

Link to Slides:

https://docs.google.com/presentation/d/1Z7PBTJaXKqWa4EDsVJ7bDHtnVpbUfuCfG5GCK_5QF_E/edit#slide=id.g5e441ba939_0_219

Appendix (Sensitive Data Support in Dataverse):

<https://docs.google.com/presentation/d/1EfHr8YjWblzEhHeUam82xtoqGcbUnZbK-FWxKMqvFF0/edit?usp=sharing>

(Participants: Please add your name at the [bottom](#) of this document.)

Gustavo: Introduction to Dataverse

- Open-source repository platform for publishing, citing, and archiving research data
- Development balances specific user needs and flexibility to accommodate broader community and use cases
- Began as virtual data center (VDC) in 2005, evolved into Dataverse in 2006
- Supports disciplinary domains beyond social sciences
- (Sub-)dataverses within dataverses, e.g. institutional ones like in DataverseNL og DataverseNO, project-based ones etc. = collections
- Branding, logo in Dataverse page header and footer
- Widgets to embed Dataverse iFrame into website
- Robust and granular permission systems / publishing workflows
- Private URL to share unpublished (versions) of dataset
- Upload and download
 - Browser
 - Dropbox
 - Rsync
 - *Remote storage (TRSAs - Trusted Remote Storage Agents) -- maintain data at its original storage (e.g. big data, sensitive data)*
- Interoperability
 - APIs
 - SWORD
 - Native
 - Metrics
 - Harvesting (OAI-PMH)
 - Client
 - Server
 - Modular External Tools
 - Explore vs. configure
 - Scope: Dataset / datafile
- Can introduce custom metadata blocks to add standardized rich descriptive metadata to datasets from specific disciplinary domains

- Request: Should make it easier to deploy metadata standards in Dataverse.
 - Philipp: Could add more domain-specific metadata schemas
- Dataverse Roadmap
 - Strategic goals
 - Implementation, planning, future

Dataverse Community

- 50 Dataverse installations around the world
- 101 contributors (documentation, bug fixes, development)
- Members include developers, researchers, librarians, data scientists
 - Dataverse Google Group to discuss and get answers to questions
 - Community calls
 - Community Meeting, usually in June at Harvard (bring soccer cleats), next one: June 17-19, 2020
 - Workshops & trainings
 - IU/UX testing & interviews
 - Global Dataverse Community Consortium
- Dataverse Ecosystem
 - Core Code + plugins (via SPIs) + External Systems (via APIs)
 - Modularity matters

Dataverse at the Odum Institute

- Services across the RDM lifecycle
- Domain repository at a research institution
- Why Dataverse for Odum?
- Were looking for tools for data that were designed to manage data
- Open source tools to leverage the community
- Needed robust software / repository community specifically focused on data around it
- Useful features, e.g. ingest of tabular files, actionable variable-level metadata
- Communities can add their extensible tools they need; a greater number of members are doing this
- Added services, levels of curation
- One of the first Data Seal of Approval repositories
- Computational reproducibility
- Looking at ways to manage person data across different institutions, through trusted remote storage agents (TRSAs)
- Dealing with sensitive data - DataTags (still in development)

DataverseNO

- National generalist repository for open research data in Norway to support Norwegian research institutions (<https://dataverse.no/>, <https://info.dataverse.no/>)
- Hosted by UiT - The Arctic University of Norway (the northernmost university in the world!)
- All data are curated by research data support staff at the 8 partner institutions, each of which has its own institutional collection (sub Dataverse) within DataverseNO
- A network of data curators are distributed across the partner institutions
- European Dataverse Workshop: 23-24 January 2020 at UiT, Tromsø (<https://site.uit.no/dataverseno/european-dataverse-workshop-2020/>)

Dataverse in the EU

- DANS-KNAW (Netherlands) for Dutch Universities
- SSHOC Dataverse project (CESSDA, DARIAH, CLARIN) to develop a production-ready Dataverse infrastructure
- Part of the SSHOC project: <https://www.sshopencloud.eu>
- Collaboration with TimeMachine project: <https://timemachine.eu>
- Services in the European Open Science Cloud (EOSC)
- EOSC requires level 8 of maturity
- “Dataverse App Store” - build different services out of tools
- Working on multilingual support. DataverseEU will run Weblate as a service.

Global Dataverse Community Consortium (GDCC)

- Wants also to reach out to communities like RDA
- Sharing code development, resources
- E.g. collaboration on DOIs from DataCite
-

-

Gustavo: DataTags

<http://datatags.org/>

See also

Input

- Francoise:

- Need to know if/how Dataverse wants to liaison with the RDA community
- E.g. in France there is a group interested in Dataverse
- Philipp: Encourage them to participate at the European Dataverse Workshop:
<https://site.uit.no/dataverseno/european-dataverse-workshop-2020/>

•

List of participants

- **Name, Affiliation, Email**
- Gustavo Durand, IQSS, Harvard University
- Philipp Conzett, UiT The Arctic University of Norway
- Thu-Mai Christian, Odum Institute, University of North Carolina at Chapel Hill
- Péter Király, Göttingen eResearch Alliance, pkiraly@gwdg.de
- Guillaume Brissebrat, CNRS/AERIS, gbrissebrat@ipsl.fr
- Geneviève Michaud, Sciences Po (FR) Genevieve.Michaud@sciences.po.fr
- Gabrielle Parent-Doliner, Swim Drink Fish, gabrielle@swimdrinkfish.ca
- Mohamed S. Yahia, Inist.CNRS, mohamed.yahia@inist.fr