versa?	No	The goal is to identify the existing methods and evaluate them.
6	Would a synthesis of insights and evidence from the industrial and academic community be useful to one or even both communities?	Yes	While predominantly intended for researchers, practitioners can still gain useful insights from the results.
7	Is there a large volume of practitioner sources indicating high practitioner interest in a topic?	No	It is not the main topic for most, but relevant information is typically found in technical specifications and implementations.

Table 4.1.: An adapted checklist for deciding the inclusion of GL in an MLR.

4.2. Conducting the Review

In line with the methodology, the procedure for carrying out an MLR consists of five distinct phases. The following section is structured accordingly and illustrated in Figure 4.2.

4.2.1. Search Process

We recognize the importance of employing carefully crafted search strings as they will determine the resources from which we will extract information. Initially, we used simple search terms and conducted full-text searches exclusively leading to non-specific results. Experimentation is therefore required, involving an iterative process for defining the search strings, syntactically adjusted to the database in which the search is conducted while making sure that the semantic meaning remains the same across all databases.

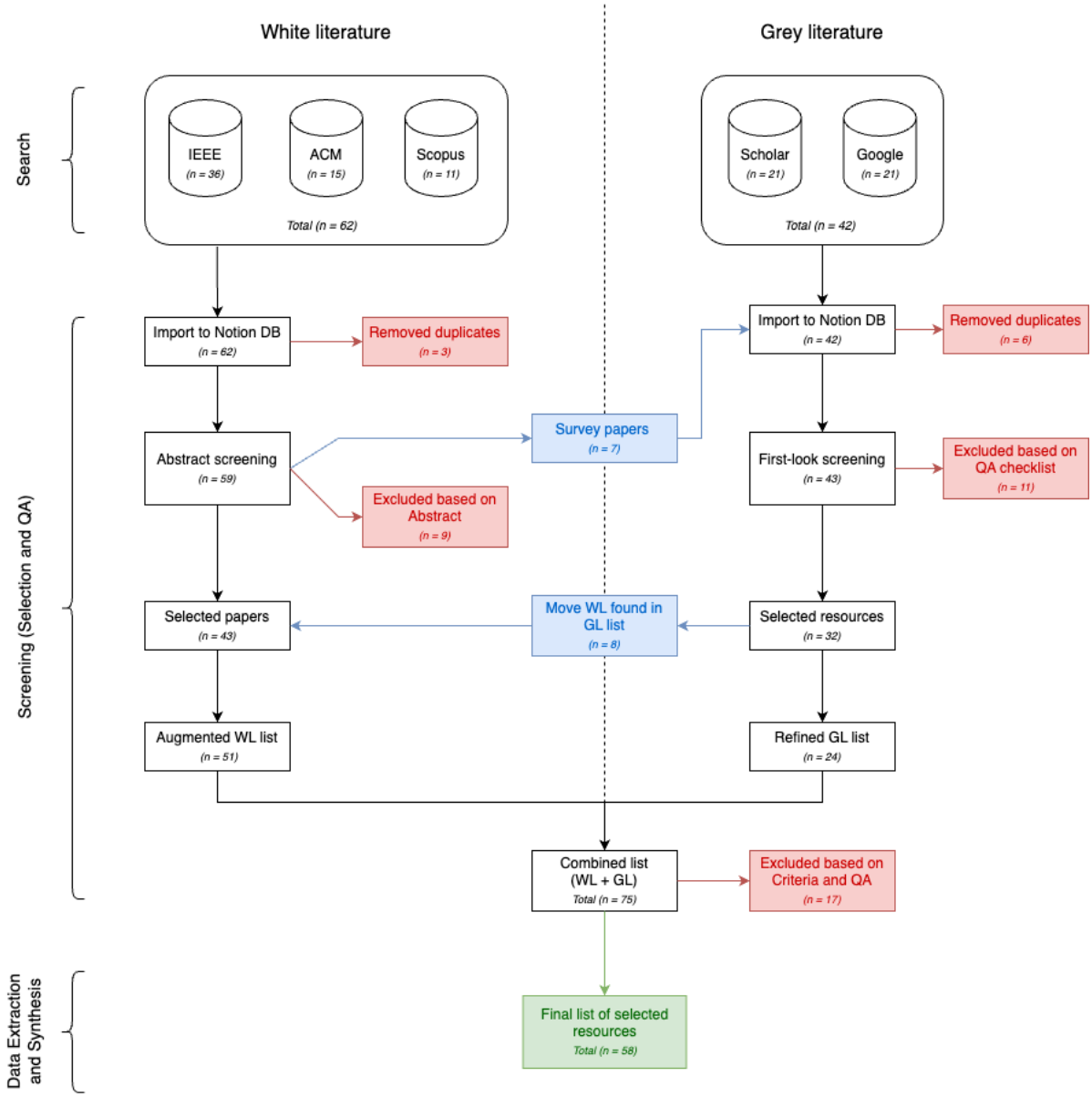


Figure 4.2.: An overview of the literature review process

The following databases along with the utilized search strings were used to find the most relevant papers on identifying natural subjects in VCs:

IEEE *(("Document Title": "Self-sovereign identity" OR "SSI") OR ("Document Title":identity management)) AND (("Full Text .AND. Metadata":natural person) OR ("Full Text .AND. Metadata":legal person) OR ("Full Text .AND. Metadata":"person")) AND ("Full Text .AND. Metadata":"verifiable credentials")*

ACM *Title:("self-sovereign identity" OR "SSI" OR identity management) AND (All-*

Field:("natural person" OR "legal person") OR AllField:("person")) AND Fulltext:("verifiable credentials")

Scopus *(TITLE ("self-sovereign identity" OR "SSI" OR "identity management") AND ALL ("natural person" OR "legal person" OR "person") AND ALL ("verifiable credentials"))*

Scholar *"verifiable credentials" “self-sovereign identity” "Identity Management" "natural person" "legal person”*

Search *verifiable credentials "natural person" self sovereign identity "Identity Management" "natural person" "legal person" person*

Databases searched for academic literature were **IEEE**, **ACM**, **Scopus**. Both Google **Search** and **Scholar** were used to search for GL. The discrepancies between the search strings can be attributed to variations in each database’s search syntax, functionality, and features. For searching grey literature, we decided to generalize the search string to gain more resources such as white papers, blog posts, technical reports, and documentation. For academic databases, our search strings were specific enough that the number of search results was manageable and relevant. As such, we collected all of the literature from the search. In contrast, the more generalized search string employed in **Google Scholar** and **Google Search** returned an unmanageable amount of resources with resources in the latter pages that weren’t necessarily relevant. We therefore decided to only extract the first two pages of the search result, employing effort-bounded stopping criteria. Furthermore, we searched for resources published in the year 2016 and onward as it was the year when Allen’s famous work on SSI [13] was first published. Filtering search results by year is supported and therefore used in all database searches.

4.2.2. Search Selection

To guarantee the relevance and significance of the collected papers, they need to be assessed for their actual relevance. For this purpose, the selection process includes determining both inclusion and exclusion criteria. The selection criteria for academic literature are listed in Table 4.2.

Table 4.2.: Inclusion and exclusion criteria for abstract and first look screening.

Inclusion Criteria	
IC-1	Papers that include natural subjects in VCs.
IC-2	Papers that are accessible through institutional login.
Exclusion Criteria	
EC-1	Papers without sufficient technical details or processes.
EC-2	Papers that focus on domains that do not involve natural subjects.
EC-3*	Survey papers.

These criteria were used for the initial review process we call *abstract screening (AS)*, where we determine the paper's relevance solely based on the contents of the *abstract* section of a paper. We do not entirely exclude survey papers from the LR. Instead, we set them aside to augment the GL sources before applying a *first-look screening* for the GL list, including the same criteria as the aforementioned AS but without **EC-3**. This is done to gather as many SSI solutions as possible. Additionally, we encountered WL during our GL search, prompting us to transfer them to the WL pool for the sake of maintaining consistency. This occurrence could be attributed to the broader search string employed in the search process.

4.2.3. Study Quality Assessment

As GL tends to be more diverse and less controlled than white literature, Garousi et al. suggest the use of a quality assessment checklist when selecting GL to further ensure its quality. They also encourage making suitable adjustments, i.e. removing or extending the suggested criteria as there is no "one-size-fits-all" quality model. The adapted criteria are listed in Table 4.3.

Table 4.3.: Quality assessment table for GL.

Criteria	Questions
Authority of the producer	Is the publishing organization or author reputable and has expertise in the area?
Methodology	Is the source supported by authoritative, documented references?
Date	Does the resource have a clearly stated date and is published during the target period?
Relevance	Does the resource describe anything related to making a connection from natural subjects to VC-based digital identities?

It is important to note that instead of using a three-point Likert scale to assign scores to assessment questions as suggested, we used the criteria to qualitatively evaluate the GL in addition to the selection criteria in Table 4.2 for the first-look screening.

4.2.4. Data Extraction and Synthesis

After concluding the previous phase, we combined the WL and GL resources into a consolidated list and removed duplicates. We then conducted a *full-text* review of all resources in this list, using the criteria outlined in Table 4.2 and 4.3, as well as taking notes on the following aspects for each solution presented in a resource:

- Identity binding
- Recipient of the digital identity

- PII location
- Financial implications
- Security and Privacy Measures
- Relevant Standards

All solutions along with the extracted information are then recorded in a Notion web application database, as depicted in Figure 4.3. In this database, SSI solutions and approaches from both WL and GL are evaluated based on their relevance to our work. This is done in the "Status" column, where a solution is excluded from the subsequent procedure if it is either deprecated, does not use VCs, is not an SSI solution, or does not provide sufficient information. We then have columns for providing backlinks to the paper(s) or source(s) in which the solution was presented that are stored in a separate database. Figure A.1 in the appendix shows the note-taking template defined earlier applied to a specific solution, in this case for a solution called *Alastria ID*. This database will serve as the dataset used to construct the taxonomy discussed in the following Chapter 5.

SSI solutions/approaches

↑ Name

Aa Name	Status	WL IDs	GL IDs	WL	GL
Abacus	Deprecated	WL-27		Blockchain-Based Id	
Abubakar et al.	Related work	WL-34		Blockchain-based Pl	
Alastria ID	Include (end-user ID)	WL-20	GL-3, GL-7	Self-Sovereign Ident	Digital Identities
AlgoCert	Include (end-user ID)	WL-36		AlgoCert: Adopt Nor	
Altme (Wallet)	Include (end-user ID)				
BanQu	Deprecated	WL-27		Blockchain-Based Id	
Belchior et al. (SSIBAC)	Include (end-user ID)	WL-6		SSIBAC: Self-Sovere	
BitID	Not SSI	WL-27		Blockchain-Based Id	
Bitnation	Deprecated	WL-27		Blockchain-Based Id	
Blockcerts	Include (end-user ID)	WL-17		A New Approach to (	
Blockchain Helix / Helix ID	Not enough info	WL-27		Blockchain-Based Id	
Blockpass IDN	Not SSI	WL-27		Blockchain-Based Id	
Blockstack	Not SSI	WL-21, WL-40		Self-Sovereign Ident	
Bloom	Not SSI	WL-27		Blockchain-Based Id	
Cambridge	Deprecated	WL-27		Blockchain-Based Id	

Figure 4.3.: Notion Database.

5. A Taxonomy of SSI Solutions: Identifying Natural Subjects in Verifiable Credentials

In this chapter, we elaborate on the methodology used to develop our proposed taxonomy following the well-established approach in the information systems domain proposed by Nickerson et al. [59]. Their approach is an iterative process that integrates both empirical and conceptual methods for defining new taxonomy dimensions.

5.1. Meta-Characteristic

The choice of meta-characteristic is a crucial aspect of taxonomy development as it will serve as the basis for the choice of taxonomy characteristics. The selection of the meta-characteristic should be guided by the taxonomy's intended purpose. We had set the goal from the start, established in our first research question. Our objective is to inform researchers of characteristics that will help them understand the technical specifics of the user identification approach in SSI ecosystems, aiding them in designing their implementation. Taking this into consideration, the meta-characteristic is defined as follows:

> *Characteristics of user identification approaches in SSI such as how the user's PII is included, the data flow, formats, identity verification, management, and storage location.*

5.2. Ending Conditions

Due to the iterative nature of the methodology, ending conditions need to be defined to determine when the process terminates. Nickerson et al. divide this into two conditions, namely objective and subjective ending conditions. Eight objective ending conditions are identified and need to be considered after each iteration. In contrast, the subjective ending conditions are used to examine the resulting taxonomy. For the method to terminate, it must be argued that all subjective conditions are met. We closely align our ending conditions with those outlined by Nickerson et al. with some slight modifications. The resulting ending conditions are detailed in Tables 5.1 and 5.2 for the objective and subjective ending conditions, respectively.

Objective Ending Condition	Comment
All SSI approaches found from the survey deemed to be relevant have been examined	The survey includes both WL and GL. Both of these sources together provide an initial total of 92 SSI approaches before further refinement. Nevertheless, it is still considered an extensive sample that encapsulates existing SSI approaches since the inception of the concept.
No object was merged with a similar object or split into multiple objects in the last iteration	If objects were merged or split, then we need to examine the impact of these changes and determine if changes need to be made in the dimensions, characteristics, or the assigned objects.
At least one object is classified under every characteristic of every dimension	Should an object be unassignable to a characteristic due to incomplete information, it will be assigned to an 'unspecified' characteristic instead of making assumptions about the object and sacrificing objectivity.
No new dimensions or characteristics were added in the last iteration	If new dimensions were found, then more characteristics of the dimensions may be identified and vice versa. Adding new dimensions might also entail the deletion of other dimensions deemed superfluous.
No dimensions or characteristics were merged or split in the last iteration	The merging or splitting of dimensions or characteristics will have effects on the rest of the taxonomy. Its impact should be considered and changes made accordingly.
Every dimension is unique and not repeated	Duplicate dimensions need to be removed as they do not increase the value of the taxonomy.
Every characteristic is unique within its dimension	The removal of duplicate characteristics is necessary as we might have several dimensions with an "unspecified" characteristic.
Each combination of characteristics is unique and is not repeated	If cells are not unique, then there is redundancy/duplication in cells that need to be eliminated

Table 5.1.: Objective Ending Conditions

5.3. Defining Objects of Interest

Our pool of objects is extracted from our MLR which includes WL and GL. As such, objects extracted from both sources of literature tend to be different. From WL, we mainly extract system designs or frameworks that might not have been implemented by the researcher. Should the system design be implemented, they are usually implemented as a proof of concept to be researched further and not as production-ready solutions. In contrast, approaches extracted from GL tend to be commercial solutions provided by so-called SSI vendors, encompassing the archetypes of SSI ecosystem players outlined by Schmidt et al. [56] called *Non-DL-operating Governance Authorities* and *Implementers*. Entities under both archetypes either provide a governance framework or implement SSI solutions.

With this in mind, we consider both sets of objects as we aim to gain insights from both theoretical and practical perspectives. We consequently define an object of interest for the taxonomy as encompassing system designs, along with commercial SSI solutions.

Subjective Ending Condition	Comment
Concise	The taxonomy should be informative without being overwhelming. We followed the suggested rule of thumb of seven plus minus two dimensions, but this isn't an objective ending condition, meaning the number can be exceeded if the addition of dimensions is deemed necessary.
Robust	The combination of dimensions and characteristics should be chosen to provide informative differentiation among objects of interest.
Comprehensive	The taxonomy is considered to be comprehensive once all dimensions of all objects of interest are identified, namely all relevant attributes of an identification approach for SSI solutions.
Extendible	Taxonomy extensibility is kept in mind during its construction to keep up with the rapid development in the SSI space. Should new information or details surface, the "unspecified" characteristic can be removed and replaced with new identified characteristics. New dimensions could also be added to extend the taxonomy.
Explanatory	We want to create a taxonomy that provides sufficient details on user identification approaches within the SSI context, including technical and non-technical information.

Table 5.2.: Subjective Ending Conditions

These approaches must possess sufficient technical details and facilitate digital identity provision to end-users, specifically through Verifiable Credentials. As a consequence of this definition, we excluded 57 approaches from an initial pool of 92.

5.4. Taxonomy Construction

We examine a total of 35 SSI approaches that provide end-user identification with VCs. To start the iterative taxonomy construction process, we started with approaches that have been mentioned in the most amount of literature with alphabetical sorting. For each iteration, three to five approaches are examined and assigned. The empirical-to-conceptual strategy was employed for the majority of the iterations. It was only in the later stages that we employed the conceptual-to-empirical strategy to refine the taxonomy dimensions and characteristics further.

5.5. Limitations

In this section, we address and acknowledge limitations imposed by the methodology on our taxonomy.

Firstly, objects were derived from the literature review based on robust search strings. However, it still cannot be guaranteed that all SSI solutions were found due to the effort-

bounded nature of the GL search that relies on the search engine to deliver the most relevant results in the first few pages. Secondly, SSI approaches that do not provide sufficient technical documentation are excluded. In most cases, these approaches also show little to no signs of development and have very few adopters.

Furthermore, information that was extracted from e.g. company blogs, documentation, whitepapers, and unofficial drafts may be inaccurate and not up to date. We tried our best to base our research on the most recent works of the respective approach.

As pointed out in our third objective ending condition, solutions that are unassignable to a characteristic are given the *unspecified* attribute, so as not to misguide researchers by making assumptions about an approach. For example in the *Projected Cost per User* dimension, information regarding pricing is often obscured, requiring people to contact these SSI vendors to get a quote on pricing. We still tried our best to limit the use of this *unspecified* characteristic, as is the case for the *Identification Data Source* dimension with only an approach being assigned under this characteristic. The taxonomy was also built with extensibility in mind to compensate for this shortcoming.

Following this, our MLR was conducted towards the end of the first half of 2023, which implies that any works published or conducted after that date are not included in the review.

Lastly, we acknowledge that researcher bias and human error cannot be entirely eliminated for works requiring human judgment, especially when having to extract information from alternative grey literature that is more inconsistent in its structure. To mitigate this issue, we adopted a systematic approach for both the MLR and taxonomy creation processes. Additionally, we sought guidance and consultation from our thesis supervisor.

5.6. Proposed Taxonomy

We now present the proposed taxonomy in Table 5.3. Each row corresponds to an SSI approach as detailed in section 5.3. In the following subsections, we will elaborate on the dimensions and their corresponding characteristics in the taxonomy that are used to describe an approach. We subsequently discuss observations derived from applying the taxonomy to our selected pool of SSI approaches.

Table 5.3.: Proposed Taxonomy of SSI Approaches

Non-exclusive/ Exclusive	E		N		N						N				N				N		N		E		E		
SSI approaches	PII Location		PII Type		Identification Data Source						Identification Authority				Projected Cost per User				VC Format			Schema Standard		Selective Disclosure		Credential Revocation		
	Standalone	Bundled	Natural	Alternative	Gov-ID	Non-Gov-ID	Biometrics	PoP	None	Unspecified	End-user-asserted	Third-party asserted	SSI-Integrator asserted	First-party asserted	Free	Per-issuance + operation	Recurring base fee	Unspecified	LDP-VC	JWT-VC	Unspecified	Standardised	Flexible	Yes	No	Yes	No	Unspecified
Civic	x		x	x	x	x	x	x				x	x			x	x		x	x			x	x		x		
Jolocom		x	x	x						x	x				x						x		x	x		x		
Midy (Evernym)	x		x		x		x					x			x						x		x		x			x
Soltani et al. [112]	x		x	x	x									x		x					x		x		x	x		
MediLinker	x		x	x	x	x								x				x			x		x		x	x		
Alastria ID	x		x		x	x							x			x	x		x			x		x		x		
Tahlil et al. [49]		x	x		x									x				x			x		x	x		x		
Blockcerts		x	x						x		x					x			x				x		x	x		
Cosmos Cash		x	x						x		x							x	x				x		x			x
Dock		x	x	x	x	x	x						x			x	x		x	x			x	x		x		
Saidi et al. [113]		x		x					x		x							x			x		x		x	x		
Herbkle et al. [114]		x	x	x		x								x				x	x			x			x			x
Hamer et al. [115]	x		x				x					x						x	x				x		x	x		
Stockburger et al. [116]	x		x	x	x									x	x						x		x	x			x	
Wang et al. [117]		x		x	x							x						x			x		x		x			x
WeIdentity		x		x	x									x				x	x				x	x		x		
Xu et al. [118]	x			x					x		x							x			x		x		x	x		
M. Morosi [119]		x	x		x									x				x	x				x	x		x		
C. Sehlke [26]	x		x		x	x								x	x				x			x			x	x		
Trinsic		x	x	x	x	x	x						x			x			x	x			x	x		x		
ValID		x	x	x					x		x				x						x		x		x	x		
Altme		x	x		x	x		x				x			x				x	x		x	x		x			x
Datakeeper	x		x		x	x	x						x			x	x		x				x	x		x		
Gataca.io		x	x	x	x	x	x	x					x			x	x		x	x		x	x	x		x		
GlobalID		x	x	x	x	x	x	x				x	x			x					x		x		x	x		
Indicio Proven		x	x	x	x	x	x	x					x			x					x		x	x		x		
Lissi		x	x	x	x	x							x			x	x				x		x	x		x		
Mattr		x	x	x	x	x	x						x				x		x			x	x	x		x		
Meeco		x	x	x					x				x				x			x			x	x			x	
Verida	x		x		x	x							x			x				x			x	x				x
VIDchain	x		x						x				x					x		x		x	x		x	x		
Walt.id	x		x	x					x				x					x	x			x	x		x	x		
Belchior et al. [120]		x	x						x					x	x						x		x	x		x		
Rahman et al. [121]	x		x		x									x				x			x		x	x				x
Satybaldy et al. [122]		x	x	x	x	x								x				x	x			x		x		x		

5.6.1. Dimensions

The proposed taxonomy encompasses a total of nine dimensions, each comprising between two to six characteristics. A dimension can be categorized as either *non-exclusive* or *exclusive*. In non-exclusive dimensions, an object may possess multiple characteristics within the same dimension. Conversely, objects assessed within exclusive dimensions are limited to having only one characteristic. We have made this distinction to avoid having too many characteristics in a single dimension. In the subsequent discussion, we will delve into each dimension and provide a detailed explanation of their respective characteristics. Furthermore, it is noted whether the dimension is exclusive (E) or non-exclusive (N).

PII Location (E) The first dimension concerns itself with where the user's identifying information is located, more specifically in what kind of VC. We adopt the definition of PII outlined by GDPR Article 4 [70], referring to any pieces of personal information that can be used to identify a particular person when viewed individually or when collected together. This includes name, email address, location data, and identification number such as a social security number (SSN). The dimension differentiates the purpose of the VC used to store such PII. Within an SSI solution, however, both types of VCs can be offered to the user. Therefore, this dimension is non-exclusive. We have identified the following characteristics:

- *Bundled*: user PII is stored within a supplementary VC that has been issued for specific purposes, specifically to serve a specific purpose within the domain. An example of this is "age-range" credentials that are issued to specifically attest to a user's age range and nothing else. They can be viewed as more privacy-preserving but are limited in their purpose.
- *Standalone*: standalone credentials are meant to specifically attest to the user's identity, containing multiple user PII that is available in a passport. Such credentials are typically needed for use cases that require a higher level of identity assurance, e.g. KYC and AML. When considered independently, they offer lower levels of privacy protection. However, such presentation is uncommon, as they are typically derived into bundled credentials or employed in conjunction with Selective Disclosure mechanisms.

PII Type (N) The second dimension further distinguishes the identifying information that is stored in a VC. Initially, we made a distinction between natural identifiers and "contact" identifiers. The latter pertained to identifiers that could be employed as a means of contacting the end-user, such as the user's email or DID. However, we found this categorization to be insufficient and replaced it with "arbitrary" identifiers, encompassing contact identifiers and assigned identifiers such as SSNs, biometric templates, and the like. Upon further examination, we realized that such identifiers can also be considered natural identifiers since they are assigned to the user, often without their input, and changing them can be challenging for the user. With this in mind, we settled on the following characteristics for this dimension:

- *Natural*: this type of identifier includes, but is not limited to name, birthdate, SSN, biometric template, place of birth, and pseudonymous identifiers [123]. This characteristic encompasses identifiers that are primarily assigned to the end-user and are typically infrequently changed or difficult, often impossible to replace and dispose of.
- *Alternative*: identifiers that can be modified by the user, created independently by the user, and can be discarded at the user's discretion. This characteristic comprises the user's DID and email addresses.

Identification Data Source (N) This dimension describes the methods and approaches utilized for user identification or the validation of user identity attributes before the issuance of VC. An SSI approach may incorporate or mandate various methods of identity validation, making this dimension non-exclusive. Moreover, an approach can be classified as having the following characteristics even if the validation approach is not mandatory or enforced. Nevertheless, the approach must acknowledge the possibility and have demonstrated intent in incorporating such forms of user identity validation:

- *Gov-ID*: user identity validation involving the use of government-issued identity documents, such as passports and driver's licenses. This validation method usually necessitates in-person validation, wherein the prospective identity VC holder must undergo validation at the issuer's physical location.
- *Non-Gov-ID*: user identity validation using documents originating from private organizations, non-governmental organizations (NGOs), and institutions, for instance, universities and corporations.
- *Biometrics*: relies on unique physical characteristics of the user for identity validation, including the use of fingerprints and face recognition to distinguish individuals from one another.
- *PoP*: Proof of Personhood pertains to techniques used to establish that an individual is a human being or natural subject rather than a computer program or bot. This characteristic encompasses methods such as CAPTCHA challenges and liveness or video selfie tests.
- *None*: approaches classified under this characteristic do not specify or mandate user identity validation, especially in the case of self-issued credentials or in academic SSI frameworks, where it is frequently left as a subject for future consideration or research.
- *Unspecified*: attributed to approaches that do not explicitly address or offer inconclusive information regarding user identification.

Identification Authority (N) The next dimension describes the entity responsible for validating the user's identity before issuing the VC containing the user's PII. This dimension is non-exclusive since an approach may provide multiple options for VC issuance, describable by these characteristics:

- *End-user asserted*: refers to self-validated VCs, trusting the user to provide accurate information about themselves, implying that identity validation is not required.
- *Third-party asserted*: user identification is carried out by a trusted third party, typically a contracted or specialized entity engaged by the SSI service provider.
- *SSI-integrator asserted*: *SSI integrators* refer to business customers looking to integrate SSI into their workflows. They differ from *SSI service providers or vendors*, who provide the necessary infrastructure and services for implementing such SSI-based workflows. SSI approaches possessing this characteristic entrust user identification processes to these integrators.
- *First-party asserted*: identification of end-users is carried out either by the SSI service provider or by institutions that built their own SSI-based process. This characteristic encompasses SSI integrators that do not have such an identification process and thus rely on SSI vendors to conduct this process for them.

Projected Cost per User (E) In practical terms, completely free services are almost non-existent, especially for integrators. This dimension specifically distinguishes costs related to credential-related operations that the end-user or integrator is required to cover. We recognize that fees vary a lot depending on the registry used, potentially increasing over time as more users are amassed. Additionally, integrators might in turn require their end-users to pay this fee. Considering this, we classify cost models rather than specific costs and have identified the subsequent models:

- *Free*: no associated costs for credential-related operations.
- *Per issuance + operation fee*: fees that need to be paid per identity-related VC issuance in addition to operational fees, such as those associated with deploying smart contracts and writing to the blockchain.
- *Recurring base fee*: pertains to fees typically structured in subscription tiers that determine the permissible number of credentials that can be issued within a specific period. This pricing model is prevalent among integrators rather than end-users.
- *Unspecified*: no information is provided regarding cost or when inconclusive pricing information is presented. This phenomenon is common among SSI vendors who request clients to engage in consultation for pricing quotations.

VC Format (N) The current dimension encompasses structures and formats used when presenting the VC from a wallet to a relying party. We distinguish the following VC formats with regard to the W3C VC data model:

- *LDP-VC*: Linked-Data Proofs (LDPs) are powered by JSON-LD formats. They ensure the integrity and ensure the authenticity of the VC. LDPs are commonly used as they enable selective disclosure, zero-knowledge proofs, as well as other benefits [124].

  - *JWT-VC*: the VC is encoded in JSON or JSON-LD, but is presented and secured in the JWT format using the Javascript Object Signing and Encryption (JOSE) Framework [125] for encryption.
  - *Unspecified*: no conclusive information is provided regarding the specific format that is used.

Schema Standard (N) This dimension addresses whether the attributes of the credential subject within a VC adhere to an established credential schema standard. We exclude adherence to the W3C VC data model, as this dimension focuses on schema standards beyond this foundational framework. We introduce the following characteristics:

  - *Standardized*: credential subject attributes are outlined by an existing credential standard such as the ISO 18013:5 and ELMO [126] standards.
  - *Flexible*: the subject attribute schema is determined by the user or SSI vendor, implying that it is customizable.

Selective Disclosure The ability to determine which attributes are shared with the relying party without revealing all identity information is an important feature of VCs to further enhance privacy. It is however not supported by all approaches, therefore providing another point of distinction:

  - *Supported*: selective disclosure is supported by the SSI approach.
  - *Unsupported*: selective disclosure is not supported by the SSI approach.

Credential Revocation (E) The final dimension focuses on the possibility of revoking the issued credential. Similar to selective disclosure, it is not a feature that is offered or even considered by all approaches. As such, we define these subsequent characteristics:

  - *Supported*: credential revocation is supported by the SSI approach.
  - *Unsupported*: credential revocation is not supported by the SSI approach.
  - *Unspecified*: no conclusive information is provided regarding the revocation of VCs.

5.6.2. Discussion and Recommendations

We now present some observations extracted from the taxonomy and explore their potential implications by discussing each dimension.

The first taxonomy dimension, *PII Location*, reveals a preference for bundled credentials in SSI approaches. This suggests that end-user identification data often become part of "purpose-driven" credentials, serving as a binding element to standalone credentials that specifically verify the end user's identity. The fragmentation in SSI solutions may stem from this approach, as specialized credentials are typically usable only within specific SSI ecosystems. However, bundled credentials promote a multi-faceted identity paradigm, where individuals can possess distinct *personas* defining their identity in various contexts,

including their professional, online, and personal identities. An ideal scenario would involve the existence of a legally recognized identity anchor within the SSI context. Recent advancements in standards and pilot projects, particularly the introduction of eIDAS 2.0 and the European Digital Wallet, suggest a potential shift towards national IDs based on Verifiable Credentials. This presents an opportunity to use such a VC as a foundational element for bundled credentials in the future.

In the second dimension, *PII Type*, most approaches incorporate natural identifiers or a combination of both natural and alternative identifiers. Only three approaches exclusively incorporate alternative identifiers. This further highlights the need for regulations to protect the usage of sensitive user information in VCs.

In the next dimension, *Identification Data Source*, it becomes evident that the majority of approaches either mandate or provide the option to use government-issued credentials or a combination of such credentials with other identification data sources. Approximately one-third of these approaches also incorporate biometrics as a means of identification. Notably, most approaches that require proof of personhood checks also make use of biometrics. Nine approaches stand out by not necessitating any form of end-user identity validation, while one approach remains unspecified in this regard. This finding implies that, in most cases, a trusted government entity is still considered the primary authority for validating identity. Nonetheless, there are initiatives, such as WorldID [53], which rely exclusively on biometrics to establish personhood. Such approaches, however, still have a long road ahead in terms of development and adoption.

Within the dimension of *Identification Authority*, only two SSI vendors offer both third-party and integrator-asserted identity validation options. Results are somewhat evenly distributed between reliance on integrators and first-party identity validation. The proportion of end-user and third-party identity validation options is equivalent, with third-party authentication services starting to creep into the SSI space, contracted especially by SSI vendors. Ultimately, the choice between these identification authorities appears to hinge on the specific use case at hand, underscoring the flexibility of SSI solutions.

During the taxonomy-building process, we often encountered challenges related to the *Projected Cost per User* dimension, primarily due to the limited disclosure of cost details and pricing models by these approaches, especially in academic contexts, posing difficulties when determining the characteristics of these approaches in terms of costs per user. Many SSI vendors opt to withhold pricing information and instead require potential clients to initiate contact for custom pricing quotes, further complicating the assessment. It is also notable that the majority of solutions with a recurring base fee also impose fees per issuance and additional operational fees. These additional fees likely include identity validation costs, which are often outsourced. Plenty of approaches also tend to suggest that the financial responsibility for these costs falls upon the integrator.

As we aim to analyze the technical aspects, the utilized *VC format* is an important aspect. Unfortunately, non-open-source solutions obscure such implementation details. Academic works that present SSI frameworks also don’t address this aspect, focusing more on the overarching architecture. Both of these approaches make up almost half of the examined

approaches. Almost the entire other half utilize LDP-VC as the format of choice, with five out of 17 offering support for either LDP-VC or JWT-VC. Only three approaches use JWT-VCs exclusively. Despite the relatively low amount of adopters, we expect an uptick in approaches that support JWT-VCs as it recently became the de facto format for the EUDI Wallet and OIDC4VC protocols for VC-related operations, considering its relative maturity compared to Linked Data formats.

Credential subject attributes are typically standardized in paper-based credentials. Given SSI's potential, we wanted to investigate whether VC subject attributes are also based on such standards in the *Schema Standard* dimension. We found that 26 out of 35 approaches are flexible in this regard and only four approaches offer exclusively standardized VCs. For the latter, such VCs are meant for specific use cases in the education and medical context. We suspect that the majority of approaches tailor their VCs to suit their specific requirements with the trade-off of possibly compromising interoperability between their SSI ecosystem and external systems in the process. This isn't necessarily a bad thing in the short term, as customizability is a positive thing in the eyes of SSI integrators. In the long term, given recent efforts in the EU aimed at promoting standardized formats, it is advisable for SSI approaches to align their schemes to such standards, particularly for use cases requiring a higher level of assurance (LoA).

Another cornerstone of SSI is the end-user's capability to selectively disclose information to verifiers, a feature we examined through the *Selective Disclosure* dimension in our proposed taxonomy. Overall, the distribution of selective disclosure among SSI approaches is quite balanced, with an inclination toward its support. Selective disclosure is often facilitated through the use of LDP-VCs, combined BBS+ or CL [127] signatures. Notably, there have been multiple draft proposals aiming to enable selective disclosure for JWT-VCs as well.

Lastly, for the final dimension *Credential Revocation*, we determine whether SSI approaches support credential revocation or not. It is apparent in our assessment that revocation is a crucial feature as it enjoys support from the majority of approaches. Only two approaches explicitly state their lack of support for this feature, while seven approaches do not provide clear specifications in this regard. Our findings here might be relevant for the upcoming chapter, where we discuss update mechanisms in VCs.

In conclusion, we have extracted some noteworthy findings from the construction of our proposed taxonomy, gaining actionable insights from the status quo of SSI approaches. It's important to note, however, that these observations are based on a limited sample size. Therefore, it would be beneficial to assess a larger number of approaches in future iterations for a more comprehensive analysis. Additionally, the use of "unspecified" characteristics within the taxonomy may be reconsidered in the future as solutions become more transparent in sharing implementation details or cease to obscure them.

6. Verifiable Credentials Update Mechanisms

In the proceeding chapter, we will discuss mechanisms aimed at enabling information updates for VC-based SSI approaches. We'll begin by outlining VC mechanisms related to information updates in this section. Typically, VCs are issued with a predefined lifespan, a deliberate choice due to the natural weakening of cryptography over time. Beyond cryptographic considerations, the information contained within VCs may also evolve, such as age-related attributes, the grade level in a student ID, or changes in a person's last name due to marriage. In most scenarios, VCs need to be updated once they reach their expiration date. The processes for updating traditional paper-based credentials in such cases are often complicated as they involve extensive form-filling and bureaucratic procedures. Below, we present some of the mechanisms in VCs attempting to address these challenges.

6.1. Short-lived Credentials

To start, we have the so-called *short-lived VCs*. The general idea is to consider reducing the validity period of VCs to ensure that their attributes remain relevant and accurate over time. This prompts the question: how brief should this period be? To answer this question, it might be worth it to examine the typical validity periods of other formats of credentials. For physical, paper-based credentials, this period lasts several years, e.g. passports and driver's licenses. This is acceptable as shorter active periods could lead to even greater administrative problems. On a more technical note, X.509 credentials may similarly have an active period for months or even years in the case of private Public Key Infrastructure (PKI). JWTs utilized in protocols like OAuth and OIDC often have relatively short lifespans, typically measured in minutes to a few hours.

Having short credential lifespans applies in scenarios in which credentials are difficult or impossible to revoke, such as in OAuth where revoking a token is not possible unless there is direct communication between the resource server and the authorization server. In a more extreme case, such credentials may also have been used for access control, i.e. granting and restricting access in high-security use cases in the form of access cards, badges, or electronic fobs. Such credentials are commonly used for guarding physical premises, even though updating or rotating these physical credentials entails more work and management.

While VCs often serve use cases that mirror physical credentials, adopting extremely short expiry periods on the scale of months rather than years may not effectively address

the core issue. Instead, it could potentially exacerbate operational challenges, resulting in a less-than-optimal digital user experience and potentially unnecessary operational overhead similar to that of physical use cases. To counter this, a mechanism that enables end users to more easily update or "refresh" their short-lived credentials is needed to make this approach feasible, which we will touch upon in another section.

6.2. Atomic Credentials

The notion of *atomic credentials* is not a new one. It has previously been explored in several forms within different contexts. A prime example of this is the EU's *Micro-Credentials* [128], aimed at certifying learning outcomes of learning activities to encourage lifelong learning. Through micro-credentials, short courses and employee training can be constructed in a way that is more targeted to the learners, while offering them a system to prove their achievements and competencies earned from such programs to future employers. The Council of the EU decided to adopt a Recommendation on a European approach to micro-credentials for lifelong learning and employability on 16 June 2022 to further push micro-credentials, which also feature in other agendas such as the European Pillar of Social Rights Action Plan and the Commission Communication on achieving the European Education Area by 2025.

Another example is atomic (Qualified) Electronic Attestations of Attributes analyzed in an ETSI technical report [129], which are single attribute claims issued by a (Qualified) Trusted Service Provider. Furthermore, the concept is also explored as a means of realizing progressive trust, a process enabling individuals to "gradually increase the amount of relevant data revealed as trust is built" [130].

The basic premise of such credentials remains the same in that issued VCs would only contain a subset of claims attesting to attribute(s) of a subject/holder. When needed altogether, e.g. for relying parties that require more attributes, these "atomic" VCs can be selected depending on the requirements and subsequently presented in a VP, effectively enabling selective disclosure.

The granular nature of such VCs also lends itself well to update mechanisms, as only the claim requiring an update would need to be refreshed or updated. This leads to much more control on the issuer's side as issuers could now specify terms applying to a VC with finer granularity, e.g. specifying different validity periods for different attributes and determining which attributes cannot be updated at all.

While the concept may appear promising, it comes with certain caveats. For instance, atomic VCs would be problematic for representing VCs with a substantial number of claims, resulting in additional service and, to a lesser extent, storage overhead. Moreover, it's possible for "fragments" from various VCs to be assembled in unintended ways, where attributes that don't naturally belong together are combined into a "Frankenstein" presentation, potentially leading to false claims [27, 131]. As such, atomic credentials alone cannot guarantee that claims are properly paired in a presentation and should not be trusted by verifiers unless additional mechanisms are introduced to prevent improper

pairing [129].

Another potential downside is the obfuscation of *negative credentials*, which are credentials containing negative information about the holder that may result in the removal of certain privileges or rights from them, e.g. points on a driver's license or points deduction on an exam result resulting in exam failure. Some possible solutions to this problem have been presented, including the binding of negative information with identity information and verifiers mandating the presentation of negative claims. These ideas are however only offered in an unofficial W3C draft [132] and should be explored further. Rigorous standardization is needed to ensure the compatibility of such credentials across a wide range of verifiers in a single domain, exemplified by the EU's micro-credentials program.

Ultimately, the concept of atomic credentials lends itself better to implementing selective disclosure, as explored by numerous works and discussions [133, 134, 135]. A noteworthy example of its application is explored by the FIDO Alliance in their white paper concerning the EUDI Wallet [136], where a combination of short-lived and atomic credentials in the form of short-lived EAAs with atomic claims for selectively disclosing credential claims is considered.

6.3. Credential Disputes

The next mechanism, referred to as *Disputes*, empowers entities to inform and urge issuers who have issued VCs containing inaccurate claims about them or another party, to rectify this information. This is a rather complicated situation since the issuer is technically the owner of the claims, but the holder has the knowledge of whether the issued claims are accurate or not. As such, the issuer needs to be convinced that the information under contention is indeed inaccurate and that it must be changed. Moreover, the issuer needs to make sure that the dispute is made by the actual holder of the disputed credential and not some impostor who might try to insert false information or remove true information [132]. Here, there are at least two different cases to consider [124].

In the first scenario, the subject/holder has been issued a VC containing claims based on false information and as such issues a `DisputeCredential`, exemplified by the code in Figure 6.1. The `id` field refers to the identifier of the disputed credential, with the value "Disputed" as the `currentStatus`. The reasoning behind the dispute must also be mentioned in the `statusReason` field.

In the second scenario, on the other hand, the `DisputeCredential` is issued by an entity other than the issuer to dispute a potentially false claim made by the issuer about a different subject. These entities are defined as "A thing with distinct and independent existence, such as a person, organization, or device that performs one or more roles in the ecosystem." [124]. Since the specific roles of these entities in SSI are not explicitly defined, we presume them to be, for instance, verifiers, watchdogs, or auditors within an SSI ecosystem. It's worth noting that these entities may choose to publish the `DisputeCredential` in a public venue to signal that the credential is under dispute, as mentioned in the editor's draft of the VC implementation guide [124]. However, this approach raises privacy concerns, as it

ß

```
{
   "@context": [
   "https://www.w3.org/2018/credentials/v1",
   "https://www.w3.org/2018/credentials/examples/v1"
   ],
   "id": "http://example.com/credentials/123",
   "type": ["VerifiableCredential", "DisputeCredential"],
   "credentialSubject": {
      "id": "http://example.com/credentials/245",
      "currentStatus": "Disputed",
      "statusReason": {
        "@value": "Address is out of date",
        "@language": "en"
      },
   },
   "issuer": "https://example.com/people#me",
   "issuanceDate": "2017-12-05T14:27:42Z",
   "proof": {...}
}
```

Figure 6.1.: An example of a `DisputeCredential` [124]

could potentially expose the identity of the credential subject, as guidelines regarding what can be disclosed as a dispute reason are not explicitly defined. Only recommendations from an unofficial W3C draft [132] on this matter have been made, recommending that VC issuers should provide subjects/holders means to obtain copies of the information held by the issuer and the right to correct erroneous information, both of which in accordance to GDPR regulations.

Nevertheless, this dispute mechanism is due for removal in the upcoming version of the VC Data Model [27] due to the lack of implementations in the previous two versions. Although it made it into the first two official renditions of the specifications, we noticed a general lack of discussion on the matter, with no noteworthy implementations of this mechanism. However, the mechanism lends itself particularly well in the aforementioned scenario where false or outdated information or claims have been made about the subject, especially when the false claim is particularly important to the issued credential or when the claim is of a sensitive nature in important credentials such as national ID cards, passports, and diplomas. Disputes additionally facilitate the rights of individuals or end users outlined by the GDPR, namely to lodge a complaint for correcting erroneous information.

6.4. VC Refresh Service

Credentials are designed to have a limited validity period. As such it would be useful for systems to have a mechanism that enables them to refresh the credential, e.g. in the case of a visa extension. The *refresh service* was therefore conceptualized, enabled by the

`refreshService` property in VCs which enables issuers to include a link to a corresponding service. Where this property is included is a point of distinction: Should the issuer wish to expose the service to the verifier, holder, or both, the property would then be included inside the VC. In contrast, if the service is only intended for the holder to use, the `refreshService` is then included in a verifiable presentation. In the latter case, it is possible for the holder to refresh the VC before presenting it to a relying party. Figure 6.2 depicts the inclusion of the property in a VC.

```
{
  "@context": [
    "https://www.w3.org/2018/credentials/v1",
    "https://www.w3.org/2018/credentials/examples/v1"
  ],
  "id": "http://example.edu/credentials/3732",
  "type": ["VerifiableCredential", "UniversityDegreeCredential"],
  "issuer": "https://example.edu/issuers/14",
  "issuanceDate": "2010-01-01T19:23:24Z",
  "credentialSubject": {
    "id": "did:example:ebfeb1f712ebc6f1c276e12ec21",
    "degree": {
      "type": "BachelorDegree",
      "name": "Bachelor of Science and Arts"
    }
  },
  "refreshService": {
    "id": "https://example.edu/refresh/3732",
    "type": "ManualRefreshService2018"
  }
}
```

Figure 6.2.: Example usage of the `refreshService` property [137]. Here, the service is accessible under the value of the `id` attribute.

This property simplifies the refresh process of a credential that is about to expire for subjects/holders in the second case. It is a bit more complicated in the first case, more specifically when the verifier triggers the service. Here, there are two main downsides as outlined in [25]. Firstly, the issuer would know that a VC has been presented to a certain verifier, a breach of privacy that goes against the core of SSI. Secondly, the issuer would need to grant authorization to this verifier for accessing the service. However, if the issuer is already authorized, it will possess the service details and won't require consulting the property again in the future, rendering it redundant.

Despite this, an unofficial W3C draft has been published on the topic called the *Verifiable Credential Refresh 2021* [138]. It proposes two further points of distinction, which are manual and automatic refresh. Both are expressed by the following attributes:

- **`url`** A mandatory URL used to initiate a credential refresh.
- **`type`** Expresses the type of refresh and must either be `MediatedRefresh2021` for

manual refresh or `UnmediatedRefresh2021` for automatic refresh.

- **`validAfter`** An optional date-time value to indicate the earliest point in time when a refresh can take place.
- **`validUntil`** An optional date-time value to indicate the latest point in time when a refresh can take place.

The refresh protocol depends on the type of the `refreshService`. The draft specifies two protocols, starting with the `MediatedRefresh2021` Protocol which is used for refresh procedures that are non-automatable. In contrast, the `UnmediatedRefresh2021` Protocol is used for automatable refresh procedures that do not require holder-issuer interaction. Both protocols follow the general workflow depicted in Figure 6.3.

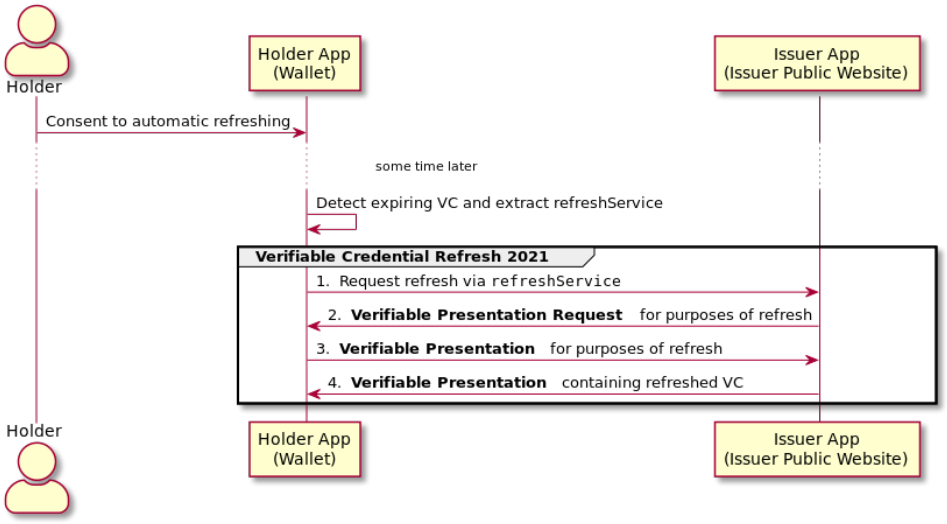


Figure 6.3.: The general workflow of the VC Refresh 2021 Protocol [138].

The `refreshService` mechanism is even more beneficial when combined with previously mentioned ones, offering upsides for both issuer and holder. Combining, for example, atomic credentials and an implementation of a refresh service offers users the possibility of refreshing atomic claims, whereas the usage of short-lived credentials and the refresh service also offers issuers a way to manage expired credentials in an efficient manner.

In general, we've identified the primary advantage of the `refreshService` as its convenience for the credential holder, especially when they are aware that the credential is about to expire. The inclusion of the property potentially simplifies matters for the issuer since they can issue an updated VC and simultaneously revoke the old one. However, it's important to note that this convenience comes with potential privacy concerns, as it could establish a connection between the verifier and the issuer, posing privacy risks for the holder. Therefore, it is recommended to consider using status mechanisms instead, as they share functionality with the `refreshService`. In fact, the latest published version of W3C's VC data model [137] explicitly states that "the refresh service is only expected to be used

when either the credential has expired or the issuer does not publish credential status information".

6.5. Conclusion

Concluding this section, our discussion on this topic reveals a common pattern: while various mechanisms have been explored in drafts and discussions, there is no strong momentum to solidify and standardize them specifically for handling VC updates. These mechanisms remain somewhat incomplete and insufficiently investigated, except for atomic credentials and disputes. Atomic credentials have gained significant attention, particularly the EU's *Micro-Credentials* in the education sector, while disputes are facing potential removal in future versions of the W3C VC specification [27], suggesting that implementers and standard setters may not currently consider them a critical feature.

A common thread among these update mechanisms is that they ultimately resort to a simple approach of revoking the credentials and then reissuing them. However, even the revocation process remains an ongoing research topic. As the SSI user base expands, revocation alone may not suffice, as stated by Bochnia et al. [57]: "While revoking and issuing an updated credential may be suitable for certain use cases, it may not be feasible in situations where the modification is performed by a party other than the original issuer. This approach can also lead to confusion for users. [...] Furthermore, revoking and reissuing a VC can impact trust in the VC, as it suggests that there was a problem with the original VC and raises questions about its reliability".

Nevertheless, our research has led us to the temporary conclusion that handling credential updates is currently not a significant concern in most use cases.

7. Engineering Effective Identity Credentials within GX-Credentials

7.1. Selective Disclosure and VC Encoding Formats

Drawing from the taxonomy insights and in a broader context, Selective Disclosure (SD) emerges as a pivotal aspect of SSI. This feature empowers users by allowing them to choose precisely which data to share, ensuring that no data is divulged without their explicit consent. As such, it is a substantial research topic that can definitely merit an entire thesis on its own. In our attempt to apply insights gathered from the taxonomy to engineer compliant and privacy-preserving identity credentials within the context of the Gaia-X project named *GX-Credentials*, we begin by exploring SD and its significance to the project.

From our survey, current SD methods can be classified into three general approaches: *atomic credentials*, *hashed values*, and *SD signatures* [124]. In the preceding chapter, we discussed atomic credentials and how they lend themselves well to enable SD, however, they are generally expensive memory-wise and computationally as they would require as many signatures as the number of claims [139]. Similarly, the computational cost associated with SD signatures is known to be notably high. That leaves us with hash-based SD approaches, which we will discuss in the following subsections after a brief overview of two SD signature schemes.

With regard to our proposed taxonomy, we found that a lot of approaches do not offer or consider SD (15/35 approaches) despite its importance. We suspect that this might be linked to the VC proof format. Historically, this feature has been enabled through LDPs and specific signature schemes that natively support SD, such as the Camenisch-Lysyanskaya (CL) [127] and Boneh, Boyen, and Shacham (BBS) [140] signature schemes. Implementing these signature schemes has been demonstrated to be intricate and demanding, which might encourage implementers to prioritize other features over SD.

Although there are other techniques for enabling SD such as *predicates* [124] and *ZK-SNARKS* [141], we will mainly explore the hash-based approach by discussing the two main VC proof formats in accordance with our taxonomy and their corresponding proof system for enabling SD.

7.1.1. Selective Disclosure Through Linked-Data Proofs

The first proof format we would like to discuss is VCs with Linked-Data Proofs, also called *LDP-VCs*. The VC is encoded in JSON-LD format with a `proof` attribute. Proofs or signatures

can be included in the VC externally (e.g. via JWT) or embedded (proof is included in the data via the `proof` attribute). Such a VC can be broadly depicted as shown in Figure 7.1, which illustrates a JSON-LD encoded VC with an embedded proof. We now look at two cryptographic schemes of particular interest, starting with the CL signatures.

```
{
   "@context": [
      "https://www.w3.org/2018/credentials/v1",
      "https://www.w3.org/2018/credentials/examples/v1",
      "https://w3id.org/security/suites/ed25519-2020/v1"
   ],
   "id": "http://example.edu/credentials/3732",
   "type": [
      "VerifiableCredential",
      "UniversityDegreeCredential"
   ],
   "issuer": "https://example.edu/issuers/565049",
   "issuanceDate": "2010-01-01T00:00:00Z",
   "credentialSubject": {
      "id": "did:example:ebfeb1f712ebc6f1c276e12ec21",
      "degree": {
         "type": "BachelorDegree",
         "name": "Bachelor of Science and Arts"
      }
   },
   "proof": {
      "type": "Ed25519Signature2020",
      "created": "2022-02-25T14:58:43Z",
      "verificationMethod": "https://example.edu/issuers/565049#key-1",
      "proofPurpose": "assertionMethod",
      "proofValue": "..."
   }
}
```

Figure 7.1.: Example of a JSON-LD encoded Verifiable Credential with LDP, adapted from [137].

Camenisch-Lysyanskaya Signatures

Jan Camenisch and Anna Lysyanskaya [127] presented a signature scheme and protocols that would be suitable as a building block for other applications, notably anonymous credential systems. As such, the signature scheme is adopted for signing AnonCreds, the default credential profile in all Hyperledger Indy implementations and libraries [61, 62]. The scheme is based on the Strong RSA assumption and relies on the difficulty of factoring the multiplication result of two large prime numbers, thus requiring lengthy keys and signatures to ensure a sufficient level of security, a requirement which unfortunately results in slow cryptographic operations [142]. To address this challenge, this signature

scheme evolved into the BBS+ signature scheme, as detailed in the following section.

BBS Signatures

In contrast to CL signatures, BBS signatures rely on the difficulty of the discrete logarithm problem and q-Strong Diffie Hellman (q-SDH) with pairing-based Elliptical Curve Cryptography (ECC). BBS requires shorter keys and signatures compared to CL signatures with its usage of elliptic curves [142] and enables selective disclosure as well as unlikable properties through elliptic curve pairings [143]. This is outlined by a pilot ZKP implementation from a company called Mattr that is based on BBS+ signatures using BLS12-381 curves, which generally require significantly smaller key and signature sizes than CL signatures [144]. Other than enabling selective disclosure, BBS signatures provide additional properties such as unlikable proofs and proof of possession, highlighted by this draft specification of the BBS signature scheme [145] that is still in the review process. A demo web application [146] developed by Dr. Greg Bernstein illustrates the usage of BBS signatures, enabling the creation and signing of VCs with this signature scheme as well as deriving and verifying selective disclosure proofs from a VC.

7.1.2. Selective Disclosure Through JWTs

The second proof format involves encoding JSON or JSON-LD VCs as a JWT. A JWT consists of three parts: the header, payload, and signature. The header contains information about the signing algorithm being used defined in the `alg` attribute, a `kid` attribute which may refer to a key in a DID document, and the type of the token, which is simply "JWT". This JSON is exemplified by the snippet in Figure 7.2 and will then be encoded as standard JOSE header parameters, forming the first part of the JWT.

```
{
   "alg": "RS256",
   "typ": "JWT",
   "kid": "did:example:abfe13f712120431c276e12ecab#keys-1"
}
```

Figure 7.2.: Example of a JWT header of a JWT-based VC using JWS as proof, adopted from [137].

The second part is the payload, which contains different types of claims that must be included to properly express the VC as well as ensuring backward compatibility with JWT processors. These claims are as follows:

sub *Subject*, represents the id property of the `credentialSubject` in the VC.

jti *JWT ID*, represents the id property of the VC.

nbf *Not Before*, represents the `issuanceDate` of the VC and encoded as a UNIX timestamp.

exp *Expiration Time*, represents the `expirationDate` property, also encoded as a UNIX timestamp.

iss *Issuer*, represents the issuer of the VC.

aud *Audience*, represents the intended audience of the VC, which can either be the holder when issuing a VC or the verifier when the holder presents a VP.

iat *Issued At*, represents the time at which the JWT was issued as a UNIX timestamp.

All of these claims are illustrated in Figure 7.3 along with the actual content of a VC or VP in the `vc` or `vp` claim. It is noteworthy that they could also be contained in the JWS part of the JWT, which makes up the final part of the JWT, which is created by signing both the encoded header and payload using the specified algorithm in the header with a secret. This JWS proves that the issuer of the JWT signed the contained payload.

```
{
   "sub": "did:example:ebfeb1f712ebc6f1c276e12ec21",
   "jti": "http://example.edu/credentials/3732",
   "iss": "https://example.com/keys/foo.jwk",
   "nbf": 1541493724,
   "iat": 1541493724,
   "exp": 1573029723,
   "nonce": "660!6345FSer",
   "vc": {
      "@context": [
         "https://www.w3.org/2018/credentials/v1",
         "https://www.w3.org/2018/credentials/examples/v1"
      ],
      "type": ["VerifiableCredential", "UniversityDegreeCredential"],
      "credentialSubject": {
        "degree": {
          "type": "BachelorDegree",
          "name": "Bachelor of Science and Arts"
        }
      }
   }
}
```

Figure 7.3.: Example of a JWT payload of a JWT-based (JSON-LD) VC using JWS as proof, adopted from [137].

The Internet Engineering Task Force (IETF) has recently published a specification for supporting selective disclosure of JWT (*SD-JWT*) claims [147]. It is a hash-based approach to enabling SD, granting the issuer the ability to predefine which claims within the VC can be selectively disclosed or not. This particular approach to enabling SD is noteworthy as even with its specification being in draft status, SD-JWT is mandated by the ARF (Architecture and Reference Framework) as the de facto JSON format for enabling SD alongside the ISO mDL MSO for expressing PID (Person Identification Data) and (Q)EAAs

within EUDI Wallet solutions [148]. Therefore, we will delve deeper into this approach in the subsequent text, in accordance with the specification.

SD-JWT is compatible with various JSON-based representations of claims, including JSON-LD. Claims are in the form of key-value pairs, with values such as strings, arrays, and objects. The general workflow plays out as follows: Firstly, the issuer uses a claim set in which it is specified which claims are selectively disclosable. To do this, the issuer creates a *Disclosure* for each claim in the form of an array, which consists of the following elements:

A salt value A base64url-encoded string with a minimum of 128 bits that is cryptographically secure.

Claim name The key of a claim in the form of a string.

Claim value The value of the corresponding claim key may be of any type that is allowed in JSON.

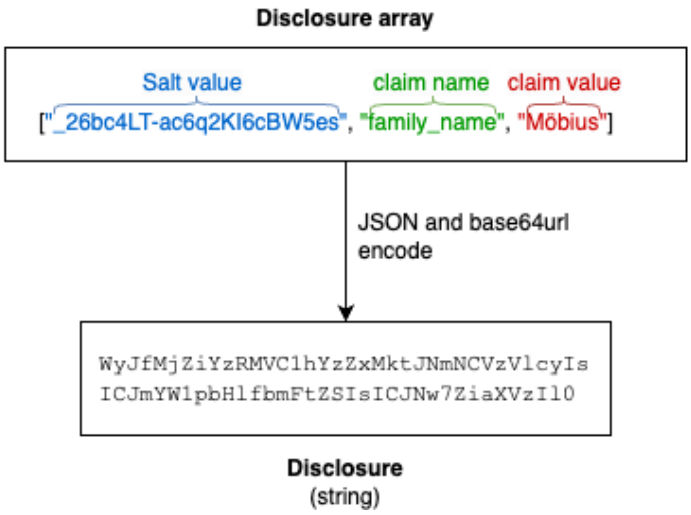


Figure 7.4.: An example of an SD-JWT Disclosure, inspired by [129].

The array takes the structure depicted in Figure 7.4. Subsequently, this array is JSON-encoded into a UTF-8 string, which is finally base64url-encoded into the disclosure. Disclosing claims in the form of elements in an array and objects is done differently, with several different options to disclose the latter selectively to handle nested data. Once all Disclosures have been created for the selected claims, they are included as an element of an array in the form of a message digest to hide the original value of the Disclosure. This array is the value of the `_sd` claim in the SD-JWT. The ordering of the elements within the array should also be randomized so that the order of the disclosures does not stay the same every time. An example of a claim set is depicted in Figure 7.5, while the payload of an SD-JWT is shown in Figure 7.6.

```
{
   "sub": "user_42",
   "given_name": "John",
   "family_name": "Doe",
   "email": "johndoe@example.com",
   "phone_number": "+1-202-555-0101",
   "phone_number_verified": true,
   "address": {
      "street_address": "123 Main St",
      "locality": "Anytown",
      "region": "Anystate",
      "country": "US"
   },
   "birthdate": "1940-01-01",
   "updated_at": 1570000000,
   "nationalities": [
      "US",
      "DE"
   ]
}
```

Figure 7.5.: An example of a claim set determined by the issuer as an input for SD-JWT, adopted from [147].

Furthermore, there is an optional Key Binding (KB) JWT mechanism that the presenter of the SD-JWT may include to prove that they are indeed the holder of the VC. It contains a public key or a reference to it which corresponds to a private key owned by the holder. The verifier would then require the holder to prove possession of the aforementioned private key when presenting the SD-JWT credential. Notably, without including key binding, the verifier would be able to confirm that the VC was issued by a specific issuer but would not prevent the credential itself from being replayed by anyone who has access to it. To send the SD-JWT, the standard requires that the SD-JWT itself and the disclosures be serialized into base64-url encoding, separated by a tilde, exemplified as follows: `<SD-JWT>~<Disclosure 1>~<Disclosure 2>~...~<Disclosure N>~<optional KB-JWT>`. For presentations, the holder would then send a selected set of attributes to disclose to the verifier.

7.1.3. A Brief Comparison of LDP and JWT-Enabled Selective Disclosure

In a general sense, the combination of JSON-LD and LDP is designed to express context and semantic richness by resolving to an external document, providing verifiers insights into how a VC should be interpreted and processed, thus ensuring semantic interoperability in an open data world [129]. Nevertheless, despite the promise of both CL and BBS signature schemes enabling SD, the security comes from a reliance on the difficulty of factoring multiplications of large prime numbers (CL signatures) and elliptical curve cryptography (BBS signatures). The analysis conducted by standards setter ETSI also noted that a CL-

```
{
   "_sd": [
      "CrQe7S5kqBAHt-nMYXgc6bdt2SH5aTY1sU_M-PgkjPI",
      "JzYjH4svliH0R3PyEMfeZu6Jt69u5qehZo7F7EPYlSE",
      "PorFbpKuVu6xymJagvkFsFXAbRoc2JGlAUA2BA4o7cI",
      "TGf4oLbgwd5JQaHyKVQZU9UdGE0w5rtDsrZzfUaomLo",
      "XQ_3kPKt1XyX7KANkqVR6yZ2Va5NrPIvPYbyMvRKBMM",
      "XzFrzwscM6Gn6CJDc6vVK8BkMnfG8vOSKfpPIZdAfdE",
      "gbOsI4Edq2x2Kw-w5wPEzakob9hV1cRD0ATN3oQL9JM",
      "jsu9yVulwQQlhFlM_3JlzMaSFzglhQG0DpfayQwLUK4"
   ],
   "iss": "https://example.com/issuer",
   "iat": 1683000000,
   "exp": 1883000000,
   "sub": "user_42",
   "nationalities": [
      {
        "...": "pFndjkZ_VCzmyTa6UjlZo3dh-ko8aIKQc9DlGzhaVYo"
      },
      {
        "...": "7Cf6JkPudry3lcbwHgeZ8khAv1U1OSlerP0VkBJrWZ0"
      }
   ],
   "_sd_alg": "sha-256",
   "cnf": {
      "jwk": {
        "kty": "EC",
        "crv": "P-256",
        "x": "TCAER19Zvu3OHF4j4W4vfSVoHIP1ILilDls7vCeGemc",
        "y": "ZxjiWWbZMQGHVWKVQ4hbSIirsVfuecCE6t4jT9F2HZQ"
      }
   }
}
```

Figure 7.6.: An example payload used in an SD-JWT, adopted from [147]. Elements in the `_sd` claim array are randomized from the original claim set and are digests of their corresponding Disclosures.

proof system exists for all NP problems. Furthermore, BBS signatures have not received approval from the Senior Officials Group Information Systems Security (SOG-IS) due to concerns that they may not be quantum-safe, quantum-safe, i.e. systems or algorithms that remain secure in the era of quantum computing.

On the other hand, JWTs are more mature than LDP and ubiquitous, used and supported by a broad range of libraries. Additionally, it is smaller in size and there is no need to resolve to an external document or an internet PKI to process JWTs, therefore offering offline support. Further distinctions arise when combining different encodings with proof formats as outlined by W3C's VC implementation guide [124].

However, as underscored by the first version of the ARF [148], a set of common standards

and technical specifications created by the European Commission that serves as the basis for the EUDI Wallet, JWTs are preferred due to their relative simplicity and maturity, as well as being post-quantum safe. Moreover, the usage of SD-JWT for PID attestations is listed as a configuration requirement for EUDI Wallet Solutions. This fact holds considerable importance, as the same document is adopted by the eIDAS Expert Group, which will impact the SSI landscape in the EU.

This does not imply a complete abandonment of LDPs, as the ARF states that EUDI Wallet Solutions will support two initial configurations: *Type 1* and *Type 2*. The first configuration type is intended for PID purposes, namely use cases where relying parties require a high level of identity assurance. In contrast, the Type 2 configuration aims to enable flexibility and additional feature support for (Q)EAA use cases that cannot be met by Type 1 configuration, for example, credentials in the education and health domains.

Despite being stated as a requirement for EUDIW Solutions by the ARF, SD-JWT is still a draft specification at the time of its inception. ETSI outlines two possible alternatives on how to secure VCs with SD-JWT given the current state of the (draft) specifications involved:

1. Including the VC based on the VCDM v1.1 in the `vc` claim and mapping VC claims to JWT claims, effectively duplicating their values

2. Using only SD-JWT and relying on a transformation algorithm so that the verifier could transform a presented SD-JWT into a VC.

The first option might seem straightforward to implement but there are conflicts that exist between the current W3C VC Data Model 1.1 and SD-JWT specifications that fail to reap the benefits from, the utilization of both LDP and JWT-based VCs. Furthermore, confusion arises for example when using SD-JWT with JWT-secured JSON-LD-based VCs as specified in VCDM v1.1. Such confusion is bound to happen as JSON-LD itself was not designed with an extension to SD-JWT in mind. Moreover, this approach would entail encoding the selectively disclosable claims twice, resulting in inefficiencies.

Ultimately when analyzing the application of SD-JWT to eIDAS 2.0, ETSI recommends the latter, i.e. the usage of SD-JWT as a standalone attestation format, due to the problems associated with the former option. An approach to mapping mandatory claims from a VC to a JWT is also specified by the *SD-JWT-based Verifiable Credentials (SD-JWT VC)* draft specification [149] for the time being.

Although the ARF document holds no legal validity and its contents may undergo revisions until the finalization of the EUDI Framework Regulation, it along with the ETSI technical report on selective disclosure offered valuable insights that have led us to the conclusion that proceeding with SD-JWT is the way forward due to its quantum-safe property, inherent simplicity in comparison with LDPs for including PID of natural subjects in VCs, and the fact that it is the EU Commission's chosen proof format for EUDIW Solutions. The acquired insights will additionally serve as the foundation for our implementation within the framework of the GX-Credentials project, outlined in the proceeding section.

7.2. Gaia-X: GX-Credentials

7.2.1. Preliminaries

We have given an overview of Gaia-X as a whole in the Related Work Chapter 3, more specifically section 3.6.3. There, we also mentioned that the Gaia-X Federation Services or GXFS constitute a major pillar of the initiative, often touted as the foundation of the Gaia-X framework [150]. GXFS is a project funded by the BMWK (German Federal Ministry of Economics and Climate Protection) that extends the Gaia-X framework by acting as a toolbox to provide a reference implementation for interested parties in supporting the inclusion of participants in the federated data ecosystem and ensure its interoperability with services from other federations [151].

Federations and *Self-Descriptions* are core concepts of GXFS. The former is a group of *Participants* that collaborate by exchanging services and data, owned by the collective based on a joint set of rules based on different industries. Functionality-wise, Federations are based on Self-Descriptions, which are simply Verifiable Credentials that represent entities or Participants in the ecosystem as well as their service offerings. As such, Self-Descriptions inherit qualities from VCs, more specifically VCs encoded in JSON-LD format.

Example implementations include projects in the project family called *Gaia-X 4 Future Mobility*, which is aimed at developing Gaia-X-based applications and services in the mobility sector. It currently comprises six projects funded by the BMWK: *Gaia-X 4 AI*, *Gaia-X 4 AMS*, *Gaia-X 4 ROMS*, *Gaia-X 4 PLC-AAD*, *Gaia-X 4 moveID*, and *Gaia-X 4 AGEDA*, run by around 80 companies and research institutions including TUM [152]. We are particularly interested in the Gaia-X 4 PLC-AAD (Product Life Cycle - Across Automated Driving) project, which aims to “[…] establish an open and distributed data ecosystem that supports future product development, manufacturing, and customer service” [153], hence encompassing the entire product life cycle of automated driving functions. In the Gaia-X 4 PLC-AAD GitHub page [154], there is a repository called *GX Credentials* [155], where we explore an approach for issuing VCs to companies and their employees in a Gaia-X ecosystem. These credentials will then be used by employees to authenticate with various services in the ecosystem. Due to their importance, it is crucial to ensure that these identity credentials are well-designed, taking both legal and technical aspects into consideration with the ultimate goal of creating an effective identity credential. In the following subsections, such aspects are explored, starting with an overview of the project.

7.2.2. Overview of GX-Credentials

The GX Credentials project is a web app-based approach that enables the issuance of VCs to companies and their employees within a Gaia-X ecosystem, enabling employees to authenticate with different services in the ecosystem using VCs. There are three main stakeholders as a part of the project’s user stories, namely the *Operator or Registrars*, *Company*, and *Employee*. Registrars can be thought of as administrators who host the web app itself, taking the role of a trusted entity within a dataspace or consortium and enabling

identity management within a dataspace, with the operator only directly certifying company identities. Once companies are certified by obtaining their company credential, they are now able to issue employee credentials to their employees. Overall, both companies and their employees are able to apply for their respective credentials, which will then be approved or rejected by the Trust Anchor or company respectively.

Architecture Components

The GX Credentials application is built upon three main architectural components:

- **Client and Server Side** a **Next.js** app that provides both a client-side user interface along with a server for session management and authenticated database operations.
- **Database** Firebase's Cloud storage service, **Firestore**, is used to store credential applications and issued credentials.
- **Smart Contract** a registry smart contract managed by the trust anchor to enable credential issuance logging and revocation.

For development purposes, the registry smart contract was deployed on the Tezos Ghostnet network to best simulate production behavior as the Tezos Mainnet network is used by Gaia-X. End-to-end testing, involving the signing and storage of issued credentials, was carried out using the *Altme Wallet*, an SSI wallet app. This choice was influenced by its compatibility with SSI and its support for the **Beacon Protocol**, a critical component employed by GX Credentials for user authentication as well as various VC and DID-related functionalities. To enable this, the DIDKit library by Spruce ID was utilized as it offers VC and DID functionality across various platforms related to the Beacon Protocol.

The selection of the technology stack and architectural components aligns with the project's experimental nature, as it is still in development and not intended for production use. For instance, the application currently stores issued credentials in Firestore, which, in its default configuration, is publicly accessible to anyone with the URL. While these credentials should ideally be deleted from the database after issuance, this behavior has not been implemented in the current state of the app. Furthermore, the development of a trusted issuer's list which makes company credentials publicly accessible through future credential verifier software is still pending. Of particular relevance to our thesis is the use of placeholder data for the credential application form fields and the subsequent credential issuance, left as a working item to be finalized at a later point. We will discuss this aspect further in the following text.

Overview of Identity Credentials in GX Credentials

All entities and their service offerings in the Gaia-X federation are expressed through Self-Descriptions, including stakeholders in the GX Credentials project. These VCs adhere to basic well-defined schema types that ensure a basic level of interoperability between federations. Within a federation, additional sets of rules can be added by defining more

VC types to better describe Gaia-X Participants. According to the basic model depicted in Figure 7.7, a Participant can either be an instance of a `LegalPerson` or a `NaturalPerson`.

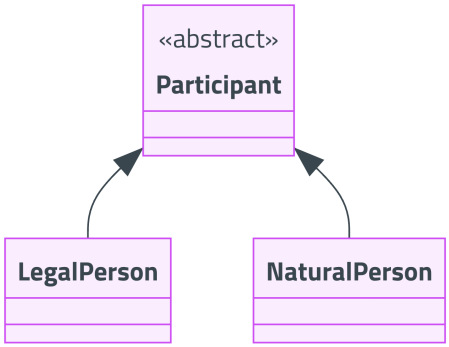


Figure 7.7.: A simple model hierarchy for Gaia-X Participants, adopted from [156].

In the Gaia-X Trust Framework, a fundamental schema is established that encompasses legally binding attributes for the former, as depicted in Figure 7.8, but the framework does not define one for the latter. It does, however, make reference to methods for remote validation of natural persons within a federation, including approaches like WebAuthn with FIDO2 dongles and the utilization of Android applications integrated with the Google Play Integrity API [156]. These methods are still in the experimental phase and will probably be finalized in the next version of the document.

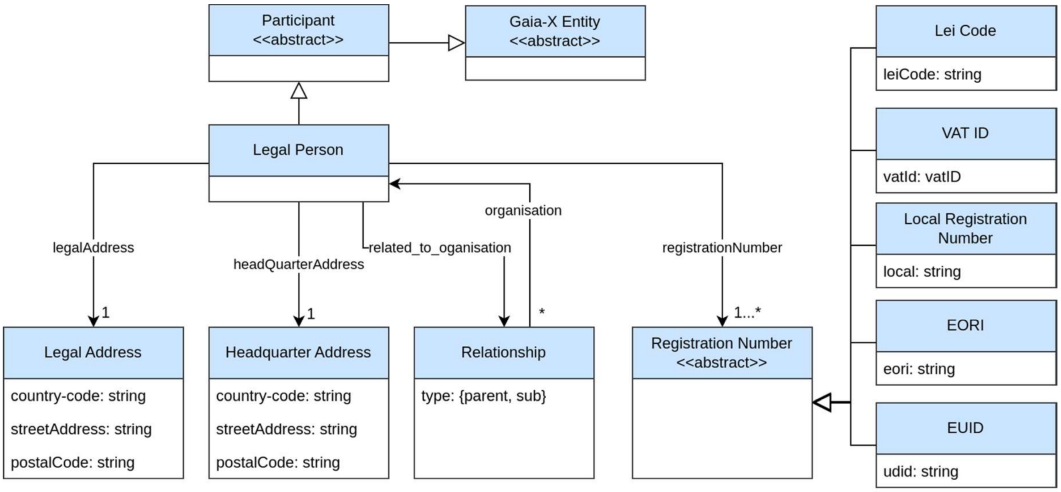


Figure 7.8.: A model describing the basic schema for Gaia-X Participants, adopted from [157].

Within the context of the GX credentials project, both the `CompanyCredential` and `EmployeeCredential` credentials follow the schema defined by the Gaia-X Registry [158] project, a prototype for a registry of trust anchors as defined in the Trust Framework, that

as of writing this thesis is accessible through a *URL*. The URL returns a JSON-LD context for all available shapes in the registry. Based on this, employee credentials include the following attributes:

- **`type`** all employee credentials possess the type `gx:LegalParticipant`
- **`gx:legalName`** the employee's legal name
- **`id`** corresponds to the employee's DID stored in their wallet
- **`gx:legalRegistrationNumber`** as defined in the Trust Framework, there are several valid entity registration numbers such as `EUID` (European Unique Identifier for businesses) and `leiCode` (a unique LEI number as defined by GLEIF). A registration number possessed by both employees and companies, the **`gx:vatID`** or VAT identification number, is used in the project.
- **`gx:issuerCompanyName`** the name of the company that issued the credential.
- **`gx:issuerCompanyID`** the DID that corresponds to the aforementioned company.
- **`gx-terms-and-conditions:gaiaxTermsAndConditions`** a SHA256 hash of the terms and conditions document that outlines the requirements for participants to provide accurate information about themselves, agreed to by the employee.

As our thesis is concerned with natural subjects, we are particularly drawn to the employee credential schema. Employee credentials in general can be considered as a rather unique case as they straddle the intersection of various identity contexts, requiring them to effectively express both natural and legal persons. This matter is discussed in the following section along with a discussion on the identity credential schema itself.

7.2.3. Considerations for Identity Credentials

In this section, we evaluate the current employee credential schema used in the GX Credentials project and identity credentials for natural subjects from various perspectives. To facilitate this assessment, we formulated three key questions:

1. How does the current GX credential fare according to the proposed taxonomy from Chapter 5?
2. What legal regulations and technical frameworks must be taken into account for the design of identity credentials?
3. Which data or attributes should be integrated into an employee's identity credential in GX Credentials?
4. What features need to be implemented in the GX Credentials project to facilitate the revised employee credential scheme?

Hence, we've divided this section into three segments, each focused on answering this query.

Assessing the GX Credential with Respect to the Proposed Taxonomy

We start by addressing the first question. While our proposed taxonomy was originally formulated to effectively differentiate between SSI approaches in terms of user identification and identity credentials, this framework can be extrapolated to employee credentials given their dual role as both natural and legal persons. As such, we assessed the GX Credentials approach for employee credentials using our taxonomy and obtained the subsequent results presented in Table 7.1, concluding that our assessment provides a good overview of where the GX Credentials project stands as an SSI approach.

Dimension	Assigned Characteristic(s)	Comment
PII Location	*Bundled*	The credential's main purpose is to identify employees within a specific context, such as for activities where the employee needs to act on behalf of the company.
PII Type	*Natural, Alternative*	The employee credential schema includes both natural and alternative identifiers as attributes such as `id` and `gx:legalRegistrationNumber` are present, referring to the employee and issuing company DID respectively.
Identification Data Source	*Gov-ID, Non-Gov-ID*	Although specifics of the employee onboarding process are not determined in the user stories as it depends on each federation and company, we assume that government-issued identity documents will be used to validate the employee's identity along with documents from other institutions.
Identification Authority	*End-user asserted*	Employees make claims about their identity attributes by filling out an application form for the credential, which will then be verified by the company.
Projected Cost per User	*Free*	Fees associated with credential issuance are most likely to be paid for by the employer.
VC Format	*LDP-VC*	GX Credentials are signed by the prospective holder as a part of the issuance process as a part of the Beacon protocol, creating a proof with the type `TezosSignature2021`.
Schema Standard	*Flexible*	Although some attributes in the credential are standardized (e.g. `vatID`), the format itself doesn't, to the best of our knowledge, adhere to a specific credential standard.
Selective Disclosure	*Unsupported*	This feature is not supported.
Credential Revocation	*Supported*	Revocation is made possible through the deployed smart registry contract. However, the revocation of company credentials does not explicitly flag corresponding employee credentials as revoked, a feature that is still to be implemented.

Table 7.1.: GX Credentials evaluated according to the proposed taxonomy

Framework and Legal Considerations

Due to the sensitive nature of attributes that are included in identity credentials, legal regulations have been established to set a baseline for securing such credentials, especially considering that governments are positioned as natural issuers of machine-readable VCs which will be utilized to represent authoritative proofs in an attempt to comply business practices and regulatory requirements [159]. This claim is further backed by our findings from the taxonomy, with the majority of examined SSI approaches requiring government-issued documents for identity validation as well as adhering to credential standards. While conducting our literature survey, two primary works emerged as particularly relevant to identity credentials: the **ARF** and the **eIDAS Regulation**.

Of particular significance is the forthcoming **EUDI Wallet based on the ARF**, with each member state mandated by the EU Commission to provide at least one implementation. A commentary article from two experts in the field namely Steffen Schwalm and Andre Kudra [160] provided us with a comprehensive overview of the matter and its potential relevance to GX Credentials, which we will discuss in the proceeding paragraph.

EUDI Wallet Solutions can be offered by governmental organizations responsible for issuing identity documents and other attestations, making them potential PID, QEAA, and EAA providers. While such solutions may not replace existing national digital ID solutions such as Estonia's *eID*, Belgium's *itsme*, and Italy's *SPID*, they can serve as a bridge or translation mechanism for transferring eID tokens from the OIDC context to Person Identification Data (PID), made possible through the use of OID4VCI and SD-JWT or mdoc format as specified by the mandatory Type 1 config of the ARF. One of the key advantages of bridging these approaches is that it addresses the issue of underutilized eID functions. For instance, in Germany, identity cards with eID capabilities have been available since 2010 but are hardly used in various processes and services, such as logging into the München website for visa extensions. The streamlined and standardized workflow offered by the EUDIW approach can help unlock the potential of eID functionalities across the EU.

Moreover, the EUDIW has the potential to simplify onboarding processes by leveraging its legal power and standardization. It is supported by all EU member states, which promotes interoperability not only for the public sector but also potentially for the private sector. This means that organizations can issue VCs to both natural and legal persons through their EUDIW. However, there is a need for clarification on the conditions under which an issuer of digital evidence, such as Attestation of Attributes, must be a qualified trust service, especially considering existing and planned liability obligations for (Q)TSPs. Currently, the onboarding processes and the derivation of national ID cards to electronic IDs are outlined by BSI TR-03159 specifications [161]. However, these specifications do not explicitly define how this process should look for VCs and EUDIW PID, highlighting the need for further development and clarification in this area.

In the current GX Credentials employee onboarding workflow described in the user story, companies are required to contact the employee "out of band" to confirm their suitability for membership, identity, and intent. The exact process for this confirmation is not explicitly

outlined and is intended to be determined by the company. Additionally, companies also act as verifiers and must facilitate verification processes related to EUDIW, making use of PID attributes to establish trust and ensure the reliability of the information provided by employees [162].

Furthermore, companies are legally acknowledged by Qualified Electronic Attestation Authorities (QEAAs). This means that they can provide VCs based on the Type 2 config to fulfill their needs, with the Type 1 configuration being reserved for PID applications. Nevertheless, companies can also be considered trust services as they allow ongoing digital transaction relationships between EU member states, natural persons, and legal entities. EUDIW, based on eIDAS 2.0, enables users to create Qualified Electronic Signatures (QESignatures) that hold the same legal effect as handwritten signatures in a court of law within the EU [163]. However, as mentioned earlier, it is still unclear under which conditions issuers would need to be a (Q)TSP. The ARF states that the Type 2 config can be used to support QEAAs that are not met by the Type 1 config, underlining the need for further clarification and alignment with existing legal and contractual frameworks.

The **eIDAS Regulation** plays a significant role in *electronic identification* matters, with the term defined in Article 3 [72] as "the process of using person identification data in electronic form uniquely representing either a natural or legal person, or a natural person representing a legal person". Naturally, given its relevance and the rise in SSI's popularity, attempts have been made to connect the two, such as the eIDAS Bridge which aims to enhance the legal certainty of any class of VC, by incorporating the issuer's advanced or qualified electronic signature [164]. This legal assurance aligns with Article 17 of eIDAS, which encourages the private sector to use eID means under a notified scheme [72, 165].

Within the context of the GX Credentials project, it makes sense for registrars or admins working at trust anchors who are responsible for onboarding companies into the ecosystem to leverage identifiers already used in digital certificates issued by TSPs. This ensures that when credentials are issued to companies, the companies will have been issued VCs that are authorized by relevant EU laws.

A similar approach is proposed in the Technical Convergence report by the Data Spaces Business Alliance (**DSBA**) [165], where they assume that the issuer, who is also a Participant in the data space, already possesses an eIDAS certificate. Consequently, every VC issued within the ecosystem is signed using digital certificates adhering to the JAdES format as outlined in ETSI TS 119 182-1. This ensures legal validity and interoperability for cross-border data-related transactions. The report also covers identifiers for identity binding for legal persons, proposing the derivation of DIDs from identifiers that are already embedded in the eIDAS certificates, certificates that conform to relevant standards. The DID derivation process takes the `organizationIdentifier` specified in Recommendation ITU-T X.520 and ETSI EN 319 412-1 V1.4.2 (2020-07) to create a DID in the following format:

`did:elsi:organizationIdentifier`

An example of how such a DID would look within the Gaia-X context is as follows: `did:elsi:VATBE-0762747721`. Here, `elsi` stands for *ETSI Legal Semantic Identifier* or

ELSI, an acronym for the name for this type of identifier used in the ETSI documents. `VAT` in the `VATBE-0762747721` or the `organizationIdentifier` refers to identification based on a national value added tax identification number, and the rest of the identifier after "-" being an identifier according to country and identity type reference. Overall, this process creates a bidirectional mechanism to derive DIDs from the eIDAS digital certificate and vice versa. Any legal person can therefore have a standard eIDAS certificate with an automatically associated DID identifier complying with the ELSI DID method specification and thus foregoing the need to invent new identifiers or have a central entity in a data space assign identifiers to participants [165].

Currently, as outlined by the Gaia-X Architecture Document, TSPs along with GX Label Issuers and Trusted Data Sources utilized by issuers to validate attribute attestations before credential issuance, establish *Trust Anchors* responsible for confirming the identity of ecosystem Participants. These Trust Anchors are nominated and operated by the Participants themselves, although the Trust Framework specifies a list of defined Trust Anchors which include eIDAS as the most prominent among them. To ensure compliance, it is imperative that the TSPs mentioned in the documentation align with those specified in eIDAS. Specifically, these TSPs should be either state-validated identity issuers or temporarily valid TSPs for Extended Validation Secure Sockets Layer (EV SSL) issuance. It is also important to note that the Trust Anchor referred to here is not the same as the one in the GX-Credentials project. As such, identifiers in the Gaia-X ecosystem can be replaced with the aforementioned derived DID from the DSBA report.

Necessary Attributes for Employee Identity Credentials

We now arrive at the final question. To try and answer the question, we first consulted W3C's VC data model, where sections relevant to the matter (e.g. Privacy Considerations) are mostly non-normative. As such, we won't be discussing it in further detail, as this matter is not the primary concern of the specifications. Nevertheless, we mainly concluded from the data model that the principle of data minimization, i.e. limiting the information included in VCs, should be adhered to and is considered a best practice that should be considered.

The more recently released **ARF** specifies a set of attributes for identifying natural persons, which considers both the eIDAS minimum and optional datasets, i.e. a set of minimum (or optional) attributes for identifying a person in an eID. The document does not however address legal persons.

EBSI or ESSIF, in contrast, addresses both legal and natural persons by offering schemes for **Verifiable ID**s for both categories. When it comes to legal persons or entities, the majority of attributes can be easily mapped to the `registrationNumber` in the schema for Gaia-X Participant credentials depicted in Figure 7.8. Furthermore, both schemes are grounded in the eIDAS minimum dataset, ensuring GDPR compliance and a strong legal foundation. As such, the use of Verifiable IDs is not only legally robust but also seems to be the preferred choice in notable reports gathered during our literature review. For instance, the SSI eIDAS legal report [73] suggests using the Verifiable ID for natural persons defined

in ESSIF as eIDAS eID means, while the DSBA report follows the principle of utilizing what is already defined by EBSI for Verifiable IDs and more in general, including in their approach for identity binding. This indicates that EBSI Verifiable IDs are well-regarded credential schema in general.

We also came across the **ETSI EN 319 412-1 - V1.4.1** [166] specification, which defines both eIDAS natural and legal person semantic identifiers that map attributes from SAML attribute profiles defined by Recommendation ITU-T X.520 to eIDAS attributes. Based on this mapping, we conclude that the current employee credential scheme in the GX Credentials project is partially compliant with the standard, as an `organizationIdentifier` can be derived from the provided `vatID` in the employee credential's `gx:legalRegistrationNumber` claim.

After considering attribute sets and credential schemes from the aforementioned standards and specifications, we conclude that the attributes present in identity credentials should adhere to the minimum and optional datasets outlined in the eIDAS Regulation, as it is the basis of all relevant identity credential schemas such as EBSI's Verifiable ID and the PID attributes specified by the ARF. This aligns with Gaia-X's principles of being built upon existing schemas that have been standardized or widely adopted, as stated by the Architecture document [167].

On a more general note, **for identifying natural persons** in the form of VCs, we recommend SSI implementers to follow EBSI's Verifiable ID data model due to the fact that it is based on the eIDAS minimum data set for describing a natural person, which includes a "unique identifier constructed by the sending Member State in accordance with the technical specifications for the purposes of cross-border identification and which is as persistent as possible in time" [72]. Additionally, considering the sensitive nature of the attributes contained in such identity credentials, the credential should be encoded in JSON format, and SD-JWT is the recommended method for transporting this credential, within the context of W3C VCs. This recommendation aligns with the Type 1 configuration for issuing PID attestations as specified in the ARF. The ARF calls for both JSON and SD-JWT or, alternatively, CBOR and Mobile Security Object, as specified in ISO/IEC 18013-5:2021. As the latter falls outside the scope of this thesis, it won't be discussed in further detail. As a result of the chosen credential format (JSON + SD-JWT + VC Data Model v1.1), identity credentials or PID attestations should employ signatures and encryption formats detailed in JOSE and COSE RFCs. To enhance the level of identity binding, additional measures can be taken. For instance, when onboarding a new Participant into a Gaia-X data space, it's advisable that the Participant accepts a digital certificate or seal if issued by any European TSP [165]. A comprehensive list of these TSPs can be found in the List of Trusted Lists (LOTL), maintained by the EU Commission.

For properly **identifying legal persons with VCs**, similar requirements should be met, following the specifications of the eIDAS minimum data set. However, there are some additional considerations to keep in mind. Although the specific conditions under which companies can issue Qualified Electronic Attestations (QEAAs) are not explicitly defined yet, identity credentials solely attesting to a legal person's identity should also be

considered as QEAAs. This implies the use of the first configuration type as specified by the ARF. However, as legal entities can also issue Electronic Attestations (EAAs), the Type 2 configuration with JSON-LD and Linked Data Proofs (LD-Proofs) can be employed for such purposes. Still, it's not suitable for solely attesting to the identity of a legal person. Additional information that describes legal entities, such as Gaia-X service offerings, should be linked to the fundamental identity credential in the form of EAAs.

Shifting our focus back to the context of GX Credentials project and its employee identity credentials, we have to reiterate that considering employees as legal persons, which they usually are, is not that simple as employees fall are both natural and legal persons. While it may be feasible to link the employee credential to the identity credential designed for a natural person, the sole means of correlation are the name and VAT ID fields, which are both natural identifiers. We argue that there should be at least another alternative identifier apart from the DID corresponding to the employee. This would facilitate better communication flows between entities and potentially enable two-factor identification. Moreover, the inclusion of alternative identifiers offers employees the flexibility to discard an identifier, giving them more control over their information. We also believe that the existing employee credential, which encompasses only three out of seven essential claims, falls short. Therefore, adhering to the principle of data minimization and considering various aspects, we propose the following attribute set for employee credentials outlined in Table 7.2.

The proposed set of attributes is primarily a combination of eIDAS minimum data sets for both natural and legal persons. EBSI's Verifiable ID, which is also compliant with the minimum data set, also inspired the current claims set. We believe that this current set is descriptive enough while still minimizing disclosed data, especially with SD-JWT which further allows employees to selectively disclose attributes to be presented to the verifier. Furthermore, our schema and recommendations are backed by multiple technical reports from significant technical reports on the matter [73, 165].

Additional Credential Features

As previously outlined, employee credentials issued by companies should maintain a level of legal assurance akin to QEAAs. This entails utilizing VCs in JSON format with SD-JWT following the Type 1 configuration in the ARF. Even without considering this aspect, incorporating selective disclosure is advantageous within an SSI framework, aligning with principles like data minimization as endorsed by the W3C VC Data Model. However, due to the significance of the context enabled by the JSON-LD format in Gaia-X Data Spaces to effectively express entities and service offerings, we consequently adhere to the implementation recommendation by ETSI that relies on verifiers to transform the SD JWT VC into a JSON-LD VC as specified in the VCDM v.1.1. As a result of this, plain JSON data with key-value pairs of claims and their values can be employed as input to generate an SD JWT VC. Additionally, given the relatively young age and the rapid evolution of the SD JWT VC specifications, a user-friendly interface is crucial but often left as an afterthought. The current state of SD-JWTs is not particularly intuitive, as selectively disclosable claims

Property	Description	eIDAS Minimum Data Set Property and Description
`id`	Defines the DID of the employee	✗
`familyName`	Defines the current family name(s) of the employee	✓, *current family name*
`firstName`	Defines the current first name(s) of the employee	✓, *current first name*
`dateOfBirth`	Defines the date of birth of the employee	✓, *date of birth*
`nationalIdentifier`	Defines the unique national identifier of the employee	✓, *uniqueness identifier*
`emailAddress`	Defines the employee's email address according to the company's domain name used for correspondence	✗
`companyId`	Defines the ID (DID) of the issuing company	✗
`companyLegalName`	Defines the legal name of the company that issued the employee credential	✓, *current legal name*
`companyLegalIdentifier`	Legal identifier as proposed by [165], in our case derived from the employee's VAT identification number	✓, as the ID may be derived from VAT registration number, tax reference number, LEI, EORI, SIC (the identifier related to Article 3(1) of Directive 2009/101/EC of the European Parliament and of the Council), SEED (excise number provided in Article 2(12) of Council Regulation (EC) No 389/2012) [102].

Table 7.2.: Identity Attributes Included in the Proposed GX Employee Credential

issued by the issuer lack context, presenting only disclosure digests. These digests are essentially hashes, which might be perplexing for users, as the claim values are presented separately in the Disclosures. Therefore, an effective user interface plays a critical role in enhancing the user experience, aiding users and in our case employees in receiving and presenting credentials to relying parties.

Unfortunately after having analyzed the GX Credentials project, we came to the conclusion that in its current state, we could not sensibly integrate SD-JWT VC-related workflows into the project. This is mainly due to the fact that the Beacon protocol is used. Although the associated DIDKit library supports rudimentary functions such as signing JWT VCs using imported keys, it does not support the signing of JWT VCs using requests to the user's Altme Wallet. Moreover, at the time of writing this thesis, Altme also does not

support displaying SD-JWT VCs properly [168]. Talao, the company behind Altme, however, is actively working on implementing OID4VCI and VP workflows [169], which will then be used as the transport protocol for VC and VP-related exchanges.

Gaia-X aims to incorporate ARF-recommended technologies and protocols, including OID4VCI and VP, in the future. Naturally, these have not been implemented in the current state of the GX Credentials project. While we considered implementing our own components for these protocols, we realized that it would require extensive effort and may warrant a thesis of its own. The available libraries for these protocols are still in their early stages, and reference implementations are scarce, often making them unsuitable for production. To address this limitation, we propose an explanation of how the existing setup can be extended to support both the Beacon protocol and OID4VCI. The OIDC4VCI protocol is an extension of the widely used OIDC protocol, which differs from its predecessor by introducing Authorized and Pre-Authorized Code Flows among other features. Of particular interest is the pre-authorized flow, which assumes that users authenticate with the issuer using an out-of-band mechanism before any credential issuance operations take place. This flow leverages pre-existing authentication mechanisms, making it a valuable transition step for developers.

In Figure 7.9, we modeled the flow based on the assumption that the employee has already logged into the GX Credentials app with their wallet for the sake of simplicity, which we will explain step-by-step:

1. Once the employee is logged in, the employee applies for an employee credential by filling in an application form. In this step, the credential issuer effectively obtains consent from the employee and data that is required for the issuance of the requested employee credential.

2. The GX Credentials stores app this application in Firestore for the company to evaluate at a later point.

3. The company can either reject (3.a) or accept (3.b) the employee application.

4. Upon the company’s approval, a pre-authorization code is generated along with an OID4VCI Initiate Issuance Request that is communicated to the wallet via a QR code.

5. The employee interacts with the wallet and scans the QR code. The wallet then sends the OID4VCI Token Request to the issuer’s corresponding token endpoint, including the pre-authorized code from the previous step.

6. The app verifies the token request and checks whether the request was pre-authorized. Upon successful verification, an OID4VCI access token is returned.

7. The Wallet sends a Credential Request along with the access token and optionally the proof of possession of the public key to which the issued VC shall be bound. The app would then return either the requested employee credential right away or an acceptance token to specify a later time point from when the Wallet can start sending a Deferred Credential Request to obtain an issued credential [92].

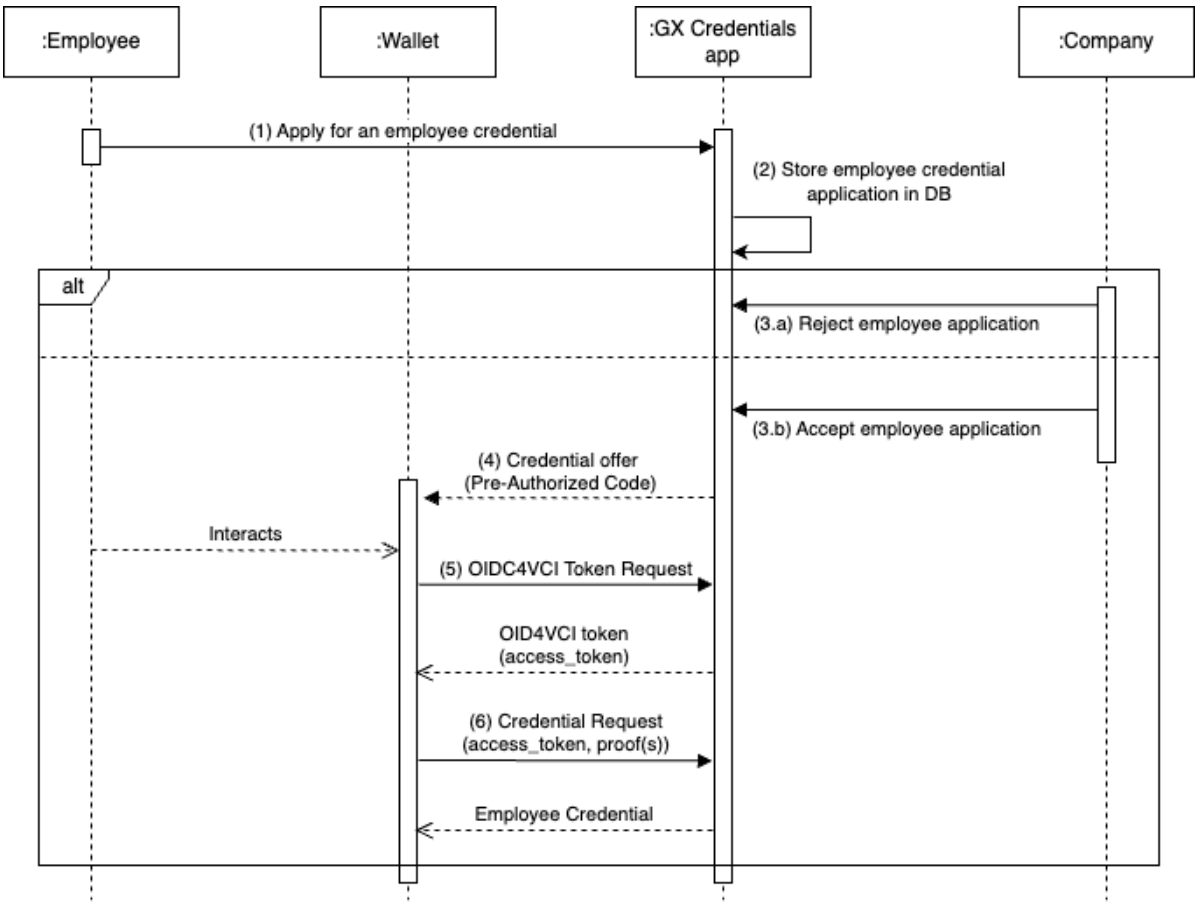


Figure 7.9.: Proposed Employee Credential Issuance using the OID4VCI Pre-Authorized Code Flow GX Credentials [92, 170].

Nonetheless, the Pre-Authorized Code Flow has certain limitations, as outlined by the OID4VCI specifications [92]. It is susceptible to replay attacks, as it's not bound to a specific device like the Authorization Code Flow. This vulnerability can, however, be mitigated using user PINs or by confirming the originating device when a token request is made. PIN code phishing concerns arise when attackers attempt to phish PIN codes sent by other services. To address this, the Wallet should interact with trusted Credential Issuers and avoid processing credential offers from untrusted issuer URLs. This should not be too much of an issue as companies themselves need to be certified by the registrar before being able to issue employee credentials.

Based on the discussions above, we conclude this section by stating that employee credentials in the GX Credentials project should be SD-JWT VCs encoded in JSON format and issued with the Pre-Authorized Code Flow to ensure interoperability with future wallet offerings from various SSI implementers, who will undoubtedly adopt the technologies and specifications outlined in the ARF that serves as the basis of EUDI Wallet Solutions. Given the status quo of the GX Credentials project, however, we are not able to implement this

without extensive effort and further analysis that may warrant a thesis in itself. Therefore, in order to educate and aid developers in comprehending the concepts and workflows related to the relatively new SD-JWT VC specifications, we opted to develop an interactive web application to demonstrate and visualize SD JWT VCs. We will provide a more detailed explanation of this implementation in the following section.

7.2.4. Development of an Interactive SD-JWT VC Demo Application

Core Functionality and Requirements

Before we started the development of the demo application, we analyzed the latest SD-JWT draft specification [147] to identify key SD-JWT features to be supported in the app. From there, we derived the following **functional requirements (FR)**:

- **FR-1: Creation of SD-JWT VCs Based on a JSON payload.** The app should enable users to create SD-JWT VCs from a JSON input. This implies the support for the creation of Disclosures for claim values in the form of an array or object properties.
- **FR-2: Presentation of SD-JWT VCs.** The app should enable users to present the created VCs to relying parties and selectively disclose which claims are presented.
- **FR-3: Holder Key Binding.** The app should support the optional holder key binding feature through the use of KB JWT.
- **FR-4: Verification of Issued or Presented Credentials.** The app should enable users to verify SD-JWT VCs that have been issued and presented.

In addition to the previously mentioned functional requirements, we establish the following two **non-functional requirements (NFR)**s. These NFRs align with the subcategories of *Usability* and *Security and Compliance*, respectively.

- **NFR-1: Intuitive User Interface.** The app should be intuitive enough to use even for users who have no prior knowledge or experience in handling SD-JWTs.
- **NFR-2: Use of Recommended Standards.** The app should facilitate the creation, presentation, and verification of SD-JWT VCs based on hash algorithms, specifications, and other general recommendations made by the latest draft of the IETF OAuth SD-JWT specification [147] to ensure compliance and address security considerations mentioned in the specifications.

Libraries

Despite SD-JWT specifications being relatively new, there are already an impressive number of libraries that support the issuance and presentation of SD-JWT VCs. Selecting the appropriate library for our demo web app is a crucial decision, as it plays a central role in

Library or Reference Implementation	SD Draft version	Ease of Use	Documentation	Test cases	Programming Language	GitHub Stars
christianpaquin/sd-jwt [171]	3	Medium	Good	✓	TypeScript	12
berendsliechtrecht/sd-jwt-ts [172]	5*	Medium	None	✓	TypeScript	6
chike0905/sd-jwt-ts [173]	2	Medium	Good	✓	TypeScript	5
Meeco/sd-jwt-vc [174]	5	Easy	Very Good	✓	TypeScript	0
transmute-industries/vc-jwt-sd [175]	5*	Medium	Good	✓	TypeScript	3
openwallet-foundation-labs/sd-jwt-python [176]	5	Medium	Very Good	✓	Python	4
authlete/sd-jwt [177]	4	Easy	Very Good	✓	Java	13
walt-id/waltid-sd-jwt [178]	4	Medium	Good	✓	Kotlin (Multiplatform)	7

Table 7.3.: Evaluation of SD-JWT Libraries

ensuring that we effectively meet our requirements. To this avail, we have compiled a list of libraries and reference implementations in Table 7.3.

The first critical decision point involves selecting the version of the draft specification upon which the implementation or library relies. Our preference is to align our app with the most recent version of the draft specification. This choice holds significance because certain concepts from earlier versions have become obsolete. For example, the *SD-JWT-Release* and *Salt/Value Container*, as defined in prior versions, have undergone renaming and workflow modifications in the proceeding versions. Notably, some libraries or implementations have adjusted their software to accommodate the new version (draft 5) during the period between our initial implementation and the writing of this section, as exemplified by [172, 175]. While this update could have potentially influenced our final decision, it occurred after our assessment was concluded, so our judgment is based on the version available at the time of our selection.

Ease of use is another important point. It is heavily correlated with the presence of documentation of the implementation. In this regard, libraries generally have an advantage as they are specifically designed for ease of integration and utilization.

Furthermore, it's evident that the majority of these implementations are developed using TypeScript and are well-documented, both of which significantly influence our library selection, especially the latter. Additionally, the presence of comprehensive test cases is crucial to ensure the reliability and functionality of the code. Considering that these specifications are relatively new, it's not surprising that many of these implementations have a low number of stars on their repositories. However, it's worth noting that the [174]

library has an extremely low star count, mainly because it was released just a few days before we began implementing our app. On the contrary, [176] has the most stars, likely because it's developed by the individuals behind the SD-JWT specifications and used to create examples in the documents. As such, the GitHub repository star count doesn't carry substantial weight in our decision-making process.

Among the libraries we considered, one library that particularly stood out is the one developed by *Meeco* [174]. Its simplicity and comprehensive documentation make it an attractive choice. Additionally, it's developed in TypeScript, which aligns well with web application development. While some other libraries, like the one by *walt.id*, are multi-platform and usable in this context, we found the developer experience and syntax to be suboptimal for our specific use case. Moreover, our prior familiarity and experience with TypeScript further bolstered our decision. Meeco's implementation is also based on the latest version of the SD-JWT draft specifications. It's worth noting that our decision might have been influenced if the recently updated implementations had been available earlier.

User Interface Implementation

To meet our web app's specific requirements, we opted for **Next.js**, a robust web framework, given that the SD-JWT VC business logic was developed in TypeScript for Node.js usage. Next.js offers developers the tools to create full-stack web applications by extending features present in React. The recent introduction of API routes in Next.js is particularly pertinent to our project as it allows us to create endpoints to handle operations related to encoding, presenting, and verifying SD-JWT VCs. Additionally, for the user interface design, we selected **Tailwind CSS**, a flexible and user-friendly CSS framework known for its customizability, which greatly facilitated the development of a clean and intuitive UI.

Screenshots of the implemented web application are attached in Section A.2. The user's journey begins on the *Create* tab, as illustrated in Figures A.2 and A.3, where they can complete a form featuring fields based on the attributes detailed in Section 7.2.3. Alternatively, users have the option to provide their custom payload. Once this step is completed, users must select a signing algorithm for creating the SD-JWT VC. Upon choosing an algorithm, the fields on this page will be populated automatically. This includes the Encoded SD-JWT VC, SD-JWT payload, the issuer key pair (which is generated automatically if not explicitly provided earlier), and, lastly, the SD Claims field.

Subsequently, users have two options: they can navigate to the *Verify* tab if they intend to verify the issued SD-JWT VC. Alternatively, by clicking on the *Create Verifiable Presentation* button, they can create an SD-JWT VP with holder key binding to present the VC to a verifier. When this button is clicked, a dialog is presented, as illustrated in Screenshot A.4, containing additional configuration settings for the presentation. In this dialog, users can specify which claims from the issued VC they want to exclude from the VP. Additionally, users can provide their key pair for the creation of the KB JWT. The public key is utilized as the value of the `cnf` claim in the VP, while the corresponding private key is employed to sign the KB JWT, which is attached to the end of the list of disclosures within the encoded VP. It's important to note that a 16-byte long nonce is randomly generated by default (ideally

by the verifier in the actual flow) to ensure the freshness of the signature.

Users can proceed to the *Verify* tab to assess the validity of the created VP, including the disclosures they have chosen to share with the verifier. It's worth noting that the titles of the cards will conditionally change depending on whether users choose to bind their keys to the VP or not. By default, the VP contains the issuer's public key, effectively binding it to the VP. For an improved understanding of the payload, users can click the icon next to the card title, such as the "Payload" card, to view a graphical representation of the JSON object, which, in this case, represents the payload.

7.2.5. Evaluation and Future Work

In this section, we conduct an evaluation of our SD-JWT demo web application implementation. We begin this evaluation by scrutinizing the implementation in the context of the established requirements. Subsequently, we discuss the existing limitations within the current implementation. Our assessment culminates with an exploration of potential directions for future work, encompassing aspects related to both the implementation and the overarching scope of the GX Credentials project.

Fulfilled Requirements

The choice of library in Section 7.2.4 plays a vital role in the fulfillment of the established requirements. Table 7.4 provides an overview of the established functional requirements, whereas the results of the non-functional requirements are depicted by Table 7.5.

Requirement	Status	Comment
FR-1	✓	Users are able to create SD-JWT VCs including for JSON inputs with claims in the form of arrays or object properties.
FR-2	✓	Users are able to create an SD-JWT VP and choose claims to selectively disclose to the verifier.
FR-3	✓	Users are able to optionally bind their public key in the VP.
FR-4	✓	Users are able to verify the created SD-JWT VC and its signature.

Table 7.4.: Adherance of SD-JWT VC Demo to Established Functional Requirements

Consequently, we can conclude that our implementation has successfully met all the established requirements, with the exception of **NFR-2** which is partially fulfilled. Nevertheless, there are some limitations that require discussion, and these will be addressed in the following section, as these might be considerations for future work.

Requirement	Status	Comment
NFR-1	✓	The app is adequately intuitive for users with no prior knowledge of SD-JWTs by e.g. providing tooltips when configuring the VP
NFR-2	✗	The app's functionality relies on a library [174] that aligns with the latest draft specification version. For our implementation, we have utilized the SHA256 algorithm, which according to [147] has raised concerns about the security suitability of SHA256 for certain applications. However, for our specific context and objectives, this choice should suffice.

Table 7.5.: Adherance of SD-JWT VC Demo to Established Non-Functional Requirements

Limitations and Future Work

As previously mentioned, we utilized Next.js API routes to create API endpoints for handling token-related operations. While we briefly explored the possibility of using web workers to enhance performance and efficiency, we encountered concerns regarding the compatibility of these workers with most of the libraries in our list, which are primarily meant to be used in server environments such as Node.js. Additionally, we realized that if our application were to be deployed, we would need to investigate alternative methods to minimize the numerous API calls required for JWT-related operations to ensure a smooth user experience and efficiency. However, given the local and experimental nature of our software, similar to the GX Credentials project, we decided to keep the functionality within Next.js' relatively new API routes as endpoints. This choice is listed as an area for potential future work, especially if deployment becomes a more significant consideration.

Moreover, a recommended feature outlined in the SD-JWT specifications that is not supported in the current implementation is the inclusion of decoy digests in the SD-JWT payload. These decoy digests are not associated with any claim in the payload and serve the purpose of making it more difficult for attackers to deduce the original number of claims contained in the SD-JWT. This feature is however already supported in the library upon which the library we used is based [179], which should help us in implementing it. As such, we have left this feature for future work.

In a broader context, the GX Credentials project should be extended in the future to accommodate the usage of SD-JWT VCs and VPs. As previously discussed in Section 7.2.3, however, it is currently not possible to extend the GX Credentials project to use SD-JWT VCs due to the usage of the Beacon protocol. To overcome this, we propose the integration of the Pre-Authorized Code Flow outlined in OID4VCI, which coincidentally is the chosen credential issuance protocol for EUDIW Solutions as specified in the ARF. To further align itself with technologies and specifications chosen for the implementation of EUDIW Solutions, the current employee credential schema should be revised to ensure legal compliance and relevance.

8. Conclusion

In this undergraduate thesis, we embarked on a comprehensive exploration of current approaches for identifying natural subjects in Verifiable Credentials. Our journey began with a systematic and rigorous literature survey, employing well-established methodologies in the fields of information systems and software engineering. We curated a final list of 58 sources, comprising both white and grey literature, to gain a holistic understanding of the SSI landscape.

From the gathered literary sources, we constructed a taxonomy of SSI approaches an initial pool of over 90 such approaches. The taxonomy spans nine dimensions and encompasses 27 distinct characteristics, which we used to distinguish a final total of 35 SSI approaches based on their incorporation of identifying information into VCs. The results of this taxonomy led to the observation that government-issued documents serve as the primary source for validating an individual's identity before the issuance of identity credentials. Furthermore, while standards significantly influence the composition of traditional paper-based credentials, only a minority exclusively employ standardized formats. This revelation underscores the fragmented nature of the field. We also observed that a substantial majority of these approaches continue to rely on VCs secured with LDPs, with an overwhelming majority supporting revocation mechanisms.

Our focus then shifted toward credential update methods, where we identified four primary approaches. Notably, atomic credentials gained attention, although their application is distinct and not primarily for updating information. However, the situation is similar to other methods; they are underdeveloped, with limited initiative in enhancing these mechanisms. In fact, the concept of "credential disputes" faces potential removal in the next version of the W3C's VCDM. Thus, considering that even credential revocation remains a significant topic in development, it is evident that managing credential updates is not a primary concern in most use cases.

Transitioning to the GX Credentials project, a proof of concept enabling VC issuance to companies and their employees within a Gaia-X ecosystem, our findings dictated that GX Employee credentials require revisions aligned with the eIDAS minimum data set. We assert that identity credentials, in general, should adhere to the recommendations outlined in the ARF, which also forms the foundation of future EUDIW Solutions. Moreover, due to the sensitive nature of PID within employee credentials and their pivotal role within the ecosystem, we advocate for endowing them with the same legal assurance as QEAAs, despite the unclear conditions for companies to issue such attestations. This necessitates the use of SD-JWT VCs as outlined in the ARF.

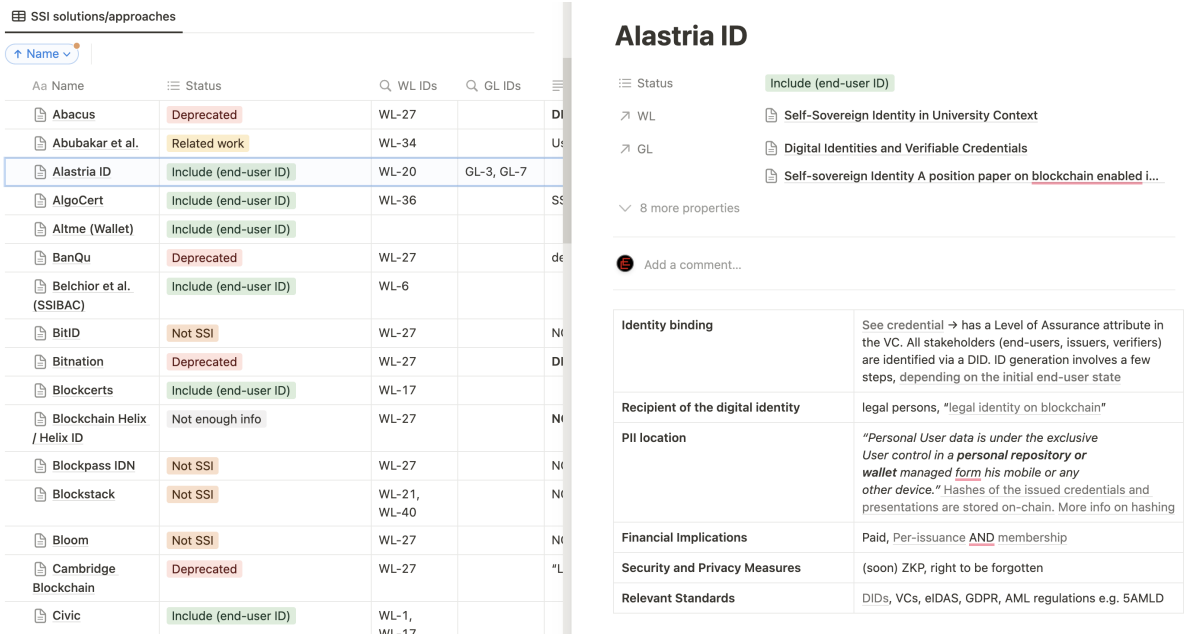
Regrettably, an in-depth analysis of the project's current state revealed that implementing SD-JWT VCs would require significant effort and warrant a dedicated thesis, exceeding

the scope of this work. To address this limitation, we explored the application of the Pre-Authorized Code Flow outlined in OID4VCI in conjunction with the Beacon protocol. In a concluding effort to educate future implementers and showcase the potential of SD-JWT VCs, we developed an interactive demo web application, which aims to serve as an intuitive demonstration of how SD-JWT VCs can be effectively managed in user interfaces.

As the GX Credentials project continues to evolve and technical specifications become more refined, we envision a decentralized future for identity management. This vision hinges on the broad adoption of Verifiable Credentials, clearly defined technical specifications, and the collaboration of private and public sectors. Notably, the EU's pursuit of empowering businesses and individuals through the EUDI Wallet adds a layer of excitement to the upcoming years. To that end, advocating for interoperability is paramount, ensuring that Verifiable Credentials can be utilized on a wider scale.

A. Addenda

A.1. Literature Review Tools



SSI solutions/approaches

↑ Name

Aa Name	Status	WL IDs	GL IDs	
Abacus	Deprecated	WL-27		DI
Abubakar et al.	Related work	WL-34		Us
Alastria ID	Include (end-user ID)	WL-20	GL-3, GL-7	
AlgoCert	Include (end-user ID)	WL-36		SS
Altme (Wallet)	Include (end-user ID)			
BanQu	Deprecated	WL-27		de
Belchior et al. (SSIBAC)	Include (end-user ID)	WL-6		
BitID	Not SSI	WL-27		NO
Bitnation	Deprecated	WL-27		DI
Blockcerts	Include (end-user ID)	WL-17		
Blockchain Helix / Helix ID	Not enough info	WL-27		NO
Blockpass IDN	Not SSI	WL-27		NO
Blockstack	Not SSI	WL-21, WL-40		NO
Bloom	Not SSI	WL-27		NO
Cambridge Blockchain	Deprecated	WL-27		"L
Civic	Include (end-user ID)	WL-1, WL-17		

Alastria ID

Status: Include (end-user ID)

WL: Self-Sovereign Identity in University Context

GL: Digital Identities and Verifiable Credentials

Self-sovereign Identity A position paper on blockchain enabled i...

8 more properties

Add a comment...

Identity binding	See credential → has a Level of Assurance attribute in the VC. All stakeholders (end-users, issuers, verifiers) are identified via a DID. ID generation involves a few steps, depending on the initial end-user state
Recipient of the digital identity	legal persons, "legal identity on blockchain"
PII location	*"Personal User data is under the exclusive User control in a* ***personal repository or wallet*** *managed form his mobile or any other device."* Hashes of the issued credentials and presentations are stored on-chain. More info on hashing
Financial Implications	Paid, Per-issuance AND membership
Security and Privacy Measures	(soon) ZKP, right to be forgotten
Relevant Standards	DIDs, VCs, eIDAS, GDPR, AML regulations e.g. 5AMLD

Figure A.1.: Notes in Notion Database

A.2. SD-JWT VC Demo Web Application for GX Credentials Screenshots

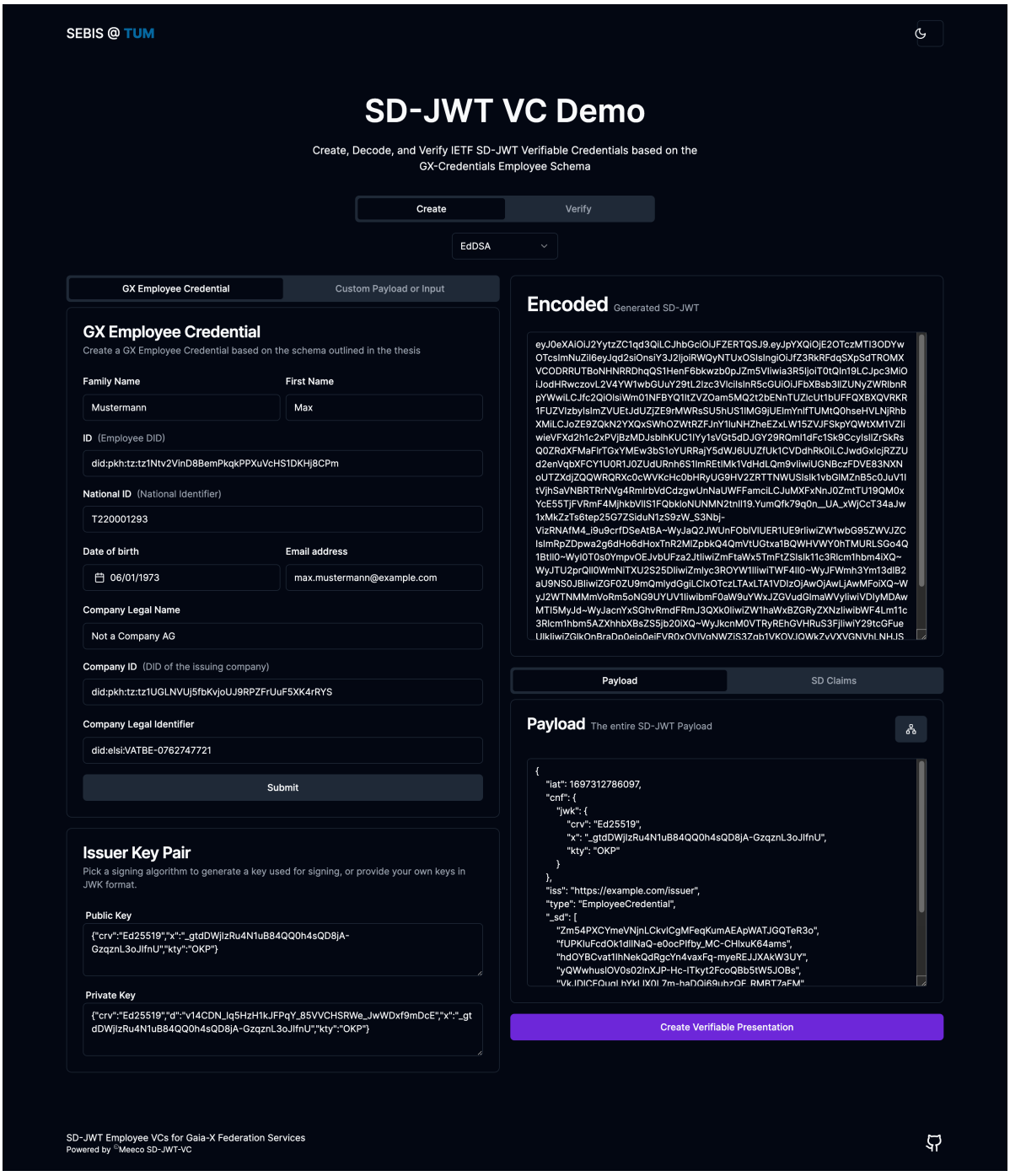


Figure A.2.: The SD-JWT VC Demo for GX Credentials home page with a form to create the payload used for the GX Employee Credential.

Figure A.3.: The SD-JWT VC Demo for GX Credentials home page with the custom payload or input tab. SD claims are also displayed.

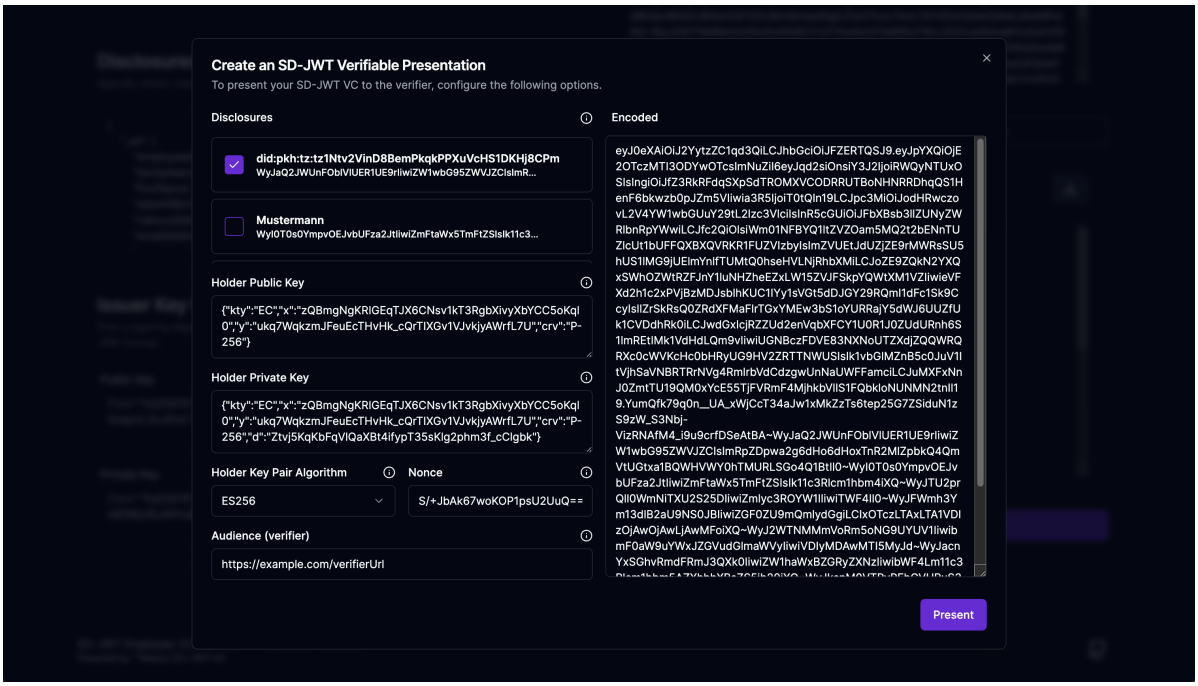


Figure A.4.: The SD-JWT VC Demo for GX Credentials VP dialog to configure the presentation and holder key binding.

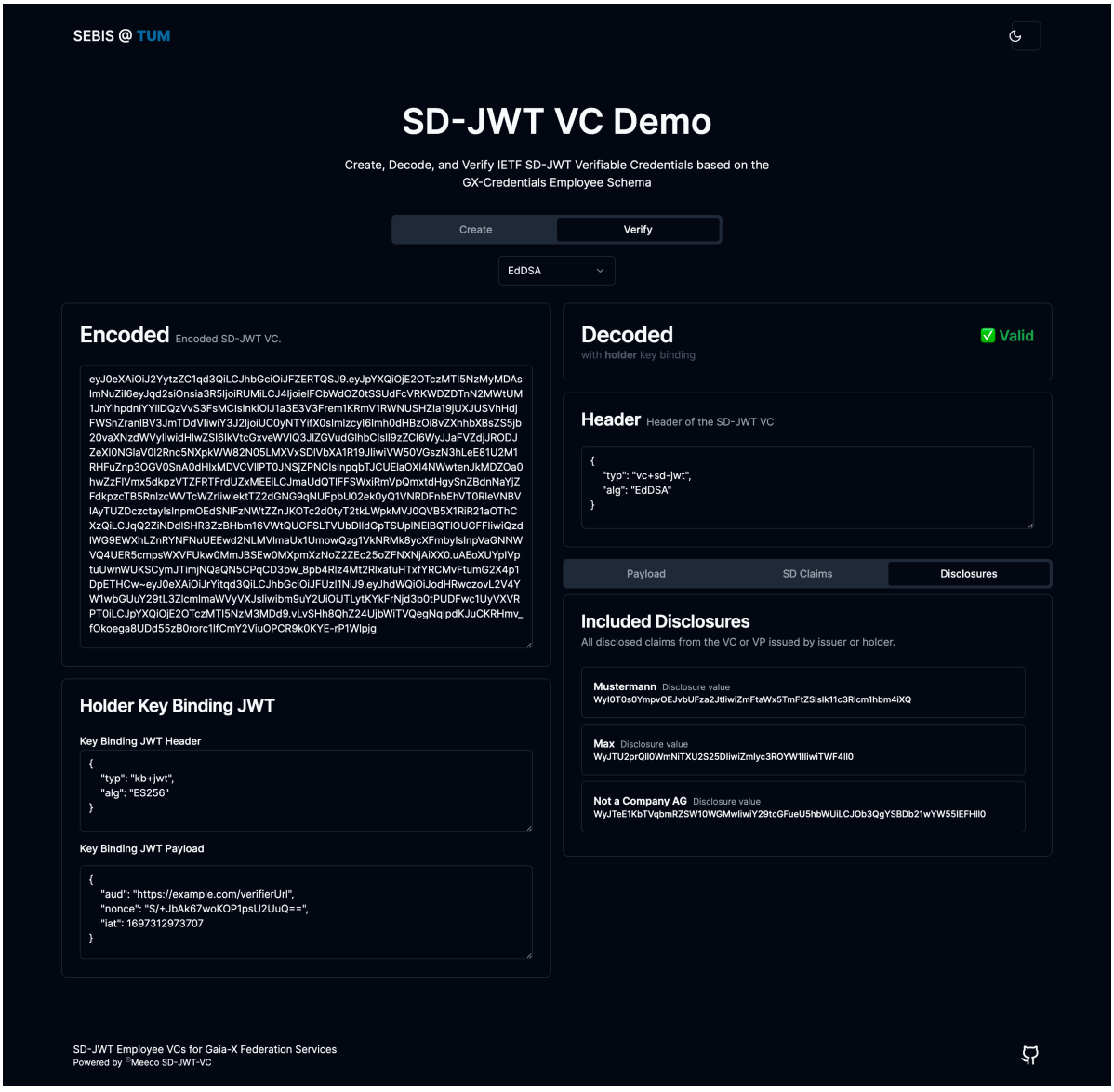


Figure A.5.: The SD-JWT VC Demo for GX Credentials verification page to validate issued or presented SD-JWT VCs.

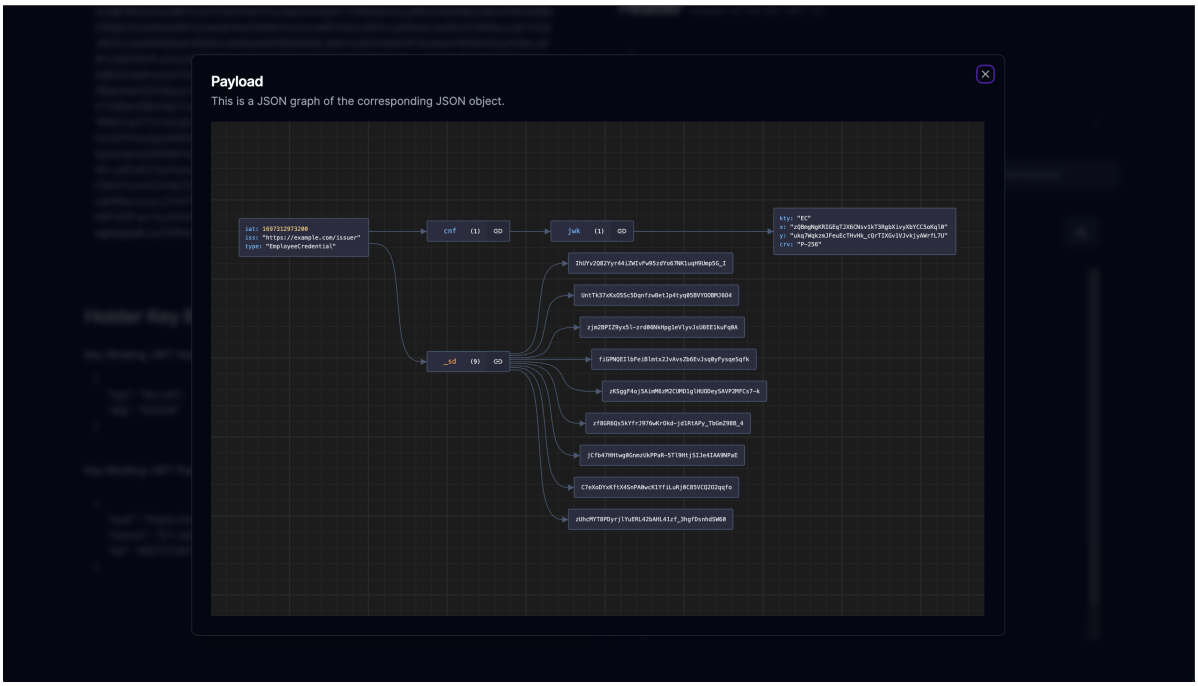


Figure A.6.: A dialog in the SD-JWT VC Demo for GX Credentials verification page that enables users to visualize JSON payloads and SD claims.

List of Figures

List of Tables

Acronyms

ARF
Architecture Reference Framework.

BBS
Boneh, Boyen, and Schacham.

BMWK
Bundesministerium für Wirtschaft und Klimaschutz.

CL
Camenisch-Lysyanskaya.

COSE
CBOR Object Signing and Encryption.

DID
Decentralized Identifier.

DLT
Distributed Ledger Technology.

DSBA
Data Spaces Business Alliance.

EAA
Electronic Attestation of Attributes.

EBSI
European Blockchain Services Infrastructure.

EIDAS
Electronic Identification, Authentication, and trust Services.

ENISA
European Union Agency for Cybersecurity.

EORI
Economic Operators Registration and Identification.

ESSIF
European Self-Sovereign Identity Framework.

ETSI
European Telecommunications Standards Institute.

EUDI
European Digital Identity.

EUDIW
European Digital Identity Wallet.

FR
Functional Requirements.

GDPR
General Data Protection Regulation.

GL
Grey Literature.

GXFS
Gaia-X Federation Services.

IdMs
Identity Management Systems.

IEC
International Electrotechnical Commission.

IETF
Internet Engineering Task Force.

ISO
International Organization for Standardization.

JOSE
JSON Object Signing and Encryption.

JSON
JavaScript Object Notation.

JSON-LD
JSON Linked Data.

JWT
JSON Web Token.

KB
Key Binding.

LDP
Linked Data Proofs.

LE
Legal Entity.

LEI
Legal Entity Identifier.

LoA
Level of Assurance.

LR
Literature Review.

mDL
Mobile Driving License.

MLR
Multivocal Literature Review.

MSO
Mobile Security Object.

NFR
Non Functional Requirements.

NIST
National Institute of Standards and Technology.

OID4VCI
OpenID for Verifiable Credential Issuance.

OID4VP
OpenID for Verifiable Presentation.

OIDC
OpenID Connect.

PID
Person Identification Data.

PII
Personal Identifiable Information.

PKI
Public Key Infrastructure.

QEAA
Qualified Electronic Attestation of Attributes.

RP
Relying Party.

SD
Selective Disclosure.

SIC
Standard Industrial Classification.

SIOP
Self-Issued OpenID Provider.

SSI
Self Sovereign Identity.

TAO
Trusted Accrediation Organization.

TI
Trusted Issuer.

TSP
Trust Services Provider.

VAT
Value-Added Tax.

VC
Verifiable Credentials.

VCDM
Verifiable Credentials Data Model.

VP
Verifiable Presentation.

W3C
World Wide Web Consortium.

WL
White Literature.

Bibliography

[1] Statista. *Digital Identity Solution Market Revenue Worldwide from 2020 to 2027*. Mar. 2023. URL: https://www.statista.com/statistics/1263580/worldwide-digital-identity-solution-market-revenue/.

[2] MarketsandMarkets. *Digital Identity Solutions Market Size, Share, Trends, Growth Drivers, Opportunities & Statistics*. urlhttps://www.marketsandmarkets.com/Market-Reports/digital-identity-solutions-market-247527694.html. Nov. 2022.

[3] O. Solon. *Credit firm Equifax says 143m Americans' social security numbers exposed in hack*. https://www.theguardian.com/us-news/2017/sep/07/equifax-credit-breach-hack-social-security. Sept. 2017.

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

[5] F. T. Commission. *Equifax to Pay $ 575 Million as Part of Settlement with FTC, CFPB, and States Related to 2017 Data Breach*. July 2019. URL: https://www.ftc.gov/news-events/news/press-releases/2019/07/equifax-pay-575-million-part-settlement-ftc-cfpb-states-related-2017-data-breach.

[6] S. McCallum. *Meta settles Cambridge Analytica scandal case for $725m*. Dec. 2022. URL: https://www.bbc.com/news/technology-64075067.

[7] Research and Markets. *Global Decentralized Identity Market Report 2022: An $8.9 Billion Market by 2028, Rising at a Staggering CAGR of 78.5%*. https://www.prnewswire.com/news-releases/global-decentralized-identity-market-report-2022-an-8-9-billion-market-by-2028--rising-at-a-staggering-cagr-of-78-5-301624096.html. Sept. 2022.

[8] A. Ibrahim and T. Dimitrakos. "Towards Collaborative Security Approaches Based on the European Digital Sovereignty Ecosystem". en. In: *Collaborative Approaches for Cyber Security in Cyber-Physical Systems*. Ed. by T. Dimitrakos, J. Lopez, and F. Martinelli. Advanced Sciences and Technologies for Security Applications. Cham: Springer International Publishing, 2023, pp. 123–144. ISBN: 978-3-031-16088-2. DOI: 10.1007/978-3-031-16088-2_6. URL: https://doi.org/10.1007/978-3-031-16088-2_6.

[9] *"Identity"*. https://dictionary.apa.org/identity. Retrieved 20.08.2023.

[10] J. E. Stets and P. J. Burke. "A Sociological Approach to Self and Identity". In: *Handbook of self and identity* 128152 (2003), pp. 23–50.

[11] K. Cameron. “The Laws of Identity”. In: *Microsoft Corp* 12 (2005), pp. 8–11.

[12] D. W. Chadwick. “Federated Identity Management”. In: *International School on Foundations of Security Analysis and Design*. Springer, 2007, pp. 96–120.

[13] C. Allen. *The Path to Self-Sovereign Identity*. 2016.

[14] M. S. Ferdous, F. Chowdhury, and M. O. Alassafi. “In Search of Self-Sovereign Identity Leveraging Blockchain Technology”. In: *IEEE Access* 7 (2019), pp. 103059–103079. DOI: 10.1109/ACCESS.2019.2931173.

[15] *How OpenID Connect Works - OpenID Foundation*. https://openid.net/developers/how-connect-works/. Retrieved 21.08.2023.

[16] F. Schardong and R. Custódio. “Self-Sovereign Identity: a Systematic Review, Mapping and Taxonomy”. In: *Sensors* 22.15 (2022), p. 5641.

[17] *FIDO Privacy White Paper*. https://media.fidoalliance.org/wp-content/uploads/FIDO__Privacy_White_Paper_Jan_2016.pdf. Retrieved 21.08.2023.

[18] *How FIDO Works - Standard Public Key Cryptography & User Privacy*. https://fidoalliance.org/how-fido-works/. Retrieved 21.08.2023.

[19] A. Mühle, A. Grüner, T. Gayvoronskaya, and C. Meinel. “A survey on essential components of a self-sovereign identity”. In: *Computer Science Review* 30 (2018), pp. 80–86.

[20] *GlobaliD 101: What is the trust triangle?* https://medium.com/global-id/globalid-101-what-is-the-trust-triangle-260e85e1c640. Retrieved 24.08.2023.

[21] *The right to erasure (Articles 17 & 19 of the GDPR) | Data Protection Commission*. https://www.dataprotection.ie/individuals/know-your-rights/right-erasure-articles-17-19-gdpr. Retrieved 23.08.2023.

[22] N. Naik and P. Jenkins. “Self-Sovereign Identity Specifications: Govern Your Identity Through Your Digital Wallet using Blockchain Technology”. In: *2020 8th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud)*. 2020, pp. 90–95. DOI: 10.1109/MobileCloud48802.2020.00021.

[23] Š. Čučko, Š. Bećirović, A. Kamišalić, S. Mrdović, and M. Turkanović. “Towards the Classification of Self-Sovereign Identity Properties”. In: *IEEE Access* 10 (2022), pp. 88306–88329. DOI: 10.1109/ACCESS.2022.3199414.

[24] *W3C*. https://www.w3.org/. Retrieved 25.08.2023.

[25] A. Preukschat and D. Reed. *Self-sovereign identity*. Manning Publications, 2021.

[26] C. Sehlke. *Transforming a Digital University Degree Issuance Process Towards Self-Sovereign Identity*. Sept. 2022.

[27] *Verifiable Credentials Data Model v2.0*. https://www.w3.org/TR/vc-data-model-2.0/#credential-subject/. Retrieved 25.08.2023.

[28] *VC Specifications Directory*. https://w3c.github.io/vc-specs-dir/#adding-a-specification-entry. Retrieved 25.08.2023.

[29] V. Papanchev. *An Interoperable Access Control System based on Self-Sovereign Identities*. Sept. 2022.

[30] *Decentralized Identifiers (DIDs) v1.0*. https://www.w3.org/TR/2022/REC-did-core-20220719/. Retrieved 26.08.2023.

[31] P. Y. Herrmann. *Verification of Digital Credentials Supporting Self-Sovereign Identity for Higher Education Institutions*. Oct. 2022.

[32] F. Hoops, A. Mühle, F. Matthes, and C. Meinel. "A Taxonomy of Decentralized Identifier Methods for Practitioners". In: *2023 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS)*. 2023, pp. 57–65. DOI: 10.1109/DAPPS57946.2023.00017.

[33] D. van Bokkem, R. Hageman, G. Koning, L. Nguyen, and N. Zarin. "Self-Sovereign Identity Solutions: The Necessity of Blockchain Technology". In: arXiv:1904.12816 (Apr. 2019). arXiv:1904.12816 [cs]. DOI: 10.48550/arXiv.1904.12816. URL: http://arxiv.org/abs/1904.12816.

[34] R. Nokhbeh Zaeem, K. C. Chang, T.-C. Huang, D. Liau, W. Song, A. Tyagi, M. Khalil, M. Lamison, S. Pandey, and K. S. Barber. "Blockchain-Based Self-Sovereign Identity: Survey, Requirements, Use-Cases, and Comparative Study". In: *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. WI-IAT '21. New York, NY, USA: Association for Computing Machinery, Apr. 2022, pp. 128–135. ISBN: 978-1-4503-9115-3. DOI: 10.1145/3486622.3493917. URL: https://dl.acm.org/doi/10.1145/3486622.3493917.

[35] Y. Bai, H. Lei, S. Li, H. Gao, J. Li, and L. Li. "Decentralized and Self-Sovereign Identity in the Era of Blockchain: A Survey". In: *2022 IEEE International Conference on Blockchain (Blockchain)*. 2022, pp. 500–507. DOI: 10.1109/Blockchain55522.2022.00077.

[36] J. Kaneriya and H. Patel. "A Comparative Survey on Blockchain Based Self Sovereign Identity System". In: *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*. 2020, pp. 1150–1155. DOI: 10.1109/ICISS49785.2020.9315899.

[37] A. Badirova, S. Dabbaghi, F. F. Moghaddam, P. Wieder, and R. Yahyapour. "A Survey on Identity and Access Management for Cross-Domain Dynamic Users: Issues, Solutions, and Challenges". In: *IEEE Access* 11 (2023), pp. 61660–61679. DOI: 10.1109/ACCESS.2023.3279492.

[38] M. Kuperberg. "Blockchain-Based Identity Management: A Survey From the Enterprise and Ecosystem Perspective". In: *IEEE Transactions on Engineering Management* 67.4 (2020), pp. 1008–1027. DOI: 10.1109/TEM.2019.2926471.

[39] M. R. Ahmed, A. K. M. M. Islam, S. Shatabda, and S. Islam. "Blockchain-Based Identity Management System and Self-Sovereign Identity Ecosystem: A Comprehensive Survey". In: *IEEE Access* 10 (2022), pp. 113436–113481. DOI: 10.1109/ACCESS.2022.3216643.

[40] Š. Čučko and M. Turkanović. “Decentralized and Self-Sovereign Identity: Systematic Mapping Study”. In: *IEEE Access* 9 (2021), pp. 139009–139027. DOI: 10.1109/ACCESS.2021.3117588.

[41] J. Sedlmeir, R. Smethurst, A. Rieger, and G. Fridgen. “Digital Identities and Verifiable Credentials”. en. In: *Business & Information Systems Engineering* 63.5 (Oct. 2021), pp. 603–613. ISSN: 1867-0202. DOI: 10.1007/s12599-021-00722-y.

[42] K. C. Toth and A. Anderson-Priddy. “Self-Sovereign Digital Identity: A Paradigm Shift for Identity”. In: *IEEE Security & Privacy* 17.3 (2019), pp. 17–27. DOI: 10.1109/MSEC.2018.2888782.

[43] *Self-Sovereign Identity for more Freedom and Privacy – SelfKey*. https://selfkey.org/. Retrieved 01.09.2023.

[44] *AI-Powered Proof of Individuality – SelfKey*. https://selfkey.org/ai-powered-proof-of-individuality/. Retrieved 01.09.2023.

[45] *A Guide to Using SelfKey iD for Digital Identity Verification*. https://selfkey.org/a-guide-to-using-selfkey-id-for-digital-identity-verification/. Retrieved 01.09.2023.

[46] *SelfKey DAO Whitepaper*. https://selfkey.org/selfkey-dao-whitepaper-en/. Retrieved 01.09.2023.

[47] M. Zichichi, C. Bomprezzi, G. Sorrentino, and M. Palmirani. “Protecting digital identity in the Metaverse: the case of access to a cinema in Decentraland”. In: (2023).

[48] U. Cali, M. S. Ferdous, E. Karaarslan, S. N. G. Gourisetti, and M. Mylrea. “SSI meets Metaverse for Industry 4.0 and Beyond”. In: *2022 IEEE 1st Global Emerging Technology Blockchain Forum: Blockchain & Beyond (iGETblockchain)*. 2022, pp. 1–6. DOI: 10.1109/iGETblockchain56591.2022.10087134.

[49] T. Tahlil, S. S. Gomasta, and A. B. M. S. Ali. “AlgoCert: Adopt Non-transferable NFT for the Issuance and Verification of Educational Certificates using Algorand Blockchain”. In: *2022 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*. 2022, pp. 1–8. DOI: 10.1109/CSDE56538.2022.10089274.

[50] M. Abubakar, P. McCarron, Z. Jaroucheh, A. Al Dubai, and B. Buchanan. “Blockchain-based Platform for Secure Sharing and Validation of Vaccination Certificates”. In: *2021 14th International Conference on Security of Information and Networks (SIN)*. Vol. 1. 2021, pp. 1–8. DOI: 10.1109/SIN54109.2021.9699221.

[51] *Worldcoin Whitepaper*. https://whitepaper.worldcoin.org/. Retrieved 01.09.2023.

[52] O. White. *Digital identification: A key to inclusive growth*. 2019.

[53] *Worldcoin Whitepaper*. https://whitepaper.worldcoin.org/proof-of-personhood. Retrieved 01.09.2023.

[54] *Understanding the Orb and why Worldcoin uses biometrics*. https://worldcoin.org/blog/worldcoin/understanding-orb-why-worldcoin-uses-biometrics. Retrieved 01.09.2023.

[55] *The wait is over: Worldcoin's World ID SDK now publicly available*. https://worldcoin.org/blog/product/wait-over-worldcoin-world-id-sdk-publicly-available. Retrieved 01.09.2023.

[56] K. Schmidt, A. Mühle, A. Grüner, and C. Meinel. "Clear the Fog: Towards a Taxonomy of Self-Sovereign Identity Ecosystem Members". In: *2021 18th International Conference on Privacy, Security and Trust (PST)*. 2021, pp. 1–7. DOI: 10.1109/PST52912.2021.9647797.

[57] R. Bochnia, D. Richter, and J. Anke. "Lifting the Veil of Credential Usage in Organizations: A Taxonomy". In: *Open Identity Summit 2023* (2023).

[58] D. Kundisch, J. Muntermann, A. M. Oberländer, D. Rau, M. Röglinger, T. Schoormann, and D. Szopinski. "An update for taxonomy designers: methodological guidance from information systems research". In: *Business & Information Systems Engineering* (2021), pp. 1–19.

[59] R. C. Nickerson, U. Varshney, and J. Muntermann. "A method for taxonomy development and its application in information systems". In: *European Journal of Information Systems* 22.3 (2013), pp. 336–359.

[60] T. Kölbel, M.-C. Härdtner, and C. Weinhardt. "Enterprise business models leveraging self-sovereign identity: Towards a user-empowering me2X economy". In: *Proceedings of the 56th Hawaii International Conference on System Sciences*. 2023.

[61] A. Kudra, T. Lodderstedt, P. Bastian, M. Mollik, M. van Leuken, and C. Roelofs. *A credential profile comparison matrix to facilitate technical and non-technical decision making*. https://github.com/WebOfTrustInfo/rwot11-the-hague/blob/master/final-documents/credential-profile-comparison.pdf. Apr. 2023.

[62] A. Kudra, T. Lodderstedt, P. Bastian, M. Mollik, M. van Leuken, and C. Roelofs. *Credential Comparison Matrix*. https://docs.google.com/spreadsheets/d/1Z4cYfjbbE-rABcfC-xab8miocKLomivYMUFibOh9BVo/edit?usp=embed_facebook. Retrieved 02.09.2023.

[63] *Organisation | TNO*. https://www.tno.nl/en/about-tno/organisation/. Retrieved 02.09.2023.

[64] M. van Leuken, P. Langenkamp, and T. Glastra. *An overview of SSI wallets and their characteristics*. https://github.com/tno-ssi-lab/wallet-overview. Retrieved 02.09.2023.

[65] M. van Leuken, P. Langenkamp, and T. Glastra. *An overview of SSI wallets and their characteristics*. https://github.com/tno-ssi-lab/wallet-overview. Retrieved 02.09.2023.

[66] *[DRAFT] SSI Standardisation Overview – GitHub*. https://github.com/tno-ssi-lab/standardisation-overview/tree/main. Retrieved 02.09.2023.

[67] *SSI Standards Overview*. https://tno-ssi-lab.github.io/standardisation-overview. Retrieved 02.09.2023.

[68] *SSI Standardisation Overview*. https://tno-ssi-lab.github.io/standardisation-overview/docs.html. Retrieved 02.09.2023.

[69] *The ToIP Model*. https://trustoverip.org/toip-model. Retrieved 02.09.2023.

[70] European Parliament and Council of the European Union. *Regulation (EU) 2016/679 of the European Parliament and of the Council*. of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). May 4, 2016. URL: https://data.europa.eu/eli/reg/2016/679/oj.

[71] A. Giannopoulou. "Data Protection Compliance Challenges for Self-Sovereign Identity". In: 3671523 (Feb. 2020). DOI: 10.2139/ssrn.3671523. URL: https://papers.ssrn.com/abstract=3671523.

[72] European Parliament and Council of the European Union. *Regulation (EU) No 910/2014 of the European Parliament and of the Council*. of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC. July 23, 2014. URL: http://data.europa.eu/eli/reg/2014/910/oj/eng.

[73] I. Domingo. "SSI eIDAS legal report; how eIDAS can legally support digital identity and trustworthy DLT-based transactions in the digital single market". In: *CEF Digital, Joinup* (Apr. 2020).

[74] E. Commission. *2030 Digital Compass: the European way for the Digital Decade*. COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS. https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52021DC0118. Retrieved 04.09.2023. Mar. 9, 2021.

[75] *Provisional political agreement EU Digital Identity Wallet*. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3556. Retrieved 04.09.2023.

[76] *EU Digital identity: 4 projects launched to test EUDI Wallet*. https://digital-strategy.ec.europa.eu/en/news/eu-digital-identity-4-projects-launched-test-eudi-wallet. Retrieved 04.09.2023. May 2023.

[77] *NIST SP 800-63 Digital Identity Guidelines*. https://pages.nist.gov/800-63-4/. Retrieved 03.09.2023.

[78] *NIST SP 800-63*. https://pages.nist.gov/800-63-4/sp800-63.html. Retrieved 03.09.2023.

[79] *NIST SP 800-63A*. https://pages.nist.gov/800-63-4/sp800-63a.html. Retrieved 03.09.2023.

[80] *NIST SP 800-63B*. https://pages.nist.gov/800-63-4/sp800-63b.html. Retrieved 03.09.2023.

[81] *NIST SP 800-63C*. https://pages.nist.gov/800-63-4/sp800-63c.html. Retrieved 03.09.2023.

[82] I. Alamillo, S. Mouille, A. Röck, N. Soumelidis, M. Tabor, and S. Gorniak. *Digital Identity Standards: Analysis of standardisation requirements in support of cybersecurity policy*. https://www.enisa.europa.eu/publications/digital-identity-standards. Report/Study. July 2023. DOI: 10.2824/28598.

[83] I. O. for Standardization. *Personal identification — ISO-compliant driving licence — Part 5: Mobile driving licence (mDL) application*. ISO/IEC 18013-5:2021. Vernier, Geneva, Switzerland: International Organization for Standardization, Sept. 2021. URL: https://www.iso.org/standard/69084.html.

[84] H. Birkholz, C. Vigano, and C. Bormann. *Concise Data Definition Language (CDDL): A Notational Convention to Express Concise Binary Object Representation (CBOR) and JSON Data Structures*. RFC 8610. June 2019. DOI: 10.17487/RFC8610. URL: https://www.rfc-editor.org/info/rfc8610.

[85] C. Bormann and P. E. Hoffman. *Concise Binary Object Representation (CBOR)*. RFC 7049. Oct. 2013. DOI: 10.17487/RFC7049. URL: https://www.rfc-editor.org/info/rfc7049.

[86] T. Bray. *The JavaScript Object Notation (JSON) Data Interchange Format*. RFC 8259. Dec. 2017. DOI: 10.17487/RFC8259. URL: https://www.rfc-editor.org/info/rfc8259.

[87] I. O. for Standardization. *Cards and security devices for personal identification — Building blocks for identity management via mobile devices — Part 1: Generic system architectures of mobile eID systems*. ISO/IEC 23220-1:2023. Vernier, Geneva, Switzerland: International Organization for Standardization, Feb. 2023. URL: https://www.iso.org/standard/74910.html.

[88] N. Sakimura, J. Bradley, M. Jones, de Medeiros. B, and C. Mortimore. *OpenID Connect Core 1.0 incorporating errata set 1*. https://openid.net/specs/openid-connect-core-1_0.html. Retrieved 04.09.2023. Nov. 2014.

[89] K. Yasuda, M. Jones, and T. Lodderstedt. *Self-Issued OpenID Provider v2*. https://openid.net/specs/openid-connect-self-issued-v2-1_0.html. Retrieved 04.09.2023. Jan. 2023.

[90] K. Young. *Where to begin with OIDC and SIOP*. https://medium.com/decentralized-identity/where-to-begin-with-oidc-and-siop-7dd186c89796. Retrieved 04.09.2023. July 2020.

[91] *OpenID for Verifiable Credentials – Specifications*. https://openid.net/sg/openid4vc/specifications. Retrieved 03.09.2023.

[92] T. Lodderstedt, K. Yasuda, and T. Looker. *OpenID for Verifiable Credential Issuance*. https://openid.net/specs/openid-4-verifiable-credential-issuance-1_0.html. Retrieved 04.09.2023. Feb. 2023.

[93] T. O, T. Lodderstedt, K. Yasuda, A. Lemmon, and T. Looker. *OpenID Connect for Verifiable Presentations*. https://openid.net/specs/openid-connect-4-verifiable-presentations-1_0-10.html. Retrieved 04.09.2023. Apr. 2022.

[94] K. Yasuda, T. Lodderstedt, K. Nakamura, S. Ganesan, and R. Narayanan. *OpenID for Verifiable Presentations over BLE - draft 00*. https://openid.net/specs/openid-4-verifiable-presentations-over-ble-1_0.html. Retrieved 04.09.2023.

[95] M. Ansari, R. Barnes, P. Kasselman, and K. Yasuda. *OpenID Connect UserInfo Verifiable Credentials - Draft 00*. https://openid.net/specs/openid-connect-userinfo-vc-1_0-00.html. Retrieved 04.09.2023. May 2023.

[96] *Digital Credentials Protocols (DCP) Working Group - Overview*. https://openid.net/wg/digital-credentials-protocols. Retrieved 04.09.2023.

[97] K. N. Chadwick and J. Vercammen. *OpenID for Verifiable Credentials*. https://openid.net/wordpress-content/uploads/2022/05/OIDF-Whitepaper_OpenID-for-Verifiable-Credentials_FINAL_2022-05-12.pdf. Retrieved 04.09.2023.

[98] European Commission. *EBSI – Home*. https://ec.europa.eu/digital-building-blocks/wikis/display/EBSI/Home. Retrieved 05.09.2023.

[99] European Commission. *EBSI – About us*. https://ec.europa.eu/digital-building-blocks/wikis/display/EBSI/About+us. Retrieved 05.09.2023.

[100] *EBSI Specifications – Issuers trust model - Accreditation of Issuers*. https://ec.europa.eu/digital-building-blocks/wikis/display/EBSIDOC/Issuers+trust+model+-+Accreditation+of+Issuers. Retrieved 05.09.2023.

[101] *EBSI Specifications–EBSI DID Method*. https://ec.europa.eu/digital-building-blocks/wikis/display/EBSIDOC/EBSI+DID+Method. Retrieved 05.09.2023.

[102] *EBSI Specifications – Verifiable ID - Legal Entity*. https://ec.europa.eu/digital-building-blocks/wikis/display/EBSIDOC/Verifiable+ID+-+Legal+Entity. Retrieved 05.09.2023.

[103] *EBSI Specifications – Verifiable ID - Natural Person*. https://ec.europa.eu/digital-building-blocks/wikis/display/EBSIDOC/Verifiable+ID+-+Natural+Person. Retrieved 05.09.2023.

[104] CORDIS. *European Self Sovereign Identity Framework Laboratory*. https://cordis.europa.eu/project/id/871932/results. Retrieved 05.09.2023. DOI: 10.3030/871932.

[105] NGI eSSIF-Lab. *eSSIF-LAB – Home*. https://essif-lab.eu/. Retrieved 05.09.2023.

[106] Validated ID. *SEB_project_summary – GitLab*. https://gitlab.grnet.gr/essif-lab/infrastructure/validated-id/seb_project_summary. Retrieved 05.09.2023.

[107] *SSI eIDAS Bridge Interoperability – GitLab*. https://gitlab.grnet.gr/essif-lab/interoperability/ssi-eidas-bridge. Retrieved 05.09.2023.

[108] *Gaia-X – Home*. https://gaia-x.eu. Retrieved 05.09.2023.

[109] *Gaia-X secure and trustworthy ecosystems with Self Sovereign Identity*. https://gaia-x.eu/wp-content/uploads/2022/06/SSI_White_Paper_Design_Final_EN.pdf. Retrieved 05.09.2023. May 2022.

[110] *Gaia-X Framework*. https://gaia-x.eu/gaia-x-framework. Retrieved 05.09.2023.

[111] V. Garousi, M. Felderer, and M. V. Mäntylä. "Guidelines for including grey literature and conducting multivocal literature reviews in software engineering". In: *Information and Software Technology* 106 (2019), pp. 101–121. ISSN: 0950-5849. DOI: https://doi.org/10.1016/j.infsof.2018.09.006.

[112] R. Soltani, U. Trang Nguyen, and A. An. "A New Approach to Client Onboarding Using Self-Sovereign Identity and Distributed Ledger". In: *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*. 2018, pp. 1129–1136. DOI: 10.1109/Cybermatics_2018.2018.00205.

[113] H. Saidi, N. Labraoui, A. A. A. Ari, L. A. Maglaras, and J. H. M. Emati. "DSMAC: Privacy-Aware Decentralized Self-Management of Data Access Control Based on Blockchain for Health Data". In: *IEEE Access* 10 (2022), pp. 101011–101028. DOI: 10.1109/ACCESS.2022.3207803.

[114] P. Herbke and H. Yildiz. "ELMO2EDS: Transforming Educational Credentials into Self-Sovereign Identity Paradigm". In: *2022 20th International Conference on Information Technology Based Higher Education and Training (ITHET)*. 2022, pp. 1–7. DOI: 10.1109/ITHET56107.2022.10031276.

[115] T. Hamer, K. Taylor, K. S. Ng, and A. Tiu. "Private Digital Identity on Blockchain." In: *BlockSW/CKG@ ISWC*. 2019.

[116] L. Stockburger, G. Kokosioulis, A. Mukkamala, R. R. Mukkamala, and M. Avital. "Blockchain-enabled decentralized identity management: The case of self-sovereign identity in public transportation". In: *Blockchain: Research and Applications* 2.2 (2021), p. 100014.

[117] D. Wang, X. Chen, L. Zhang, Y. Fang, and C. Huang. "A Blockchain-Based Human-to-Infrastructure Contact Tracing Approach for COVID-19". In: *IEEE Internet of Things Journal* 9.14 (2022), pp. 12836–12847. DOI: 10.1109/JIOT.2021.3138971.

[118] J. Xu, K. Xue, H. Tian, J. Hong, D. S. L. Wei, and P. Hong. “An Identity Management and Authentication Scheme Based on Redactable Blockchain for Mobile Networks”. In: *IEEE Transactions on Vehicular Technology* 69.6 (2020), pp. 6688–6698. DOI: 10.1109/TVT.2020.2986041.

[119] M. Morosi. “Study of authentication models and implementation of a prototype by using eID and Distributed Ledger Technologies.” PhD thesis. Politecnico di Torino, 2022.

[120] R. Belchior, B. Putz, G. Pernul, M. Correia, A. Vasconcelos, and S. Guerreiro. “SSI-BAC: Self-Sovereign Identity Based Access Control”. In: *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. 2020, pp. 1935–1943. DOI: 10.1109/TrustCom50675.2020.00264.

[121] M. Rahman, M. M. Azam, and F. S. Chowdhury. “An Anonymity and Interaction Supported Complaint Platform based on Blockchain Technology for National and Social Welfare”. In: *2021 International Conference on Electronics, Communications and Information Technology (ICECIT)*. 2021, pp. 1–8. DOI: 10.1109/ICECIT54077.2021.9641269.

[122] A. Satybaldy, A. Subedi, and M. Nowostawski. “A Framework for Online Document Verification Using Self-Sovereign Identity Technology”. In: *Sensors* 22.21 (2022), p. 8408.

[123] N. C. S. R. Center. *Pseudonymous Identifier - Glossary*. https://csrc.nist.gov/glossary/term/pseudonymous_identifier. Retrieved 11.09.2023.

[124] *Verifiable Credentials Implementation Guidelines 1.0*. https://www.w3.org/TR/vc-imp-guide/. Retrieved 11.09.2023.

[125] *Javascript Object Signing and Encryption (JOSE) — jose 0.1 documentation*. https://jose.readthedocs.io/en/latest/. Retrieved 11.09.2023.

[126] *The ELMO XML format – GitHub*. https://github.com/emrex-eu/elmo-schemas. Retrieved 11.09.2023.

[127] J. Camenisch and A. Lysyanskaya. “Signature schemes and anonymous credentials from bilinear maps”. In: *Annual international cryptology conference*. Springer. 2004, pp. 56–72.

[128] *A European approach to micro-credentials – European Education Area*. https://education.ec.europa.eu/education-levels/higher-education/micro-credentials. Retrieved 14.09.2023.

[129] ETSI Technical Committee Electronic Signatures and Infrastructures (ESI). *Electronic Signatures and Infrastructures (ESI); Analysis of selective disclosure and zero-knowledge proofs applied to Electronic Attestation of Attributes*. Tech. rep. ETSI TR 119 476 V1.1.1. ETSI, Aug. 2023.

[130] B. Zundal, C. Allen, M. Sabadello, E. Chachkarova, and F. Tagliaferro. *Selective Correlation*. https://github.com/WebOfTrustInfo/rwot11-the-hague/blob/master/draft-documents/selective-correlation.md. Retrieved 14.09.2023. Sept. 2022.

[131] *Holder Bundling of Separate Claims – issue#75 – GitHub*. https://github.com/w3c/vc-use-cases/issues/75. Retrieved 14.09.2023.

[132] *Verifiable Credentials Lifecycle 1.0*. https://w3c-ccg.github.io/vc-lifecycle. Retrieved 16.09.2023.

[133] *Selective Content Generation*. https://research.csiro.au/blockchainpatterns/general-patterns/self-sovereign-identity-patterns/selective-content-generation/. Retrieved 14.09.2023.

[134] *credentialSchema and Selective Disclosure – issue#890 – GitHub*. https://github.com/w3c/vc-data-model/issues/890. Retrieved 14.09.2023.

[135] A. Flamini, S. Ranise, G. Sciarretta, M. Scuro, A. Sharif, and A. Tomasi. *A First Appraisal of Cryptographic Mechanisms for the Selective Disclosure of Verifiable Credentials*. May 2023. DOI: 10.13140/RG.2.2.16534.11843.

[136] S. Elfors. *FIDO Alliance White Paper: Using FIDO for the EUDI Wallet*. https://media.fidoalliance.org/wp-content/uploads/2023/04/FIDO-EUDI-Wallet-White-Paper-FINAL.pdf. Retrieved 14.09.2023.

[137] *Verifiable Credentials Data Model v1.1*. https://www.w3.org/TR/vc-data-model. Retrieved 16.09.2023.

[138] *Verifiable Credential Refresh 2021*. https://w3c-ccg.github.io/vc-refresh-2021. Retrieved 16.09.2023.

[139] A. De Salve, A. Lisi, P. Mori, and L. Ricci. "Selective Disclosure in Self-Sovereign Identity based on Hashed Values". In: *2022 IEEE Symposium on Computers and Communications (ISCC)*. 2022, pp. 1–8. DOI: 10.1109/ISCC55528.2022.9913052.

[140] D. Boneh, X. Boyen, and H. Shacham. "Short group signatures". In: *Annual international cryptology conference*. Springer. 2004, pp. 41–55.

[141] *What are zk-SNARKs?* https://z.cash/learn/what-are-zk-snarks. Retrieved 19.09.2023.

[142] N. Kongsuwan and R. Tosirisuk. *Anonymous Credential Part 3: BBS+ Signature*. https://medium.com/finema/anonymous-credential-part-3-bbs-signature-26797721ca74. Retrieved 19.09.2023. Oct. 2020.

[143] G. Bernstein. *BBS for Verifiable Credentials - Basics*. https://grotto-networking.com/Presentations/BBSforVCs/BBSforVCsBasics.html. Retrieved 20.09.2023. May 2023.

[144] N. Helmy. *A solution for privacy-preserving Verifiable Credentials*. https://medium.com/mattr-global/a-solution-for-privacy-preserving-verifiable-credentials-f1650aa16093. Retrieved 19.09.2023. May 2020.

[145] T. Looker, V. Kalos, A. Whitehead, and M. Lodder. *The BBS Signature Scheme*. https://identity.foundation/bbs-signature/draft-irtf-cfrg-bbs-signatures.html. Retrieved 20.09.2023. July 2023.

[146] *BBS Signature Demo*. https://www.grotto-networking.com/BBSDemo. Retrieved 20.09.2023.

[147] D. Fett, K. Yasuda, and B. Campbell. *Selective Disclosure for JWTs (SD-JWT)*. draft-ietf-oauth-selective-disclosure-jwt-05. June 2023.

[148] *The Common Union Toolbox for a Coordinated Approach Towards a European Digital Identity Framework*. https://digital-strategy.ec.europa.eu/en/library/european-digital-identity-wallet-architecture-and-reference-framework. Jan. 2023.

[149] O. Terbu and D. Fett. *SD-JWT-based Verifiable Credentials (SD-JWT VC)*. draft-ietf-oauth-sd-jwt-vc-00. Aug. 2023. URL: https://datatracker.ietf.org/doc/draft-ietf-oauth-sd-jwt-vc.

[150] Gaia-X. *What is Gaia-X ENG*. https://youtu.be/DrG-EBBFniw?si=ME0wLn_F-qk3dN5q. Retrieved 29.09.2023. 2022.

[151] *Gaia-X Federation Services*. https://www.gxfs.eu/. Retrieved 29.09.2023.

[152] *Gaia-X 4 Future Mobility*. https://www.gaia-x4futuremobility.de/en/home. Retrieved 29.09.2023.

[153] *Gaia-X 4 Future Mobility | PLC-AAD*. https://www.gaia-x4futuremobility.de/en/projects/plc-aad. Retrieved 29.09.2023.

[154] *GAIA-X 4 PLC-AAD | GitHub*. https://github.com/GAIA-X4PLC-AAD. Retrieved 29.09.2023.

[155] *GX Credentials | GitHub*. https://github.com/GAIA-X4PLC-AAD/gx-credentials. Retrieved 29.09.2023.

[156] *Gaia-X Trust Framework | 22.10 Release*. https://docs.gaia-x.eu/policy-rules-committee/trust-framework/22.10/. Retrieved 12.10.2023.

[157] *GXFS Connect 2023*. https://www.gxfs.eu/gxfs-connect-2023/. Retrieved 12.10.2023.

[158] *Gaia-X Registry | GitLab*. https://gitlab.com/gaia-x/lab/compliance/gx-registry. Retrieved 12.10.2023.

[159] DIACC. *BC Government's Verifiable Credential Issuer Kit Proof of Concept Report*. https://diacc.ca/2021/10/20/bc-governments-verifiable-credential-issuer-kit-proof-of-concept-report/. Retrieved 12.10.2023.

[160] S. Schwalm and A. Kudra. *ARF and EUDIW – a community commentary | LinkedIn*. https://www.linkedin.com/pulse/arf-eudiw-community-commentary-andre-kudra/. Retrieved 13.10.2023.

[161] Bundesamt für Sicherheit in der Informationstechnik. *BSI TR-03159 Mobile Identities*. de. https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/Technische-Richtlinien/TR-nach-Thema-sortiert/tr03159/tr-03159.html?nn=460676.

[162] H. Taylor. *Outline of the Players in the EUDI Wallet Ecosystem*. https://www.cryptomathic.com/news-events/blog/outline-of-the-players-in-the-eudi-wallet-ecosystem. Retrieved 13.10.2023.

[163] D. McNeal. *Which Trust Service Providers Support Remote QES Services?* https://www.cryptomathic.com/news-events/blog/qualified-remote-electronic-signatures-provided-by-trust-service-providers-part-1. Retrieved 13.10.2023.

[164] European Commission Directorate-General Informatics. *eIDAS Bridge WP2 - Use cases and technical specifications*. https://joinup.ec.europa.eu/sites/default/files/document/2020-04/SSI%20eIDAS%20Bridge%20-%20Use%20cases%20and%20Technical%20Specifications%20v1.pdf. Retrieved 13.10.2023.

[165] European Commission Directorate-General Informatics. *Data Spaces Business Alliance Technical Convergence - Discussion Document*. https://data-spaces-business-alliance.eu/wp-content/uploads/dlm_uploads/Data-Spaces-Business-Alliance-Technical-Convergence-V2.pdf. Retrieved 13.10.2023. Apr. 2023.

[166] European Commission Directorate-General Informatics. *ETSI EN 319 412-1 V1.4.1. Electronic Signatures and Infrastructures (ESI); Certificate Profiles; Part 1: Overview and common data structures*. https://www.etsi.org/deliver/etsi_en/319400_319499/31941201/01.04.01_60/en_31941201v010401p.pdf. June 2020.

[167] *Gaia-X Architecture Document | 22.10 Release*. https://docs.gaia-x.eu/technical-committee/architecture-document/22.10/. Retrieved 12.10.2023.

[168] *Support for SD-JWT format*. https://github.com/TalaoDAO/AltMe/issues/1935. Retrieved 12.10.2023.

[169] *TalaoDAO/oidc4vc-server*. https://github.com/TalaoDAO/oidc4vc-server. Retrieved 12.10.2023.

[170] *OpenID for Verifiable Credential Issuance - Client and Issuer | GitHub*. https://github.com/Sphereon-Opensource/OID4VCI. Retrieved 14.10.2023.

[171] C. Paquin. *Simple selective disclosure for JSON Web Tokens | GitHub*. https://github.com/christianpaquin/sd-jwt. Retrieved 14.10.2023.

[172] B. Sliedrecht. *Selective Disclosure JWT Draft 05 implementation | GitHub*. https://github.com/berendsliedrecht/sd-jwt-ts. Retrieved 14.10.2023.

[173] Ryosuke. *sd-jwt-ts | GitHub*. https://github.com/chike0905/sd-jwt-ts. Retrieved 14.10.2023.

[174] Meeco. *SD-JWT-VC | GitHub*. https://github.com/Meeco/sd-jwt-vc. Retrieved 14.10.2023.

[175] Transmute Industries. *@transmute/vc-jwt-sd_2023 | GitHub*. https://github.com/transmute-industries/vc-jwt-sd. Retrieved 14.10.2023.

[176] O. F. Labs. *SD-JWT Reference Implementation | GitHub*. https://github.com/openwallet-foundation-labs/sd-jwt-python. Retrieved 14.10.2023.

[177] Authlete, Inc. *Java Library for SD-JWT | GitHub*. https://github.com/authlete/sd-jwt. Retrieved 14.10.2023.

[178] walt.id. *Kotlin Multiplatform SD-JWT library | GitHub*. https://github.com/walt-id/waltid-sd-jwt. Retrieved 14.10.2023.

[179] Meeco. *Meeco/sd-jwt: Selective disclosure for a signed JWT | GitHub*. https://github.com/Meeco/sd-jwt. Retrieved 14.10.2023.