

UiT

THE ARCTIC
UNIVERSITY
OF NORWAY

UiT Open Research Data

Helene N. Andreassen, Philipp Konzett, Stein Høydalsvik, Leif Longva, Obiajulu Odu
University Library, UiT The Arctic University of Norway
helene.n.andreassen@uit.no, philipp.conzett@uit.no, stein.hoydalsvik@uit.no,
leif.longva@uit.no, obiajulu.odu@uit.no

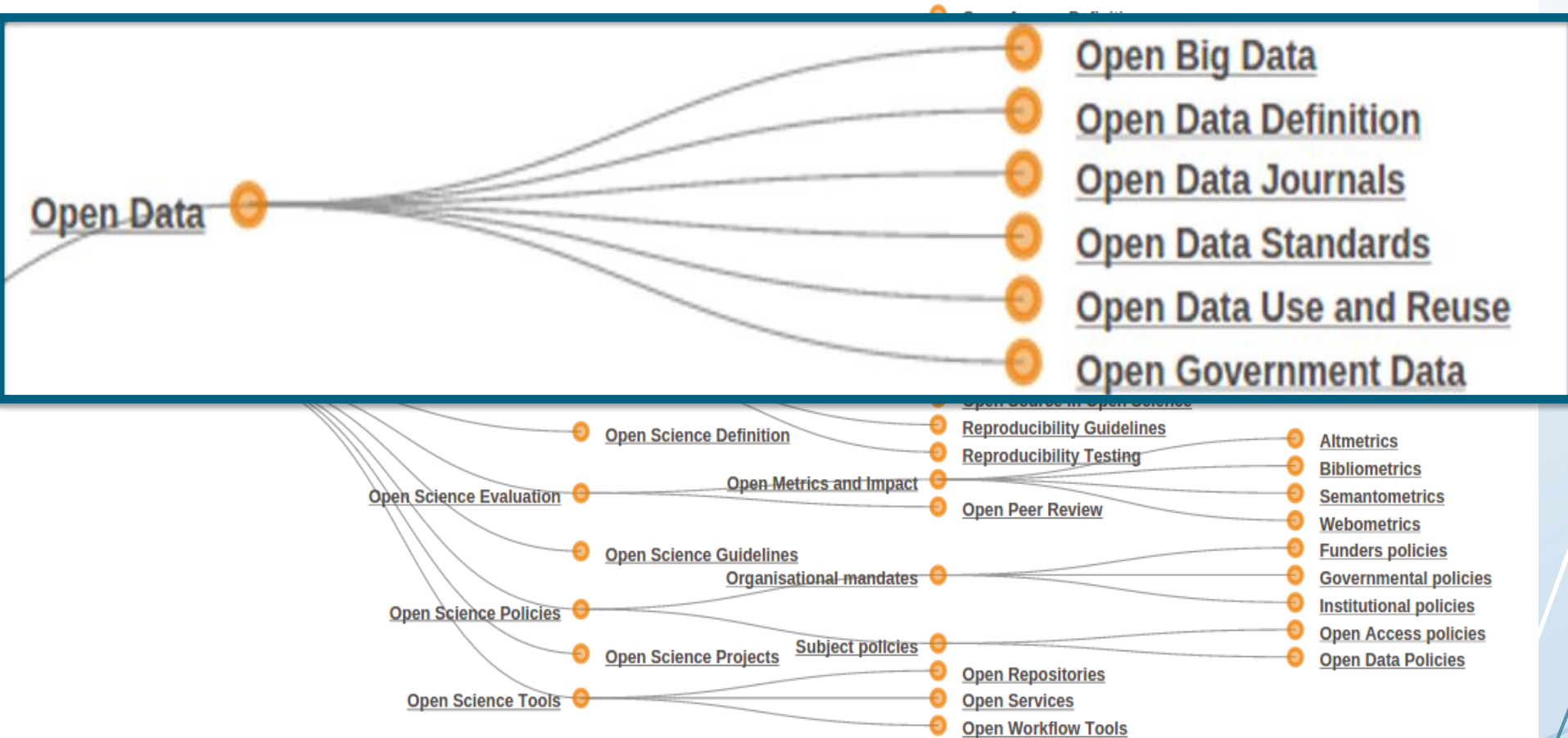
Emerging technologies in academic libraries - emtacl15 | 20-22 April 2015 | Trondheim



Open Data – part of Open Science

- The advantages of **Open Access (OA)** to research literature have been acknowledged for a long time.
- In recent years the focus has been widened to cover all scientific activities = **Open Science**.
- “Open science is the movement to make scientific research, data and dissemination accessible to all levels of an inquiring society [...]” (Wikipedia)
- Open Science includes
 - Open Access (to research literature)
 - Open Data
 - Open Science Evaluation
 - Open Science Policies
 - Open Science Tools
 - Open Education(FOSTER - Facilitate Open Science Training for European Research)
- Guidelines and (eventually) requirements from research funders to make science open

Open Science Taxonomy Tree



Expanding Open Access to Research Data

- Academic libraries have long experience in
 - handling research publications, and
 - providing OA to research literature
- But little experience in handling research data
- How may we, an academic library, be considered a reliable partner to take care of research data?
- We *have* proved ourselves trustworthy by
 - providing well-maintained institutional repositories
 - initiating and running publishing services for many years, and
 - actively promoting these and other OA services to the research community
- The research community members are thus aware of the library as an interesting partner for running new services

Collaboration with Research Community

- The UiT Library had for quite some time had plans to get started with open access to research data.
- In 2013 we were asked by the linguistics community at UiT to help them establishing (and running) a service for open access to linguistic research data and statistical code.
 - Just what we needed: A kick to get us started working on the issue of Open Research Data
- A project group was established:
 - The contributors listed for this presentation
 - Plus the initiating members from the linguistic community
- [TROLLing](#) – The Tromsø Repository of Language and Linguistics
 - Officially launched on 18th June 2014.

TROLLing - what

- Research data and statistical code in the field of language and linguistics
- Data contributors may come from anywhere in the world
- Limited to open research data
 - Archiving other types of data (e.g. sensitive data) imposes strict security requirements and need to be handled by other providers (e.g. Norsk samfunnsvitenskapelig datatjeneste NSD or Norstore)
- TROLLing is our starting point into open research data
 - A pilot to be followed by other academic disciplines and research communities at UiT
 - The platform must have a customizable metadata scheme
 - And otherwise be extendable to cover any research discipline

TROLLing - how

- First we needed to find a suitable tool or platform
 - Screening for available software
 - The platform needed to comply with international standards for metadata and metadata exchange (OpenAIRE+, DataCite, OAI-PMH).
- Chosen platform: [The Dataverse Network](#) software initially developed at Institute for Quantitative Social Science at Harvard University.

TROLLing and UiT Open Research Data

- TROLLing is part of something bigger (than itself):
- **UiT Open Research Data** (<http://opendata.uit.no>)
 - A general open access service for UiT:
 - To share, cite, preserve, discover, analyze and reuse research data
 - TROLLing is the first member (the first dataverse) of UiT Open Research Data
- Gradual extension of the service to cover other subject areas
 - Looking for the low hanging fruit: Which of UiTs research communities are in need of a service
 - We need to build success stories in order to promote the service

UiT Open Research Data

Share, Cite, Reuse, Archive Research Data
Scientific data for reproducible research

UiT Dataverse Network


    [Create Account](#) [Log In](#)

[Search](#)

[Advanced Search](#) [Tips](#)

Dataverses

1 Dataverses

 A **Dataverse** is a container for research data studies, customized and managed by its owner.

RECENTLY RELEASED DATAVERSES

[TROLLing](#)

As of 9th April 2015:

- 98 registered contributors from 32 different institutions (3 in Norway, 29 abroad)
- 25 uploaded datasets consisting of
- 191 data files
- 568 downloads

Studies

25 Studies, **191** Files, **568** Downloads

 A **study** is a container for a research data set. It includes cataloging information, data files and complementary files.

TUDIES

Old Church Slavonic by Eckhoff, Hanne; Haug, Dag	Mar 30, 2015
th of a construction: Old Church Slavonic touch verbs	Mar 30, 2015
h L3: Learners' production of liaison and 1, Helene N.; Lyche, Chantal	Mar 23, 2015
morphs of French de in coordination: a reproducible	Dec 29, 2014
sian nu-drop verbs by Makarova, Anastasia; Nessel,	Dec 10, 2014

[View More >](#)

TROLLing is special – or is it?

- TROLLing is an international service open to researchers from all over the world
 - New TROLLing contributors need to be approved manually
 - Scientific Advisory Board decides on strategic and executive issues.
- Expanding the service to other subject areas will likely be done at the institutional level
 - UiT Open Research Data will primarily invite the various research communities at UiT to make use of the service
 - However, we will listen to the needs and wishes of the interested UiT research communities

Metadata and data requirements

Best practice guides:

- Metadata requirements:
 - Sufficient metadata sets: Defining the metadata scheme, adapted to the needs of the individual dataverses
 - Metadata quality
- Data requirements:
 - Comprehensible data structure
 - Good description of the data sets, so others are able to understand what the data are – what they are measuring
 - Persistent file format (e.g. csv instead of xls)

The role of the library

- Curation carried out by subject librarians before publishing:
 - Contributor approval
 - Relevance of uploaded data
 - Quality and comprehensiveness of metadata
 - Structure of uploaded data
- In cooperation with the academic community at UiT:
 - How-to guides and data quality guide
 - Promotion: Conferences, contact with editors, video

Data management plan

- Do it properly – define your DMP
- All research projects should develop a data management plan, describing:
 - What data will be collected or generated in the project, and how?
 - What will be the data formats?
 - How will the data be handled?
 - What metadata and descriptions will you apply?
 - How will the data be kept safe during collection and research?
 - How will the data be preserved?
 - How will the data be made available for interested users?
- The library should build up competence on this in order to assist researchers

On the road to the future

- We gain very useful experiences through our TROLLing project
- Each separate dataverse (each collection) in UiT Open Research Data, within its academic fields, will have its own special requirements
- Competence building step by step, dataverse by dataverse
- An upcoming new member of UiT Open Research Data: Sami Language Technology at UiT
 - Corpus data
 - Tools and software for language analysis
 - In need of a data repository





Are we up to it?

- The library's needs for resources will reveal gradually, dataverse by dataverse
 - Future capacity challenges: Will we be able to serve all interested research groups at UiT?
- Management of open research data may become an important task for academic libraries
 - If we are up to it and show that we can cope with this role.
- Other related services and todos:
 - Research data management plans
 - Best practice guides
 - Teaching curricula for Research Data Management
- Raise awareness of the benefits of sharing scientific information in the academic community and among students, and thus contributing to the venture of Open Sciences

UiT Open Research Data

Share, Cite, Reuse, Archive Research Data
Scientific data for reproducible research

UiT Dataverse Network


    [Create Account](#) [Log In](#)

[Search](#)

[Advanced Search](#) [Tips](#)

Dataverses

1 Dataverses

 A **Dataverse** is a container for research data studies, customized and managed by its owner.

RECENTLY RELEASED DATAVERSES


TROLLing

May 19, 2014

[View More >](#)

Studies

25 Studies, **191** Files, **568** Downloads

 A **study** is a container for a research data set. It includes cataloging information, data files and complementary files.

RECENTLY RELEASED STUDIES

- | | |
|---|--------------|
|  Aspect and prefixation in Old Church Slavonic by Eckhoff, Hanne; Haug, Dag | Mar 30, 2015 |
|  Replication data for: Death of a construction: Old Church Slavonic touch verbs by Eckhoff, Hanne Martine | Mar 30, 2015 |
|  Norwegian L1 and French L3: Learners' production of liaison and enchainement by Andreassen, Helene N.; Lyche, Chantal | Mar 23, 2015 |
|  Replication data for: Allomorphs of French de in coordination: a reproducible study by Zuraw, Kie | Dec 29, 2014 |
|  Replication data for: Russian nu-drop verbs by Makarova, Anastasia; Nessel, Tore | Dec 10, 2014 |

[View More >](#)

opendata.uit.no