

# From synchronisation to curation

Research data stored, shared, published and preserved

*Frederik Orellana, DeIC/DTU*

*Lars Christensen, DeIC/DTU*

*In collaboration with Lars Holm Nielsen, CERN/IT*

# Before I start...

---

- ❖ Still in conceptualisation phase
- ❖ Input/ideas very welcome

# What we have

---

- ❖ **data.deic.dk** - national (Danish) ownCloud-based data service for researchers
  - Intended for 'active' research data
  - Not for long-term archival
- ❖ ownCloud metadata app - intended for tagging and organising datasets

Places ▾

- Home
- documents
- Shared by me
- Shared with me
- Trashbin

Tags ▾

- mytag
- music

Apps ▾

- Metadata
- Teams
- Pictures
- Importer
- Uploader



[Upload](#) [New ▾](#)

<input type="checkbox"/>	Name ▲	Size	Modified
<input type="checkbox"/>	📁 Data	0 kB	3 months ago
<input type="checkbox"/>	📁 documents	6.3 MB	last month
<input type="checkbox"/>	📁 music	3.6 MB	9 months ago
<input type="checkbox"/>	📁 photos	663 kB	9 months ago
<input type="checkbox"/>	📁 testfolder	0 kB	6 days ago
<input type="checkbox"/>	📄 IPMIViewSuperBlade.pdf <span>mytag</span>	1.3 MB	last month
<input type="checkbox"/>	📄 msymdoc.pdf	1.8 MB	last month
<input type="checkbox"/>	📄 ownCloudUserManual.pdf	1.8 MB	9 months ago
5 folders and 3 files		15.4 MB	

Places ▾

- Home
- documents
- Shared by me
- Shared with me
- Trashbin

Tags

- mytag
- music

Apps

- Metadata**
- Teams
- Pictures
- Importer
- Uploader

New tag

Name	Color	On left menu	Public	Files
mytag		<input checked="" type="checkbox"/>	<input type="checkbox"/>	1
sound		<input type="checkbox"/>	<input type="checkbox"/>	0
				0
				2

Tag name:

Meta-data fields:

 ✕

✕

Add meta-data field

OK Cancel



# What CERN has

---

- ❖ **Zenodo** - an archiving, publishing and sharing service for documents *and* research data
- ❖ Open to researchers from all member states
- ❖ Easy and free assignment of Datacite DOIs
- ❖ Based on well-engineered, scalable, well-maintained infrastructure - EOS data store, tape backup, active research on bit-rot prevention etc.
- ❖ Long-term commitment and general trust with researchers



Search

Communities

Browse ▾

Upload

Get started ▾

frederik.orellana@cern.ch ▾

 Search

## Filter by types ▾

## Publications (27288)

Books (2115) Book sections (0) Conference papers (9228) Journal articles (11391) Patents (2) Preprints (188)

Project Deliverables (48) Project Milestones (1) Proposals (17) Reports (572) Theses (209)

Technical notes (55) Working papers (172) Other (787)

## Posters (406)

## Presentations (797)

## Datasets (1785)

## Images (397)

Figures (330) Plots (7) Drawings (11) Diagrams (24) Photos (17) Other (8)

## Videos/Audio (86)

## Software (6020)

## Lessons (54)

## Using GitHub?



Check out our GitHub integration. Software Preservation Made Simple!

## New to Zenodo?

- **Research. Shared.** – all research outputs from across all fields of science are welcome!
- **Citeable. Discoverable.** – uploads gets a Digital Object Identifier (DOI) to make them easily and uniquely citeable.
- **Community Collections** – accept or reject uploads to your own community collections (e.g workshops, EU projects or your complete own digital repository).
- **Funding** – integrated in reporting lines for research funded by the European Commission via OpenAIRE.
- **Flexible licensing** – because not everything is under Creative Commons.

## Recent Uploads

20 October 2015

Presentation

Open access

View

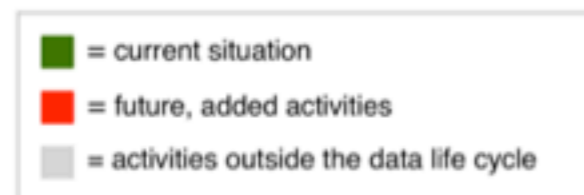
## Big and Smart Data Analytics - Possible Advantages to Clinical Practice

Di Meglio, Alberto ; Manca, Marco

The growing availability of large quantities of data from medical devices, laboratory notes, digital simulations, doctors' notes, images, and even social networks are opening

# Data lifecycle

---





# The vision

---

- ❖ Promoting:
  - ❖ data reuse, discoverability and provenance
  - ❖ open access
  - ❖ dataset ‘citability’ - credits to valuable datasets
- ❖ CERN/Zenodo acting as ‘terminus’ for end-of-lifecycle European research data
- ❖ Regional/volatile data stores get *seamless*, long-term publication channel
- ❖ Curation is *not* done by CERN, but by regional data stores

Places ▾

- 🏠 Home
- 📁 documents
- 👤 Shared by me
- 👤 Shared with me
- 🗑️ Trashbin

Tags ▾

- 📌 mytag
- 📌 music

Apps ▾

- 📌 Metadata
- 👤 Teams
- 📷 Pictures
- 📥 Importer
- 📤 Uploader

**Upload** New ▾

<input type="checkbox"/>	Name ▲	Size	Modified
<input type="checkbox"/>	📁 cernfolder	105.1 MB	11 minutes ago
<input type="checkbox"/>	🔗 📁 documents	23 kB	3 months ago
<input type="checkbox"/>	📁 hello	0 kB	13 days ago
<input type="checkbox"/>	📁 music	3.6 MB	3 months ago
<input type="checkbox"/>	📁 photos	663 kB	3 months ago
<input type="checkbox"/>	📁 testfolder	< 1 kB	7 days ago
<input type="checkbox"/>	📄 owncloud	12 MB	last month
<input type="checkbox"/>	📄 owncloudcmd <span style="border: 1px solid green; padding: 2px;">mytag</span>	646 kB	last month
<input type="checkbox"/>	📄 ownCloudUserManual.p	1.8 MB	3 months ago
<input type="checkbox"/>	📄 sync-exclud	< 1 kB	last month
<input type="checkbox"/>	📄 test.txt	< 1 kB	28 days ago
6 folders and 5 files		123.8 MB	

zenodo

ORCID: 0000-0001-8135-3489 (choose a different ID?)  
 Zenodo token: none yet

Publish in sandbox
Publish in production

# The complications - TBD

---

- ❖ **ownCloud ORCID app**

- ❖ Assign ORCID ID to ownCloud account
- ❖ Login with ORCID

- ❖ **ownCloud Zenodo app**

- ❖ Generate archive from selected files/folders
- ❖ Integrate with ownCloud metadata app (create deposition metadata schema)
- ❖ Upload to Zenodo

- ❖ **Zenodo API extensions**

- ❖ DOI assignment - primary and secondary URL
- ❖ Delegate Zenodo access to ownCloud server

# Conclusion

---

- ❖ No observed show stoppers (?) :-)
- ❖ Plenty of goodwill
- ❖ High priority and manpower assigned on both ends

# Proposal: delegating Zenodo access to ownCloud server

---

When clicking “Publish on Zenodo” on data.deic.dk:

- ❖ data.deic.dk will redirect to  
[https://orcid.org/oauth/authorize?...&client\\_id=0000-0001-8135-3489&redirect\\_uri=https://zenodo.org/oauth/authorized/orcid/&scope=/authenticate&state=https://data.deic.dk/index.php/apps/files\\_zenodo?datatoken=...](https://orcid.org/oauth/authorize?...&client_id=0000-0001-8135-3489&redirect_uri=https://zenodo.org/oauth/authorized/orcid/&scope=/authenticate&state=https://data.deic.dk/index.php/apps/files_zenodo?datatoken=...)
- ❖ orcid.org will ask to log in or create account, then redirect to  
[https://zenodo.org/oauth/authorized/orcid/&scope=/authenticate&state=https://data.deic.dk/index.php/apps/files\\_zenodo?datatoken=...](https://zenodo.org/oauth/authorized/orcid/&scope=/authenticate&state=https://data.deic.dk/index.php/apps/files_zenodo?datatoken=...)
- ❖ zenodo.org will match with, ask to log in or create account, and redirect to  
[https://data.deic.dk/index.php/apps/files\\_zenodo?datatoken=...&token=...](https://data.deic.dk/index.php/apps/files_zenodo?datatoken=...&token=...)
- ❖ data.deic.dk will use the token to upload to zenodo.org