Dezentrales Identitätsmanagement am Beispiel eines GAIA-X Projekts zur Simulation autonomer Fahrfunktionen

Felix Hoops

09.12.2024

Chair of Software Engineering for Business Information Systems (sebis)
Department of Computer Science
School of Computation, Information and Technology (CIT)
Technical University of Munich (TUM)
wwwmatthes.in.tum.de

About me

PhD

- Working full-time at the chair of Prof. Dr. Florian Matthes since June 2021
- Research interests evolved from Distributed Ledger Technology to Self-Sovereign Identity
- Involved in maintaining, updating, and teaching our “Blockchain-based Systems Engineering” lecture (ca. 500 students)

Prof. Dr. Florian Matthes
Head of sebis

Felix Hoops
Research Associate
& PhD Candidate

Current Industry Project

GAIA-X 4 Production, After-Sales and PLC - Across Automated Driving

The Gaia-X project aims to build a federated data infrastructure for Europe. As part of the GAIA-X 4 Future Mobility project family, this project focuses on the secure implementation of digital twins for the automotive sector in the context of automated driving through an open distributed data ecosystem (ODDE). Spanning the entire product lifecycle, these twins are envisioned to improve product verification, validation, and update strategies. more here ...

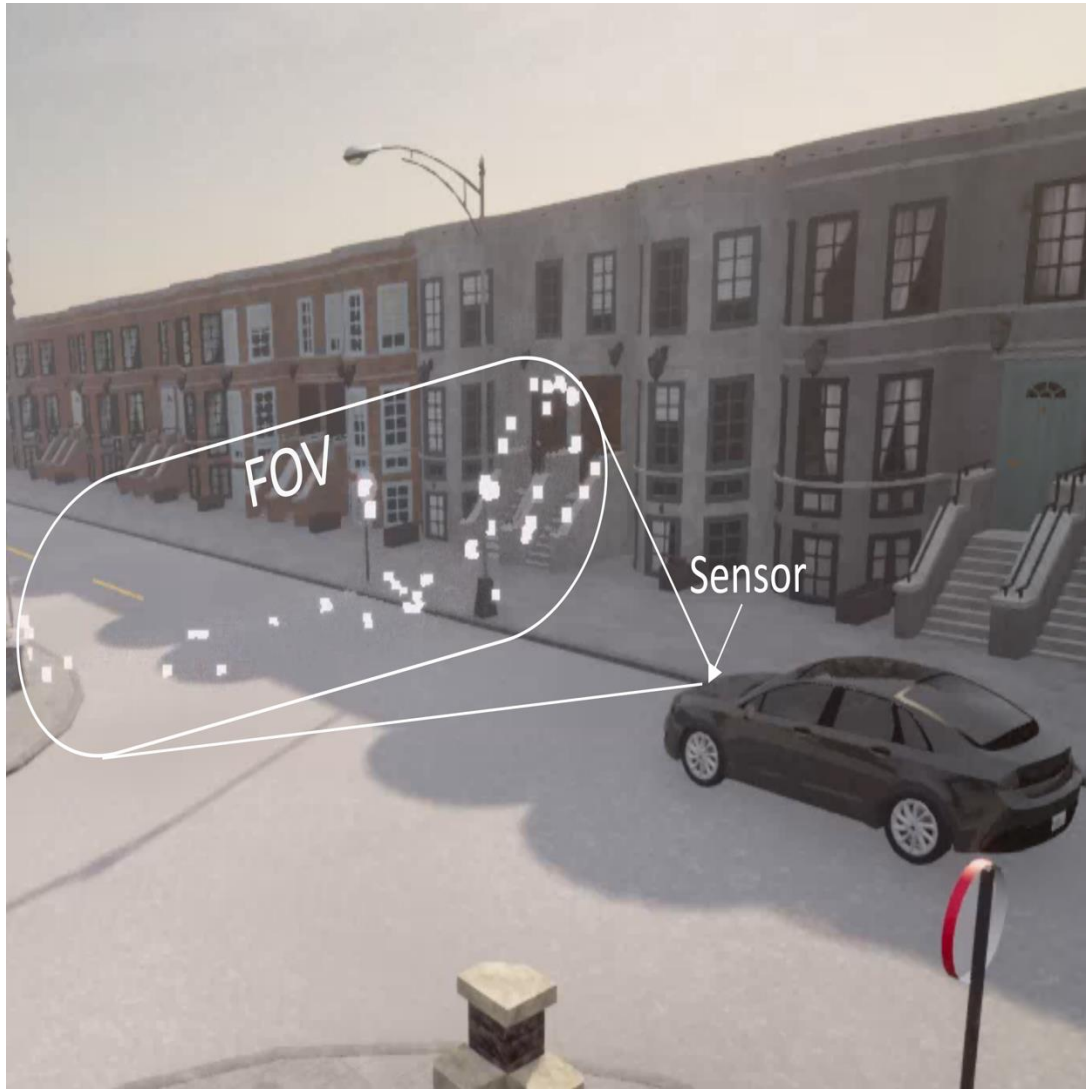
Outline

1. Motivation
2. Introduction to Self-Sovereign Identity
3. Issuing Verifiable Credentials for Dataspace Participants
4. Authentication & Authorization with Verifiable Credentials
5. Conclusion

Development of Autonomouse Driving Functions

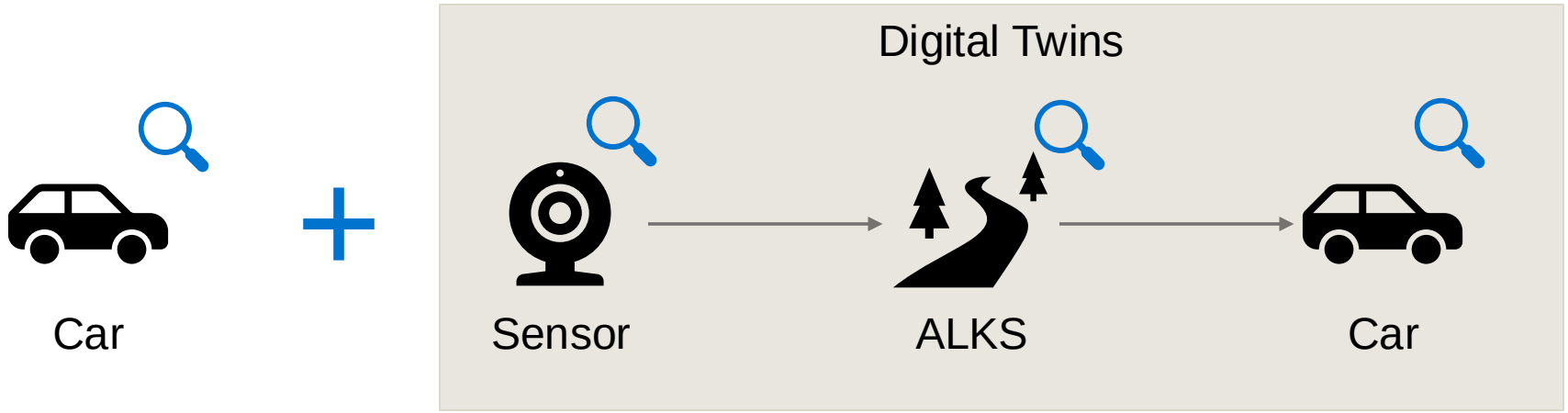
- Customers demand more and more autonomous driving functions.
  - e.g., Automatic Lane Keeping System
- These **complex systems** require lots of testing in different conditions. They are also **updated more frequently**.
- In Germany, new car models must be approved to be street legal.
  - Full test coverage required.

OEMs cannot get the number of track kilometers driven in specific weather conditions in reasonable time.



Scenario-based Testing

- Combine real-world driving tests with simulations.
- Simulations can be done for most parts and subassemblies.



OEMs need a lot of quality data and know-how to perform all necessary tests.

Gaia-X

- Gaia-X introduces a concept for connecting suppliers in a data space.
- It is decentralized to ensure neutrality and impartiality.

Gaia-X is a high-level specification that leaves a lot of things open.

"TOGETHER TOWARDS A FEDERATED AND SECURE DATA INFRASTRUCTURE"

Funded by the European Union
NextGenerationEU

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

Gaia-X 4 PLC-AAD

GAIA-X 4 Product Life Cycle – Across Automated Driving

- We design and build a decentralized marketplace for simulation data and services based on Gaia-X principles.
- Data assets are always sold with certificates:

Data
e.g., 3D Map

Certificate(s)
e.g., Certified Level of Detail

gaia-x future mobility
Gaia-X Lighthouse Project

gaia-x KI
Daten- und Dienste-Ökosystem für das Trainieren und Validieren von KI-Anwendungen in der Automobilindustrie

gaia-x AMS
ODD-konforme Bereitstellung von Routen für automatisierte Fahrzeuge und sichere Rettungskorridore für Einsatzfahrzeuge

gaia-x ROMS
Remote-Operation automatisierter Fahrzeuge und Fahrzeugflotten

gaia-x PLC-AAD
Product-Life-Cycle und Digital Twin beim automatisierten Fahren

gaia-x moveID
Sichere und dezentrale digitale Identitäten im Straßenverkehr

gaia-x AGEDA
Software-Architektur für Fahrzeuge, die datengetriebene, anpassbare Anwendungen by Design ermöglicht

A market place provides participants with **certified data**, fosters **common open standards**, and **promotes reuse**.

Funded by the European Union
NextGenerationEU

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

Certification

- Every data asset can be traced back to the processes and data assets used in its creation.
- Every process step and every input can be verified through certificates.
- We always know what is guaranteed by whom.

It becomes possible to prove data and process quality for any asset, building trust for cooperation and homologation.

How to identify companies, people, data assets, …?
For a decentralized system, we need decentralized identity.

- The need for **authentication and authorization** is universal.

Outline

1. Motivation
2. Introduction to Self-Sovereign Identity
3. Issuing Verifiable Credentials for Dataspace Participants
4. Authentication & Authorization with Verifiable Credentials
5. Conclusion

Motivating Self-Sovereign Identity (SSI)

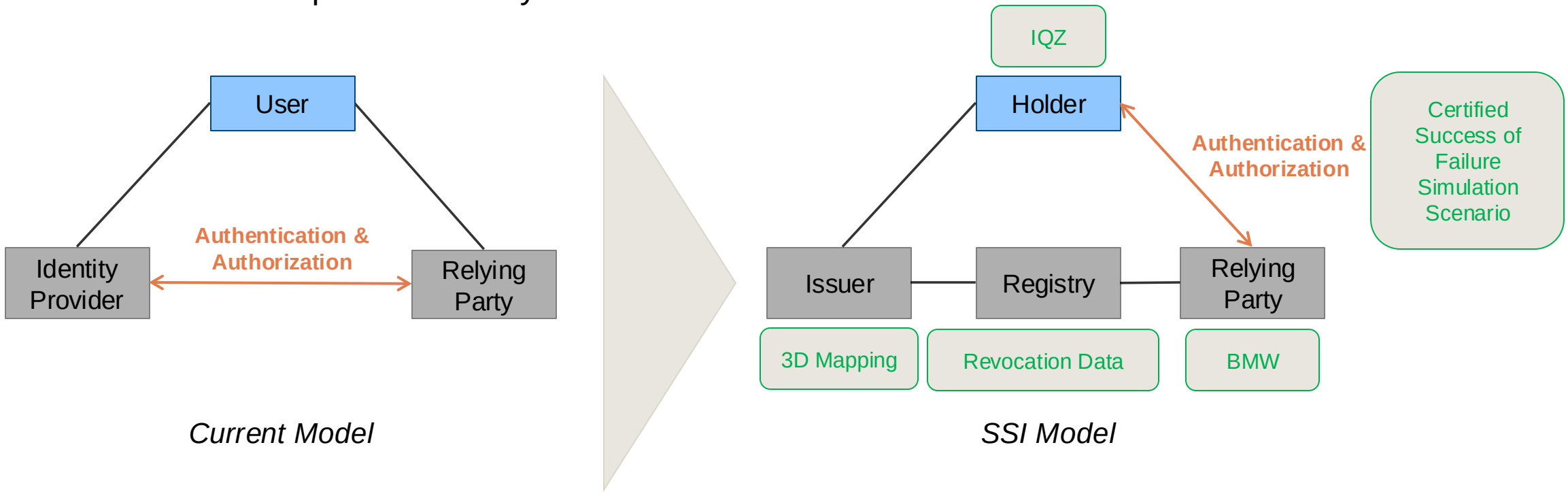
- Today's identity providers have immense amounts of power over us and metadata about us.

- Offline, you receive a state-issued ID card that you and only you control after issuance:
  - You decide when and to whom you identify yourself.
  - Everyone accepts your ID card.
  - No one can prevent you from physically presenting your ID card.
  - The act of physically presenting an ID card is not trackable by any third party.

We should strive to make **online identity better than paper-based identity across the board**. Digitalization should not just be about making a process faster but should also retain important properties of the old process.

Motivating SSI (cont.)

- **Self-Sovereign Identity** (SSI) is the term used to describe this target architecture.
- With SSI you can instantly create an account (i.e., identifier) without anyone being able to prevent it.
  - You alone control that account, which means no one can shut it down or take it over.
  - The account is compatible with any online service.

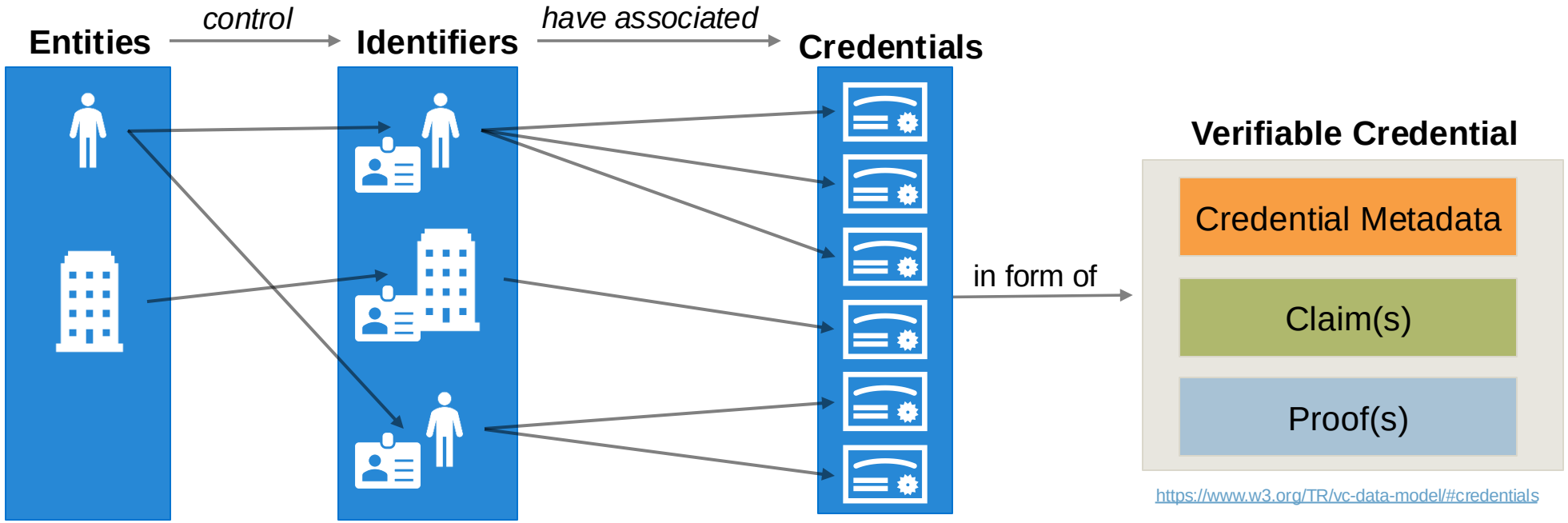


For industrial cooperation, this means a proof about an asset can be made without being dependent on the original issuer.

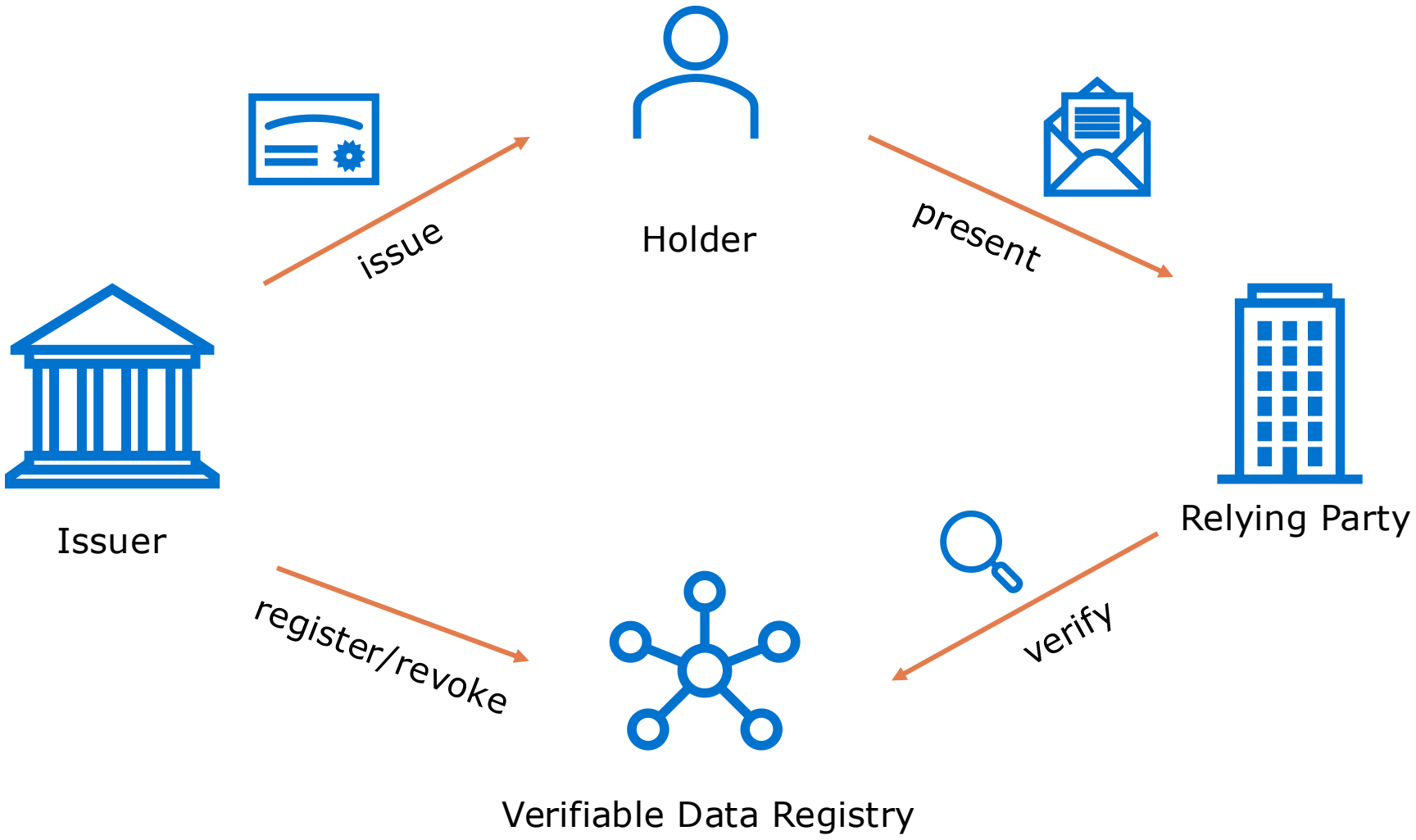
Rough SSI Definition

- Self-Sovereign Identity (SSI) is still a very new approach, and definitions vary slightly.

- Rough definition:
  - It is a model in which entities are represented by digital identities, and every entity has **sole ownership** over the ability to control their identity data.
  - An identity can be seen as an account consisting of a pseudonymous identifier and an arbitrary number of attributes that are confirmed by some witness.
  - It is a paradigm that emphasizes privacy. Entities can decide what attributes to present to whom.



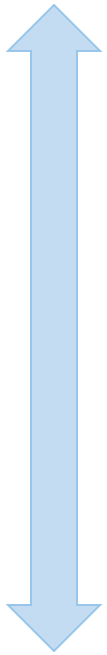
Lifecycle of a Verifiable Credential



SSI Stack

Preukschat and Reed[1] propose a four-layer technology stack for SSI that goes from technical trust to human trust:

Human Trust



Technical Trust

Layer 4	**Governance**	Solving standardization issues, legal issues, and human trust.
Layer 3	**Credentials**	Defining Credential format and how credentials are exchanged via the interfaces.
Layer 2	**Interfaces**	Orchestrating secure communication between parties.
Layer 1	**Identifiers**	Managing identifiers and their associated keys.

Adopting SSI for a use case means that challenges on all of these layers must be addressed. It is by no means only a technical problem.

[1] https://livebook.manning.com/book/self-sovereign-identity/chapter-5/v-8/8

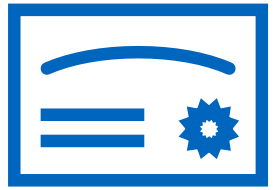
Identifier and Credential Standards

The World Wide Web Consortium (W3C), which creates guidelines and standards for the Internet, developed two core specifications for SSI:

Word Wide Web Consortium

- **Decentralized Identifier** (DID): A DID is a globally unique identifier for every entity in the SSI ecosystem. It does not need the use of a centralized authority. An example of a DID is *did:example:123456abcdef*.

- **Verifiable Credential** (VC): A VC is a means of making verifiable claims about an identity. This can be a government authority stating that a DID belongs to a certain Person's Name, Date of Birth, etc. But it could also be an entry pass for a building. Or a diploma.



Decentralized Identifier Examples

A DID has the following structure:

did:ethr:0xb9c5714089478a327f09197987f16f9e5d936e8a		
Scheme This part is static.	**Method** Usually short and identifies one specific publicly documented DID method. This one provides high decentralization and a low barrier to entry.	**Method-Specific Identifier** Arbitrarily long identifier. In this case, it is an Ethereum account address. Creating one is as easy as creating an Ethereum account.
did:web:tum.de		
Scheme This part is static.	**Method** Usually short and identifies one specific publicly documented DID method. This one enables easy adoption for institutions via an existing web presence.	**Method-Specific Identifier** Arbitrarily long identifier. In this case, it is a domain hosting a DID document at default relative path “/.well-known/did.json”.

DIDs are very flexible due to their reliance on **custom DID methods**. A custom DID method defines how to

- **create** an identifier
- **retrieve** information about the identifier (i.e., resolving to a DID document)
- **update** and **delete** information about the identifier (optional)

Resolution of Decentralized Identifiers to DID Documents

- Decentralized Identifiers (DIDs) are unique identifiers created and controlled by the individual, organization, or device they identify, rather than being issued and controlled by a central authority.

DID Resolution:

- A DID **resolver** is software that resolves a DID into a DID document by following a pre-defined algorithm specific to the DID's method.
- The resolution process may depend on external data sources like a blockchain.

DID Document:

- A DID document is a document that is accessible to anyone by resolving a DID and contains information related to a specific decentralized identifier, such as the currently used public key and usage conditions.

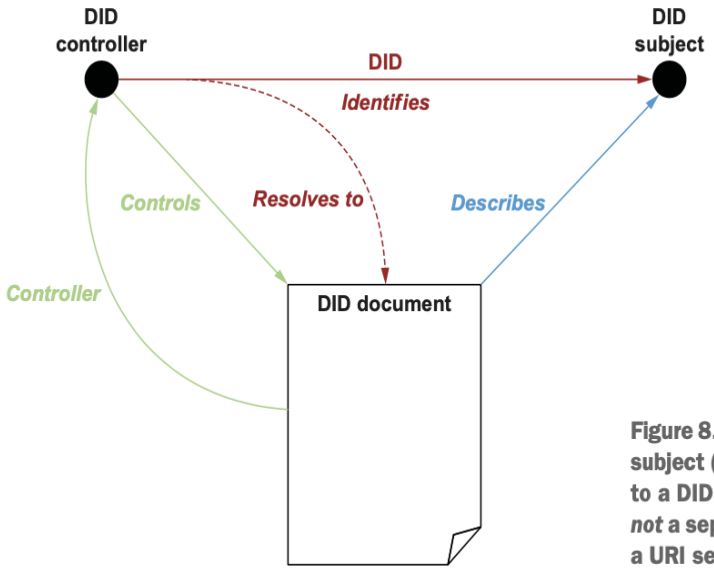


Figure 8.29 A DID always identifies a DID subject (whatever it may be) and resolves to a DID document. The DID document is *not* a separate resource and does *not* have a URI separate from the DID.

Image source: https://www.w3.org/TR/did-core/

The main advantage of DIDs over just using an asymmetric key pair is the flexibility provided by the DID document. Because of it, keys can be updated, and additional meta information can be included in a standardized way.

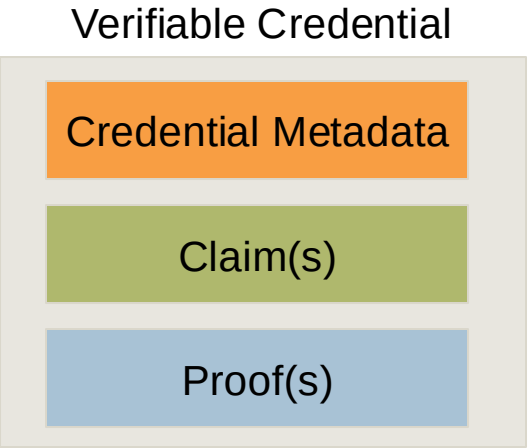
Verifiable Credentials

- A Verifiable Credential (VC) is a digital representation of a set of claims that a third party can verify without needing a trusted intermediary. The key elements of a verifiable credential are:

  Credential Metadata: Issuer as well as information about the format and purpose of the credential.

  Credential Claims: Credential Subject and claims about the subject made by the issuer.

  Credential Proof: A digital signature produced by the issuer that enables the credential to be verified by a third party.



https://www.w3.org/TR/vc-data-model/#credentials

VCs can theoretically work with different types of identity. However, DIDs are the preferred solution to identify issuer, subject, and potential other involved entities.

Verifiable Data Registry

- The identity data itself is stored in each user's wallet.

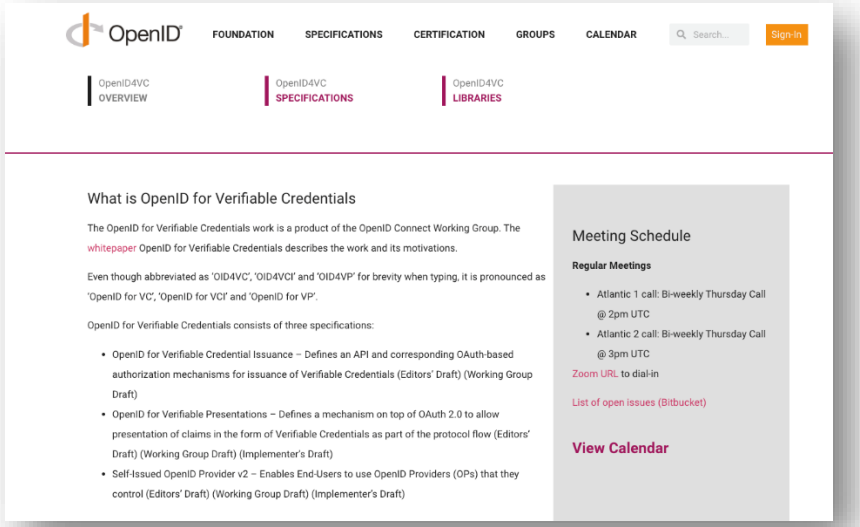
- Some public metadata is required for specific functionalities.
  1. **Status Lists**
     - Support the revocation or suspension of credentials.
  2. **DID Documents**
     - Depends on the DID method used.
  3. **Trusted Issuer List**
     - One or more lists created by trust anchors that designate trustworthy issuers.
  4. **Logging** (optional)
     - Provides auditability (e.g., to detect fraudulent activity).

- Different implementations can be used for each of these functionalities. Thus, the Verifiable Data Registry is a concept and not necessarily one single infrastructure.

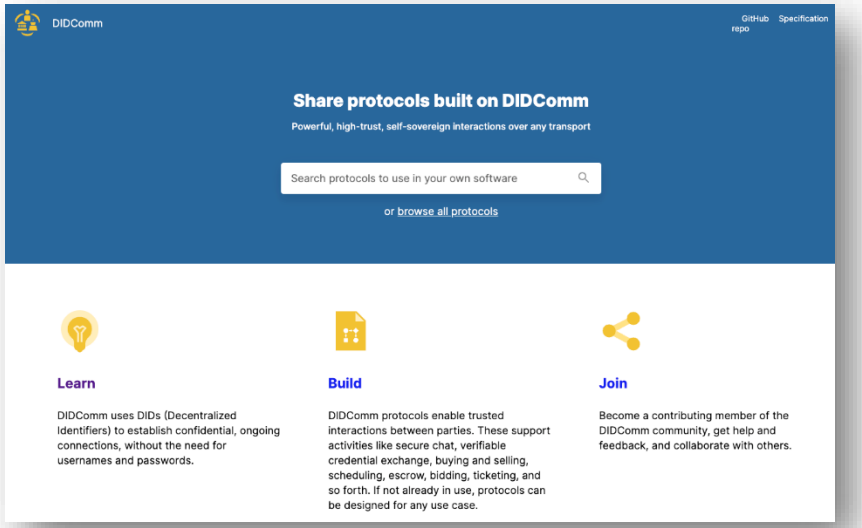
The publicly readable Verifiable Data Registry should never expose private information (e.g., credentials themselves). **Data stored there is typically minimal, such as serial numbers or hashes of credentials.** Exact data and data structure are implementation-specific.

Protocols for Verifiable Credential Exchange

- Issuance and presentation processes need standardized protocols.
- There are two fundamental approaches:
  - Client-server
    - Web-based exchange via HTTP
    - *Example*: OID4VC
  - Peer-2-peer
    - Direct and symmetric message exchange
    - *Example*: DIDComm



https://openid.net/sg/openid4vc/



https://didcomm.org/

Protocols for Verifiable Credential exchange have been developed, and there are very different philosophies behind them. For truly interoperable SSI, everyone must support the same set of protocols.

SSI Criticism

Google, Apple, and Mozilla filed official objections to the acceptance of the W3C DID 1.0 specification in September 2021. So, what was the reason for it?

Four main reasons were given:

- The DID 1.0 specification standardizes DIDs in general but does not standardize any specific DID methods.
- The DID 1.0 specification encourages many different DID methods instead of just a few, which might limit interoperability.
- The DID 1.0 specification does not prohibit centralized DID methods.
- The DID 1.0 specification promotes the use of blockchains, about which environmental concerns have been raised.

But…

- Currently, there is no serious alternative to DIDs.
- Diversification means “plug and play,” ensuring interoperability and easier adoption for existing systems.
- Besides, all the objecting companies have a significant interest in staying a federated identity provider.

There was also some criticism of SSI in general from tech influencers who argue that most SSI use cases can be solved more easily using existing central authority database systems. While that is generally true, there are arguably benefits in researching and designing systems that do not needlessly centralize control and data. Technical criticism is rare.

Pause for Questions!

Outline

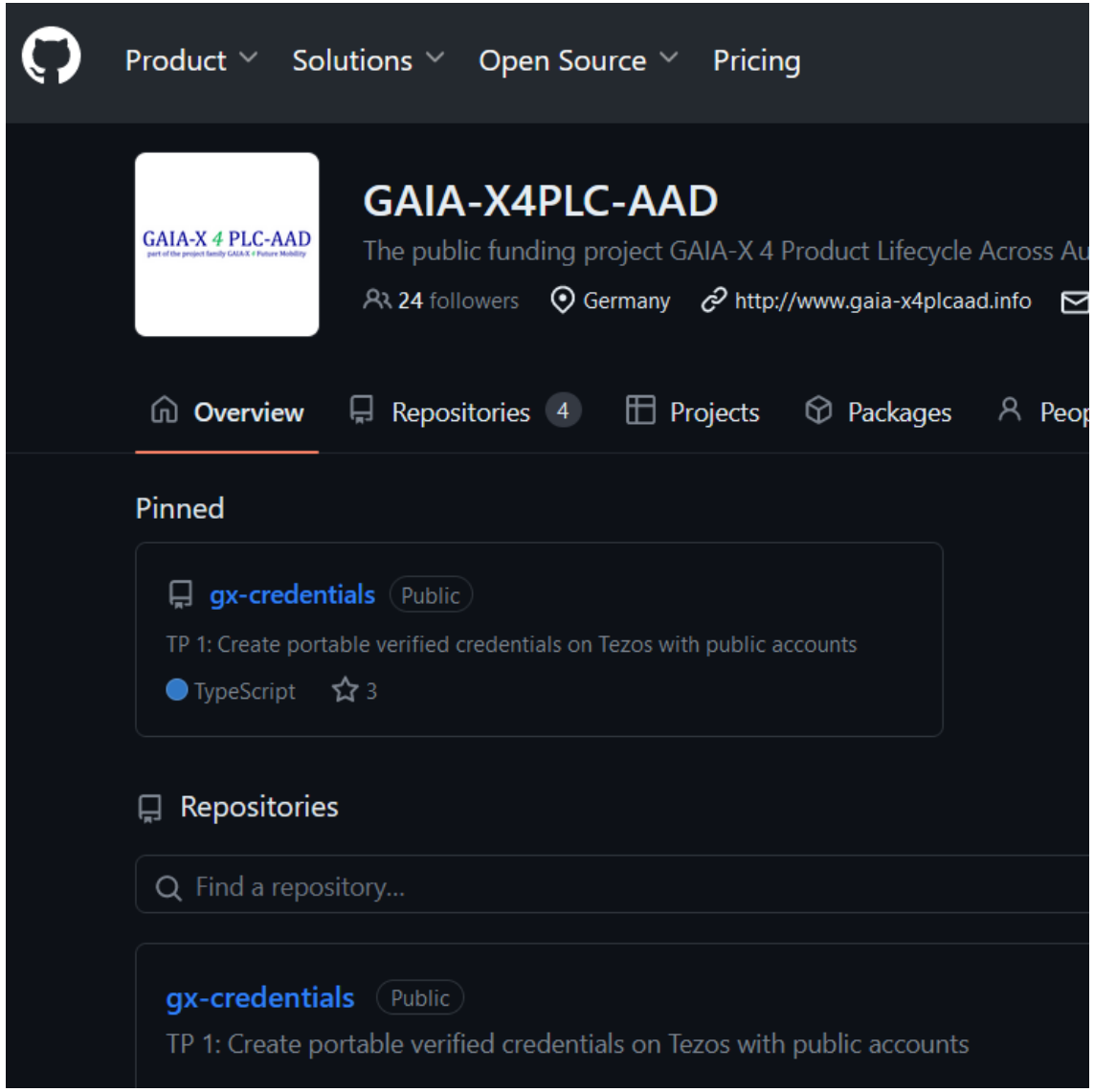
1. Challenges of Developing Autonomous Driving Systems
2. Introduction to Self-Sovereign Identity
3. Issuing Verifiable Credentials for Dataspace Participants
4. Authentication & Authorization with Verifiable Credentials
5. Outlook

What Is a Data Space?

- We consider a data space as a set of services open to a common set of entities.
- Gaia-X mandates that participants (i.e., companies) are identified by verifiable credentials called “participant credentials”.
- We also issue similar credentials to employees of these companies, since humans ultimately interact with most systems.
- Individually identifying participants is technically possible but makes little sense operationally.

There needs to be at least one trust anchor (comparable to TLS root certificates) in a data space.

Verifiable Credential Issuance for Organizations

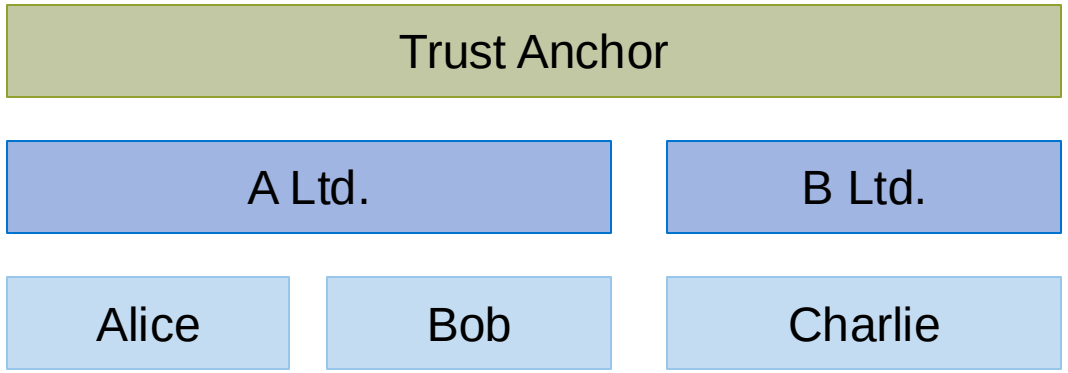


Main goals:

- Provide employee identification for a dataspace.
- No keys on live servers. All signatures happen on user devices.

One webservice that can be run by anyone who wants to offer their service as a trust anchor.

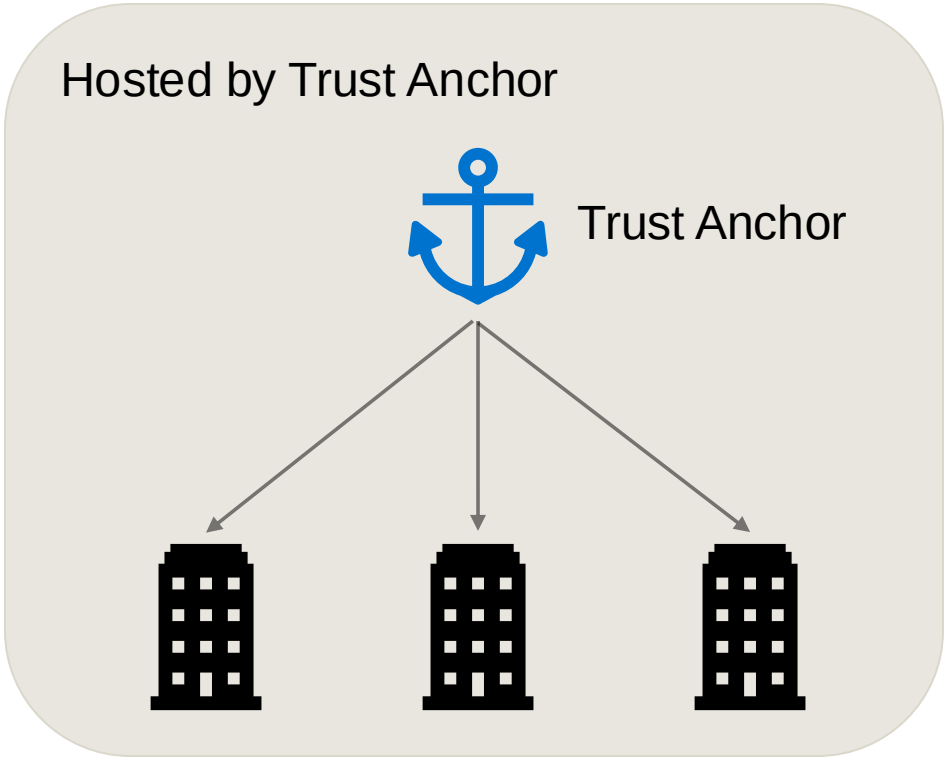
Data space services can then make use of these credentials.

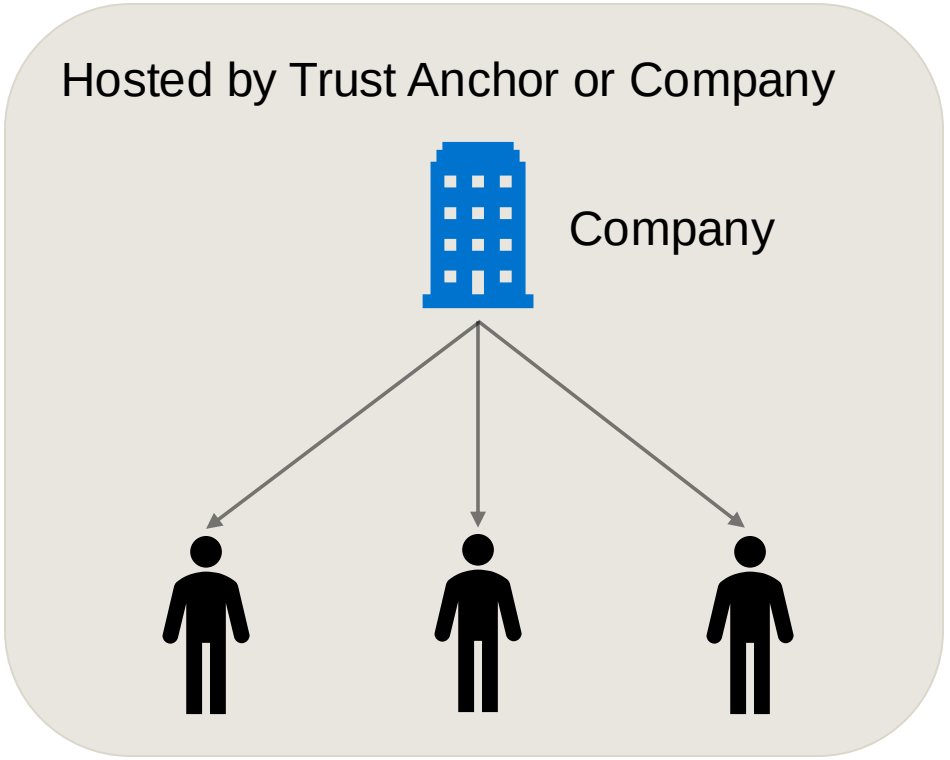


Verifiable Credential Issuance for Organizations (cont.)

- The two issuance scenarios are close to identical from a technical perspective.

- Companies do not have to host their own issuance software, but they can if they want increased privacy.
  - Companies can also use the trust anchor's web app.
  - This is possible without compromising security because no keys are in the live system.





Low barrier to entry for companies.

Altme Wallet & The Power of Leveraging Existing Communities

- SSI requires wallet software that can handle Verifiable Credentials, not just keys.
  - receiving, storing, and presenting

- Crypto wallets are usually older and more mature, but do not support Verifiable Credentials.

- Altme Wallet is the perfect exception:
  - relatively mature
  - full support for Crypto & SSI
  - free and open-source software
  - roots in Tezos blockchain community
  - very active developer

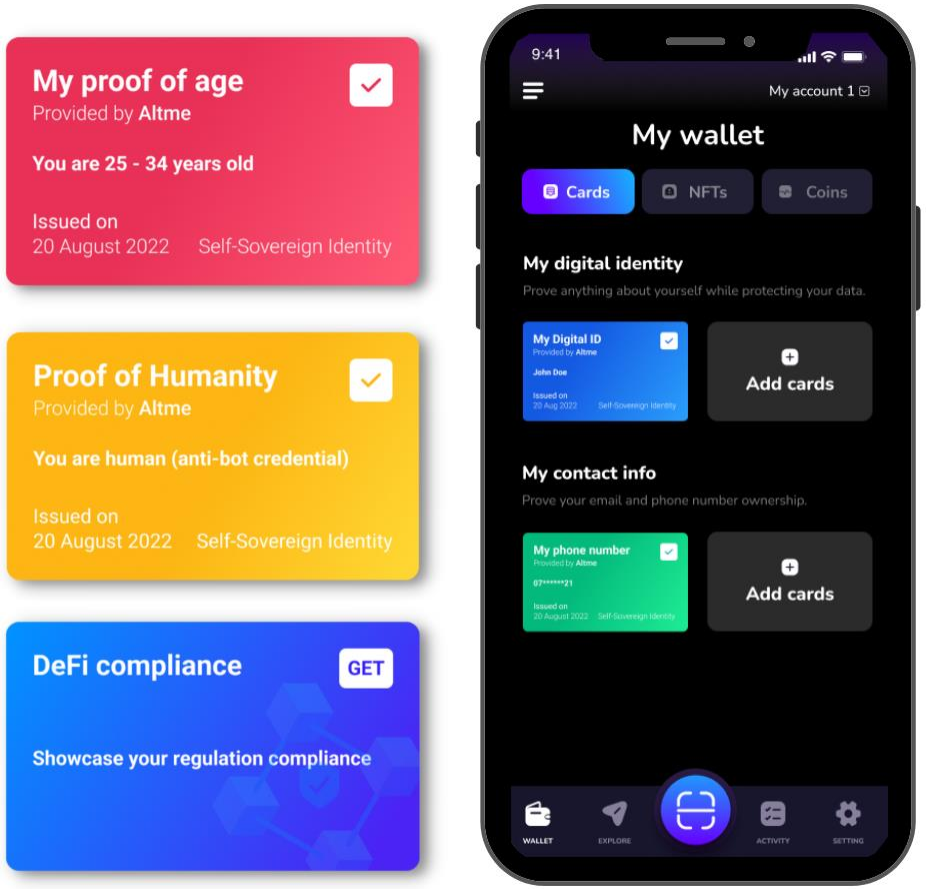
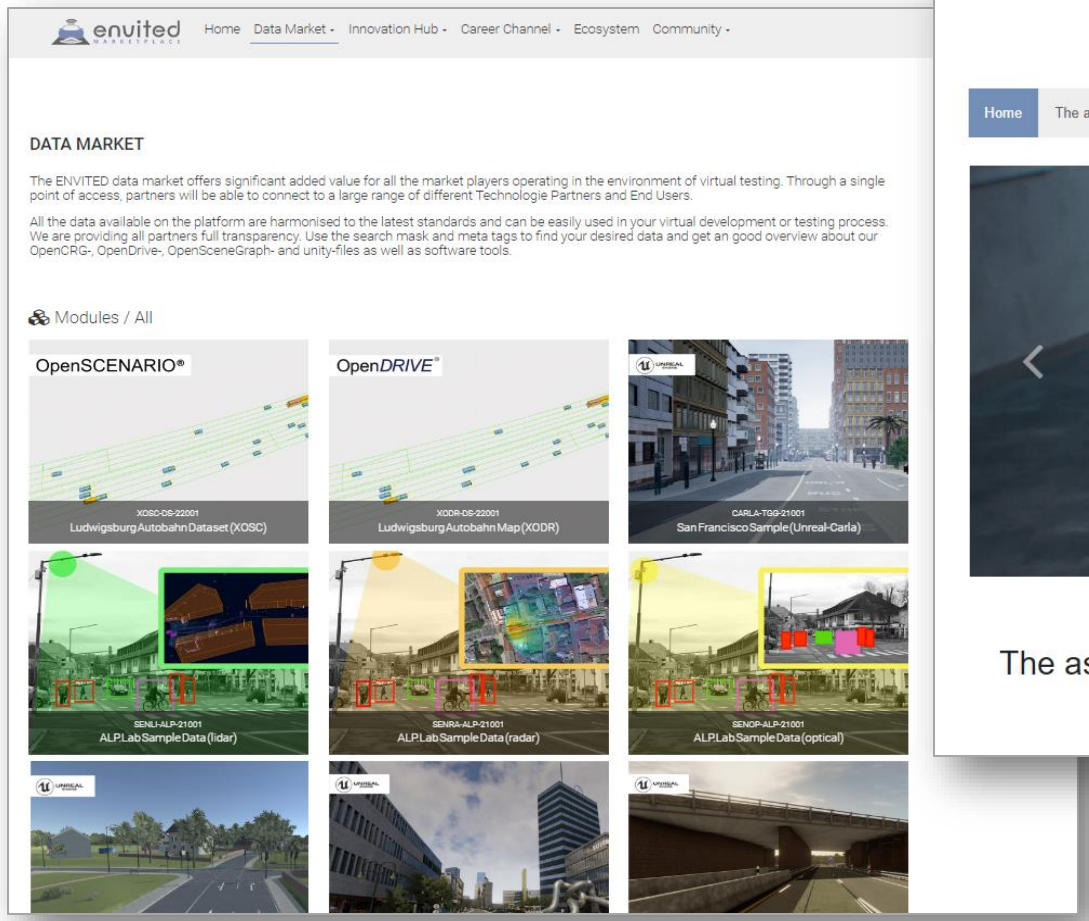
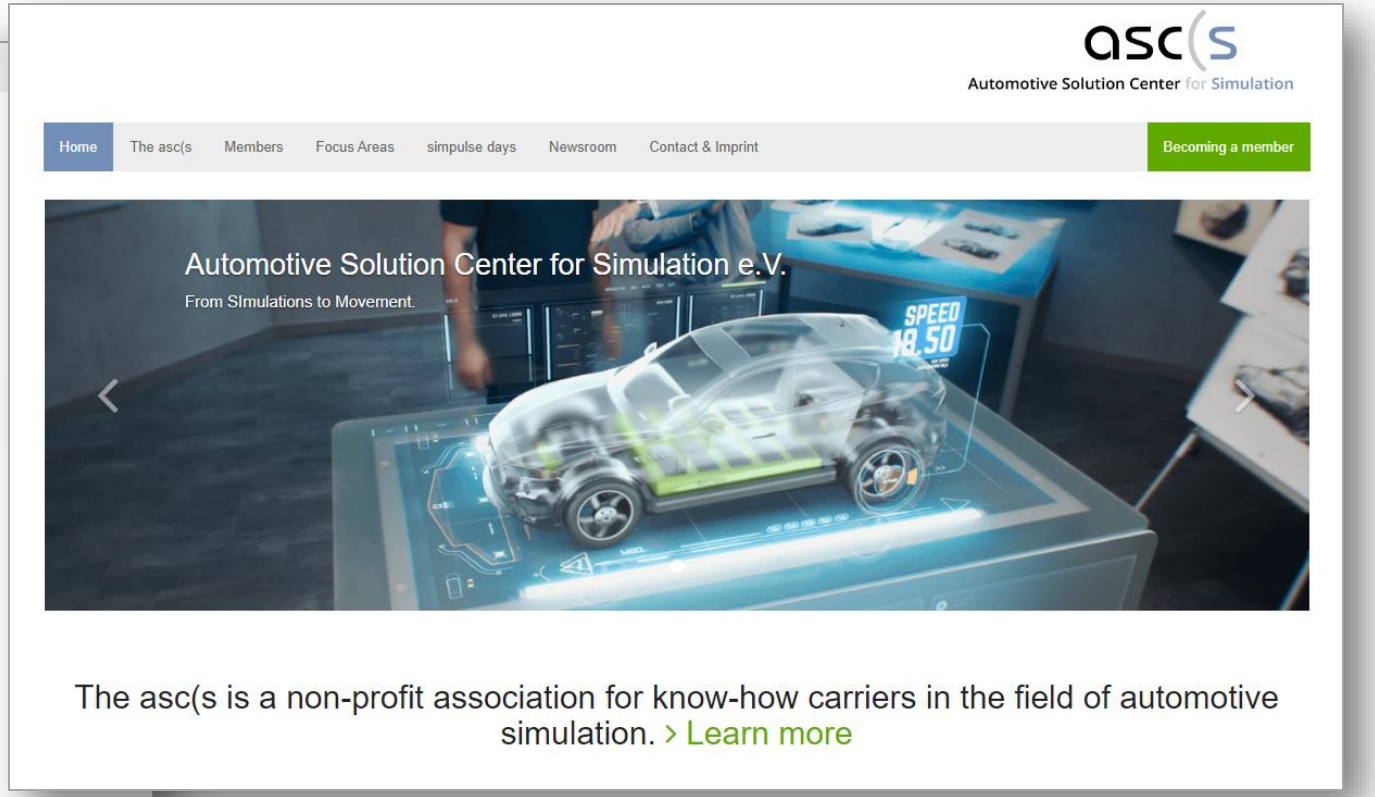


Image source: https://altme.io/

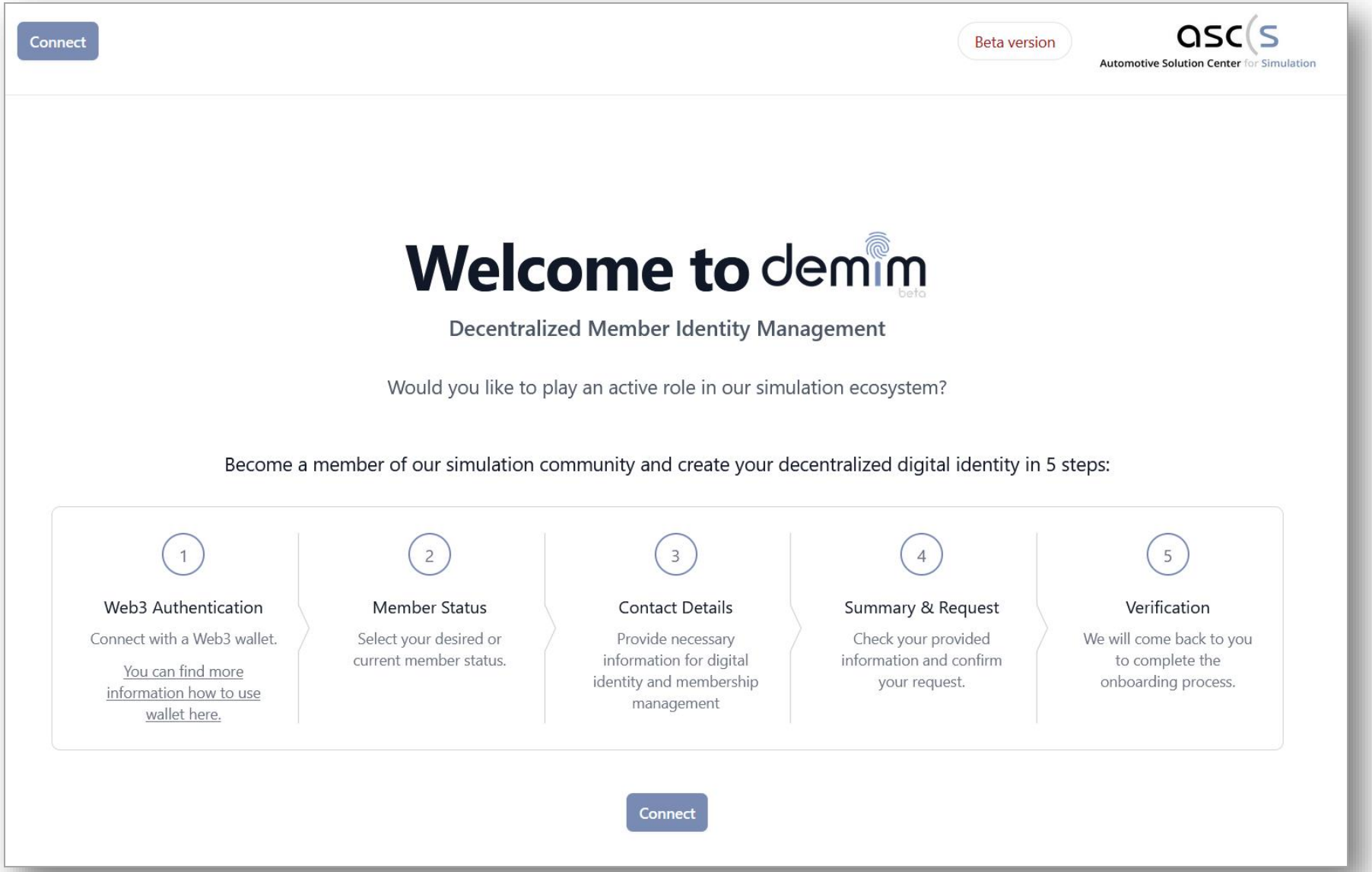
Choosing Trust Anchors

- Very difficult governance question.
- One option: existing industry associations.





Running Opt-in Beta



https://identity.ascs.digital

Example of a Participant Credential

- Simplified example of a company's participant credential:

- Verifiable Credentials take the form of a JSON (or technically JSON-LD) document and typically contain:
  - Context
  - Issuer
  - Issuance timestamp
  - Expiry timestamp (optional)
  - Type
  - Subject
  - Subject identity attributes
  - Cryptographic proof to ensure the integrity and authenticity of the VC

- Full context and proof section are omitted.

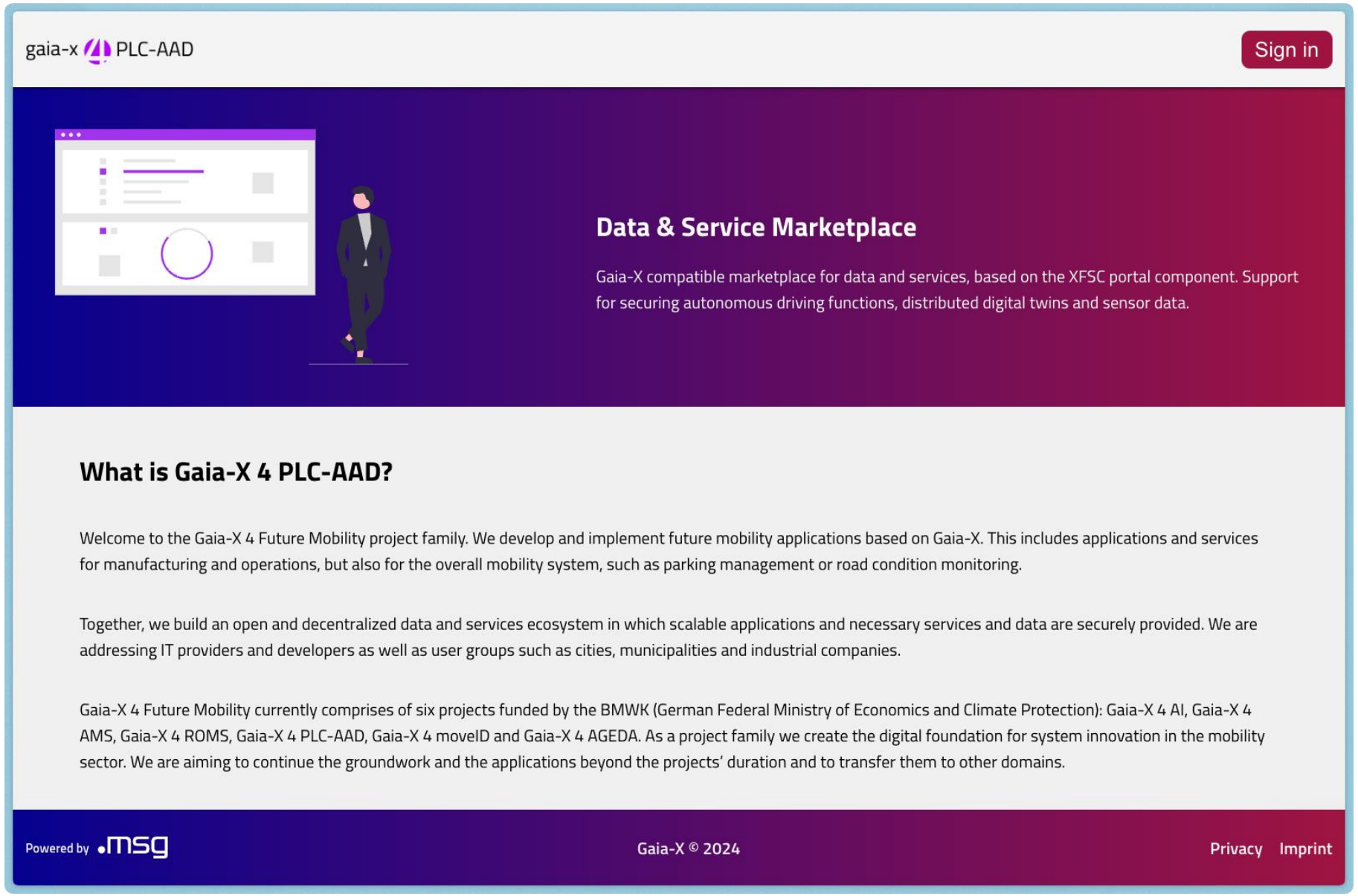

```
{
    "@context": [
        "https://www.w3.org/2018/credentials/v1",
        "https://schema.ascs.digital/AscsMemberCredential/v1"
    ],
    "type": [
        "VerifiableCredential",
        "AscsMemberCredential"
    ],
    "issuanceDate": "2023-11-22T17:14:33Z",
    "expirationDate": "2102-09-15T17:14:33Z",
    "id": "urn:uuid:576fbefb-35e8-4b71-bb1a-53d1803c86de",
    "issuer": {
        "id": "did:pkh:tezos:NetXnHfVqm9iesp:tz1ggujjYjA7oYoaZBzTg1tYSXn3VMjcgDuv",
        "type": "AscsIssuer",
        "name": "Automotive Solution Center for Simulation e.V.",
        "url": "https://identity.ascs.digital/"
    },
    "credentialSubject": {
        "id": "did:pkh:tezos:NetXnHfVqm9iesp:tz1bpeJArd7apJyTUryfXH1SD6w8GL6Gwhj8",
        "type": "AscsMember",
        "name": "Testcompany GmbH",
        "url": "https://test.de/",
        "address": {
            "type": "PostalAddress",
            "streetAddress": "Teststraße 1",
            "postalCode": "12345",
            "addressLocality": "Munich",
            "addressCountry": "DE"
        },
        "vatID": "DE123456789",
        "isAscsMember": true,
        "isEnvitedMember": true,
        "privacyPolicy": "https://media.ascs.digital/terms/ascs_privacy_policy_2020-07-08.pdf#cidv1",
        "articlesOfAssociation": "https://media.ascs.digital/terms/ascs_articles_of_association_2021-09-17.pdf#cidv1",
        "contributionRules": "https://media.ascs.digital/terms/ascs_contribution_rules_2020-07-08.pdf#cidv1"
    }
}
```

Pause for Questions!

Outline

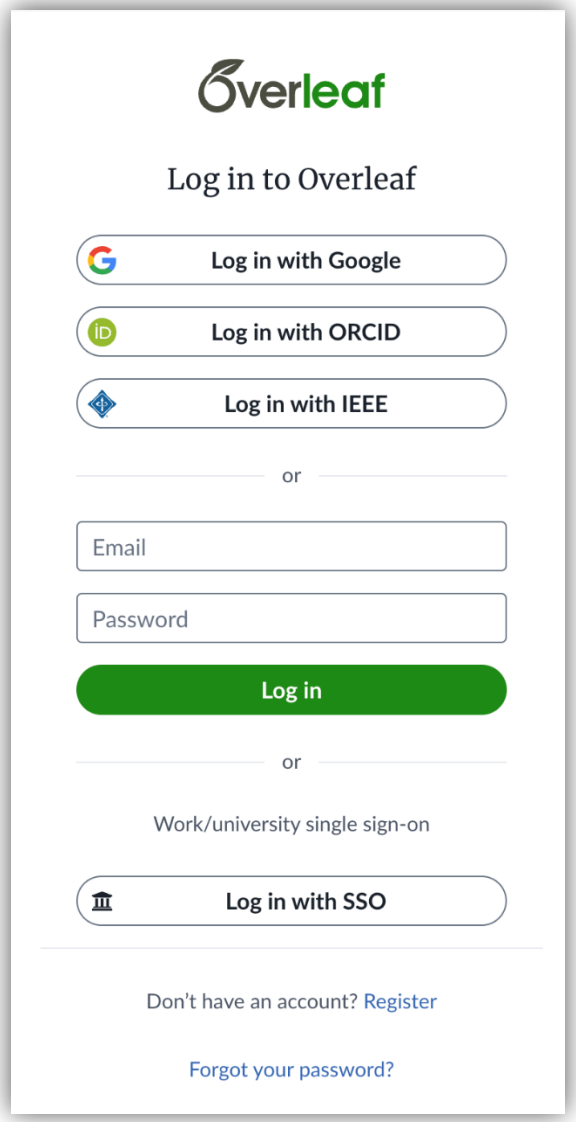
1. Motivation
2. Introduction to Self-Sovereign Identity
3. Issuing Verifiable Credentials for Dataspace Participants
4. Authentication & Authorization with Verifiable Credentials
5. Conclusion

Accessing Data Space Services



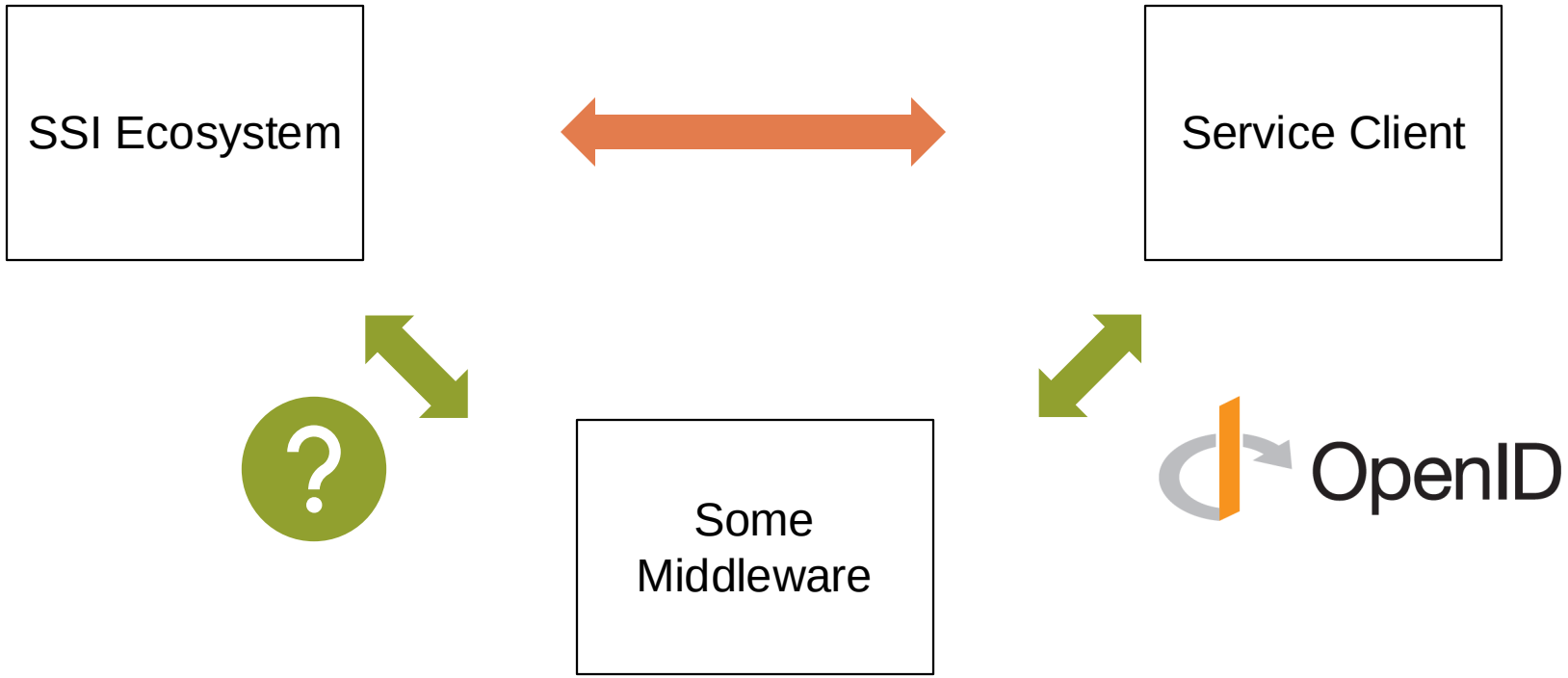
https://portal.gxfs.gx4fm.org/

Adopting SSI-based Authentication & Authorization



- Companies must make services accessible with the Verifiable Credentials.
- But, **adopting SSI as a verifier is not trivial.**
  - verification itself
  - use case dependent constraints
  - exchange of VPs
- While some services need more complicated authorization processes, **most services just need a sign-in** that is simple:
  - navigate to sign-in
  - scan QR code with phone (or press link if on phone)
  - choose (preselected) VC(s) to present and confirm on phone
  - page refreshes with session

The Idea of Bridging to Established Systems



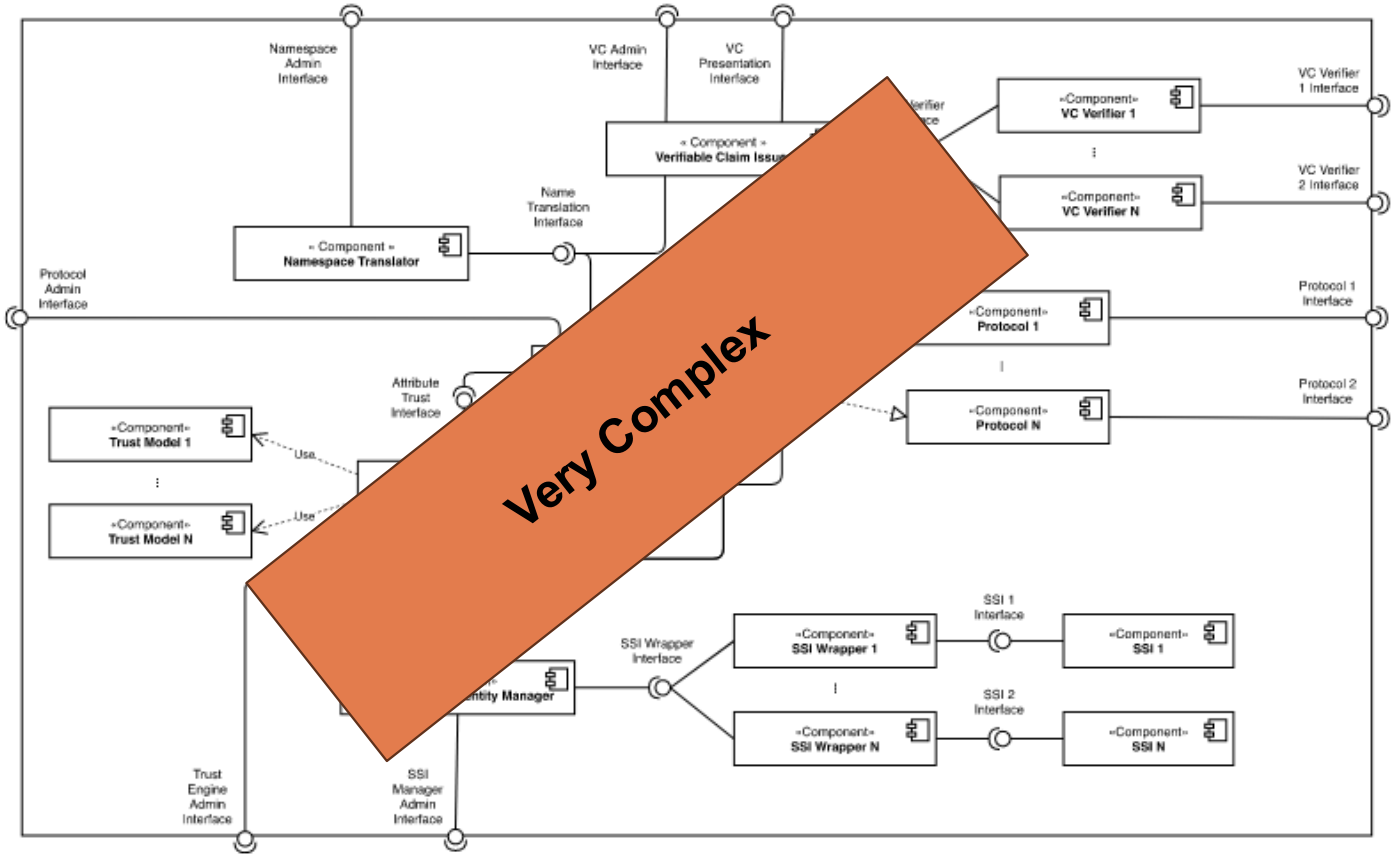
A general survey of the state of SSI from 2022 by Schardong et al. notes that protocol integration into established IAMs is vital to pave the adoption for SSI.

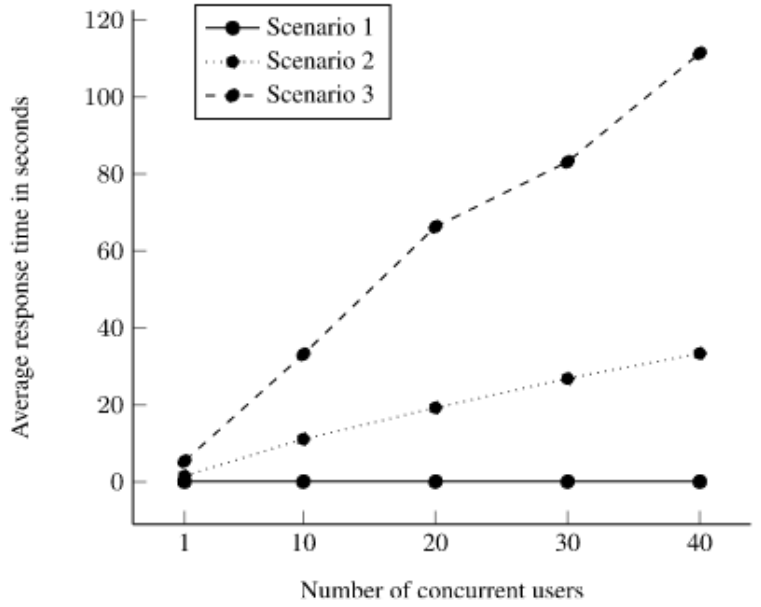
But…

A survey by Kuperberg et al. (2022) looking at SSI integrations for established IAM protocols identifies only **seven relevant candidates**.
The majority are commercial, and only **2 of them are available as open-source software**.

Complexity of Existing IAM Bridges

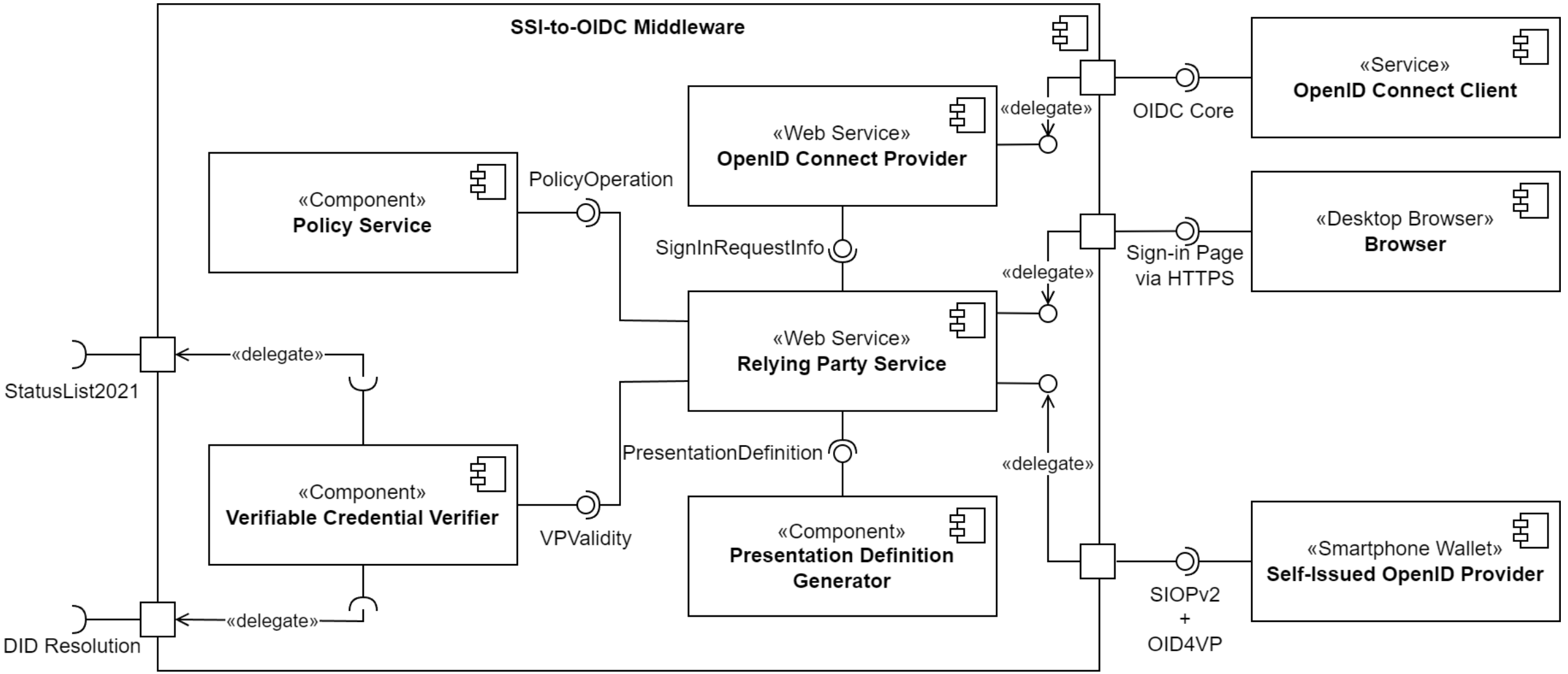
Grüner et al. (2019, 2021) have created one of the most extensive works so far. However, there is still no adoption.



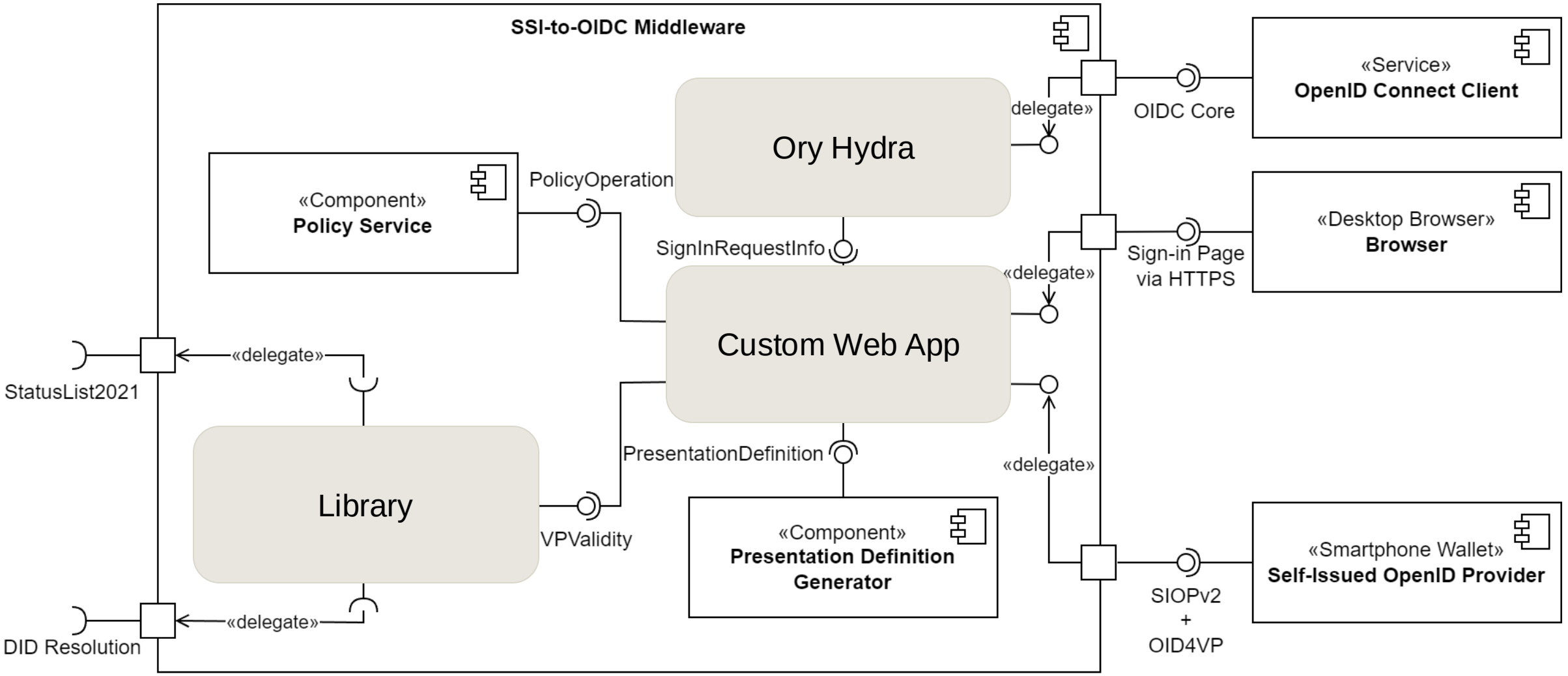


Grüner, A., Mühle, A., & Meinel, C. (2021). ATIB: Design and evaluation of an architecture for brokered self-sovereign identity integration and trust-enhancing attribute aggregation for service provider. *IEEE Access*, *9*, 138553-138570.

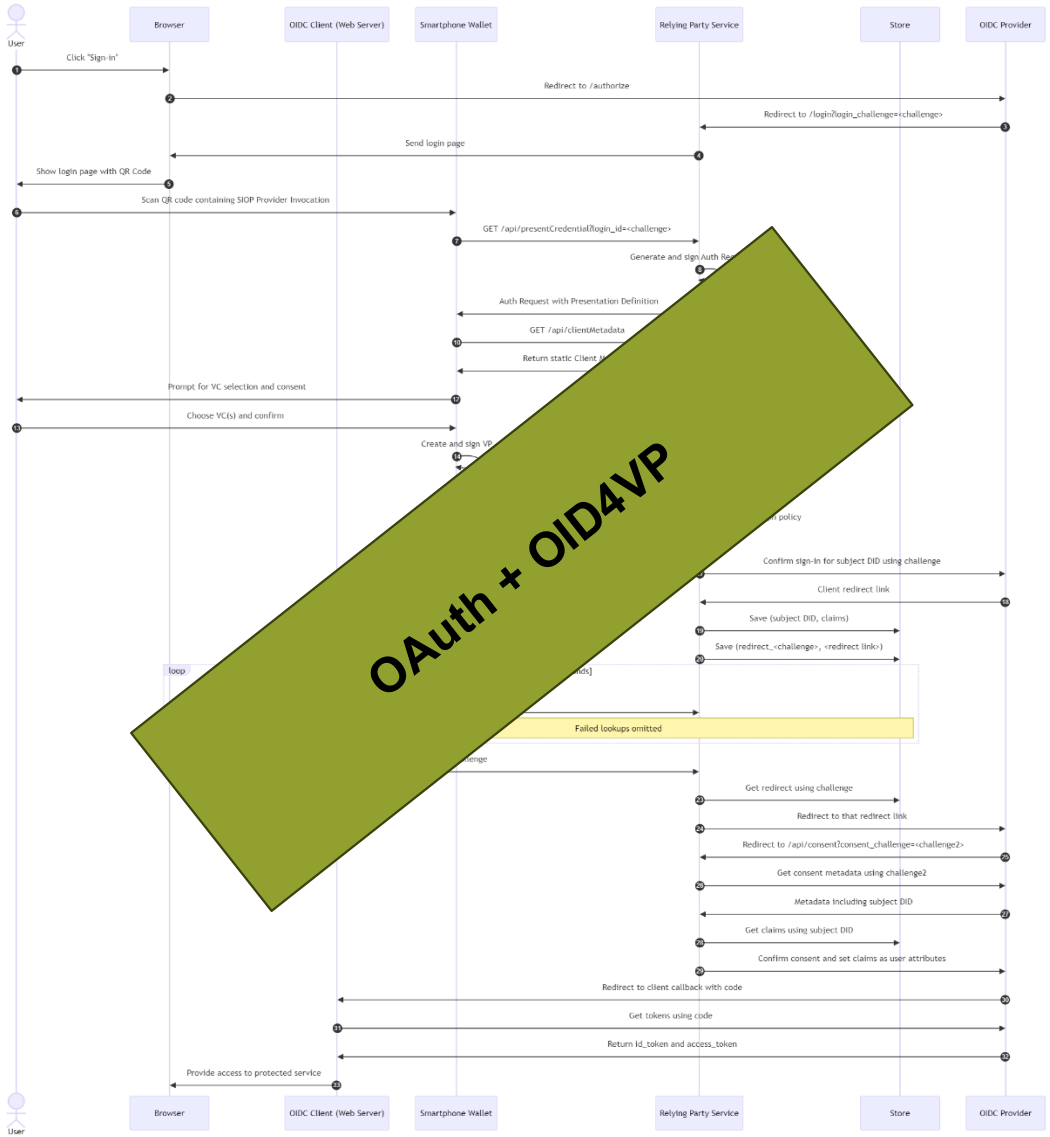
Our Design



Our Design



The Protocol Flow

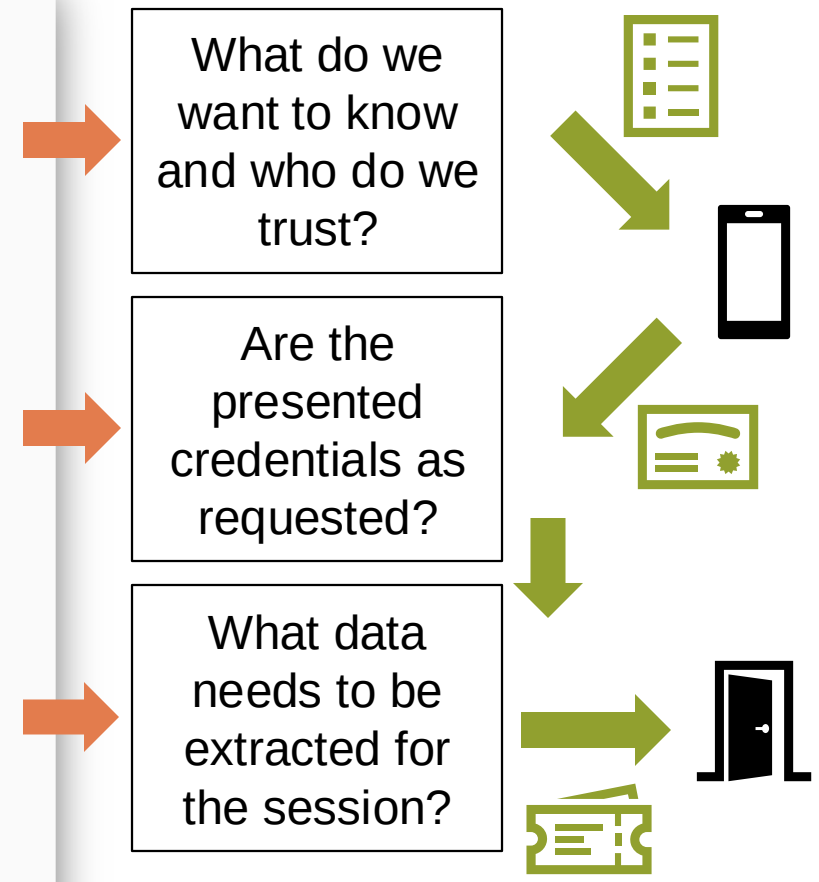


- Modified OAuth 2.0 authorization code flow.
- The login challenge from OIDC (in shorter form) is used throughout the entire process.
  - We pass it to the phone by encoding it as part of the QR code.
- Sign-in page actively polls for completion.
- Consent stage of the authorization code flow is skipped by confirming it automatically.
  - Users already give consent by sharing VP.
- No refresh tokens are issued.
  - The **bridge remains stateless in the grand scheme.**

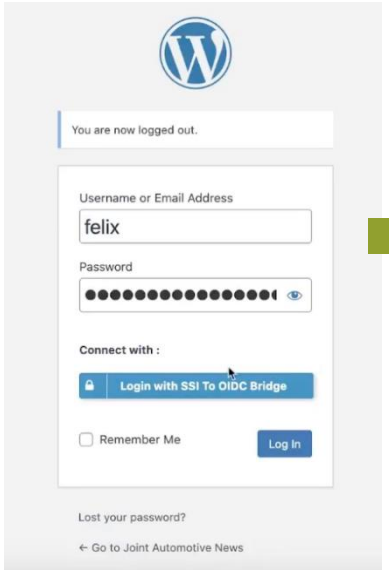
Login Policy

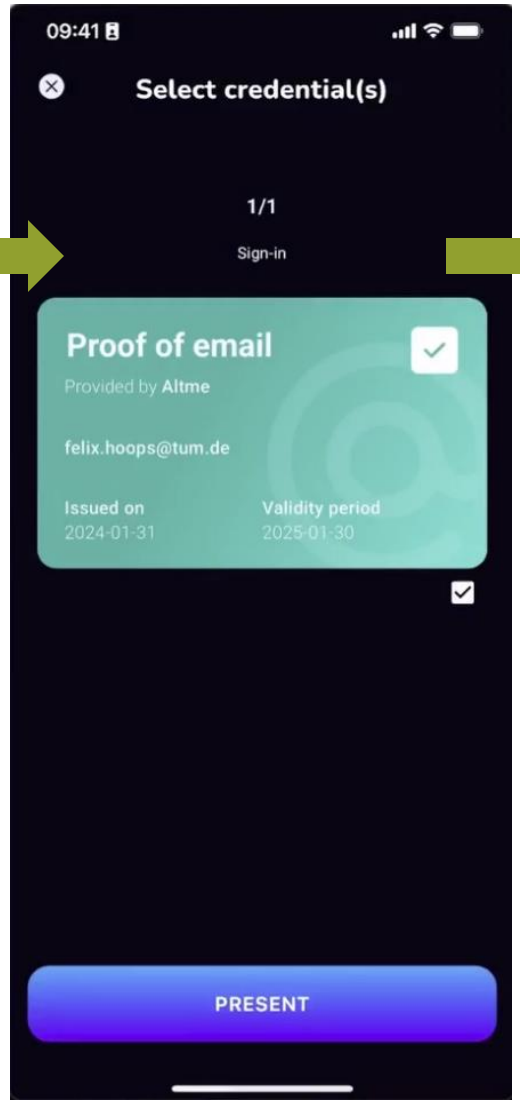
To not repeat what prior works did, we wanted one central configuration file, that combines trust, required values, and value mapping. For that, we had to develop a custom format.

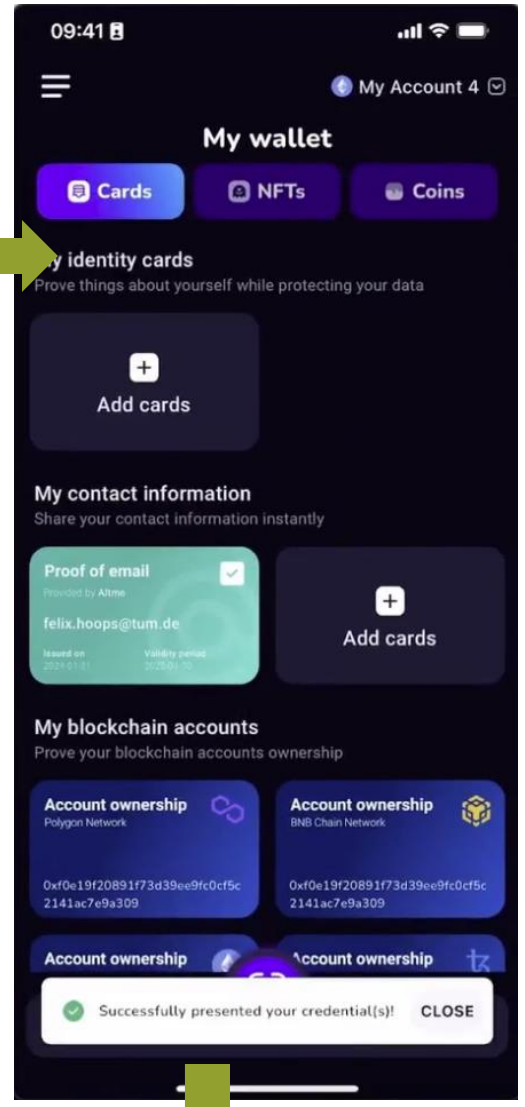
```
[{
  "credentialID": "expected_credential_for_email",
  "patterns": [{
    "issuer": "did:example:123",
    "claims": [{
        "claimPath": "$.credentialSubject.e_mail",
        "newPath": "$.email",
        "token": "id_token"
    }]
  },
  {
    "issuer": "did:example:456",
    "claims": [{
        "claimPath": "$.credentialSubject.email",
        "token": "id_token"
    }]
  }]
}]
```



End-User Experience







Usually takes about 6 seconds
(technical bound much lower)

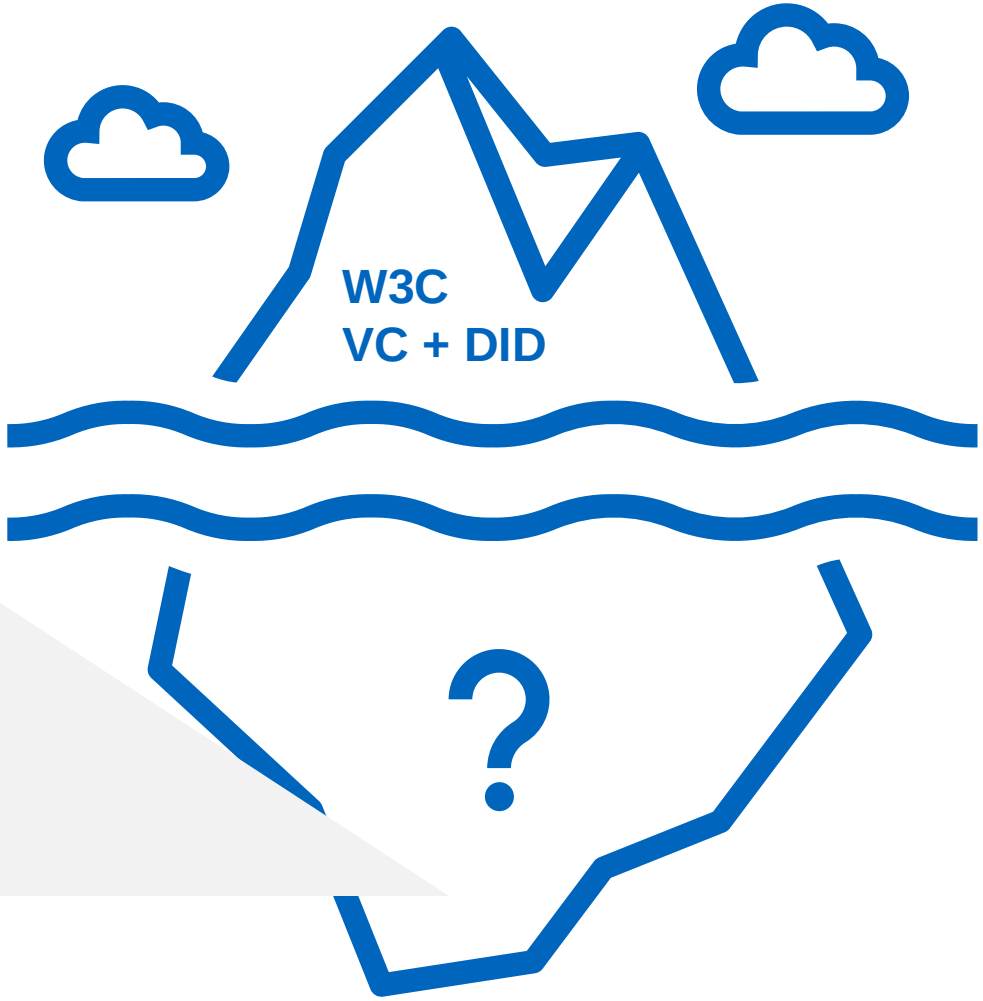
Pause for Questions!

Outline

1. Motivation
2. Introduction to Self-Sovereign Identity
3. Issuing Verifiable Credentials for Dataspace Participants
4. Authentication & Authorization with Verifiable Credentials
5. Conclusion

Open Challenges

- Many libraries are still in a very experimental state.
- Many “VC-adjacent” functionalities have no viable standards.
  - Status lists only have one standard that is arguably not sufficient (“Bitstring Status List v1.0”).
- Wallet software is also still in its infancy.
  - Usability is key and needs to be improved.
- Loss of private keys is generally not recoverable.
  - Backup solutions are simple if present at all.
  - Should privacy be traded for usability (e.g., through cloud wallets)?
- The choice of DID methods is overwhelming, even for technical experts.
  - Roughly 170 methods exist.[1]
  - Very different characteristics spanning cost, features, and security.
- General governance is hard, and fundamental questions need solid answers that go beyond just technical contributions:
  - *If we encounter a credential from an unknown issuer, how do we know if that issuer DID is actually who it appears to be?*
  - *And who is able/allowed/trusted to decide which issuers are trusted?*



[1]Hoops, F., Mühle, A., Matthes, F., & Meinel, C. (2023, July). A taxonomy of decentralized identifier methods for practitioners. In *2023 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS)* (pp. 57-65). IEEE.

Learnings & Outlook

- Nobody likes JSON-LD.
- Building the SSI-to-OIDC Bridge has made a big difference in SSI adoption in the project.
- Standards have converged a lot in the past few years.

Everything is starting to come together.

M. Sc.

Felix Hoops

Technical University of Munich (TUM)
TUM School of CIT
Department of Computer Science (CS)
Chair of Software Engineering for Business Information Systems (sebis)

Boltzmannstraße 3
85748 Garching bei München

+49.89.289.17126
felix.hoops@tum.de
wwwmatthes.in.tum.de