Digital identity is evolving from centralized systems to a decentralized approach known as Self-Sovereign Identity (SSI). SSI empowers individuals to control their digital identities, eliminating reliance on third-party data custodians and reducing the risk of data breaches. However, the concept of trust in SSI remains complex and fragmented. This paper systematically analyzes trust in SSI in light of its components and threats posed by various actors in the system. As a result, we derive three distinct trust models that capture the threats and mitigations identified across SSI literature and implementations. Our work provides a foundational framework for future SSI research and development, including a comprehensive catalogue of SSI components and design requirements for trust, shortcomings in existing SSI systems and areas for further exploration.

KEYWORDS
digital identity, self-sovereign identity, identity data privacy

1 INTRODUCTION

Digital identity encompasses not only usernames and passwords but also our online presentation, interactions with others, and how the world perceives us [5]. In 2023, Louisiana passed a law requiring all websites containing adult content to verify the user’s age. Louisiana’s law mandated users present their state-issued digital identities, such as driver’s licenses, to access these websites [39]. This is not a singular event; digital identities are becoming ubiquitous tools in many applications, including age verification, know-your-customer compliance [121, 127], and academic credential verification [146]. Digital identity has expanded beyond mere name and email addresses in an online account.

In the current ecosystem, digital identity is managed centrally through siloed (one identity per service) or federated (single sign-on) identity systems [129]. These centralized identity management systems typically rely on a single entity, such as a government agency or a private entity, to issue and manage identities. This concentration of power has created a single point of failure, making these systems susceptible to data breaches and misuse. In recent years, there have been numerous high-profile breaches that highlight the inherent risks of centralized identity management [21, 23, 58, 118].

Furthermore, the reliance on centralized authorities has placed undue trust on these data custodians. Users are forced to relinquish control over their personal identity information, with limited transparency and accountability regarding how their data is used. This lack of control raises concerns about potential privacy violations, misuse of identity data, and the possibility of the central authority gaining access to sensitive information about the usage of an identity [73, 131].

In response to these concerns, a new identity paradigm has emerged: Self-Sovereign Identity (SSI), aiming to revolutionize identity management by empowering individuals to take control of their digital identity in the absence of a centralized authority [5]. The essence of SSI lies in the issuance and verification of user credentials. These credentials, issued by trusted entities, attest to the authenticity of a particular attribute of a user’s identity. By holding the credentials, users eliminate reliance on third-party data custodians, reduce the potential for data breaches, and ensure they retain control over their identity.

While SSI bears the same underlying principle of fostering trust between unknown parties as previous systems [29], the extent of what “trust” means in this context has undergone considerable discussion. In both conceptual and implementation level discussions of SSI, there are disagreements on what trust is already available between participants and what trust needs to be built up through the design of the system [10, 72]. The discussions and summaries in this paper will illustrate how complex and fragmented the field of SSI has become. Without a clear agreement on the threats the participants face, there can be no guidance on what SSI must address in its design and requirements.

The paper aims to systematically analyze the elements of SSI in light of trust requirements and assumptions. We do this by breaking down SSI into its individual components to understand the underlying trust assumptions (or alternatively threats) that motivated their design requirements. This analysis provides insights into the design of an SSI system and identifies areas for further research and development. A detailed description of the data collection and analysis process is included in Appendix A.
1.1 Contributions
We claim the following three major contributions:

1. A comprehensive analysis of trust in SSI, introducing three distinct trust models to capture the varied levels of assumed trust across literature and implementations. These models are built by identifying the components of SSI and the threats they face from various actors in the system.

2. An analysis of known implementations of SSI, both in academic literature and industry products, detailing their trust assumptions made and the implementation approaches taken to address the assumed threats identified in (1).

3. A foundational framework for future SSI research and development by providing a comprehensive catalogue of SSI components and design requirements according to trust models. We identify shortcomings in existing SSI systems and highlight areas for further exploration.

The remainder of this paper is structured as follows: Section 1.2 provides an overview of related work. Section 1.3 introduces the core concepts of SSI, laying the foundation for subsequent discussions. Section 2 uses these core concepts to establish a minimal trust model for SSI, highlighting the design requirements necessary to mitigate associated threats. Section 3 delves into the threats that issuers, identity owners, and service providers face, extending beyond the minimal trust model. Section 4 consolidates these threats into two additional threat models. Section 5 evaluates the state-of-the-art academic literature and industry implementations, and Section 6 concludes with a discussion of open research directions.

1.2 Related Work
We situate our work within the landscape of well-established surveys on SSI and closely related topics. Through six categorizations, we differentiate our contribution from existing literature and present a summary in Table 1.

**Academic Literature and Industry Implementations:** We present a comprehensive analysis of both academic literature and industry implementations. There have been several surveys that collect and analyze academic literature on SSI. Soltani et al. [131] performed a thorough survey and provided a detailed and in-depth survey of SSI. They provided compelling motivation for SSI and then introduced core components with a strong focus on academic literature, briefly considering a subset of industry implementations. Schardong and Custódio [126] extensively covered publishing venues and provided a taxonomy of academic literature. Zhu and Badr [154] described a sub-set of academic and industry contributions focused on blockchain identity for the internet-of-things. Mühl et al. [108] did not consider any industry contributions, only surveying academic contributions. Ernstberger et al. [51] and Bistarelli et al. [17] both broadly collected DID methods from academia and industry. Lim et al. [95] and Satybaldy et al. [123] evaluated a sub-set of industry blockchain identity systems. Grüner et al. [65] only evaluated a single industry implementation.

**Catalogue of Components:** This work presents a catalogue of components of SSI and their design requirements. Schardong and Custódio [126] independently collected a large catalogue, while Ernstberger et al. [51] presented a limited catalogue.

**Threats and Trust Assumptions:** A key focus of our contribution is trust in SSI and the associated threats. Hardman [70] presented the most complete SSI threat model focused on preventing credential fraud. Grüner et al. [65] provided a threat model for a single implementation, while Ernstberger et al. [51] had limited discussion on threats faced. Lim et al. [95] introduced a trust model for blockchain implementations.

**Foundation Framework:** Satybaldy et al. [123] is the only contribution presenting a framework to evaluate future work. They evaluated five SSI constructions with a framework inspired by [5].

**Not Limited to Blockchain Identity:** A key distinction with our contribution is that we do not present a survey focused on blockchain-based SSI contributions. Schardong and Custódio [126] were the only other work not to have a scope limited to blockchain identities.

**Table 1: Related Works Comparison**

<table>
<thead>
<tr>
<th>Academic Literature</th>
<th>Industry Implementations</th>
<th>Catalogue of Components</th>
<th>Threats and Trust Assumptions</th>
<th>Foundational Framework</th>
<th>Not Limited to Blockchain Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schardong &amp; Custódio</td>
<td>Lim et al. [95]</td>
<td>Soltani et al. [131]</td>
<td>Ernstberger et al. [51]</td>
<td>Bistarelli et al. [17]</td>
<td>Mühle et al. [108]</td>
</tr>
<tr>
<td>Provided a framework</td>
<td>Evaluated a sub-set</td>
<td>Evaluated a single industry implementation</td>
<td>Highlighted areas for future research</td>
<td>Provided a comprehensive catalogue</td>
<td>Extensively covered publishing venues and provided a taxonomy of academic literature</td>
</tr>
</tbody>
</table>

Although not a survey, the closest work to ours is Hardman [70] who introduced a credential fraud threat model in a blockchain-based Aries RFC. Our review of existing SSI surveys identified a lack of consensus on SSI components and their underlying design motivations. No existing work provides a comprehensive catalogue of SSI components, an in-depth analysis of trust assumptions, and an extensive collection of academic literature and industry implementations presented in this paper. Consequently, no existing work serves as a sufficient foundational framework for future research.

1.3 Overview of SSI and Terminology
Self-sovereign identity (SSI) is founded on the issuance, storage, and presentation of identity credentials [56, 57]. Three distinct roles are involved in SSI:

**Identity Owners** require identification and collect credentials that represent their identities.

**Issuers** are authorized entities that issue credentials to identity owners (government registries issuing driver’s licenses).

**Service Providers** request credentials from identity owners to verify their identities.
To facilitate further discussion, we clarify the definition of an identity. Aligned with the Sovrin Glossary [57], this work defines identity as a collection of attributes, where each attribute represents a particular characteristic, such as name or age. Attributes are verified through attestations from issuers. These attestations are termed claims. A collection of claims compiled by an issuer to describe an identity owner is a credential.

Credentials are stored in a personal wallet, typically on the identity owner’s device. The exchange of credentials between parties is illustrated in Figure 1. Issuers share verifiable credentials with identity owners, who can then share verifiable presentations of these credentials with service providers. The additional term “verifiable” indicates that the credentials may be verified by the service providers; this concept is further discussed in Section 2.1.3. Service providers rely on the credential issuer’s authority without direct interaction. They infer the issuer’s identity from the credential and use their existing knowledge of the issuer and their authority to determine whether to accept the credential (only a government authority can issue a driver’s licence). This separation of credentials and isolation of the issuer from the service provider during credential presentations mitigate the concentration of issuer power that challenges the centralized identity models [5, 56].

This process of issuing, storing, and presenting identity credentials, known as the credential exchange, is often portrayed as the entirety of SSI [65, 126]. While it represents a crucial component of SSI, it does not capture all of the components. For that, we will turn to the Trust Over IP Stack.

2 TRUST OVER IP AND THE TRUSTFUL MODEL FOR SSI

The Trust Over IP (ToIP) Stack introduced by the ToIP Foundation is a four-layer architectural framework designed to establish trust between users over the Internet and other digital networks. It is inspired by the TCP/IP stack, which standardized packet exchange and enabled the development of the Internet. The ToIP Stack, however, emphasizes human trust through establishing digital identities and secure communication.

We outline the structure of SSI through the trust layers of the ToIP stack to analyze fundamental concepts and components of SSI in the context of forming trust. Following the analysis, we introduce a minimal trust model, our baseline trustful model.

2.1 SSI through Trust Over IP Layers

Our adaptation of ToIP layers tailored to SSI with the appropriate terminologies [57] is shown in Figure 2. We walk through each layer, introduce the relevant components of SSI, and summarise implementation choices available for the components.

2.1.1 Layer 1: Trust Support. Layer 1 provides a way to identify participants and verify the provenance of these identities through trust anchors. This layer provides a trusted foundation upon which the subsequent layers are constructed.

Identification. Every party represents themselves with a globally unique self-generated identity handle [51]. Recipients communicating with the party use this identifier to derive public attributes

2DID methods used for short-lived DIDs, where a trust anchor is not required, frequently require that the DID is the public key (did : key) [101], did : peer, [45]). In these instances, DIDDoc resolution involves extracting the public key directly from the DID itself.
implementation, building upon DIDs (Layer 1), DIDComm [40] establishes a suite of protocols for inter-entity messaging. DIDComm utilizes the public keys associated with DIDs to encrypt messages exchanged between entities.

2.1.3 Layer 3: Trust Tasks. Layer 3 facilitates the credential exchange. It defines the standardized, interoperable credential formats (e.g., driver’s licenses). It also specifies verification mechanisms (e.g., signatures), credential management protocols (collection, storage, presentation), and optional extensions to the core process.

Credential Formats. Credentials adhere to a standardized format comprising credential metadata, claims, and proof. Credential metadata encompasses the issuer’s identity (e.g., DID), credential type (e.g., driver’s license), and issuance date, among other details [132]. Claims represent the identity attributes verified by the issuer (e.g., name, birthdate, address, licence no.). The proof ensures the credentials are tamper-proof for integrity and provenance. Presentations are temporary, ephemeral documents containing verifiable credentials that service providers request. Some works consider presentations synonymous with verifiable credentials [56]. Regarding implementation, W3C verifiable credentials are the predominant data model, employing JSON-LD to define the credential structure [63, 151]. The AnonCreds JSON format links credential definitions to DIDs stored on VDR (Layer 1) [79].

Verification Mechanisms (Proofs). Credentials contain a verification mechanism as a proof of the credential’s authenticity. This proof may be provided through a signature, where the issuer signs the entire credential (metadata and claims) with their private key. Service providers verify the credential’s authenticity by retrieving the issuer’s public key from their DID. For implementation, ECDSA signatures are a common choice [87]. All surveyed credential profiles use signatures as a verification mechanism, although some do not present the signature to the service provider; instead, the identity owner presents a zero-knowledge proof of knowledge of a valid signature [28, 141] (discussed further in Sections 3.2.1 and 3.2.2).

Protocols. Protocols define the procedures controlling the exchange of credentials, encompassing how identity owners request credentials from issuers, how service providers request presentations from identity owners, and how identity owners present credentials to service providers. Many proposals implement various protocols. For instance, multiple related projects from OpenID address the issuing and presentation of verifiable credentials [98, 140, 149].

Extensions. Extensions add additional SSI features to the credential exchange. Our discussion utilizes these extensions to introduce threats and analyze their risk and mitigations to derive various trust models (Section 3).

2.1.4 Layer 4: Trust Applications. Layer 4 depicts the identity applications ecosystem (e.g., government, health, banking). Each application is considered to have Trust Assumptions. They define trusted and untrusted parties, interactions, and components. For instance, an application may distrust identity owners and assume they may present fraudulent credentials.

Deriving Trust Requirements from Trust Assumptions. Intuitively, a trust assumption implies a threat — e.g., suggesting the credential presented to the service provider could be fraudulent.

In fact, we can derive that a trust assumption (or lack thereof) implies the application accepting the presence of a threat. Threats pose a risk to some participants and must be addressed through constraints. That is, a Layer 4 application that distrusts identity owners and assumes they may present fraudulent credentials mandates the implementation of tamper-proof credentials in Layer 3.

Based on this observation, in our analysis of trust requirements for trust models, we consider that the human trust assumptions made in Layer 4 are stated as threats, which defines constraints that need to be addressed by the components of Layers 1-3.

As a summary, Table 2 outlines common implementations or standardization initiatives for each SSI component discussed in this section. Trust assumptions in Layer 4 determine the use of these components.

Table 2: Common SSI Trust Over IP Layer Implementations

<table>
<thead>
<tr>
<th>Layer 4 Application</th>
<th>Various trust assumptions impacting choices Layer 1-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Format</td>
<td>WSC, VC [132], JSON-LD [63], JVT/VC [109], SD-JVT/VC [139], AnonCred JSON [79], MDOIC [147], ICOA DTC [84], s-509, BIMA XML [55]</td>
</tr>
<tr>
<td>Verification Mechanism</td>
<td>ECDSA [87], RSA5, CL [28], BBS [141], BoundBBS [145], zk-SNARK [114], GGM-Merkle [109]</td>
</tr>
<tr>
<td>Protocol</td>
<td>CHAPI [31], WACI [8], Aries Protocols [77], OpenD4VCI [98], OpenD4VP [140], SOIPv2 [149]</td>
</tr>
<tr>
<td>Extensions</td>
<td>Detailed in Section 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Layer 2 Communication</th>
<th>DIDComm [40]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer 1 Identification</td>
<td>DID (indy [44], sov [97], ion [32], ebsi [48], key [101]), Any Public Key (x 509 [74])</td>
</tr>
<tr>
<td>Trust Anchor</td>
<td>Blockchain (indy DLT [83], Bitcoin [32], EBSI Trust Registry [49], Ethereum [53], Purity [138], Iota [82], Fabric [80]), ROTA [54], GNU Name System [123], Any PKI (CA, FIDO [6])</td>
</tr>
</tbody>
</table>

2.2 The Trustful Model

Here, we introduce the first trust model: trustful SSI. Trustful SSI encapsulates the fundamental principles of SSI, establishing the minimum responsibilities for each component. As mentioned, the model is represented as a collection of threats accepted as posing risks under the application’s trust assumptions. As a minimal model, it considers the threats relating to potential credential integrity and external actors, along with mitigation strategies. To clarify the design and responsibilities of the SSI components, we group the threats by components as a functional unit that implements a group of features.

2.2.1 Identity Owner Identification.

Threat 2.1. If identity owners can present credentials without a verification mechanism, they may present credentials (e.g., a driver’s licence) that were never issued or do not represent their identity (e.g., the identity owner is not licensed to drive).

Risk: A service provider may authorize an identity owner who should not have been authorized.

Constraint: Credentials (and the presentations that encompass them) must be verifiable. Credentials must be issued alongside some cryptographic proof that the service provider can use to verify the integrity of credentials.
Mitigations: **Identity owner identification** is implemented through the verifiable credentials in Layer 3. When the proof portion of verifiable credentials uses a signature to ensure the integrity of credentials the minimum requirement from the signature is three basic algorithms that provide key generation, signing, and verification functions. These signatures must satisfy correctness and security [87] such that a service provider will always accept a valid signature by an issuer, and importantly, an adversary cannot forge a signature that passes the verification function but was not created by the issuer.

### 2.2.2 Issuer Identification.

**Threat 2.2.** An issuer may not have the authority to issue credentials of any type (e.g., may only issue state driver’s licenses).

**Risk:** A service provider may accept a credential issued by an entity that did not have the authority to attest to the identity claims.

**Constraint:** The issuer of a credential must be identifiable from a verifiable presentation such that service providers can be assured that the claims they have been presented were issued by an issuer that they have some human trust in (trust that a government entity may issue driver’s licenses).

**Mitigations:** Layers 1 and 3 work together to address this constraint. Issuers embed and sign their DID in the Layer 3 verifiable credentials. Service providers then use this DID to derive what entity issued the credential and verify their authority. The immutability of credentials (Threat 2.1) ensures that once issued, the issuer’s DID can not be modified. The DID is then de-referenced to the publicly recorded DIDDoc to verify the identity and rights of the issuer.

### 2.2.3 Non-Transferability.

**Threat 2.3.** The identity owner may present a correctly issued credential; however, it was not issued to them. Rather, it was collected through collusion with another identity owner or stolen from an unknowing identity owner (e.g., Alice presents Bob’s driver’s licence as if it was hers).

**Risk:** A service provider may authorize an identity owner by incorrectly believing claims that do not actually describe them.

**Constraint:** Verifiable credentials and their presentations must be tied to the identity owner such that they can not be transferred incorrectly to another identity owner.

**Mitigations:** The identity owner’s DID is included in the verifiable credential. The immutability (Section 2.2.1) ensures that the DID cannot be modified after issuance. Ownership of the credential is provided by proving knowledge of the secret key associated with the public key of the DID (e.g., a Sigma protocol [41]). The issuer ensures that the identity owner knows the secret key when they embed the DID in the credential. Service providers perform the same verification when receiving presentations. For an adversary to present a verifiable credential as their own, they must have knowledge of the secret key associated with the embedded DID.

### 2.2.4 Protected Communication.

**Threat 2.4.** An outside entity may intercept a credential in transit, either passively or actively. Passive interception would allow the outside entity to learn about the contents of the presentation, while active interception would allow the outside entity to modify the contents of the presentation.

**Risk:** For identity owners, their credentials in transit may unknowingly have their private identity information stolen or modified.

**Constraint:** The credential exchange of Layer 2 must be secured to protect communication between entities from outside entities.

**Mitigations:** The encryption in DIDComm protocols (Layer 2) ensures that no eavesdropper may learn the contents of verifiable presentations. Additionally, credentials cannot be modified by a third party if they are encrypted in transit. This includes the threats of replay [136], and man-in-the-middle attacks [60] with protections such as token-binding in the Layer 3 verifiable presentations [100].

**Threat 2.5.** An outside entity may interfere with DID resolution, resulting in an incorrect DIDDoc being de-referenced.

**Risk:** A service provider may attempt to verify a credential using an incorrect DIDDoc of the issuer, resulting in either the authorization of an inauthentic or rejection of an authentic credential.

**Constraint:** Layer 1 must provide assurances in the provenance of resolved DIDDocs.

**Mitigations:** When de-referencing a DID, service providers must be sure to use protected communication channels. Additionally, the registries publicly hosting the DIDDocs should provide provenance guarantees that the DIDDoc has not been modified. At the very least, these registries need to provide some level of assurance specified by the Layer 4 application (e.g., blockchain-backed DID [53, 97], certificate authority-backed DID [62]).

# 3 Extending Trust Requirements

While adequate in many scenarios, trustful SSI may not fully address the intricacies and security demands of all digital identity applications. This is evident in government-issued digital identities, where the identity data may be highly sensitive. The trustful model’s assumption of service provider trustworthiness overlooks potential misuse or insecure storage of identity data.

This motivates more cautious approaches to SSI, where a broader set of trust requirements should be considered. For instance, since applications that follow the trustful model do not consider the service providers a threat, identity owners may share their entire driver’s licence credentials to prove their age. This is problematic as a dishonest service provider could collect the additional claims in a driver’s licence for malicious purposes or collect them innocently but store them insecurely.

Here, we broaden the trust requirements by introducing more threats perceived by issuers, identity holders and service providers.

## 3.1 For Issuers

The trustful model did not include threats that impacted the issuers. Here, we introduce an extended set of threats for issuers.

### 3.1.1 Identity On-Boarding.

New identity owners may be brought into an SSI system through a process known as on-boarding.
**3.2 For Identity Owners**

In the *trustful model*, the only threat to identity owners was the collection of credentials during transit. Here, we expand with the introduction of four new threats.

3.2.1 **Selective Disclosure. Data Minimalization**, collecting only necessary identity claims by the service providers, is a fundamental principle of SSI [5]. It is realized through the extension of *selective disclosure* at Layer 3.

**Threat 3.3.** A service provider may collect more identity claims than required, either maliciously or without harmful intentions, such as a bar collecting an identity owner’s name and address during age verification.

**Risk:** Excessive collection of claims by service providers elevates the risks faced by identity owners when presenting credentials. Malicious service providers may exploit identity data for financial gain or profiling purposes [23], while even well-intentioned service providers inadvertently elevate identity loss risks with their over-collection of identity data [156].

**Constraint:** Credential presentations must allow identity owners to present a subset of the claims in their credentials where service providers may not learn anything beyond what was presented.

**Mitigations:** Selective disclosure mechanisms can be achieved either through claim redaction or predicates.

**Claim Redaction** enables identity owners to remove claims from a verifiable credential when generating a presentation [157]. Service providers can verify the authenticity and integrity of the revealed claims without having access to the entire credential. Claim redaction describes an outcome where the service providers are only able to learn a subset of the claims that were initially included in a verifiable credential (e.g., a birthdate on a driver’s license). Some approaches require that the identity owner only share the claims they want to disclose along with the accompanying proofs [47]; even though claims are not strictly being redacted, the outcome remains the same.

**Predicates** allows identity owners to provide a predicate proof (TRUE/FALSE) on a claim (e.g., their birthdate claim indicates they are over 18) without revealing the actual claim value. The extent of the predicates available depends on the implementation chosen. Some implementations limit predicates to numerical ranges (e.g., age > 18) [28], while others enable more complex assertions (e.g., the listed address is within some defined city boundary) [94].

There are many proposals for implementing SSI selective disclosure categorized under three approaches: *atomic* (signing each claim individually), *hash records* (signing a hash of each claim rather than the claim itself), *zero-knowledge proofs* (either as a custom proof designed for a specifically signature scheme [28, 141], or a generic zero-knowledge proof on the claim and credential [94, 114]). While all of these approaches enable claim redaction, only *zero-knowledge proofs* support genuine predicate selective disclosure. This functionality can be emulated by having the issuer sign a pre-computed predicate claim (i.e. *age_over_18 = TRUE*) during credential creation.
3.2.2 Unlinkability.

Threat 3.4. Selectively disclosed claims provide limited privacy preservation if service providers can link presented claims to the same identity owner across consecutive presentations or through collusion with other service providers. Consent to a venue using presentations to track patronage or a car rental service linking claims with information from an insurance agent without consent.

Risk: If a presentation provided to a service provider contains an identifier unique to the identity owner, service providers can easily link an identity owner’s actions [35]. Service providers may then learn more about an identity owner’s identity than what the identity owner consented to.

Constraint: Verifiable presentations should not include any correlatable identifier the service provider can derive, such as identity owner DID or signature.

Mitigations: Single-show [14, 120] or k-show [93] credentials offer unlinkability only for a fixed (one or k) number of presentations of a credential. For effective SSI integration, credentials must enable identity owners to generate an unlimited number of unlinkable presentations from the same credential. Two unlinkability levels exist: complete and limited.

Complete Unlinkability: To achieve complete unlinkability, where no presentations may ever be linked to the same credential, two transformations must be applied. The removal of correlating factors addresses the vulnerability of a correlating DID identifier in every credential. Some works suggest managing the scope of DIDs, where the identity owner creates multiple DID pseudonyms, each used for a small subset of their credentials [101]. This still enables correlation for credentials issued to the same DID. Instead, verifiable credentials should not have DIDs (or other identifiers) embedded in them; DIDs remain only as a means for DIDComm communications and encryption (Layer 2) [40] and issuer identification (Layer 1) [143]. At times, this is the extent of unlinkability afforded (Coelho et al. [38], Takemiya and Vanieiev [137]).

Still, the signature of the credential remains a correlating factor. With the transformation of signatures, when the identity owner is preparing a credential for presentation, they generate a new proof that they then provide in place of the signature from the issuer. This new proof is an unlinkable zero-knowledge proof of knowledge of the credential’s signature. Here, anonymous credential signatures are used (BBS+) [24, 141], CL [26], PS [122]).

Limited Unlinkability: Complete unlinkability is not always desirable. In some situations, the service provider must be able to identify repeat presentations of identity owners, such as to prevent Sybil attacks [107]. If a service provider relies on SSI credentials for authentication into their service, such as through user accounts, then limited unlinkability is required, allowing service providers to identify repeat presentations; otherwise, a single credential could be used to create many accounts. The limited unlinkability construction from Zhang et al. [153] included a unique identifier derived from a combination of the identity owner and the service provider’s identifiers. Service providers may use this persistent identifier to identify repeat presentations but can not use it to correlate presentations made to other service providers.

3.2.3 Key and Wallet Management.

Threat 3.5. A malicious party could compromise and steal credentials or keys, or the identity owner may lose their keys.

Risk: An identity owner who loses their keys or credentials (or has them compromised by a malicious party) then loses control over their personal identity information and its use.

Constraint: Credentials and keys must be protected in wallets, ensuring they are not vulnerable to being stolen or lost.

Mitigations: The risk of losing keys and, in turn, access to the credentials in the identity wallet are not concerns unique to SSI [106]. Potential solutions are:

- Trusted Third Party Key Recovery. Soltani et al. [130] proposed a key recovery method using trusted third parties for SSI. The wallet of the identity owner is still in control of the user’s credentials. As a precaution, third-party providers back up the keys in escrow. Approaches vary across implementations, but all either rely on some form of trusted third-party service [90, 96] or friend [96, 111].

- Multi-Factor Authentication (MFA): With multi-factor authentication for credential presentations, the compromise of a wallet does not allow a malicious actor to present the compromised credentials as their own. Zhang et al. [153] implemented single sign-on using anonymous credentials following a structure compatible with Layer 3. Service providers could require that identity owners present the same credentials from multiple devices.

- Multiple Devices: The same design used by Zhang et al. [153] for multi-factor authentication also enables multiple devices. Since credentials can be stored on any number of devices, identity owners could present their credentials from any registered device.

Threat 3.6. Identity owners cannot always be expected to follow best practices regarding safeguarding and maintaining access to their wallets. Without proper consideration, identity owners could rely on insecure wallet protection mechanisms (e.g., credentials stored on devices without access protection through passwords).

Risk: Sacrificing credential security and sovereignty for usability places identity owners at risk of damaging their self-sovereignty (losing control of its use) or leakage of their credential data.

Constraint: Identity owners must maintain control over their credentials and keys regardless of how they are stored [5].

Mitigations: A solution such as Cloud Wallets places the entire responsibility of credential management with a cloud provider. Naively trusting a cloud provider with control over identity claims and their disclosure would violate the self-sovereignty of the identity owners (the cloud provider can dictate credential use). Furthermore, Lagutin et al. [91] identified that constrained (e.g., embedded) devices might be unable to handle the resource overhead required by verifiable credentials. Their work ensured the constrained devices maintained control over their DIDs, but the credential processing was delegated to an authorization server. Notably, this implementation requires potentially unfounded trust in the authorization server and is only appropriate for SSI deployments in limited scopes.

To address these trust concerns, Krenn et al. [88], later improved by Haböck and Krenn [68] presented anonymous credentials for the cloud. The proposed cloud wallets could never access the stored attributes in the clear, even when presenting the credentials. Their implementation utilized proxy re-encryption [11],...
structure-preserving signatures [59], and zero-knowledge proofs of knowledge [11]. To provide transparency of a cloud wallet’s actions, Chase et al. [33] proposed a credential transparency system that allowed an identity owner to audit the actions of their cloud wallet.

### 3.3 For Service Providers

Section 2.2 introduced several threats service providers face; here, we expand on the threats faced by service providers and introduce new threats.

#### 3.3.1 Non-Transferability

Previously, as a means to address non-transferability (Threat 2.3), the concept of embedding the identity owner DID in the credential during issuance was introduced. However, this is not acceptable when unlinkability needs to be preserved. Here, we expand the mitigation approaches towards non-transferability through proving what you know or what you own with linkable and unlinkable variations of each.

**What You Know:** DIDs embedded in a verifiable credential provide a suitable identifier to prove possession of a credential. This approach is the simplest and most often used [109]. This is visible non-transferability, providing service providers with a linkable identifier in every credential presentation.

A more robust solution is required to preserve unlinkability; we classify these as hidden. Camenisch and Lysyanskaya [28] presented protocols where the identity owners commit to a link secret when requesting a credential. The issuer verifies that the identity owner knows the committed link secret without learning the secret and signs the commitment. When identity owners present their credentials, they recommit to their link secret and share this pseudonym with the service provider. The service provider checks that the identity owner owns the committed secret. The identity owner also proves that the issuer signed a commitment on the same hidden secret. BBS signatures follow a similar principle [104]. Alternatively, the BoundBBS proposal uses BLS signatures [19] to bind credentials through knowledge of the secret BLS key without using the blind signing function of BBS. Using BLS signatures removes the need to manage the representation of the link secrets and simplifies the BBS specification [145].

**What You Have:** Physical ID cards often have photos of the holder to provide these same proof-of-possession assurances using biometrics. Othman and Callahan [119] proposed the Horcrux Protocol, an SSI construction based on biometrics. This construction established DIDs linked to biometric factors, tying credentials issued to a specific identity owner. However, this implementation did not provide the unlinkability features and is a visible biometric solution.

As a hidden implementation of biometric non-transferability, Impagliazzo and More [85] introduced a scheme that ties anonymous credentials to biometric identifiers. Their construction enabled non-transferability without compromising the privacy of the identity owner. Adams [4] and Gerdes et al. [61] presented improvements on this scheme.

Credentials may also be tied to the identity owner’s specific device (i.e., the device that hosts the credential wallet). These identifiers do not always have to be public/private key pairs; in the Internet of Things (IoT) context, Niya et al. [116] exploited the manufacturing variability of devices to produce a unique device identifier to embed in credentials.

#### 3.3.2 Credential Validity (Revocation)

**Threat 3.7.** Credentials have a lifespan; identity owners can lose the privileges associated with a credential (e.g., having the right to drive revoked or a university identity card expiring after a student graduates).

**Risk:** A service provider who accepts a credential that the issuer has revoked is vulnerable to accepting identity owners who no longer have authority over the credentials they present.

**Constraint:** Service providers must be able to determine if the credentials they are presented are still valid.

**Mitigations:** We can broadly categorize the mitigating approaches as either expiring or immediate.

With expiring credential validity, the issuer specifies a validity period. When a service provider is presented with a credential, they reject any credentials for which the validity period has passed. Credential validity can be renewed through the re-issuance of the credentials. Rather than requiring the whole credential to be re-issued, issuers can instead generate lightweight attestations to the validity of the credential that accompanies the presentation. Work by Abraham et al. [2] proposed a revocation scheme where the Layer 1 network provides these attestations. Camenisch et al. [25] provided a similar construction, where issuers regularly update and publish values that identity owners retrieve and use to re-validate their expired credentials.

These short lifespans allow for efficient verification of credentials. However, they are unable to provide immediate revocation, and a balance must be struck between having a short enough validity period that credentials can be revoked in an acceptable time frame while not being too short that issuers are over-burdened re-issuing or attesting to credential validity.

With immediate revocation, the validity of a credential is altered with direct effect. Commonly, this is accomplished through a publicly available list of credential statuses (either listing the revoked or valid credentials). Issuers update this record if a credential must be revoked; service providers reference the list when verifying a credential. Abraham et al. [3] proposed using a revocation list utilizing the existing Layer 1 distributed ledger.

Rather than maintaining a list of revoked or valid credentials, accumulators can collect the identifiers of all valid credentials into a single value that is published [36]. When a credential is revoked, the accumulator is updated to no longer include the revoked credential. To preserve the unlinkability requirement, the identity owner proves in zero-knowledge that their credential is registered in the dynamic accumulator [27].

Instead of a list or accumulator – but still with direct effect – reclaimID from Schanzenbach et al. [124] used the online nature of their identity records to update the record’s status.

Compared to expiry credential validity, immediate revocation removes the burden of re-issuing credentials or re-publishing values for the issuers. In exchange, the identity infrastructure must accommodate the revocation or status list.

#### 3.3.3 Identity Owner Identification

**Threat 3.8.** Identity owners may not always be collecting credentials from trusted issuers. There may be situations where identity owners self-published claims of their identity.
Risk: Service providers are provided with no indication of the trustworthiness of an identity owner’s self-published credentials.
Constraint: Presentations of self-published credentials should provide some confidence for service providers in the claims truth.
Mitigations: Reputation of Self-Published Claims is an approach where reputation is built by collecting attestations to the validity of credentials in a web-of-trust approach [135]. Aside from variations across implementations and how reputation is collected and measured, implementations vary in their motivations as to why there is no trusted issuer.

Threat 3.9. Identity owners may need to delegate the authority of their credentials to another identity owner in an exceptional circumstance (e.g., in healthcare).
Risk: Without delegation capabilities, identity owners who are in situations where they are unable to present their credentials are essentially left without their identity.
Constraint: Credentials should enable their authority and claims to be shared with another identity owner when required.
Mitigations: For Delegation of Credentials, the identity owner takes on the role of an issuer, passing on the rights of their credential to a new identity owner [110]. Delegation can follow a chain, such as in the example of a hospital that issues a credential to a doctor who then issues the rights to their associates when they are on leave [43] or a patient who delegates a letter of their authority to a friend or family member [110]. Chase and Lysyanskaya [34] (and Demuynck et al. [43]) introduced delegatable anonymous credentials as a further constraint that the delegating issuer should not be identifiable from the presentation of a delegated credential.

Threat 3.10. The ownership of a credential issued by a trustworthy issuer does not indicate that the identity owner will behave appropriately. From Sections 3.2.1 (selective disclosure) and 3.2.2 (unlinkability), the service provider may not even be able to identify a misbehaving identity owner. Consider a case where SSI is used when renting a car; a driver’s license credential, while proving identity and a right to drive, offers no guarantee of safe driving.
Risk: Service providers may enter into a relationship with untrustworthy identity owners, and become unacceptably vulnerable.
Constraint: Service providers must be afforded some level of recourse or protection against identity owners who misbehave1.
Mitigations: Using Blacklisting and Reputation, when unlinkability is not considered, a service provider can easily blacklist a misbehaving identity owner from their service through their DID. With unlinkability, this becomes a more complex problem. Additionally, one service provider may wish to report their experiences to other service providers while maintaining the identity owner’s privacy. Tsang et al. [144] proposed a scheme that supported the blacklisting of anonymous credentials. Nakanishi and Kanatani [112] expanded on this by allowing service providers to contribute to a reputation score of identity owners rather than a boolean blacklist.

Anonymity Revocation: Another challenge introduced by unlinkability is how service providers can take recourse (beyond just blacklisting from their service) against misbehaving identity owners. With roots in group signatures [18], anonymity revocation enables a service provider, supported by a mutually trusted third party between the identity owner and service provider, to revoke the anonymity provided by unlinkability. Camenisch and Lysyanskaya [26] proposed that identity owners include some traceable identifier encrypted against the public key of a trusted revocation manager. The encrypted identifier included de-anonymization conditions to ensure the service provider could not tamper with the agreed-upon de-anonymization policy between the service provider and the identity owner.

3.3.4 Issuer Identification. The authority to issue credentials is not enforced through SSI; this falls on the human trust and governance layer of SSI (Layer 4) [56]. As a brief expansion, some SSI proposals provide a means for the service provider to build trust in issuers that may not be traditionally recognized via reputation.

Issuer Reputation: From the same motivation of Threat 2.2, issuer reputation mechanisms enable service providers to learn some calculated reputation scores about an issuer. This reputation aims to measure the trustworthiness of an issuer’s attestations. Bhattacharya et al. [16] used the number of credentials issued by the issuer in a time window. Additionally, reputation may be measured against the specific claims of a credential rather than the issuer [142].

3.3.5 Offline Verification.

Threat 3.11. The identity owner’s control of credentials in SSI lends itself well to offline operations where the identity owner or service provider may not have reliable Layer 1 network access (e.g., verification of a credential in a remote region).
Risk: Service providers may be denied from accepting (or identity owners presenting) valid credentials if their Layer 1 network connection is interrupted or unavailable.
Constraint: The Layer 3 credential exchange between identity owners and service providers must only rely on Layer 2 peer-to-peer communications and not require access to the Layer 1 network.
Mitigations: The self-contained proofs (signatures) in verifiable credentials permit offline verification. The peer-to-peer communications of Layer 2 and pair-wise DIDs ensure neither the service provider nor the identity owner is required to reference the Layer 1 network during credential exchange. What is lost is the service provider’s ability to de-reference issuer DIDs, find credential structures, and check revocation registries. Solutions to the DID and structure challenges are agnostic of implementation but introduce further constraints on the situations in which offline verification could be performed. If the service provider is offline, they must know the credential structure they are requesting and have dereferenced the issuer’s DID to their public key before being offline.

In many cases, this is acceptable, especially if the service provider has intermittent access to the network to update their local cache. What poses a more significant challenge is Threat 3.7 (credential validity). Without immediate access to a revocation or status list, immediate credential validity is incompatible. Credential status options are likely constrained to expiry [2]. Furthermore, key and wallet management cannot rely on possibly inaccessible cloud wallets (Threat 3.6). Some biometric implementations for non-transferability (Threat 2.3) rely on server access for secure verification of credentials [119], making them incompatible with offline verification.

---
1The concept of misbehaving is situational to the identities use and the service provider’s policies.
4  THE TRUST MODELS
This section analyses the extended trust requirements, presented as threats and mitigations in Section 3, and introduces two additional models as alternatives to the trustful model (Section 2.2).

4.1  Threat Dependencies
The discussions in Sections 3.1, 3.2, and 3.3 revealed implicit dependencies between threats as they emerged. Figure 3 summarizes these dependencies.

![Figure 3: SSI Threat Dependencies](image)

The privacy protection provided by selective disclosure is compromised if a malicious or colluding service provider may correlate presentations. This necessitates the inclusion of Threat 3.4 (unlinkability) if Threat 3.3 (selective disclosure) is included in the trust assumptions.

Threat 2.3 (non-transferability) necessitates non-transferable credentials. Approaches that utilize an embedded identifier visible to service providers introduce a correlatable identifier that enables linking. In the presence of Threat 3.4 (unlinkability), hidden non-transferability constructions must be employed to mitigate this risk.

Similarly, for Threat 3.7 (credential validity), approaches that employ credential validity lists must exclude correlatable identifiers from the lists and instead adopt an approach where identity owners prove their membership in an accumulator.

The anonymity potentially provided by unlinkability (Threat 3.4) may place the service providers at risk, suggesting Threat 3.10 (blacklisting, reputation, and anonymity revocation).

The incorporation of Threat 3.11 (offline verification) imposes limitations on Threat 3.7 (credential validity), restricting implementations that require Layer 1 access (immediate via lists).

The consideration of these dependencies is relevant to discussions involving alternative models and the range of trust assumptions within each model. Omitting to address one threat may require incorporating another related threat mitigation.

4.2  The Trust Models of SSI
The design of SSI components is influenced by underlying trust assumptions, often implicitly embedded in academic literature and industry implementations. We identify and categorize these implicit trust assumptions into three distinct trust models for SSI, representing the overarching approaches to trust. These trust models define the scope of what is untrusted in the Layer 4 application; effectively, the models identify the threats users face. The three models represent varying levels of trust, with each subsequent model assuming less trust, accepting less risk, and subsequently capturing more threats than the previous. We further explore the implications of these assumptions and provide recommendations on what threats are captured by each model in the Implications on Threat Mitigation sections. Table 3 summarizes the recommended threats captured by each trust model, providing a reference of the remaining discussions in this section.

4.2.1  Trustful Model.

**General Concept.** The trustful model (from Section 2.2) assumes as few threats as possible. Implementation approaches that follow this trust model often do not consider the sensivity of the identity data to be high. This model assumes complete trust in issuers regarding credential issuance and indicates no trust concerns about service providers’ behaviour.

**Implications on Threat Mitigation.** As the most trusting model, the captured threats must provide a functional identity system. Credentials should be verifiable by service providers (Threat 2.1 identity owner identification). Service providers should be able to ascertain the issuer of a credential (Threat 2.2 issuer identification) and the ownership of credentials (Threat 2.3 non-transferability). Additionally, the Layer 2 communication channels employed and the Layer 1 identification should provide assurances of security (Threats 2.4 and 2.5 protected communication).

Implementing an SSI system under this model necessitates a rudimentary Layer 1 DID infrastructure for issuer identification, Layer 2 protocols like DIDComm [40] to safeguard communications, and a credential format coupled with a simple verification mechanism such as an ECDSA signature [87].

4.2.2  Intermediate Trust Model.

**General Concept.** The intermediate trust model extends upon the trustful model with a lower risk tolerance, making it more suitable for applications handling sensitive data. Service providers are no longer trusted, and all entities adopt a more cautious approach to potential threats. Implementations adhering to the intermediate trust model often target large-scale identity systems, such as those used for government-issued credentials, as opposed to internal company authentication tools.

**Implications on Threat Mitigation.** To safeguard identity owners from the now untrusted service providers, Threat 3.3 (selective

---

### Table 3: Trust Models and Recommended Threats

<table>
<thead>
<tr>
<th>Extension / Feature</th>
<th>Threat</th>
<th>Trustful</th>
<th>Intermediate Trust</th>
<th>Zero Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identity Owner Identification</td>
<td>2.1</td>
<td>●</td>
<td>○</td>
<td>●</td>
</tr>
<tr>
<td>Issuer Identification</td>
<td>2.2</td>
<td>●</td>
<td>○</td>
<td>●</td>
</tr>
<tr>
<td>Non-Transferability</td>
<td>2.3</td>
<td>●</td>
<td>○</td>
<td>●</td>
</tr>
<tr>
<td>Protected Communication</td>
<td>2.4</td>
<td>●</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Issuance of Sensitive Claims</td>
<td>3.2</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Selective Disclosure</td>
<td>3.3</td>
<td>●</td>
<td>○</td>
<td>●</td>
</tr>
<tr>
<td>Unlinkability</td>
<td>3.4</td>
<td>○</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Key and Wallet Management</td>
<td>3.5</td>
<td>○</td>
<td>€</td>
<td>€</td>
</tr>
<tr>
<td>Credential Validity (Revocation)</td>
<td>3.7</td>
<td>○</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Offline Verification</td>
<td>3.11</td>
<td>○</td>
<td>●</td>
<td>○</td>
</tr>
</tbody>
</table>

- ● required  ○ likely  € optional extension  @ forbidden
disclosure) should be considered. The lower risk tolerance also indicates that Threat 3.7 (credential validity) should be included. Following the dependencies outlined in Figure 3, Threat 3.4 (unlinkability) would also need to be included. The implementation of credential validity (Threat 3.7) and Threat 2.3 (non-transferability) must be restricted to their unlinkable/hiden variants, and Threat 3.10 (blacklisting, reputation, and anonymity revocation) may be considered viable.

Implementations do not necessitate modifications to Layer 1 or 2. Layer 3 would require the selection of a signature scheme that supports selective disclosure and unlinkability. Furthermore, the format and protocols may need to be extended to accommodate the additional extensions.

4.2.3 Zero-Trust Model.

General Concept. The zero trust model challenges the notion of a predefined set of trusted issuers with absolute authority to verify identities. Under this assumption, any form of central authority is considered a threat that demands mitigation. These systems embody a fully decentralized identity system.

Implications on Threat Mitigation. To eliminate the concept of trusted issuers, Threat 2.2 (issuer identification) is removed and prohibited from the model. Threat 3.1 (identity on-boarding) is also incompatible. Threat 3.8 (self-attestation of claims) is introduced to augment identity owner identification. Beyond these changes, the low-risk tolerance implies the same trust assumptions as the intermediate trust model.

In fully trustless systems, the centralized authority of traditional public key infrastructure (PKI) is no longer appropriate and, if present, is replaced with decentralized trust anchors, typically in the form of a verifiable data registry (VDR) such as a blockchain.

The absence of established issuers necessitates reconsidering the Layer 3 credential exchange structure. Credentials and their corresponding proof mechanisms must support alternative assurance mechanisms, such as web-of-trust and self-attestation of claims.

5 THE STATE OF SSI

This section examines the current landscape of self-sovereign identity (SSI) from the perspective of the components and threat assumptions discussed earlier. Both academic literature and industry implementations are evaluated. The component selections and the approaches adopted for the extensions are presented. Columns were omitted in instances where no information was available.

Table 5 presents a compilation of academic contributions gathered through the literature review outlined in Appendix A. The double horizontal line separates works that present a complete SSI construction (top) from those that only introduce a variation of an extension (bottom). Table 6 lists industry implementations. These implementations are commonly referred to as credential profiles [89]. In both tables, (−) indicates that the component or extension was not considered in the implementation. Table 4 further categorizes the approaches employed for the extensions in the credential profiles.

Several notable distinctions between academic and industry SSI implementations underscore the divergence between the two domains. Most prominently, the choice of trust anchors differs significantly. Academic constructions predominantly rely on blockchains, such as the identity-specific Indy platform [81], as trust anchors. This extensive use of blockchains has led some to erroneously categorize SSI as a blockchain identity. However, this is inaccurate, as the majority of credential profiles in Table 6 utilize traditional public key infrastructure. This fundamental difference stems from varying trust assumptions and differing risk priorities. None of the industry implementations adheres to the zero-trust model, which disallows the existence of trusted authorities.

Still, there are commonalities across domains. Both academic and industry approaches often share fundamental goals, such as agreeing on broad trust models for privacy protection (intermediate and zero-trust models) or aiming for a practical digital identity solution (trustful model). Despite these shared objectives, differing risk perspectives lead to variations in component choices within the overarching framework. Notably, these implementation differences do not indicate significant changes in the framework or architecture. Instead, adapting to address potential threats often requires simply swapping specific components within the established structure.

The trust assumptions listed for each contribution of Tables 5 and 6 reflect the authors’ implicit threat assumptions, aligning with the models presented in Section 4. Comparing the implemented constraints or mitigations through the selected components and extensions against the recommended threats for each model category reveals that few works fully address the threats implied by the intermediate or zero-trust models.

These discrepancies do not imply any fault in the authors’ implementations; however, they do emphasize the need for explicit threat definitions and careful consideration of the requisite extensions and components. Ultimately, achieving mutual agreement on a framework to capture SSI as well as a catalogue of threats and mitigations available can foster a more consistent approach to SSI across all domains.

### Table 4: Credential Profile Extension Categorizations

<table>
<thead>
<tr>
<th>Selective Disclosure</th>
<th>None</th>
<th>ECDSA [87], BBG+C [141], BoundBBS [145], GGM-Merkle [209]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Redact</td>
<td>CL [28], ZKP (zkSNARK) [143]</td>
<td></td>
</tr>
<tr>
<td>Predicates</td>
<td>None</td>
<td>ECDSA, RS256, GGM-Merkle, BLS</td>
</tr>
<tr>
<td></td>
<td>Limited</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Complete</td>
<td>CL, BBG+C, BoundBBS, ZKP (zkSNARK)</td>
</tr>
<tr>
<td>Non-Transferability</td>
<td>What you know Linkable ID Key, JWT [13], Any Key</td>
<td>Hidden Link Secret, BLS Key (BoundBBS)</td>
</tr>
<tr>
<td></td>
<td>What you have Linkable Biometric or Device</td>
<td>Hidden -</td>
</tr>
<tr>
<td>Revocation</td>
<td>None</td>
<td>Expiry Epoch Expiry</td>
</tr>
<tr>
<td></td>
<td>Immediate</td>
<td>AnonCreds Revocation [78], W3C Status List 2021 [99], EBSI Status List [50], JWT and CWT Status List [102]</td>
</tr>
</tbody>
</table>

* When used with SD, JWT/VC, redaction is possible [139]. ECDSA does not offer selective disclosure with the other credential formats in Table 6.
Table 5: Academic Contributions Comparison

<table>
<thead>
<tr>
<th>Trust Anchor</th>
<th>Identification</th>
<th>Signature</th>
<th>Trust Assumption</th>
<th>Selective Disclosure</th>
<th>Unlinkability</th>
<th>Non-Transparency</th>
<th>Credential Validity</th>
<th>Key &amp; Wallet Management</th>
<th>Identity On-Boarding</th>
<th>Issuance of Sensitive Claims</th>
<th>Identity Owner Identification</th>
<th>Offline Verification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooiho et al. [38]</td>
<td>Blockchain Public Key</td>
<td>-</td>
<td>Intermediate</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lin and Ren [115]</td>
<td>Ethereum Public Key</td>
<td>-</td>
<td>Intermediate</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Stincklein et al. [134]</td>
<td>Any PKI Public Key</td>
<td>ZKP</td>
<td>Intermediate</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lahorde et al. [90]</td>
<td>FIDO PKI</td>
<td>FIDO</td>
<td>Trustful</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Borce et al. [20]</td>
<td>Ethereum Public Key</td>
<td>ZKP</td>
<td>Intermediate</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Muits et al. [109]</td>
<td>Parity DID</td>
<td>GM-Merkle</td>
<td>Intermediate</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Luxeck et al. [185]</td>
<td>IOTA DID</td>
<td>Zero-Trust</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Stincklein and Pouwelse [185]</td>
<td>Blockchain Public Key</td>
<td>ZKP</td>
<td>Zero-Trust</td>
<td>≤</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Alsayed Kassem et al. [7]</td>
<td>Ethereum Public Key</td>
<td>-</td>
<td>Intermediate</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lagitin et al. [91]</td>
<td>Indy DID</td>
<td>-</td>
<td>Trustful</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Abraham et al. [11]</td>
<td>Blockchain DID</td>
<td>RLS</td>
<td>Intermediate</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lee et al. [94]</td>
<td>Blockchain DID</td>
<td>zkSNARK</td>
<td>Intermediate</td>
<td>≤</td>
<td>√</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ham et al. [65]</td>
<td>Indy DID</td>
<td>-</td>
<td>Intermediate</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Abraham et al. [3]</td>
<td>Blockchain DID</td>
<td>RLS</td>
<td>Intermediate</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Scharenbreich et al. [124]</td>
<td>GNU Name System Public Key</td>
<td>-</td>
<td>Zero-Trust</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Abraham et al. [2]</td>
<td>Blockchain DID</td>
<td>RLS</td>
<td>Intermediate</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Takemiya and Yoneyue [137]</td>
<td>Iroha DID</td>
<td>-</td>
<td>Intermediate</td>
<td>×</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cho et al. [58]</td>
<td>Indy DID</td>
<td>CL</td>
<td>Intermediate</td>
<td>≤</td>
<td>√</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>


Key and Wallet Management: MD = multi-device TTP = trusted third-party MFA = multi-factor authentication FIDO = key management offloaded to FIDO components

Table 6: Credential Profiles Comparison

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>JSON-LD + Bone DID Claims [15]</td>
<td>CA PKI</td>
<td>did:web</td>
<td>RLS Public Key</td>
<td>JSON-LD</td>
<td>BoneBLS</td>
<td>Intermediate</td>
<td>×</td>
<td>√</td>
<td>RLS Key</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>JSON-LD + Bone DID Claims [15]</td>
<td>CA PKI</td>
<td>did:web</td>
<td>RLS Public Key</td>
<td>JSON-LD</td>
<td>BoneBLS</td>
<td>Intermediate</td>
<td>×</td>
<td>√</td>
<td>RLS Key</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>JWTC VC [189]</td>
<td>BitBox</td>
<td>did:web</td>
<td>did:web</td>
<td>JWTC VC</td>
<td>EBSI</td>
<td>Trustful</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>JWTC VC [189]</td>
<td>BitBox</td>
<td>did:web</td>
<td>did:web</td>
<td>JWTC VC</td>
<td>EBSI</td>
<td>Trustful</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>NGT Atlantic for Open BBACs [42]</td>
<td>CA PKI</td>
<td>did:web</td>
<td>did:web</td>
<td>NGT Atlantic for Open BBACs</td>
<td>EBSI</td>
<td>Trustful</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>NGT Atlantic for Open BBACs [42]</td>
<td>CA PKI</td>
<td>did:web</td>
<td>did:web</td>
<td>NGT Atlantic for Open BBACs</td>
<td>EBSI</td>
<td>Trustful</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>Dutch Decentralized Identity [17]</td>
<td>CA PKI</td>
<td>did:web</td>
<td>did:web</td>
<td>JWTC VC</td>
<td>EBSI</td>
<td>Trustful</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>Dutch Decentralized Identity [17]</td>
<td>CA PKI</td>
<td>did:web</td>
<td>did:web</td>
<td>JWTC VC</td>
<td>EBSI</td>
<td>Trustful</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>Microsoft Extra did:web [128]</td>
<td>CA PKI</td>
<td>did:web</td>
<td>did:web</td>
<td>JWTC VC</td>
<td>EBSI</td>
<td>Trustful</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>Microsoft Extra did:web [128]</td>
<td>CA PKI</td>
<td>did:web</td>
<td>did:web</td>
<td>JWTC VC</td>
<td>EBSI</td>
<td>Trustful</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>Selective Disclosure JWT [139]</td>
<td>Any PKI</td>
<td>Any Public Key</td>
<td>Any Public Key</td>
<td>SD-JWT</td>
<td>EBSI</td>
<td>Trustful</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>EBSI [133]</td>
<td>DID Key</td>
<td></td>
<td></td>
<td>EBSI</td>
<td>Trustful</td>
<td>Intermediate</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>OpenBBAC High Assurance [156]</td>
<td>Any PKI</td>
<td>Any Public Key</td>
<td>Any Public Key</td>
<td>JWTC VC</td>
<td>EBSI</td>
<td>Trustful</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>DID mTLS (RFC 8673) [147]</td>
<td>x.509 + CA PKI</td>
<td>X.509 Public Key</td>
<td>X.509 Public Key</td>
<td>X.509</td>
<td>EBSI</td>
<td>Trustful</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>IAM Direct [15]</td>
<td>Any PKI</td>
<td>Any Public Key</td>
<td>Any Public Key</td>
<td>OpenID4VC</td>
<td>EBSI</td>
<td>Trustful</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>EBSI [133]</td>
<td>DID Key</td>
<td></td>
<td></td>
<td>EBSI</td>
<td>Trustful</td>
<td>Intermediate</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
<tr>
<td>IBSA (Trusted Wallet) [22]</td>
<td>Any PKI</td>
<td>Any Public Key</td>
<td>Link Secret</td>
<td>BMA XML</td>
<td>CL</td>
<td>Intermediate</td>
<td>×</td>
<td>√</td>
<td>JWTC VC</td>
<td>W3C Status List 2023</td>
<td>-</td>
</tr>
</tbody>
</table>


Selectitive Disclosure: ≤ = predicate × = claim reduction Unlinkability: √ = complete - = none Key and Wallet Management: MD = multi-device TTP = trusted third-party CW = cloud wallets Offline Verification: × = supported × = possible

Aggregation of Signatures, Identity On-Boarding, Issuance of Sensitive Claims, Identity Owner Identification excluded from table as no credential profile considered them.

6 OPEN WORK

Post-Quantum Resistance; Existing signature schemes employed in SSI systems, as outlined in Table 5 and 6, lack quantum-safe resilience. This vulnerability extends beyond Layer 3 signatures to encompass the cryptographic underpinnings of Layers 1 and 2, potentially jeopardizing the entire cryptographic architecture depicted in Figure 2. While post-quantum secure solutions have emerged for public-key management in Layers 1 and 2, Layer 3 solutions remain limited. Though Jeudy et al. [36] proposed a quantum-secure anonymous credential scheme suitable for Layer 3 signatures, Dutto et al. [47] highlighted the continued challenge of achieving efficient selective disclosure (predicates or redaction).

Interoperability of Credentials: Without broad interoperability of SSI credentials, identity owners will again be siloed into identity systems that have chosen compatible constructions. Interoperability must be considered for each of the layers. Layer 1 constructions that use DIDs conforming to the standards of [133] are, by design, interoperable. Any entity that needs to dereference a DID can follow the specification of the defined method and collect the standardized DIDDoc. Credential and presentation incompatibilities represent the most significant obstacle to interoperability. The protocols, signatures, credential format and definitions, claims canonicalization, and encoding methods must all be compatible.

Verifiable credentials based on JSON-LD, such as the W3C VC Data Model [132], claim compatibility with JSON credentials [152],
but this has been disputed [148]. AnonCreds employs a proprietary credential format [79], deviating from W3C verifiable credential standards. The ISO 18013-5 mobile Driver’s License proposal also adheres to a proprietary format [147].

SSI interoperability remains an ongoing challenge, and to date, there is no definitive consensus on what this should look like. The disparity of implementations in Tables 5 and 6 serves as evidence.

**Key Management:** Section 3.2.3 introduced the concept of multi-factor authentication (via presentations). The proposal by Zhang et al. [153] required that identity owners present their credentials multiple times from distinct devices. This is inefficient as all that is required is that identity owners are protected against a single on-device wallet compromise.

In addition, the protection of keys used by issuers when signing credentials through key rotation has been a source of contention within the standards [71, 76]. The online nature of resolved DIDDocs provides many methods for issuers to rotate or revoke their signing keys. With rotation, a credential issued under a key pre-rotation is considered valid if issued when the key was valid. Revocation of a signing key indicates that the key has been compromised, and existing credentials signed under the key should not be accepted. This is further complicated as assuming an issuer is malicious and holds a compromised key, the issue date embedded in the credential cannot be trusted. AnonCreds relies on the Layer 1 node attesting to the issuance date, further complicating issuance and verification and introducing further challenges to offline verification (Section 3.3.5). Protecting these keys and their lifecycles is critical to the resilience of SSI, and yet, there is a gap in addressing this challenge.

**Transparency of Service Providers:** Service providers establish trust in issuers through either pre-existing trust relationships (using DIDs to link the trusted issuer with the credential signer) or credential reputation (Threat 3.10). They build trust in identity owners through the Layer 3 credential exchange. What is missing is the transparency of service providers. Identity owners must hope that the service provider they interact with is honest. Often, it is up to the identity owner to evaluate the risk of disclosure.

Some discussions of digital identities, particularly those from government authorities, propose that the issuing entity accredit service providers (see Australia’s proposed digital identity law [117]). However, this approach risks further siloing identities and hindering interoperability. The current state-of-the-art for protecting identity owners from malicious service providers offers two options: limiting service providers to a pre-approved list of accredited entities or relying on selective disclosure and unlinkability to safeguard identity owners from unfettered service providers. The effectiveness of these measures remains uncertain, and there is a strong likelihood that they may not be sufficient. Consider a scenario where a service provider must collect a claim that uniquely identifies the identity owner (i.e., a student number id) because the service provider does not trust the identity owner, then any attempt at unlinkability mitigation’s is destined to fail. To address this, the discussion of Threat 3.10 introduced some mitigations (blacklisting, reputation, and anonymity revocation) that may enable a service provider to find trust in an unlinkable identity owner. Yet, this balance between the competing motivations of trust and risk priorities remains underexplored, with no clear solution to appease both parties.

**Unlinkability Remains Ambiguous:** Section 3.2.2 delves into the complexities surrounding unlinkability in SSI. Correlating unlinkability and anonymity is a typical response when presented with unlinkability. Anonymous credentials – one of the core technologies in many SSI constructions – even bear the name. However, this does not adequately communicate the nuances. Anonymity encompasses a broader scope than unlinkability. To provide anonymity would require that the SSI construction prevent tracking of identity owners through means beyond signatures in the presentations. Correlatable identifiers can be extracted from network protocols. TOR [46] may be able to address this in an online credential exchange; peer-to-peer credential exchanges proposed on top of layers such as NFC, Bluetooth, and QR Codes also face this challenge and will require consideration if total anonymity is required.

None of this matters, if the disclosed claims place identity owners in small identity, sets such that they could be profiled regardless. It remains disputed if applications should take on the burden of unlinkability if anonymity cannot be provided (evident by the number of credential profiles in Table 6 that provide selective disclosure without unlinkability). At a minimum, future contributions must clarify their position on unlinkability, communicating what level of unlinkability their construction provides to identity owners.

### 7 CONCLUSION

As digital identities become unavoidable, Self-Sovereign Identity (SSI) offers a promising direction for individual control and privacy. Through a systematic analysis of SSI’s components and examining trust assumptions and threats posed by various actors, we developed three distinct trust models. These models capture the diverse levels of trust found across different SSI implementations and literature. We highlight the multifaceted nature of trust in SSI and how various stakeholders perceive and approach trust differently. Our comprehensive catalogue of SSI components and trust-related design requirements serves as a valuable resource for practitioners and researchers.

**For practitioners.** Practitioners tasked with designing an SSI deployment are provided with a framework to build their construction (Section 2.1). The trust models from Section 4 define a broad categorization and reference of the assumed risks. The catalogue of components, design requirements, and existing implementations in Sections 2.2, 3, and 5 may be referenced to select the appropriate mitigations based on the specific threats identified in the chosen trust model.

**For researchers.** Researchers can use this same framework and catalogue of components to locate their work in the field of SSI. Additionally, we support the justification of further work through our trust models and threats and identify open, unaddressed work.

### ACKNOWLEDGMENTS

The work has been supported by the Cyber Security Research Centre Limited whose activities are partially funded by the Australian Government’s Cooperative Research Centres Programme. This research has been supported by an Australian Government Research Training Program (RTP) Scholarship.
We started our systematization with the seed article *The Path to Self-Sovereign Identity by Allen [3]*. These articles were collected through ResearchRabbit. These articles were filtered and included if they met the following requirements:

1. The work proposed an SSI implementation.
2. Or; the work proposed a feature for an existing SSI implementation.
3. Or; the work met the prior two requirements for some close adjacent field.

We also defined the following hard exclusion requirements:

1. The work did not propose any SSI features or implementations.
2. Or; the work discussed a use case of SSI, that is, an implementation of existing SSI tooling for a specific scenario.
3. Or; the work was purely focused on cryptographic protocols without practical application.
4. Or; on point 3 from the inclusion requirements, the adjacent implementation did not have a clear analog to SSI (federations, PKI, authentication without attributes).

We collected all of the selected papers’ references and citations from the same database and repeated the filtering process. We continued these steps until the filtering converged.

This systematization resulted in two broad categories of SSI work: academic literature and industry implementations. We included extra consideration for the latter with additional sources of industry credential profiles as they have received more attention in development and deployment in recent years than purely academic discussions. Losing their contributions would not provide a complete image of the state-of-the-art.

### A METHODOLOGY

We started our systematization with the seed article *The Path to Self-Sovereign Identity by Allen [3]*. These articles were collected through ResearchRabbit. These articles were filtered and included if they met the following requirements:

1. The work proposed an SSI implementation.
2. Or; the work proposed a feature for an existing SSI implementation.
3. Or; the work met the prior two requirements for some close adjacent field.

We also defined the following hard exclusion requirements:

1. The work did not propose any SSI features or implementations.
2. Or; the work discussed a use case of SSI, that is, an implementation of existing SSI tooling for a specific scenario.
3. Or; the work was purely focused on cryptographic protocols without practical application.
4. Or; on point 3 from the inclusion requirements, the adjacent implementation did not have a clear analog to SSI (federations, PKI, authentication without attributes).

We collected all of the selected papers’ references and citations from the same database and repeated the filtering process. We continued these steps until the filtering converged.

This systematization resulted in two broad categories of SSI work: academic literature and industry implementations. We included extra consideration for the latter with additional sources of industry credential profiles as they have received more attention in development and deployment in recent years than purely academic discussions. Losing their contributions would not provide a complete image of the state-of-the-art.

### A METHODOLOGY

We started our systematization with the seed article *The Path to Self-Sovereign Identity by Allen [3]*. These articles were collected through ResearchRabbit. These articles were filtered and included if they met the following requirements:

1. The work proposed an SSI implementation.
2. Or; the work proposed a feature for an existing SSI implementation.
3. Or; the work met the prior two requirements for some close adjacent field.

We also defined the following hard exclusion requirements:

1. The work did not propose any SSI features or implementations.
2. Or; the work discussed a use case of SSI, that is, an implementation of existing SSI tooling for a specific scenario.
3. Or; the work was purely focused on cryptographic protocols without practical application.
4. Or; on point 3 from the inclusion requirements, the adjacent implementation did not have a clear analog to SSI (federations, PKI, authentication without attributes).

We collected all of the selected papers’ references and citations from the same database and repeated the filtering process. We continued these steps until the filtering converged.

This systematization resulted in two broad categories of SSI work: academic literature and industry implementations. We included extra consideration for the latter with additional sources of industry credential profiles as they have received more attention in development and deployment in recent years than purely academic discussions. Losing their contributions would not provide a complete image of the state-of-the-art.