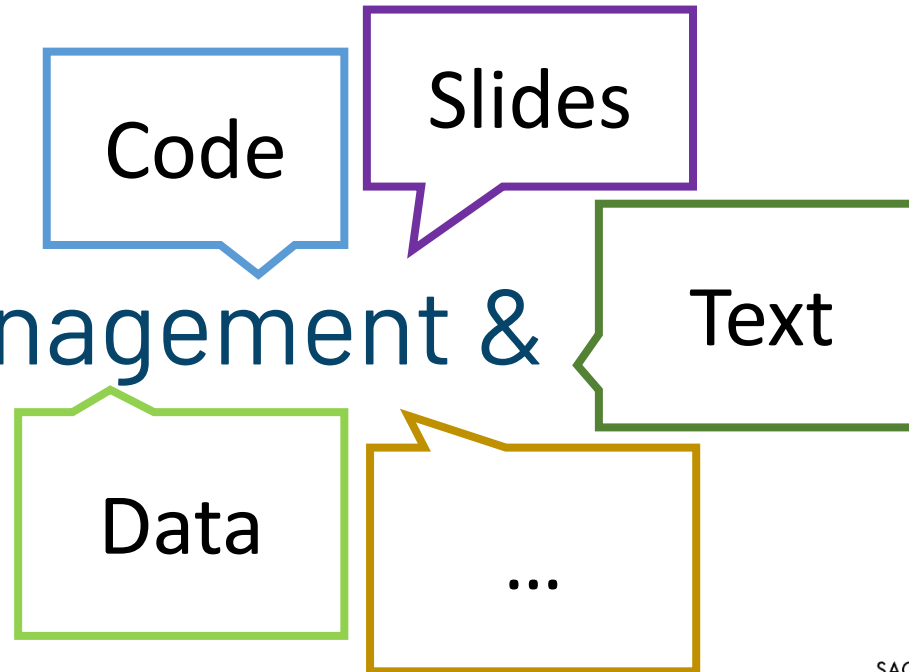


Research Data Management & Open Science

Robert Haase



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



Diese Maßnahme wird gefördert durch die Bundesregierung aufgrund eines Beschlusses des Deutschen Bundestages. Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des von den Abgeordneten des Sächsischen Landtags beschlossenen Haushaltes.

Quiz

- When you shared materials over the internet, which *platform* did you use?

Onedrive/Google
cloud/Dropbox/etc.



Zenodo/Figshare/
arxiv/F1000



Email



Other



Closed science

Why are some science-related materials/data/code not shared?

- Risk of being scooped
- Fear of blaming oneself (imposter syndrome)
- Lack of awareness (who is allowed to publish *my work*?)
- Assumption: it's not worth the effort.

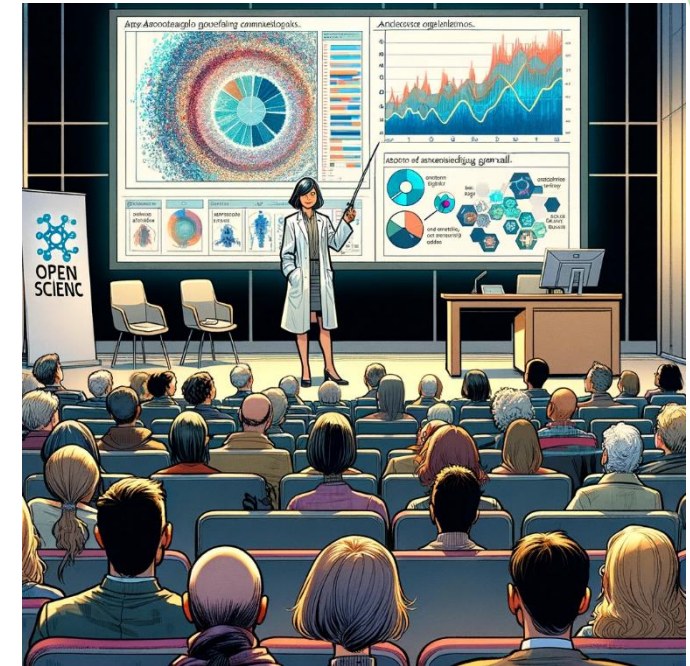


Open Science

- Research related
(hot topics)
- Often tailored towards
general audience
(science communication)
- Earliest at the time a
manuscript is published
(e.g. as preprint)

Open Training

- Routine tasks
(colder topics)
- Transfer of
domain-specific
knowledge



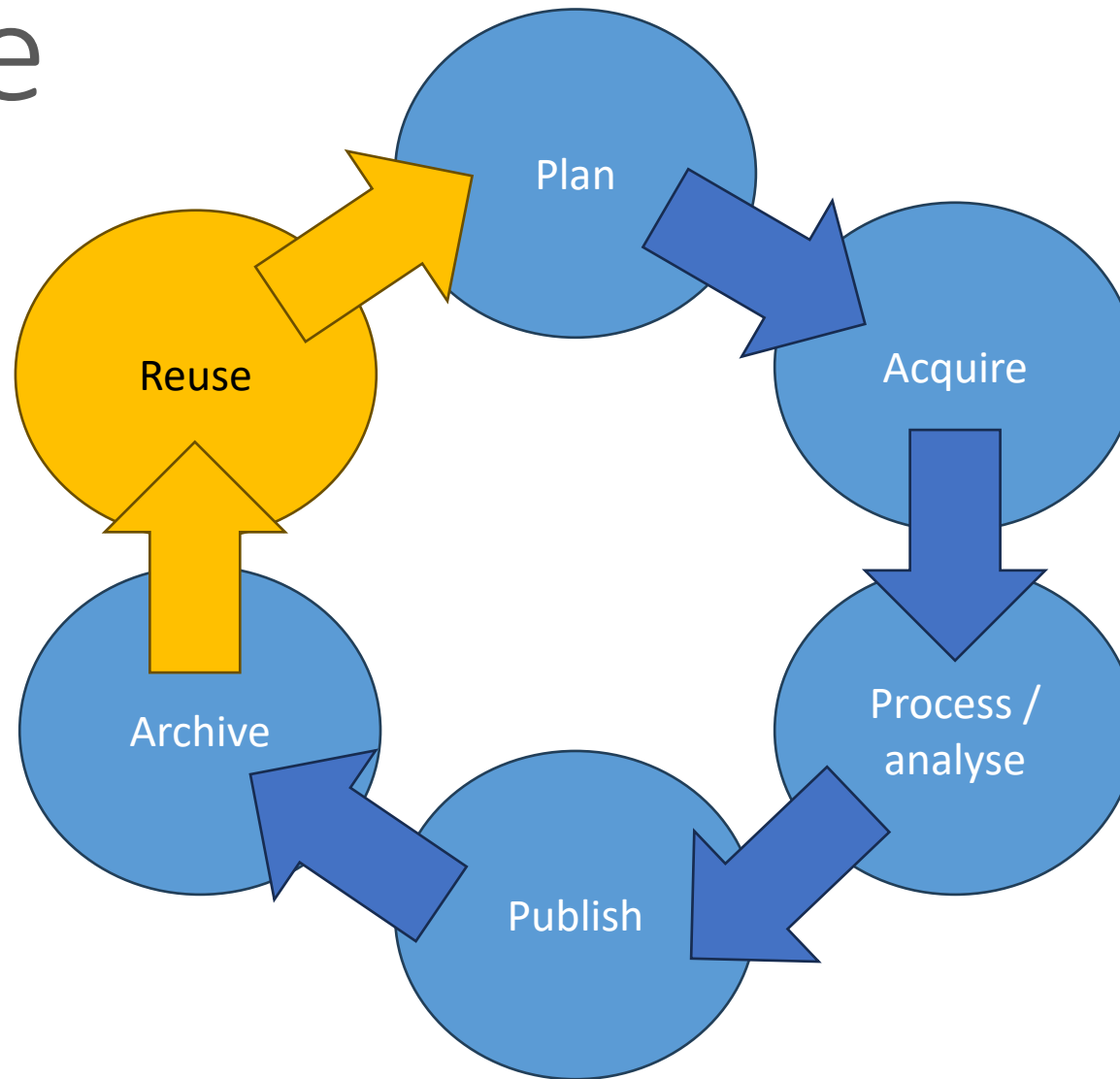
Research Data Management (RDM)

- All activities, processes, terms, persons which have relationships with data
 - Processing
 - Storage
 - Organization
 - Publication
 - ...
- In routine: working with data



RDM Life Cycle

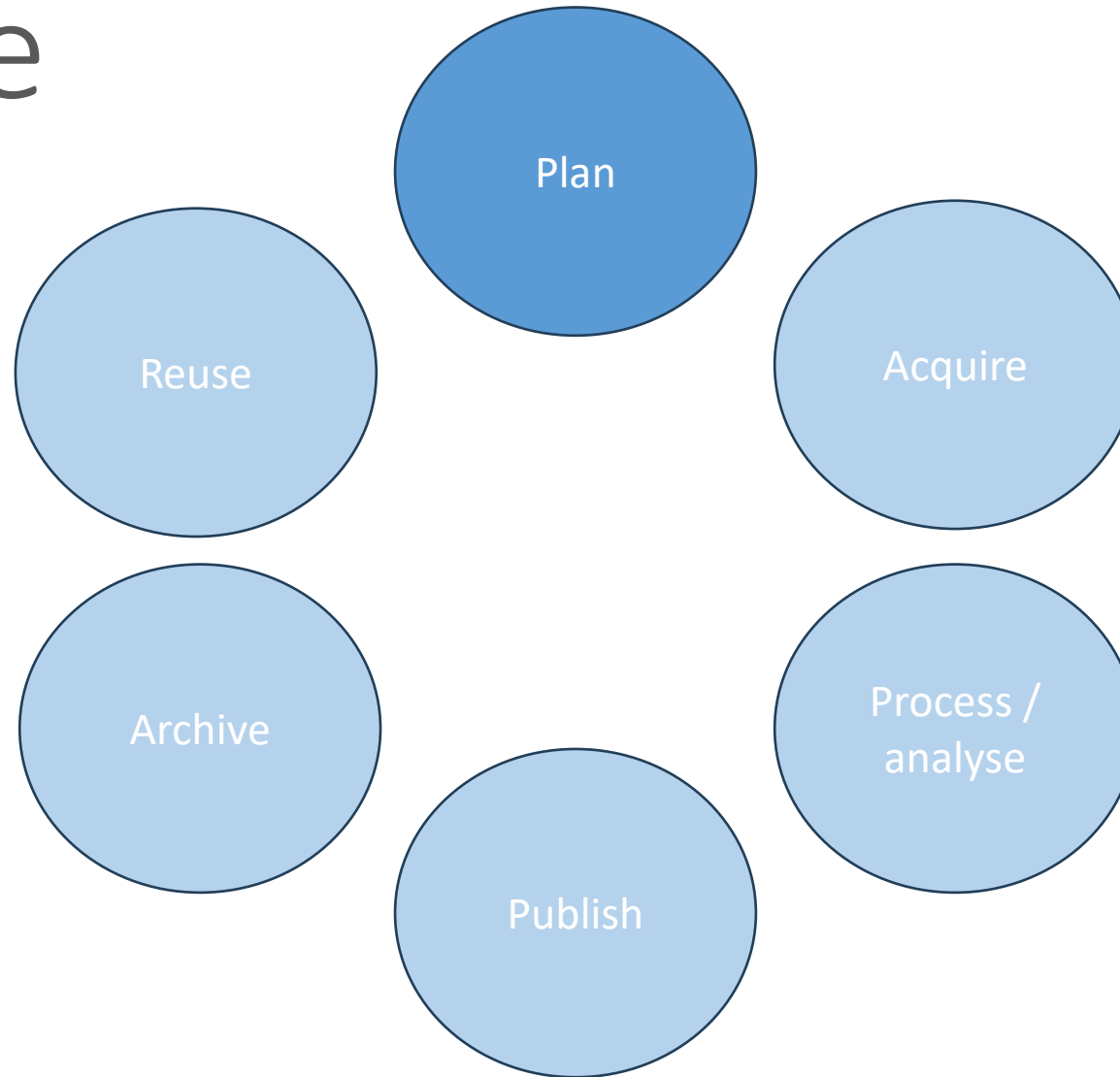
- Processes are ideally cyclic
- Closing the loop is a major the challenge



RDM Life Cycle

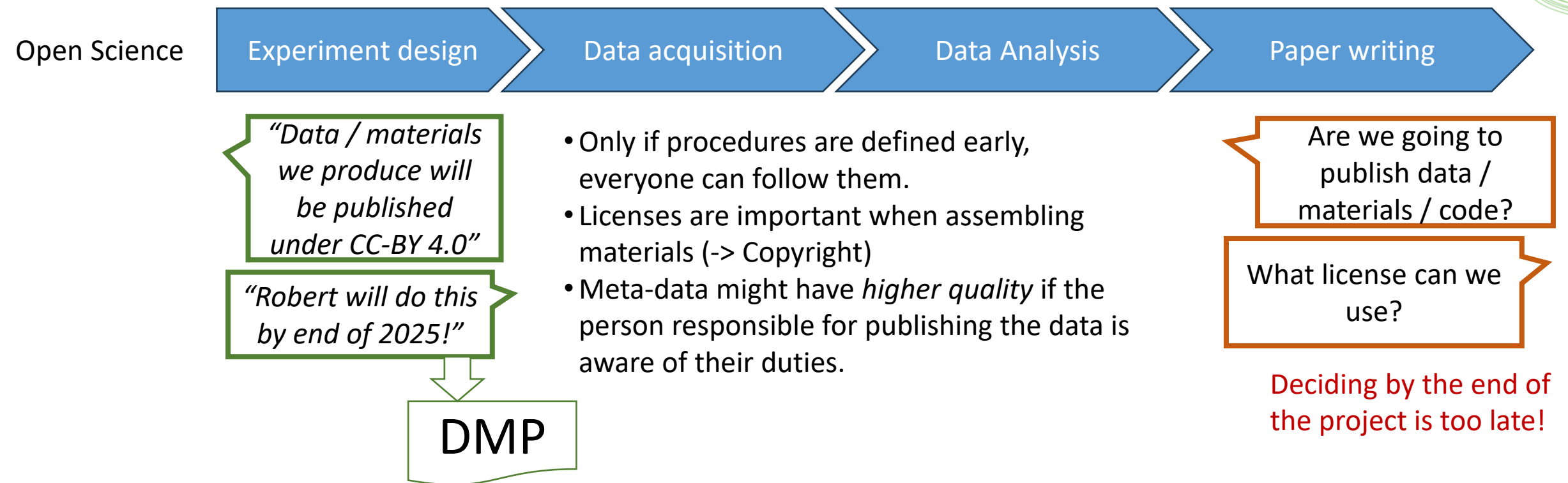
Plan

- Cost
- Benefit
- Quality
- Strategic decisions



Data Management Plans (DMPs)

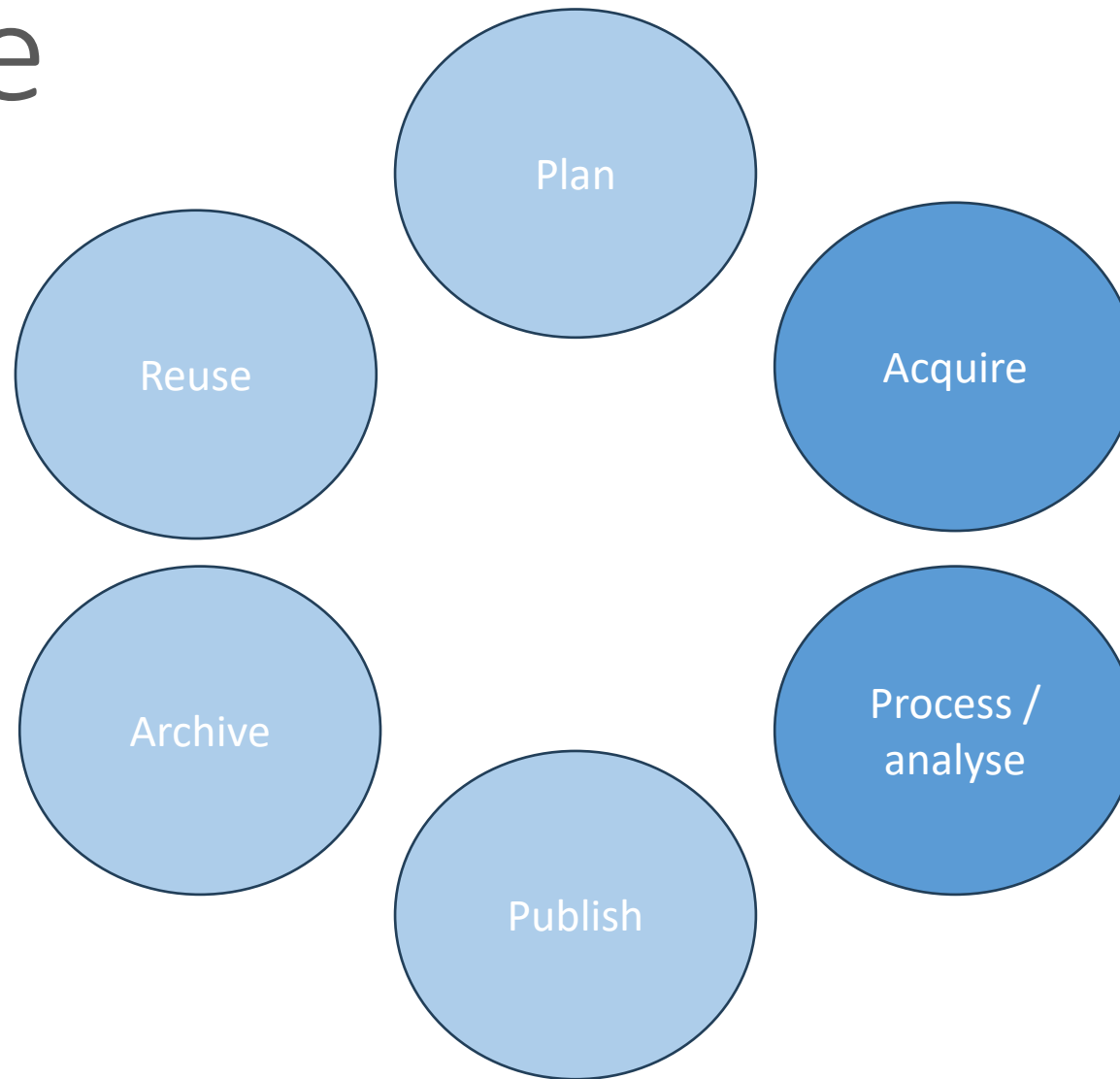
- Define responsibilities and procedures early!



RDM Life Cycle

Acquisition + Processing

- Types of data
- Terms and conditions
 - Usage rights
 - Copyright
- IT infrastructure
 - HPC
 - USB-Drives
- Backup



Types of data

- Structured data
 - Tables, databases
- Unstructured data
 - Texte, emails, videos, pictures
- Semi-structured data
 - Surveys
 - Scientific images



Types of data

- Openly accessible data
 - „open data“
 - „open source“ software
- Business data
- Research data
 - Hot / cold
- Personal data
- Secret data

In need of
protection
(schutzbedürftig)



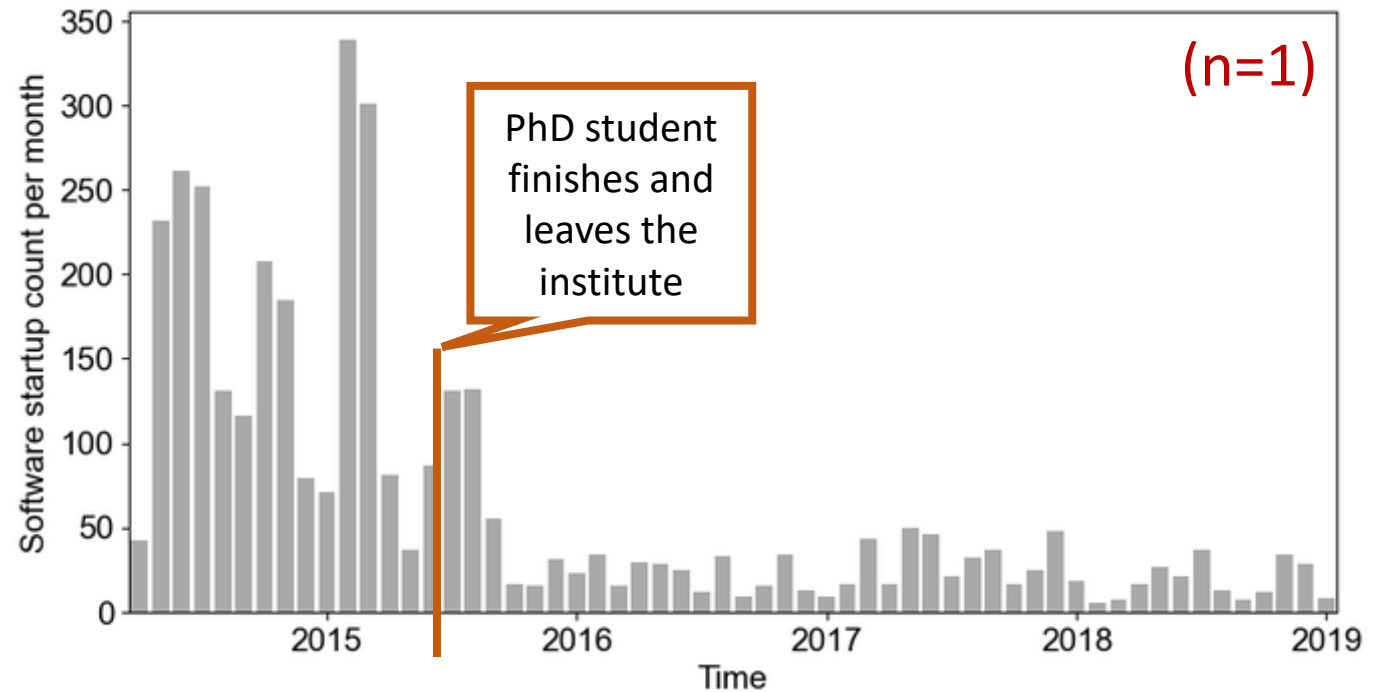
Version control [for code]

- Me, during my Phd:

„Geisterr“ usage statistics

- | | |
|--------------------------|--------------------|
| geisterr - Kopie (2).exe | 2/27/2013 4:43 PM |
| geisterr - Kopie.exe | 3/26/2012 9:53 PM |
| geisterr.exe | 2/18/2015 3:09 PM |
| geisterr_alpha.exe | 2/20/2014 8:44 AM |
| geisterr_alpha2.exe | 1/7/2014 1:41 PM |
| geisterr_alpha3.exe | 1/8/2014 11:26 AM |
| geisterr_alpha4.exe | 1/15/2014 2:09 PM |
| geisterr2013.exe | 12/17/2013 9:55 AM |
| geisterrbilder.exe | 3/22/2013 4:38 PM |
| histogeisterr.exe | 4/29/2013 2:19 PM |

(n=1)



Version control [for code]: git

Github contributions

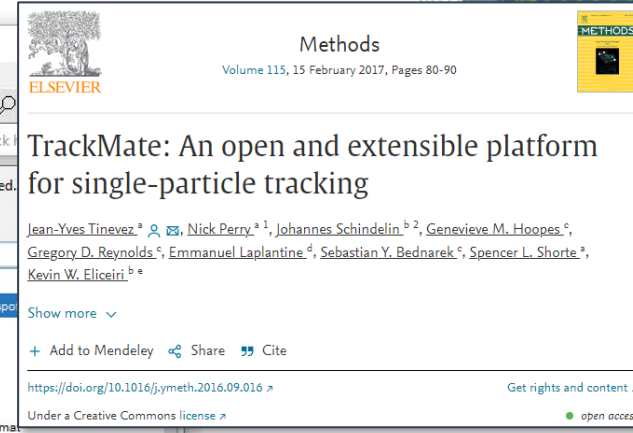
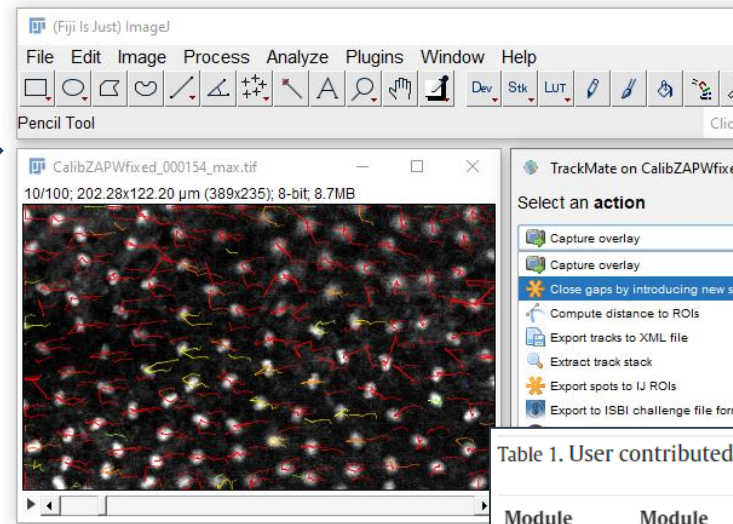
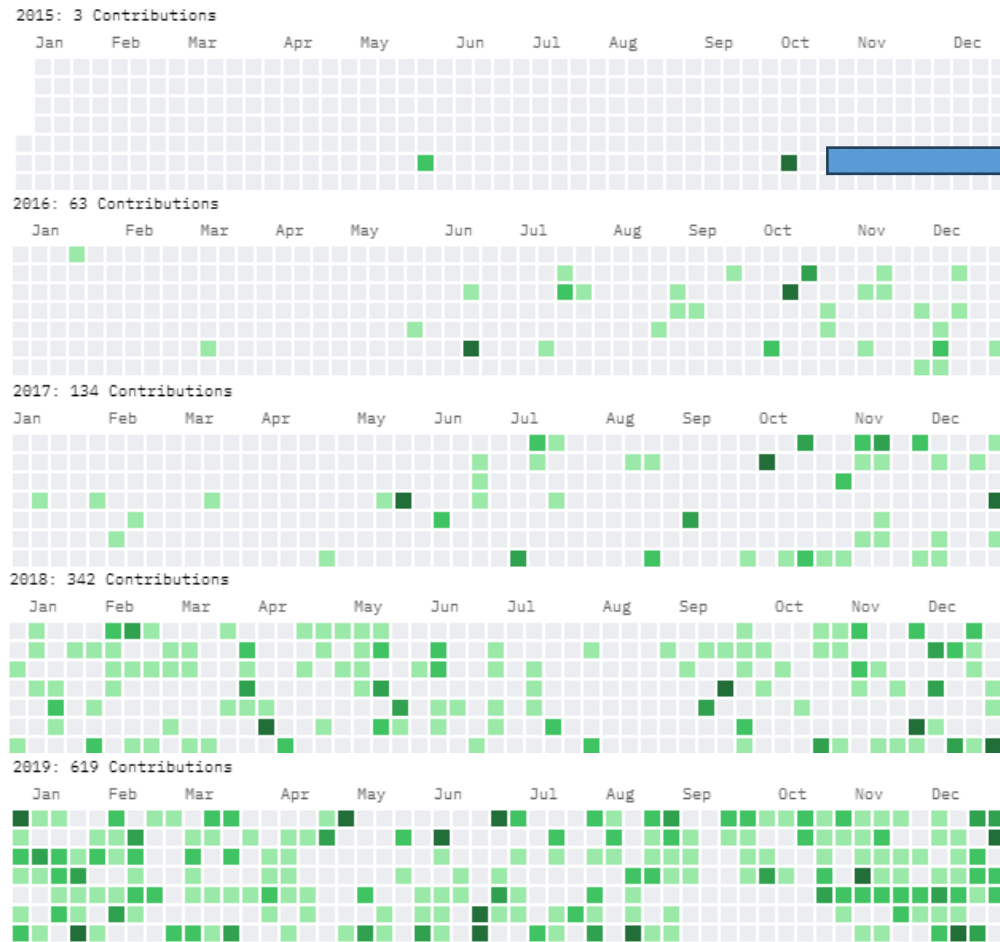


Table 1. User contributed modules of TrackMate v3.4.0.

Module name	Module type	Purpose	Author	Location
Linear tracker	Particle-linking	Linking transported particles by extrapolating their velocity	Ronny Sczech	https://github.com/chicoranny/RonnyTrackMate/
Batch mode	Plugin	Runs TrackMate in batch reading		
Close gaps	Generic action	Close gaps in tracks by creating spots in missing frame by linear interpolation of their coordinates	Robert Haase	Integrated into TrackMate v3.4.0

Note: The colour-code is not consistent between the years

Version control [for code]: git

• Me, today:

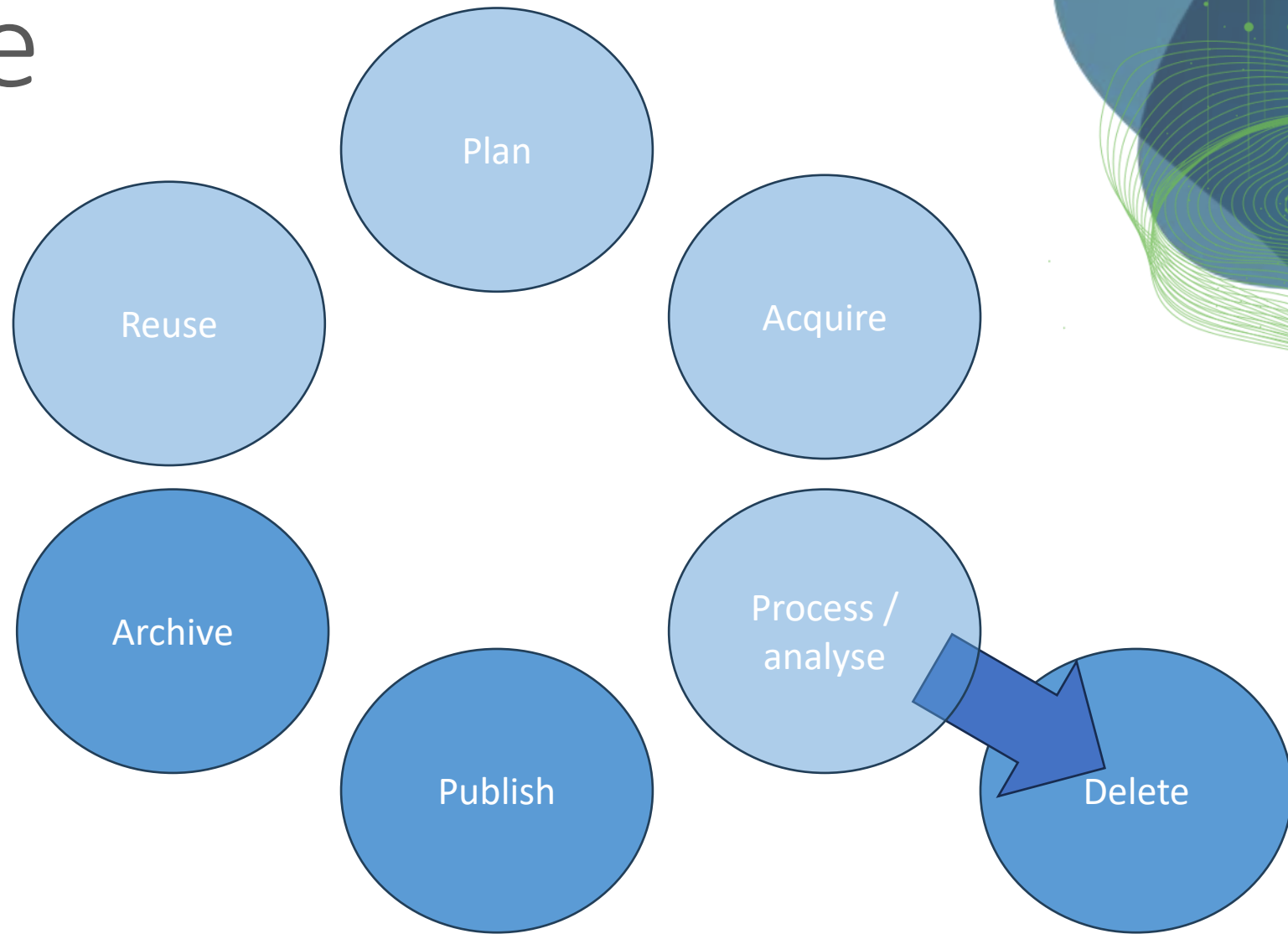
The screenshot shows the GitHub repository page for `haesleinhuepf/stackview`. The repository is public and has 26 branches and 47 tags. The commit history shows several recent commits, including one that added a DOI badge. The repository description states: "Interactive image stack viewing in jupyter notebooks based on ipycanvas and ipywidgets". The repository is licensed under the BSD-3-Clause license. A preview of the stackview application is shown at the bottom, displaying a dark image with a bright, circular, multi-colored pattern.

The screenshot shows the Zenodo record page for `haesleinhuepf/stackview: 0.12.1`. The record was published on December 5, 2024, and is version 0.12.1. The authors listed are Robert Haase, Jordão Bragantini, Oren Amsalem, and yiqunma. The record is categorized as Software and is available for download. The "What's Changed" section lists a new release of `wordcloudplot` by @haesleinhuepf. The "Full Changelog" link is provided. The "Files" section shows the download link for `haesleinhuepf/stackview-0.12.1.zip` and the directory structure, including `haesleinhuepf-stackview-08a048d`, `.github`, and `workflows`.

The screenshot shows the PyPI Download Stats page for the `stackview` package. The page displays the package name, author (Robert Haase), and a summary: "Interactive image stack viewing in jupyter notebooks". The latest version is 0.12.1. The required dependencies are listed: `imageio`, `ipycanvas`, `ipyevents`, `ipykernel`, `ipympi`, `ipywidgets`, `matplotlib`, `numpy`, `scikit-image`, `toolz`, and `wordcloud`. The download statistics show 278 downloads in the last day, 1,513 in the last week, and 5,372 in the last month. A line graph titled "Daily Download Quantity of stackview package - Overall" shows the download trend over time, with a peak in late October/early November.

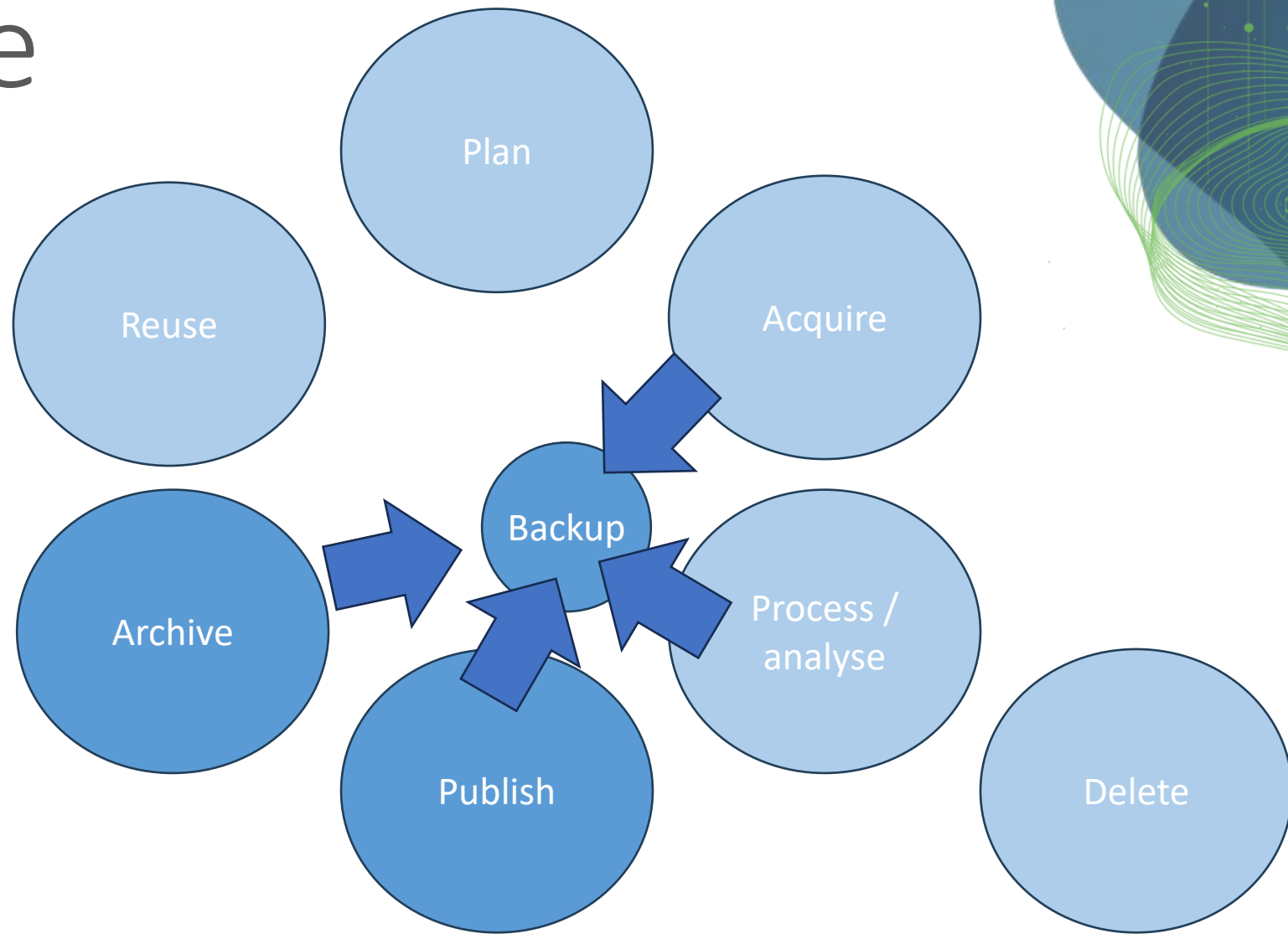
RDM Life Cycle

- Right to publish
- Regulatory aspects
 - Research data: archive 15 years
- Authorship
- Registration (-> Findable)



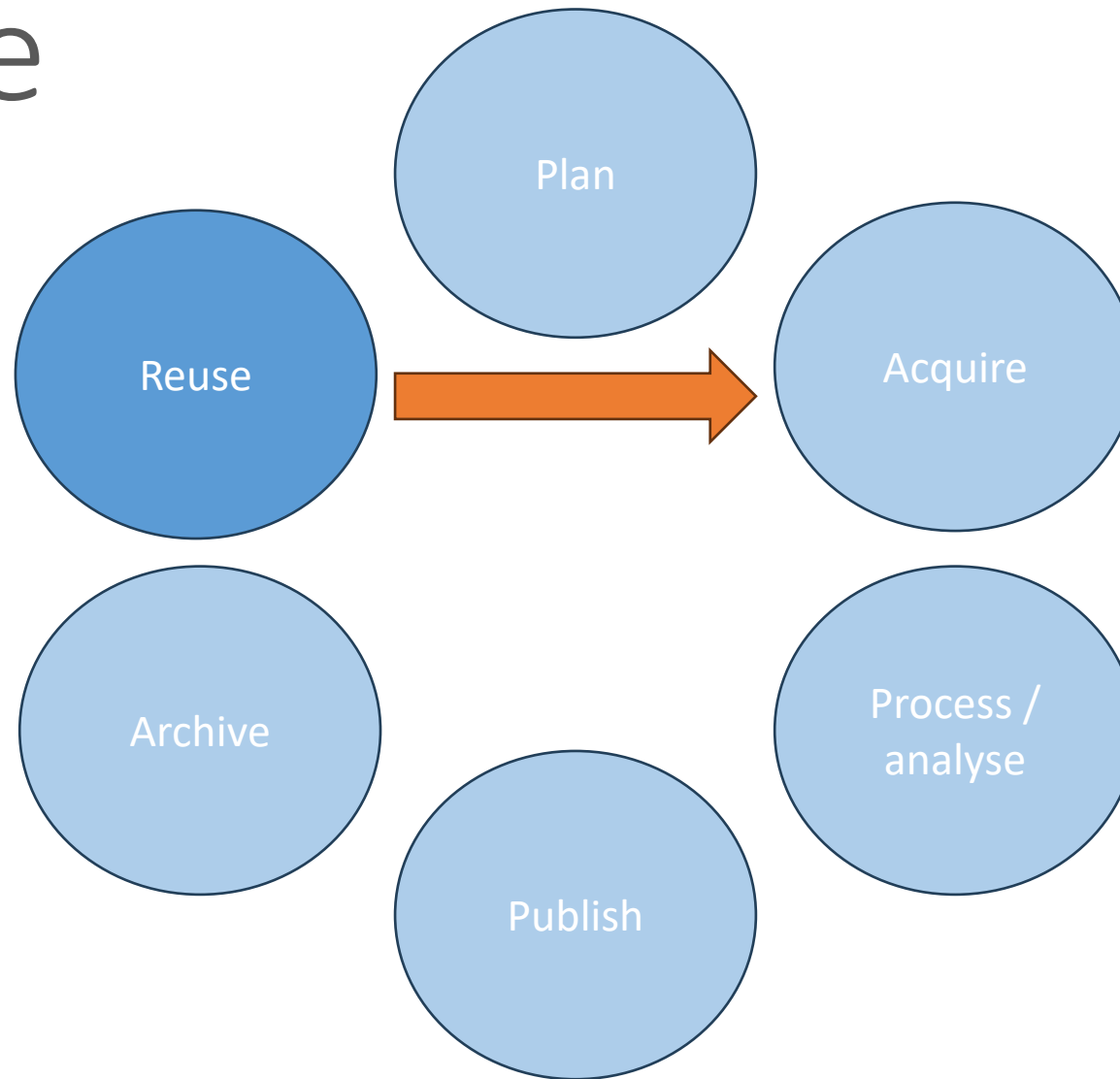
RDM Life Cycle

- Backup: Regular copy of (all) files
 - Conflict: Some personal data needs to be deleted, should not go into backup
- Publication: Selection of curated files for external use
- Archive: Selection of curated files for later use (not necessarily public)



RDM Life Cycle

- Potential future benefit
- Sustainability
- Important: **Licensing**
 - Has impact on next cycle / acquisition



Scientific culture

Public access to research results -> Reusability



“Codex”

Guidelines for Safeguarding Good Research Practice

Code of Conduct

Guideline 13: Providing public access to research results

► As a rule, researchers **make all results available** as part of scientific/academic discourse. In specific cases, however, there may be reasons not to make results publicly available (in the narrower sense of publication, but also in a broader sense through other communication channels); this decision must not depend on third parties. **Researchers decide autonomously** – with due regard for the conventions of the relevant subject area – whether, how and where to disseminate their results. If it has been decided to make results available in the public domain, researchers describe them clearly and in full. Where possible and reasonable, this includes making the research data, materials and information on which the results are based, as well as the methods and software used, available and fully explaining the work processes. Software programmed by researchers themselves is made publicly available along with the source code. Researchers provide full and correct information about their own preliminary work and that of others.

Explanations:

In the interest of transparency and to enable research to be referred to and **reused by others**, whenever possible researchers make the research data and principal materials on which a publication is based available in recognised archives and repositories **in accordance with the FAIR principles** (Findable, Accessible, Interoperable, Reusable). Restrictions may apply to public availability in the case of patent applications. If self-developed

Scientific culture



Information for Researchers, No. 61 | September 1, 2022

Package of Measures to Support a Shift in the Culture of Research Assessment

DFG changes proposal forms and introduces mandatory CV template / The aim is to support a shift in the culture of research assessment / Improvement of equal opportunity practices

DFG changes proposal forms and introduces mandatory CV template / The aim is to support a shift in the culture of research assessment / Improvement of equal opportunity practices

Successful science and research require suitable framework conditions. The Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) ensures these conditions by regularly conducting analyses, providing the relevant information and adapting its procedures accordingly. The DFG set out the challenges and fields of action in a position paper on academic publishing published in May of this year: it sees both the academic community as a whole and itself as a funding organisation as being responsible for initiating a cultural shift towards research assessment that is geared more towards equal opportunity and attaches even greater importance to the substance of research. In the interests of bringing about such a shift, it is up to research funding organisations to **broaden the spectrum of accepted publication formats, to attach greater value to content-based evidence of achievement** and to strengthen the recipient side of publishing. The DFG has launched a comprehensive and far-reaching package of measures in order to fulfil this mandate.

Binding CV template across all funding programmes

For this reason, the assessment of a researcher's accomplishments must be holistic and based on substantive qualitative criteria. In order to strengthen qualitative evaluation criteria over quantitative indicators, the DFG will be introducing a curriculum vitae template that will be mandatory across all programmes from 1 March 2023 (the template will be adapted shortly for proposals under the Collaborative Research Centre and Research Training Group programmes; information will be provided separately in this regard). The template adopted by the DFG Senate allows applicants to provide both narrative and tabular information, thereby facilitating a holistic view of the applicant's academic career in the review and evaluation process.

In addition to the mandatory information required in order to assess eligibility, applicants may also provide details of special circumstances or additional services to scholarship such as committee activities or the establishment of research infrastructures. As such, the template provides the basis for a qualitatively sound assessment of academic performance that takes greater account of the respective stage of the individual's life and career. Accordingly, reviewers are now instructed to consider applicants' academic performance in the context of their individual curriculum vitae and career stage.

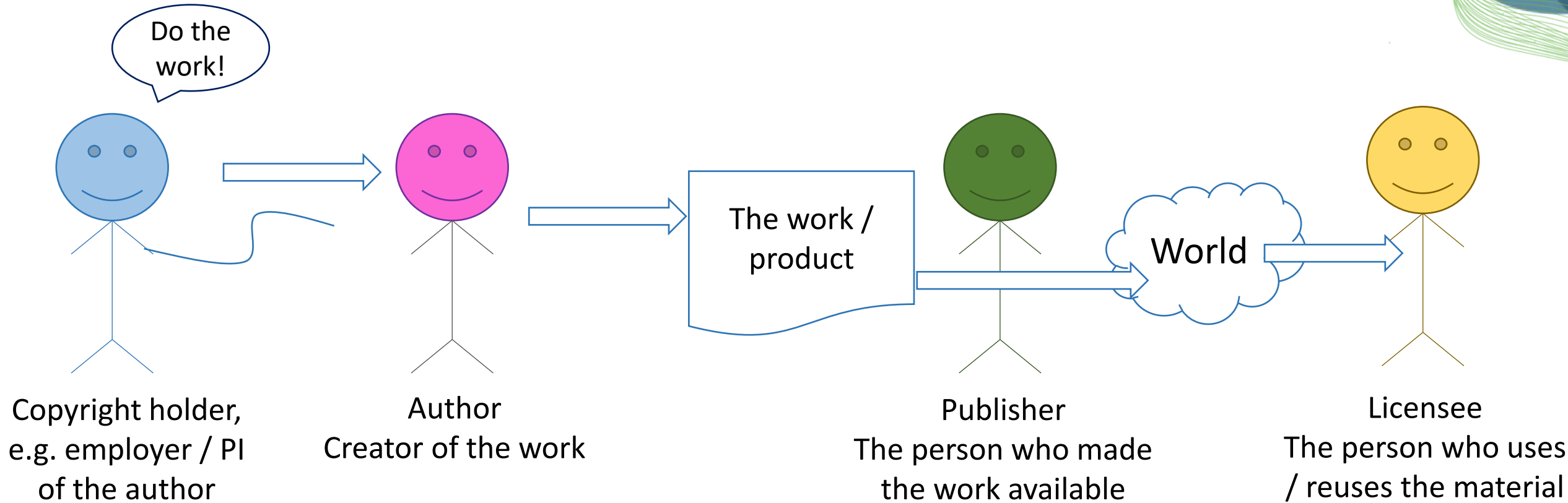
Publication details in proposals and CVs

Performance assessment based on content-related qualitative criteria also explicitly includes ensuring that the entire spectrum of academic publication types are equally displayed and acknowledged in funding proposals and CVs. In addition to a maximum of ten publications in the more common publication formats, the **CV can therefore now list up to ten further sets of research outcomes and findings that have been publicised in a variety of other ways, including articles on preprint servers, data sets or software packages, for example.** In DFG proposals, the project-specific list of publications will be included in the general bibliography. The intention here is to shift the focus of the review and the evaluation of a proposal away from the list of publications and towards the substance of the applicant's accomplishments. In order to document their own published preliminary work, applicants can typographically highlight (e.g. in bold) a maximum of ten of their own publications in the bibliography that are important for the project. **No information on quantitative metrics such as impact factors and h-indices is required in the CV or the proposal, and such information is not to be considered in the review.** The relevant details are included in DFG forms and review instructions.

These modifications and innovations reflect the fact that the DFG is continuing to promote the cultural shift in research assessment that was advocated in May with the publication of the position paper on academic publishing. The DFG hopes that this refocus – away from quantitative indicators and towards the substance of scholarship – will lead to improved equality of opportunity and a higher-quality basis for review overall.

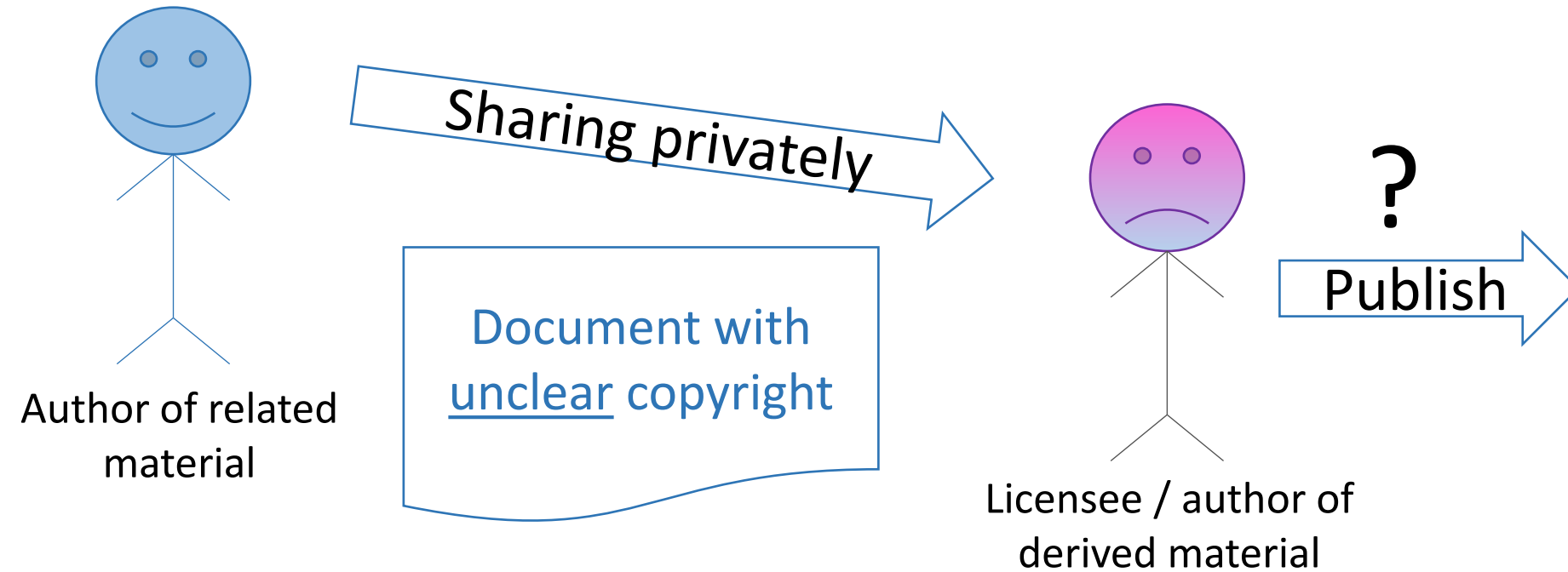
Am I allowed to publish my stuff?

- ... it depends... on who is responsible



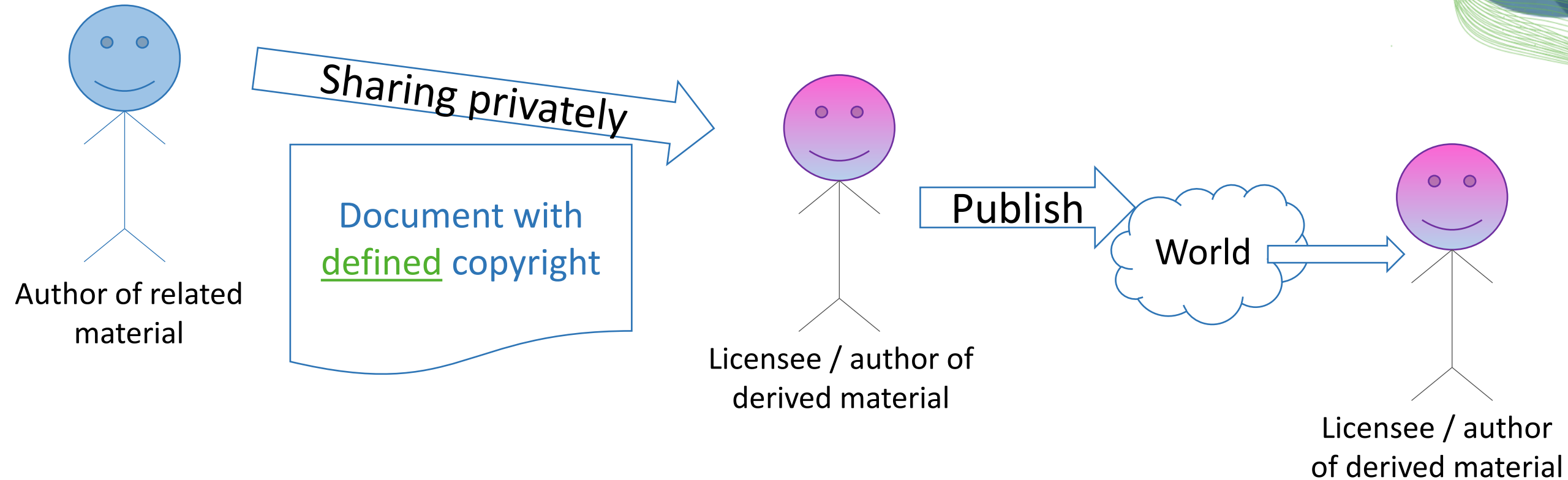
Am I allowed to publish my stuff?

- ... it depends... on what materials served as basis



Am I allowed to publish my stuff?

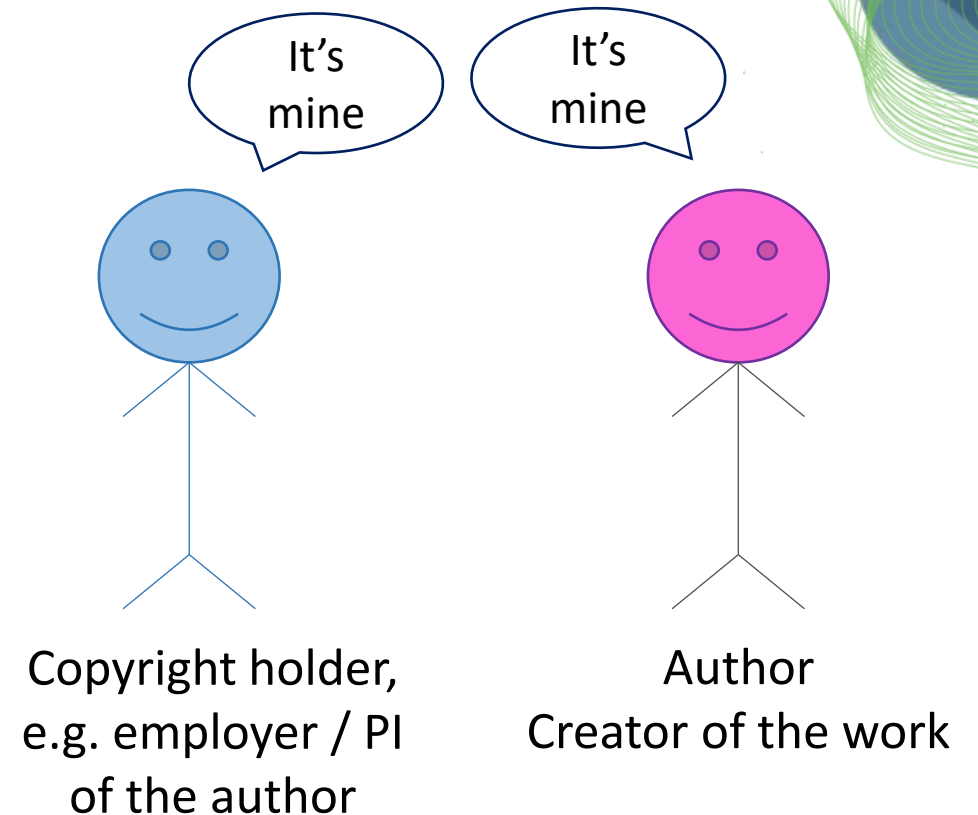
- ... it depends... on what materials served as basis



Public versus institutional repositories

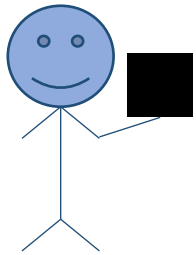
Conflict of interest between employers and scientists

- Universities seek to keep things secret and potentially exploit them commercially (licenses, startups, ...)
- Scientists need to publish to advance their career.
- Hints:
 - Decide early during the project what will be published and by who (-> DMP)
 - Check your job description! (Is “Programming” or “model training” part of it?)



Openness of software / projects

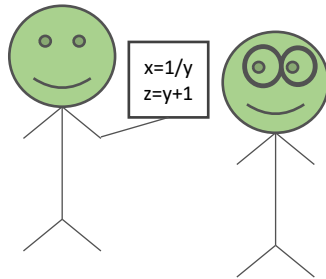
Closed source



- Open to collaborations
- “Black box”
- Compiled code (e.g. C/C++)
- Good for protecting intellectual properties (\$\$\$)

Hardware device drivers

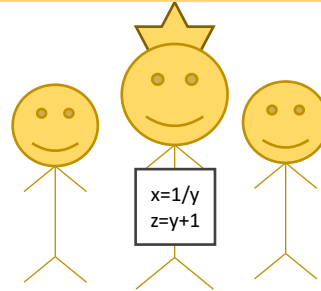
Open source



- Code available to read
- Not necessarily executable code
- No maintenance / support efforts

Custom data analysis scripts

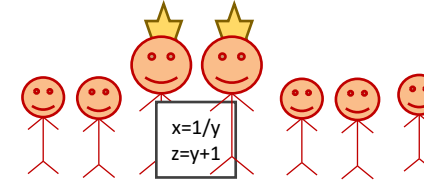
Benevolent dictatorship



- Open to contributions
- Single maintainer, often overwhelmed
- Efficient decision making
- Bus factor ≈ 1

scikit-image, scipy, OpenCL

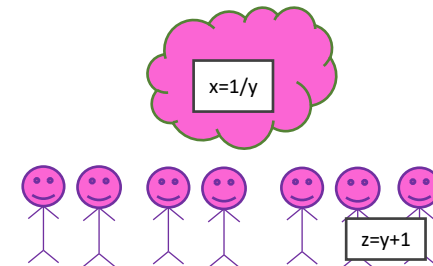
Community driven



- Open to contributions
- Partially democratic
- Board of maintainers (core developers)
- Long-winded decision making

Python, numpy

Openly extensible



- Openly extensible; without maintainers involved
- Partially community driven

Standard for sharing: The FAIR-principles

- Findable
- Accessible
- Interoperable
- Reusable



Findability

Domain-specific

- Search repository registries for your field!

Guidelines

- Publish where your community publishes
- Publish where everyone publishes (beyond your community)
- Publish in your local institute's infrastructure

Your
individual
interest



Your
Institute's
interest

re3data.org

World Data Center for Solid Earth Physics
WDC for SEP

Subject(s) Natural Sciences Geosciences Geophysics and Geodesy

Repository type(s) disciplinary

Provider type(s) dataProvider

FAIRsharing.org

search through all content

STANDARDS DATABASES POLICIES COLLECTIONS ORGANISATIONS ADD CONTENT STATS

Databases

A registry of knowledgebases and repositories of data and other digital assets.

Wednesday, April 23rd 2025, 10:02: The FAIRsharing application is currently unable to send emails, which will prevent various features from operating. We are working with our service provider to fix this issue. Please accept our apologies for the inconvenience.

seismology

MATCH ALL TERMS MATCH ANY TERM

MAINTAINED NOT MAINTAINED

RECOMMENDED NOT RECOMMENDED

READY DEPRECATED UNCERTAIN IN DEV.

Registry

Search APPLY

Record Type

Clear All Registry: Database Query string: seismology

Displaying 1 to 14 of 14.

IRIS Data

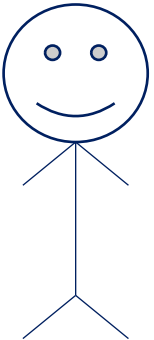
Incorporated Research Institutions for Seismology Data

IRIS provides management of, and access to, observed and derived data for the global earth science community. IRIS membership comprises virtually all US universities with research programs in seismolog...

Environme... Earth Scien... Not applic...

Incentives: Findability

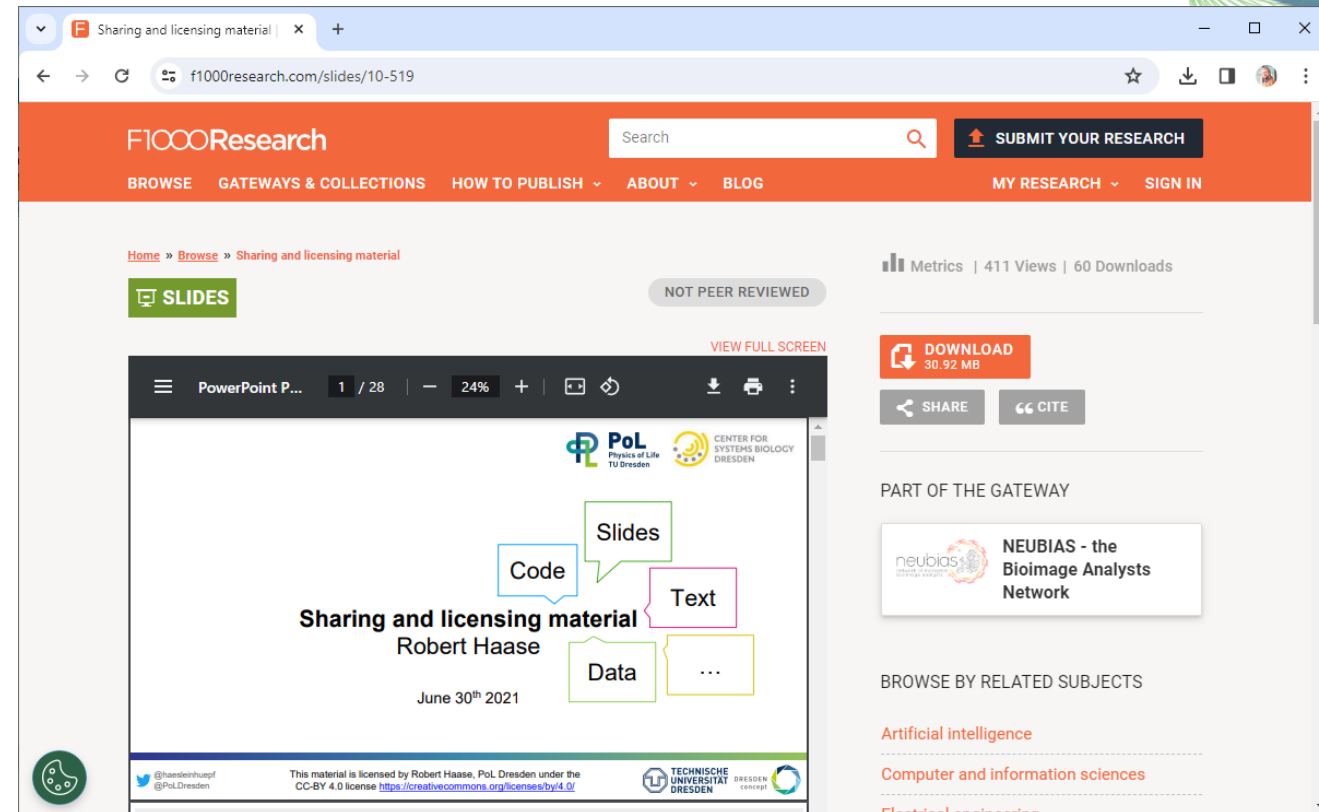
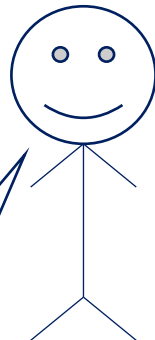
- Your *future-self* will thank you, because they will find your work



You remember
that talk you gave
in 2021?

Where are the
slides?

Online,
open access!



Accessibility

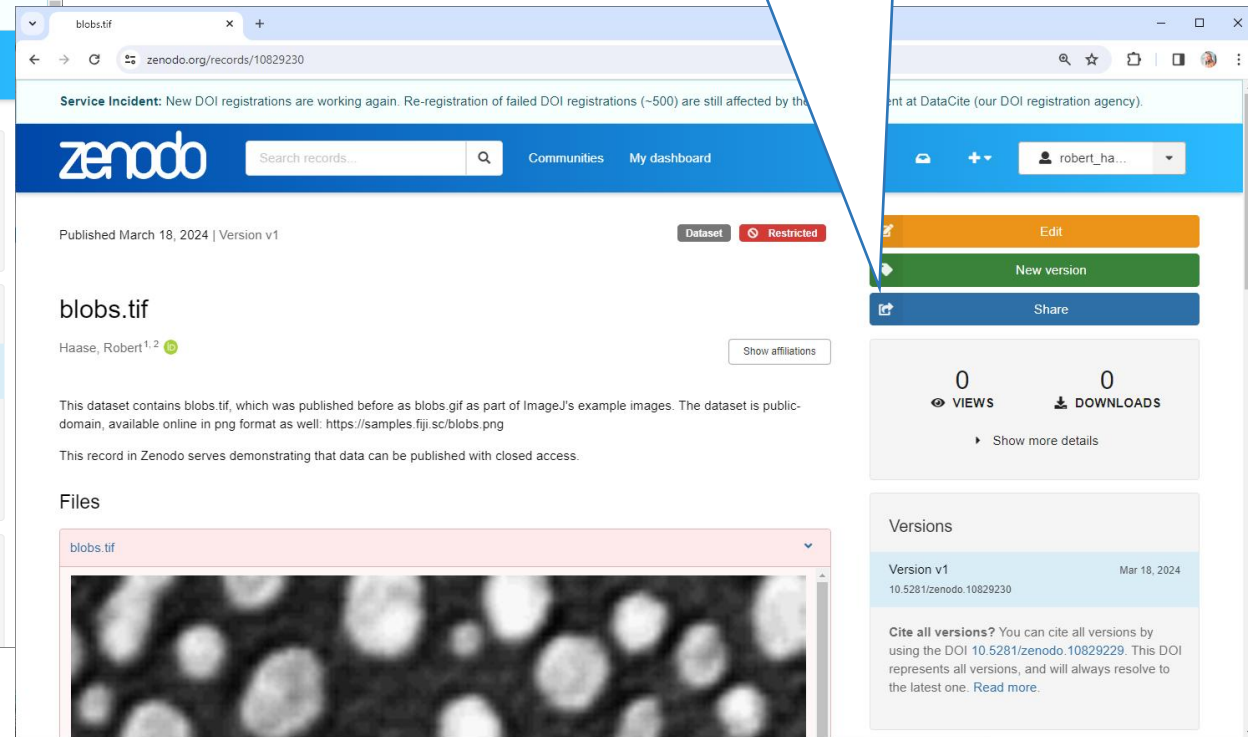
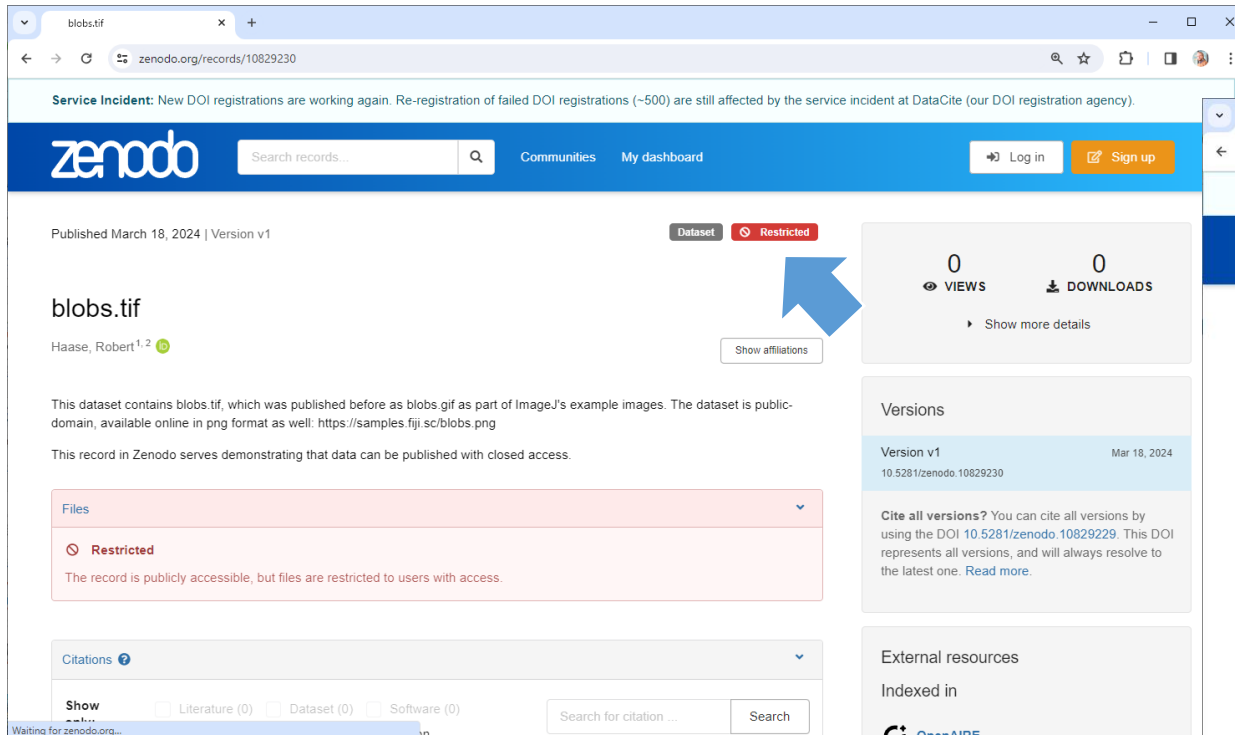
- The A in FAIR does not necessarily stand for Open Access

April 7th 2024:

26
👁 VIEWS

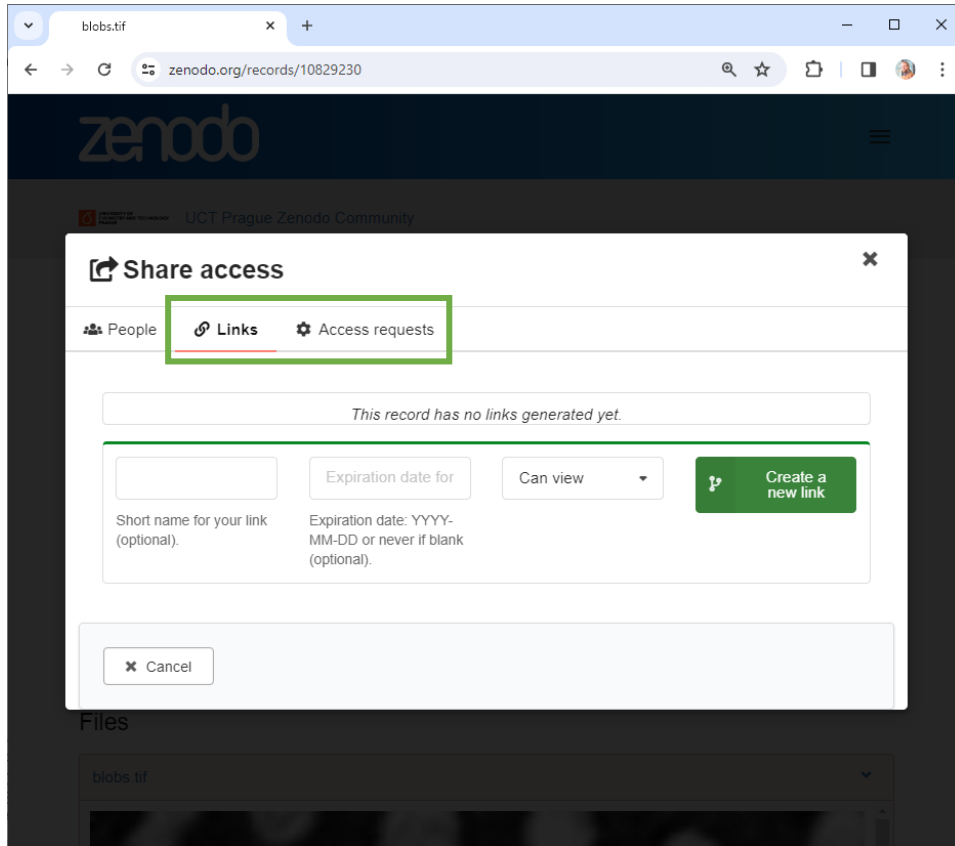
0
📄 DOWNLOADS

► Show more details



Accessibility

- The A in FAIR does not necessarily stand for Open Access



Accessibility

- The A in FAIR does not necessarily stand for Open Access

blobs.tif

zenodo.org/records/10829230

Share access

People Links **Access requests**

☒ Allow authenticated users to request access to restricted files.

☒ Allow non-authenticated users to request access to restricted files.

Enable users and guests to request access to your record's files. When access is requested by someone, you will get an e-mail asking for approval. After you approve a request, users will be granted access and guests will receive a secret link.

Accept conditions

Optional. Specify conditions under which you approve access. This message will be visible for any user when requesting access to this record.

Paragraph B I Link Quote Table List Bulleted List Numbered List Indent List More

Advanced options

blobs.tif

Published March 18, 2024 | Version v1

Dataset Restricted

26 VIEWS 0 DOWNLOADS

Show more details

Haase, Robert^{1,2} Schätz, Martin

Show affiliations

This dataset contains blobs.tif, which was published before as blobs.gif as part of ImageJ's example images. The dataset is public-domain, available online in png format as well: <https://samples.fiji.sc/blobs.png>

This record in Zenodo serves demonstrating that data can be published with closed access.

Files

Restricted

The record is publicly accessible, but files are restricted to users with access.

Request access

If you would like to request access to these files, please fill out the form below.

You are currently not logged in. Do you have an account? [Log in here](#)

Your email address * Your full name *

Email address Full name

Request message

I agree to that my full name and email address is shared with the owners of the record

Request access

Versions

Version v1	Mar 18, 2024
10.5281/zenodo.10829230	

Cite all versions? You can cite all versions by using the DOI [10.5281/zenodo.10829229](https://doi.org/10.5281/zenodo.10829229). This DOI represents all versions, and will always resolve to the latest one. [Read more](#).

External resources

Indexed in

OpenAIRE

Communities

UCT Prague Zenodo Community

Details

DOI [10.5281/zenodo.10829230](https://doi.org/10.5281/zenodo.10829230)

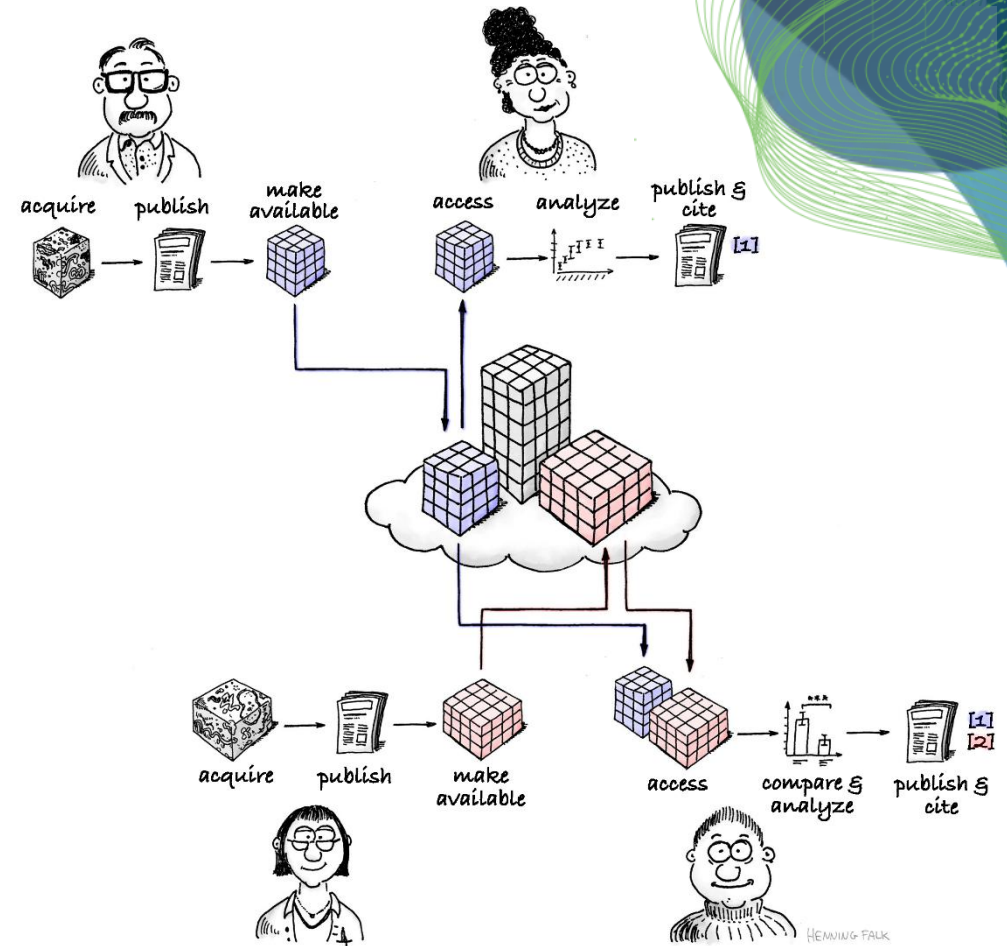
Interoperability

- I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (Meta)data use vocabularies that follow FAIR principles
- I3. (Meta)data include qualified references to other (meta)data



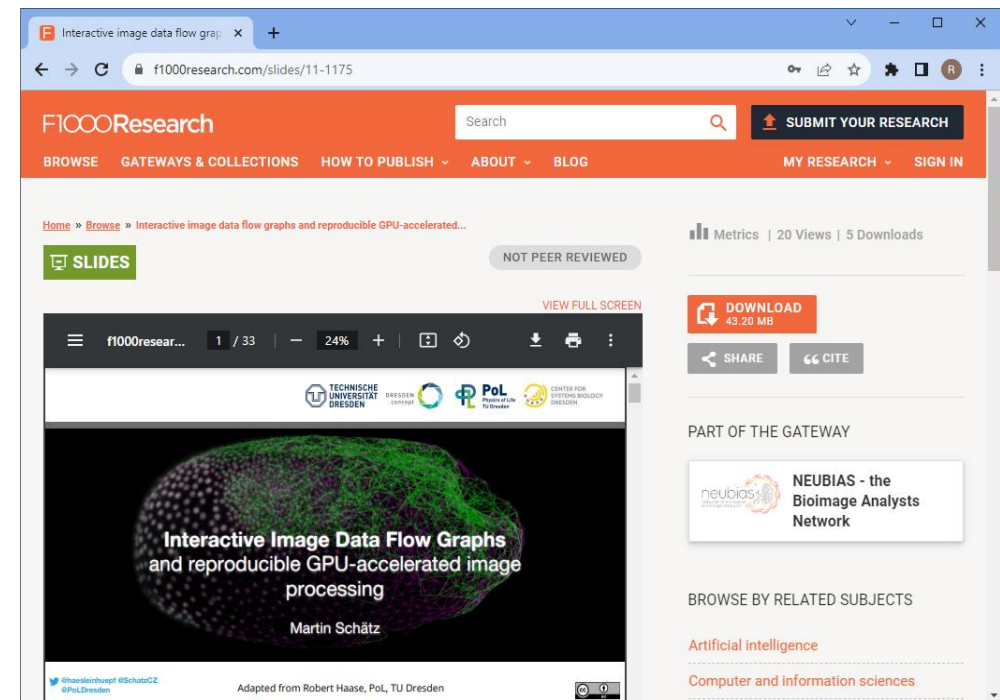
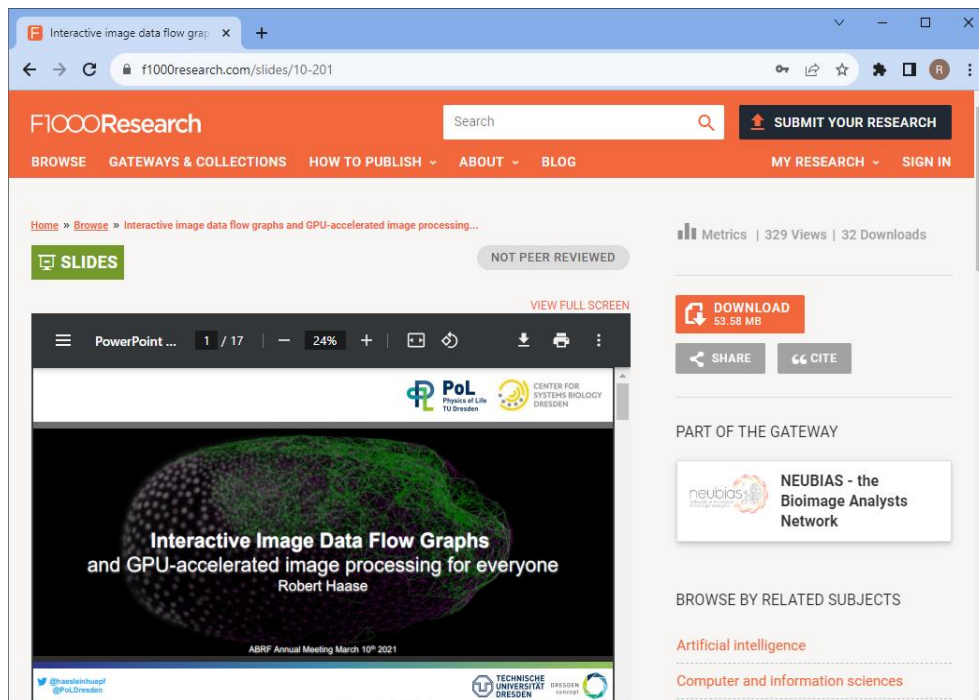
Reusability

- R1. (Meta)data are richly described with a plurality of accurate and relevant attributes
- R1.1. (Meta)data are released with a clear and accessible data usage license
- R1.2. (Meta)data are associated with detailed provenance
- R1.3. (Meta)data meet domain-relevant community standards



Incentives: Reusability

- Open Educational Resources
-> Others teach how to use your *tools & methods*



Incentives: Citability

Dataset associated with article: x +

zenodo.org/records/7516192

zenodo Search records... Communities My dashboard Log in Sign up

SEISMICA Seismica

Published September 22, 2022 | Version Uploaded at 2023-01-09

Dataset Open

Dataset associated with article: Self-sufficient seismic boxes for monitoring glacier seismology in Greenland

Ana Nap¹; Fabian Walter²; Martin P. Lüthi¹

Show affiliations

Dataset associated with article: Self-sufficient seismic boxes for monitoring glacier3 seismology in Greenland

Contains:

- Seismic data of both SG-boxes and regular geophones from Gornergletscher fieldtest, 2021(Seismic_Data_Gorner_Fieldtest.zip)
 - > SG-box data naming: GO*station_number*SG
 - > Geophone data naming: GO*station_number*GP
- Weather data Gornergletscher fieldtest from Monte Rosa, Meteo Swiss Weather station (Weather_data_Gorner_Fieldtest_2021.zip)
 - > 1hr wind averages
 - > 1hr temperature averages
- MSR logger data from SG-boxes from Gornergletscher fieldtest, 2021. Every 5 min these log battery power, tilt (along three axes, temperature and humidity inside the box and light strength on two sides of the SG-box. (MSR logger data SGboxes Gorner Fieldtest.zip)

240 VIEWS 131 DOWNLOADS

Show more details

Versions

Version	Uploaded at
Version Uploaded at 2023-01-09	Sep 22, 2022
10.5281/zenodo.7516192	
Version v1	Sep 22, 2022
10.5281/zenodo.7105052	

View all 2 versions

Cite all versions? You can cite all versions by using the DOI 10.5281/zenodo.7105051. This DOI represents all versions, and will always resolve to the latest one. [Read more.](#)

Weather_station_data_Greenland_2021.zip

md5:90d1bae30bde05352d97e09ee6897af1

Rights

Creative Commons Attribution 4.0 International

Citation

Ana Nap, Fabian Walter, & Martin P. Lüthi. (2022). Dataset associated with article: Self-sufficient seismic boxes for monitoring glacier seismology in Greenland (Uploaded at 2023-01-09) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7516192>

Style APA

Export

JSON Export

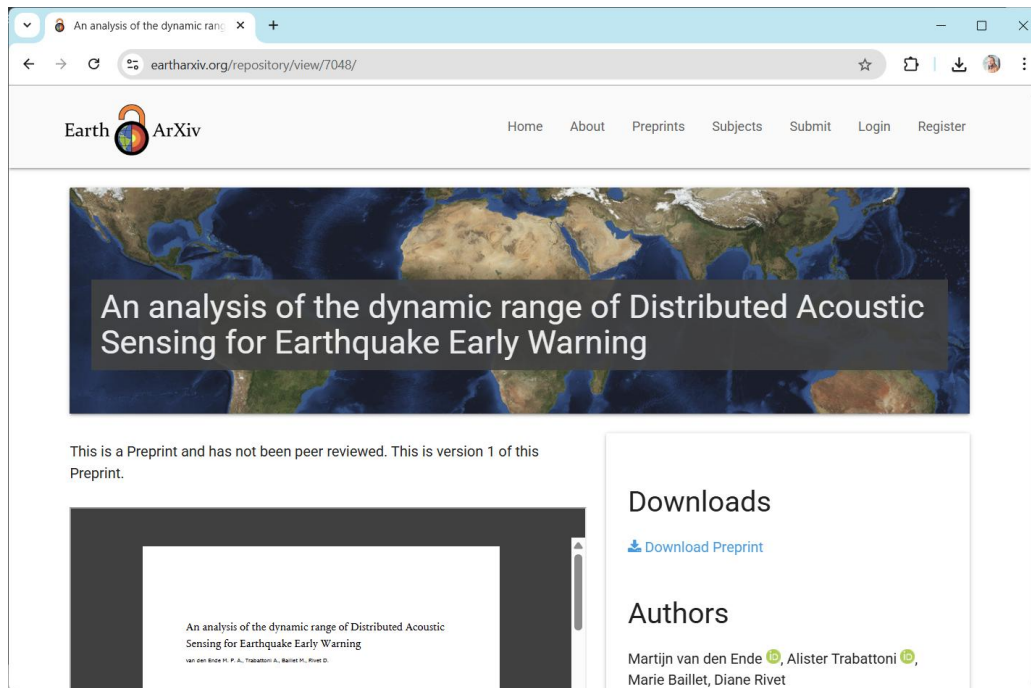
Technical metadata

Created January 9, 2023
Modified April 3, 2023

Where to share what?

Manuscripts

- <https://eartharxiv.org/>
- <https://arxiv.org/>

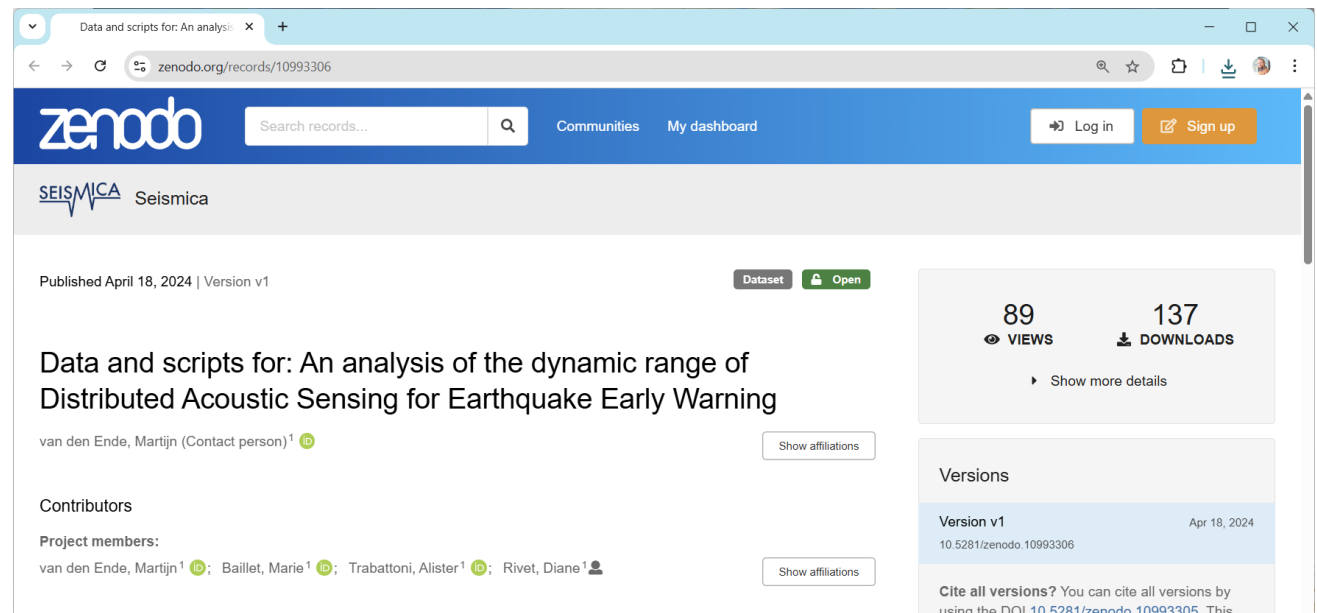


Code, Models & Data

- <https://github.com>
- <https://huggingface.co>
- <https://zenodo.org>

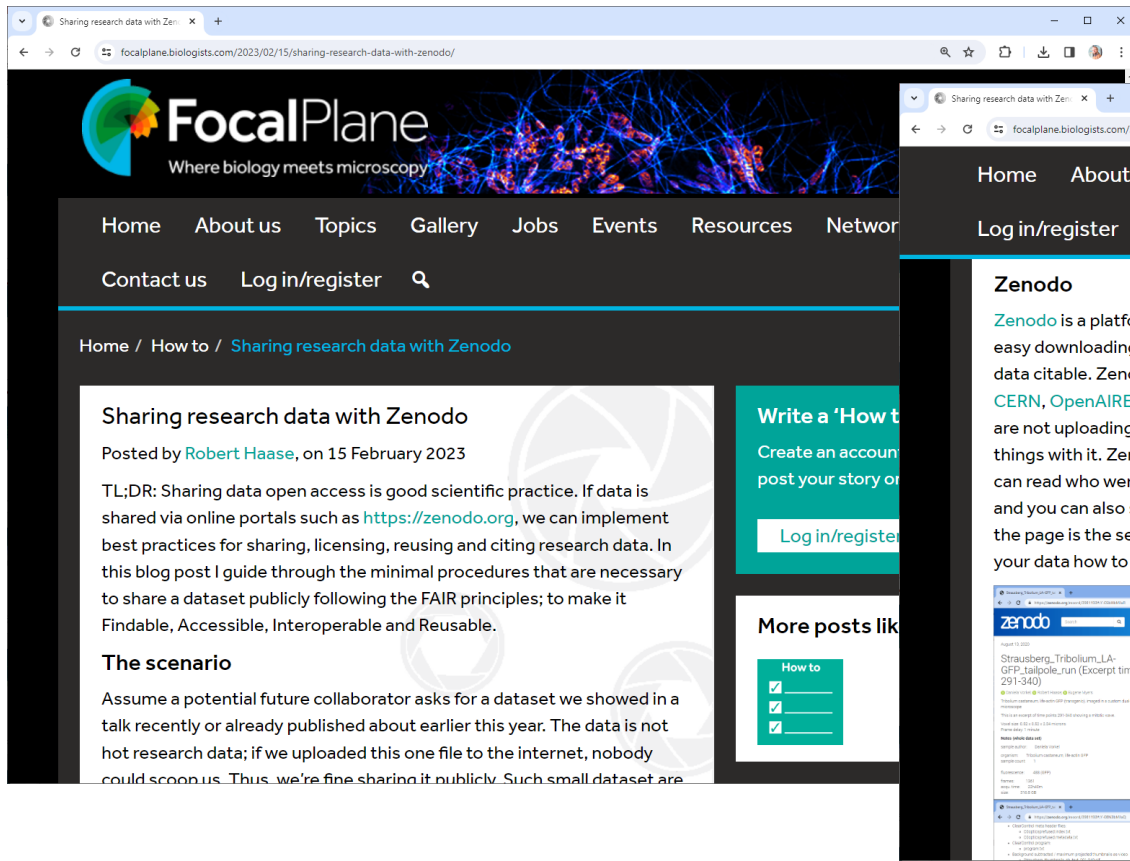
Ongoing work on code and models

Publicly funded archive



Sharing files on Zenodo

- ... is easier than you think



FocalPlane
Where biology meets microscopy

Home About us Topics Gallery Jobs Events Resources Network

Contact us Log in/register

Home / How to / **Sharing research data with Zenodo**

Sharing research data with Zenodo

Posted by **Robert Haase**, on 15 February 2023

TL;DR: Sharing data open access is good scientific practice. If data is shared via online portals such as <https://zenodo.org>, we can implement best practices for sharing, licensing, reusing and citing research data. In this blog post I guide through the minimal procedures that are necessary to share a dataset publicly following the FAIR principles; to make it Findable, Accessible, Interoperable and Reusable.

The scenario

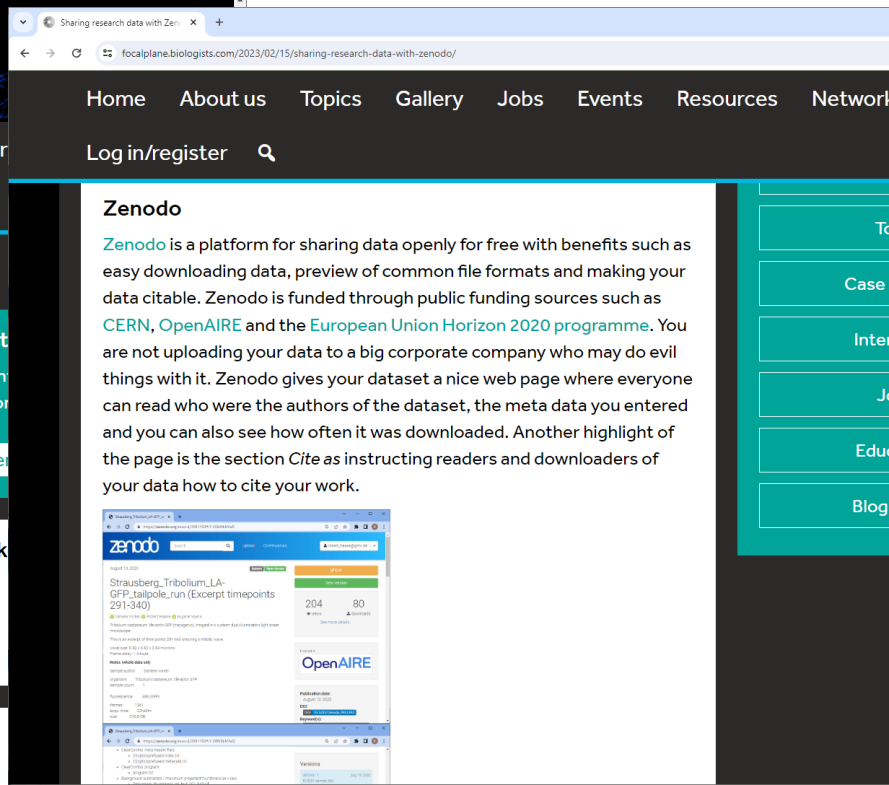
Assume a potential future collaborator asks for a dataset we showed in a talk recently or already published about earlier this year. The data is not hot research data; if we uploaded this one file to the internet, nobody could scoop us. Thus, we're fine sharing it publicly. Such small dataset are

Write a 'How to' post
Create an account
post your story on

Log in/register

More posts like this

- How to
- ✓
- ✓
- ✓



Home About us Topics Gallery Jobs Events Resources Network

Log in/register

Zenodo

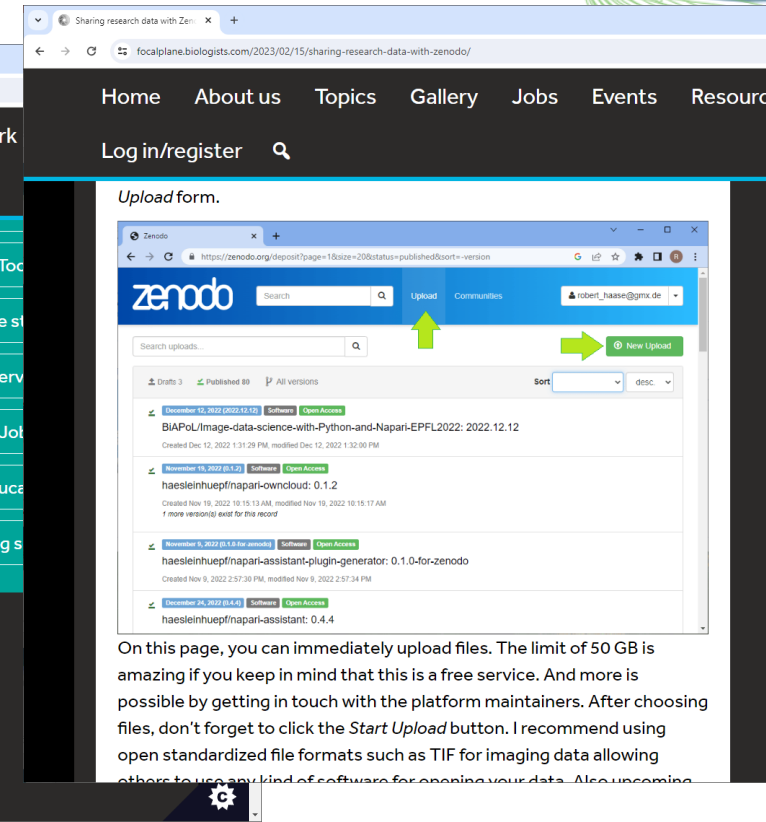
Zenodo is a platform for sharing data openly for free with benefits such as easy downloading data, preview of common file formats and making your data citable. Zenodo is funded through public funding sources such as CERN, OpenAIRE and the European Union Horizon 2020 programme. You are not uploading your data to a big corporate company who may do evil things with it. Zenodo gives your dataset a nice web page where everyone can read who were the authors of the dataset, the meta data you entered and you can also see how often it was downloaded. Another highlight of the page is the section *Cite* as instructing readers and downloaders of your data how to cite your work.

zenodo

Strausberg, Tribolium_LA-GFP_tailpole_run (Excerpt timepoints 291-340)

204 Downloads 80 Citations

OpenAIRE



Home About us Topics Gallery Jobs Events Resources Network

Log in/register

Upload form.

zenodo

Search uploads...

Upload

Communities

robot_haase@gmx.de

Search uploads...

Sort: desc

November 12, 2022 (0.1.0) Software Open Access

BIAPoL/image-data-science-with-Python-and-Napari-EPFL2022: 2022.12.12

Created Dec 12, 2022 1:31:29 PM, modified Dec 12, 2022 1:32:00 PM

November 19, 2022 (0.1.2) Software Open Access

haeseinhuepf/napari-owncloud: 0.1.2

Created Nov 19, 2022 10:15:13 AM, modified Nov 19, 2022 10:15:17 AM

1 more version(s) exist for this record

November 9, 2022 (0.1.0 for zenodo) Software Open Access

haeseinhuepf/napari-assistant-plugin-generator: 0.1.0-for-zenodo

Created Nov 9, 2022 2:57:30 PM, modified Nov 9, 2022 2:57:34 PM

November 24, 2022 (0.4.0) Software Open Access

haeseinhuepf/napari-assistant: 0.4.4

On this page, you can immediately upload files. The limit of 50 GB is amazing if you keep in mind that this is a free service. And more is possible by getting in touch with the platform maintainers. After choosing files, don't forget to click the **Start Upload** button. I recommend using open standardized file formats such as TIF for imaging data allowing others to use any kind of software for opening your data. Also upcoming

Research Software Management

Robert Haase

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



SACHSEN
Diese Maßnahme wird gefördert durch die Bundesregierung aufgrund eines Beschlusses des Deutschen Bundestages. Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des von den Abgeordneten des Sächsischen Landtags beschlossenen Haushaltes.

Quiz

- How do you typically install Python packages on your computer?

pip install ...



conda install ...

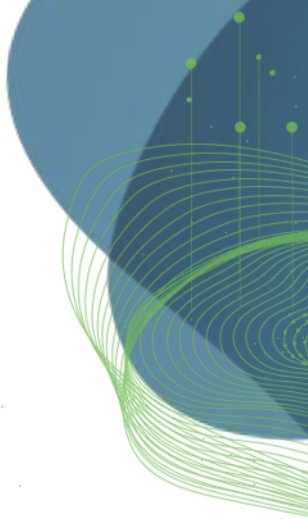


uv add ...



<other>





Conda environments

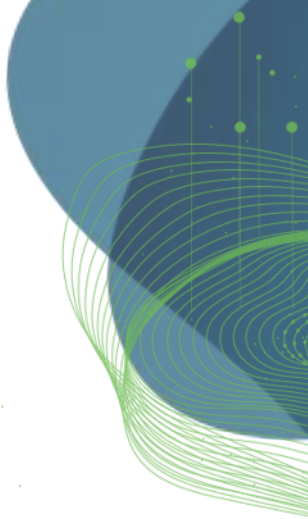
- What is the difference between

`pip install ...`

and

`conda install ...`

?

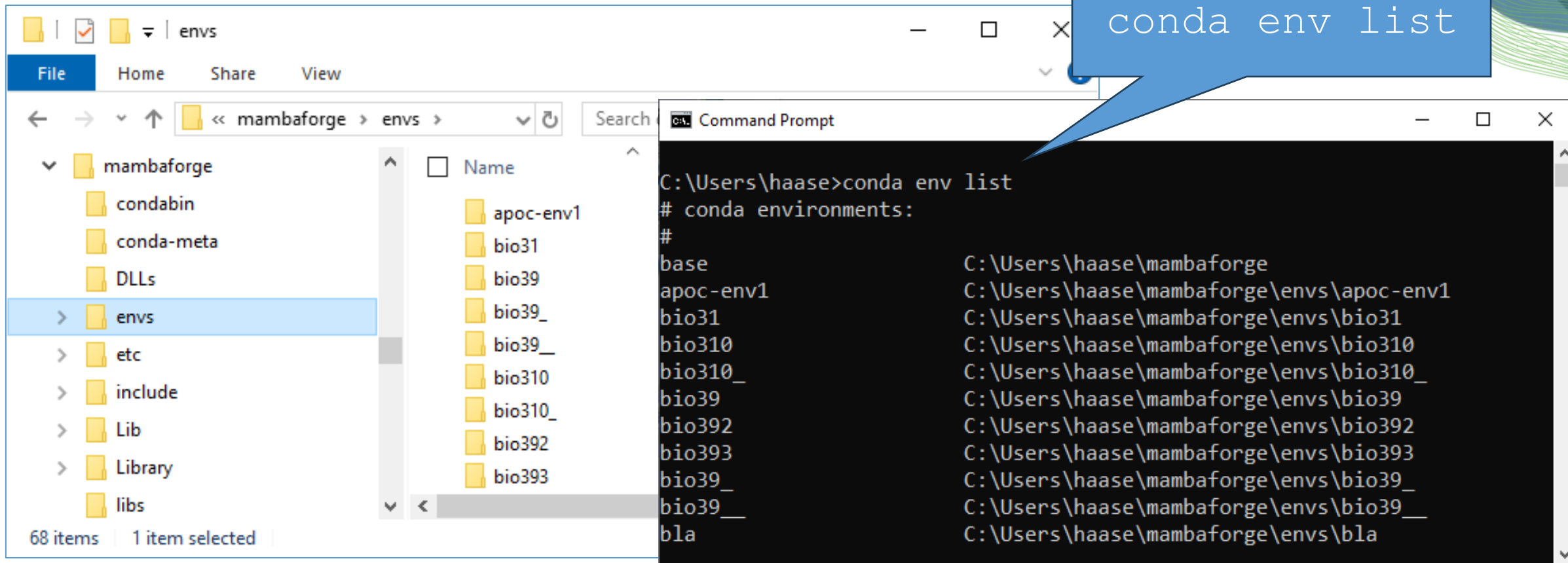


Conda environments

- Conda is a package manager
 - Allows to install Python packages
 - Allows to install other stuff (git, JDK)
- Conda is an environment manager
 - Virtual environments
 - Import/export/distribute environments

Conda environments

- A conda environment is *just a folder*



conda env list

File Explorer: mambaforge > envs

- condabin
- conda-meta
- DLLs
- envs
- etc
- include
- Lib
- Library
- libs

68 items | 1 item selected

Command Prompt:

```
C:\Users\haase>conda env list
# conda environments:
#
base                        C:\Users\haase\mambaforge
apoc-env1                   C:\Users\haase\mambaforge\envs\apoc-env1
bio31                       C:\Users\haase\mambaforge\envs\bio31
bio310                      C:\Users\haase\mambaforge\envs\bio310
bio310_                     C:\Users\haase\mambaforge\envs\bio310_
bio39                       C:\Users\haase\mambaforge\envs\bio39
bio392                      C:\Users\haase\mambaforge\envs\bio392
bio393                      C:\Users\haase\mambaforge\envs\bio393
bio39_                      C:\Users\haase\mambaforge\envs\bio39_
bio39__                     C:\Users\haase\mambaforge\envs\bio39__
bla                         C:\Users\haase\mambaforge\envs\bla
```

Conda environments

- Packages within the folders are just linked
- If you change a file in env1, the same file in env2 is changed.

Different virtual environments

Same physical file

NEVER edit those files!

Name	Date modified
aiosqlite	5/8/2023 5:20 PM
aiosqlite-0.19.0.dist-info	5/8/2023 5:20 PM
alabaster	5/8/2023 3:58 PM
alabaster-0.7.13.dist-info	5/8/2023 3:57 PM
anyio	5/8/2023 5:19 PM
anyio-3.6.2.dist-info	5/8/2023 5:19 PM
apiclient	4/29/2023 2:25 PM
apoc	5/8/2023 5:20 PM
apoc-0.12.0.dist-info	5/8/2023 5:20 PM
appdirs-1.4.4.dist-info	5/8/2023 3:57 PM
arbol	5/4/2023 6:49 AM
arbol-2021.5.27.641.dist-info	5/4/2023 6:49 AM
argparse	5/8/2023 5:20 PM

Broken [conda] environments

- Common case of confusion in the Python ecosystem
- Happens to everyone – it's just a matter of time.

The image is a collage of three screenshots illustrating a common issue in the Python ecosystem: a broken conda environment.

Left Screenshot: Shows the MAMBA logo and the command prompt output for installing 'stackview' using mamba. The output indicates that the installation is successful, with 142 packages installed and a total download of 0 B.

Middle Screenshot: Shows a list of installed packages and their versions, along with a summary of the environment. The packages listed include toolz, tornado, traitlets, typing_extensions, tzdata, ucrt, unicodedata2, vc, vs2015_runtime, wcwidth, wheel, widgetsnbextension, xorg-libxau, xorg-libxdmcp, xz, yaml, zeromq, zfp, zipp, zlib-ng, and zstd.

Right Screenshot: Shows a JupyterLab interface with a 'ModuleNotFoundError: No module named 'stackview'' error. The error message is displayed in a red box, indicating that the module 'stackview' is not found.

Background

- The base-environment is special.
- You install packages into the base environment, if you do not run `conda activate my_env`
- You can run base-software from within other environments

Base

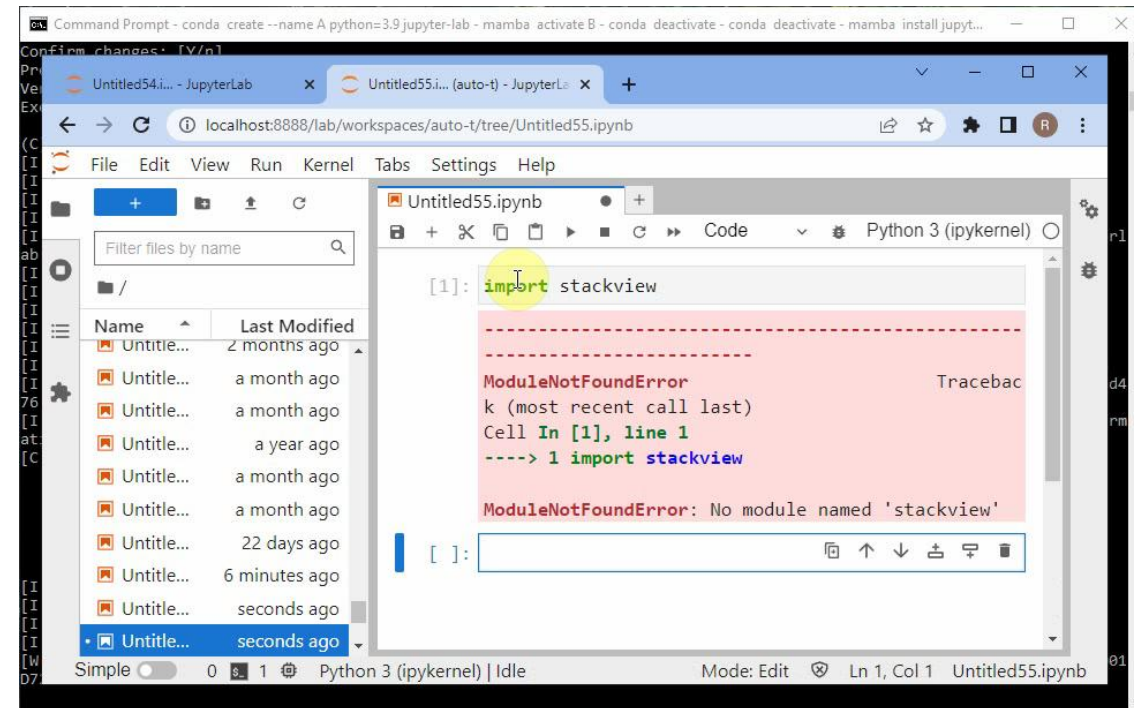
Installed packages:

Jupyter lab

my_env

Installed packages:

stackview



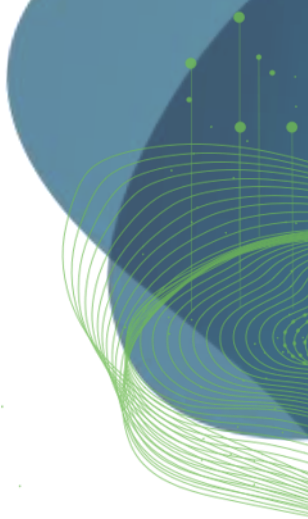
Broken [conda] environments

- The only cure: Uninstall and reinstall [Ana]conda.

Huge effort.

Little effort.

- Alternatively:
 - Create a new environment for every project.
 - `conda activate my_env`
 - Do not install project-specific stuff in the base-environment.



Broken [conda] environments

Do not install anything in the base environment

Always call **conda activate ...** before doing anything

Documenting dependencies

- Recommendation: Maintain a document with the dependencies (and versions) you need in your project
- The conda way
- The pip way

Your future self will thank you

```
environment.yml - Notepad
File Edit Format View Help
name: cellpose-test
channels:
  - conda-forge
  - pytorch
dependencies:
  - python=3.8
  - pytorch=1.12
  - devbio-napari
  - cellpose-napari
```

In case your environment is screwed up, you can rebuild it any time.

```
conda env create -f environment.yml
```

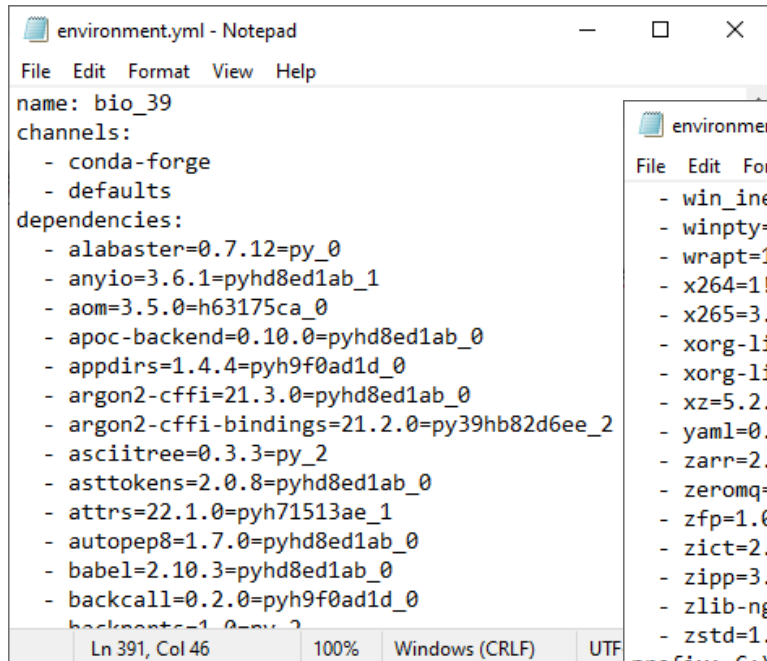
```
requirements.txt - Notepad
File Edit Format View Help
numpy>=1.21
andas~=1.4.0
napari~=0.4.11
matplotlib~=3.5.0
scikit-image~=0.19
umap-learn~=0.5.1
hdbscan~=0.8.27
scikit-learn~=1.0
numba~=0.55.1
```

```
pip install -r requirements.txt
```

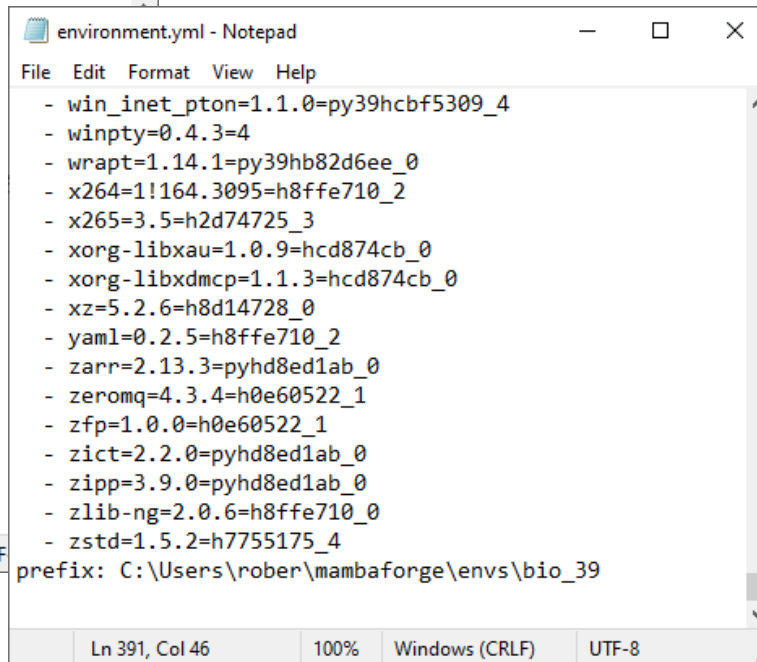
Documenting dependencies

- Exporting all installed libraries + versions

```
conda env export > environment.yml
```



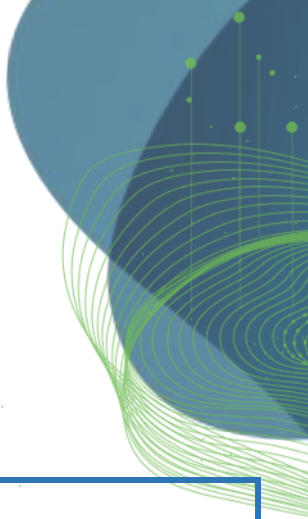
```
environment.yml - Notepad
File Edit Format View Help
name: bio_39
channels:
- conda-forge
- defaults
dependencies:
- alabaster=0.7.12=py_0
- anyio=3.6.1=pyhd8ed1ab_1
- aom=3.5.0=h63175ca_0
- apoc-backend=0.10.0=pyhd8ed1ab_0
- appdirs=1.4.4=pyh9f0ad1d_0
- argon2-cffi=21.3.0=pyhd8ed1ab_0
- argon2-cffi-bindings=21.2.0=py39hb82d6ee_2
- asciitree=0.3.3=py_2
- asttokens=2.0.8=pyhd8ed1ab_0
- attrs=22.1.0=pyh71513ae_1
- autopep8=1.7.0=pyhd8ed1ab_0
- babel=2.10.3=pyhd8ed1ab_0
- backcall=0.2.0=pyh9f0ad1d_0
- backports=1.0=py_2
```



```
environment.yml - Notepad
File Edit Format View Help
- win_inet_pton=1.1.0=py39hcbf5309_4
- winpty=0.4.3=4
- wrapt=1.14.1=py39hb82d6ee_0
- x264=1!164.3095=h8ffe710_2
- x265=3.5=h2d74725_3
- xorg-libxau=1.0.9=hcd874cb_0
- xorg-libxdmcp=1.1.3=hcd874cb_0
- xz=5.2.6=h8d14728_0
- yaml=0.2.5=h8ffe710_2
- zarr=2.13.3=pyhd8ed1ab_0
- zeromq=4.3.4=h0e60522_1
- zfp=1.0.0=h0e60522_1
- zict=2.2.0=pyhd8ed1ab_0
- zipp=3.9.0=pyhd8ed1ab_0
- zlib-ng=2.0.6=h8ffe710_0
- zstd=1.5.2=h7755175_4
prefix: C:\Users\rober\mambaforge\envs\bio_39
```

Excellent way to
document which
dependencies were
actually used...

It is *questionable* if re-
creating an environment
from this yml file works.



Installing dependencies

- A difference between pip and conda:

```
pip install package_a
```

```
conda install package_a
```

Depends on:
numpy<1.22.0



Dependency
conflict

```
pip install package_b
```

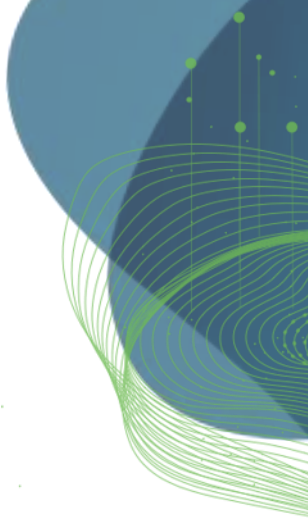
```
conda install package_b
```

Depends on:
numpy>=1.22.0

passes

fails

**Because the environment
cannot be solved**



Installing dependencies

- A difference between pip and conda:

```
pip install git
```

fails

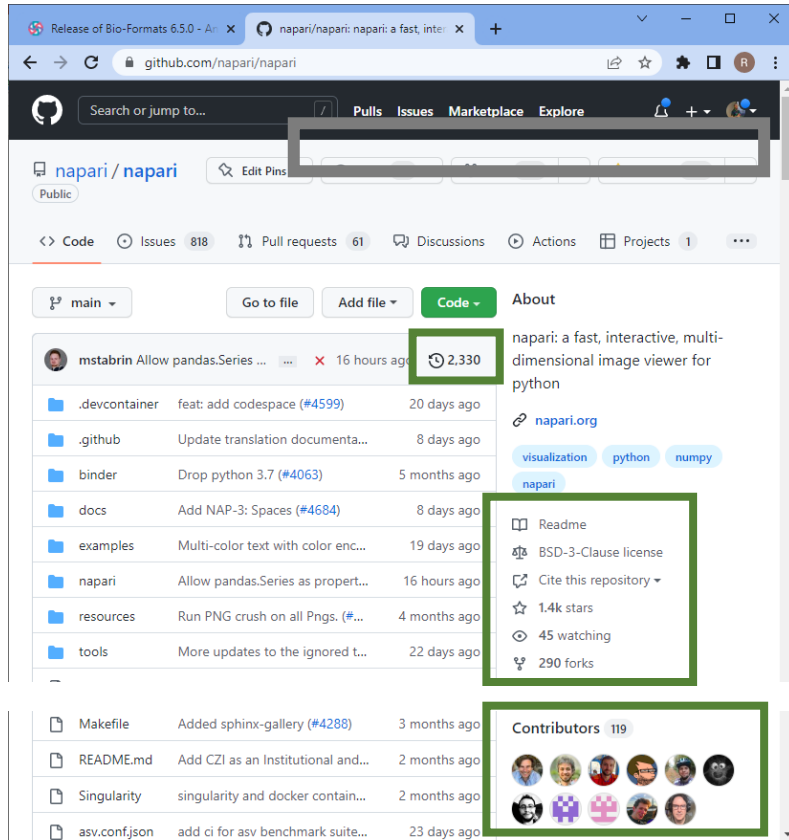
Because git is not a
python package

```
conda install git
```

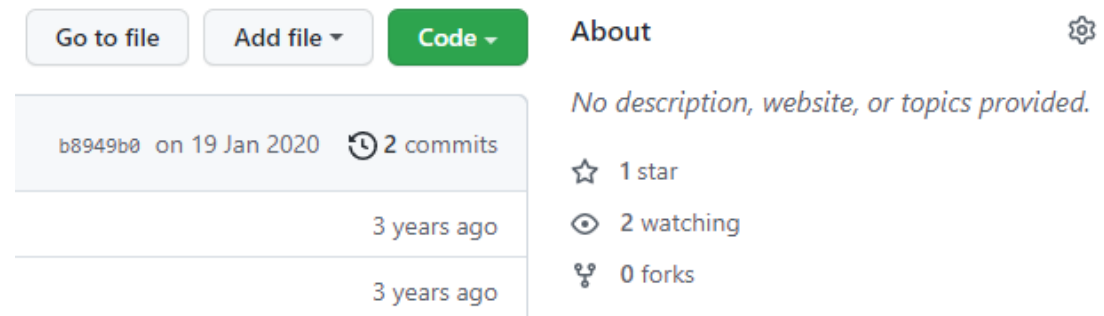
passes

Software quality indicators

- Visit the project's github or gitlab page and review indicators.

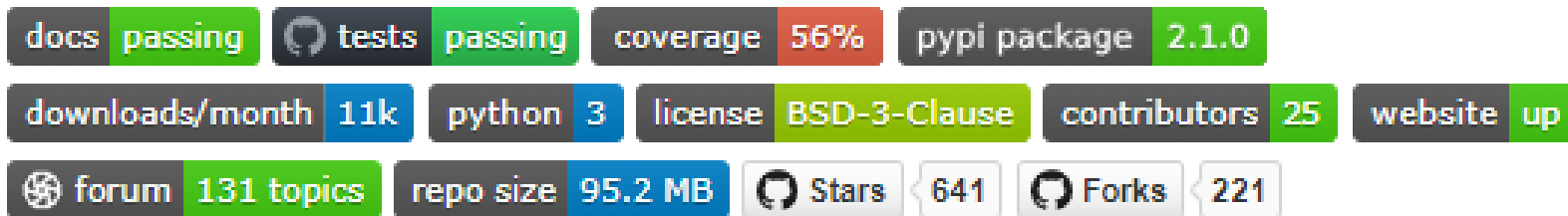


- **Stars:** People like software, similarly to posts on Bluesky
- **Watching:** People receive updates for new releases
- **Forks:** People made a copy of the code, e.g. to contribute to the project
- **Contributors:** People who contributed to the code
- **Commits:** Changes to the code



Software quality indicators

- Visit the project's github or gitlab page and review indicators.

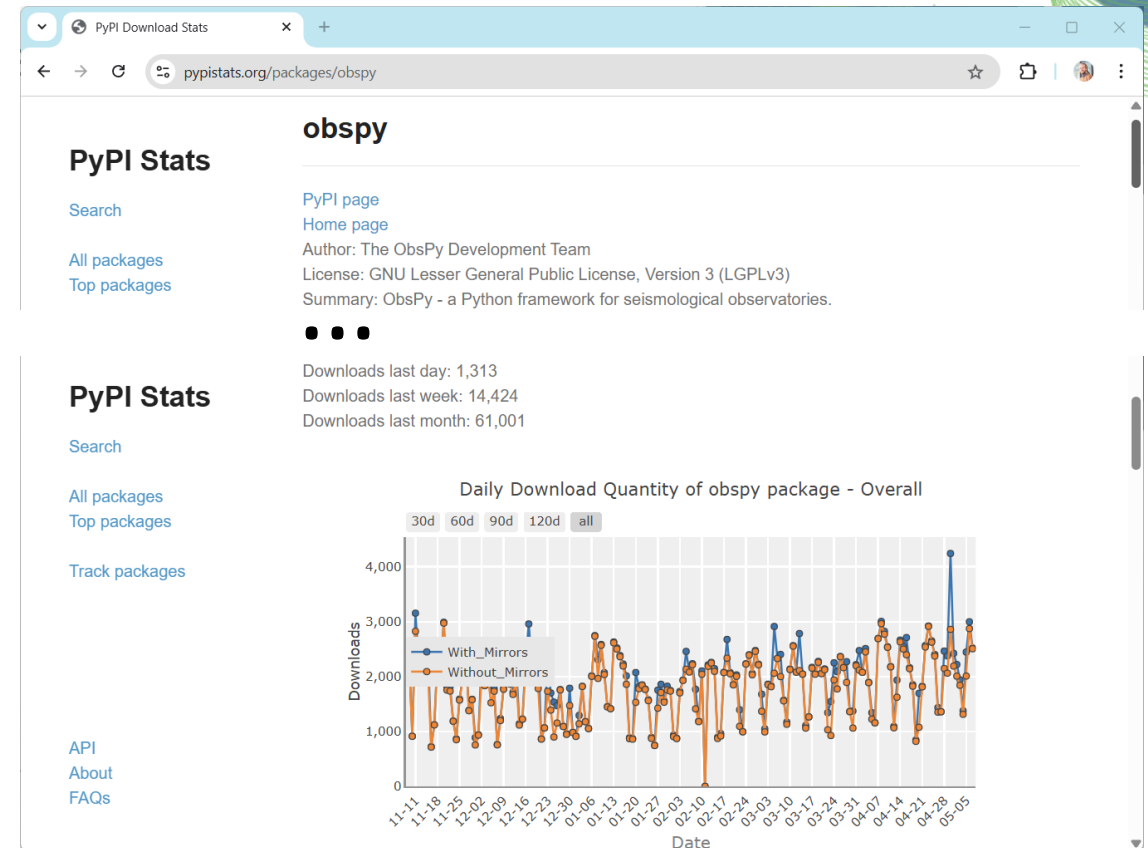
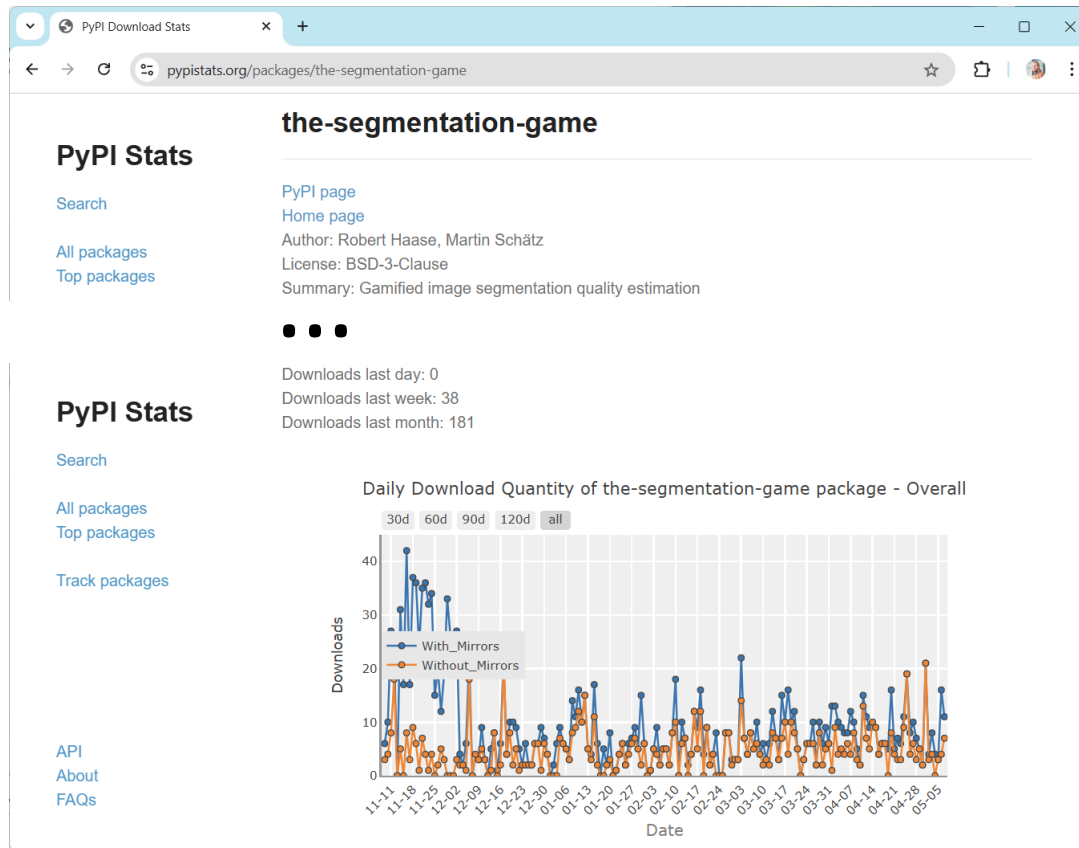


Note, github badges cannot be *deserved*. Developers put them there



Software quality indicators

- Download statistics: pypi



Software quality indicators

- Scientific publications

MouseLand/cellpose: a generalist algorithm for cellular segmentation

README BSD-3-Clause license

CITATION

If you use Cellpose 1, 2 or 3, please cite the Cellpose 1.0 [paper](#):
Stringer, C., Wang, T., Michaelos, M., & Pachitariu, M. (2021). Cellpose: a generalist algorithm for cellular segmentation. *Nature methods*, 18(1), 100-106.

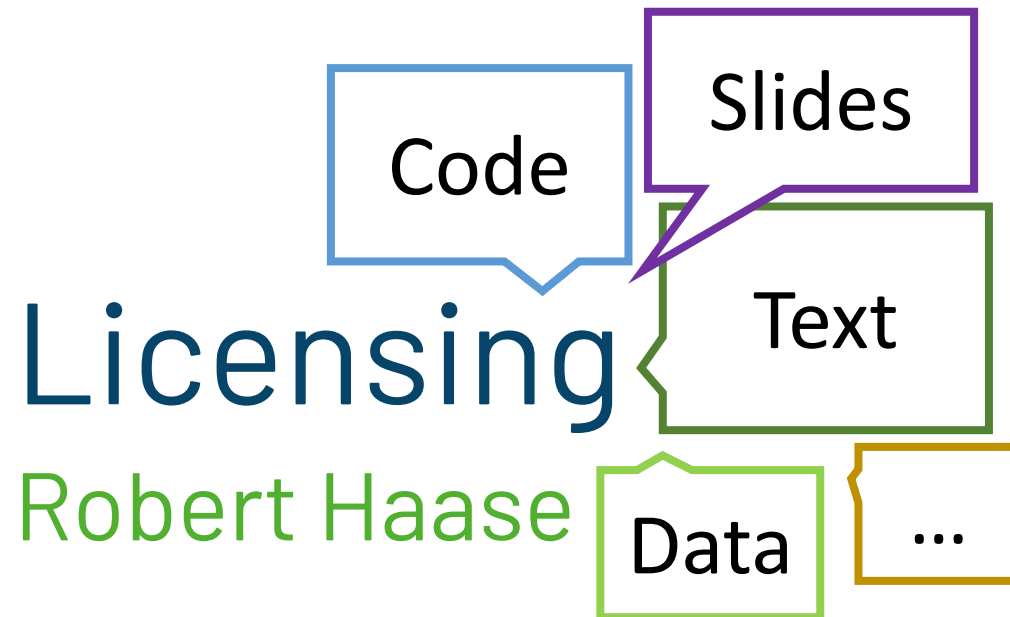
If you use the human-in-the-loop training, please also cite the Cellpose 2.0 [paper](#):
Pachitariu, M. & Stringer, C. (2022). Cellpose 2.0: how to train your own model. *Nature methods*, 1-8.

Cellpose 2.0: how to train your own model

Marius Pachitariu & Carsen Stringer

Nature Methods 19, 1634–1641 (2022) | [Cite this article](#)

55k Accesses | 120 Citations | 116 Altmetric | [Metrics](#)



Licensing

Robert Haase

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



SACHSEN
Diese Maßnahme wird gefördert durch die Bundesregierung
aufgrund eines Beschlusses des Deutschen Bundestages.
Diese Maßnahme wird mitfinanziert durch Steuermittel auf
der Grundlage des von den Abgeordneten des Sächsischen
Landtags beschlossenen Haushaltes.



Code

Slides

Text

Licensing

Robert Haase

What are the
consequences of
this sentence?

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



Diese Maßnahme wird gefördert durch die Bundesregierung
aufgrund eines Beschlusses des Deutschen Bundestages.
Diese Maßnahme wird mitfinanziert durch Steuermittel auf
der Grundlage des von den Abgeordneten des Sächsischen
Landtags beschlossenen Haushaltes.

Quiz

- When you shared materials publicly on the internet, which *license* did you use?

None



Public
Domain



Creative
Commons

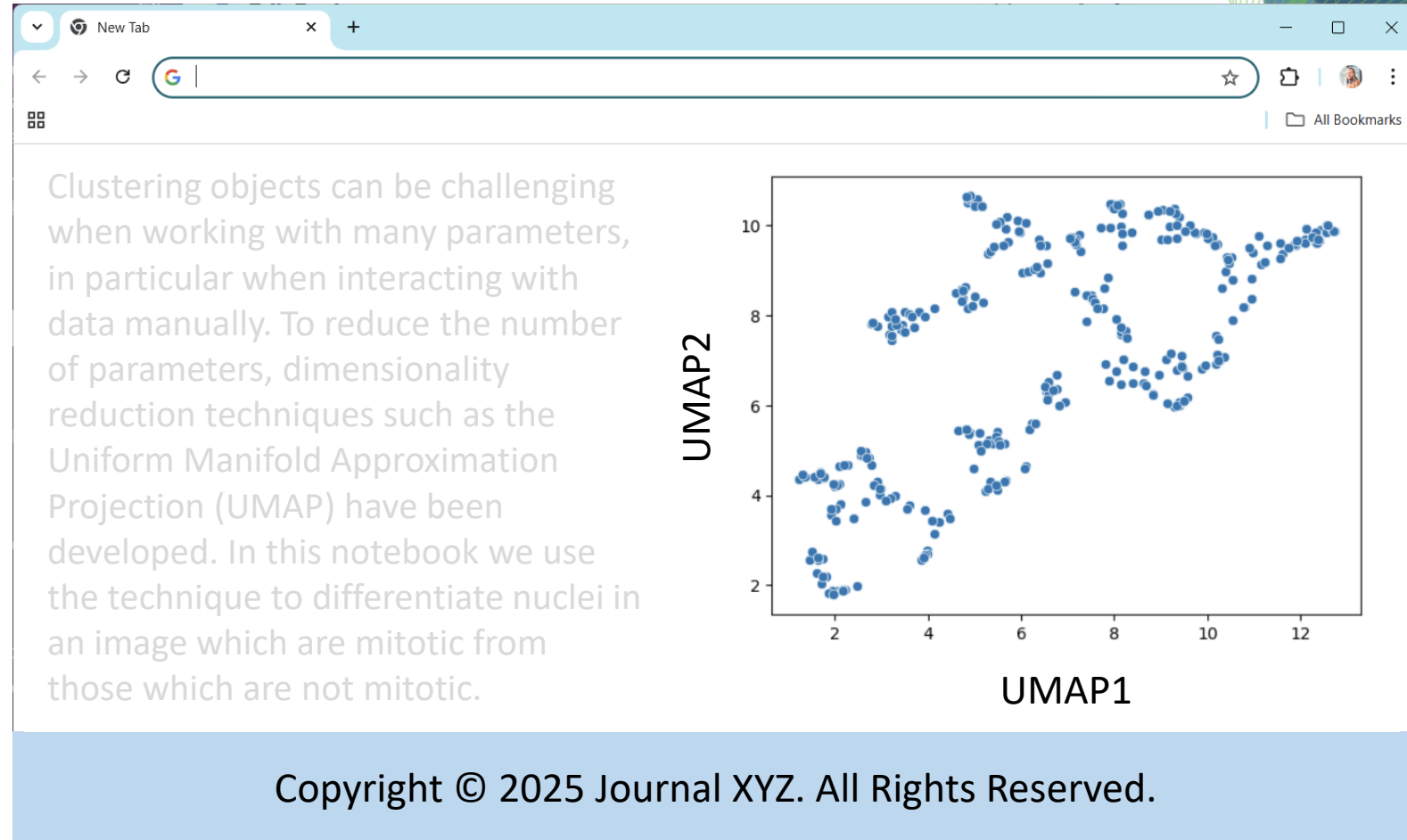


BSD/GPL/
MIT/...



Quiz

How can you reuse
this plot?
What is allowed?



Hint: Search for pre-prints

In case a journal doesn't allow reusing figures from a paper, search for the corresponding preprint!

nature.com/articles/s41592-019-0650-1

nature methods

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

Sign up for alerts RSS feed

nature > nature methods > correspondence > article

Correspondence | Published: 18 November 2019

CLIJ: GPU-accelerated image processing for everyone

Robert Haase, Loic A. Royer, Peter Steinbach, Deborah Schmidt, Alexandr Dibrov, Uwe Schmidt, Martin Weigert, Nicola Maghelli, Pavel Tomancak, Florian Jug, Eugene W. Myers

Nature Methods 17, 5–6 (2020) | [Cite this article](#)

8204 Accesses | 132 Citations | 87 Altmetric | [Metrics](#)

Access to this article via Universitätsbibliothek Leipzig AG eMedien is not available.

Change institution

Buy or subscribe

Sections: Data availability, Code availability, References, Acknowledgements, Author information, Ethics declarations, Supplementary information, Source data, Rights and permissions, About this article, This article is cited by

Access Nature and 54 other Nature Portfolio journals

Get Nature+, our best-value online-access subscription

24,99 € / 30 days
cancel any time

[Learn more](#)

biorxiv.org/content/10.1101/660704v2

bioRxiv

THE PREPRINT SERVER FOR BIOLOGY

HOME | SUBMIT | FAQ | BLOG | ALERTS / RSS | RESOURCES

CLIJ: GPU-accelerated image processing for everyone

Robert Haase, Loic A. Royer, Peter Steinbach, Deborah Schmidt, Alexandr Dibrov, Uwe Schmidt, Martin Weigert, Nicola Maghelli, Pavel Tomancak, Florian Jug, Eugene W. Myers

doi: <https://doi.org/10.1101/660704>

Now published in *Nature Methods* doi: [10.1038/s41592-019-0650-1](https://doi.org/10.1038/s41592-019-0650-1)

Abstract

Graphics processing units (GPU) allow image processing at unprecedented speed. We present CLIJ, a Fiji plugin enabling users with entry level experience in programming to benefit from GPU-accelerated image processing. Freely programmable workflows can speed up image processing in Fiji by factor 10 and more using high-end GPU hardware and on affordable mobile computers with built-in GPUs.

Copyright: The copyright holder for this preprint is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Figure 1: Performance comparison of CLIJ GPU-accelerated image processing. (a) Diagram of CLIJ architecture showing CPU and GPU components. (b) Line graphs showing processing time (s) vs. size (MB) for GaussianBlur3D and Minimum3D. (c) Table of relative speedup compared to Laptop GPU.

Operation	GPU	CPU	GPU
AddImagesWeighted2D	3	0.8	8
AddImagesWeighted3D	4	0.7	28
AddScalar2D	6	0.5	10
AddScalar3D	3	0.8	18
AutoThreshold2D	2	0.7	2
AutoThreshold3D	3	0.8	6
BinaryAnd2D	2	0.7	4
BinaryAnd3D	4	1.0	29
Erode2D	11	0.7	19
Erode3D	2	0.5	12
FixedThreshold2D	2	0.7	5
FixedThreshold3D	4	0.9	32
Flip2D	17	0.5	37
Flip3D	16	0.6	110
GaussianBlur2D	3	0.8	9
GaussianBlur3D	3	0.3	33
MaximumZProjection	0.1	0.8	38
Mean2D	0.06	0.6	11
Mean3D	0.07	4	172
Median2D	0.5	2	36
Median3D	2	4	42
Minimum2D	7	1.0	19
Minimum3D	25	4	196
MultiplyScalar2D	11	0.9	22
MultiplyScalar3D	6	1.0	38
RadialReslice	17	0.9	48
Rotate2D	3	0.5	32
Rotate3D	1.0	0.5	15

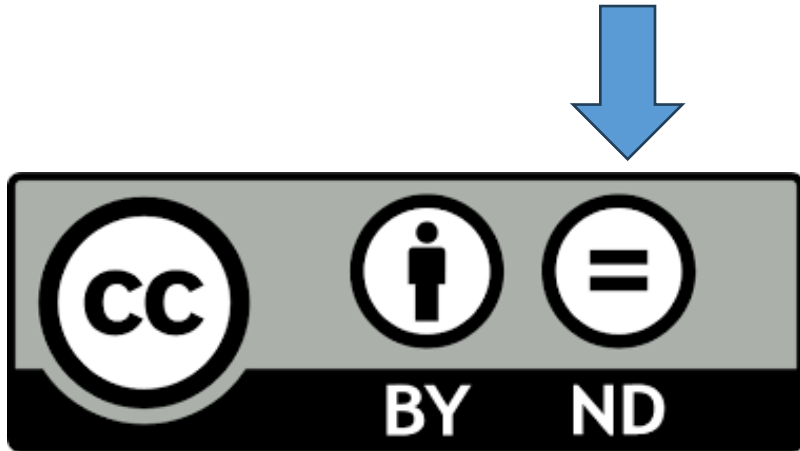
Relative speedup compared to Laptop GPU

GPU Laptop Workstation

Haase et al (2020), licensed [CC-BY 4.0](#)

Licensing: Permissive versus restrictive

- Who knows what the ND stands for?



You are free to:

Share — copy and redistribute the material in any medium or format for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

“permissive”

Under the following terms:

 **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

 **NoDerivatives** — If you remix, transform, or build upon the material, you may not distribute the modified material.

“restrictive”

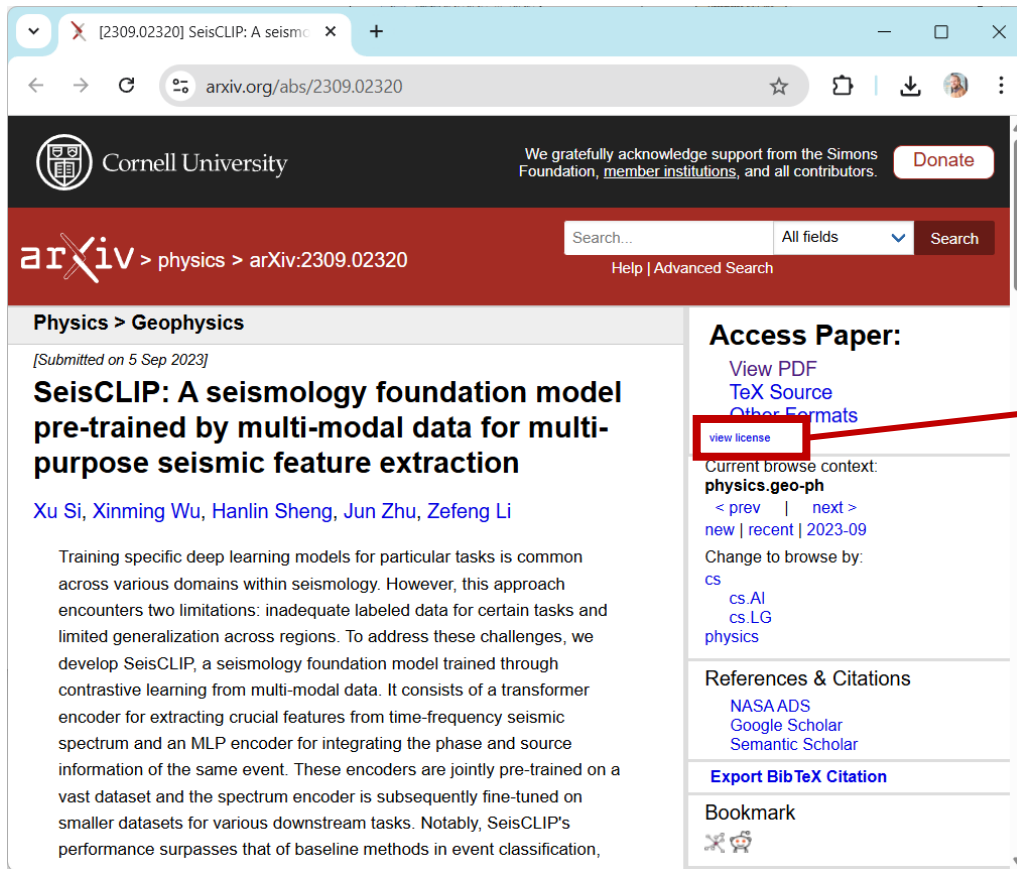
Licensing: Permissive versus restrictive

I hope nobody feels hurt
by the following slides.

I just would like to
make a point.

Licensing: Permissive versus restrictive

Example



I would love to show you a Figure from this paper!

arXiv.org - Non-exclusive license to distribute

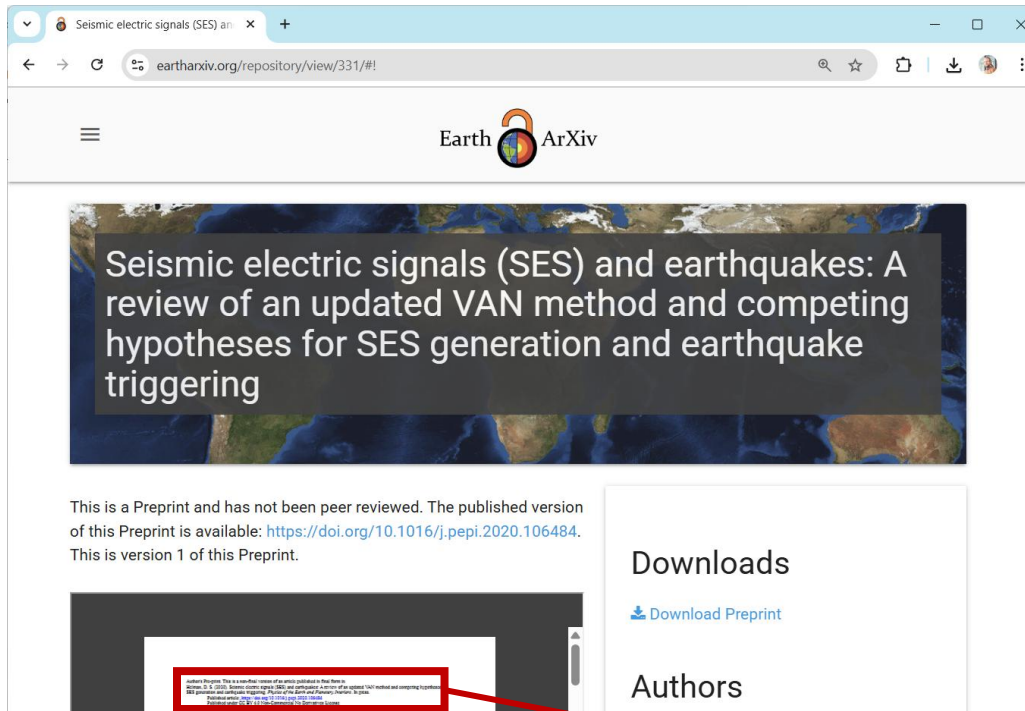
The URI <http://arxiv.org/licenses/nonexclusive-distrib/1.0/> is used to record the fact that the submitter granted the following license to arXiv.org on submission of an article:

- I grant arXiv.org a perpetual, non-exclusive license to distribute this article.
- I certify that I have the right to grant this license.
- I understand that submissions cannot be completely removed once accepted.
- I understand that arXiv.org reserves the right to reclassify or reject any submission.

But I'm not allowed!

Licensing: Permissive versus restrictive

Example



I would love to show you a Figure from this paper!

But I'm not allowed!

Author's Pre-print. This is a non-final version of an article published in final form in Helman, D. S. (2020). Seismic electric signals (SES) and earthquakes: A review of an updated VAN method and competing hypotheses for SES generation and earthquake triggering. *Physics of the Earth and Planetary Interiors*. In press. Published article: <https://doi.org/10.1016/j.pepi.2020.106484> Published under CC BY 4.0 Non-Commercial No Derivatives License

Licensing: Permissive versus restrictive

Example

SPRINGER LINK

Find a journal Publish with us Track your research Search



Annual Conference on Artificial Intelligence

→ KI 2011: **KI 2011: Advances in Artificial Intelligence** pp 111–121 | Cite as

Home > KI 2011: Advances in Artificial Intelligence > Conference paper

Swarm Intelligence for Medical Volume Segmentation: The Contribution of Self-reproduction

Robert Haase, Hans-Joachim Böhme, Daniel Zips & Nasreddin Abolmaali

Conference paper

1628 Accesses | 2 Citations | 3 Altmetric

Part of the [Lecture Notes in Computer Science](#) book series (LNAI, volume 7006)

Abstract

For special applications in diagnostics for oncology the analysis of imaging data from Positron Emission Tomography (PET) is obfuscated by low contrast and high noise. To deal with this issue we propose a segmentation algorithm based on Ant Colony Optimization (ACO) and evolutionary selection of ants for self reproduction. The self reproduction approach is no standard for ACO, but appears to be crucial for volume segmentation. This investigation was focused on two different ways for reproduction control and their contribution to quantity and

Quick Price Estimate

This service provides licences or permission to reuse content only.

If you do not have a copy of the article you are reusing content from, you may copy and paste the content from any legally accessible source and reuse according to the terms of your licence.

To purchase the article's PDF please visit either [Link.springer.com](#) or [Nature.com](#).

Adaptations/modifications - Springer Nature allows adaptation of figures for style and formatting purposes under this license under the condition that this does not alter the meaning of the content.

I would like to...?

reuse in training materials/CME material

Circulation/distribution?

30 - 99

I am a/an...?

academic/university or research institute

Are you the author of this Springer Nature content? ?

no

Is this reuse sponsored by or associated with a pharmaceutical or a medical products company?

no

I will be translating... ?

no

My format is... ?

electronic

My currency is...

EUR - €

I would like to use... ?

figures/tables/illustrations

Quick Price

100.91 EUR

Number of figures/tables ?

1

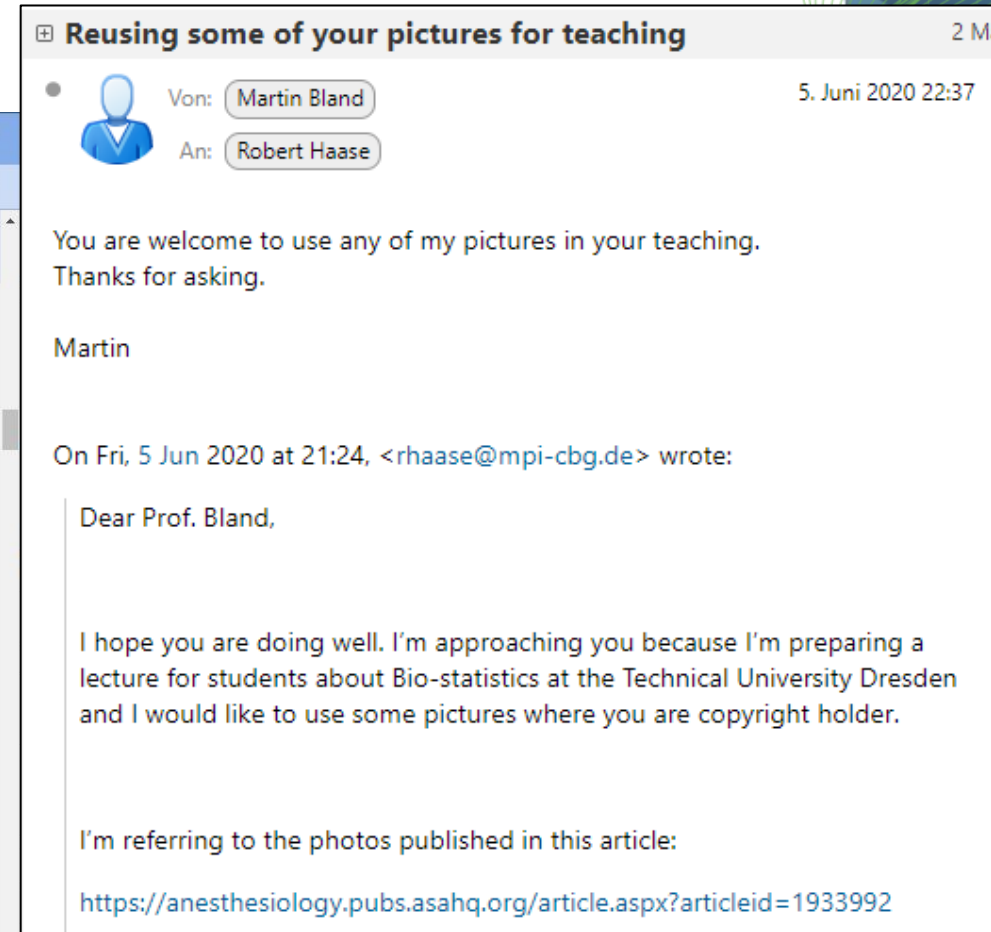
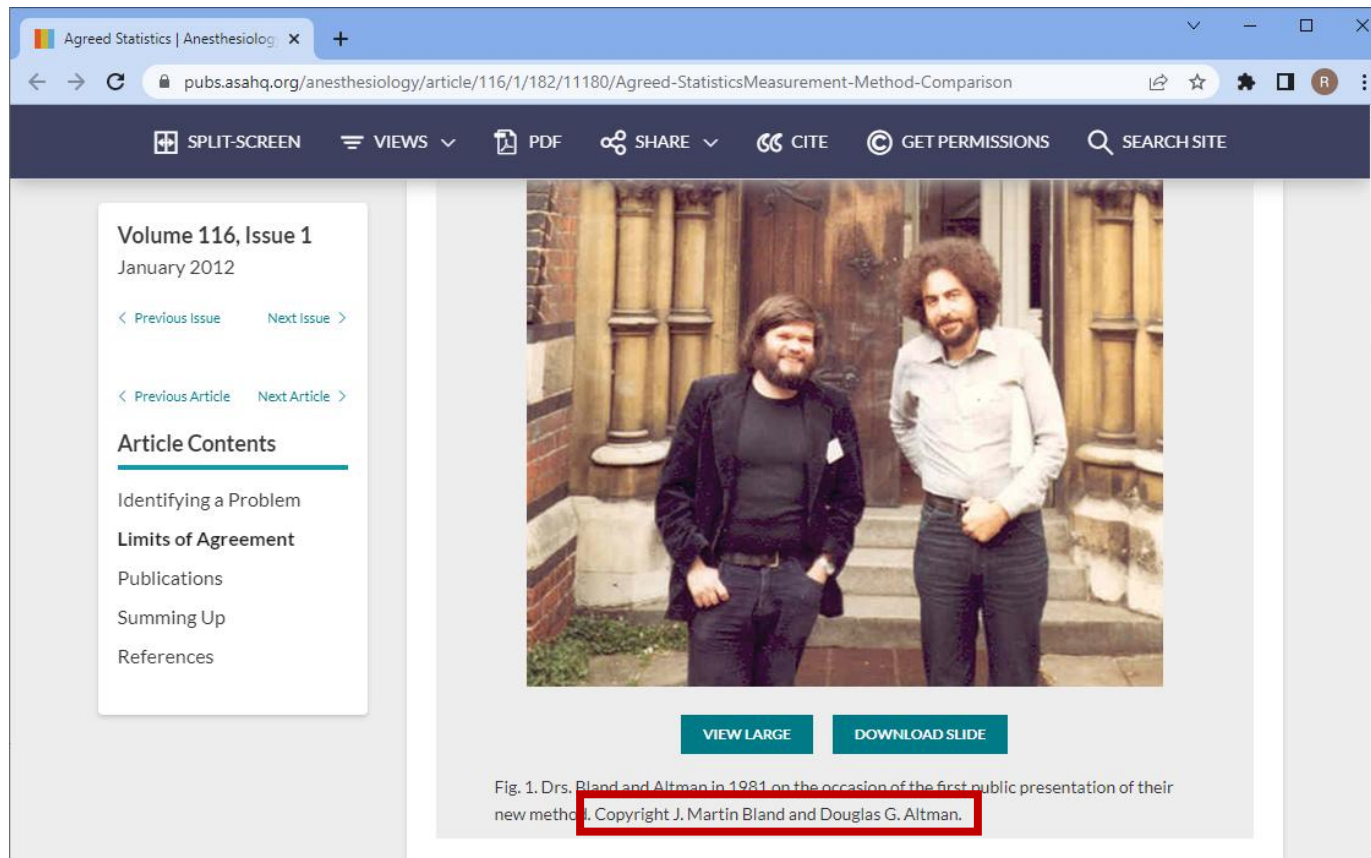
QUICK PRICE

CONTINUE

The issue is not so much paying 100 Eur, but the related administrative effort.

Licensing: Permissive versus restrictive

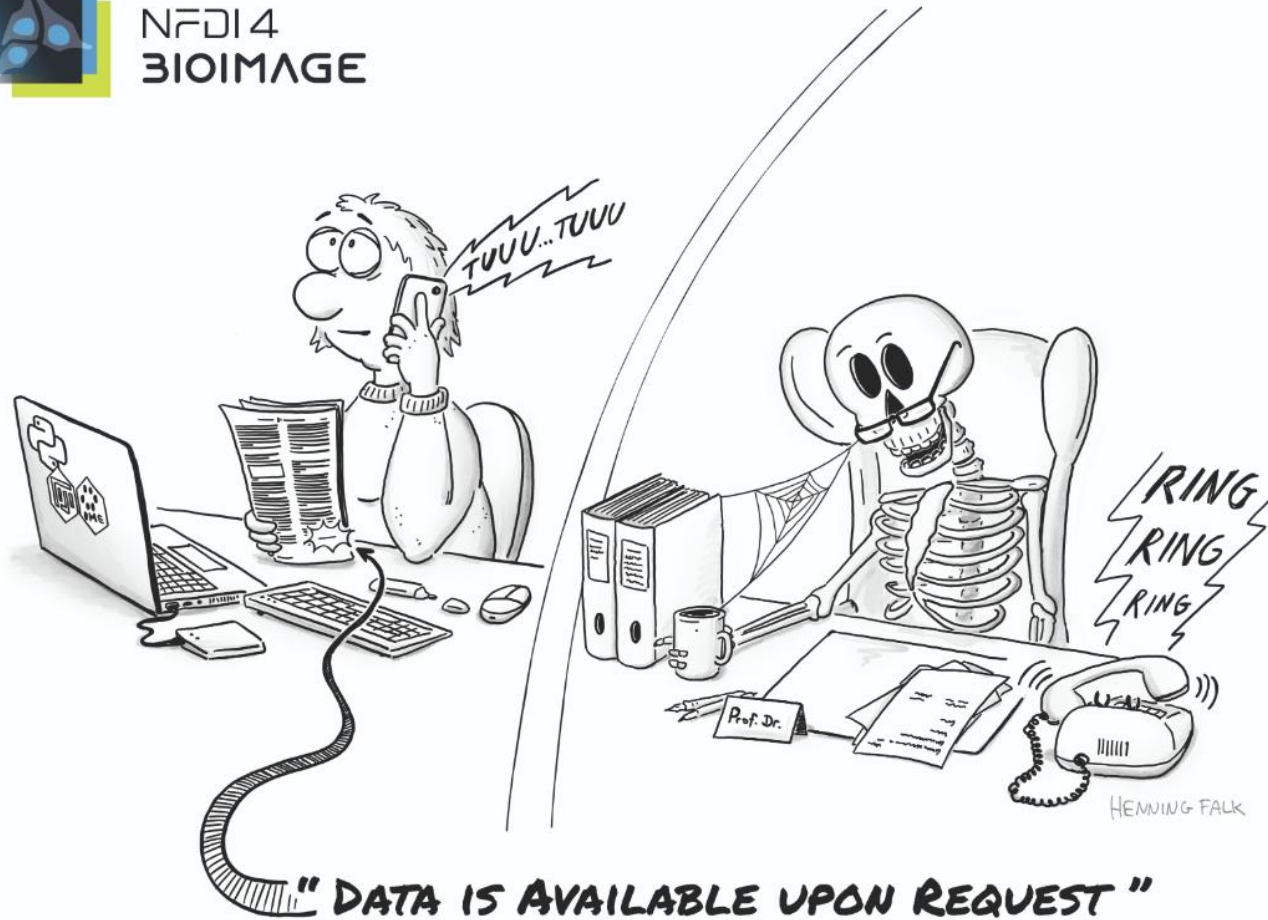
Example when you need to ask



Licensing: Permissive versus restrictive



NFDI 4
BIOIMAGE



Licensing: Permissive versus restrictive

Restrictive licensing is
a community-wide issue.

I presume due to lack of
awareness & training

Train the trainers!

Licensing: Permissive versus restrictive

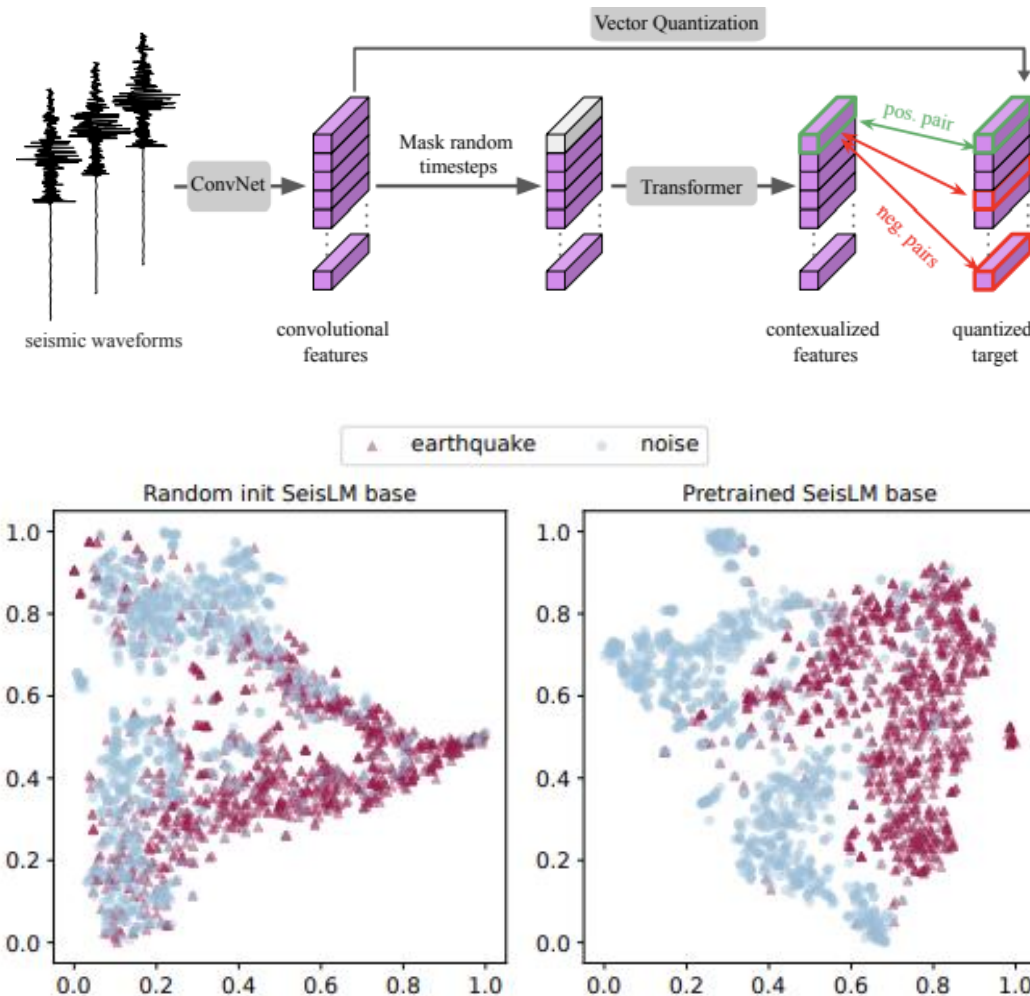
	Download and share for free	Reuse parts, e.g. Figures	Reuse parts, e.g. in paid training
CC-BY	✓	✓	✓
CC-BY-SA	✓ Only under CC-BY-SA	✓ Only under CC-BY-SA	✓ Only under CC-BY-SA
CC-BY-NC	✓	✓ (if free of charge)	✗
CC-BY-ND	✓	✗	✗
CC-BY-NC-ND	✓	✗	✗

Bad for the progress of science

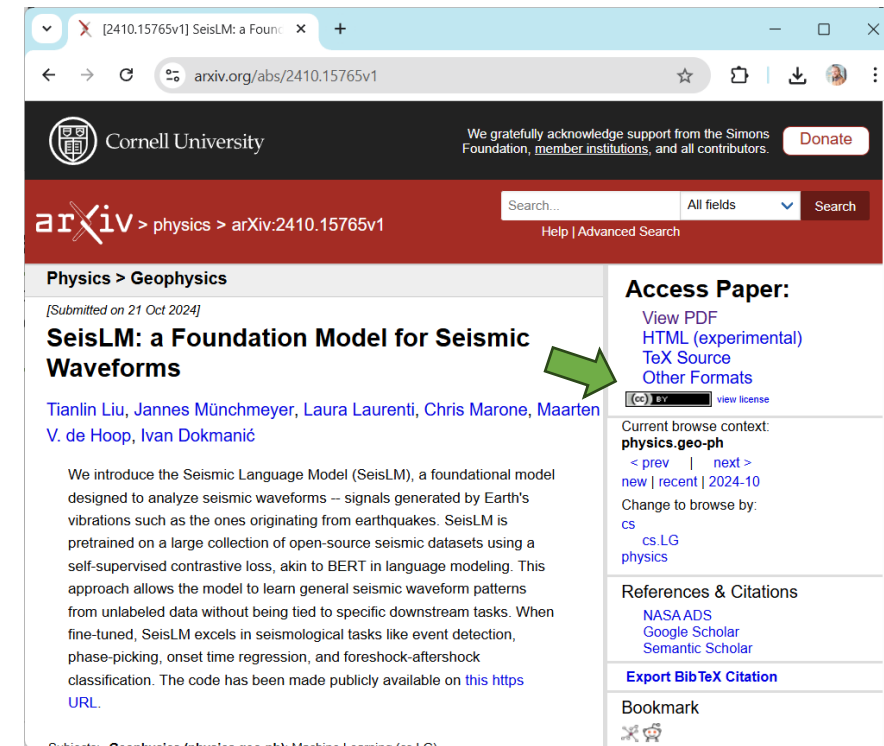
In particular in the context of training

Licensing: Permissive versus restrictive

Example



Look at these great figures! They are from <https://arxiv.org/abs/2410.15765v1> licensed [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/) by T. Liu et al.



Licensing: Permissive versus restrictive

Example

Look at this great figure! It's taken from M. Hartley et al.

BioImage Archive Search BioImages

Home Browse Submit About us Feedback Login

Release Date: 31 May 2021

The glucosylceramide synthase inhibitor PDMP causes lyso-somal lipid accumulation and mTOR inactivation

Pia Hartwig¹, Doris Höglinger¹

¹Heinrich Heine University

Accession **S-BIAD144**

Description We investigated subcellular sphingolipid distribution using a functionalized sphingosine analogue (pacSph) by confocal microscopy in control, PDMP and NB-DNJ (Miglustat) treated WT and SGPL1 knock-out cells (experimentA_11_WT_Miglustat.czi, experimentA_12_SGPL1_PDMP.czi, experimentA_13_SGPL1_PDMP.czi)

Data files

Name	Size	Section	staining	cells	labelling	treatment	Channel 1	Channel 2	timepoint
experimentA_11_WT_Miglustat.czi	1.6 MB	Study Component	click chemistry and IF	WT	pacSph	50 µM NB-DNJ (Miglustat)	pacSph	Lamp1	continuous labelling
experimentA_12_SGPL1_PDMP.czi	1.6 MB	Study Component	click chemistry and IF	SGPL1-/-	pacSph	20 µM PDMP	pacSph	Lamp1	continuous labelling
experimentA_13_SGPL1_PDMP.czi	1.6 MB	Study Component	click chemistry and IF	SGPL1-/-	pacSph	20 µM PDMP	pacSph	Lamp1	continuous labelling

Showing 11 to 15 of 223 entries

bioRxiv THE PREPRINT SERVER FOR BIOLOGY

HOME | SUBMIT | FAQ | BLOG | ALERTS / RSS | ABOUT | CHANNELS

Search Advanced Search

New Results Follow this preprint

The Biolmage Archive - building a home for life-sciences microscopy data

Matthew Hartley, Gerard J. Kleywegt, Ardan Patwardhan, Ugis Sarkans, Jason R. Swedlow, Alvis Brazma

doi: <https://doi.org/10.1101/2021.12.17.473169>

Now published in *Journal of Molecular Biology* doi: 10.1016/j.jmb.2022.167505

Abstract Full Text Info/History Metrics Preview PDF

ARTICLE INFORMATION

doi <https://doi.org/10.1101/2021.12.17.473169>

History February 11, 2022.

ARTICLE VERSIONS

Version 1 (December 21, 2021 - 20:06).

You are viewing Version 2, the most recent version of this article.

Copyright The copyright holder has placed this preprint in the Public Domain. It is no longer restricted by copyright. Anyone can legally share, reuse, remix, or adapt this material for any purpose without crediting the original authors.

COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv

Subject Area Bioinformatics

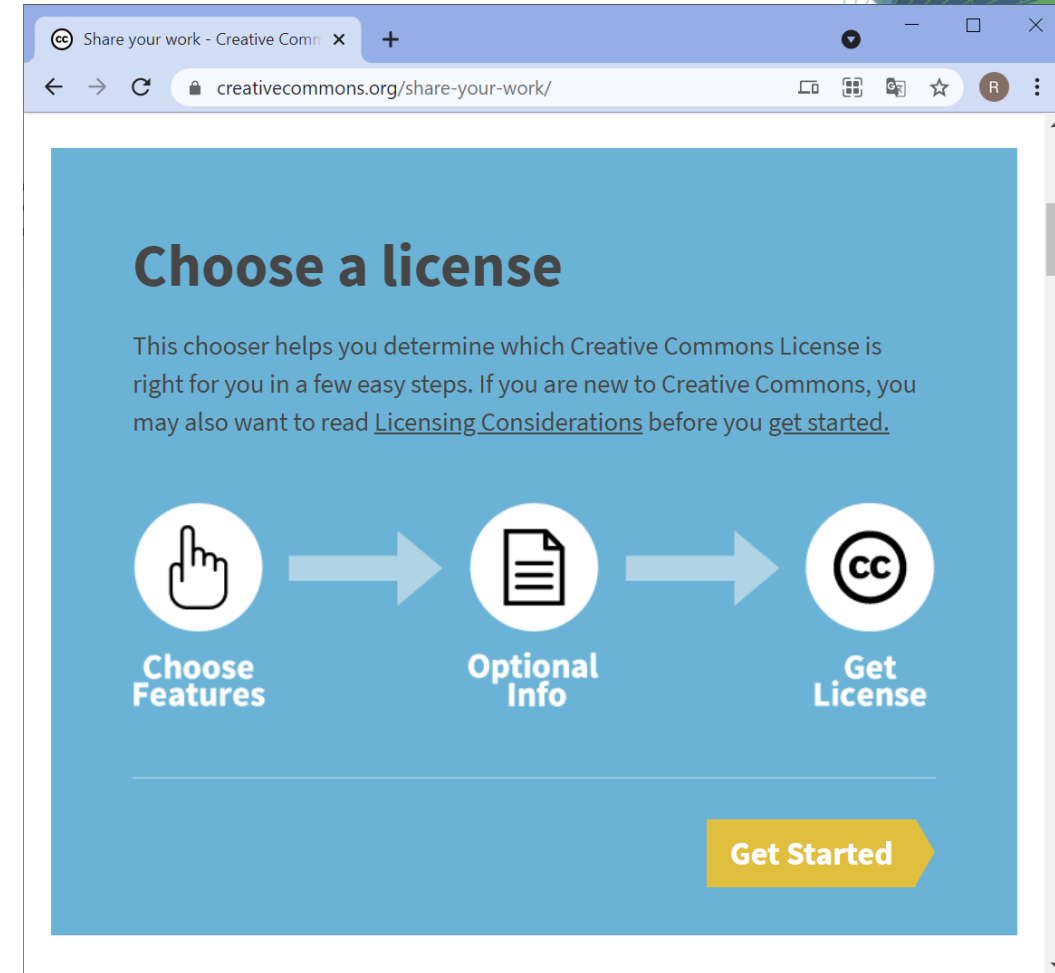
Subject Areas

- Animal Behavior and Cognition
- Biochemistry
- Bioengineering
- Bioinformatics

Licensing: Creative Commons (CC)

- Public domain (CC0)
- Attribution International (CC-BY)
- Attribution ShareAlike Int. (CC-BY-SA)
- Attribution Non-Commercial Int. (CC-BY-NC)
- Attribution NoDerivatives Int. (CC-BY-ND)

+ Combinations, e.g. CC-BY-NC-ND



Content on this site is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Icons by The Noun Project.

Licensing: Creative Commons (CC)



- **Public domain (CC0)**
- Everyone can reuse without mentioning the source or author of the shared resource.
- Public domain licenses cannot be revoked.
- The author must own the right to copy (copyright) the resource.
 - If you authored work as part of your job, you may not be the copyright holder. (check your employers' guidelines)

Employers don't like this one because you give away the rights to *exploit* your work.

Content on this slide was adapted from <https://creativecommons.org/about/cclicenses/> which is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Icons by The Noun Project.

Licensing: Creative Commons (CC)



“By attribution”

CC BY: This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license allows for commercial use.

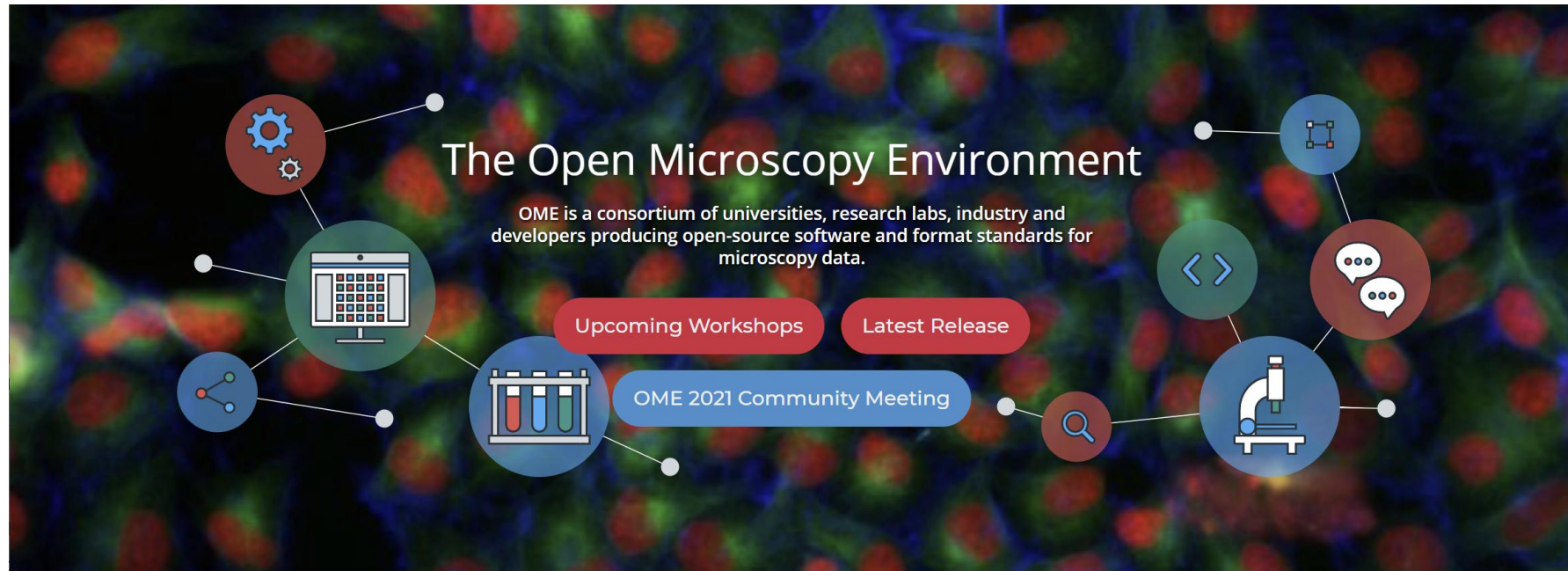
CC BY includes the following elements:

BY  – Credit must be given to the creator

Content on this slide was adapted from <https://creativecommons.org/about/ccllicenses/>, which is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Icons by The Noun Project.

Licensing: Creative Commons (CC)

Example



You must put such a sentence and keep the link to CC-BY

Figure adapted from <https://www.openmicroscopy.org/> licensed by University of Dundee & Open Microscopy Environment under [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

Licensing: Creative Commons (CC)



“Share alike”

CC BY-SA: This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license allows for commercial

use. If you remix, adapt, or build upon the material, you must license the modified material under identical terms.

CC BY-SA includes the following elements:

BY  – Credit must be given to the creator

SA  – Adaptations must be shared under the same terms

“Restrictive”
licensing

Content on this slide was adapted from <https://creativecommons.org/about/ccllicenses/> which is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by-sa/4.0/). Icons by The Noun Project.

Licensing: Creative Commons (CC)



CC BY-NC: This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format for noncommercial purposes only, and only so long as attribution is given to the creator.

It includes the following elements:

BY  – Credit must be given to the creator

NC  – Only noncommercial uses of the work are permitted

“Restrictive”
licensing

Content on this slide was adapted from <https://creativecommons.org/about/ccllicenses/> which is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/). Icons by The Noun Project.

Licensing: Creative Commons (CC)



CC BY-ND: This license allows reusers to copy and distribute the material in any medium or format in unadapted form only, and only so long as attribution is given to the creator. The license allows for commercial use.

CC BY-ND includes the following elements:

BY  – Credit must be given to the creator

ND  – No derivatives or adaptations of the work are permitted

“Restrictive”
licensing

Content on this slide was adapted from <https://creativecommons.org/about/ccllicenses/> which is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/). Icons by The Noun Project.

Licensing: Creative Commons (CC)



CC BY-NC-ND: This license allows reusers to copy and distribute the material in any medium or format in unadapted form only, for noncommercial purposes only, and only so long as attribution is given to the creator.

CC BY-NC-ND includes the following elements:

BY  – Credit must be given to the creator

NC  – Only noncommercial uses of the work are permitted

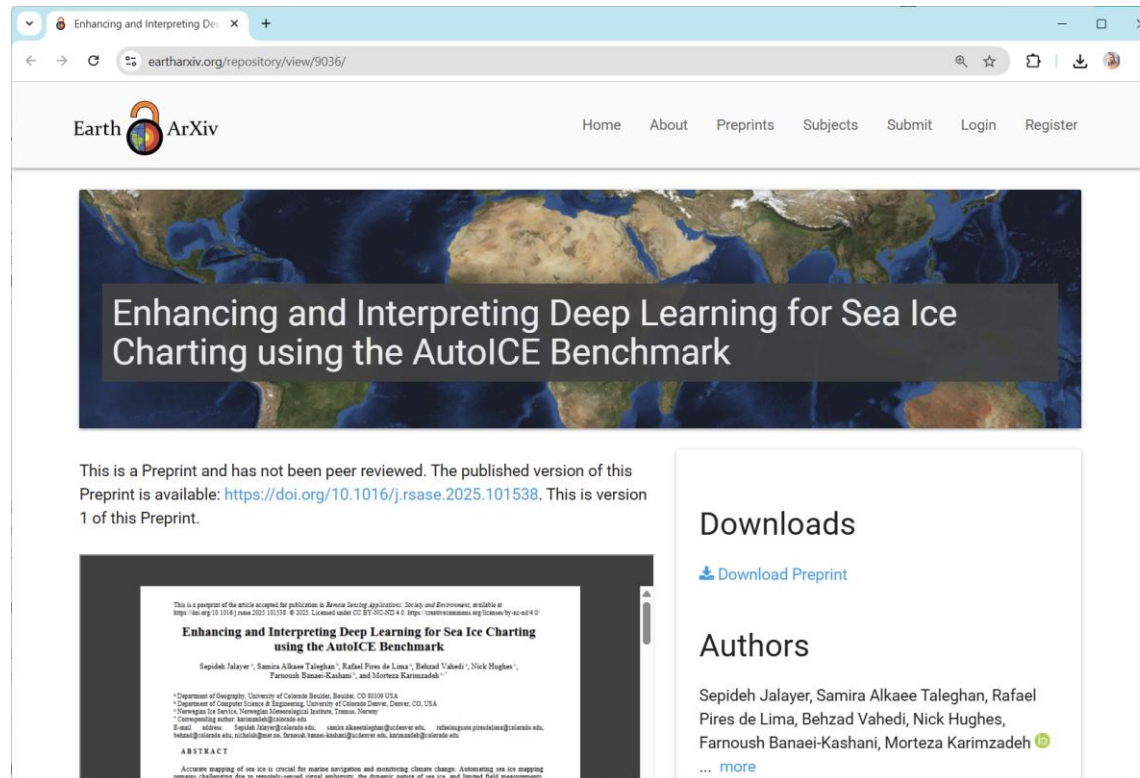
ND  – No derivatives or adaptations of the work are permitted

“Restrictive”
licensing

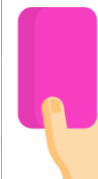
Content on this slide was adapted from <https://creativecommons.org/about/ccllicenses/> which is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Icons by The Noun Project.

Quiz

- May I use one of the Figures from this preprint?



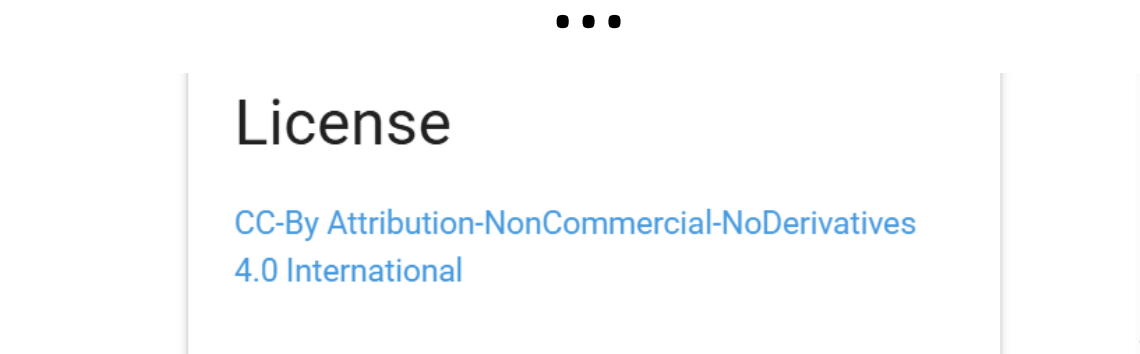
Yes



No



- May I download and redistribute this preprint to students of a course for free?



Yes

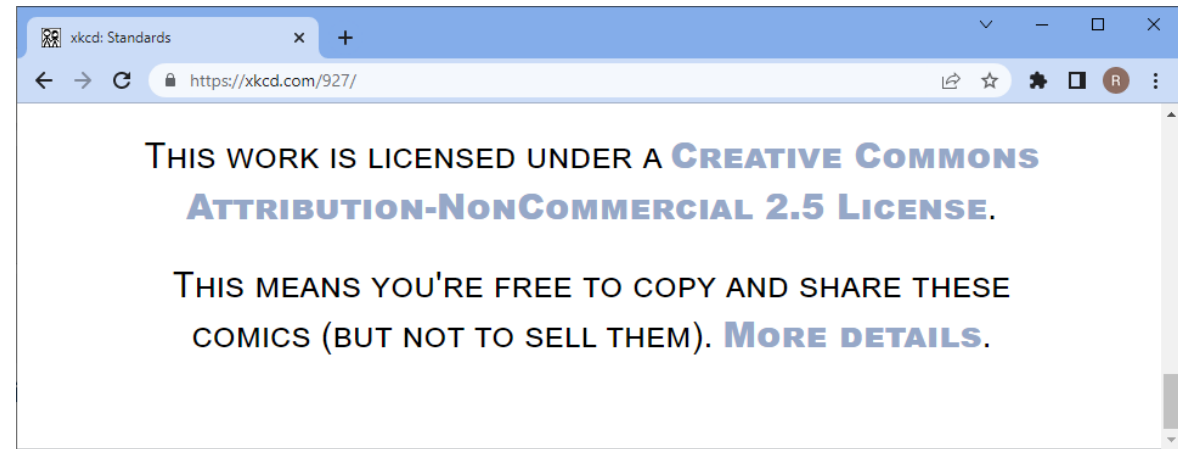
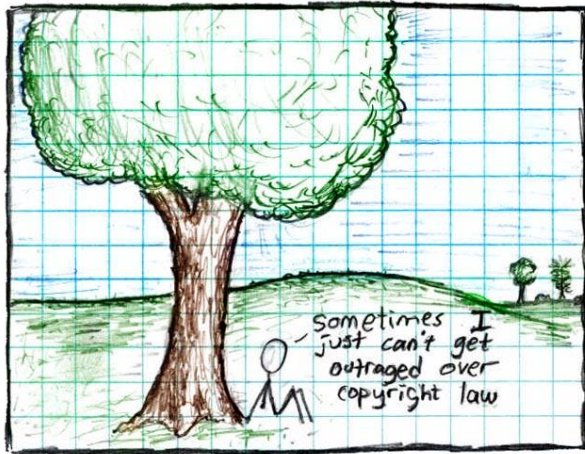


No



Quiz

- It's ok to reuse Figures from XKCD.com if...



Mention
author



Ask the
authors



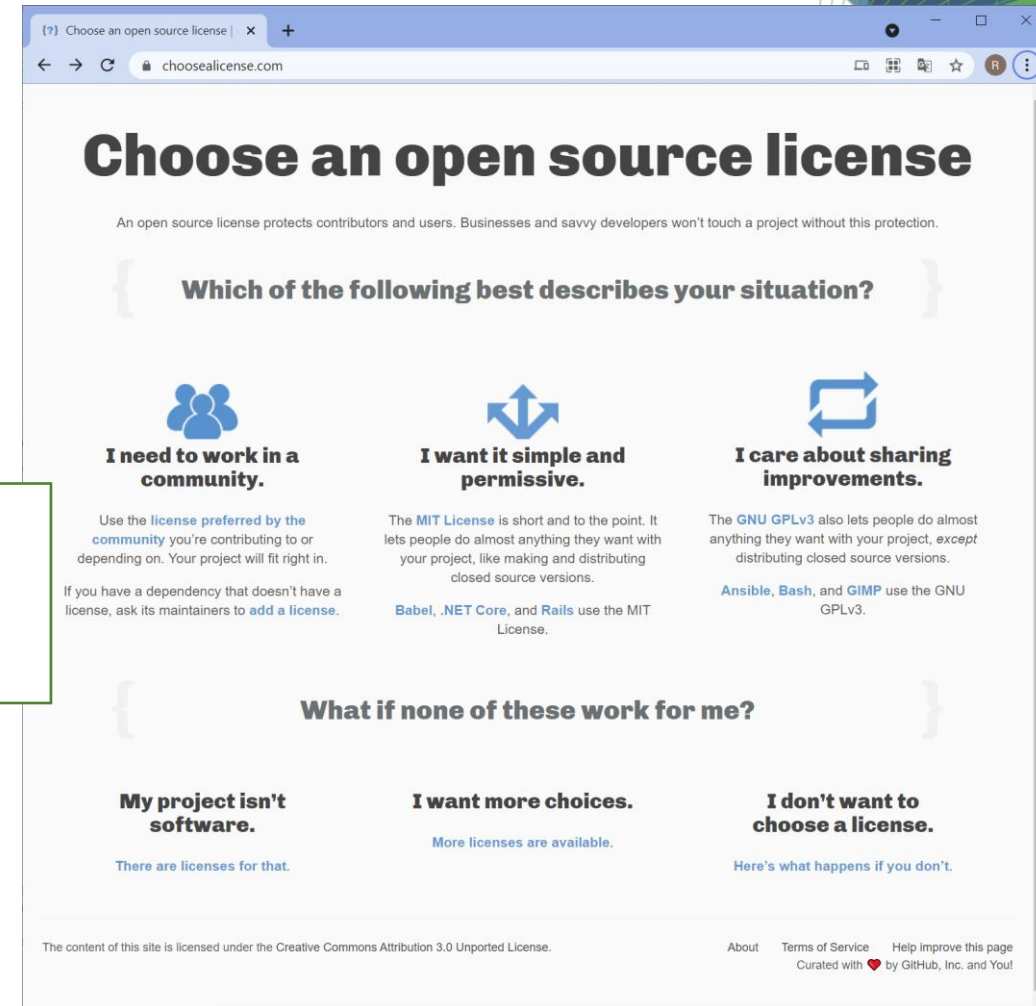
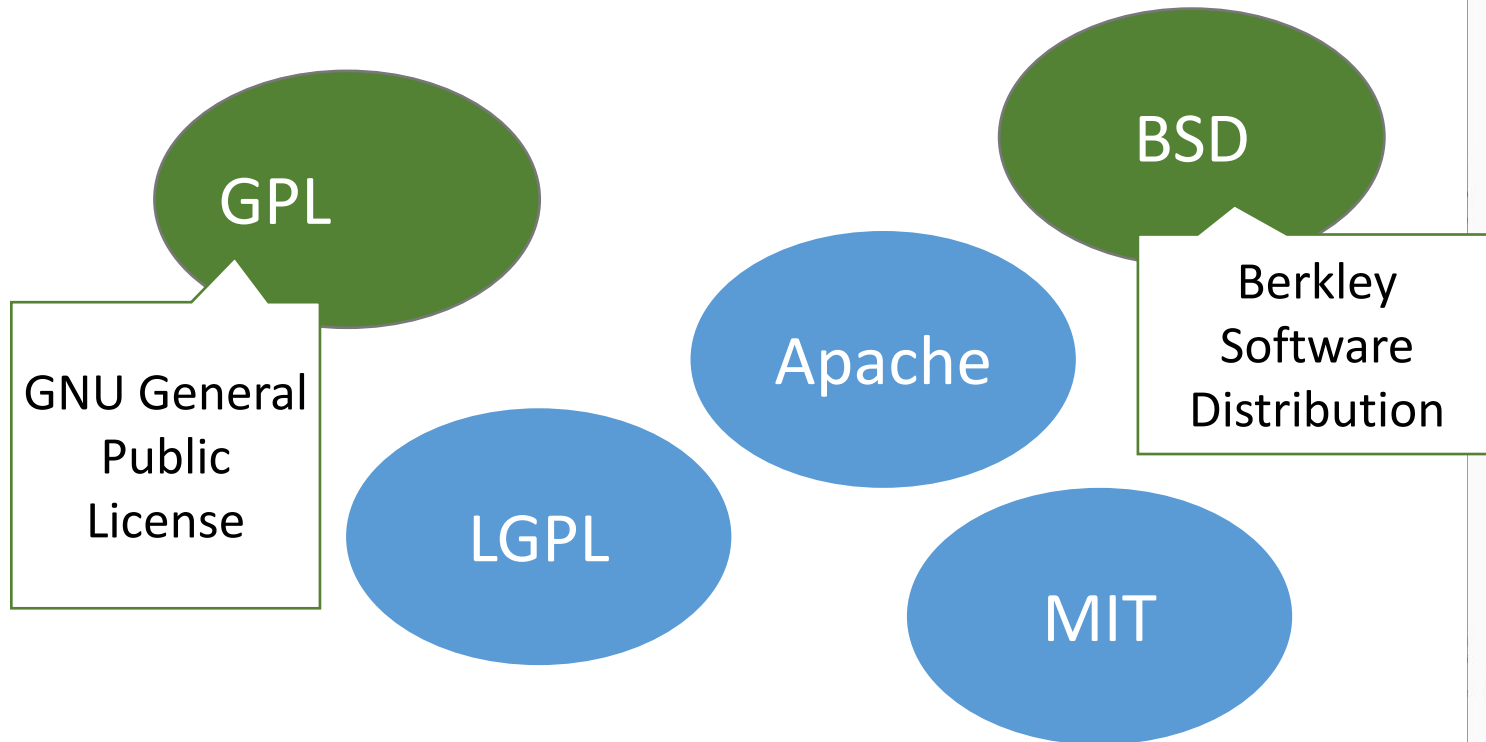
Link to the
license



Copy the
copyright
statement

Licensing Software

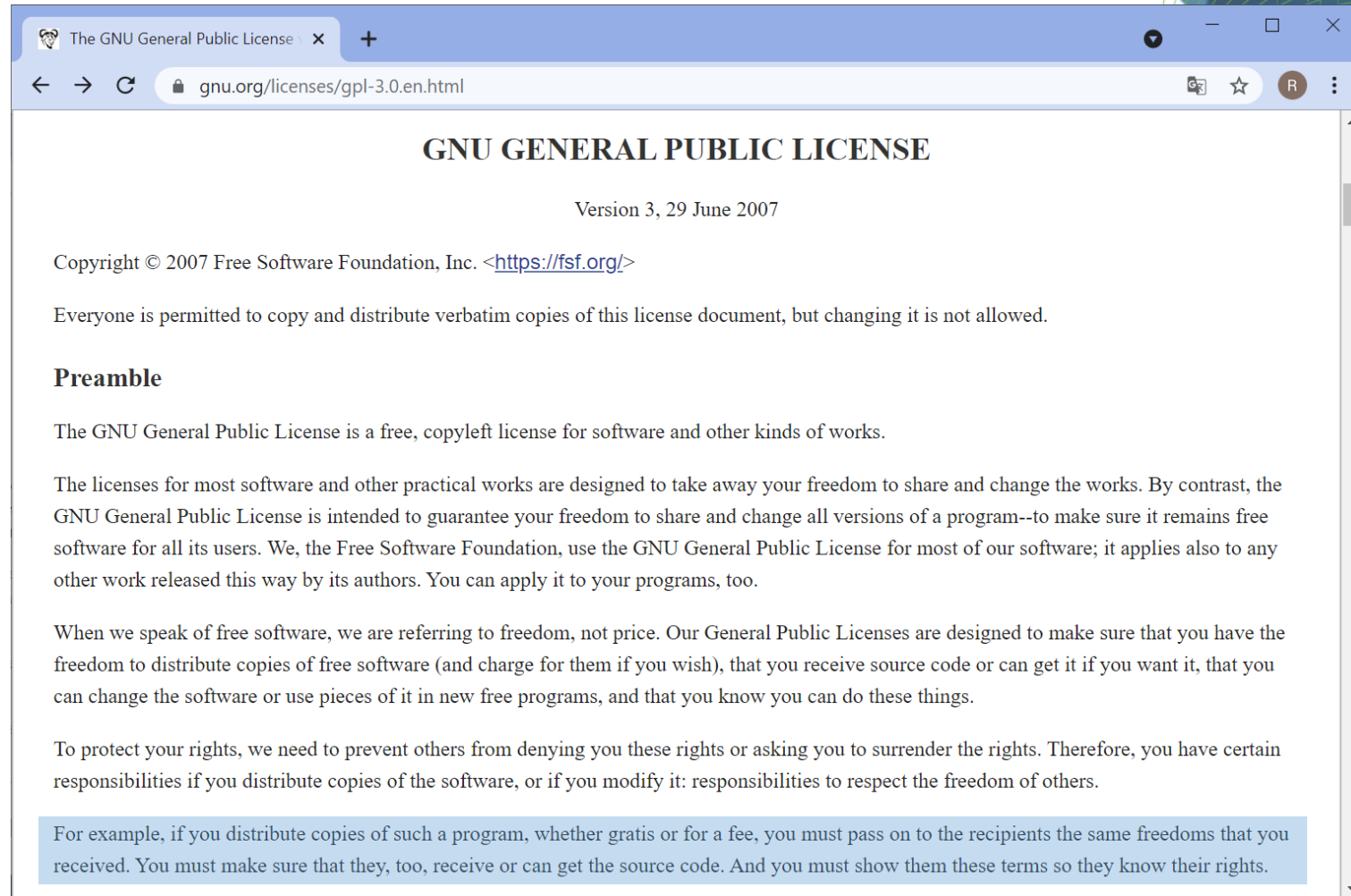
In the software world, other licenses are more popular, historically grown.



Licensing Software: GPL

GPL-derivatives must also be GPL-licensed

“Restrictive”
licensing



Quiz

Can I build a commercial product on the basis of GPL-licensed code?

Yes



No



Do I have to release the code for this commercial product?

Yes



No



Licensing Software: BSD0

Copyright (C) [year] by [copyright holder] <[email]>

Permission to use, copy, modify, and/or distribute this software for any purpose with or without fee is hereby granted.

THE SOFTWARE IS PROVIDED "AS IS" AND THE AUTHOR DISCLAIMS ALL WARRANTIES WITH REGARD TO THIS SOFTWARE INCLUDING ALL IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS. IN NO EVENT SHALL THE AUTHOR BE LIABLE FOR ANY SPECIAL, DIRECT, INDIRECT, OR CONSEQUENTIAL DAMAGES OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THIS SOFTWARE

Disclaimer

Whatever you do with it, we [the authors] are not liable

Licensing Software: BSD2

Copyright (c) <year>, <copyright holder>

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

} Similar to
CC-BY

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE

Licensing Software: BSD3

Copyright <year> <copyright holder>

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
3. Neither the name of the copyright holder nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE

Similar to
CC-BY

You must not use the
copyright holder's name
to endorse your
derivative of the work.

Licensing: Permissive versus restrictive

- Restrictive
 - You can reuse our stuff, but only if you ...
 - License your work with the same license we do
 - Make your stuff openly available
 - Make no money with derivatives of our work
 - Examples: **GPL, CC-BY-SA, CC-BY-NC, CC-BY-ND**
- Permissive licensing:
 - Do whatever you like with our stuff, just make sure to mention / cite us ...
 - Examples: **BSD, MIT, Apache, CC-BY**

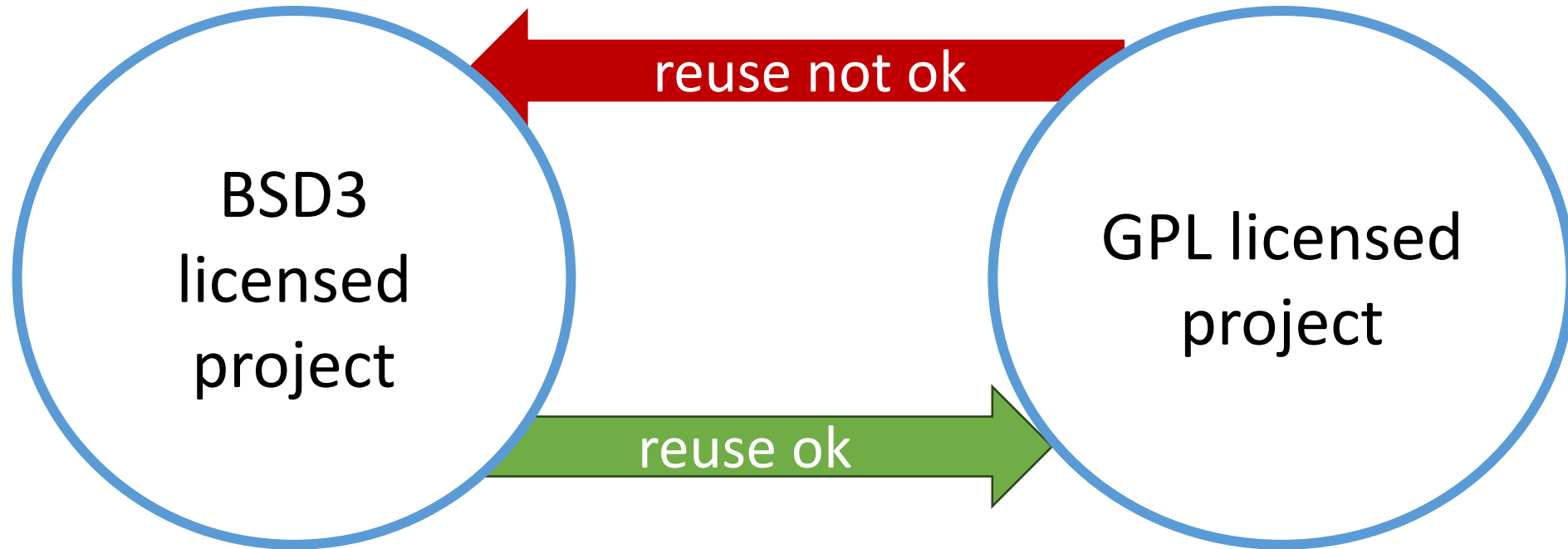
I conclude,
these are
less open in
a sense



Permissive versus restrictive

Use a license which

- is compatible to other projects you're collaborating with and
- fits to your needs / role.



Quiz

May I reuse code
from this repository
in my own BSD-
licensed work?

Yes



No



pyrocko/pyrocko: An official read-only mirror of <https://git.pyrocko.org/pyrocko/pyrocko>.

Search Type / to search

Code Pull requests 5 Actions Security Insights

pyrocko Public Watch 24 Fork 84 Star 225

master Go to file Code

emolch	gnss_unr_campaign example: update url	deee904 · 27 days ago	5,777 Commits
.gitea	fixup	2 years ago	
debian	remove python 3.7 support	10 months ago	
doc	install docs: fix venv command usa...	5 months ago	
docker	docker: add build for ubuntu 24.04	10 months ago	
examples	gnss_unr_campaign example: upda...	27 days ago	
extras	Fix bash_autocomplete for dirnames	9 years ago	
maintenance	conda: update packaging	3 months ago	

About

An official read-only mirror of <https://git.pyrocko.org/pyrocko/pyrocko>.
A seismology toolkit for Python.

pyrocko.org

python seismology geophysics
earthquake geophysical-inversions

Readme
GPL-3.0 license
Cite this repository
Activity
Custom properties

Quiz

May I use this library in a software product of a company?

Yes



No



pyrocko/pyrocko: An official read-only mirror of <https://git.pyrocko.org/pyrocko/pyrocko>

pyrocko / pyrocko

Type / to search

Code Pull requests 5 Actions Security Insights

pyrocko Public

Watch 24 Fork 84 Star 225

master Go to file Code

emolch gnss_unr_campaign example: update url deee904 · 27 days ago 5,777 Commits

.gitea	fixup	2 years ago
debian	remove python 3.7 support	10 months ago
doc	install docs: fix venv command usa...	5 months ago
docker	docker: add build for ubuntu 24.04	10 months ago
examples	gnss_unr_campaign example: upda...	27 days ago
extras	Fix bash_autocomplete for dirnames	9 years ago
maintenance	conda: update packaging	3 months ago

About

An official read-only mirror of <https://git.pyrocko.org/pyrocko/pyrocko>.
A seismology toolkit for Python.

pyrocko.org

python seismology geophysics
earthquake geophysical-inversions

Readme
GPL-3.0 license
Cite this repository
Activity
Custom properties

Quiz

May I reuse code
from this repository
in my own GPL-
licensed work?

Yes



No



The screenshot shows the GitHub repository page for 'fatiando/verde'. The repository is public and has 22 watchers, 74 forks, and 626 stars. It contains 38 issues, 6 pull requests, and 333 commits. The repository description is 'Processing and gridding spatial data, machine-learning style'. The license is BSD-3-Clause. The repository includes a README, a Code of conduct, and Custom properties. A large orange arrow points to the 'BSD-3-Clause license' link in the 'About' section.

File	Commit Message	Time Ago
.github	Bump pypa/gh-action-pypi-publis...	2 months ago
data	Move the sources of gallery/tutoria...	4 years ago
doc	Minor fixes to docs (#473)	5 months ago
env	Replace Sphinx napoleon for nump...	11 months ago
paper	Add acknowledgements and SOES...	7 years ago
verde	Minor fixes to docs (#473)	5 months ago
.codecov.yml	Start the repository with basic tem...	7 years ago
.coveragerc	Update GitHub Actions configurati...	4 years ago



Take home message

If you share material (openly or not)

license it

and it'll be harder to steal it

Summary

- If you want to make your stuff reusable:
 - Use permissive licenses
 - Share it on community-wide platforms (preferably not institutional servers)
 - Register your data and software in search-indices
- Read more:
 - Sharing on Zenodo
<https://focalplane.biologists.com/2023/02/15/sharing-research-data-with-zenodo/>
 - Sharing on Figshare
<https://focalplane.biologists.com/2023/07/26/sharing-your-poster-on-figshare/>
 - Collaborative work on github
<https://focalplane.biologists.com/2021/09/04/collaborative-bio-image-analysis-script-editing-with-git/>
 - Licensing
<https://focalplane.biologists.com/2023/05/06/if-you-license-it-itll-be-harder-to-steal-it-why-we-should-license-our-work/>

Acknowledgements

Communities & platforms

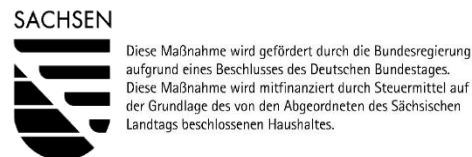


BIDS team UL



Lea Gihlein, Lea Kabjesz, Mara Lampert

Funding



Chan
Zuckerberg
Initiative



Some Figures were generated using

