Decentralized Digital heritage network

Bootcamp open culturele data 11/6/21

Miel Vander Sande
Data architect at meemoo & NDE

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

dutch digital heritage network

Decentralized Digital heritage network

Solid as basis for flexible cultural heritage data exchange

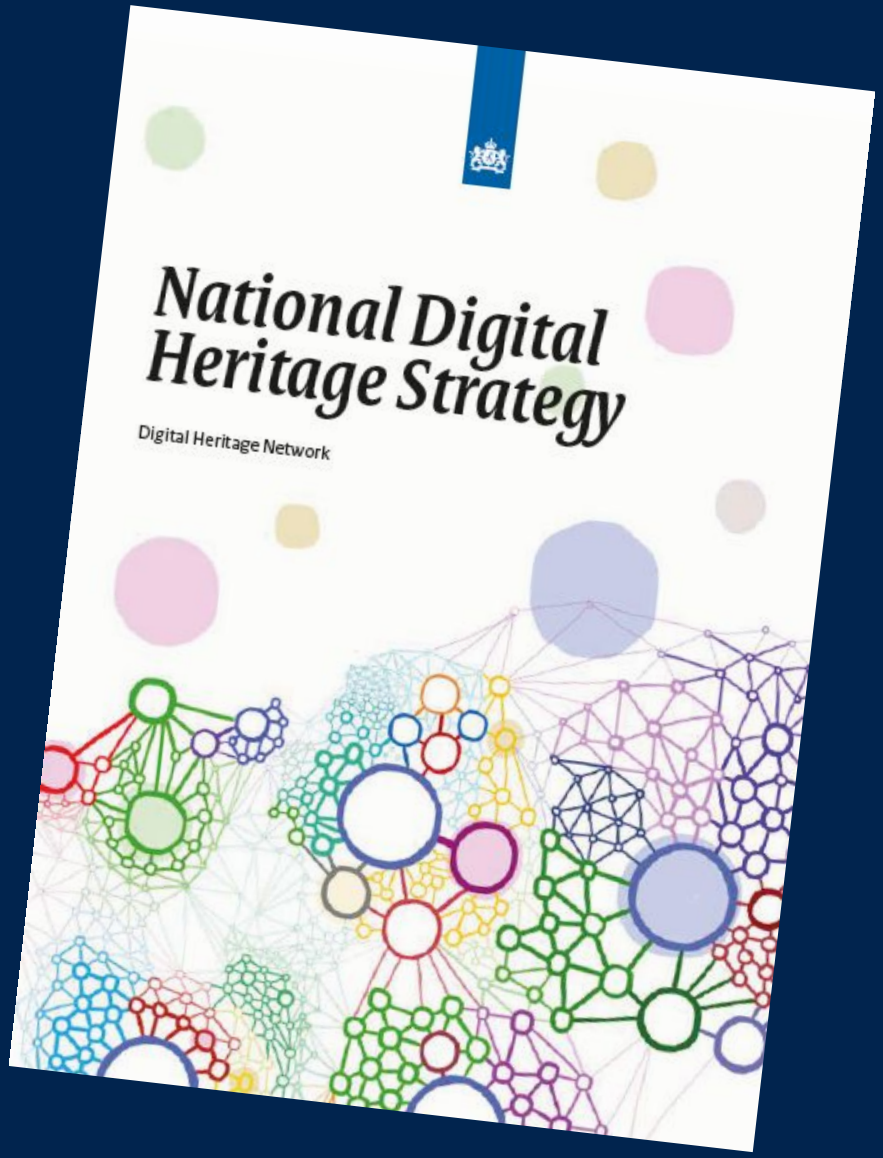
1. Usable program of Dutch Digital Heritage Network
2. Solid as implementation strategy
3. Building blocks for a decentralized heritage network

Decentralized Digital heritage network

Solid as basis for flexible cultural heritage data exchange

1. Usable program of Dutch Digital Heritage Network

The Dutch Digital Heritage (NDE) Programme



Overview of the NDE programme

Visible
Service providers

Usage profiles

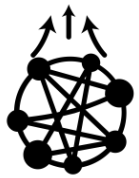
Campagne & channel

Rights & Usage

Usable
Infrastructure providers

Data & Terminology Sources (LOD)

Registries



Aggregators

Knowledge Graph

Sustainable
Source providers

PID

Preservation Policy & Certification

Cost model Preservation

Index Preservation Services

Overall

Supporting Network

Training & Education

Body of Knowledge

Services Toolbox

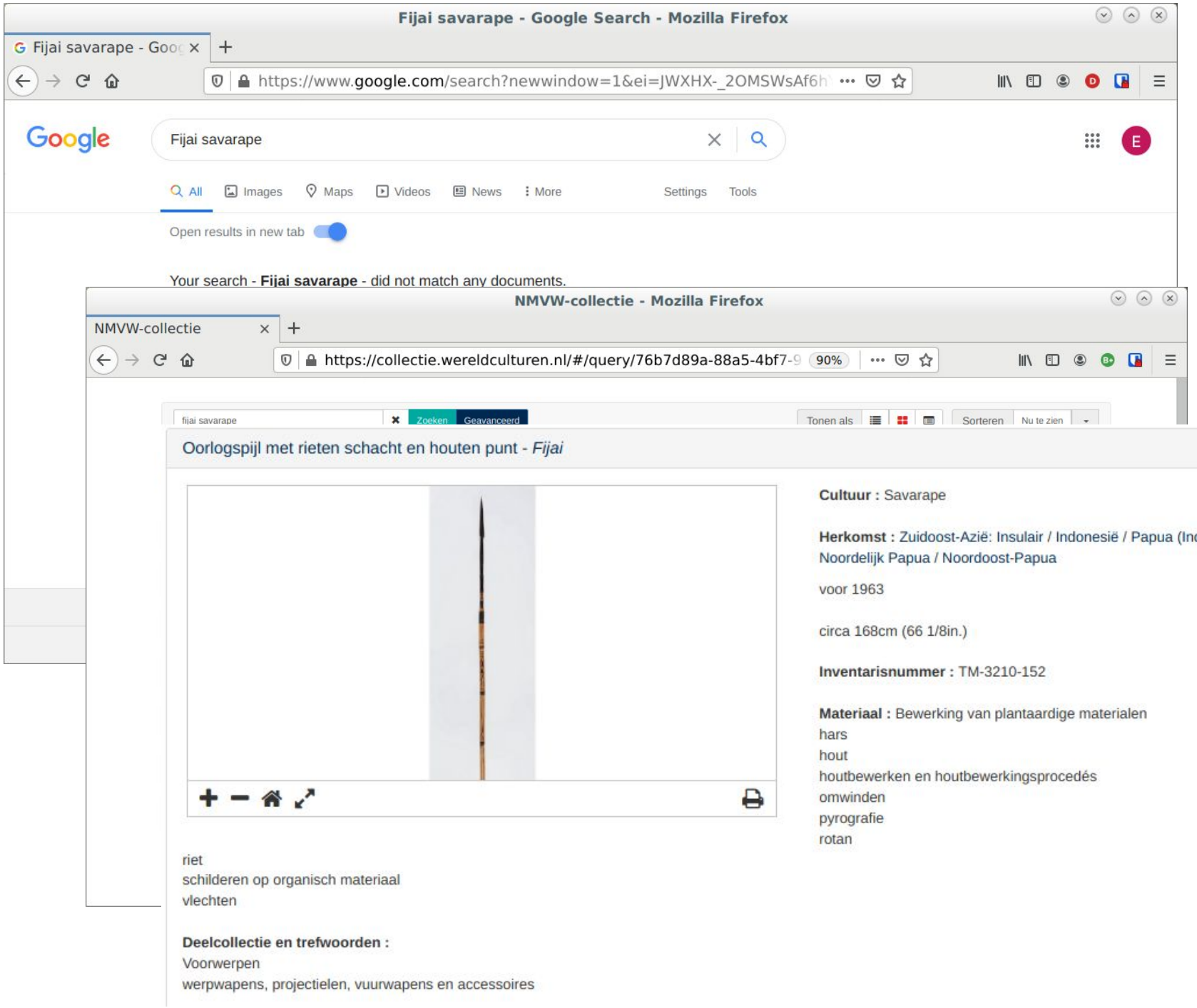
Service implementation & management

For detailed information: [www.netwerkdigitaalerfgoed.nl (also in English)](www.netwerkdigitaalerfgoed.nl)

Focus on users

They start at Google…

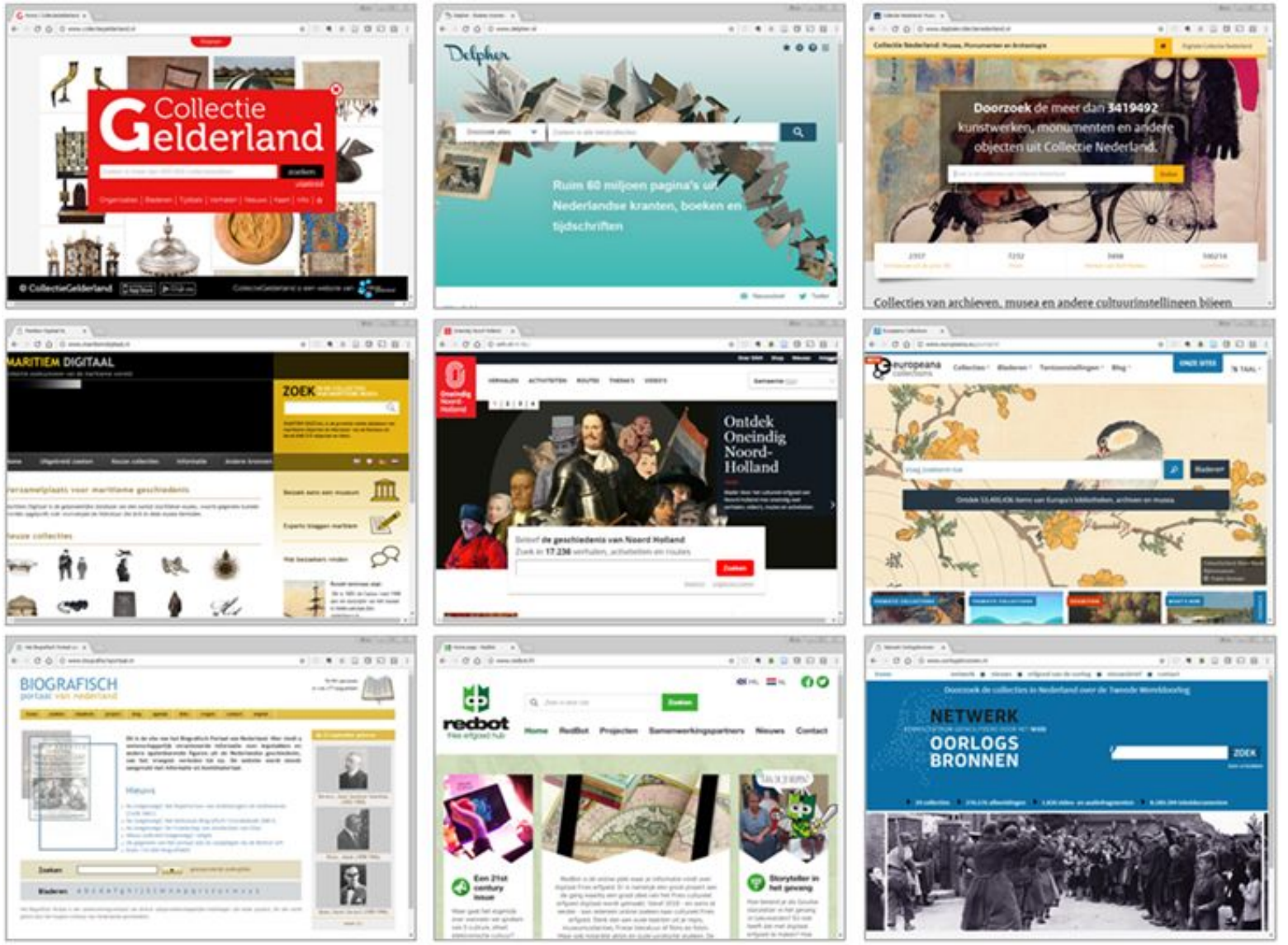
but many (long tail) topics do not show up…



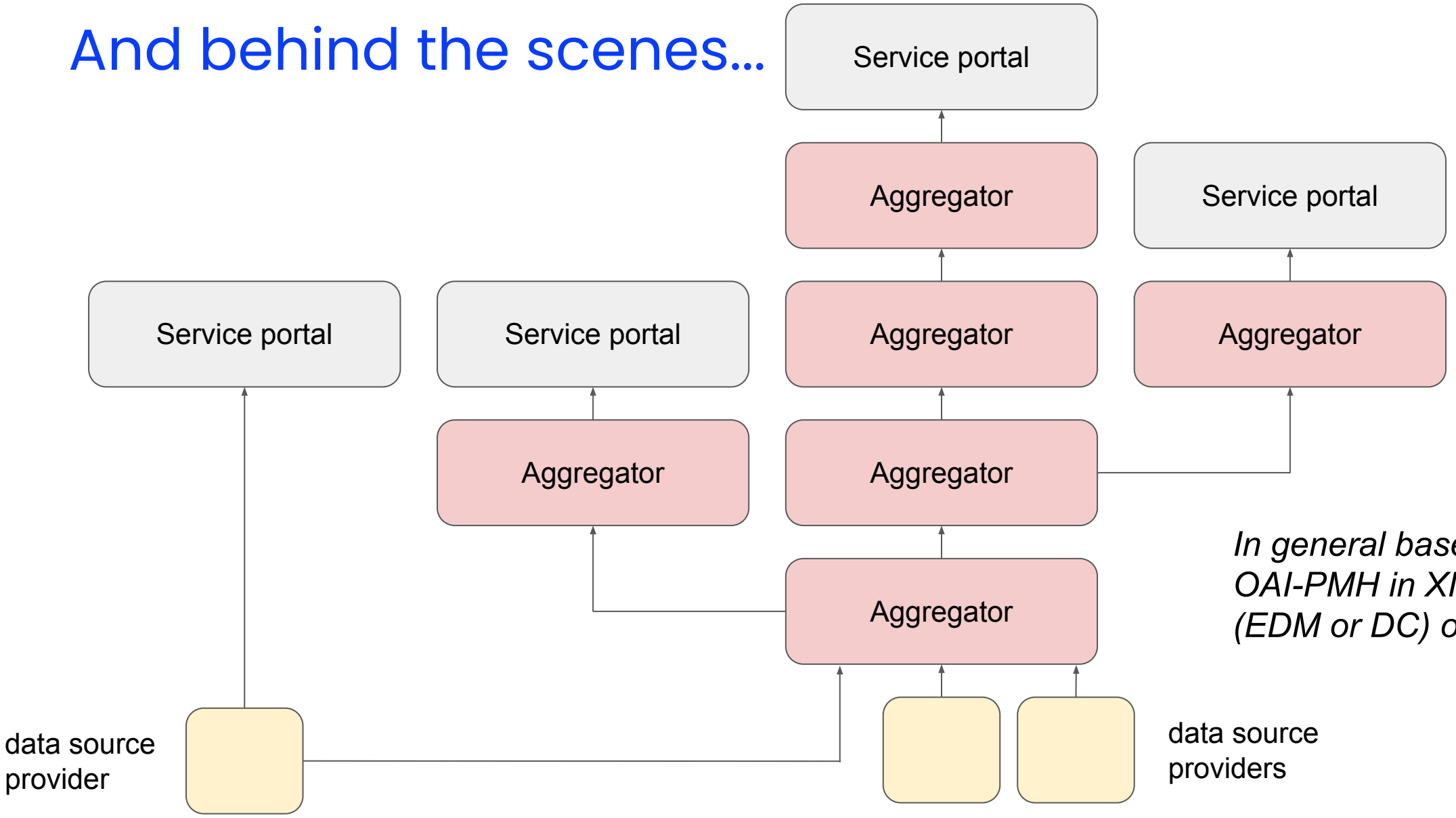
There are many portals!

but:

- which one to choose?
- can I browse from one portal to another?

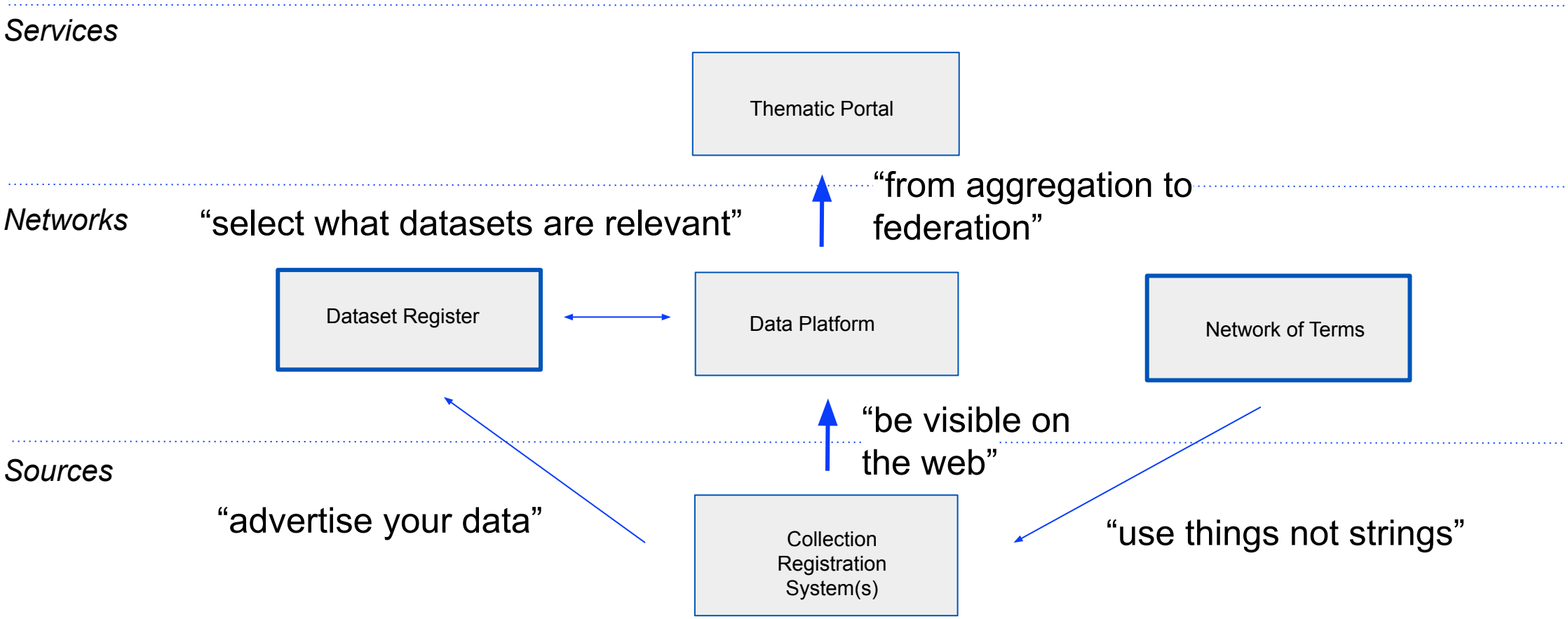


And behind the scenes...



In general based on OAI-PMH in XML (EDM or DC) or CSVs

Rebuilding the network step-by-step



meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

Decentralized Digital heritage network

Solid as basis for flexible cultural heritage data exchange

2. Solid as implementation strategy

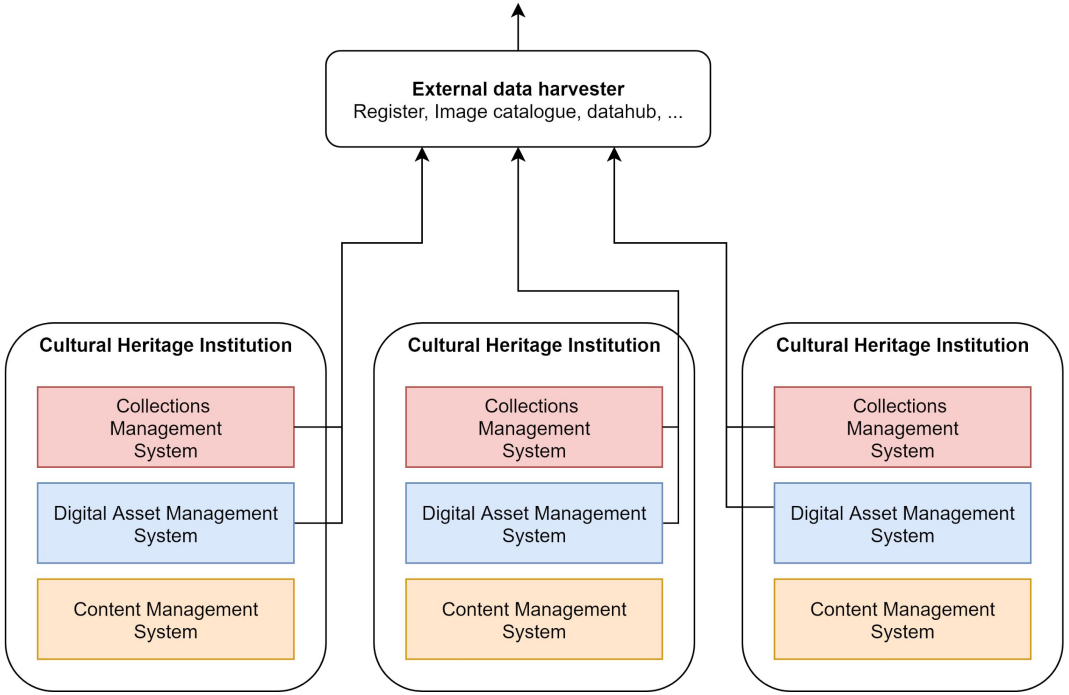
Why data aggregation layers?

Institutions often need to scale-up their operations

- short-term regional, thematic, or project-based efforts
- exchanging collection data
- approach new audiences or new forms of dissemination

Integrations are short-lived
& reasons are non-technical
eg. budget ends, know-how disappears, or change in strategy

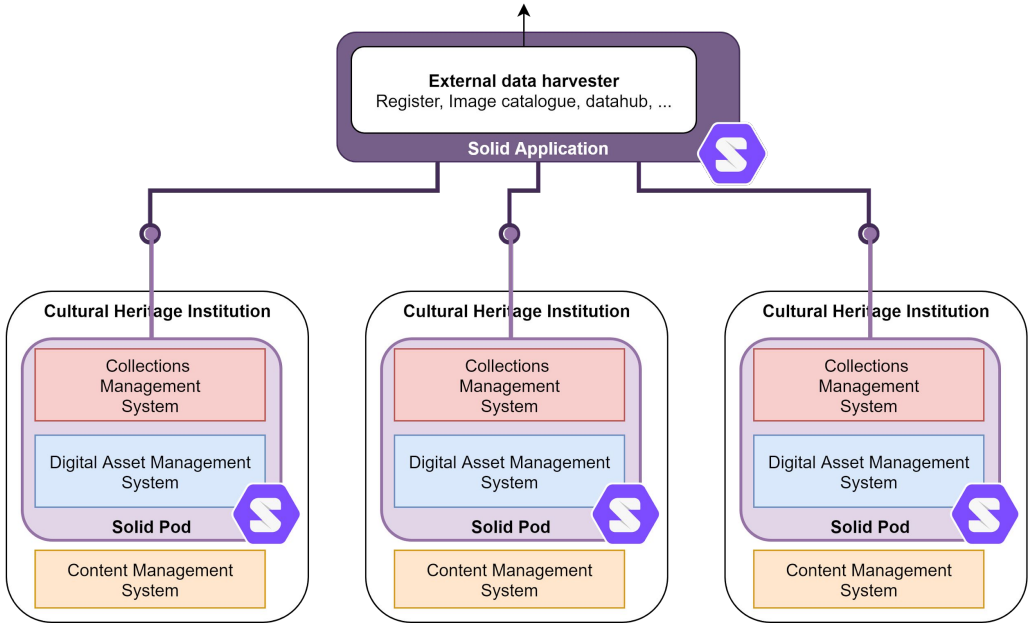
For technology,
the question is **not how to stop the constant reorientation** of collection data, but how to **support it**



Choice of technology facilitates reorientation

The technical solution should introduce flexibility

Facilitate a fast reorientation of institutions towards new integrations or networks



P2P integrations are a common, but costly solution

- each institution/harvester has a different system/data model
- custom integration to a central data harvester or aggregator
- more or less rebuild from scratch every time

Solid can possibly make a difference

integrate systems & data in a loose-coupled interoperable network

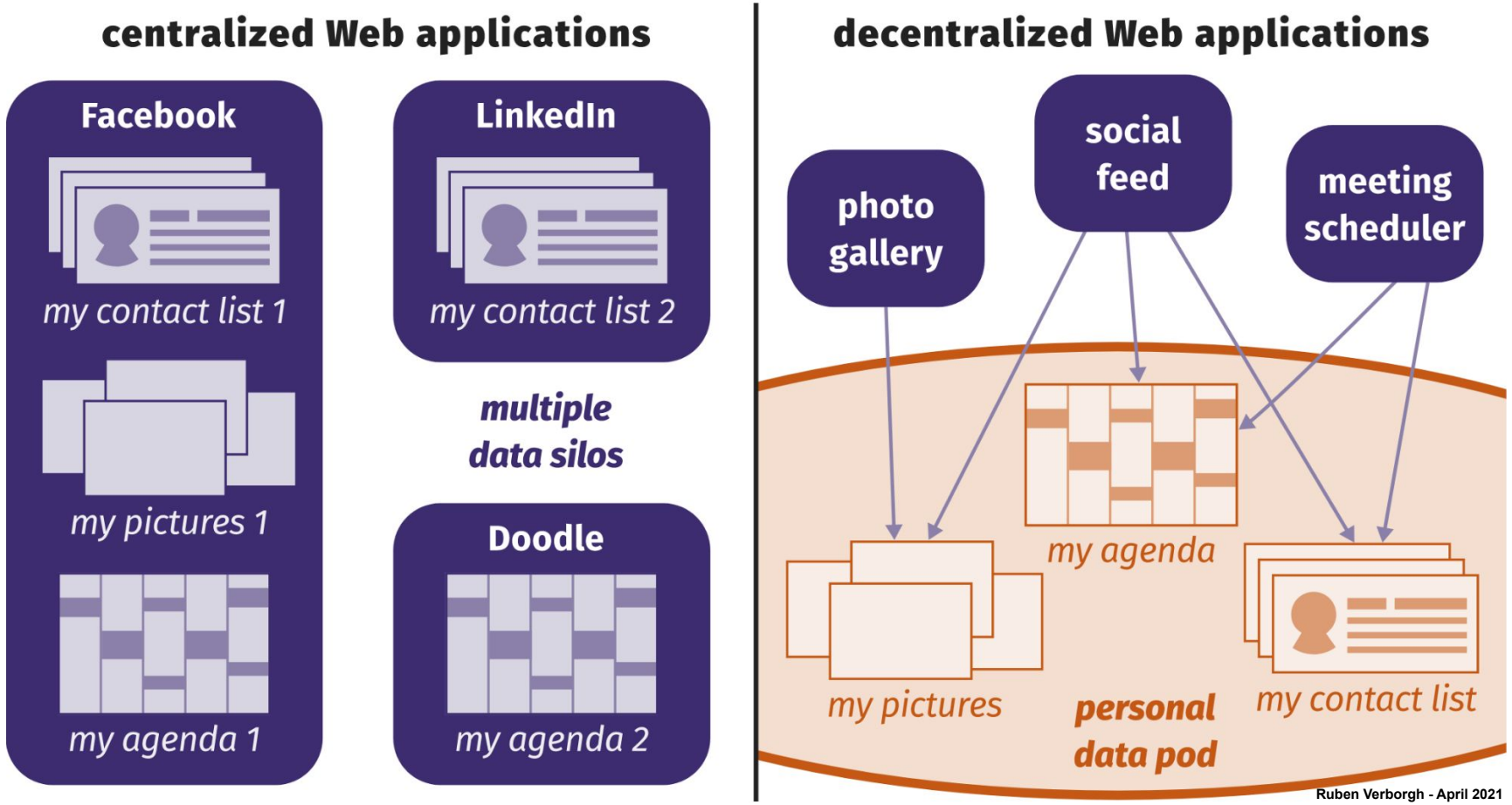
Solid (Socially Aware Cloud Storage)

Started as **alternative vision at MIT** for the growing dominance of personal data on the Web by a limited number of big players (Facebook, Google, Twitter, etc.).

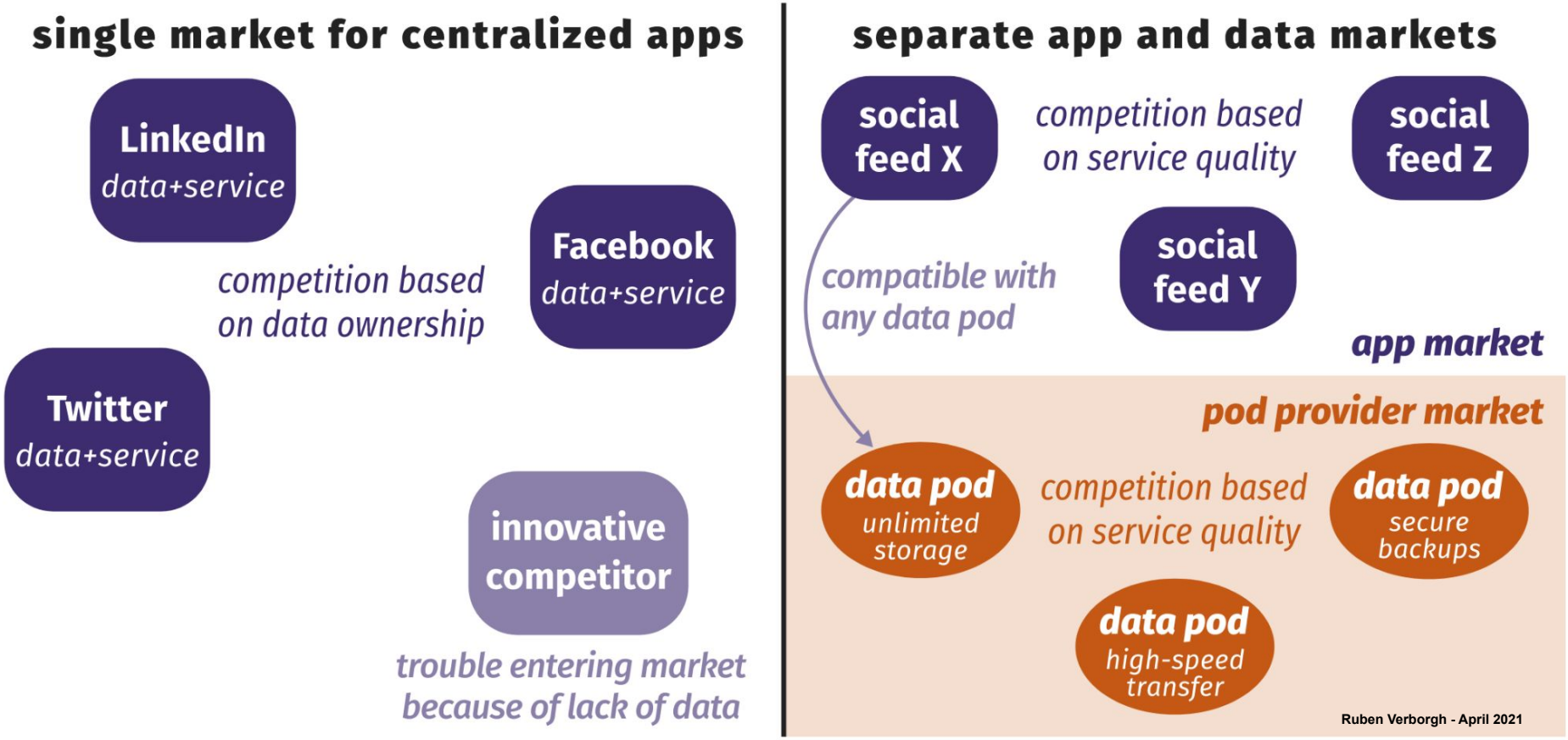
Builds the necessary principles, best practices, standards, and software components to **redecentralize data and services**.

Starts from a personal data pod from which users **manage and control their own data**, and grant applications selective access to the contents of the pod.

Data is hosted in a (personal) data pod



Separating app and storage competition drives innovation and increases mobility.



Decentralized Web as implementation strategy?

ErfgoedPOD project – NDE & meemoo & Ghent University – 2020 / 2023

Are the principles of a decentralized social network - actors announcing, sharing & following their own information - a good basis for implementing a sustainable digital heritage network?

Feasibility study on the applicability of Solid to digital heritage use cases

- not people, but institutions, service providers and applications
- not personal data, but metadata about collections
- provide missing building blocks: protocols, architecture & component blueprints
- support the Dutch (Digital Heritage Network) and the Flemish context

Decentralized Digital heritage network

Solid as basis for flexible cultural heritage data exchange

3. Building blocks for a decentralized heritage network

Project structure

Collaboration of three projects

- a shared generic foundation and output: protocols, architecture & components
- individual use cases, business processes & network profiles

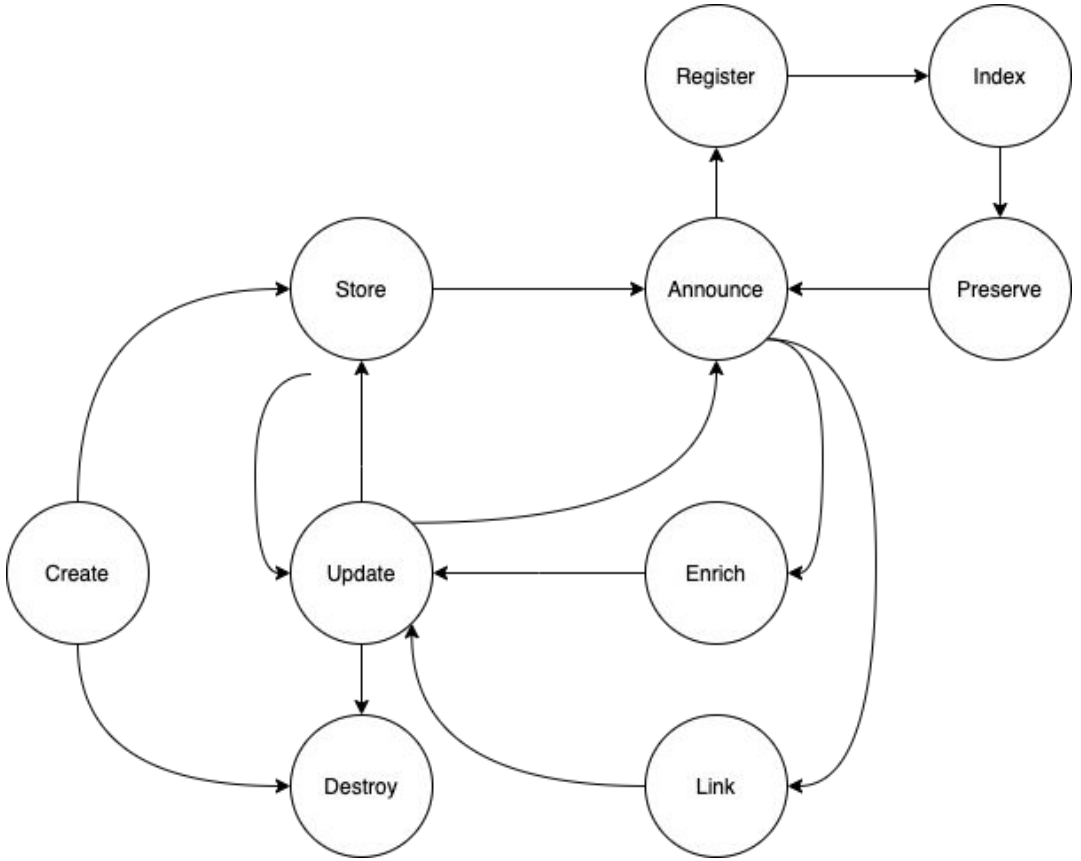
Network for exchanging artefacts

An artefact is the smallest unit that can be exchanged on the network

In digital heritage:

- dataset
- dataset metadata
- collection metadata record
- media
- ...

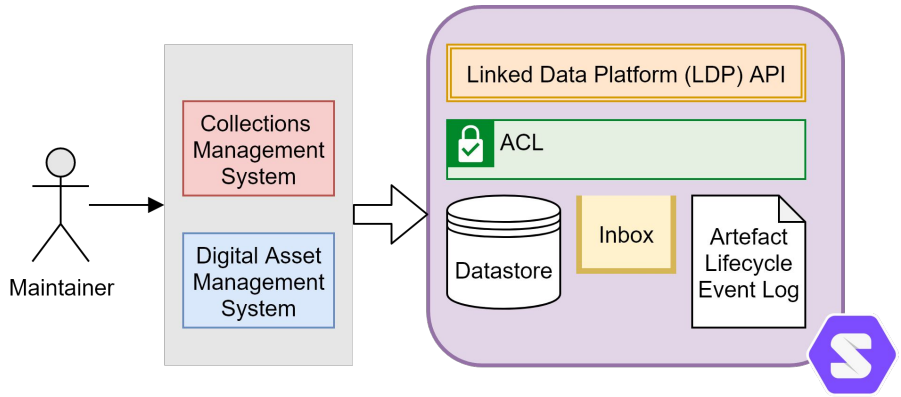
An artefact has a **lifecycle** composed of **lifecycle events**



Data Pod as main metadata exchange hub

Cultural Heritage Pod:
Collection metadata is stored in a Solid Data Pod

- API to access resources (Linked Data Platform)
- Inbox to receive notifications
- ACL layer to manage access control
- (new) immutable log resource that records all events related to artefacts known to the pod



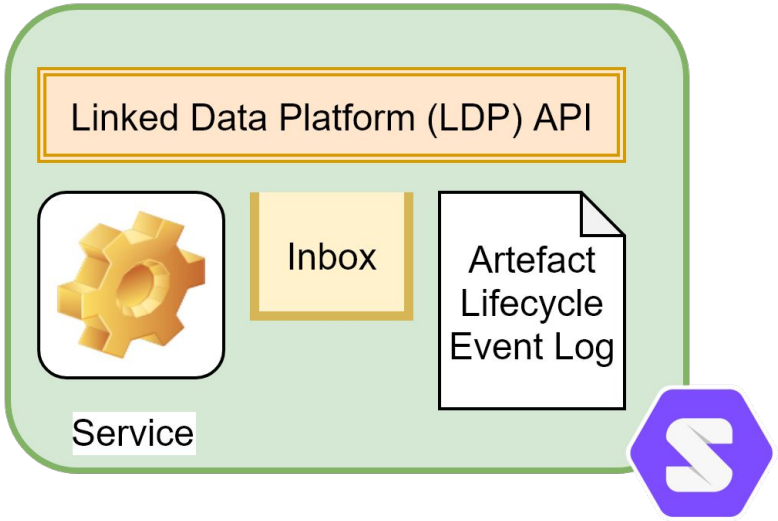
Institutions grant external services (registers, archives, indexes ...) & other institutions selective access to artefacts, metadata about artefacts or other resources.

Services as provided through Service Hubs

Service Hubs encapsulate services, which are considered black boxes

But they have a minimal interface to participate in the Solid network

- Inbox to receive notifications (exposed through Linked Data Platform API)
- immutable log resource that records all events related to artefacts processed by the service

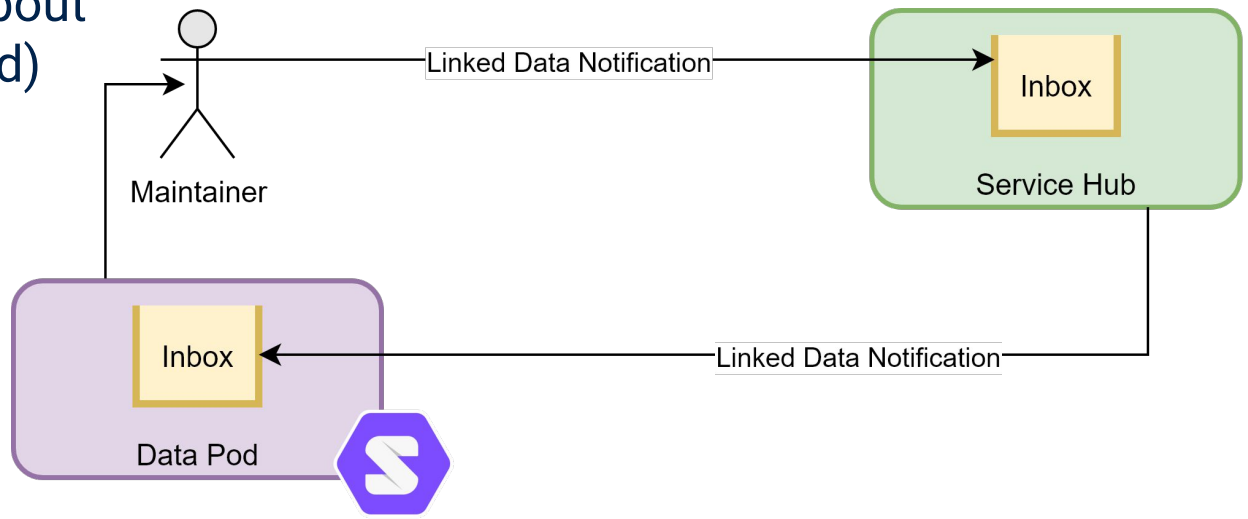


Communication through notifications

Actors communicate by notifying other actors about artefact lifecycle events (eg. Dataset was created)

Non-committal interaction to ensure loose coupling and scalability

Service Hubs pick up relevant notification and process its contents



Builds on Linked Data Notifications and Activity Streams

- working on a more concrete profile to provide guarantees
- testing decentralized social networking protocol ActivityPub

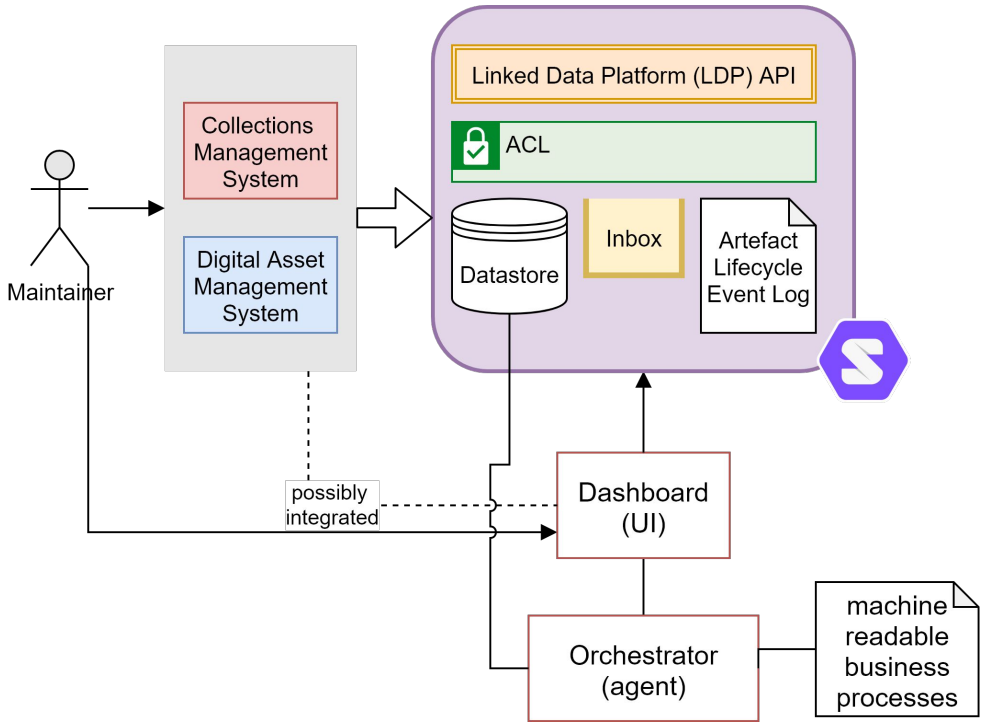
Orchestration of business processes

A Maintainer can use a **Dashboard** to manually and actively interact with the network
(eg. do actions and create notifications)

An **Orchestrator** can interact with the network on behalf of the Maintainer + Data Pod

- operates autonomously
- responds to trigger (eg. notifications)
- driven by machine-readable business processes

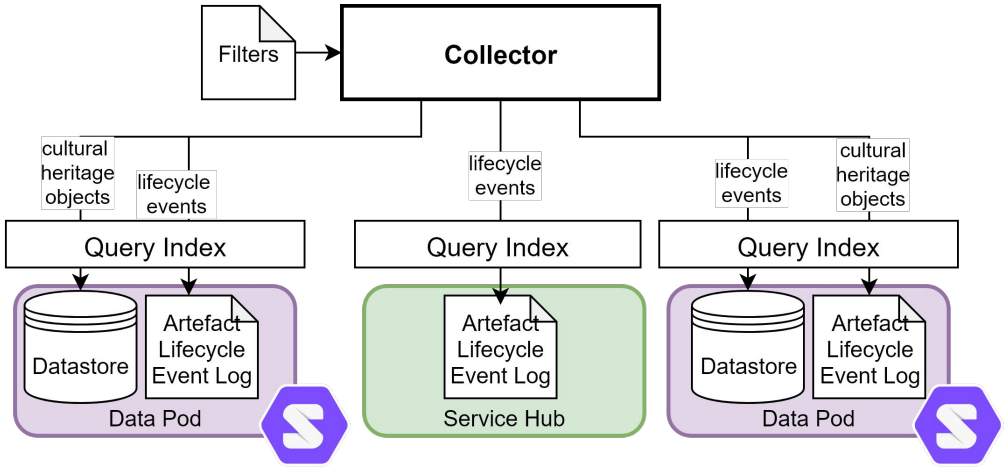
Business process policies consolidate personal preference, institution policy and ecosystem policy



Collecting distributed information

The Collector collects information about an specific artefact from distributed sources in a async manner → targeted crawling

- lifecycle reconstruction (“what happened to …”)
- object discovery (“artworks using oilpaint”)



How ensure integrity of information?

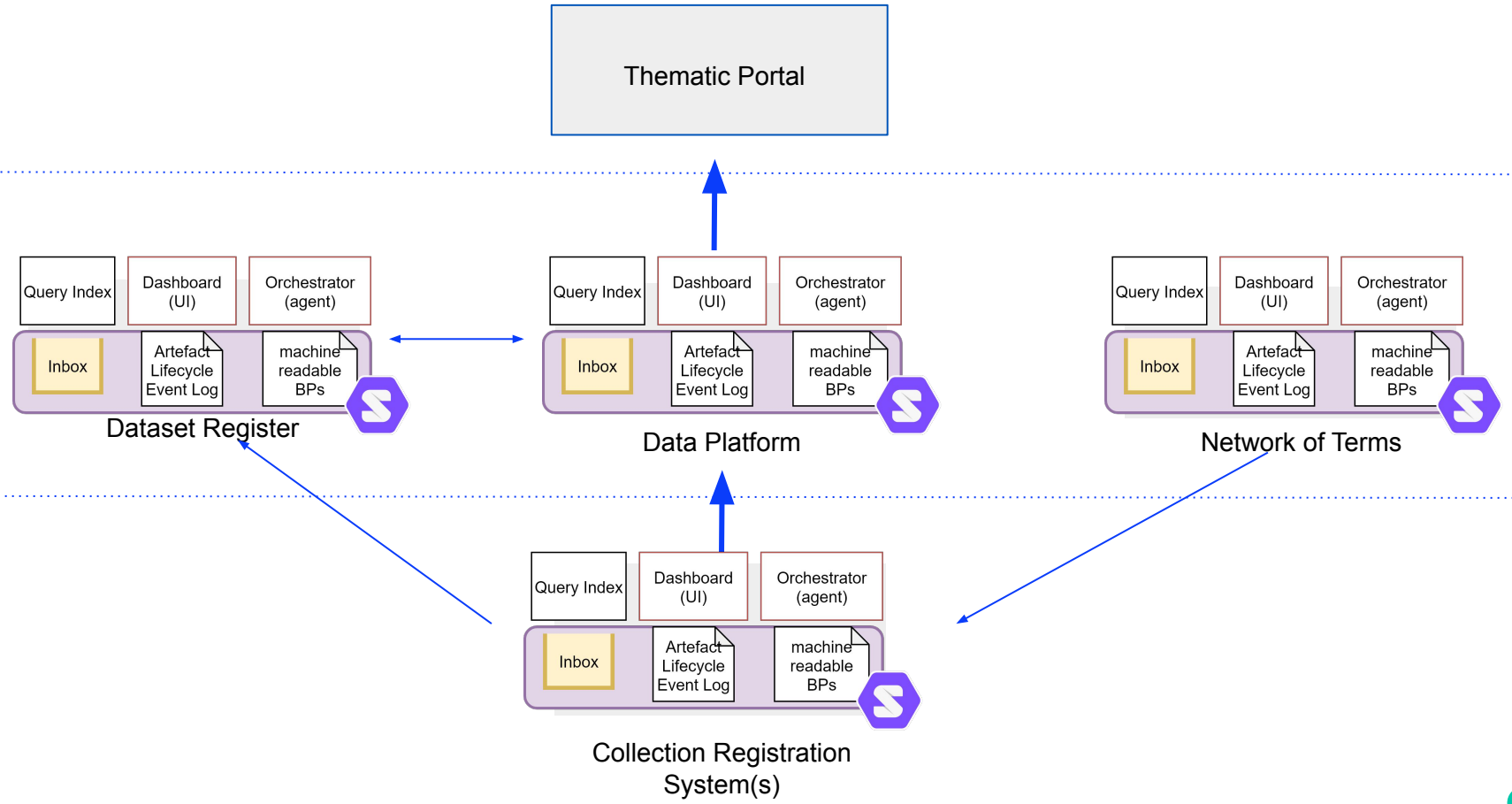
- focus on practical, best-effort solutions
- create redundancy in the distributed Artefact Lifecycle Event Logs
- sign individual events or notifications (eg. Linked Data signatures)

Apply building blocks to the NDE design

Services
Service Hub
+ Solid App

Networks
Service Hub +
Solid Data Pod

Sources
Solid Data Pod



Specifications and deliverables

ErfgoedPod

Common setups	Common software setups in Cultural Heritage institutions to set baseline
Use cases	Reference business processes in the digital heritage domain
Architecture	Implementation specification of a decentralized digital heritage network

Shared

Orchestrator	Specification of the Orchestrator component.
Data Pod	Implementation guidelines and additional requirements for Solid data pods.
Rule language	Specification of the rule language to create executable business processes.
Artefact Lifecycle Event Log	Implementation requirements for the Artefact Lifecycle Event Log.
Notifications	Specification of the possible notifications that can be used in the network.

Digital heritage as social network

STLAB 3/6/21

Miel Vander Sande
Data architect at meemoo & NDE

@miel_vds

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

dutch digital heritage network