UIT Open Research Data and our experiences using Dataverse

Helene N. Andreassen, Philipp Conzett, Stein Høydalsvik, Leif Longva, Obiajulu Odu

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

Dataverse Community Meeting 2016
July 11, 12, 13 at Harvard Medical School
5 Campuses: Alta, Harstad, Hammerfest, Narvik, Tromsø
Background & Discipline specific collaboration

• Researchers at the Department of Language and Linguistics needed an OA platform where linguists worldwide can deposit, discover and share datasets and statistical methods
• The UiT Library has a long history of working with OA, since 1999
• Pilot project collaboration (2013 – 2014)
• Gave birth to TROLLing - The Tromsø Repository of Language and Linguistics
Institutional Data Archive collaboration

• New project between the IT Dept. and the Library (2015 - May 2016)

• Some major project goals:
  – software framework to enable UiT to deposit/discover/share research data
  – develop support services that will be needed

• Built upon the TROLLing experiences
Administrative organization - UiT ORD - cross institutional collaboration

- **UiT Library**
  - support services, metadata & quality standards, etc

- **UiT ORD Repository**

- **IT Dept.**
  - technical standard & data archiving infrastructures

- **Dept. of Research & Development (ADM)**
  - policy & strategy development for research data
What have we done?

- **Migration** – Have migrated ORD system from Dataverse version 3.6.2 to 4.3

- **Institution Log In** - Shibboleth single sign-on system is configured with Feide as Identity Provider (IDP).
  - Feide is a centralized identity management solution for the higher educational sector of Norway

- **Data citation** - DataCite DOI
  - migrated datasets have HDLs
What we have done & are still doing?
- Promoting

- Promoting the use of the UiT ORD archive to capture past and ongoing research work by PhD students & researchers at UiT
- Promoting the benefits of data sharing
- Promoting the archive on social media
- Raising awareness of best practices and trends in data management
What we have done & are still doing?
- *Workshops*

- Tailoring workshops to walk researchers/students through the submission process, etc
  – Workshops are well attended and we are consistently receiving positive reviews
What we have done & are still doing?
- Surveying

• Conducting institutional survey of faculty students and researchers to gather information about existing perspectives and practices managing research data
What we have done & are still doing?
- competence building

- Curations are currently carried out manually by the subject librarians
- Improving skills and competencies among staff:
  - Seminars
  - Conferences
  - Workshops
  - Meetings
Improvements in the realm of Dataverse 4.3?

- It is important that Dataverse distinguishes between open data and data with access restrictions, so that harvesting services can harvest open data in the global open access infrastructures
  - Proper identifications (via metadata & file-tagging)
- Below are some areas we would like some changes made & new functionalities added
## Some improvement areas

<table>
<thead>
<tr>
<th>Suggestion</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>System upgrade should be easier from 4.x (cf. upgrade from v3 to v4).</td>
<td>System</td>
</tr>
<tr>
<td>Apply two or more licenses on one dataset. May be useful for datasets</td>
<td>License</td>
</tr>
<tr>
<td>containing aggregated data from different sources. But be aware of</td>
<td></td>
</tr>
<tr>
<td>license stacking!</td>
<td></td>
</tr>
<tr>
<td>Configuration of metadata schemas and search fields: System should</td>
<td>Metadata</td>
</tr>
<tr>
<td>prevent selection of search field with no underlying metadata fields.</td>
<td></td>
</tr>
</tbody>
</table>
## Some improvement areas

<table>
<thead>
<tr>
<th>Suggestion</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy way to move datasets from root to a sub-dataverse after they have been created.</td>
<td>Administration</td>
</tr>
<tr>
<td>Option for choosing whether file containers (zip, tar) should be unpacked or not.</td>
<td>Data ingest</td>
</tr>
<tr>
<td>When downloading datasets with restricted files, &quot;Request access&quot; is also shown for data files without restrictions.</td>
<td>Download</td>
</tr>
<tr>
<td>Suggestion</td>
<td>Domain</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Choose between different, pre-defined licenses</td>
<td>License</td>
</tr>
<tr>
<td>Automatic embargo handling for files with restrictions.</td>
<td>Access</td>
</tr>
<tr>
<td>Metadata field for link to article in institutional repository in addition to published version.</td>
<td>Metadata</td>
</tr>
</tbody>
</table>
## Some improvement areas

<table>
<thead>
<tr>
<th>Suggestion</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Option for adding tags to ALL the files in current upload batch, i.e. before saving the files.</td>
<td>Metadata</td>
</tr>
<tr>
<td>Advanced metrics, e.g. according to country</td>
<td>Metrics</td>
</tr>
<tr>
<td>Make Dataverse remember previously used/visited sub-dataverse after visiting user account information page.</td>
<td>Navigation</td>
</tr>
<tr>
<td>Get files listed in result list when searching on file level.</td>
<td>Search</td>
</tr>
</tbody>
</table>
Learn more …

UiT Open Research Data
support transparent and reproducible research

Archive: opendata.uit.no
Blog: site.uit.no/opendata
Contact: opendata@ub.uit.no
Facebook: facebook.com/ubuit/posts/1522345541350852
Youtube: youtube.com/channel/UCA_5WIRImOgx9cWMr7gLljQ
youtube.com/watch?v=lCJw-Qd3ZzgLaura janda
Thanks for listening!

«Open research data is a natural step towards open science»
Further Links/References

• UiT Open Research Data
  [link](http://hdl.handle.net/10037/7879)

• TROLLing: defining, building, and operating an open archive for linguistic data
  [link](http://hdl.handle.net/10037/9293)

• Pilotprosjekt infrastruktur for forskningsdata ved UiT (in Norwegian) -
  [link](http://dx.doi.org/10.7557/sr.2016.4)

• Feide - [link](https://www.feide.no/)

• Dataverse - [link](http://dataverse.org/)