Research Identifier Ecosystem

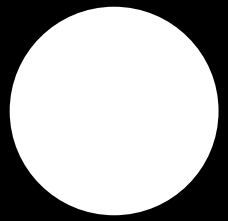
A DID is a PID with Benefits

Carly Huitema

A research identifier ecosystem is centred on

Objects

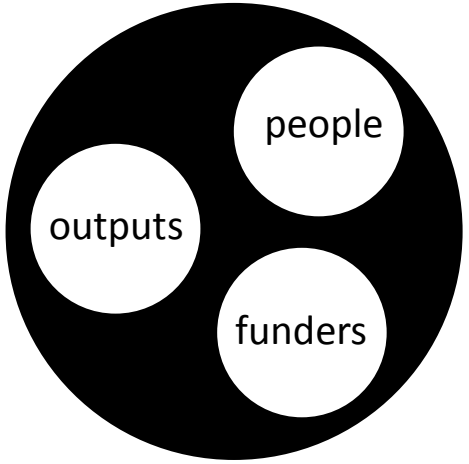
Single objects



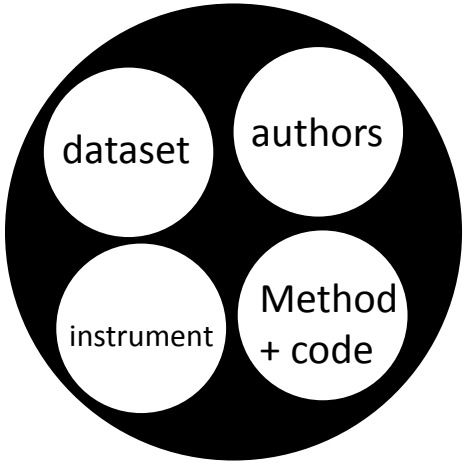
e.g. Instrument, dataset, method

Compound objects

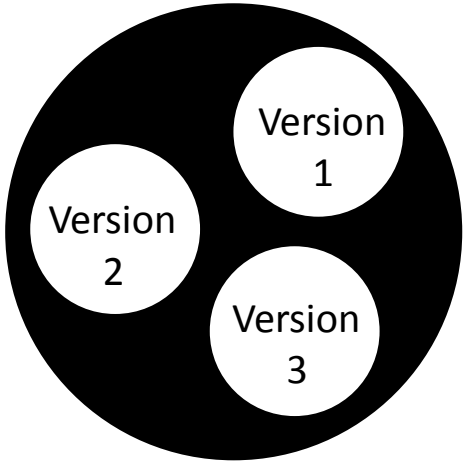
Project object



Publication object



Versioned object



Research objects can be linked

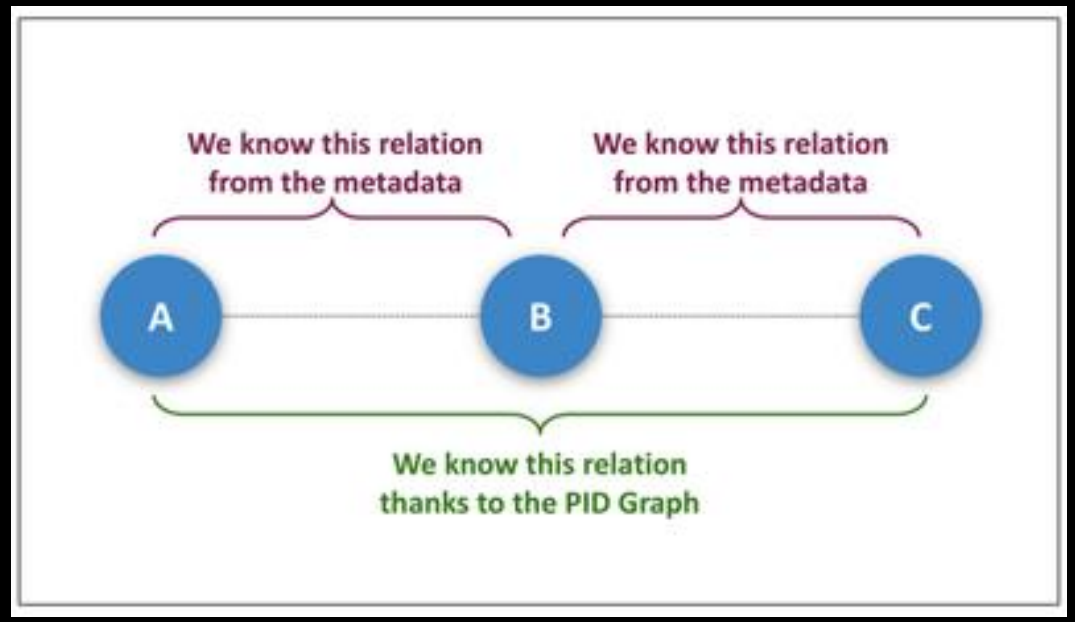
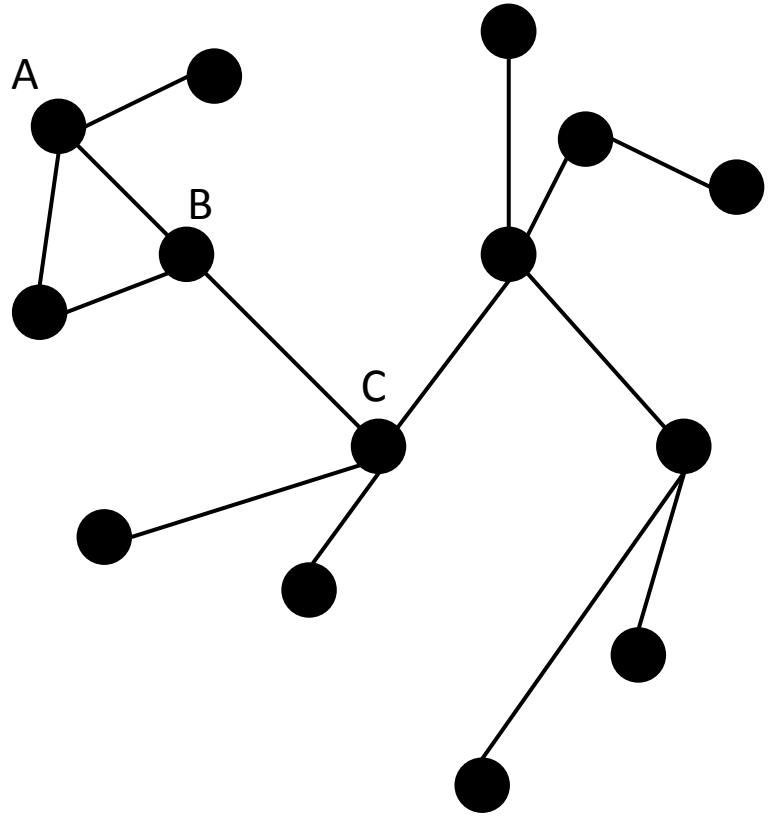


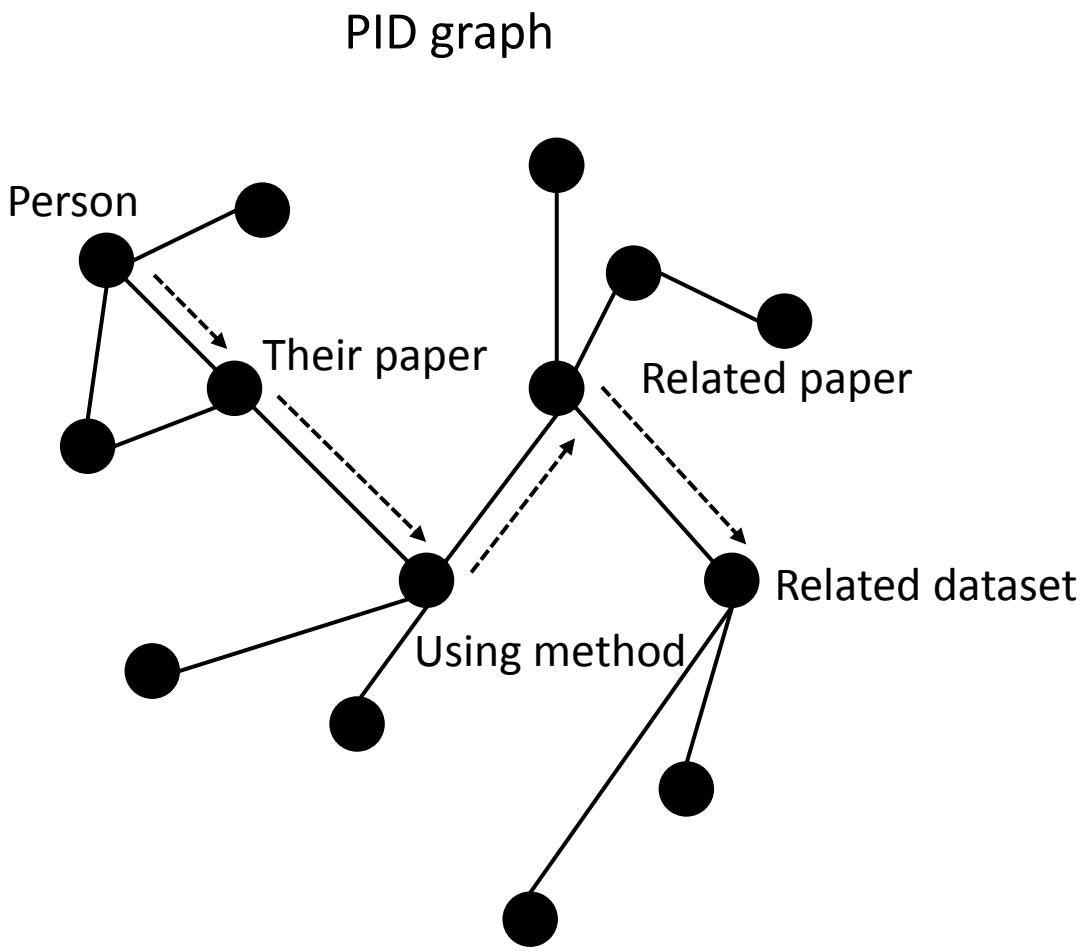
Image – project Freya

PID Graph – Research can be viewed as a network



Following the graph we can jump from person to paper to method

Generate insight
Follow funding
Find what the community values
Explore relationships
Discover new objects
Assemble new sets of related things
Locate gaps in knowledge



To…

Uniquely identify the object
Find the description of the object
Trust the result

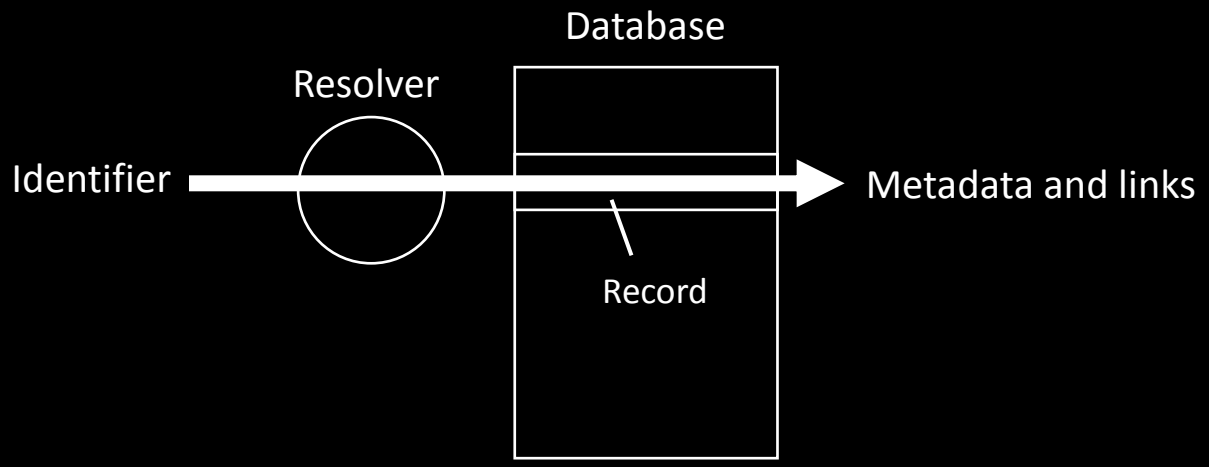
We need…

Globally unique identifier
Globally resolvable to something useful
Within a secure ecosystem

In research, these are called

PIDs

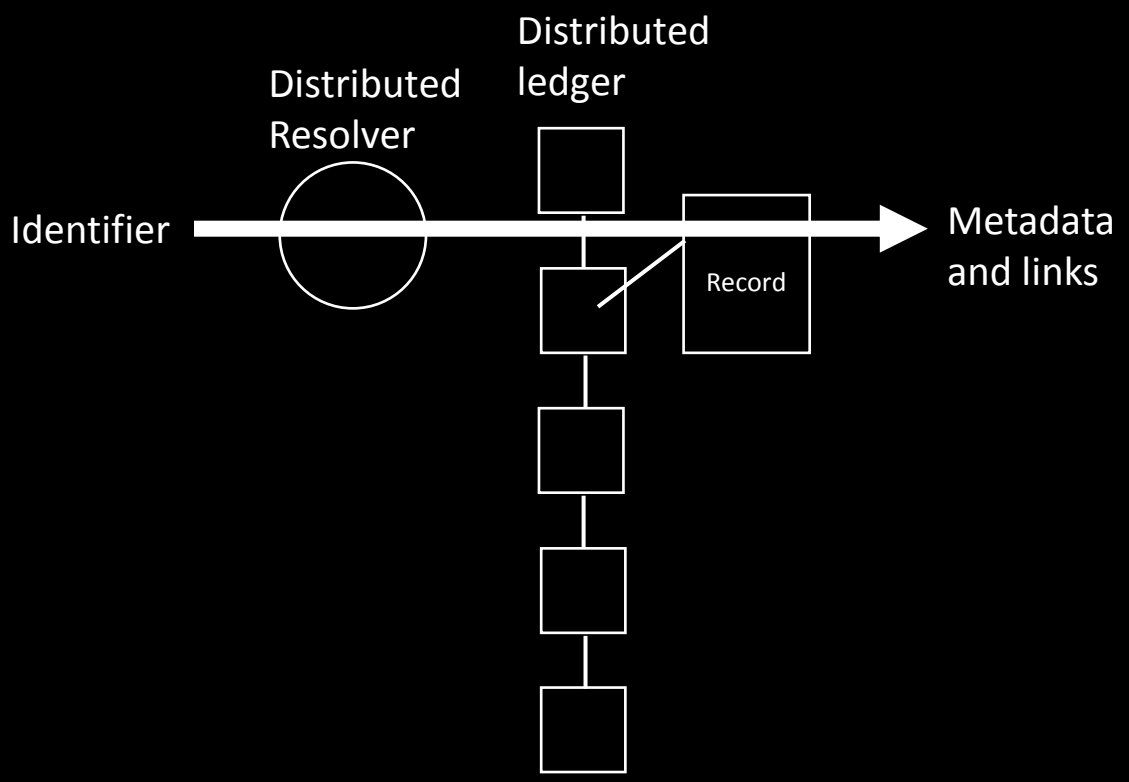
Current state



Requires:

- Centralized database(s)
- Centralized control
- Predetermined schema choice(s)
- Limits of what can be identified

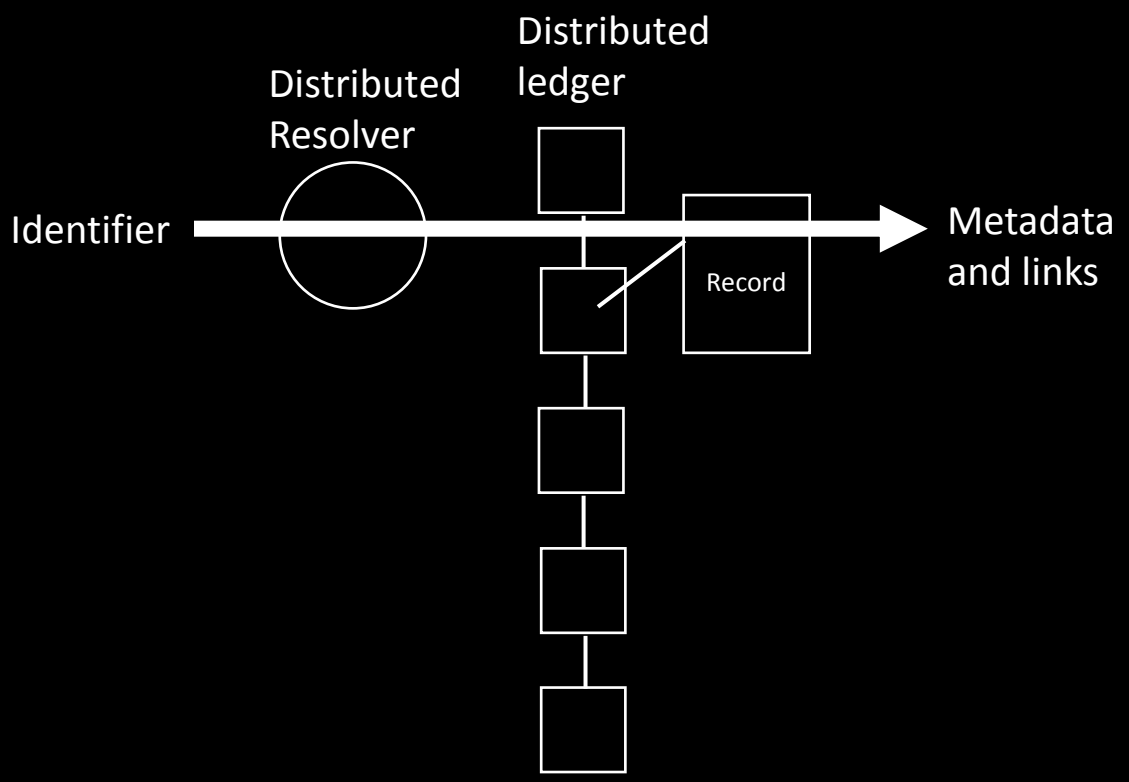
A new improvement: Decentralized Identifiers (DIDs)



Benefits:

- **Robust** decentralized network
- **Transparency, trust,** and **traceability**
- Digital signatures and **auditability**
- **Extensible** to new use cases
- **Opportunity** for more decentralized **innovation**
  - *People can create new things in the ecosystem while maintaining high trust*

Governed Ecosystem



Governance decides:

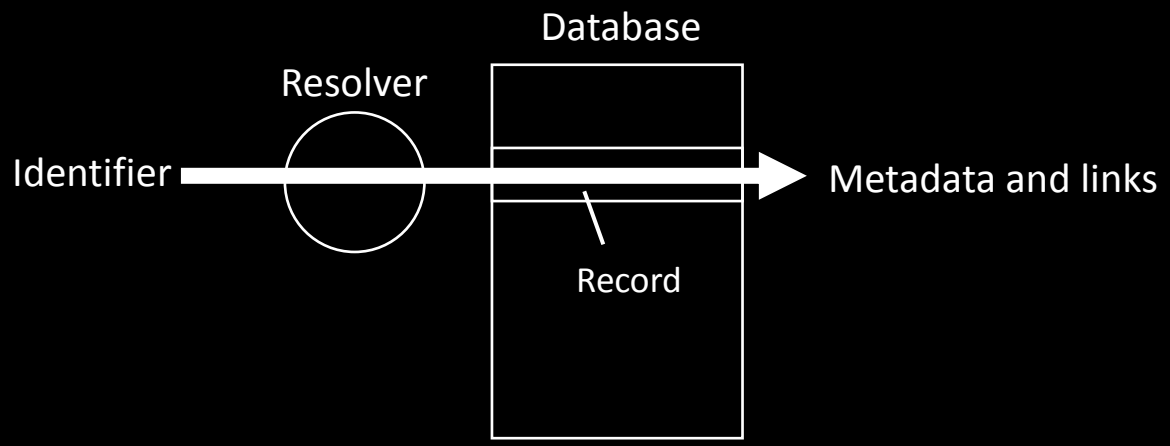
- How much of the ecosystem governed
- Who can create schemas
- Who can register identifiers
- Where can identifiers resolve to
- Role of endorsements
- Funding model
- Who can join the ledger
- Delegates

What about trust?...

How do we trust the results?

Current state

We trust the centralized authority who controls records in the database



Decentralized systems have trust built on:

Asymmetric cryptography

Asymmetric cryptography relies on keys

Anyone can independently create a private/public key pair.

Identifiers signed with the private key are mathematically confirmed using the public key

Anyone can create and sign identifiers – but no one can create an identifier with your signature

A new way to trust

We move from passwords and central authorities to…
private keys in digital wallets with public keys available on a ledger.

Basic Ecosystem Parts

1. DID Documents
   - DID Documents are signed by the creators and represent the research object

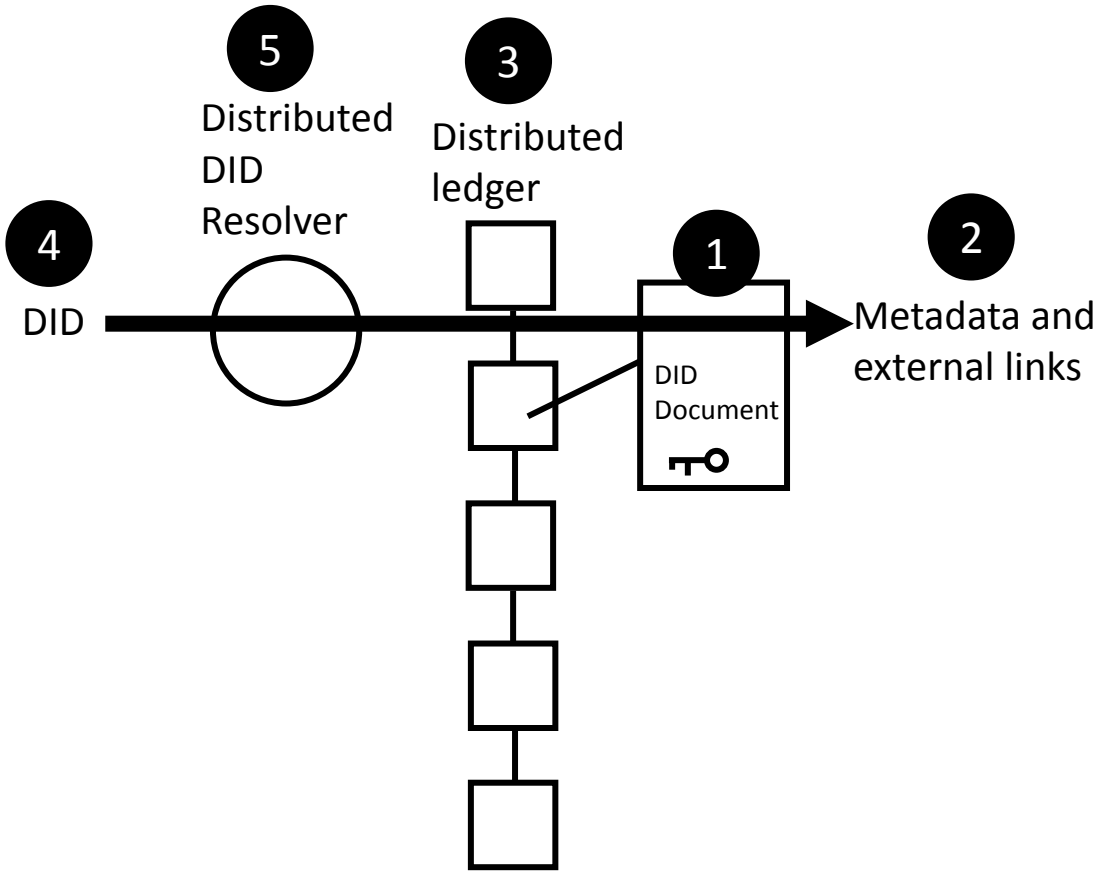
2. The DID Document can point to downstream resources, like the metadata and URL of a database

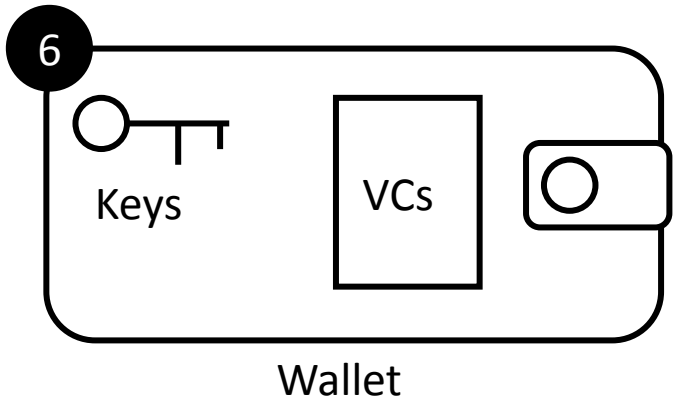
3. A location for DID Documents
   e.g. on a distributed ledger

4. The DID – identifier text e.g. DID:res:8dfj39gk

5. A DID resolving service

6. Wallets holding
   - Private keys
   - Signed credentials aka Verified Credentials (VCs)





Let’s test out some use cases….

Illustrate the possibilities

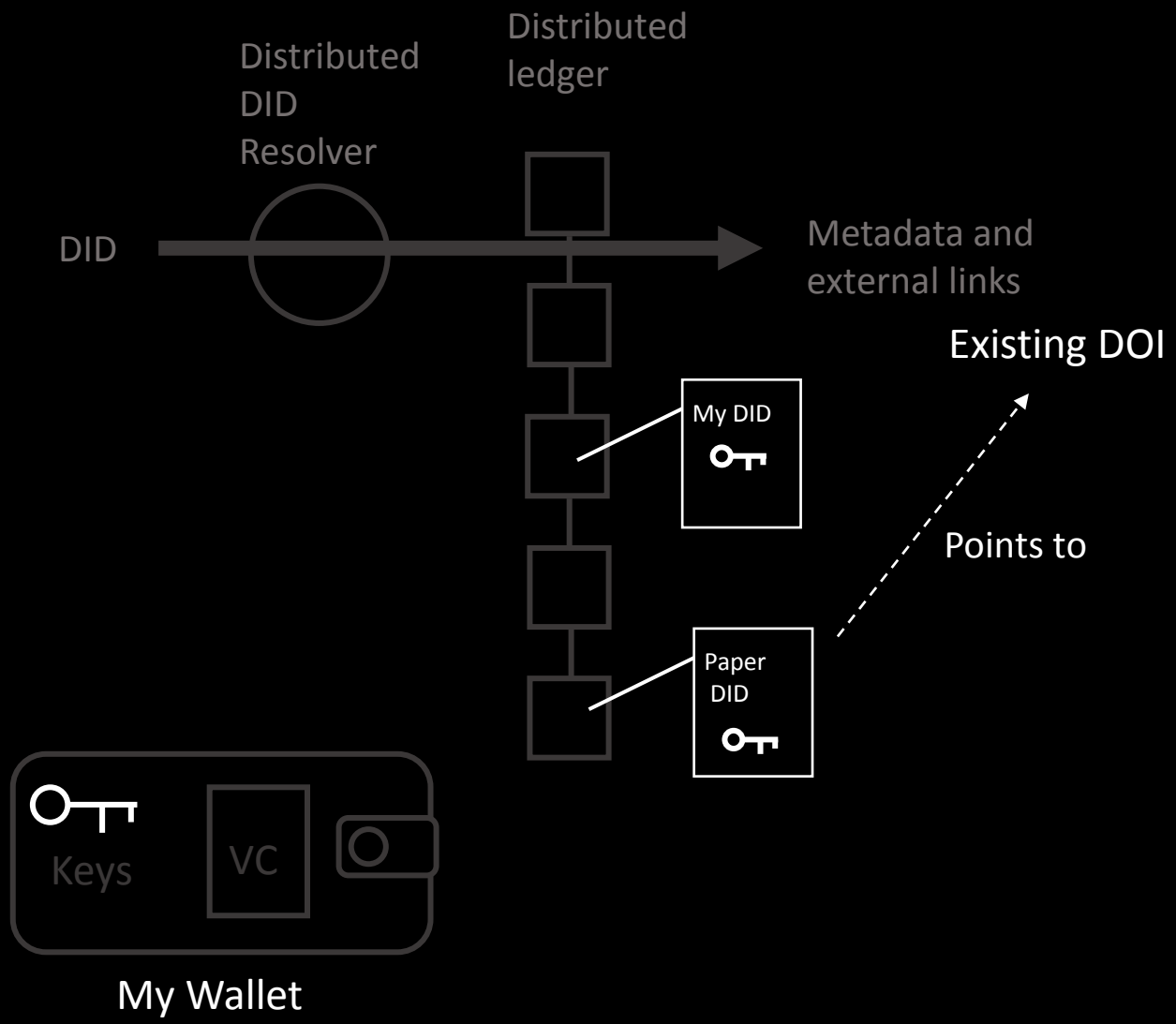
Point to an existing Identifier

My object already has a PID

I can create a DID that just points to the PID

It will be cryptographically signed by me and anyone can confirm I authored the DID

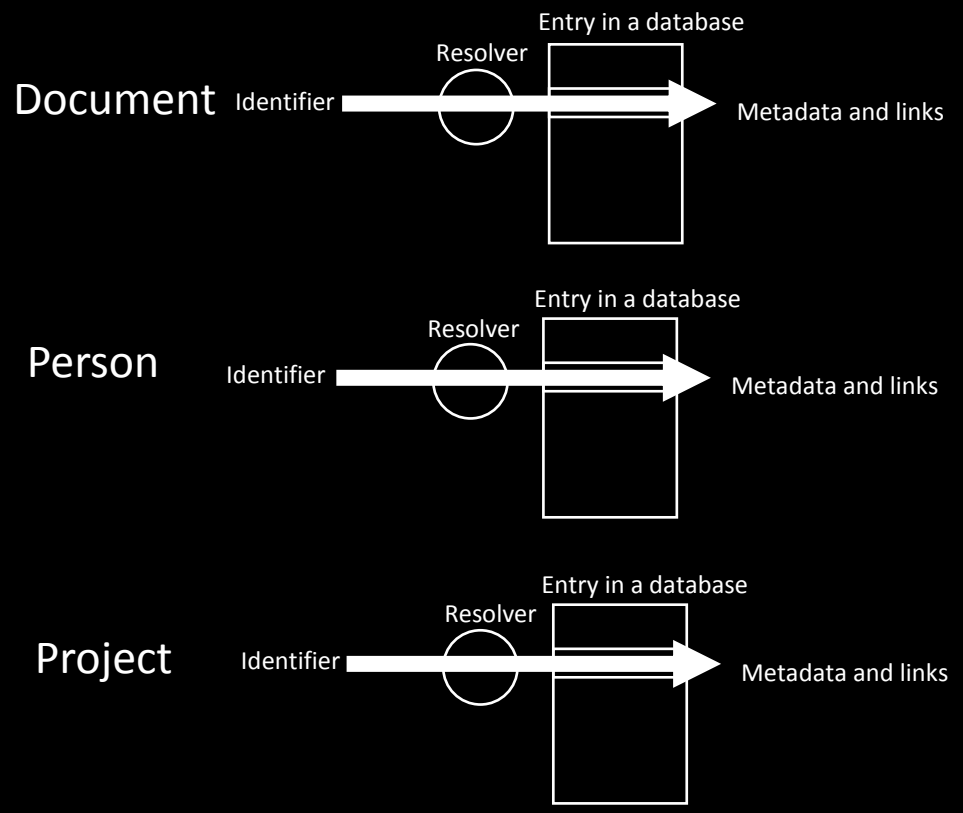
Point to an existing identifier



I don’t need a central authority to secure and trust my DID identifier

Create types of identifiers

Hard to do in the current system

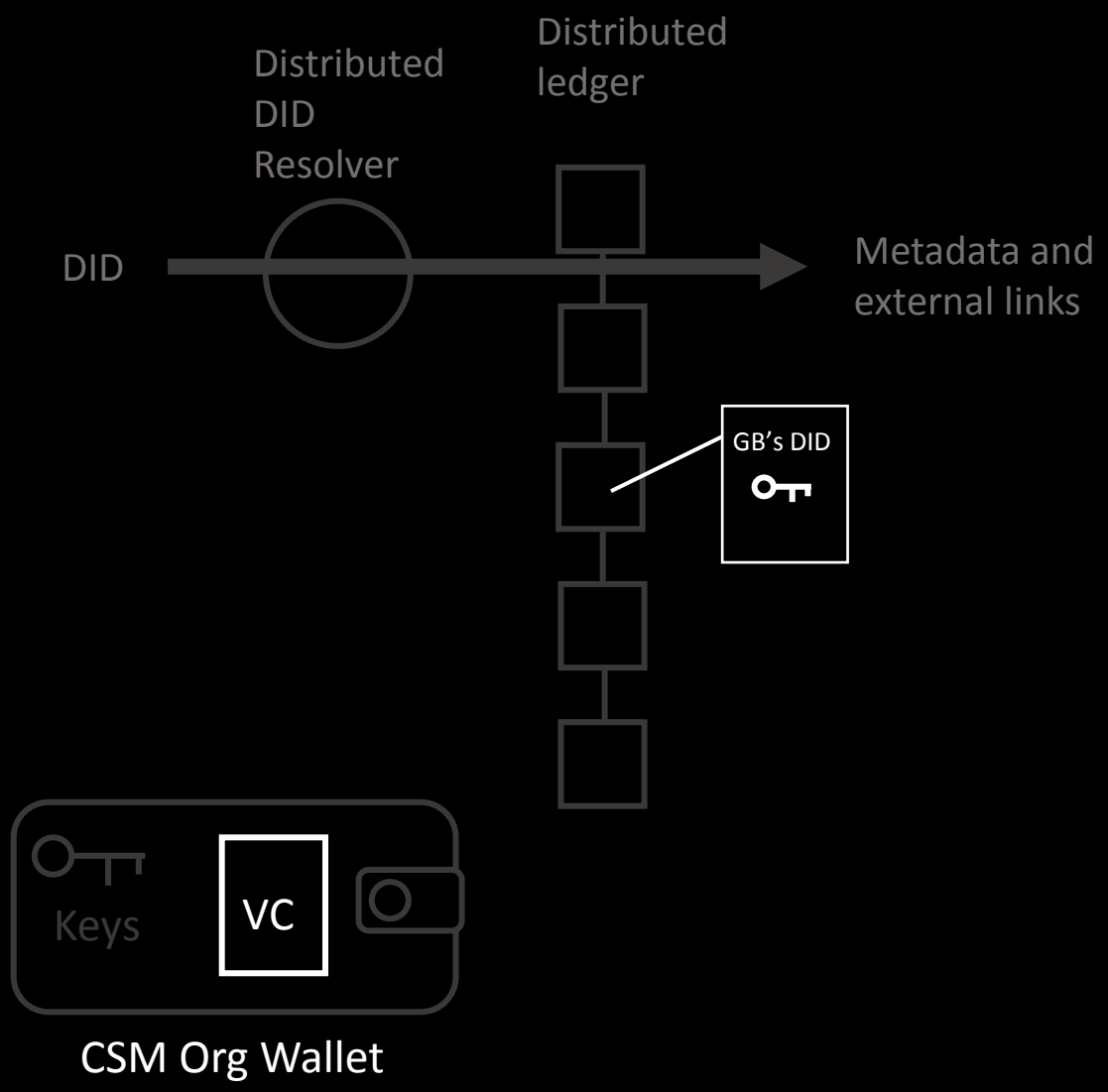


Next?

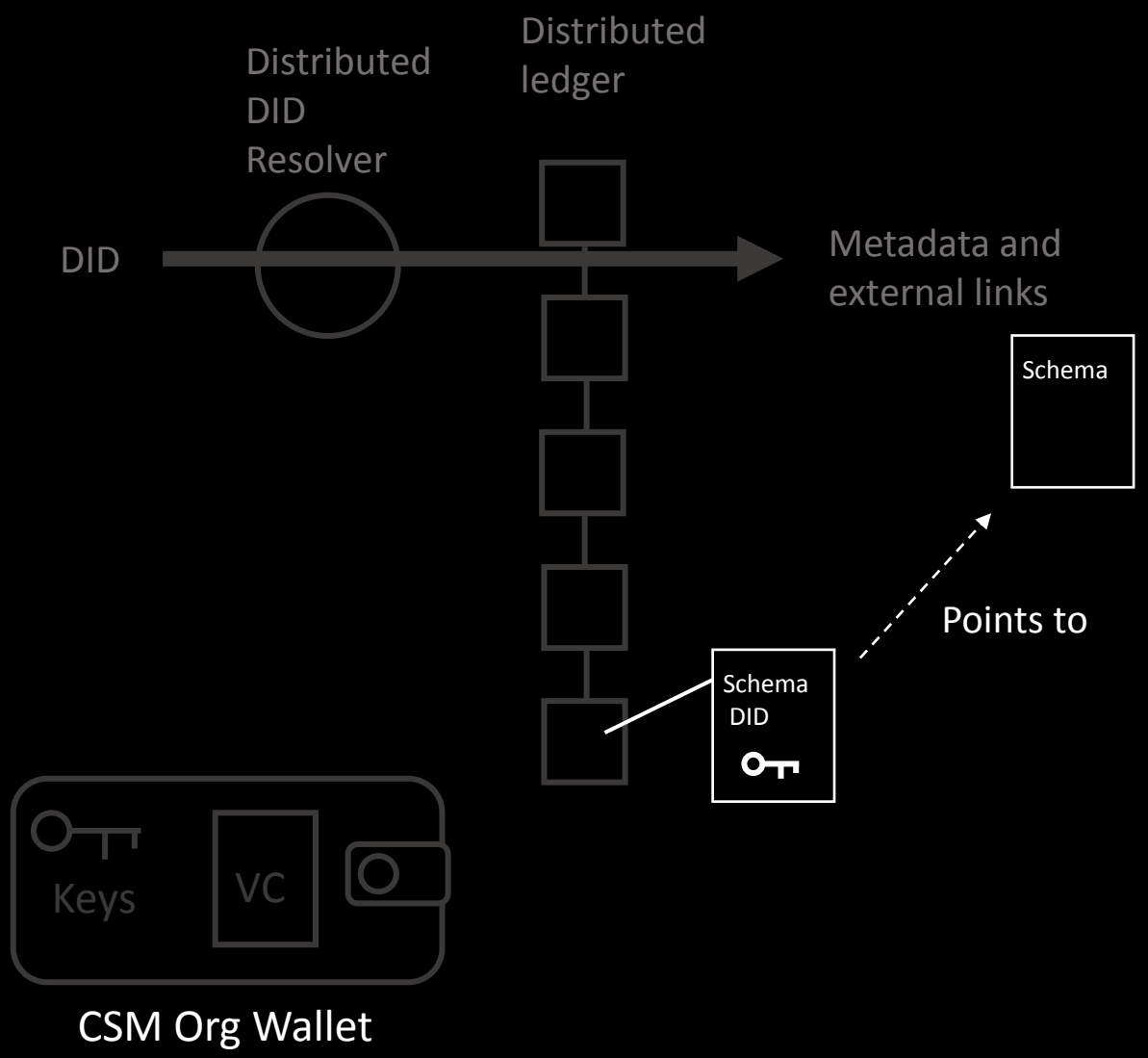
It can be easy to create a new identifier type in the DID system with high trust

Example:

The CSM wants to create a new "**Microbiome identifier type**" so that people can give their microbiome research object PIDs and describe them using this new community schema.



The Governing Body grants the CSM organization a VC granting the right to publish schemas in the ecosystem.



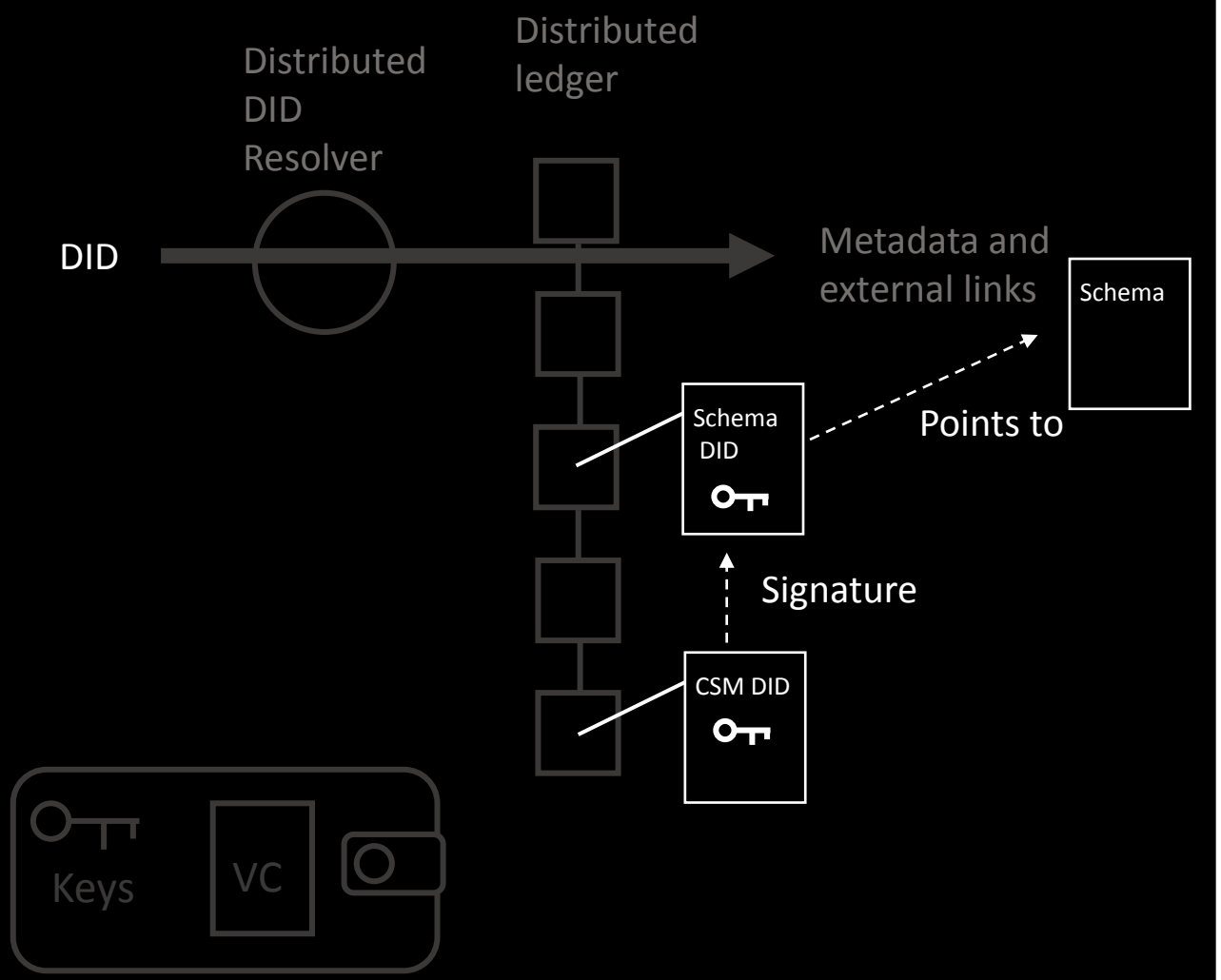
The CSM creates and registers a schema

They have permission to do this because of the VC given to them by the governing body

Use the new Microbiome Schema and mint a new identifier

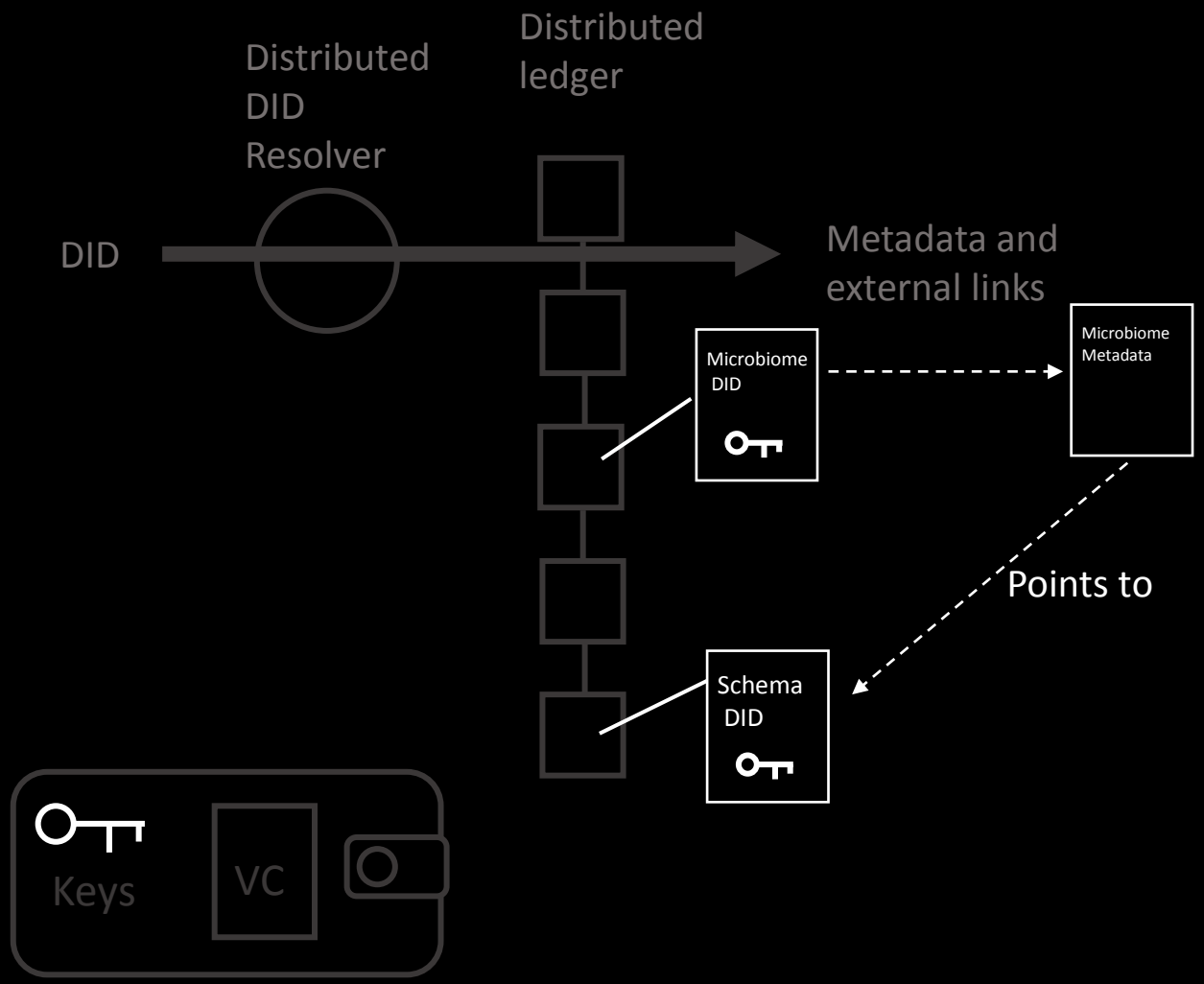
Alice has a microbiome freezer sample.

She wants to give it an identifier and describe it using the CSM microbiome schema.

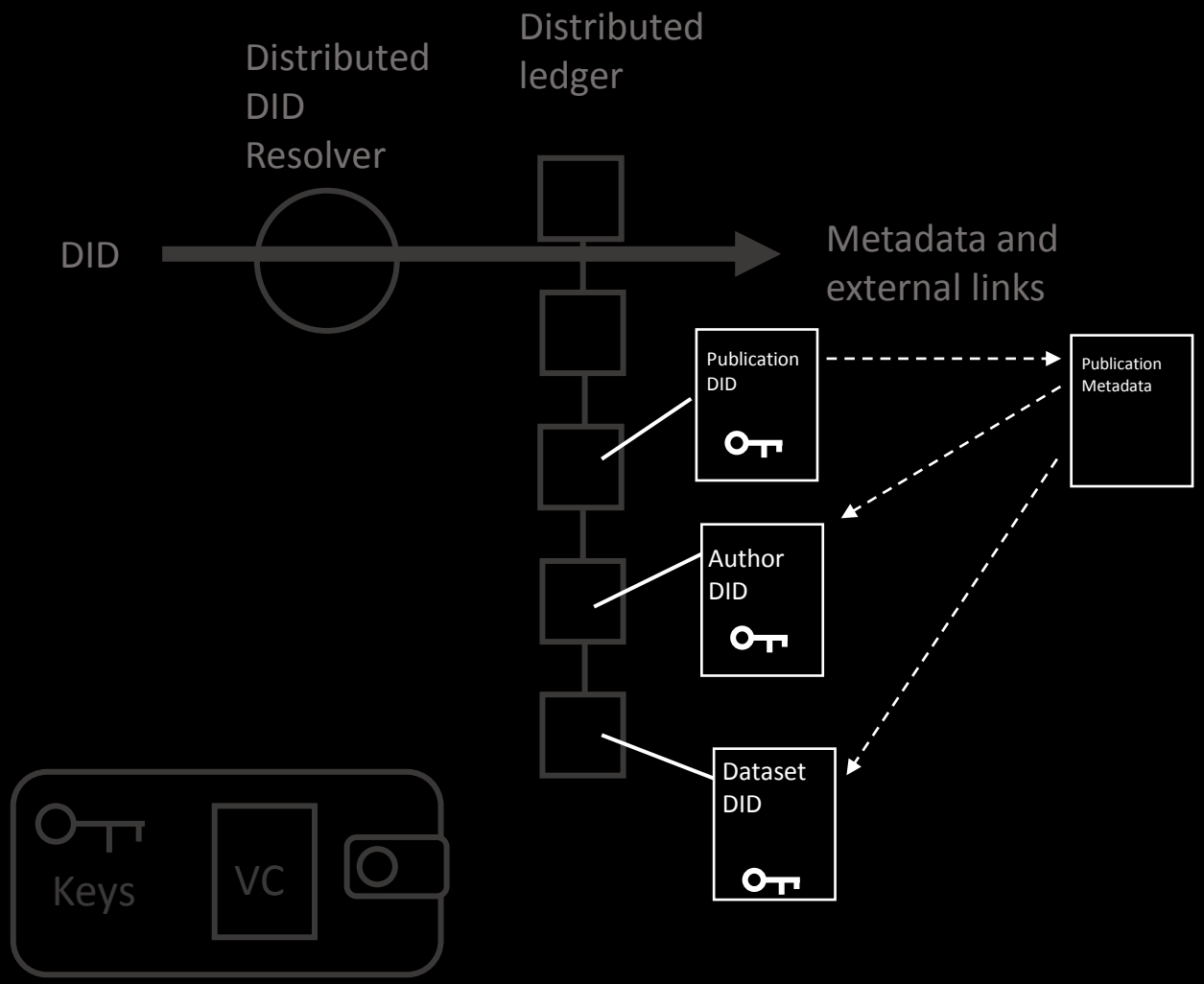


Alice finds the CSM Microbiome schema

She can confirm that the CSM digitally signed the schema DID document and trusts it is the correct one.

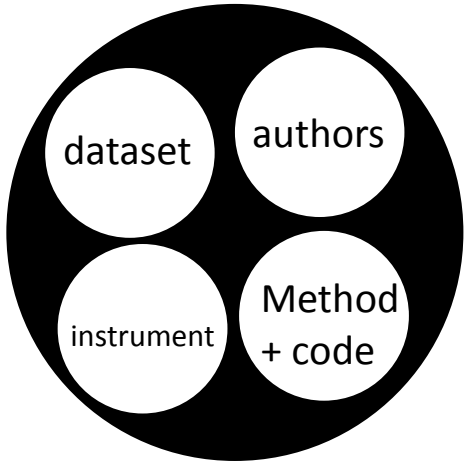


Alice describes her object using the schema and registers her new identifier.



Compound objects can be created with a schema that references DIDs

Publication object



Compound objects can reference other DIDs

Interoperable with existing identifiers

Easy to add functionalities

A flexible system that enables creativity

Researcher control

Since everything is signed, there is **full provenance**

Enhances **reproducibility**

The technology has even more to offer than I have discussed

- Open wallets
- Key management options like delegation, quorum, and recovery
- Revocation lists
- Endorsements

DIDs are PIDs with Benefits

Benefits:

- **Robust** decentralized network
- **Transparency, trust,** and **traceability**
- Digital signatures and **auditability**
- **Extensible** to new use cases
- **Opportunity** for more decentralized **innovation**
  - *People can create new things in the ecosystem while maintaining high trust*