



Introduction to Zenodo

Tibor Šimko

@tiborsimko

CERN-UNESCO School on Digital Libraries, Kumasi, Ghana · November 2016

Why preserve?



<https://www.flickr.com/photos/dmh650/4031607067/in/gallery-wlef70-72157633022909105/>

Why carefully?



<https://www.flickr.com/photos/herzogbr/6756173595/in/gallery-wlef70-72157633022909105/>

Why remotely?



http://commons.wikimedia.org/wiki/File:Brand_bouwkunde_-_TU_Delft_-_13_Mei_2008.jpg

User habits


20%
store data
in a digital archive

1,000,000(,000+) GBs



What to archive?



 Mars Mission, Viking 1 on Titan III Centaur Rocket Launch
NASA Langley Research Center 8/27/1975 Image # EL-1996-00220

“Our goal should be to save everything.
Without tomorrow’s context
we don’t know what is valuable today.”
– Arthur I. Zygielbaum, NASA

Publish or perish culture

Hard



54%



40%



24%

No Credit



92%

zenodo [Upload](#) [Communities](#) [Log in](#) [Sign up](#)

Recent uploads

November 24, 2016 [Presentation](#) [Open Access](#)

[View](#)

Wissenschaftliche Software pflegen – aber wie?

Zeller, Andreas

Präsentation im Rahmen vom Helmholtz Open Science Workshop „Zugang zu und Nachnutzung von wissenschaftlicher Software“ #hgfos16 am Helmholtz-Zentrum Dresden-Rossendorf am 22. und 23. November 2016.

Uploaded on November 24, 2016.

November 24, 2016 [Software](#) [Open Access](#)

[View](#)

PIConGPU 0.2.0: Beta Release - Full Multiple Species Support & openPMD

Axel Huebl; René Widera; Alexander Grund; Richard Pausch; Heiko Burau; Alexander Debus; Marco Garter; Benjamin Wortitz; Erik Zenker; Frank Winkler; Carlichristian Eckert; Stefan Tietze; Benjamin Schneider; Maximilian Knespe; Michael Bussmann

This release of PIConGPU, providing "beta" status for users, implements full multi-species support for an arbitrary number of particle species and refactors our main I/O to be formatted as openPMD (see <http://openPMD.org>). Several major features have been implemented and stabilized, highlights ...

Uploaded on November 24, 2016.

Sep 12: Major update



Welcome to the improved Zenodo. See [what's new](#) and [known issues](#).

Using GitHub?



Just [Log in](#) with your GitHub account and [click here](#) to start preserving your repositories.

Zenodo in a nutshell

- **Research. Shared.** – all research outputs from across all fields of research are welcome! Sciences and Humanities, really!
- **Citeable. Discoverable.** – uploads gets a Digital Object Identifier (DOI) to make them easily and uniquely citeable.
- **Communities** – create and curate your own

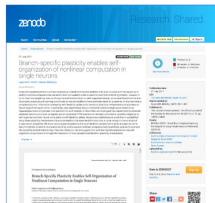
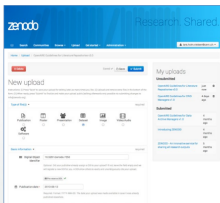
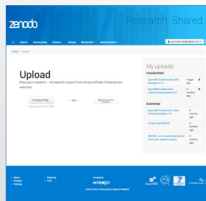
<https://zenodo.org/>

Step 1: Capture research data

 Upload

 Describe

 Publish



Upload

The image shows a screenshot of the Zenodo website's upload interface. The main heading is "Upload" with a large blue arrow icon. The Zenodo logo and "Research. Shared" are visible at the top. The page content includes a "Choose files..." button with a note "Maximum file size 2GB", a "My uploads" section, and a "Selected files" table. A "Choose from Dropbox" button is also present. Callout boxes highlight the "Choose files..." button, the "Sign in with GitHub" and "Sign in with ORCID" buttons, and the "Choose from Dropbox" button. A URL "http://www.dropbox.com" is shown below the Dropbox button.

zenodo
Research. Shared

Upload

zenodo Research. Shared

Home / Upload

Upload

Share your research — all research outputs from across all fields of science are welcome.

My uploads
You currently have no uploads.

Choose files...
Maximum file size 2GB

Choose files...
Maximum file size 1GB

— OR —

Choose from Dropbox

Selected files

4198_201.pdf	12 KB	0
--------------	-------	---

Sign in with GitHub

Sign in with ORCID

Dropbox

Choose from Dropbox

<http://www.dropbox.com>

Describe



Describe



Access right Open Access
 Embargoed Access
 Restricted Access
 Closed Access

Required: Open access uploads have considerably higher visibility on ZENODO.

Embargo date

Required only for Embargoed Access uploads. Format: YYYY-MM-DD. The date your upload will be made publicly available in case it is under an embargo period from your publisher.

License

Required: The selected license applies to all of your files displayed in the bottom of the form. If you want to upload some files under a different license, please do so in two separate uploads. If you think a license missing in the list, please inform us at info@zenodo.org.

My uploads
Unpublished 27 May 2013, 11:28

New upload

Type of file

File information

Digital Object Identifier

Required: All your uploads should contain a DOI in your upload. This has to be the only one and will appear on the DOI page. If you already have a DOI, please use it.

Publication date

Required: Please enter YYYY-MM-DD. This date your publication will be made publicly available.

DOI

Required: Publicly available. This is your identifier (one author per line).

Embargoed

Required: Embargoed uploads will be made available after a period.

Restricted

Required: Only approved users.

Private

Required: Private uploads are not public.

Additional notes

Pre-reserve DOI

Consider

License required ▾

Access right +

- Open Access
- Embargoed Access
- Closed Access

Required. Open access uploads have considerably higher visibility on ZENODO.

● License +

Required. The selected license applies to all of your files displayed in the bottom of the form. If you want to upload some files under a different license, please do so in two separate uploads. If you think a license missing is in the list, please inform us at info@zenodo.org.

Funding recommended ▾

ZENODO is integrated into reporting lines for research funded by the European Commission via OpenAIRE (<http://www.openaire.eu>). Specify grants which have funded your research, and we will let your funding agency know!

🆔 Grants

Optional. Note, a human ZENODO curator will validate your upload before reporting it to OpenAIRE, and you may thus experience a delay before your upload is available in OpenAIRE.

Related datasets/publications recommended ▾

Specify identifiers of related publications and datasets. Supported identifiers include: DOI, Handle, ARK, PURL, ISSN, ISBN, PubMed ID, PubMed Central ID, ADS Bibliographic Code, arXiv, Life Science Identifiers (LSID), EAN-13, ISTC, URNs and URLs.

🆔 Related identifiers ⬇ ⬆ ✕

[+ Add another related identifier](#)

access rights · licenses · grants · formats

Software

Zenodo ↔ GitHub bridge

The image displays a composite of screenshots illustrating the Zenodo ↔ GitHub bridge setup. On the left, the Zenodo 'Settings' page shows the 'GitHub Repositories' section with a list of repositories. The 'decouple' repository is highlighted with a blue checkmark, and an arrow points to an 'ON' toggle switch. On the right, the GitHub 'Releases' page for the 'decouple' repository is shown. A callout box labeled 'Releases' highlights the 'v1.1.3' release with a DOI of '10.5281/zenodo.8345'. Another callout box labeled '.zenodo.json' shows a JSON snippet:

```
{ "name": "Jahn, Ilmar", "affiliation": "Institut für Theoretische Ph...", "description": "This repository contains the soft...", "access_right": "open", "license": "mit-license", "related_identifiers": [{" "identifier": "arXiv:1401.0088", "relation": "isCitableBy"}]}
```

 A third callout box shows the DOI 'DOI: 10.5281/zenodo.8345'.

<https://guides.github.com/activities/citable-code>

Example: results after one year



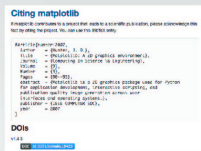
Flip the switch

Select the repository you want to preserve, and toggle the switch to turn on automatic preservation of your software.



Create a release

Go to GitHub and create a release. Zenodo will automatically download a .zip-ball of all new releases and register a DOIs for them.



Get the badge

After your first release, include our DOI badge in your README to advertise your DOI and have your software cited.

zenodo

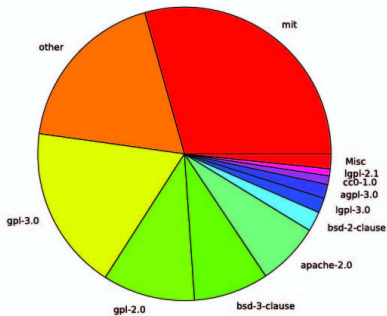
Research. Shared.

1 YEAR OF GITHUB INTEGRATION

Example: results after one year

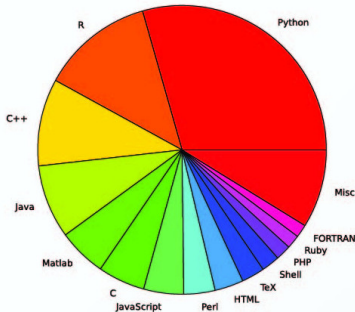
Licenses

Overview of software licences. 63% of packages did not contain any license.



Programming languages

Overview of primary programming language of software packages archived on Zenodo.



Step 2: Disseminate research data

Communities

The diagram illustrates the flow of research data dissemination. It starts with a GitHub repository for a "Microscopic vehicular mobility trace of Europarc roundabout, Creteil, France (vehicular-mobility-trace.github.io: v1.0)". This data is then uploaded to Zenodo, which is part of the "Research. Shared." network. From Zenodo, the data is disseminated to various research communities, such as the "Astronomy Thesis Collection". The diagram also shows the "Accept/reject uploads" interface and the "Direct community upload" process, which includes a "Want your upload to appear in this community?" form with an "Upload" button. Additionally, there is a "Harvesting API: OAI-PMH Interface" and an "Export" section.

15 April 2015 **Released** **Open access** ✓ Accept ✗ Reject

Microscopic vehicular mobility trace of Europarc roundabout, Creteil, France (vehicular-mobility-trace.github.io: v1.0)
Frédéric La Mouli ; Marie-Ange Labre ; Eric Ménéard

First release of the Europarc roundabout micro mobility dataset, Creteil, France.
<http://vehicular-mobility-trace.github.io/>

Uploaded by femousel on 16 April 2015.

Accept/reject uploads

Harvesting API:
OAI-PMH Interface

Export

Want your upload to appear in this community? **Upload**

Direct community upload

zenodo Research. Shared.

Astronomy Thesis Collection

Recent Uploads

Spectroscopy, Instrumentation and Applications to the Galaxy and to the Large Scale Structure of the Universe
Boris, Viktor A., Fabrikant, Daniel, Gerke, Margret
Identify the design and construction of modern spectrographs and use spectrographs to study the properties of galaxies, cool galaxy research areas and to study the evolution of our own Galaxy from a stellar population history. [2016-01-01]

Simulation and Analysis of Magnetic Reconnection in a Laboratory Plasma
Astrophysical Experiment
Magnetic reconnection is an inherently multi-scale process in which small scale physics and large scale dynamics both play important roles. To address the missing between local and global effects, we used magnetohydrodynamic (MHD) simulations of the...
Originally submitted on 16 April 2015.

The Structure and Evolution of Massive Star and Cluster Forming Regions
Santopau, Carlo, Daily, John

Astronomy Thesis Collection

This community, formed by astronomy libraries, aims to bring together astronomy and astrophysics papers and observations. Share the full text of your publication research with the world by uploading it to open access in this subject based community. Once added to the community, it can be indexed by ADS described by its abstracts, and featured in a community page. The ADS indexes this community page a month.

Community address

Astronomy Thesis Collection

Title: Astronomy Thesis Collection
Created by: community
Duration policy: This collection includes Previews and observations on the rights of accessibility and subscriptions. Other publication types and topics will not be accepted; however, documents that supplement your Previews, full-text articles or citations, are welcome. Address: Thesis access: all the

Example: Astronomy theses

The screenshot shows the Zenodo website interface. At the top, the Zenodo logo and the tagline "Research. Shared." are visible. Below the navigation bar, there is a search bar and a list of "Recent Uploads". Three uploads are listed:

- 04 April 2016** | **View** | **Download**
Spectroscopy, Instrumentation and Applications to the Galaxy and to the Large Scale Structure of the Universe
Strom, Warren B., Fabian, Daniel, Sellar, Margaret
I study the design and construction of modern spectrographs and use spectrographs to study the properties of galaxies in a local galaxy redshift survey and to study the structure of our own Galaxy from a stellar radial velocity survey. (perform four ...)
- 04 April 2016** | **View** | **Download**
Simulation and Analysis of Magnetic Reconnection in a Laboratory Plasma Astrophysics Experiment
Marty, Nicholas, Quenne, Carl, Zurek, Ellen
Magnetic reconnection is an inherently multi-scale process in which small scale physics and large scale dynamics both play important roles. To address the interplay between local and global effects, extended magnetohydrodynamic (MHD) simulations of the ...
- 14 August 2015** | **View** | **Download**
The Structure and Evolution of Massive Star and Cluster Forming Regions
Sternberg, David, Dale, John

On the right side of the page, there is a section for the "Astronomy Thesis Collection" community edition. It includes a title "Astronomy Thesis Collection" with a logo, a description of the community's goals, and contact information for the curator, Nathan Edward.

The screenshot shows a Zenodo page for a thesis. At the top, there is a navigation bar with "Home", "Search", "Feedback", "Help", and "Tour this Page". Below the navigation bar, the title "2014PHD.....115" is displayed. The page includes a navigation menu with "Abstract", "References (5)", "Citations", "Co-Reads", "Similar Articles", "Graphs", and "Table of contents".

The main content area features the title "Core-collapse Supernova Progenitors in the Era of Untargeted Transient Searches" by Sanders, Nathan Edward. It includes a link to "show affiliations" and the following information:

Ph.D. thesis, *Harvard University* (2014), 557 pages
Published in Apr 2014

The abstract text reads:

Core-collapse supernovae (SNe) are the highly energetic explosions of massive stars (> 8 solar masses) that are pervasive in their influence throughout astrophysics. They are the phenomenon with primary responsibility for enriching the universe with many of the heavy elements (like carbon and oxygen) that are needed for life, provide a critical feedback pressure which helps to shape the galaxies that host them, and are the likely formation mechanism for stellar mass black holes. In the past decade, the study of these explosions has been revolutionized by the advent of wide field, untargeted transient searches like Pan-STARRS1 (PS1). These new searches permit the discovery of SNe at unprecedented rates, and absent of many of the selection effects that have enforced biases on past, targeted transient searches. This thesis presents a broad survey of core-collapse SN phenomenology exhibited in the discoveries of untargeted searches, and statistically quantifies population properties of these explosions that link them to distinct classes of progenitor stars. Through studies of the host galaxy and explosion properties of extreme PS1-discovered events, and controlled samples of specific classes of core-collapse objects, we constrain the effect of progenitor star chemical composition (metallicity) on their eventual death states. We provide a new observational, photometric tool which lowers the cost of precisely and accurately measuring the metallicities of distant galaxies and supernova host environments. Moreover, we develop and apply a novel, multi-level Bayesian model for optical transient light curves which we apply to simultaneously interpret more than 20,000 PS1 images. This study illustrates how population-level modeling of data from large photometric surveys can yield improved physical inference on their progenitor stars through comparison to physical models. In the coming era, as next-generation facilities like the Large Synoptic Survey Telescope come online, the supernova discovery rate will accelerate, fomenting the community's capacity for detailed individual observational follow-up. New observational and statistical tools like those presented here will be critical to enable the next generation of studies in supernova astrophysics.

Step 3: Cite to give credit

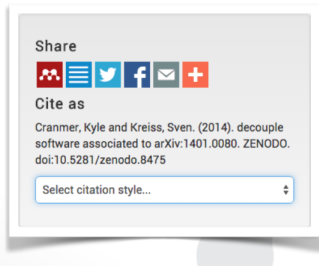
Digital Object Identifier

<http://dx.doi.org/10.5281/zenodo.8475>

<http://data.datacite.org/10.5281/zenodo.8475>

Machine actionable Globally unique

Long-term persistence






“Data ↔ Code ↔ Paper”

The image shows three overlapping screenshots illustrating the relationship between data, code, and paper. The top-left screenshot is a Zenodo page for 'decouple software associated to arXiv:1401.0080'. The top-right screenshot is an INSPIRE page for the paper 'A Novel Approach to Higgs Coupling Measurements'. The bottom screenshot is a larger view of the INSPIRE paper page, showing the title, authors, abstract, and a plot of Higgs coupling measurements.

data (DATAVERSE) ↔ code (ZENODO) ↔ paper (arXiv)

arXiv:1401.0080 · hep-ex/0011057



 <https://zenodo.org/>
 <https://github.com/zenodo>
 @zenodo_org



(Acknowledgements: Lars Holm Nielsen, Krzysztof Nowak)