a vault for your data

DANS Open Day
Open data, open science

continuity in a changing world

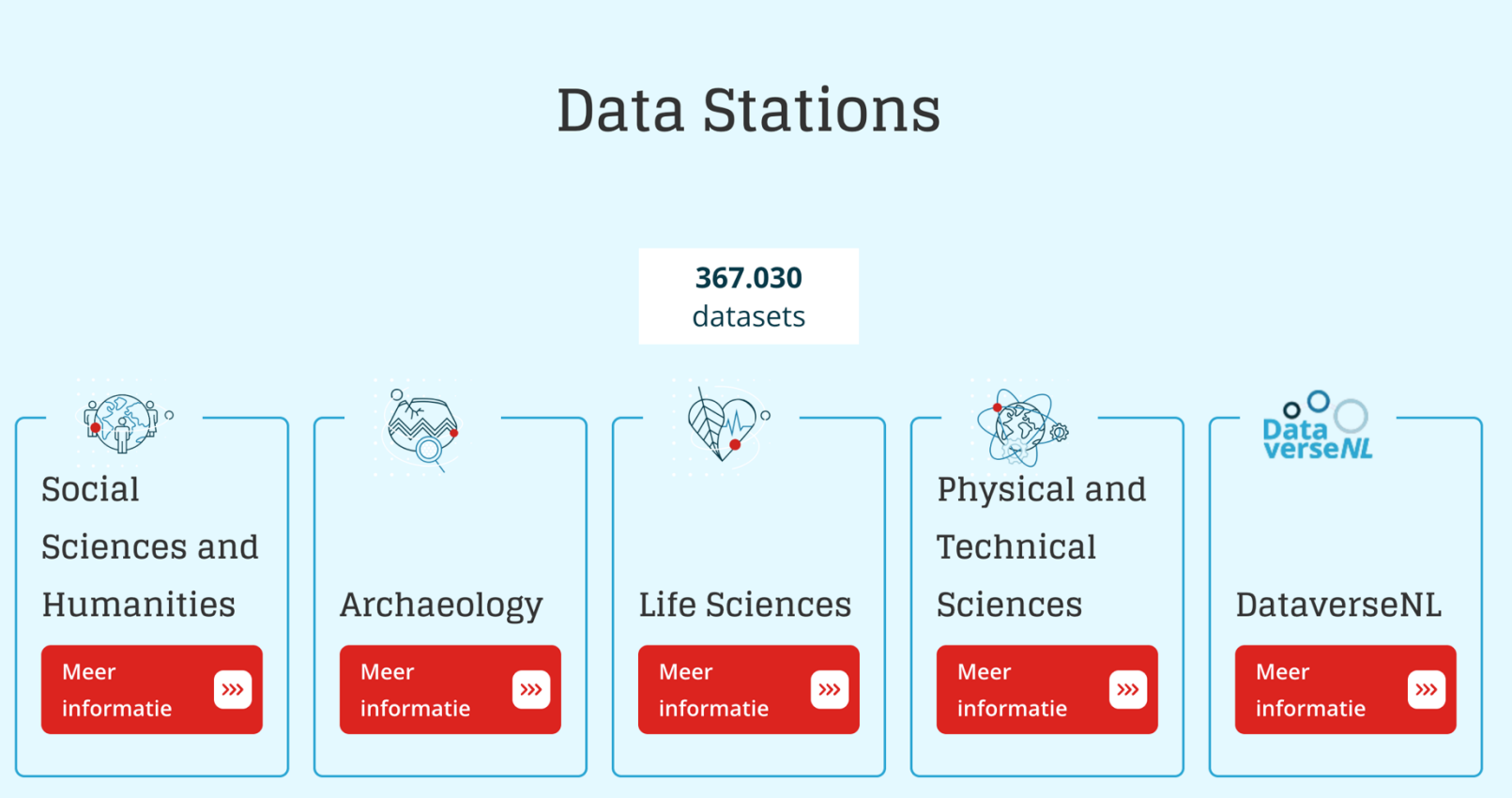
linda reijnhoudt | R&D engineer

2026-06-11

DANS

Content

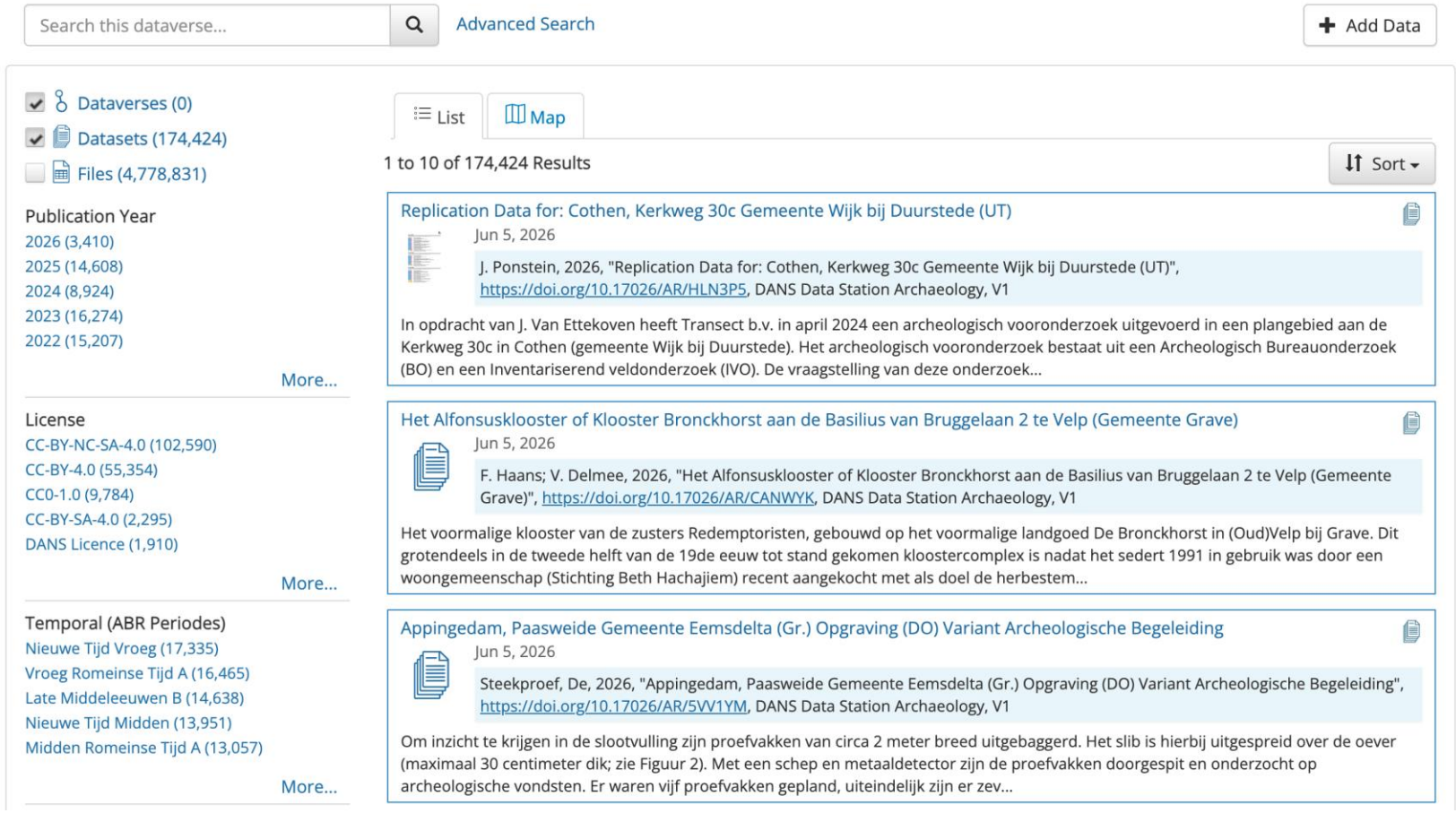
1. architecture
2. design principles
3. components
4. dataset journey
5. quality controls
6. long-term preservation
7. governance variants



DANS

Data Stations & DataverseNL

- deposit
- curate
- share
- search
- reuse



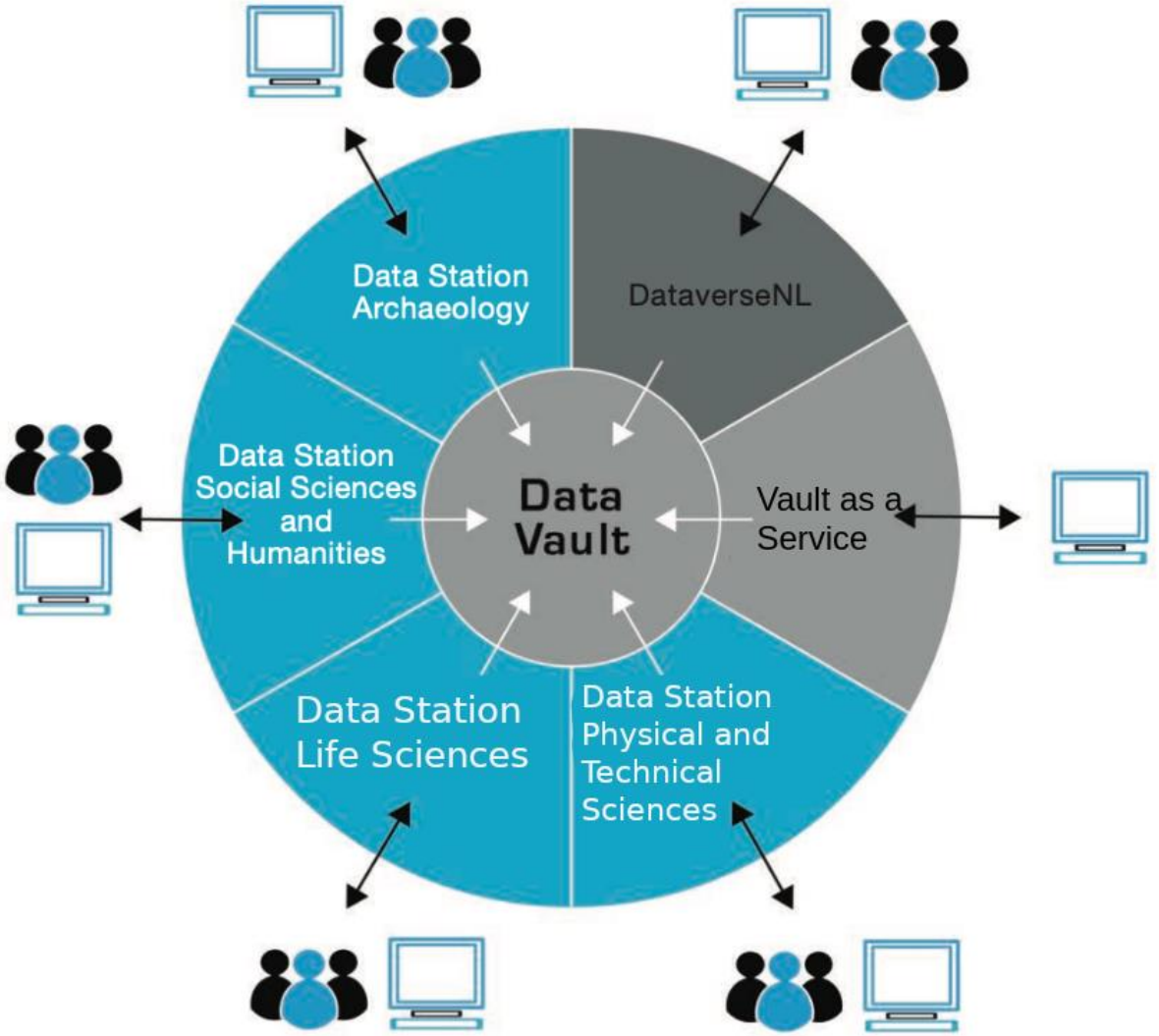
architecture

outer layer:

- deposit
- dissemination

Data Vault:

- preservation



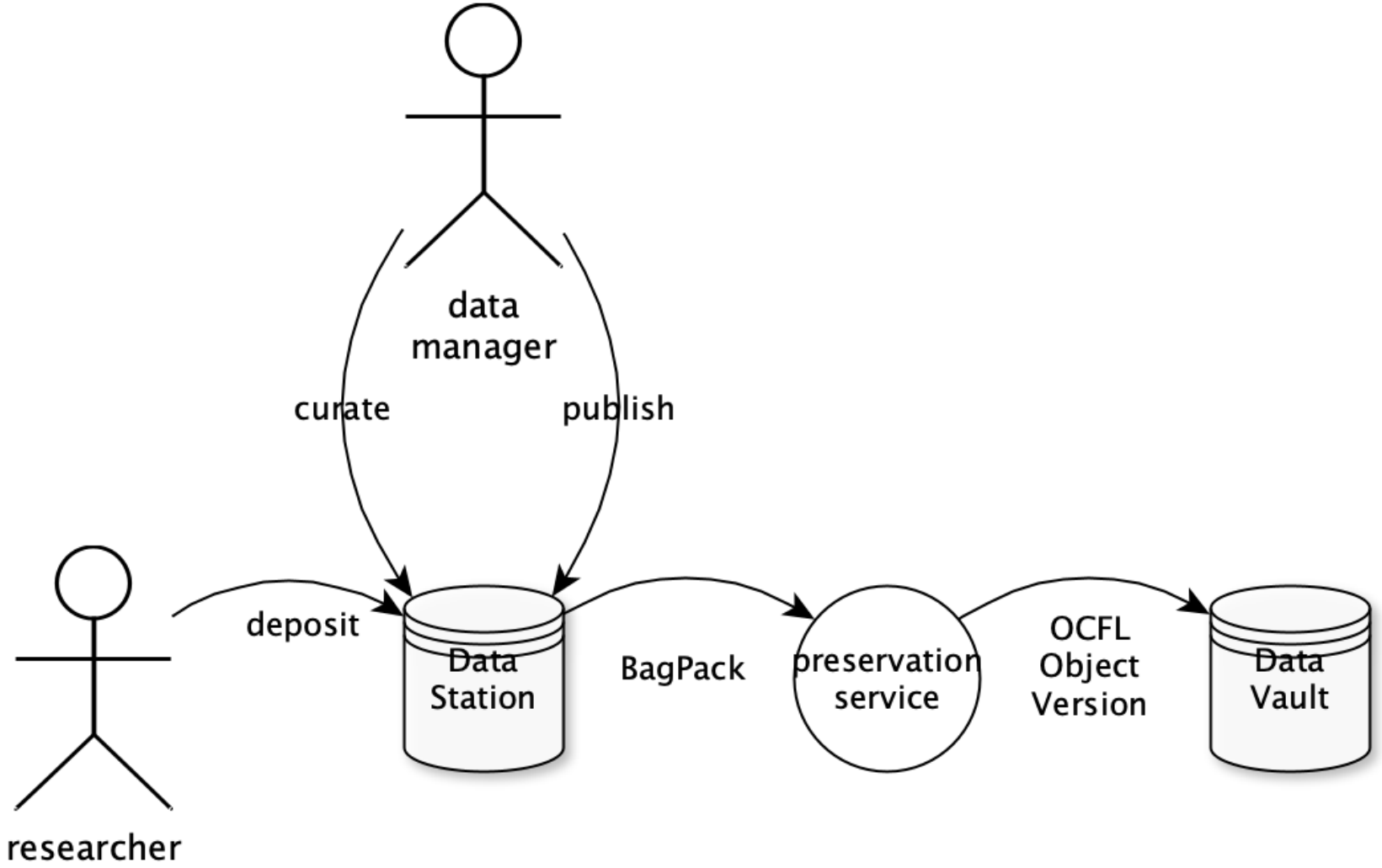
design principles

- modular by design
  - many small services instead of one giant system
- open standards
  - Dataverse, SWORD2, OCFL, BagIt/BagPack
- preservation-first mindset
  - long-term copies are a core function, not an afterthought
- future readability
  - storage root documented as an exchange format for future users

components

- repository (Dataverse)
- DANS microservices
  - ingest, validation, transfer, metadata, catalog, virus scan
- preservation
  - Data Vault + LOB store
- supporting infrastructure
  - NBN resolver, SURF Data Archive

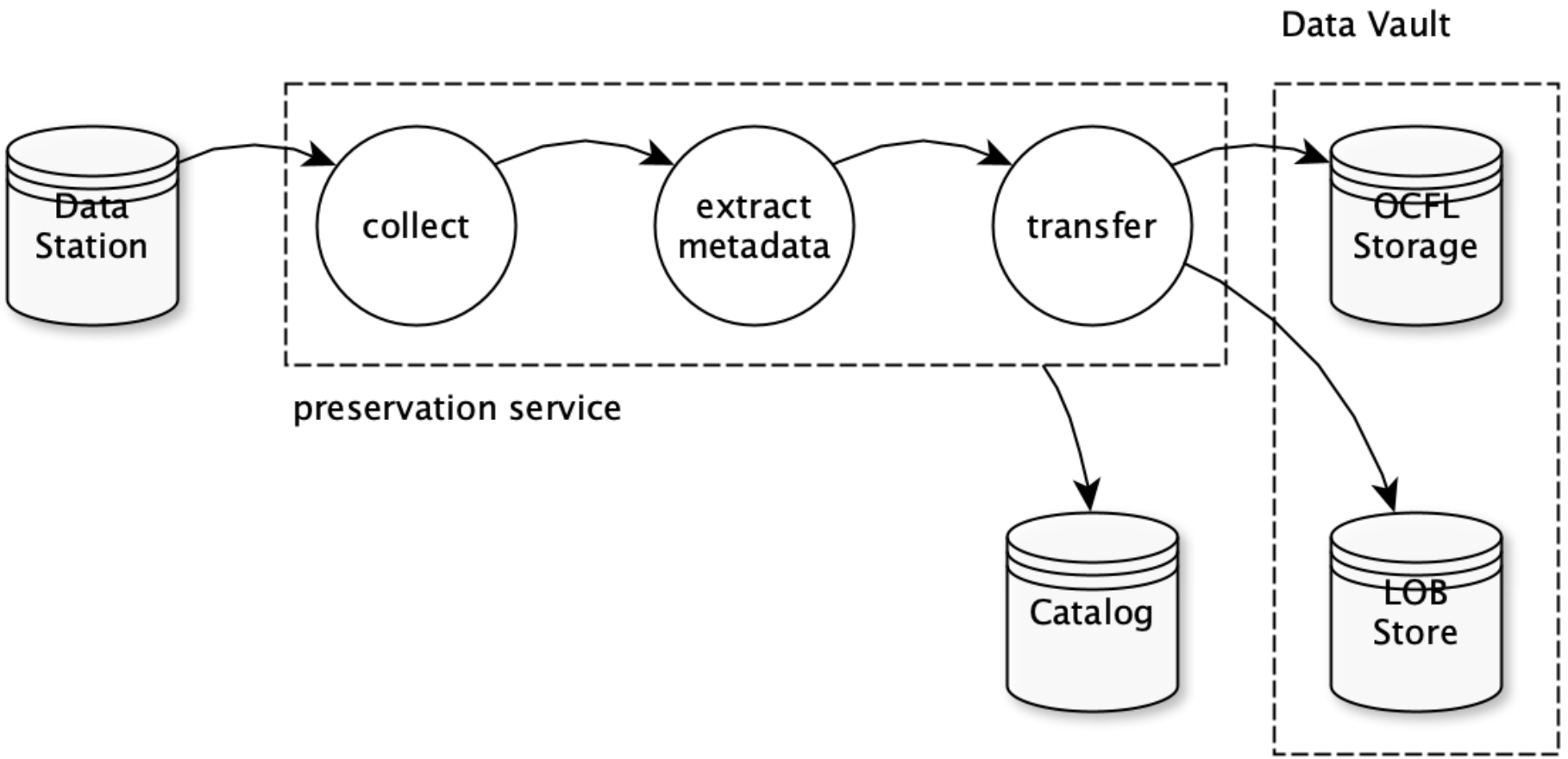
journey of a dataset



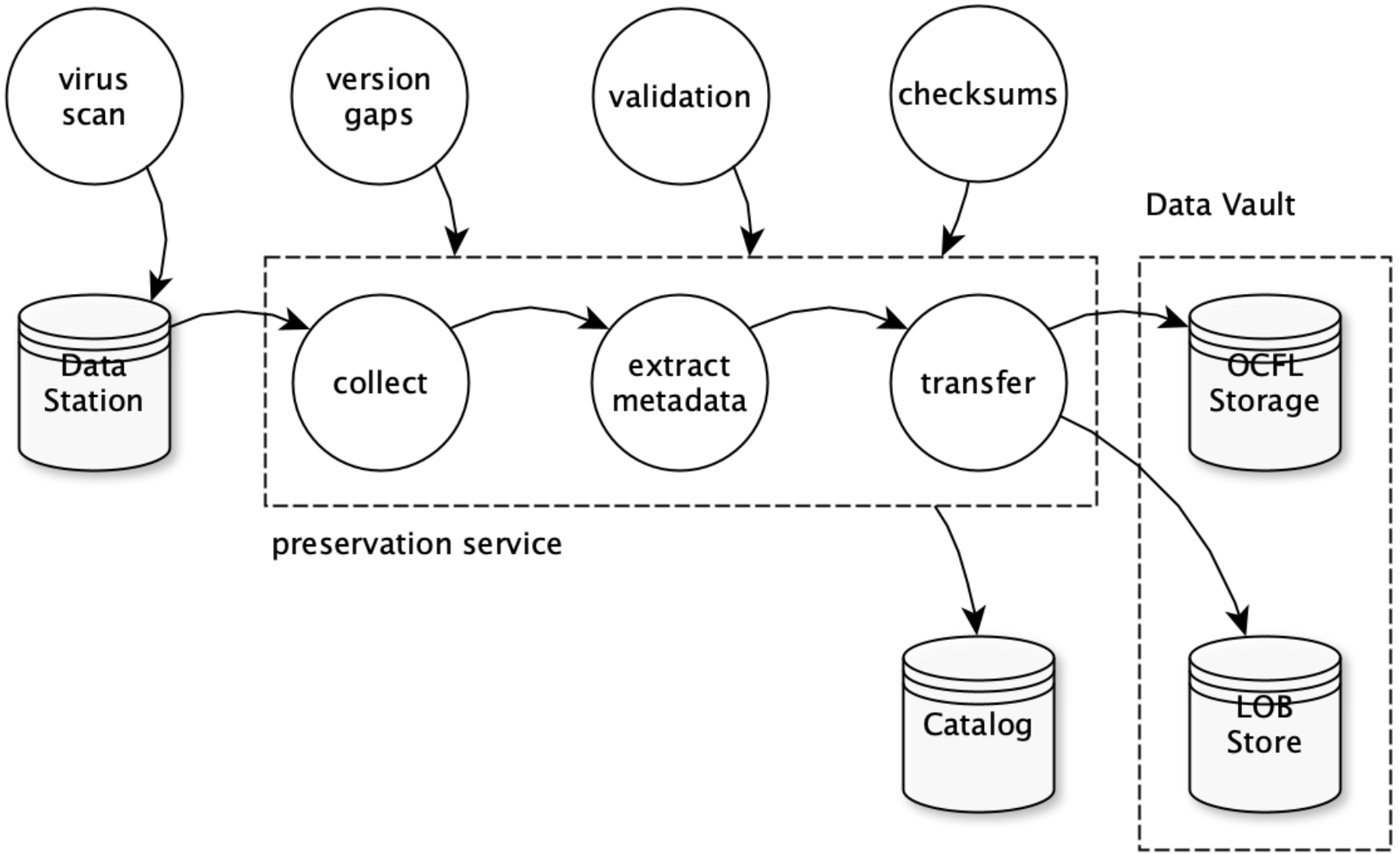
long-term preservation strategy

- Data Vault uses OCFL: versioned, structured archival objects
- Each dataset maps to one OCFL object so versions remain traceable
- Large files handled via companion Large Object Store (LOB Store)

more in-depth look of this journey



trust and quality controls



Data Vault using OCFL Storage

- Future-proof: uses open standards, not a hidden custom format.
- Restorable: datasets can be reconstructed in correct version order.
- Reliable at scale: very large files can be stored separately but still linked.

identifiers

- Datasets registered with NBN persistent identifiers
- Resolver points users/services to current location over time
- *Identifiers outlive platform changes*

Data Vault Catalog

DANS Data Vault Catalog summary for urn:nbn:nl:ui:13-881d42bd-c9fa-420b-bf52-626c66a37512

LISA Dutch Company Data

NBN	urn:nbn:nl:ui:13-881d42bd-c9fa-420b-bf52-626c66a37512
OCFL Storage Root	ssh
Dataverse PID	doi:10.17026/SS/AQNXQ9

v5

Title	LISA Dutch Company Data
Bag ID	urn:uuid:5e9dae2e-2299-4d01-b792-1664a90ba50f
OCFL Object version	5
Creation timestamp	2026-03-12T17:29:02.431+01:00
Archival timestamp	2026-03-12T17:34:25.837111+01:00
DV PID version	5.0
Exporter	Dataverse
Exporter version	6.9 build DANS-DataStation-PATCH-7

Filepath	SHA-1	File size
20240513 DataAccessProtocol_DANS_ErasmusUPT.pdf	b9138ca24195a29292b9f887e7bc51a5244ee0fb	148992
20250311 Erasmus UPT - request form LISA data.docx	b305cf9a5f918cd2fef5072da306e0423e6382dc	302100
Erasmus UPT - licence agreement LISA data.pdf	7b498b5acb02def07eb2d2c3624dd3465fef6369	145397
LISA_2022_sample.csv	2fb5807892a9b6106c94a83ba58bfd9a91621ab4	7850
LISA_2022_sample.rds	599397b859468ec1aa329aca699d1f6a589bfeca	11449
README.txt	67f420d5c8be6d655a51855b1967c7a639ae0320	1642

v4

Title	LISA Dutch Company Data
Bag ID	urn:uuid:df6b10e1-f798-4e52-b708-441d4a03ede3
OCFL Object version	4
Creation timestamp	2026-03-12T17:28:54.074+01:00
Archival timestamp	2026-03-12T17:34:22.366844+01:00
DV PID version	4.0
Exporter	Dataverse
Exporter version	6.9 build DANS-DataStation-PATCH-7

Service Variants for Different Needs

- **Data Stations**: full workflow incl. curation and dissemination
- **DataverseNL**: similar tech, different governance model
- **Vault as a Service**: direct archival pipeline for organizations with their own front-end workflows

keeping research data useful across decades

- Architecture combines usability, automation, and preservation
- Open standards reduce lock-in and support interoperability
- Modular services make evolution easier while keeping core trust guarantees

Thank you for your attention

DANS Open Day
Open data, open science

DANS

Anna van Saksenlaan 51 | 2593 HW The Hague | The Netherlands | +31 (0)88 003 46 66
National centre of expertise and repository for research data | An institute of the KNAW and NWO

DataLink LinkedIn BlueSky Mastodon www.dans.knaw.nl