

*7º Fórum de Gestão de Dados de Investigação | 7th Portuguese RDM Forum
November 17-18, 2020*

Workshop: Data repositories: services, use cases and certification

DataverseNO: A national repository for research data from Norway



Philipp Conzett

UiT The Arctic University of Norway

ORCID: <https://orcid.org/0000-0002-6754-7911>

Twitter: @PhilippConzett --- #forumgdi

Outline

- 1 **WHAT** is DataverseNO?
 - Basic features and facts
- 2 **WHY** DataverseNO?
 - Background and rationale
- 3 **HOW** does DataverseNO work?
 - Organization, structure, workflows including strengths and weaknesses, CoreTrustSeal certification process
- 4 Questions and discussion

1 WHAT is DataverseNO?

Basic features and facts

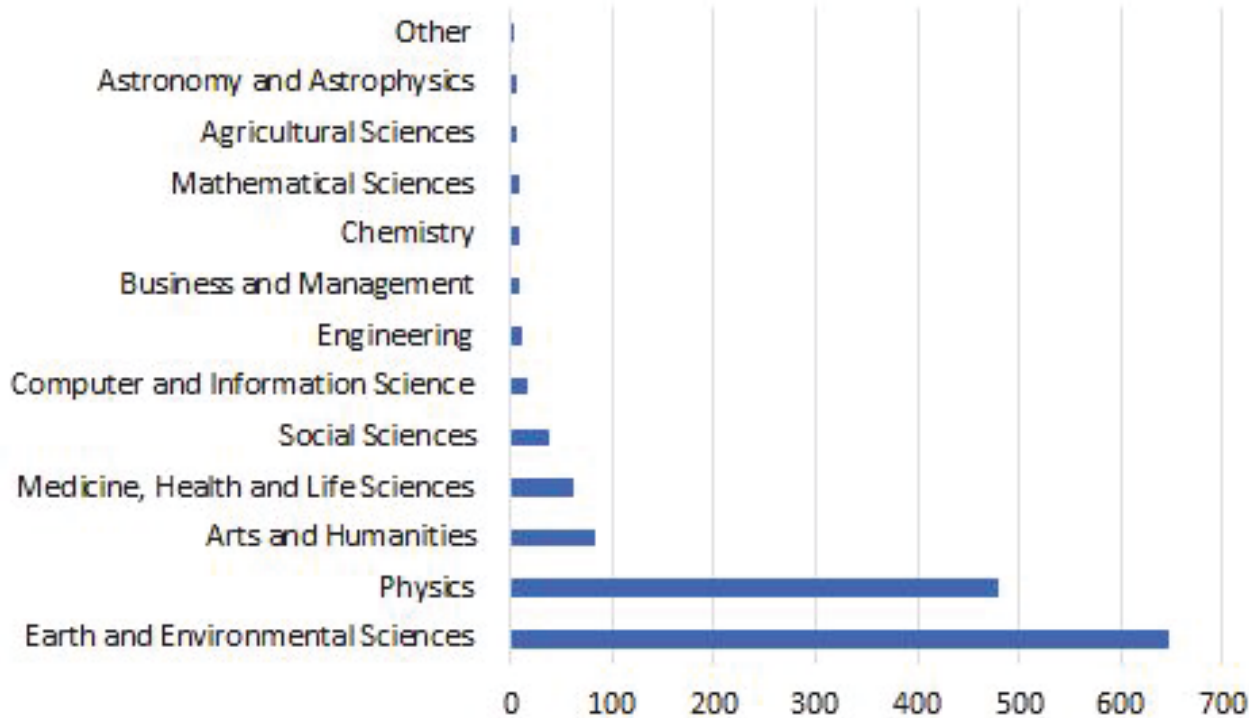
- ❑ a **national, generic** repository for **open** research data
- ❑ aligned with the **FAIR** principles and **CoreTrustSeal**-certified
- ❑ runs on the **Dataverse** software
- ❑ operated at **UiT The Arctic University of Norway**, and thereby
- ❑ the **northernmost** Dataverse repository in the world

... a **national** repository

- ❑ Currently **9 partner institutions**
- ❑ Universities and university colleges (all but 3 of the universities)
- ❑ Open for (individual) researchers from **other Norwegian research organizations**
- ❑ Contains currently data from researchers affiliated with **37 Norwegian organizations**



... a **generic** repository



- ❑ Data from **all domains** of science
- ❑ Graph shows distribution across domains
- ❑ High numbers within Physics and Earth Sciences are due to large time series.

Numbers as of 2020-11-18; Note: Many datasets are classified as belonging to more than one domain.

2 WHY DataverseNO?

WHY DataverseNO? --- Background and rationale

The Tromsø Repository of Language and Linguistics (TROLLing)

- Linguists at UiT needed a repository to share their data worldwide
- Result: **TROLLing** launched in 2014
- Based on Dataverse
- Part of CLARINO (Norw. node of CLARIN)

2015

UiT Open Research Data

- UiT needed an institutional repository for research data
- **UiT Open Research Data** established in 2015 based on experiences and solutions used in TROLLing

DataverseNO

- Other universities became interested in UiT Open Research Data
- UiT decided to expand service to become a national repository
- **DataverseNO** launched in 2017 (inspired by DataverseNL) with UiT and the University of Agder (UiA) as the first partners

2014

2017

3 HOW does DataverseNO work?

... basically as outlined in the DataverseNO
Organization Chart:

Organization of DataverseNO

ORGANIZATIONAL DOCUMENTS

DataverseNO Policy Framework

DataverseNO Steering Documents

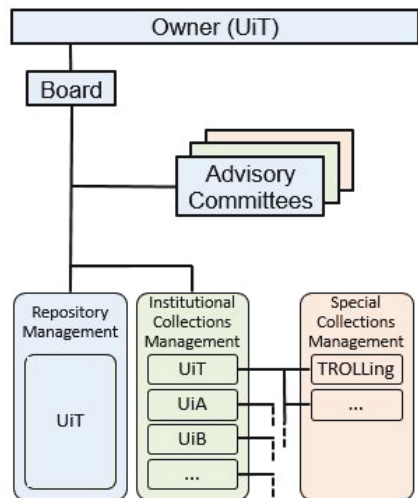
DataverseNO Administrator Guidelines

DataverseNO Curator Guidelines

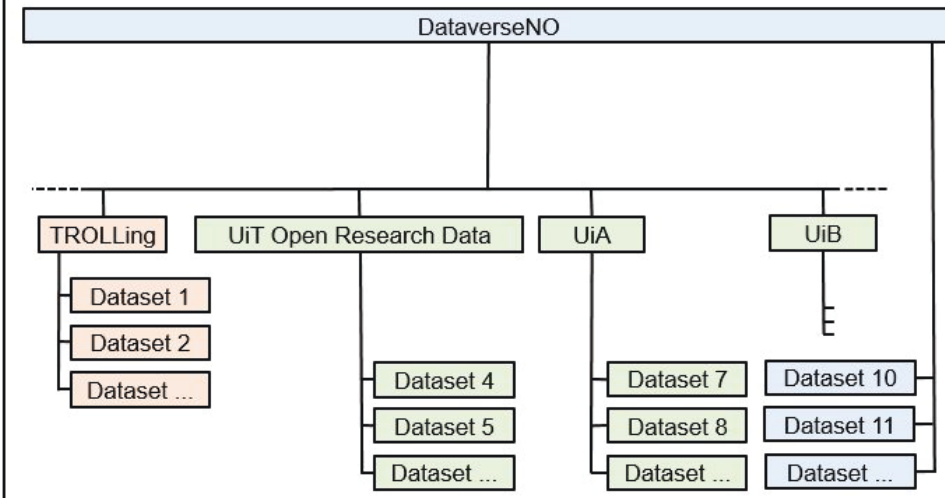
DataverseNO Deposit Guidelines

...

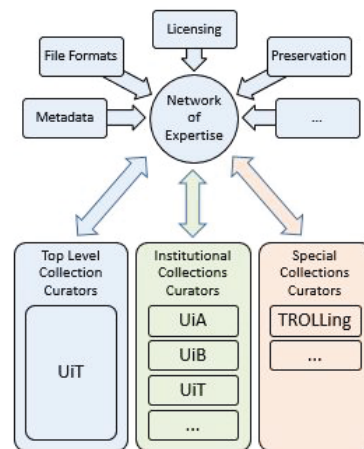
GOVERNANCE



REPOSITORY STRUCTURE



DATA CURATION



DESIGNATED COMMUNITY

Researchers within User Groups of Special Collections

Linguists

...

Researchers from Partner Institutions in Norway

HVL

INN

NMBU

NORD

NTNU

UiA

UiB

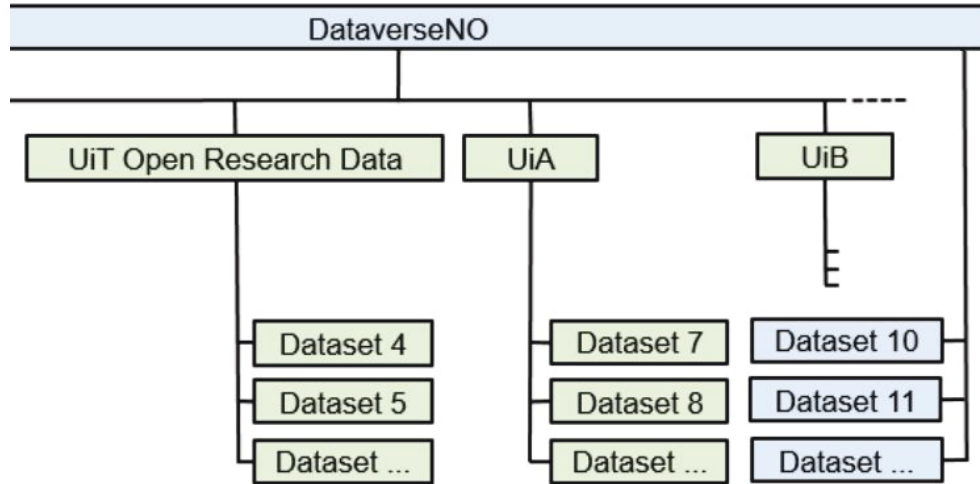
UIS

UiT

...

Researchers from Non-Partner Institutions in Norway

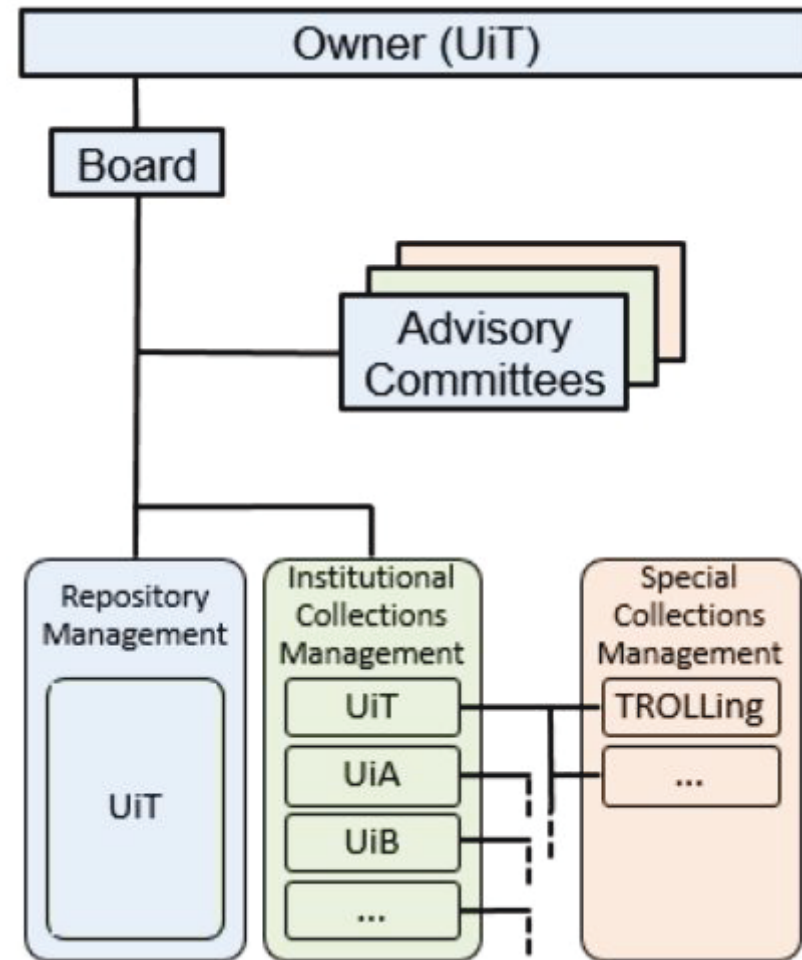
HOW? --- Repository structure



- ❑ Each partner institution has its own **collection** (sub-dataverse)
- ❑ Researchers from non-partner institutions publish their data in the top-collection

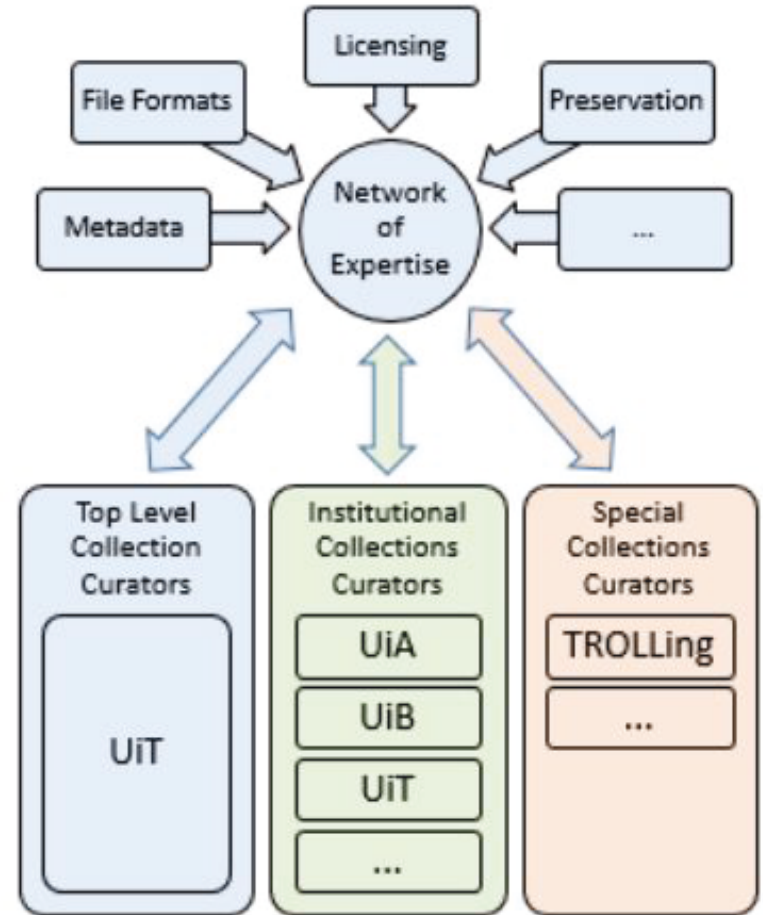
HOW? --- Governance

- ❑ UiT is responsible for management and development of **technical and functional core** of the repository and training and support for collections managers.
- ❑ Each partner is responsible for management of **institutional collection**.
- ❑ The **Board** of DataverseNO is the highest management body of the repository.
- ❑ An **advisory committee** discusses RDM and collaboration issues and gives advice to the board.



HOW? --- Data curation

- ❑ In order to make published data as FAIR as possible, each dataset is **curated** by research support staff at the partner institutions before publication.
- ❑ Only curators can publish datasets.
- ❑ Curators are usually subject librarians.
- ❑ Curators share knowledge and experience in **curator network** across partner institutions.



HOW? --- Organizational documents

The organization of DataverseNO is based on a set of **well-defined documents** describing the **responsibilities** of partners and the FAIR-aligned **stewardship of data** handled in the repository:

DataverseNO Policy Framework:

- ❑ Access and Use Policy
- ❑ Accession Policy
- ❑ Deposit Agreement
- ❑ Preservation Policy

DataverseNO Guidelines:

- ❑ DataverseNO Administrator Guidelines for Repository
- ❑ DataverseNO Administrator Guidelines for Collections
- ❑ DataverseNO Curator Guidelines
- ❑ DataverseNO Deposit Guidelines

DataverseNO Steering documents:

- ❑ Establishment of a Board for DataverseNO
- ❑ Mandate Board for DataverseNO
- ❑ Steering Document for DataverseNO

DataverseNO Partner Agreement (including data processor agreement):

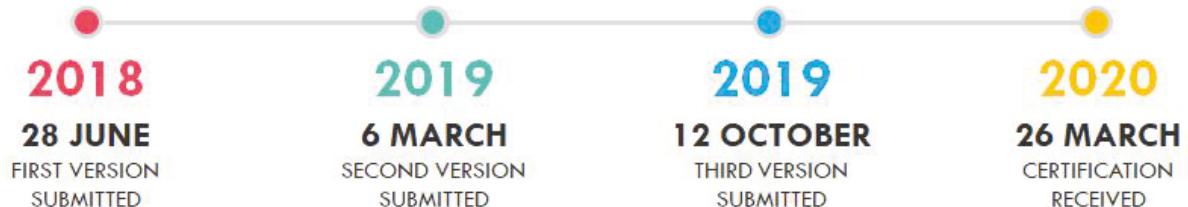
- ❑ Partners commit to manage their collections **according to DataverseNO policies and guidelines**

To demonstrate its **commitment to FAIR data stewardship and trustworthy and sustainable repository management**, DataverseNO has documented its approaches and workflows to obtain **CoreTrustSeal certification**.



CoreTrustSeal Certification

- ✓ in March, DataverseNO earned the **CoreTrustSeal**
- ✓ **long process**: about 2 years from start to certification
- ✓ **challenging issues** due to multi-institutional organization of DataverseNO
- ✓ one **common** set of **policies** and guidelines applied to all collections and datasets
- ✓ all datasets are **curated**



Strengths and weaknesses/challenges

Strengths:

- ❑ Based on **approved technical solutions**
- ❑ Provides **strong user support** co-located researchers embedded at partner institutions
- ❑ Part of **strong international collaborative networks**, e.g. Dataverse community (Harvard, DANS, ...), FAIRsFAIR, ...

Challenges:

- ❑ Many subjects/domains to be covered
- ❑ Challenging for **small organizations** to provide extensive user curation support

But: What are the alternatives in cases where no other, more appropriate (e.g. domain-specific) repositories are available?

Thank you for listening!



dataverse.no



@DataverseNO



info.dataverse.no

4 Questions and discussion