Zenodo:
Making Open Data and Open Software an Achievable Reality for all Scholars

Tim SMITH

Data and computing infrastructures for open scholarship
NOT Just Another Data Repository

zenodo

INVENIO

120 PB
120,000,000 GB
120,000,000,000,000,000 Bytes
Why?

Open Access Mandate
Open Data Pilot
Scholarship isn't Linear Anymore

Proposal
    Capture
        Processing
            Analysis
                Write-up
                    Publish
                            Share
                                Archive
                                        Disseminate
Open Scholarship

- Proposal
- Data Capture
- Processing
- Analysis
- Write-up
- Disseminate

Share
Publish

Archive
Preserve

Share
Publish

Archive
Preserve

IP[y]: Interactive Computing

Google docs
R
OpenSci

Authorea
ShareLaTeX

PLOS
arXiv.org

Other Tools:

GitHub
Git
Subversion

Zenodo

Dataset
Software
Image
Publication
Video/Audio
Poster

Presentation
Lesson
Share ≠ Publish ≠ Preserve

Collaborative development
- Descriptive programmatic logic overview
- Contributor list
- Documentation
- Unit and regression tests

Publication
- Narrative
- Standardized terminologies +

Code Share Repository

Archive / Research Repo

GitHub
sourceforge
Bitbucket
Sphinx
Read the Docs
Software Sustainability Institute
Zenodo – GitHub bridge

Languages of Software Published in Zenodo

API lego

```json
{
    "name": "Pfehn, Tilman",
    "affiliation": "Institut für Theoretische Physik",
    "description": "This repository contains the software for the decoupling and recoupling of two coupled oscillators."
}
```
Open APIs for Interoperable Services

A Ruby wrapper for the Zenodo API. https://zenodo.org/api/. Zenodo is a simple and innovative service that enables researchers, scientific institutions to share and showcase multidisciplinary publications that are not part of the existing institutional research communities.

PyRDM is a Python-based library for research data management (RDM). It facilitates the automated publication of scientific software and associated input and output data.

https://sandbox.zenodo.org
Closed to be Open

- Publication moment is too late!
- Capture before the knowledge evaporates
  - While data is live
- Share unpublished data with reviewers
- Share sensitive data with controlled list

- Open ≠ Simple
  - Logins & Roles
  - Fine grained ACLs
  - Sophisticated security model
  - Life-cycles and workflows
Choice

• Scholarly Communication Futures debate
  – Ideal services, business models, custodians?
• University, National, Subject, Memory Institute
  – Obligation, Trust

• EUDAT installable instances
  – Local and customised services for communities

• OpenAIRE – EUDAT federated instances
  – Safe, not anchored
Conclusions

• Deconstructing Open Scholarship
  – Interoperable elements
  – APIs
• Data not anchored, owned

• Separate creative, engagement, sharing services

• Safe underlying repositories
  – Proffer a scientific memory institution as a viable alternative
Thank You! Questions?

Tim.Smith@cern.ch
0000-0002-1567-7116
http://cern.ch/tim.smith

http://zenodo.org
http://www.openaire.eu
http://invenio-software.org
https://b2share.eudat.eu
Lars.Nielsen@cern.ch
CERN’s discoveries already have a considerable impact on society, from pure knowledge to cutting-edge technological applications like the World Wide Web, touch screens, and PET scanners.

With your support, we can do more to spread the CERN spirit of scientific curiosity for the inspiration and benefit of society.

CERN has developed a programme of projects made possible by philanthropic support, with activities in these areas: