CoreTrustSeal certification of DataverseNO: Lessons learned and impact

Philipp Conzett
UiT The Arctic University of Norway
ORCID: https://orcid.org/0000-0002-6754-7911
Twitter: @PhilippConzett @DataverseNO   #dataverse2021

Outline of presentation

- ❏ Key facts about DataverseNO
- ❏ Benefits and impact of certification*
- ❏ Some lessons learned from certification process*
- ❏ Long-term evolution of the repository and its effect on certification*

* Thanks to Katie Mika for suggesting these topics as panel discussion questions at Open Repositories 2021.

Key facts about DataverseNO

Key facts about DataverseNO

DataverseNO …

- ❏ is a **national**, **generic** repository for **open** research data from researchers from Norwegian organizations;
- ❏ is **curated**, aligned with the **FAIR principles** (cf. Conzett 2020), and **CoreTrustSeal**-certified since 2020;
- ❏ runs on the **Dataverse software**;
- ❏ is operated at **UiT The Arctic University of Norway**, and thereby
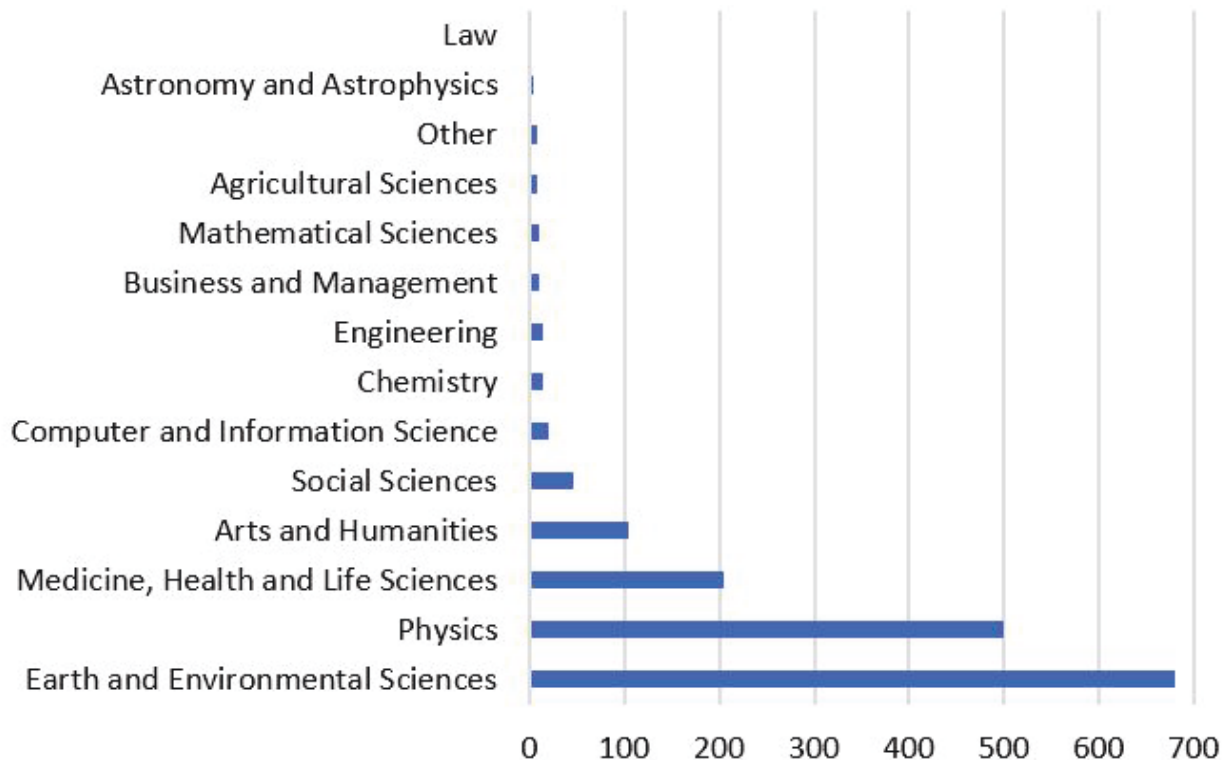- ❏ the **northernmost** Dataverse repository in the world.

... a national **repository**

- ❏ **Institutional Focused** (cf. Schlatter & Ji, 2021)
- ❏ Currently **10 partner institutions** (a new one has just joined us this month...)
- ❏ Universities and university colleges
- ❏ But also open for (individual) researchers from **other Norwegian research organizations**
- ❏ Contains currently data from researchers affiliated with approx. **40 Norwegian organizations**

DataverseNO institutions

... a generic **repository**

- ❏ Accepting data from **all domains** of science
- ❏ Graph shows distribution of currently published data across domains
- ❏ High numbers within Physics and Earth Sciences are due to large **time series**.
- ❏ Apart from time series: Mostly **background data** for publications.

Numbers as of May 15, 2021

Note: Many datasets are classified as belonging to more than one domain.

Benefits and impact of certification

How has CoreTrustSeal certification benefited our work?

❑ We started our certification process soon after we established the repository. The certification work was a good opportunity to **do some things right from the beginning** or at least in the early phase of our repository.

❑ In particular, the certification work forced us …

   ❑ to allocate time and resources to **address more fundamental issues** about our repository (and not only daily business issues);

   ❑ to take a **more integrated approach to decision making and documentation**; we started working on a policy framework rather than creating a lot of ad-hoc documentation whenever meeting a new challenge.

How has CTS improved the value of DataverseNO for users?

- ❏ For our **existing partner organizations**, the achieved CoreTrustSeal certification makes it possible to **promote the repository** as a trustworthy and sustainable service.
- ❏ For our **designated community**, that is researchers or depositors, the fact that DataverseNO is certified may give them some **confidence** that by publishing their data in the repository they will **meet** the **requirements** of funders and publishers.
- ❏ Certification makes DataverseNO **more attractive** for potential **new partner organizations**.

Some lessons learned from certification process

What did we learn from the application process?

❏ Also before we started our application process we had been focusing on best practice and international standards, but the certification process **revealed some gaps** in our approach, in particular the **need for explicit and transparent documentation**.

❏ To have a solid **technical infrastructure** in place is important for a trustworthy service, but trustworthiness and sustainability are to a great extent also about **organizational infrastructure** including long-term commitment and human resources.

Long-term evolution of the repository and its effect on certification

How may the long-term evolution of DataverseNO affect certification?

❏ We expect both **growth** and **increased variety** in terms of the number of **depositors** and **partner organizations**, and also in terms of served **disciplines** and deposited **data types**.

❏ This may imply some **scalability challenges**, which we'll need to address in a timely manner to keep meeting the CoreTrustSeal requirements. But I think given the **long-term commitment** by the partner organizations, I think we'll be well-placed to meet scalability challenges.

References

About DataverseNO. https://site.uit.no/dataverseno/about/.

Conzett, Philipp. 2020. «DataverseNO: A National, Generic Repository and Its Contribution to the Increased FAIRness of Data from the Long Tail of Research». *Ravnetrykk*, 39, 74–113. https://doi.org/10.7557/15.5514.

Schlatter, Tania & Jonathan Ji. 2021. Personas for software? How and why we created archetypes for installation of an open source product. Poster presented at The information architecture conference (IAC21). Available at https://drive.google.com/file/d/1SA2W7MKMRXTAzFrZmjVYM-E6o9tT1OQm/view?usp=sharing.

Thank you for listening!

dataverse.no

@DataverseNO

info.dataverse.no