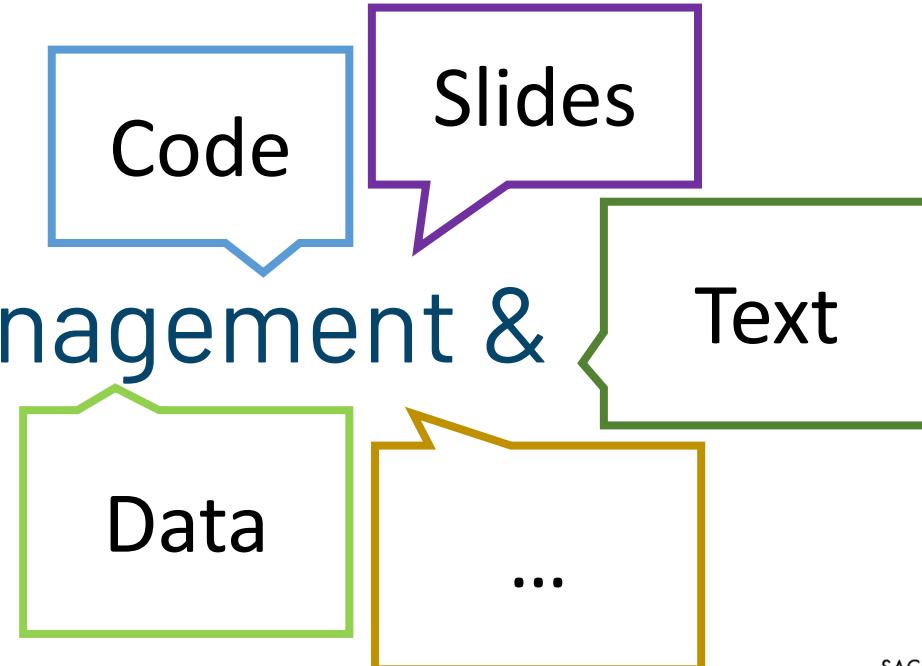




Research Data Management & Open Science

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.

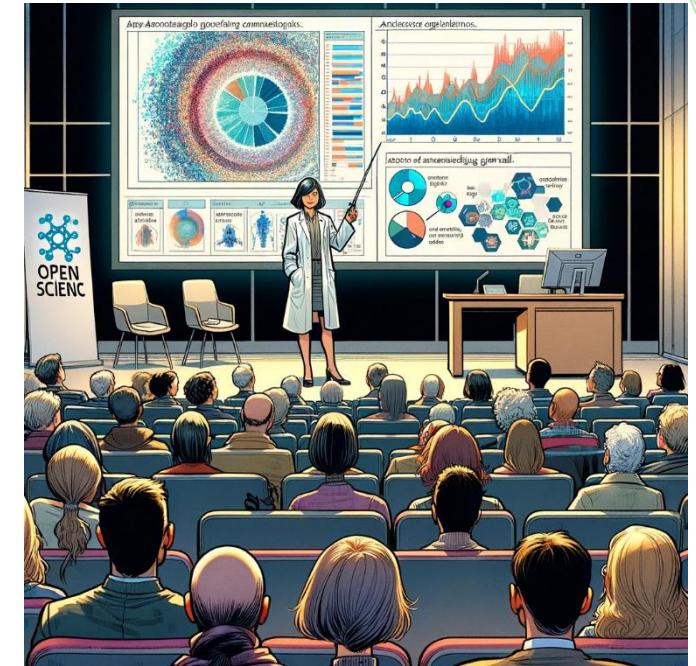


Open Science

- 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



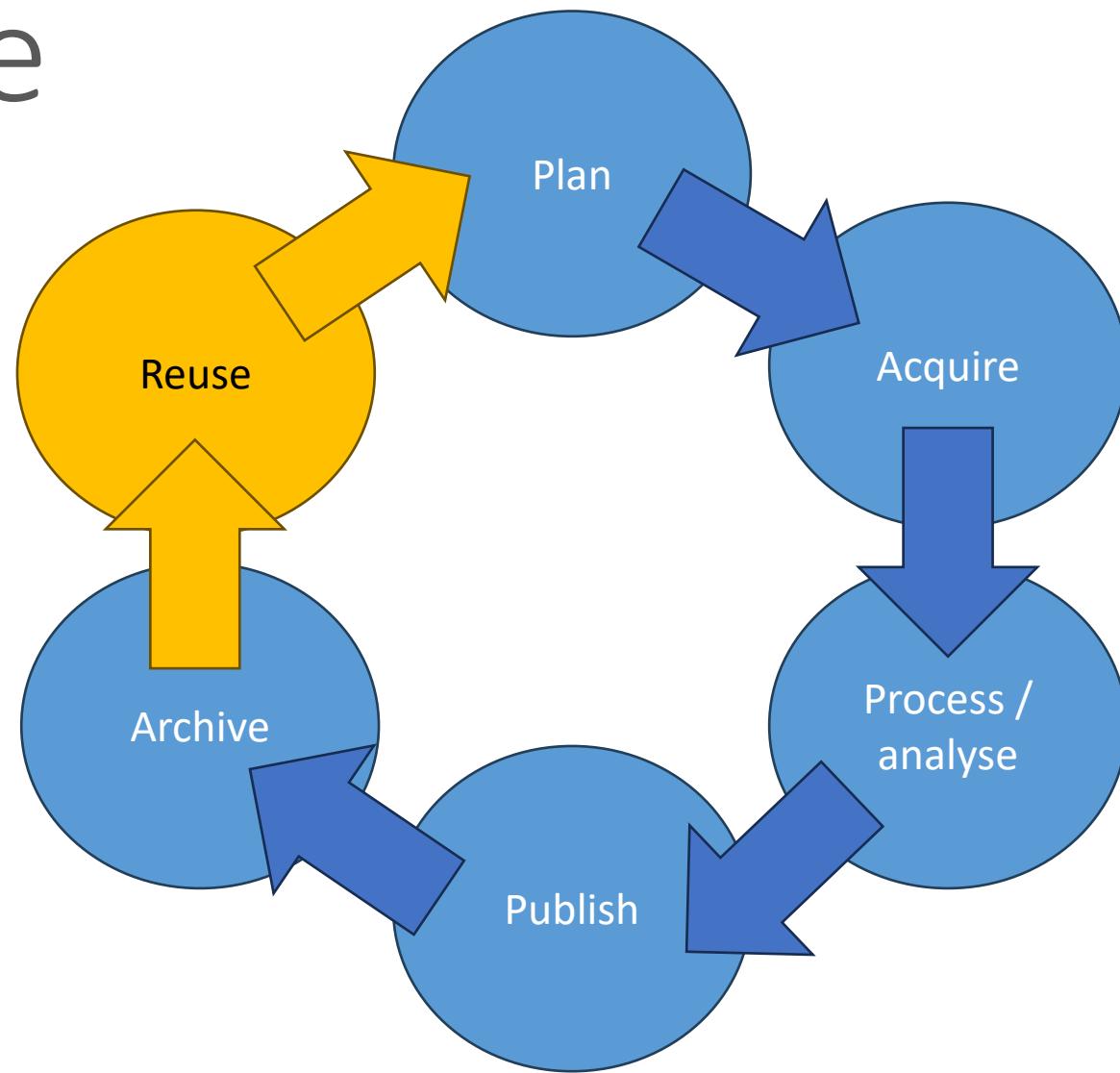
Research Data Management (RDM)

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



RDM Life Cycle

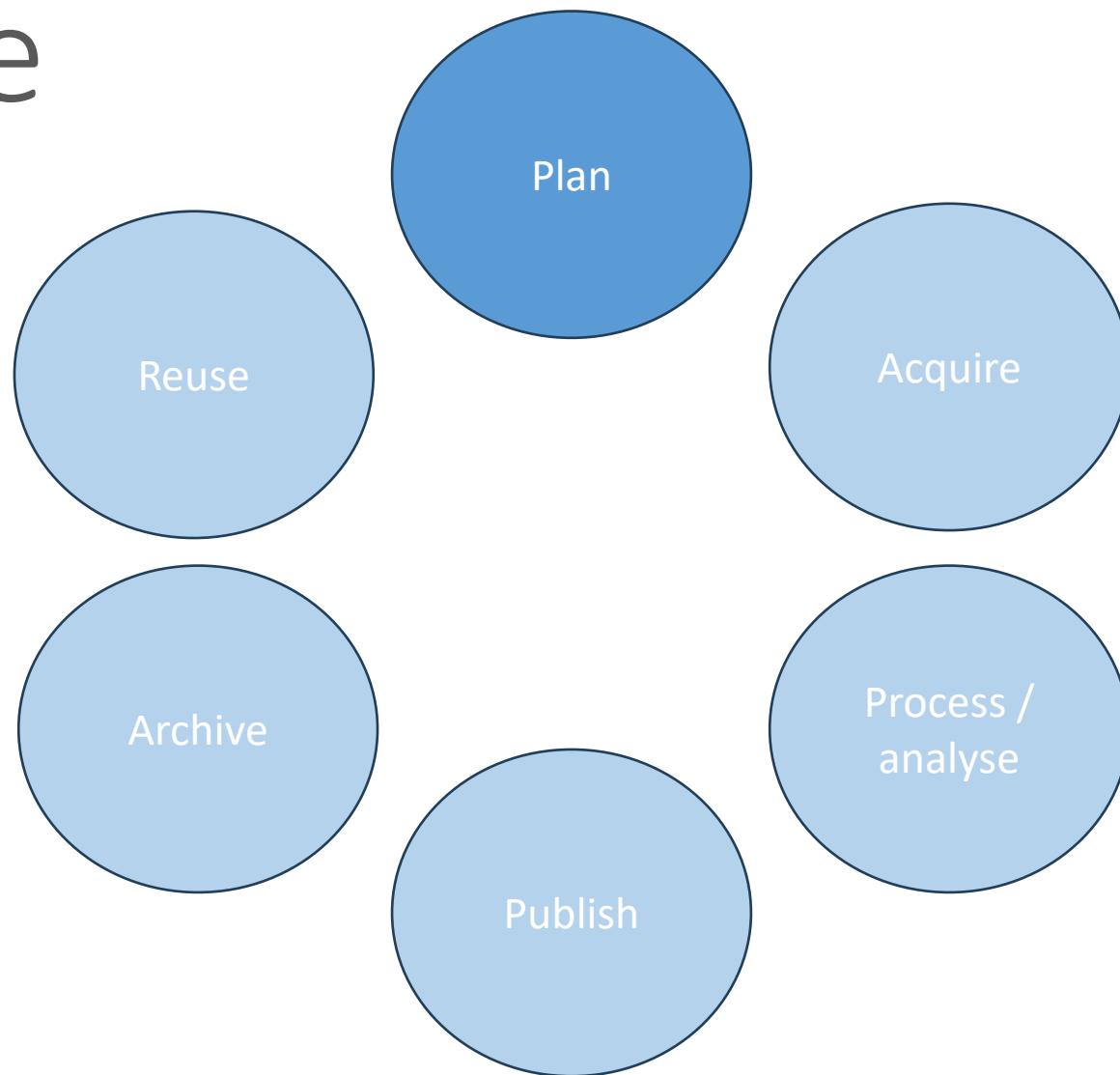
- Processes are ideally cyclic
- Closing the loop is a major challenge



RDM Life Cycle

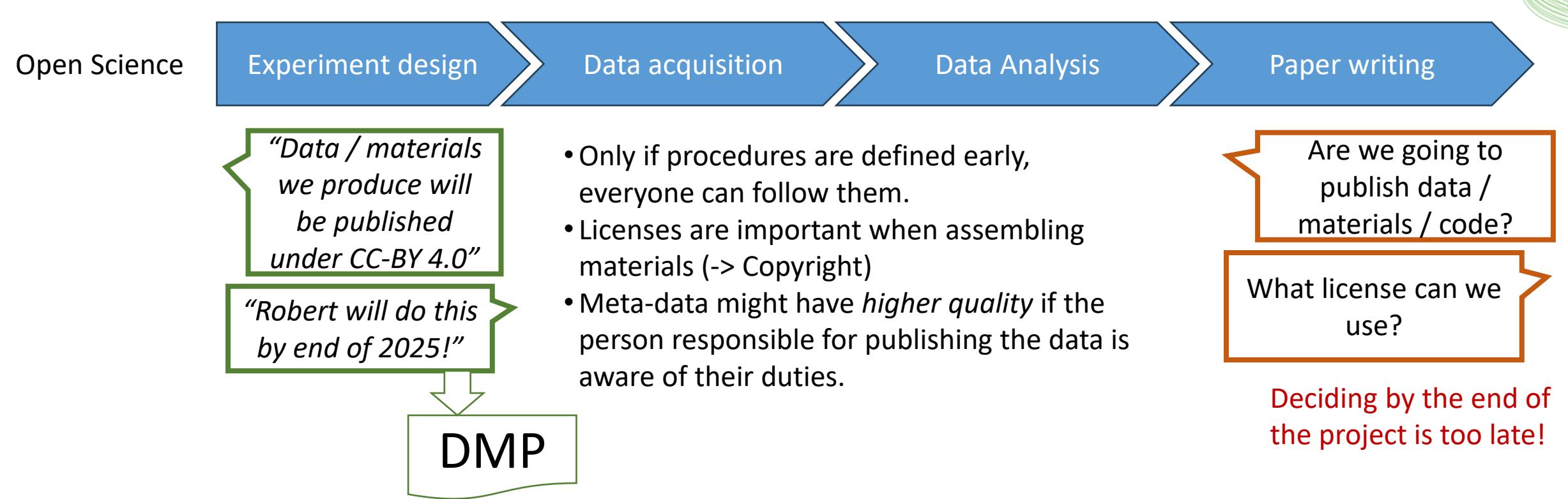
Plan

- Cost
- Benefit
- Quality
- Strategic decisions



Data Management Plans (DMPs)

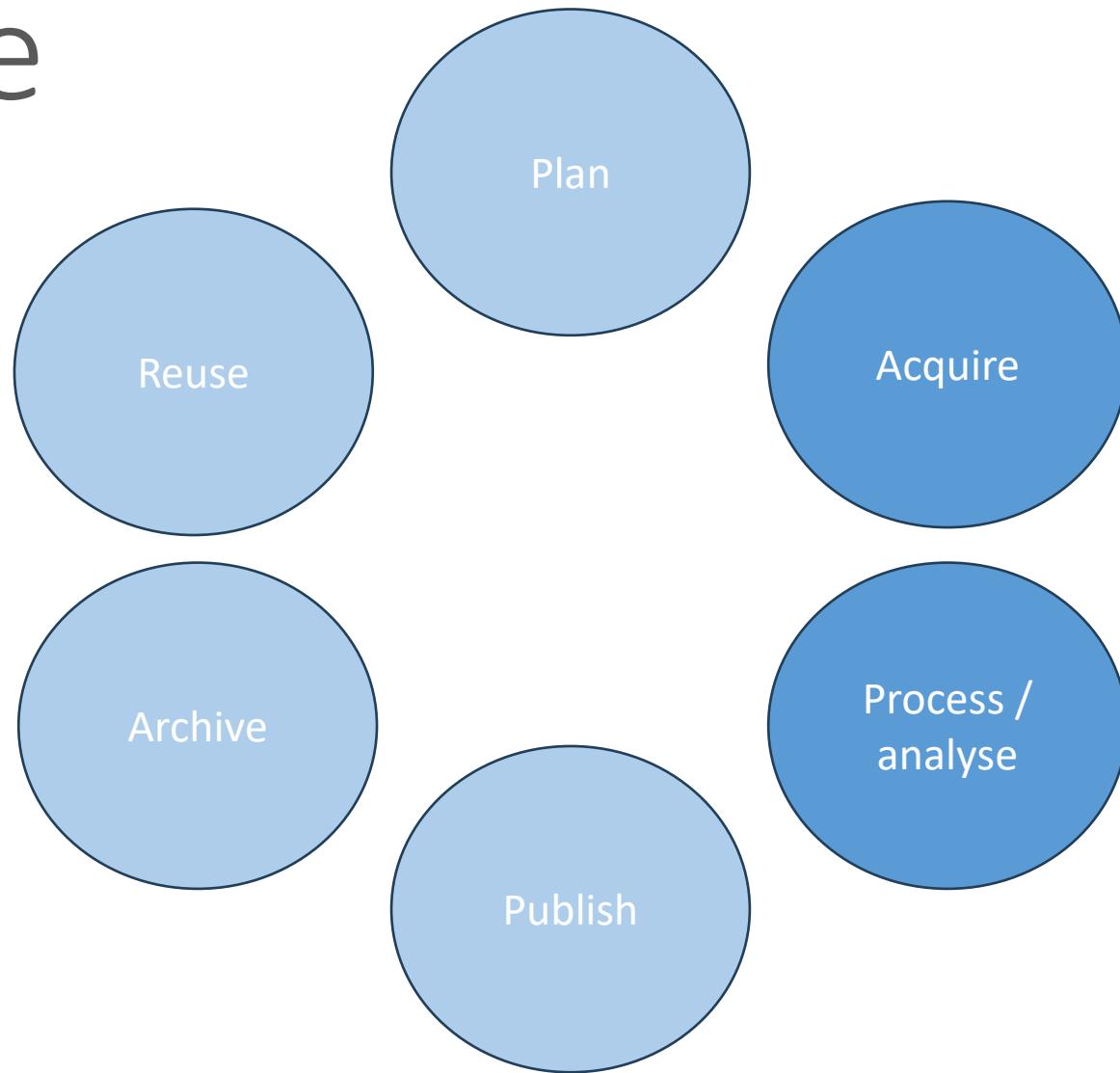
- Define responsibilities and procedures early!



RDM Life Cycle

Acquisition + Processing

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



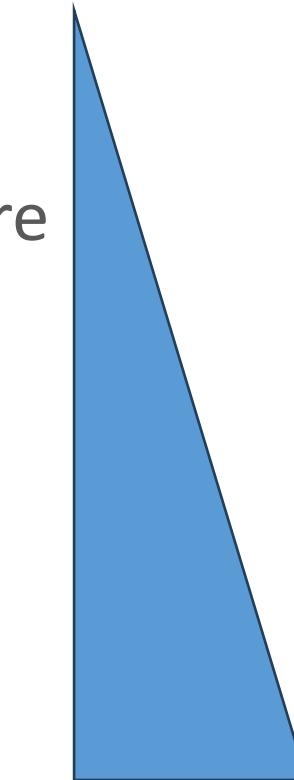
Types of data

- 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



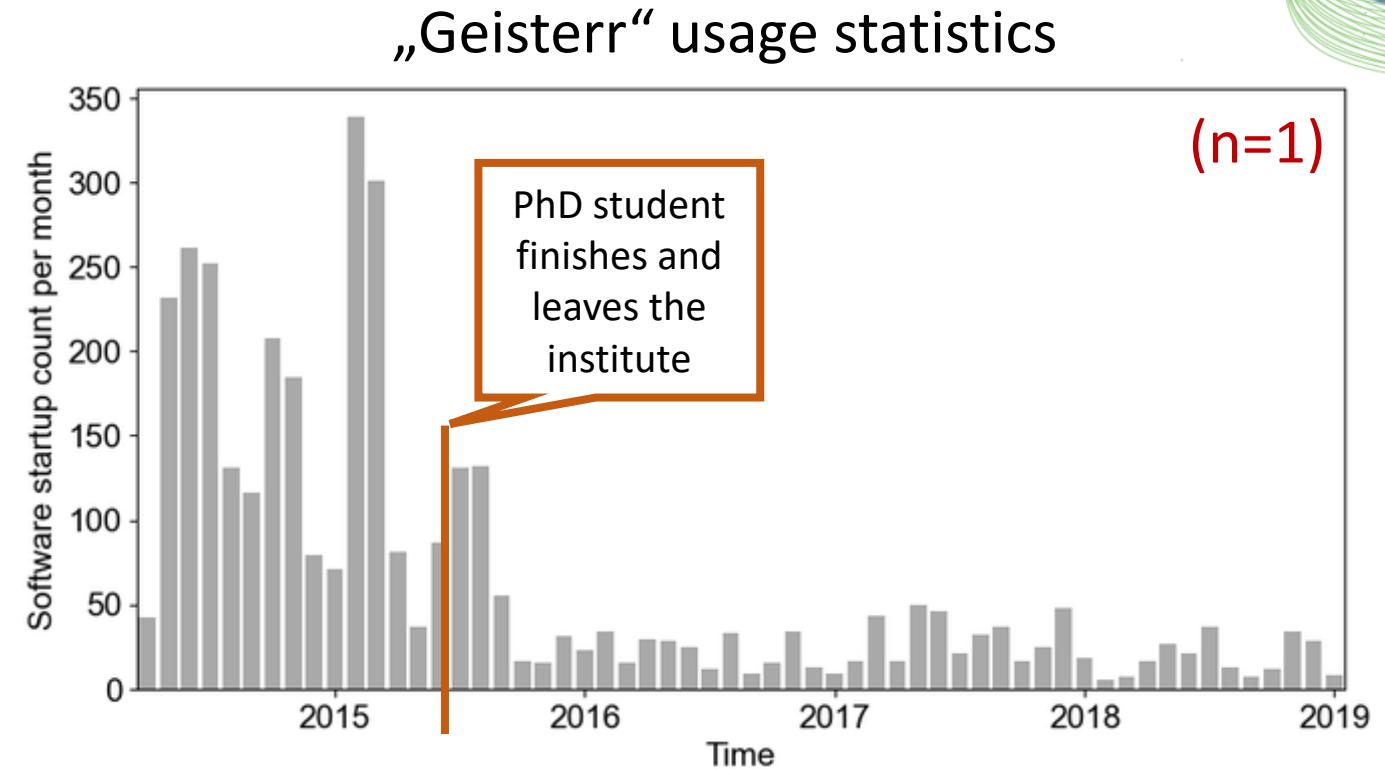
In need of
protection
(schutzbedürftig)



Version control [for code]

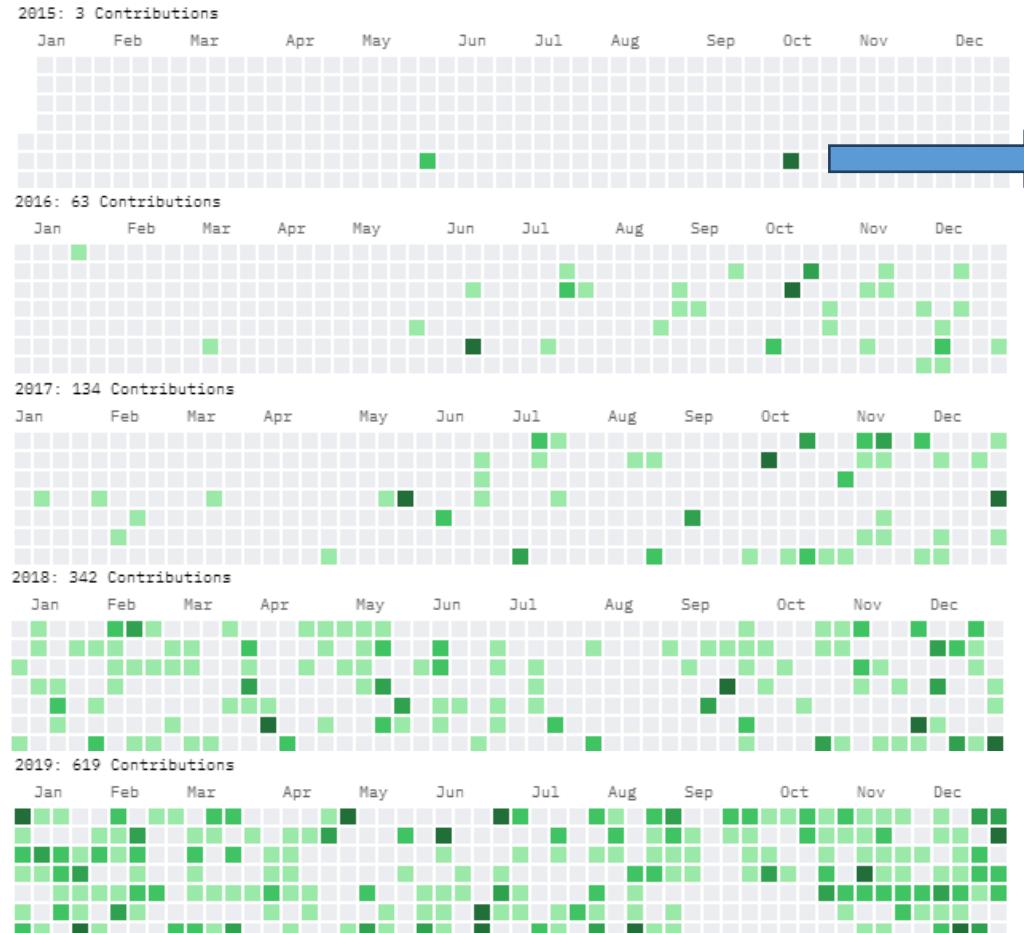
- Me, during my Phd:

geisterr - 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



Version control [for code]: git

Github contributions



Fiji (Fiji Is Just) Image

Pencil Tool

CalibZAPWfixed_000154_max.tif

Select an action

- Capture overlay
- Capture overlay
- Close gaps by introducing new spots
- Compute distance to ROIs
- Export tracks to XML file
- Extract track stack
- Export spots to IJ ROIs
- Export to ISBI challenge file format

Methods
Volume 115, 15 February 2017, Pages 80-90

TrackMate: An open and extensible platform for single-particle tracking

Jean-Yves Tinevez^a, Nick Perry^a, Johannes Schindelin^b, Genevieve M. Hoopes^c, Gregory D. Reynolds^c, Emmanuel Laplantine^d, Sebastian Y. Bednarek^c, Spencer L. Shute^a, Kevin W. Eliceiri^b

Show more ▾

+ Add to Mendeley Share Cite

<https://doi.org/10.1016/j.jymeth.2016.09.016>

Get rights and content ▾

open access

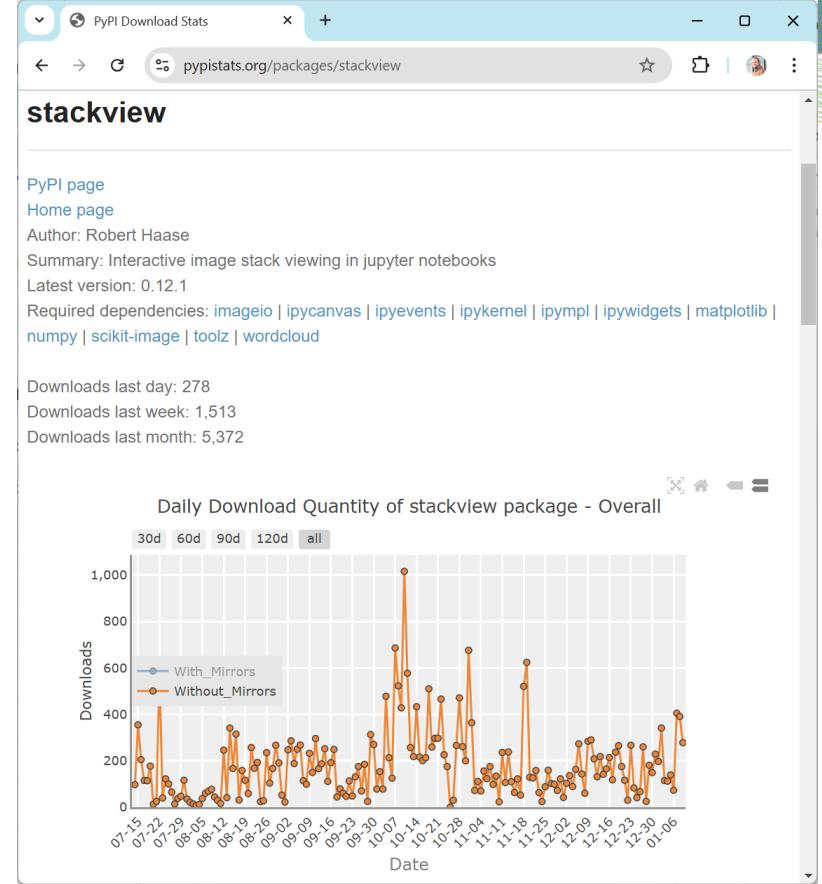
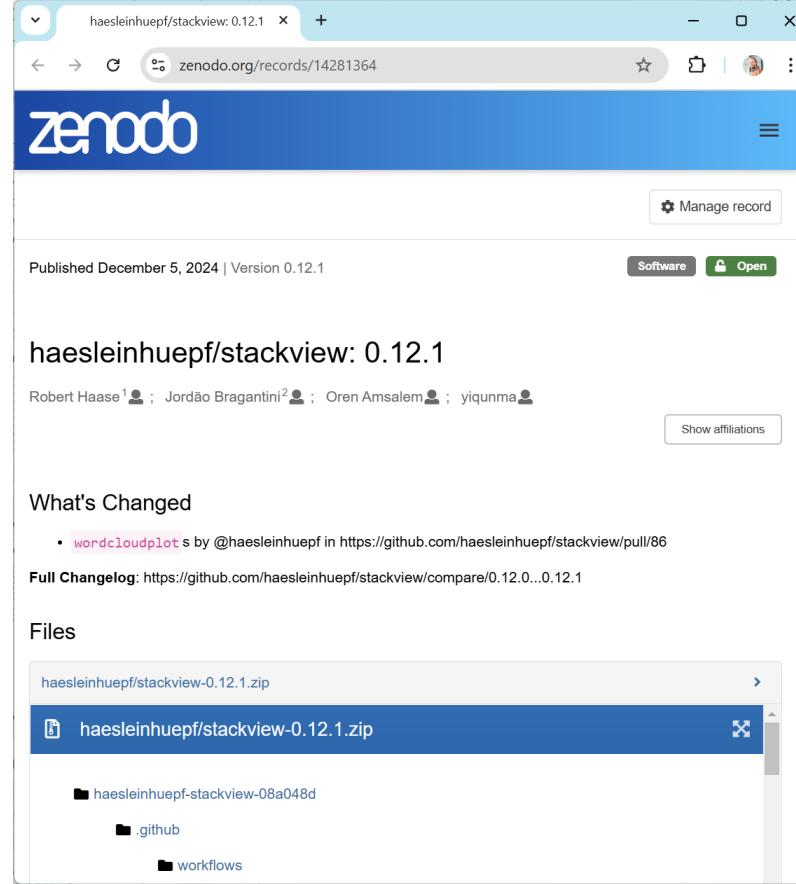
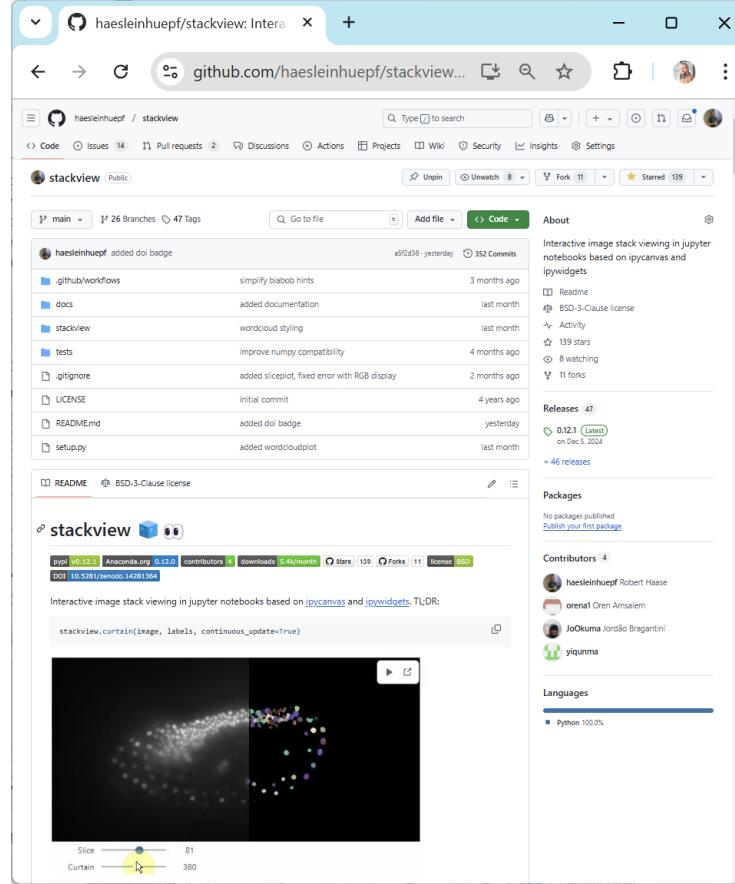
Table 1. User contributed modules of TrackMate v3.4.0.

Module name	Module type	Purpose	Author	Location
Linear tracker	Particle-linking	Linking transported particles by extrapolating their velocity	Ronny Sczech	https://github.com/chicorrony/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

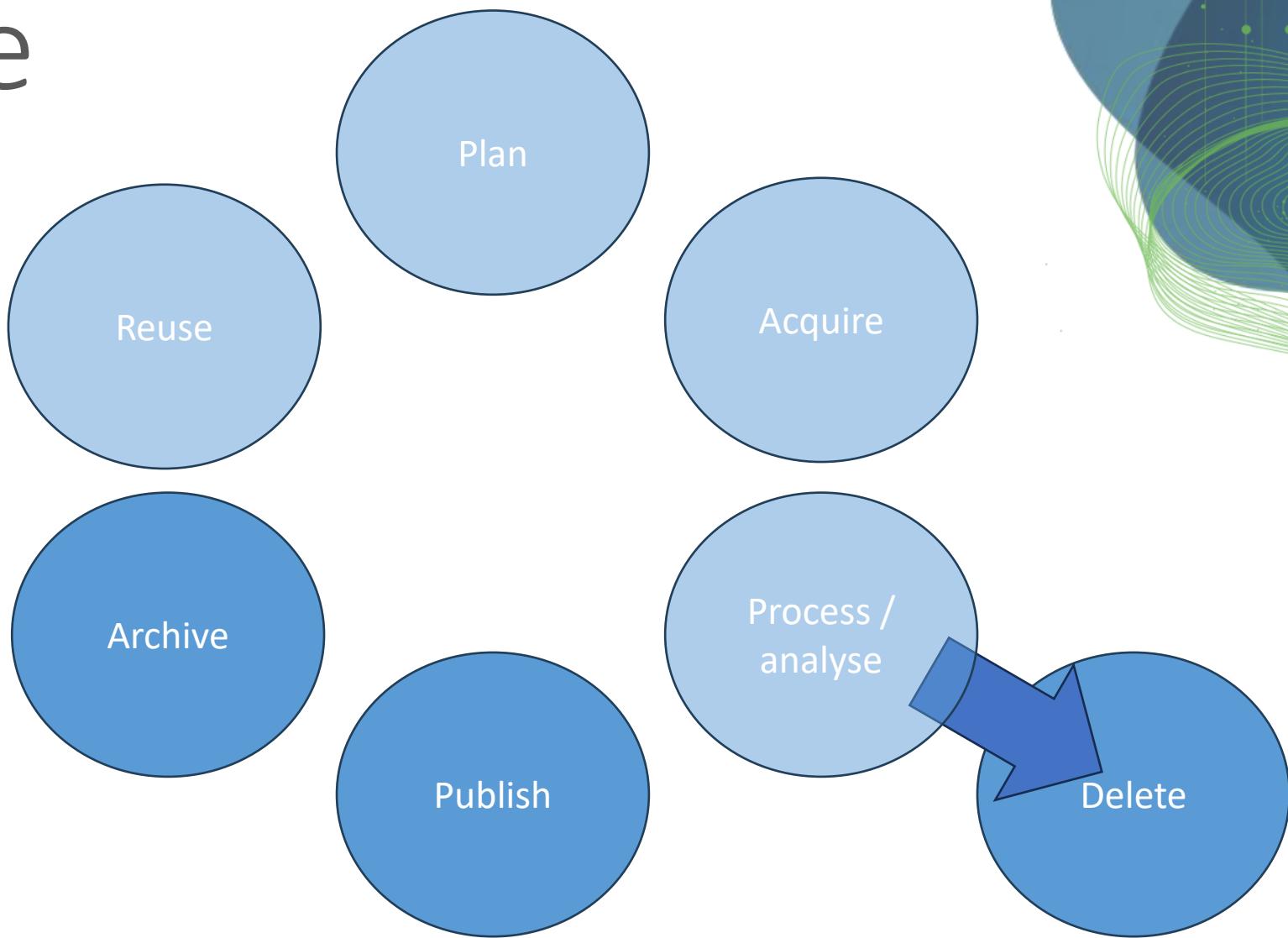
Version control [for code]: git

- Me, today:



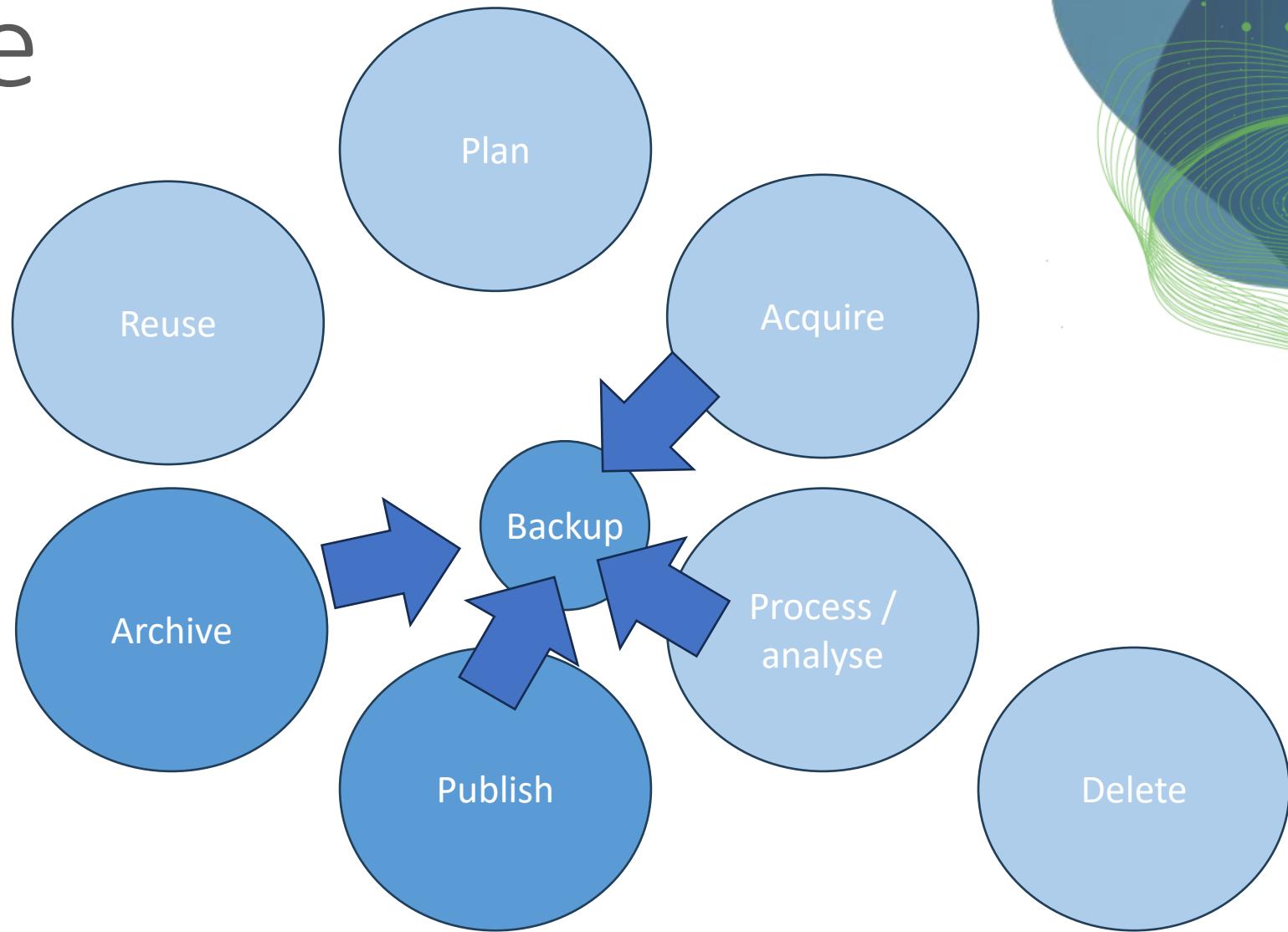
RDM Life Cycle

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



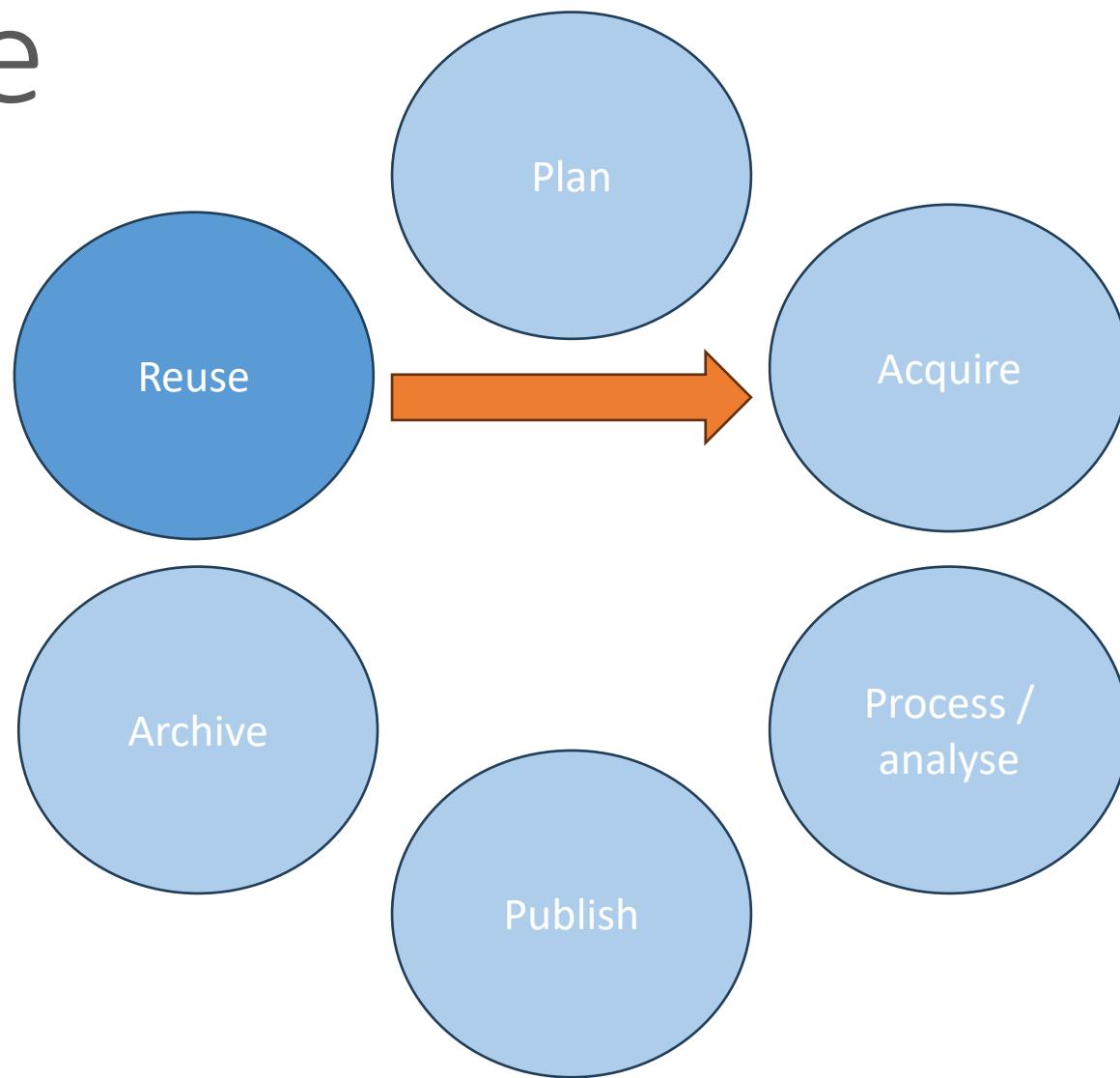
RDM Life Cycle

- 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”

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



About Us ▾ Funding ▾ Basics and Topics ▾ Funded Projects ▾ News ▾

DFG > News > News and Current Topics > Information for Researchers > Package of Measures to Support a Shift in the Culture of Research Assessment

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.

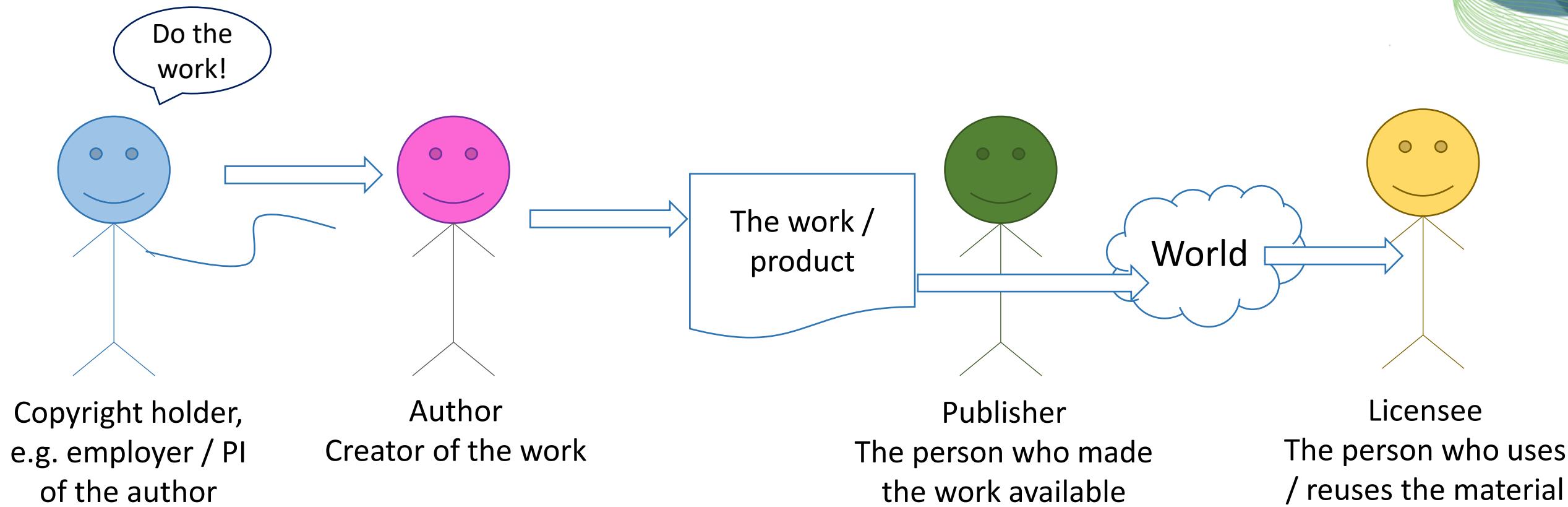
Publication details in proposals and CVs

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.

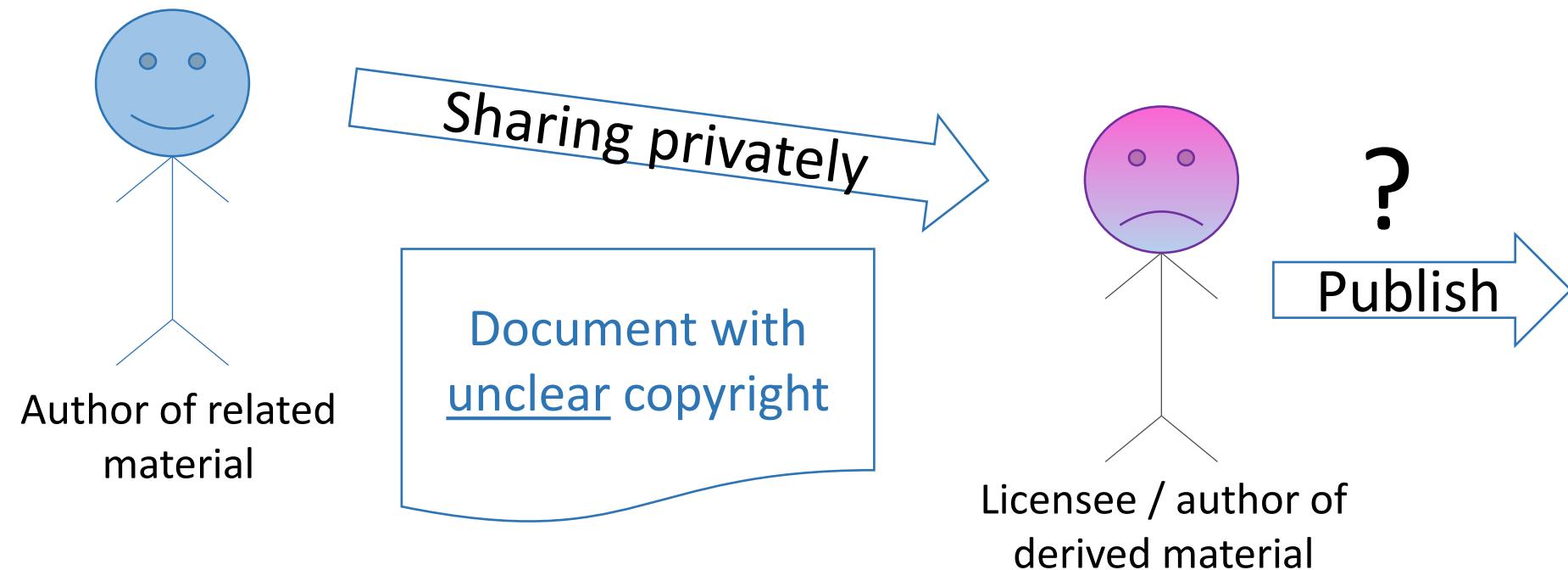
Am I allowed to publish my stuff?

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



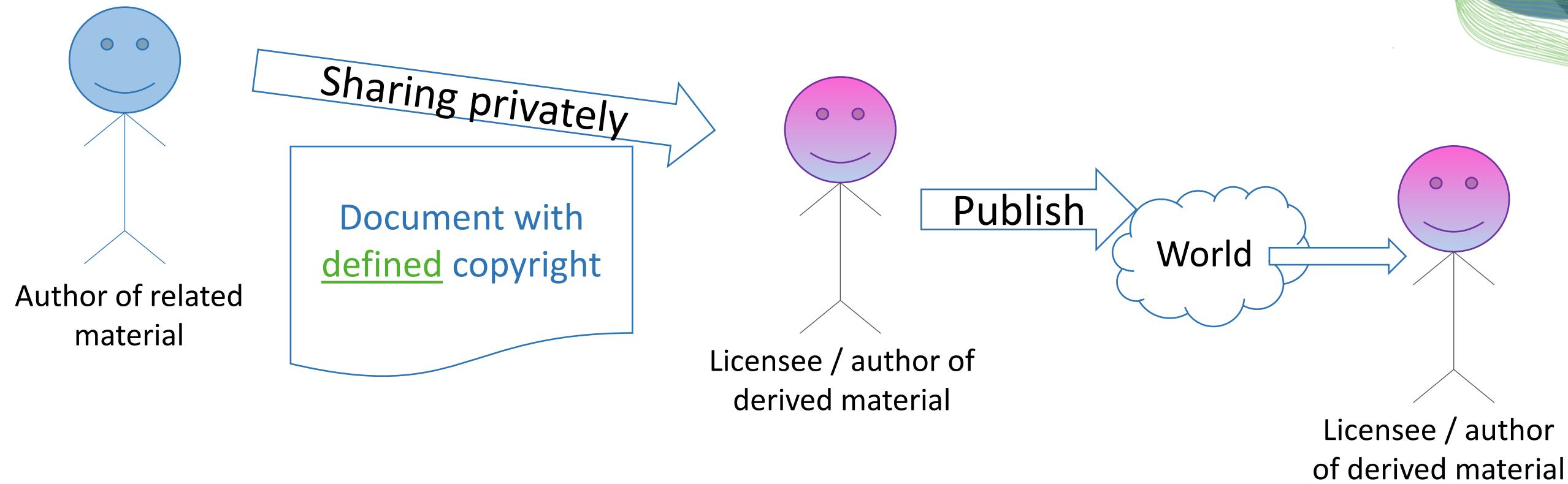
Am I allowed to publish my stuff?

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



Am I allowed to publish my stuff?

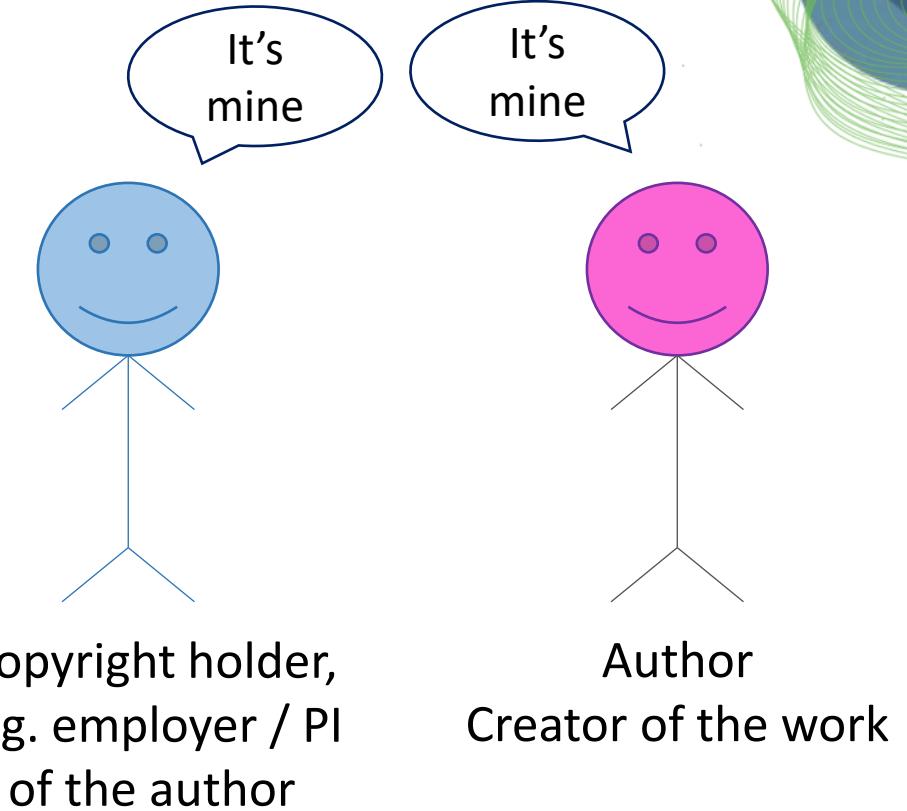
- ... it depends... on what materials served as basis



Public versus institutional repositories

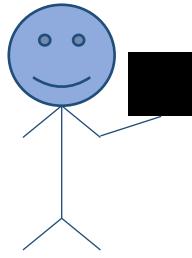
Conflict of interest between employers and scientists

- 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

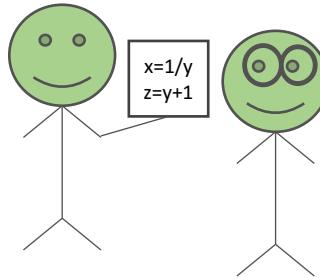
Closed source



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

Hardware device drivers

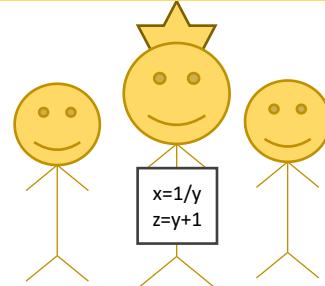
Open source



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

Custom data analysis scripts

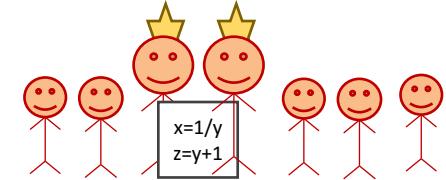
Benevolent dictatorship



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

scikit-image, scipy, OpenCL

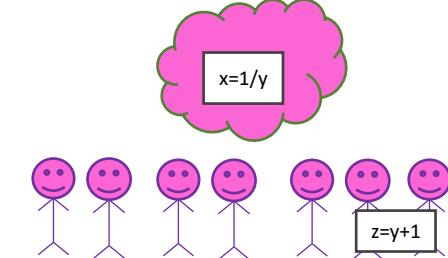
Community driven



- 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

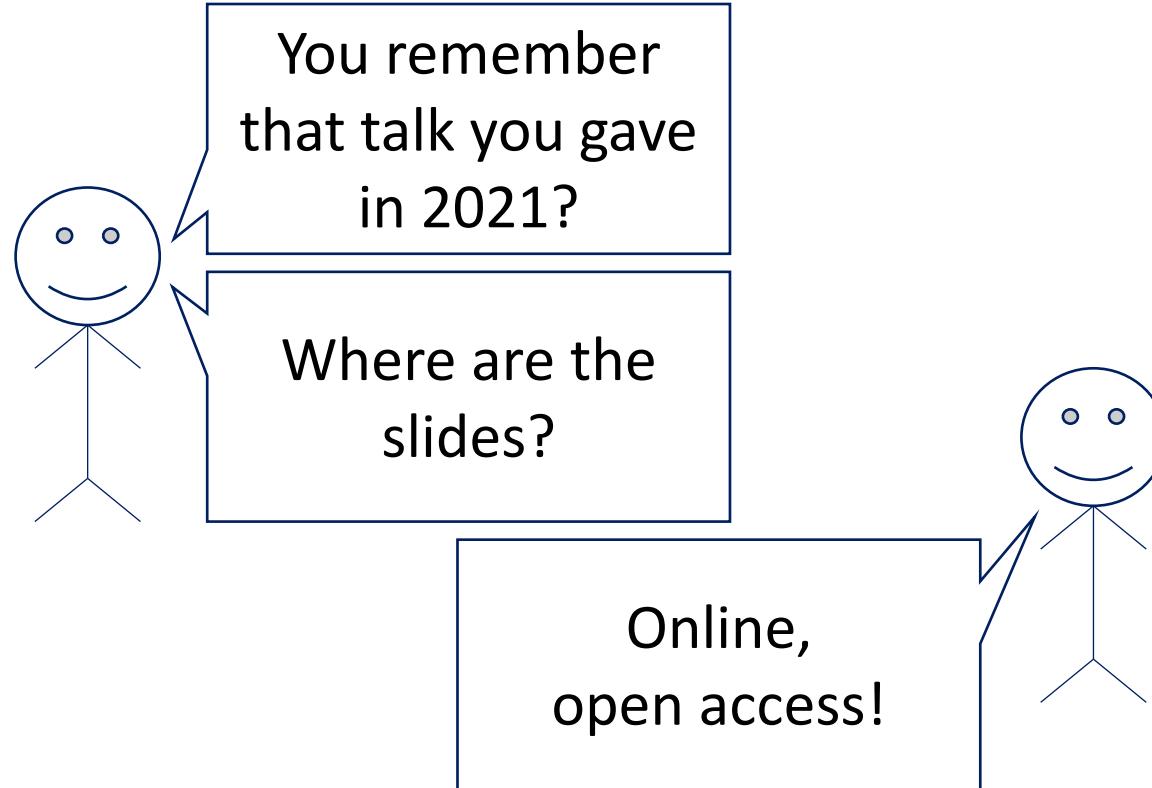
- Publish where your community publishes
- Publish where everyone publishes (beyond your community)
- Publish in your local institute's infrastructure



The top screenshot shows the re3data.org search results page for 'seismology'. The search bar at the top has 're3data.org/search?query=seismology'. Below it, the page title is 're3data.org' and the subtitle is 'World Data Center for Solid Earth Physics WDC for SEP'. There are three main filter categories: 'Subject(s)' (Natural Sciences, Geosciences, Geophysics and Geodesy), 'Repository type(s)' (disciplinary), and 'Provider type(s)' (dataProvider).
The bottom screenshot shows the FAIRsharing.org databases search results page for 'seismology'. The search bar at the top has 'fairsharing.org/search?fairsharingRegistry=Database&q=seismology'. Below it, the page title is 'FAIRsharing.org standards, databases, policies' and the subtitle is 'Databases A registry of knowledgebases and repositories of data and other digital assets'. A message at the top states: '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.' The search results show a single entry for 'IRIS Data' under the 'Incorporated Research Institutions for Seismology Data' category. The IRIS Data entry includes a logo of a database icon with a green 'R', a brief description: '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...', and three buttons: 'Environment...', 'Earth Scien...', and 'Not applic...'.

Incentives: Findability

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



Sharing and licensing material | f1000research.com/slides/10-519

f1000Research

BROWSE GATEWAYS & COLLECTIONS HOW TO PUBLISH ABOUT BLOG MY RESEARCH SIGN IN

Metrics | 411 Views | 60 Downloads

DOWNLOAD 30.92 MB

SHARE CITE

PART OF THE GATEWAY

neubias - the Bioimage Analysts Network

BROWSE BY RELATED SUBJECTS

Artificial intelligence

Computer and information sciences

Electrical engineering

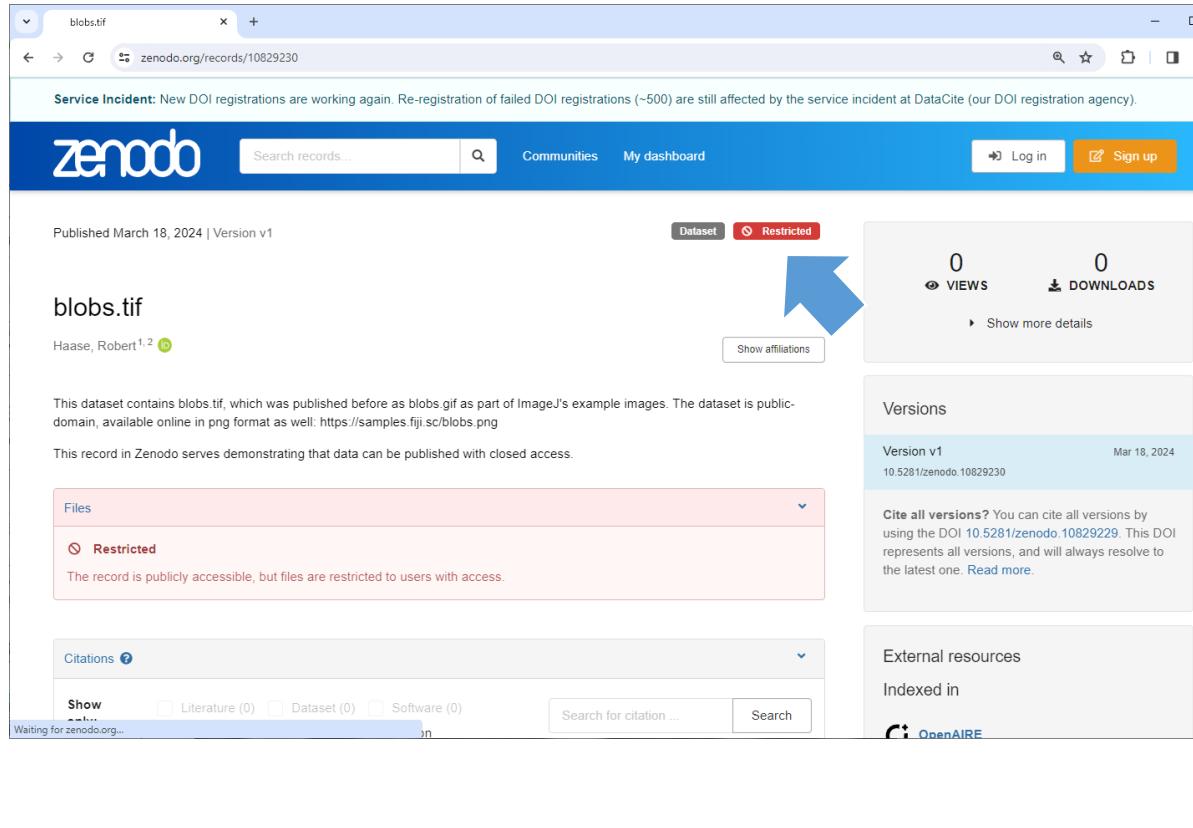
Sharing and licensing material
Robert Haase
June 30th 2021

Code Slides Text Data ...

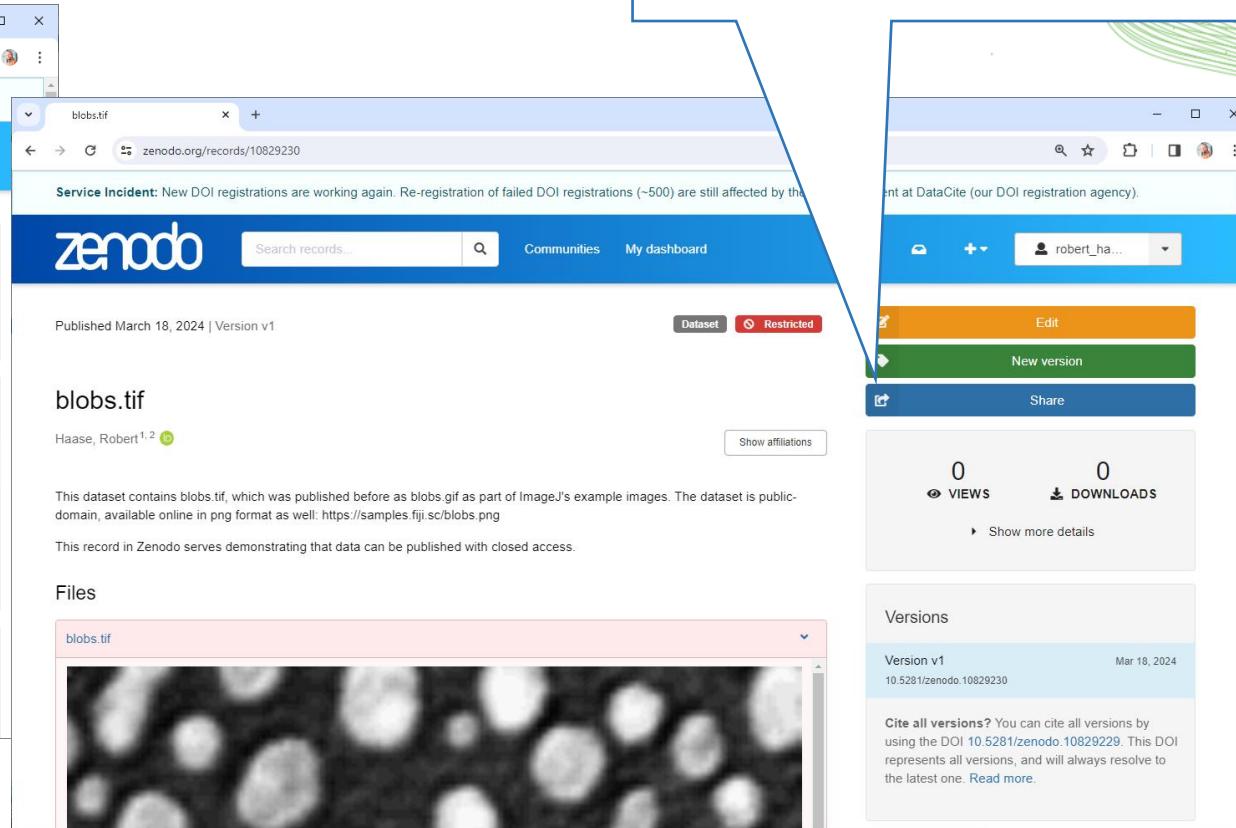
This material is licensed by Robert Haase, PoL Dresden under the CC-BY 4.0 license <https://creativecommons.org/licenses/by/4.0/>

Accessibility

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



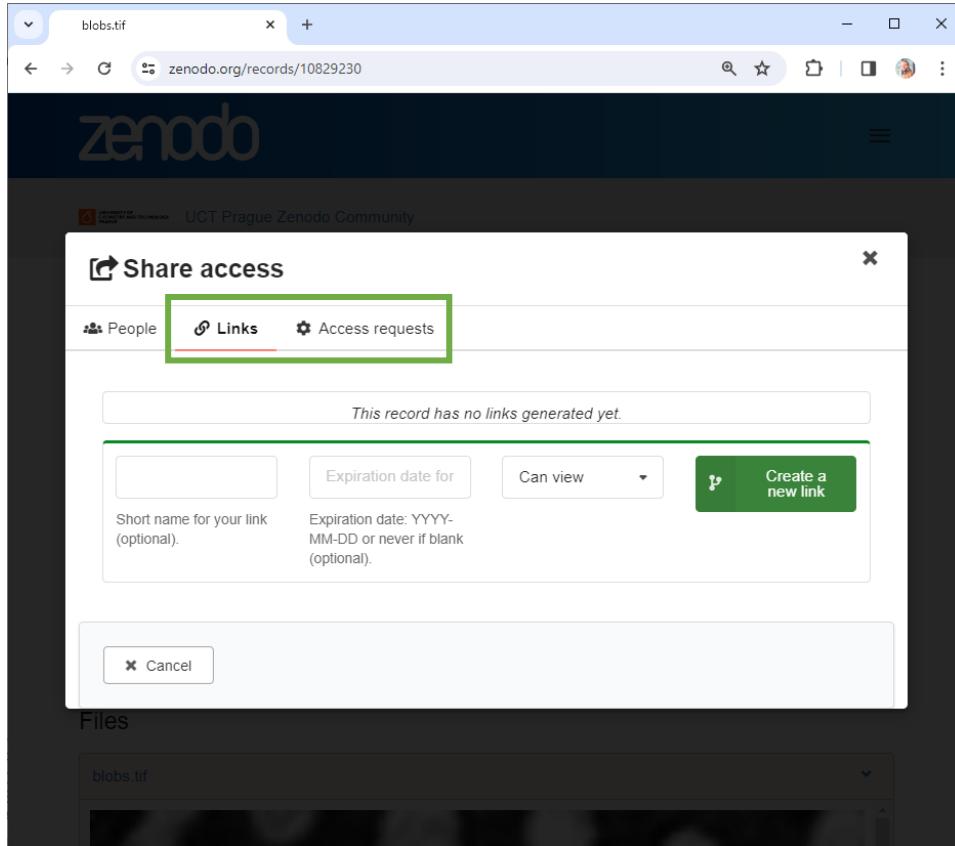
A screenshot of a Zenodo dataset page for "blobs.tif". The page shows basic metadata: Published March 18, 2024 | Version v1. The dataset is labeled as "Dataset" and "Restricted". It has 0 views and 0 downloads. A blue arrow points from the "Show more details" link under the stats to the "Show affiliations" button below it. The "Files" section is highlighted with a pink background, showing a "Restricted" status message: "The record is publicly accessible, but files are restricted to users with access." The "Citations" section shows zero results. The "External resources" section lists "Indexed in" OpenAIRE.



A screenshot of the same Zenodo dataset page for "blobs.tif", but with a different access setting. The "Dataset" and "Restricted" buttons are now grayed out. The "Files" section is now highlighted with a pink background, showing an "Open" status message: "The record is publicly accessible, but files are restricted to users with access." The "Citations" section shows zero results. The "External resources" section lists "Indexed in" OpenAIRE. The right side of the screen shows a summary box with 26 views and 0 downloads, and a "Show more details" link.

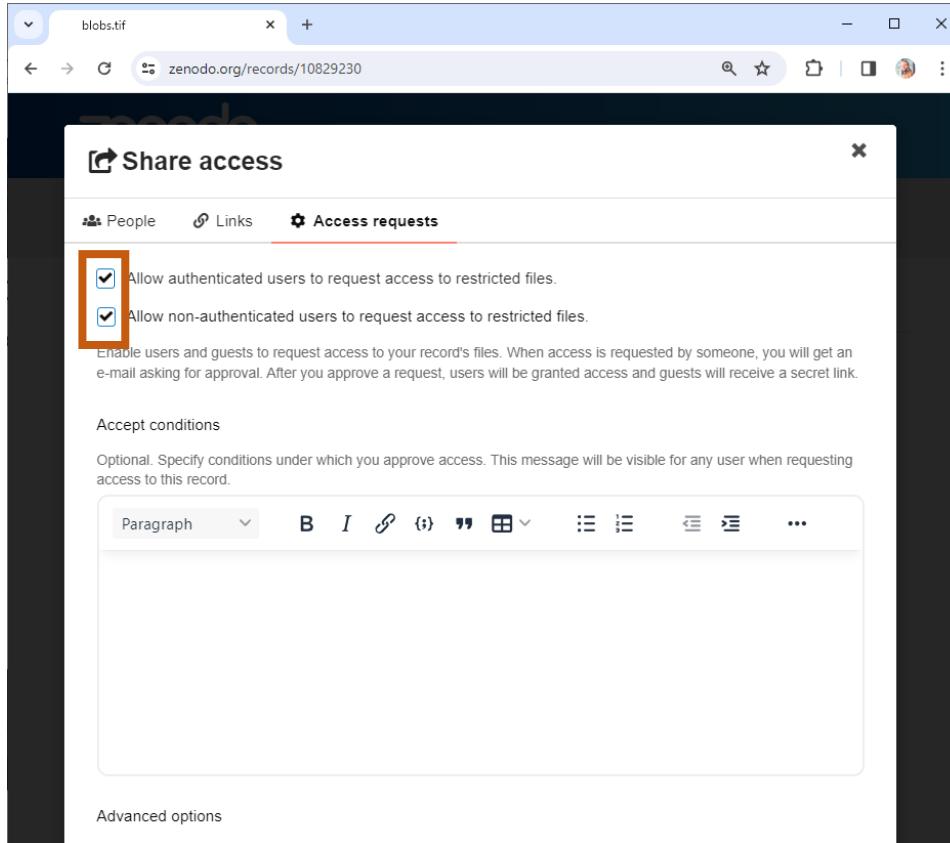
Accessibility

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



Accessibility

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



A screenshot of the Zenodo dataset page for blobs.tif. The top navigation bar shows the URL zenodo.org/records/10829230. The dataset was published on March 18, 2024, and has version v1. The status is labeled as 'Dataset' and 'Restricted'.

The main content area shows the file blobs.tif and its details:

- Published March 18, 2024 | Version v1
- Dataset
- Restricted
- blobs.tif
- Haase, Robert^{1,2} [ORCID], Schätz, Martin
- Show affiliations

The description notes: "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>".

The record serves as an example of data publication with closed access.

On the right side, there are sections for Views (26), Downloads (0), Versions (Version v1), External resources (Indexed in OpenAIRE), Communities (UCT Prague Zenodo Community), and Details (DOI: DOI 10.5281/zenodo.10829230).

The central part of the page features a 'Request access' form:

- Files: Restricted
- 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*
- Request message
- I agree to that my full name and email address is shared with the owners of the record
- Request access

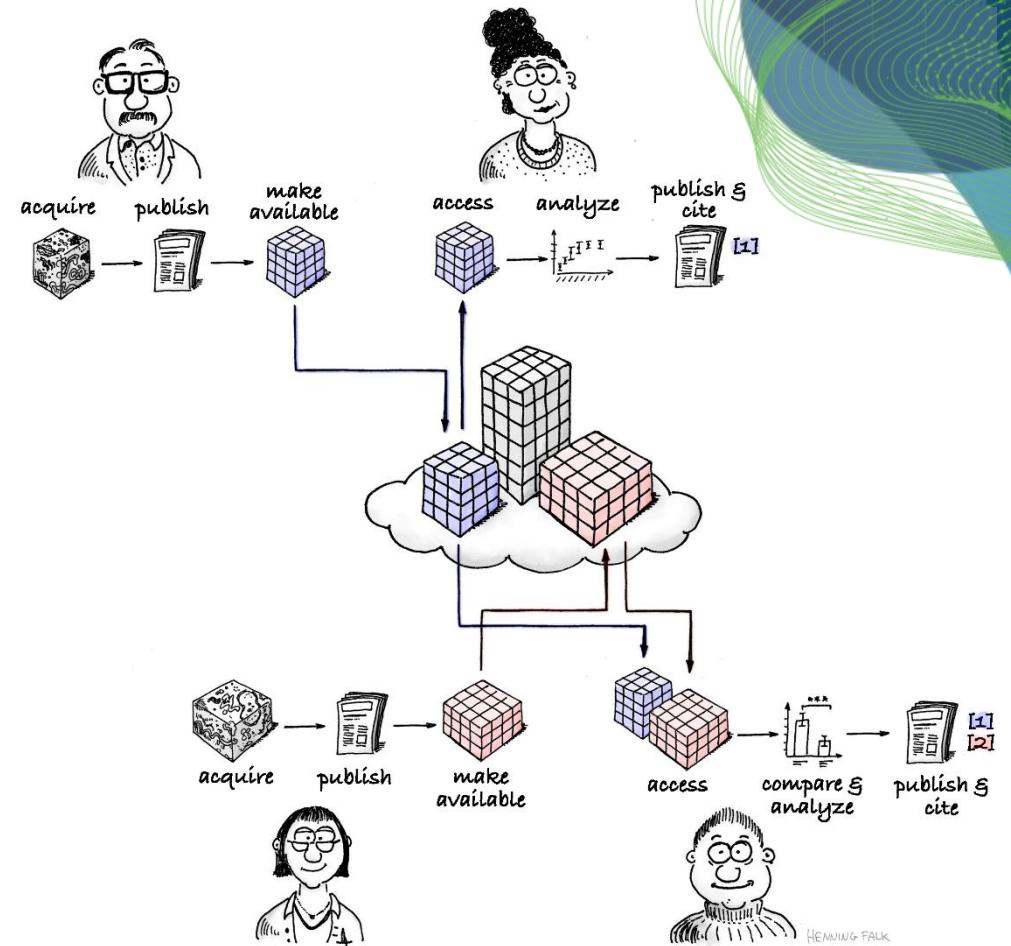
Interoperability

- 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



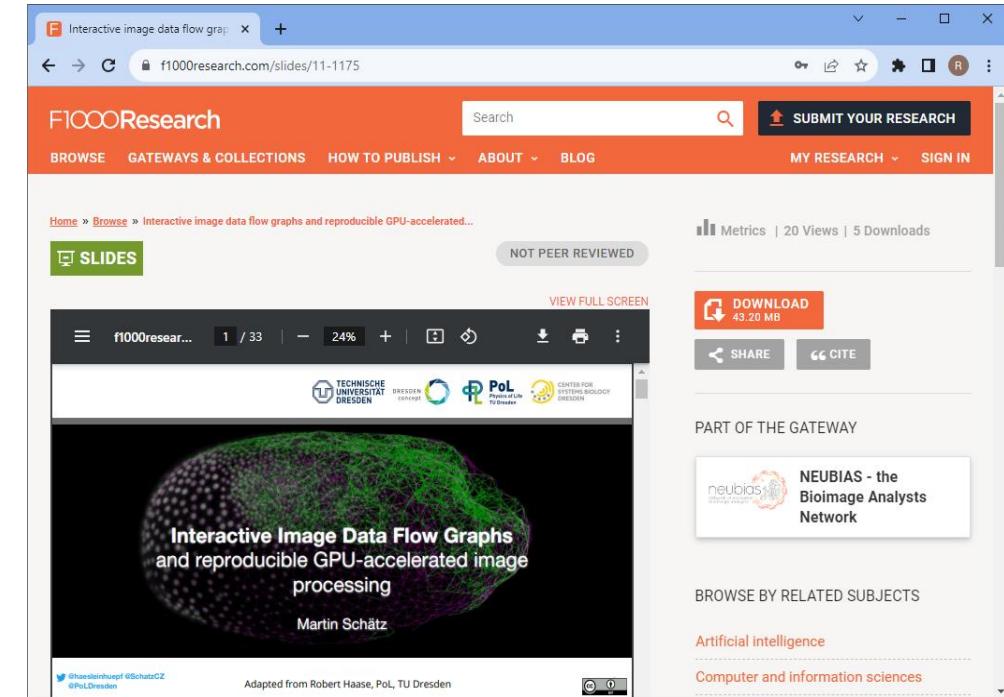
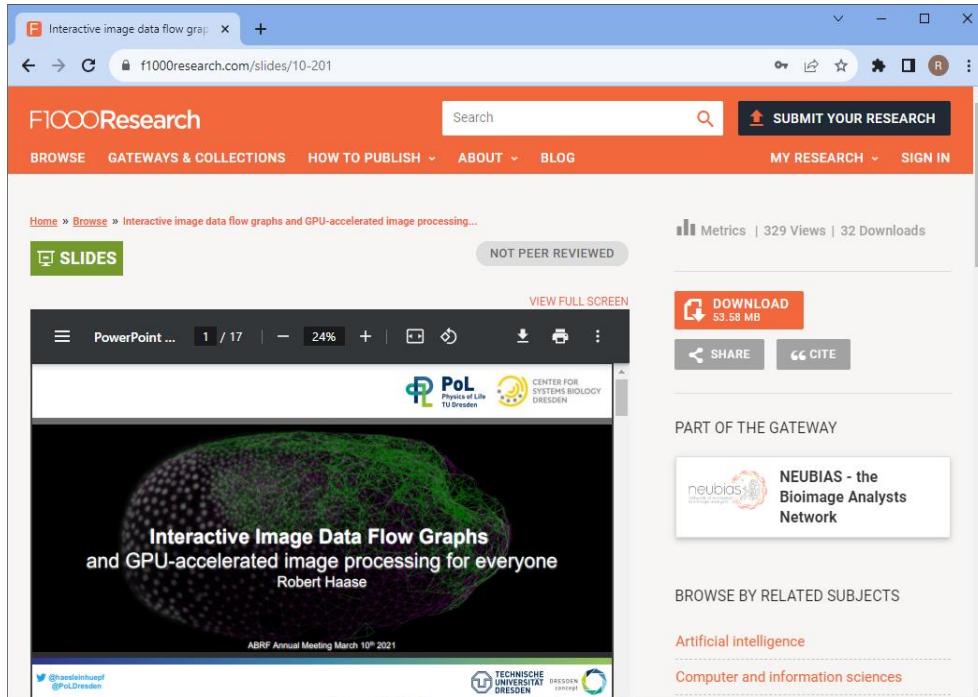
Reusability

- 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

The screenshot shows a Zenodo dataset page for a dataset associated with an article. The page includes sections for dataset statistics, versions, and citation/export options.

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

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

Views: 240 | **Downloads:** 131

Versions:

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

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/>

This screenshot shows a web browser displaying a preprint from Earth ArXiv. The title of the preprint is "An analysis of the dynamic range of Distributed Acoustic Sensing for Earthquake Early Warning". Below the title, it states: "This is a Preprint and has not been peer reviewed. This is version 1 of this Preprint." The page includes sections for "Downloads" (with a "Download Preprint" button) and "Authors" (listing Martijn van den Ende, Alister Trabattoni, Marie Baille, and Diane Rivet). A small thumbnail image of the preprint's content is visible on the left.

Code, Models & Data

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

Ongoing work on code and models

Publicly funded archive

This screenshot shows a web browser displaying a dataset record on Zenodo. The title of the dataset is "Data and scripts for: An analysis of the dynamic range of Distributed Acoustic Sensing for Earthquake Early Warning". It was published on April 18, 2024, by Martijn van den Ende (Contact person). The dataset has 89 views and 137 downloads. The page also lists contributors and project members, including Martijn van den Ende, Alister Trabattoni, Marie Baille, and Diane Rivet. A DOI link is provided at the bottom: 10.5281/zenodo.10993305.

Sharing files on Zenodo

- ... is easier than you think

The figure consists of three side-by-side screenshots of web pages. The left screenshot shows a blog post on the FocalPlane website titled "Sharing research data with Zenodo". The middle screenshot shows the Zenodo landing page with a brief introduction and a screenshot of a dataset page. The right screenshot shows the Zenodo "Upload form" interface.

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 datasets are

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.

Upload form.

zenodo

Search Upload Communities

Upload

New Upload

Sort desc

Draft 3 Published 89 All versions

December 12, 2022 (2022.12.12) 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 haesleinhuepf/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-zendodo) Software Open Access haesleinhuepf/napari-assistant-plugin-generator: 0.1.0-for-zendodo Created Nov 9, 2022 2:57:39 PM, modified Nov 9, 2022 2:57:34 PM

December 24, 2022 (0.1.4) Software Open Access haesleinhuepf/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



DRESDEN LEIPZIG

CENTER FOR SCALABLE DATA ANALYTICS
AND ARTIFICIAL INTELLIGENCE

Research Software Management

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

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

pip install ...



conda install ...



uv add ...



<other>



Conda environments

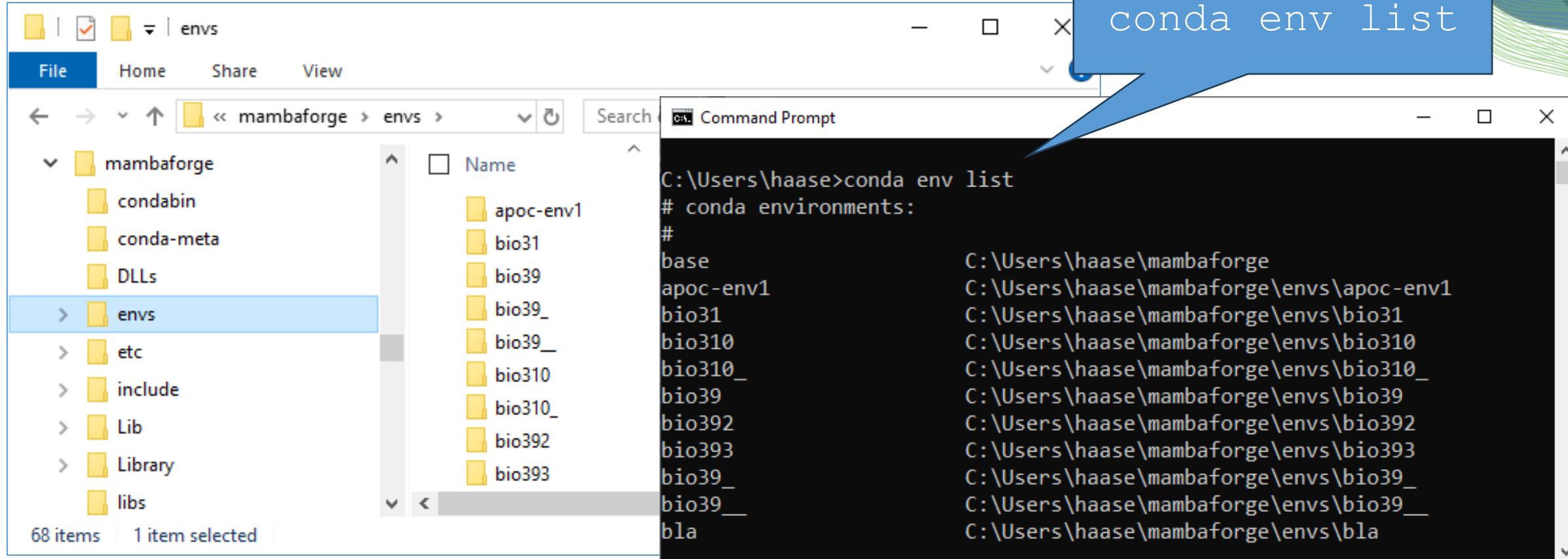
- What is the difference between
 pip install ...
 and
 conda install ...
 ?

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

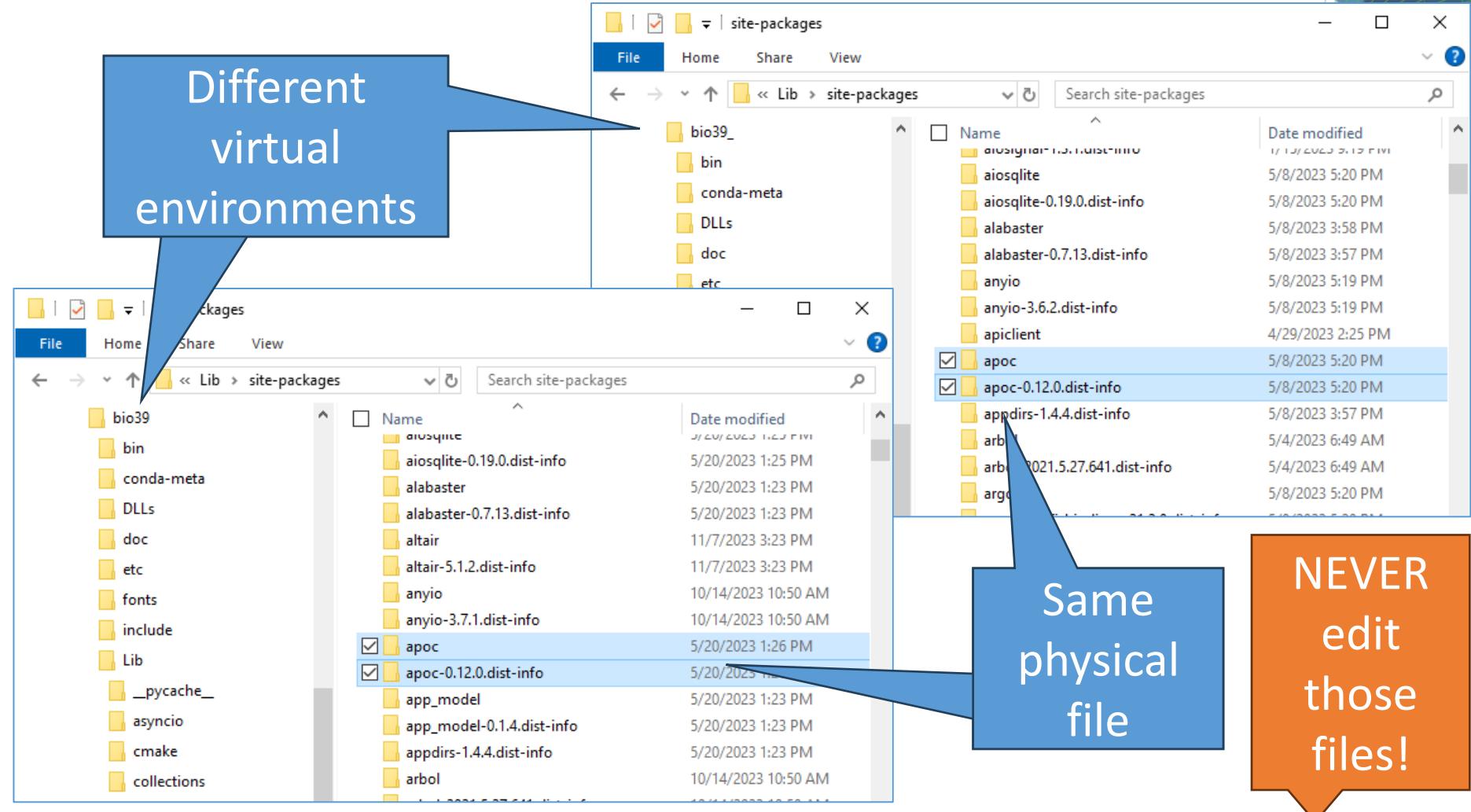
Conda environments

- A conda environment is *just a folder*



Conda environments

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



Broken [conda] environments

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

The screenshot shows a Windows desktop with three command prompt windows and a Jupyter Notebook interface.

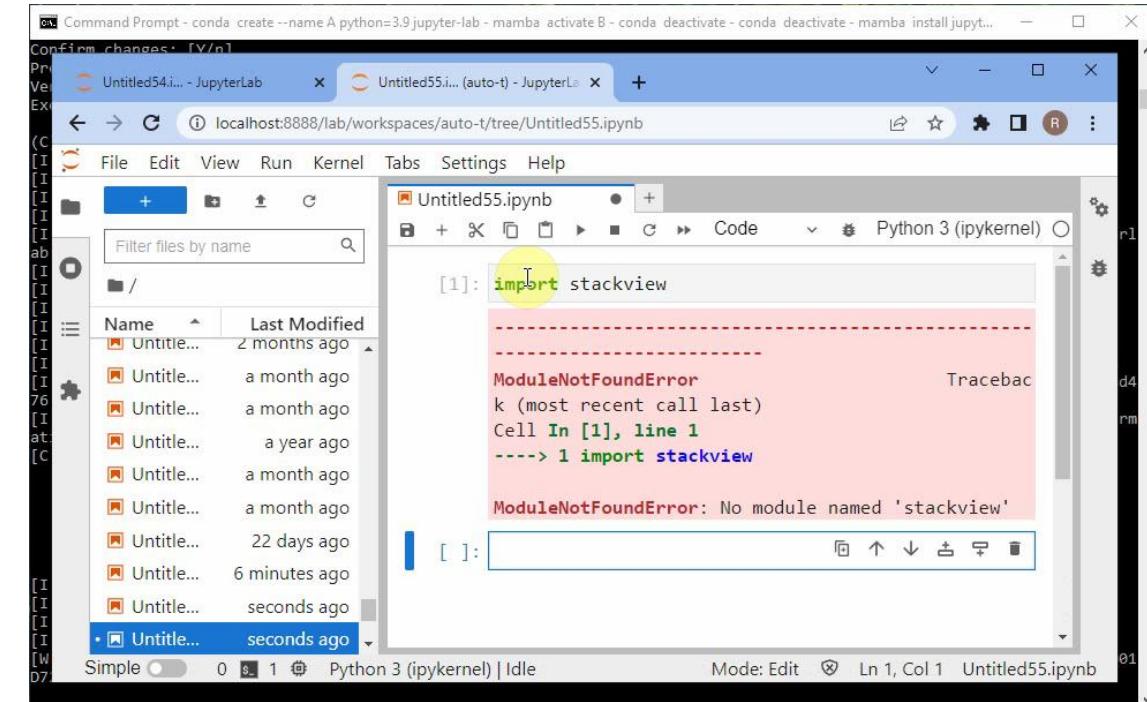
- Left Command Prompt:** Shows the output of running `mamba install stackview -c conda-forge`. It includes the MAMBA logo and version information: "mamba (0.25.0) supported by @QuantStack".
- Middle Command Prompt:** Shows the output of running `conda create --name A python=3.9 jupyter-lab - mamba activate B - conda deactivate`. It lists many packages being installed from the conda-forge channel, such as toolz, tornado, 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 Command Prompt:** Shows the output of running `conda create --name A python=3.9 jupyter-lab - mamba activate B - conda deactivate - conda deactivate - mamba install jupyter-lab`. It shows a "Confirm changes: [Y/n]" prompt and transaction logs.
- Jupyter Notebook:** Shows a cell with the code `[1]: import stackview` which results in a `ModuleNotFoundError`.

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



The screenshot shows a Jupyter Notebook interface with two tabs: 'Untitled54.ipynb' and 'Untitled55.ipynb'. The 'Untitled55.ipynb' tab is active, displaying a code cell with the following content:

```
[1]: import stackview
```

An error message is visible in the output area:

```
ModuleNotFoundError
k (most recent call last)
Cell In [1], line 1
----> 1 import stackview

ModuleNotFoundError: No module named 'stackview'
```

The 'File Explorer' sidebar on the left shows a list of files in the current directory, including several 'Untitled...' files and one file named 'stackview'.

Broken [conda] environments

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



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

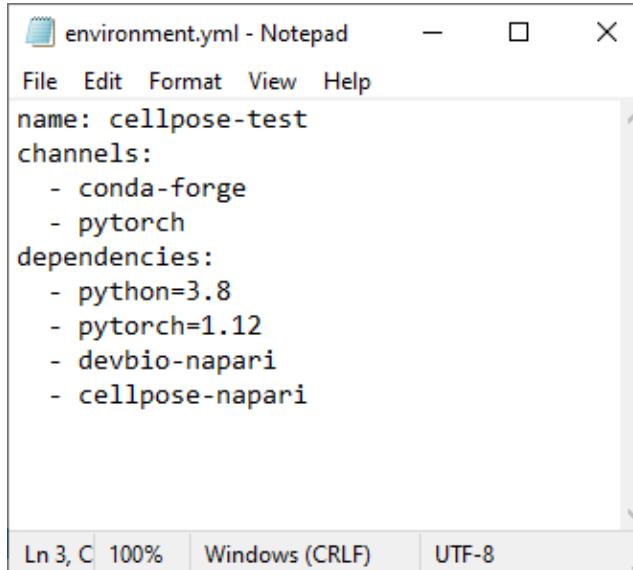
Broken [conda] environments

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

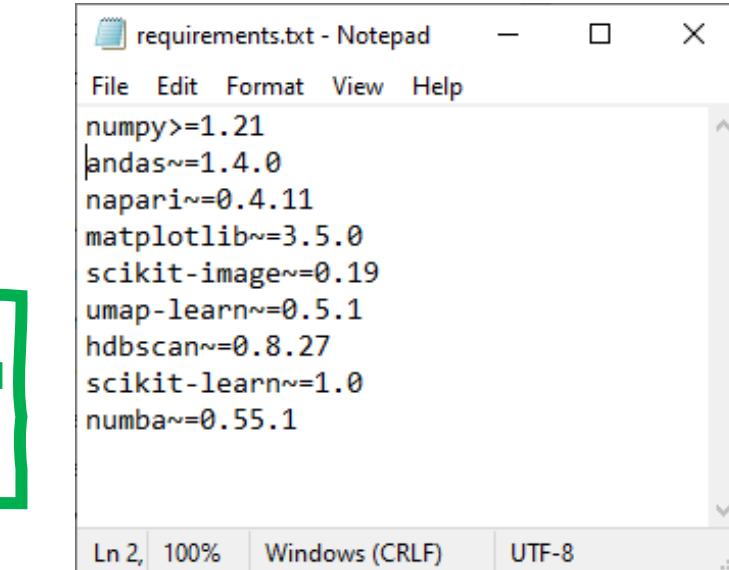


```
environment.yml - Notepad
File Edit Format View Help
name: cellpose-test
channels:
- conda-forge
- pytorch
dependencies:
- python=3.8
- pytorch=1.12
- devbio-napari
- cellpose-napari

Ln 3, C 100% Windows (CRLF) UTF-8
```

conda env create -f environment.yml

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



```
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

Ln 2, 100% Windows (CRLF) UTF-8
```

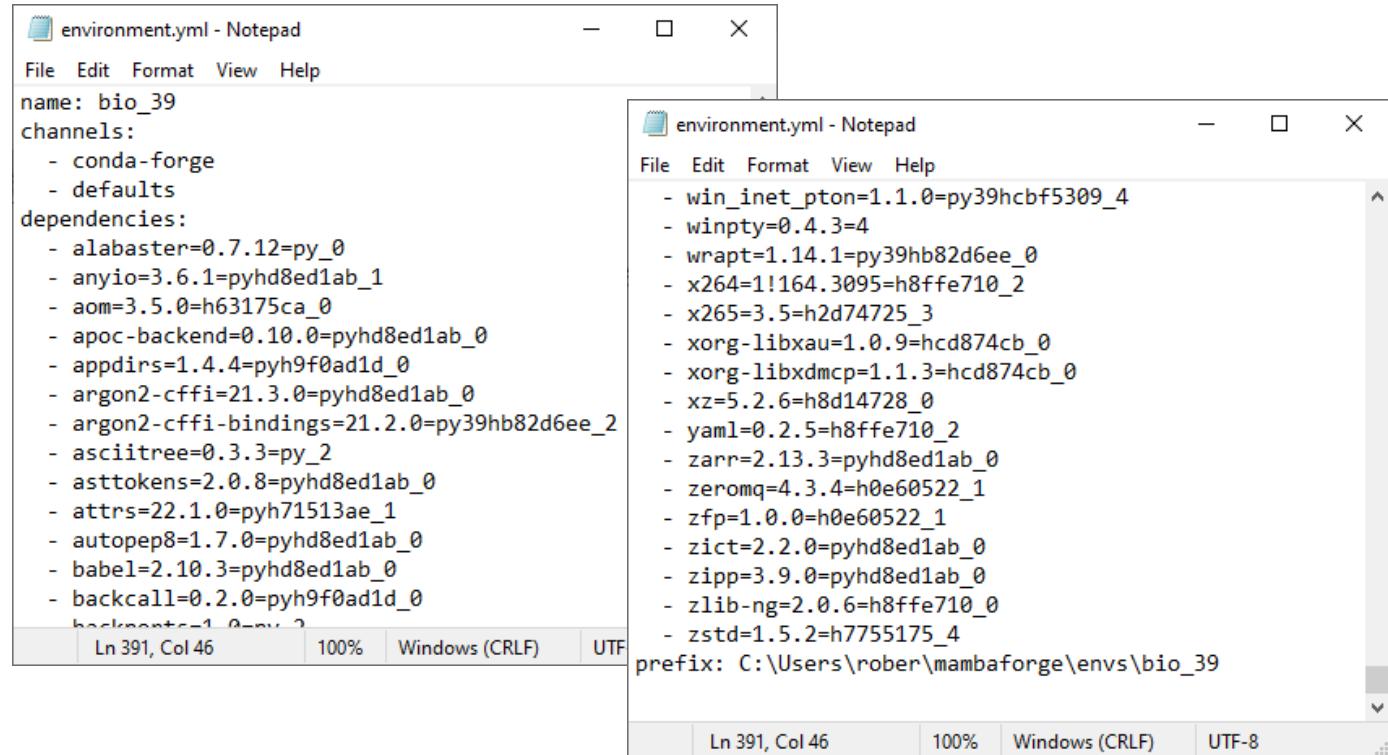
pip install -r requirements.txt

Your future
self will
thank you

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
backgrounds=1.0=py_2
Ln 391, Col 46 100% Windows (CRLF) UTF-8
```



```
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
Ln 391, Col 46 100% Windows (CRLF) UTF-8
```

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:

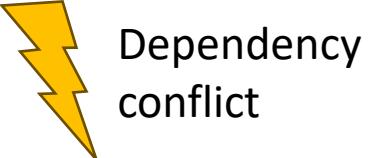
pip install package_a

conda install package_a

Depends on:
numpy<1.22.0

pip install package_b

conda install package_b



Dependency
conflict
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

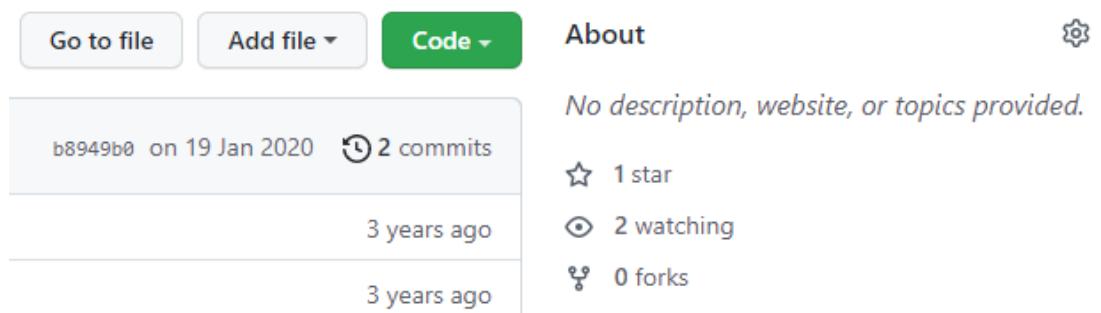
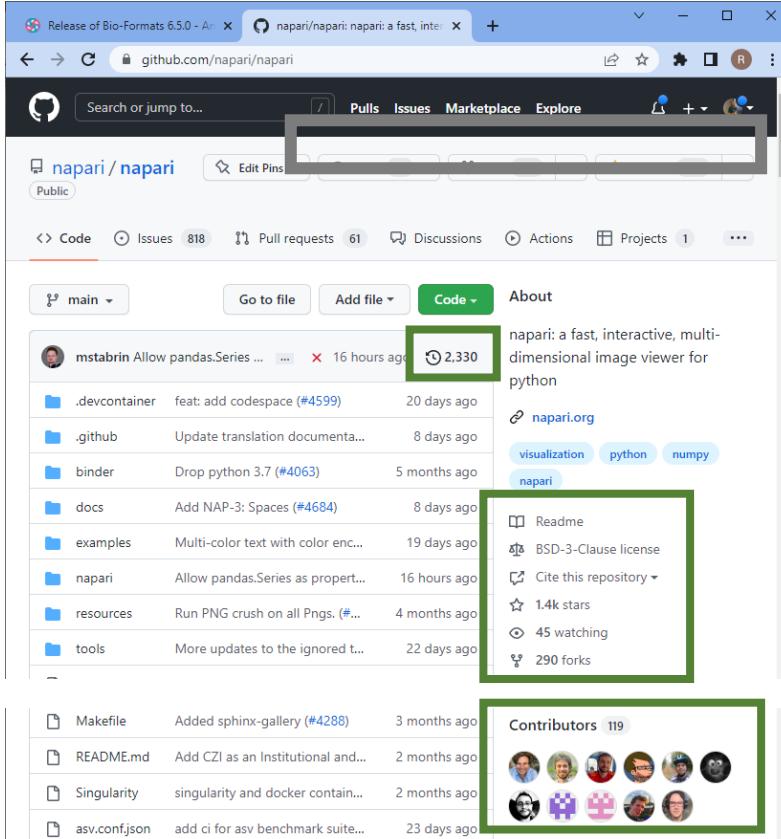
Because git is not a
python package

conda install git

passes

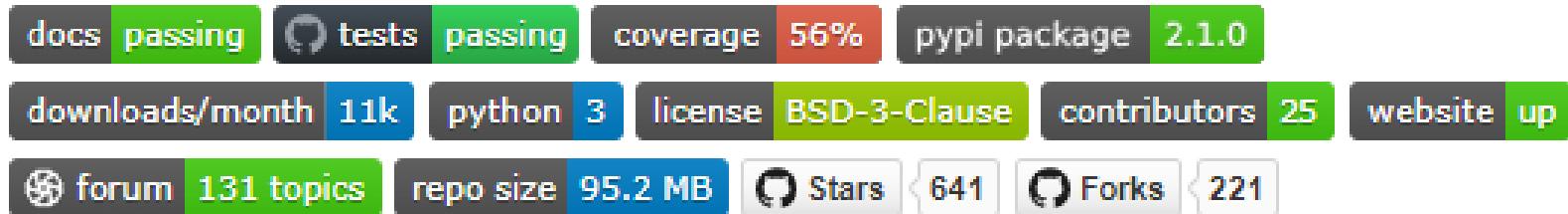
Software quality indicators

- 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.

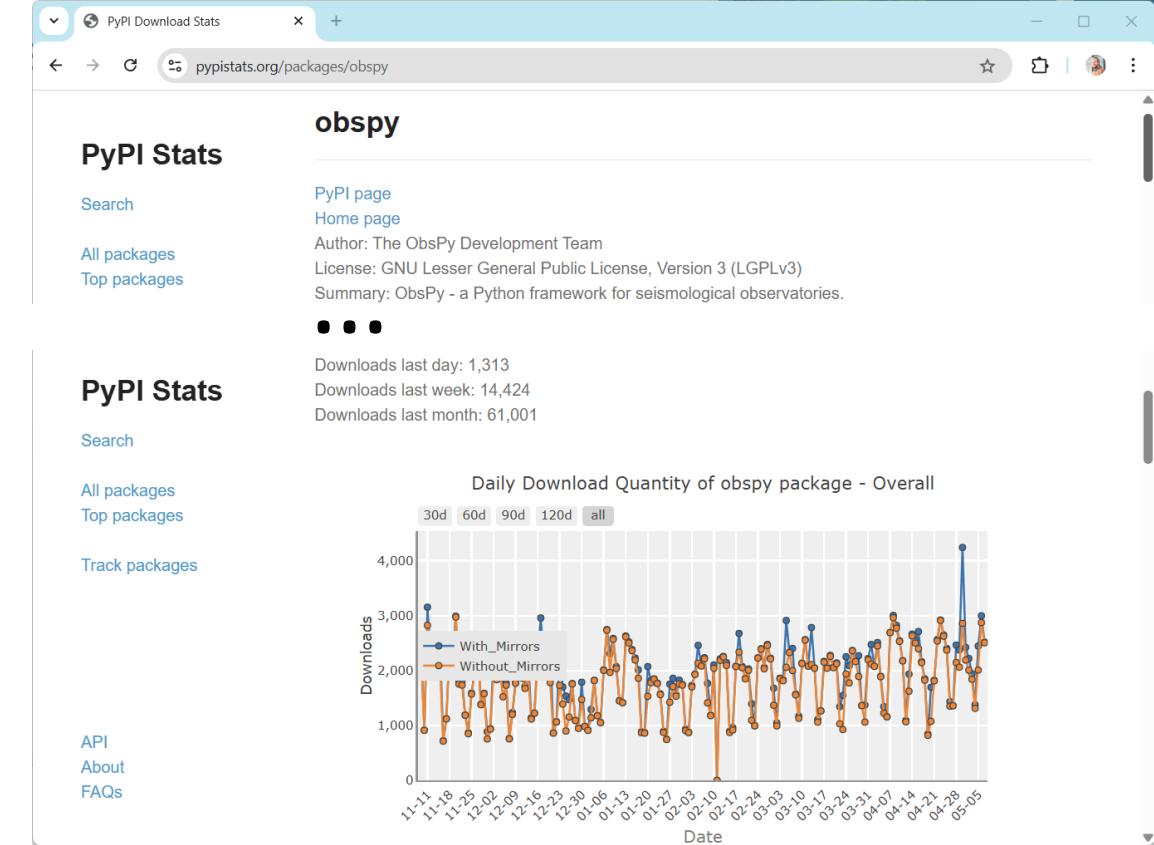
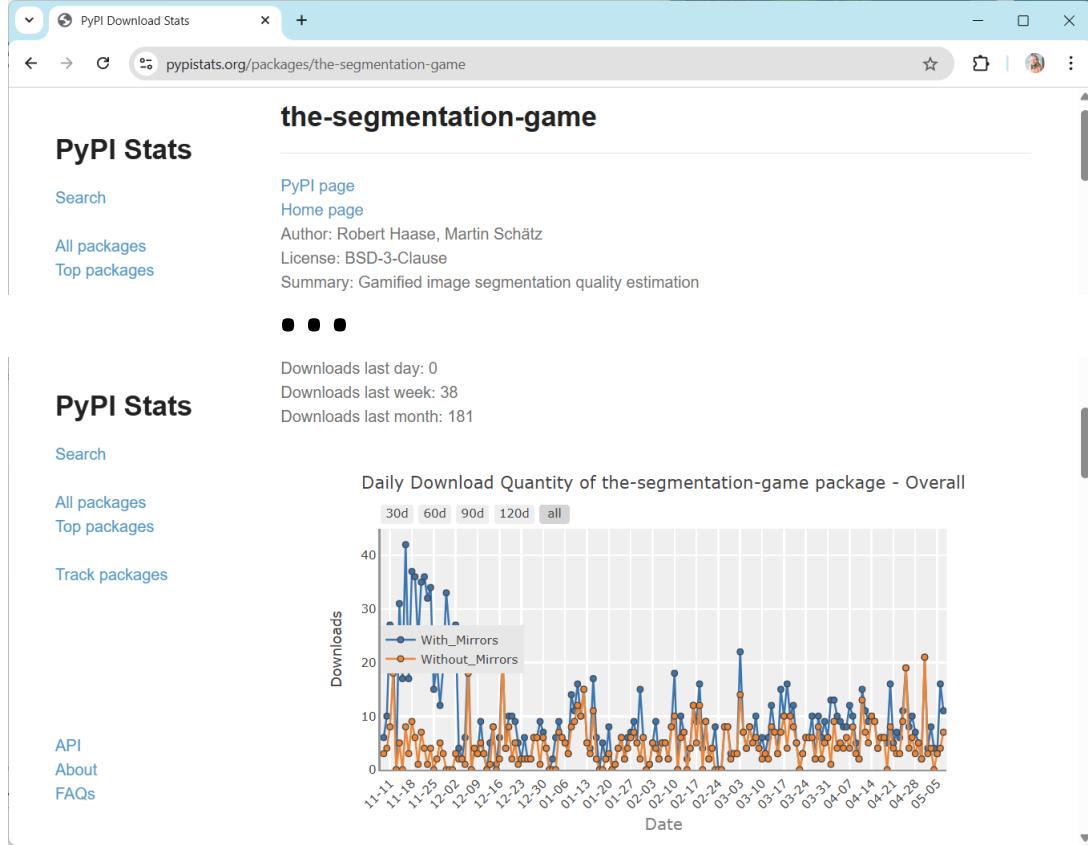


Note, github badges
cannot be *deserved*.
Developers put them
there



Software quality indicators

- Download statistics: pypi



Software quality indicators

- Scientific publications

MouseLand/cellpose: a generalist algorithm for cellular segmentation

github.com/MouseLand/cellpose

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

nature methods

Explore content About the journal Publish with us

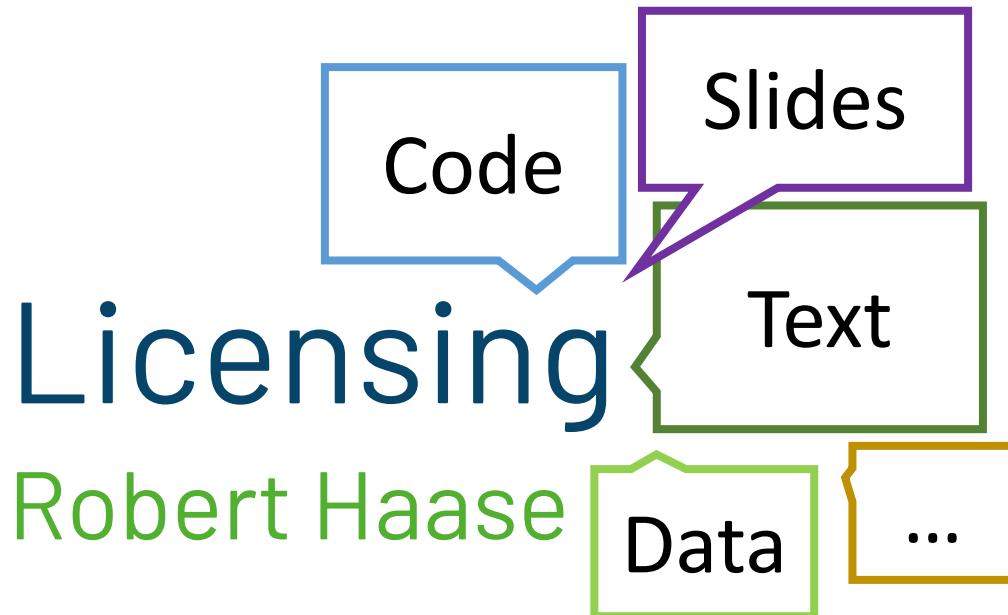
nature > nature methods > articles > article

Article | Open access | Published: 07 November 2022

Marius Pachitariu & Carsen Stringer

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

55k Accesses | 120 Citations | 116 Altmetric | Metrics



Licensing

Robert Haase



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.

Licensing
Robert Haase



What are the
consequences of
this sentence?

Quiz

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

None



Public
Domain



Creative
Commons



BSD/GPL/
MIT/...

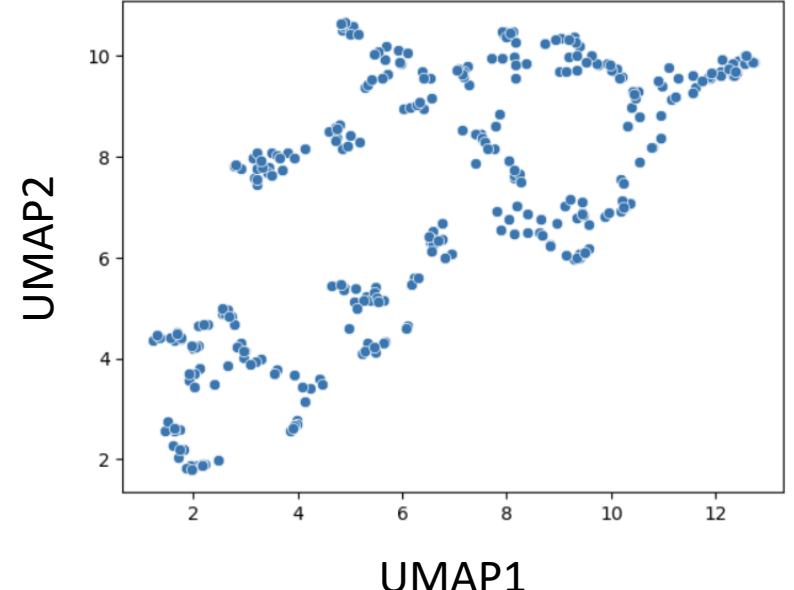


Quiz

How can you reuse
this plot?
What is allowed?

New Tab

Clustering objects can be challenging when working with many parameters, in particular when interacting with data manually. To reduce the number of parameters, dimensionality reduction techniques such as the Uniform Manifold Approximation Projection (UMAP) have been developed. In this notebook we use the technique to differentiate nuclei in an image which are mitotic from those which are not mitotic.



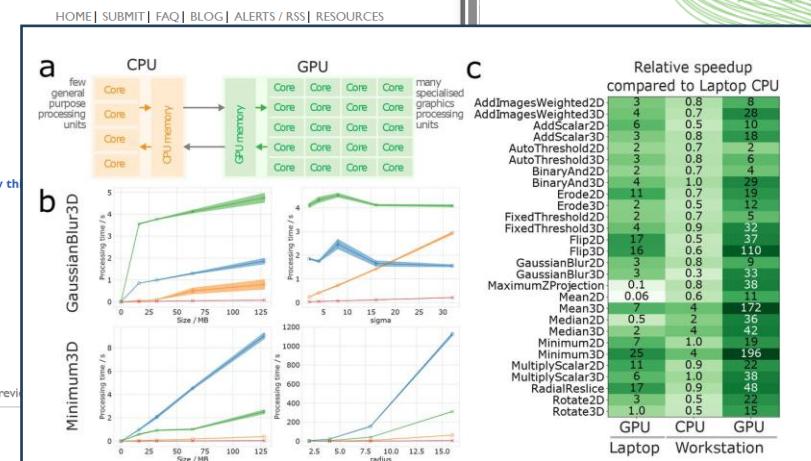
Copyright © 2025 Journal XYZ. All Rights Reserved.

Hint: Search for pre-prints

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

The screenshot shows the Nature Methods website. The main article title is 'CLIJ: GPU-accelerated image processing for everyone'. It includes author information (Robert Haase, Loic A. Royer, Peter Steinbach, Deborah Schmidt, Alexander Dibrov, Uwe Schmidt, Martin Weigert, Nicola Maghelli, Pavel Tomancak, Florian Jug & Eugene W. Myers), publication details (Nature Methods 17, 5–6 (2020) | Cite this article), and access options. A note states 'Access to this article via Universitätsbibliothek Leipzig AG eMedien is not available.' Below the main content, there's a sidebar for 'Access options' and a promotional section for 'Nature and 54 other Nature Portfolio journals'.

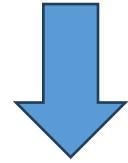
The screenshot shows the bioRxiv preprint server. The article title is 'CLIJ: GPU-accelerated image processing for everyone'. It includes the same author and publication information as the Nature Methods version. The bioRxiv interface includes sections for 'Abstract', 'Full Text', 'Info/History', and 'Metrics'. A note at the bottom states 'Now published in Nature Methods doi: 10.1038/s41592-019-0650-i'.



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

The screenshot shows the arXiv.org interface for a physics paper. The title is "SeisCLIP: A seismology foundation model pre-trained by multi-modal data for multi-purpose seismic feature extraction". The authors listed are Xu Si, Ximeng Wu, Hanlin Sheng, Jun Zhu, Zefeng Li. The abstract discusses the limitations of training specific deep learning models across domains and introduces SeisCLIP, a transformer encoder for seismic spectrum and an MLP encoder for integrating phase and source information. The page includes a "view license" link under the "Access Paper" section, which is highlighted with a red box.

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

arXiv.org - Non-exclusive license to distribute

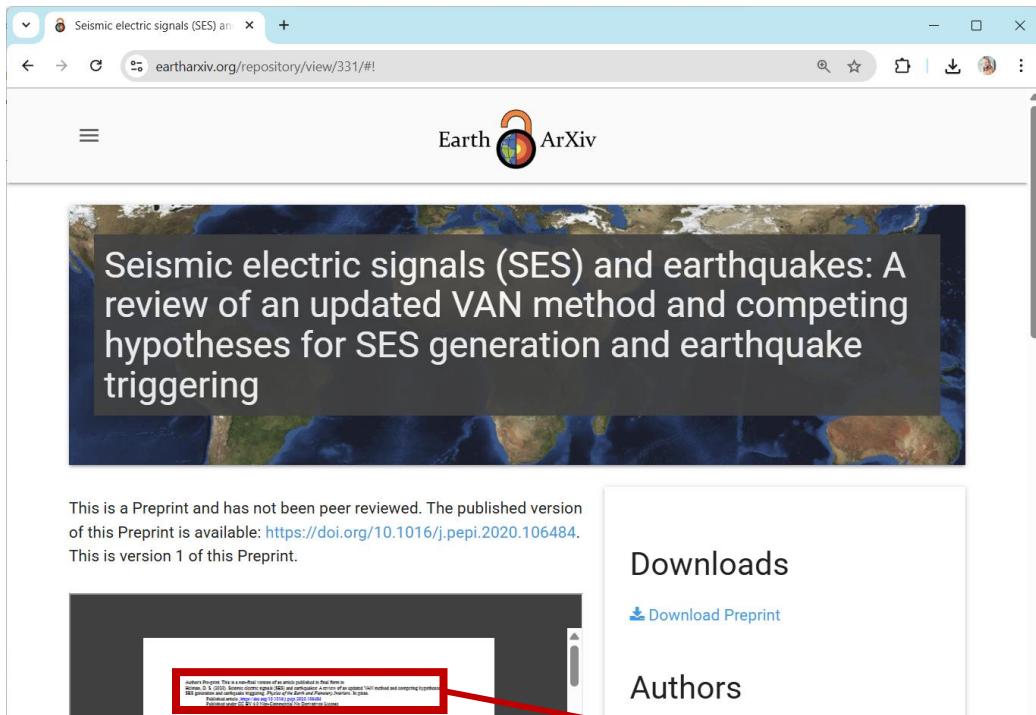
The URI <http://arxiv.org/licenses/nonexclusive-distrib/1.0/> is used to record the fact that the submitter granted the following license to arXiv.org on submission of an article:

- 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



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

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

But I'm not
allowed!

Licensing: Permissive versus restrictive

Example

SPRINGER LINK

Find a journal Publish with us Track your research

Search



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 (LNCS, 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.

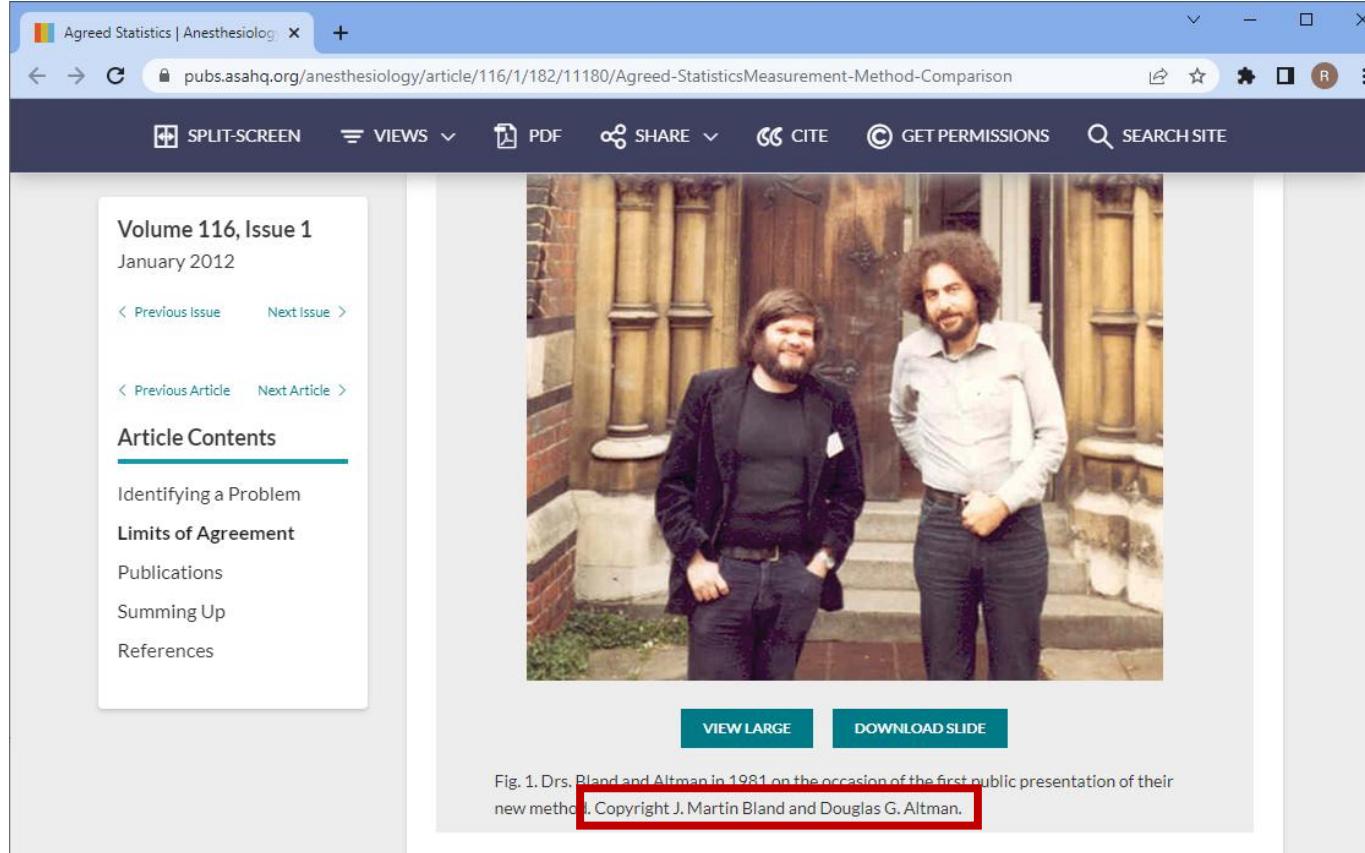
Adaptations/modifications - Springer Nature allows adaptation of figures for style and formatting purposes under this license under the condition that this does not alter the meaning of the content.

I would like to... reuse in training materials/CME materials	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



Agreed Statistics | Anesthesiologist

SPLIT-SCREEN VIEWS PDF SHARE CITE GET PERMISSIONS SEARCH SITE

Volume 116, Issue 1
January 2012

Previous Issue Next Issue < Previous Article Next Article >

Article Contents

- Identifying a Problem
- Limits of Agreement
- Publications
- Summing Up
- References

VIEW LARGE DOWNLOAD SLIDE

Fig. 1. Drs. Bland and Altman in 1981 on the occasion of the first public presentation of their new method. Copyright J. Martin Bland and Douglas G. Altman.

Reusing some of your pictures for teaching

Von: Martin Bland
An: Robert Haase

5. Juni 2020 22:37

You are welcome to use any of my pictures in your teaching.
Thanks for asking.

Martin

On Fri, 5 Jun 2020 at 21:24, <rhaase@mpi-cbg.de> wrote:

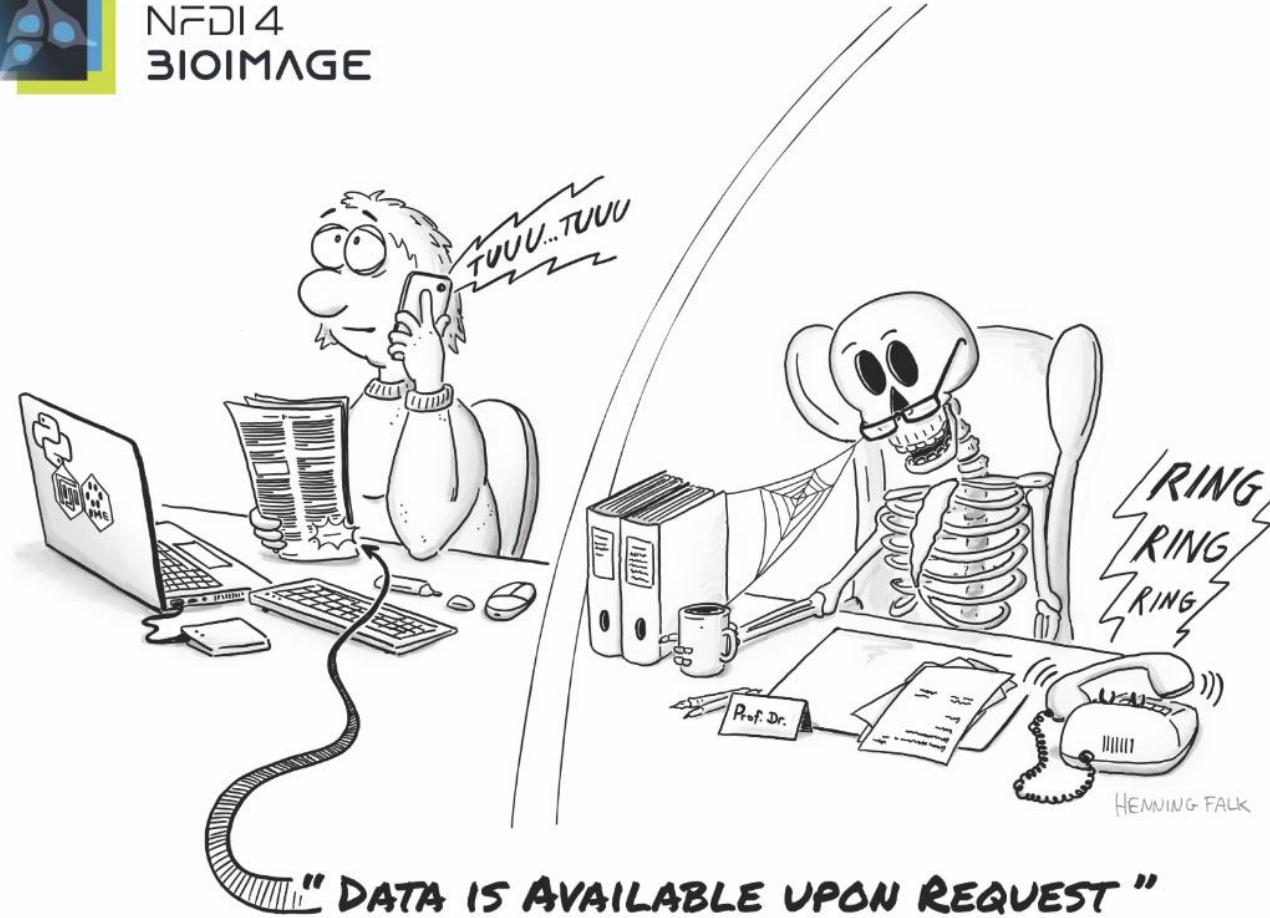
Dear Prof. Bland,

I hope you are doing well. I'm approaching you because I'm preparing a lecture for students about Bio-statistics at the Technical University Dresden and I would like to use some pictures where you are copyright holder.

I'm referring to the photos published in this article:

<https://anesthesiology.pubs.asahq.org/article.aspx?articleid=1933992>

Licensing: Permissive versus restrictive



"DATA IS AVAILABLE UPON REQUEST"

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

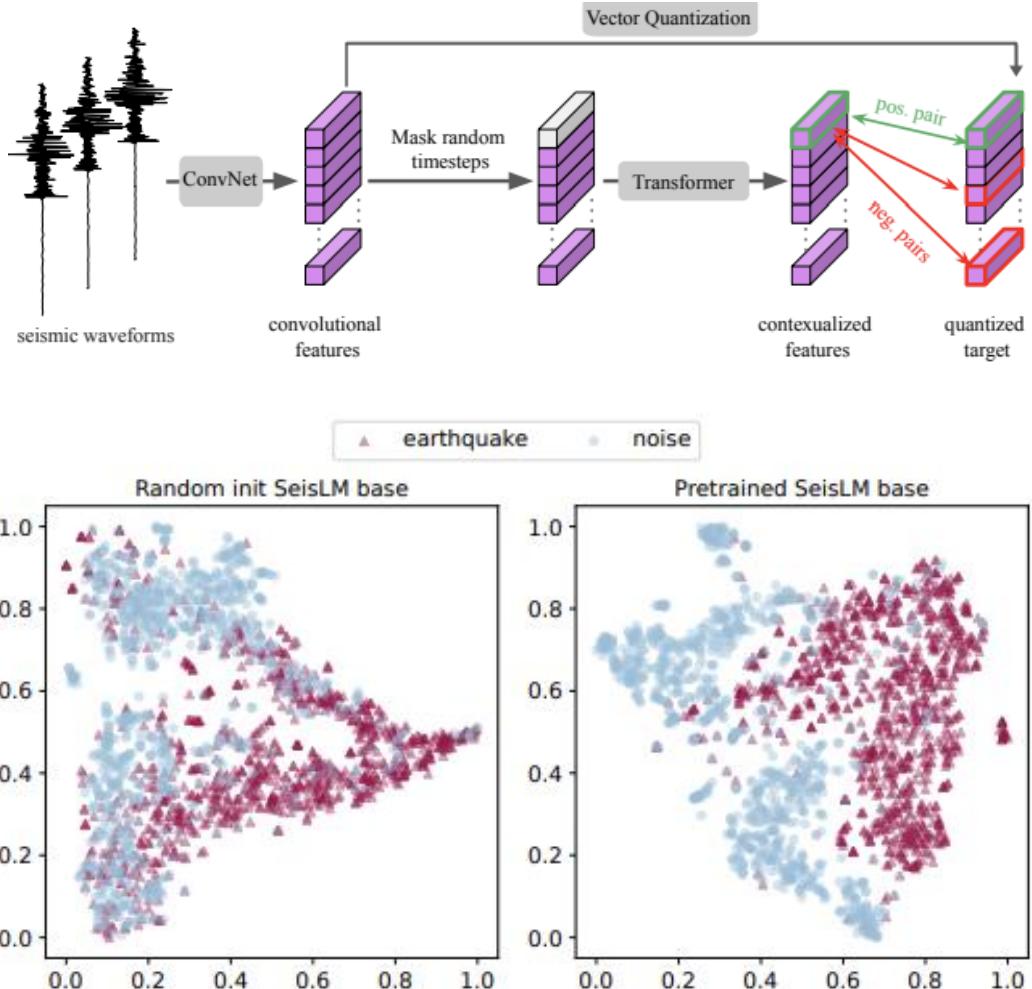
	Download and share for free	Reuse parts, e.g. Figures	Reuse parts, e.g. in paid training
CC-BY	✓	✓	✓
CC-BY-SA	✓ Only under CC-BY-SA	✓ Only under CC-BY-SA	✓ Only under CC-BY-SA
CC-BY-NC	✓	✓ (if free of charge)	✗
CC-BY-ND	✓	✗	✗
CC-BY-NC-ND	✓	✗	✗

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](#) by T. Liu et al.

A screenshot of a web browser displaying an arXiv preprint. The title of the paper is 'SeisLM: a Foundation Model for Seismic Waveforms' by Tianlin Liu, Jannes Münchmeyer, Laura Laurenti, Chris Marone, Maarten V. de Hoop, Ivan Dokmanić. The paper was submitted on 21 Oct 2024. The arXiv URL is arxiv.org/abs/2410.15765v1. A green arrow points to the 'view license' link under the 'Access Paper:' section, which is associated with the CC-BY 4.0 license mentioned in the text above.

Licensing: Permissive versus restrictive

Example

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

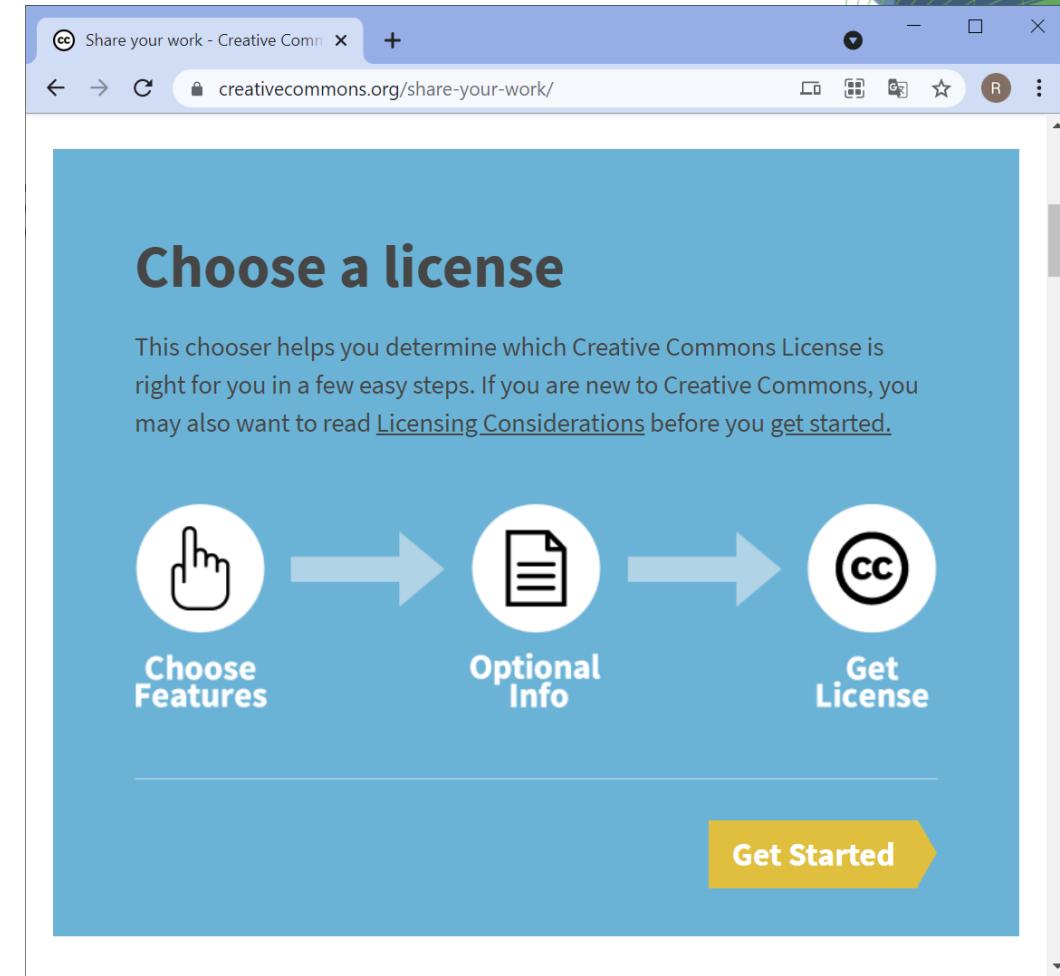
The screenshot shows the BioImage Archive homepage with a study details page overlaid. The study title is "The glucosylceramide synthase inhibitor PDMP causes lyso-somal lipid accumulation and mTOR inactivation". The study is led by Pia Hartwig and Doris Höglunger from Heidelberg University. The accession number is S-BIAD144. A data file table is shown, listing three entries: experimentA_11_WT_Miglustat.cz (1.6 MB), experimentA_12_SGPL1_PDMP.cz (1.6 MB), and experimentA_13_SGPL1_PDMP.cz (1.6 MB). All files are categorized as "Study Component" and show "continuous labelling".

The screenshot shows the bioRxiv preprint server page for the article "The BioImage Archive - building a home for life-sciences microscopy data" by Matthew Hartley, Gerard J. Kleywegt, Ardan Patwardhan, Ugis Sarkans, Jason R. Swedlow, and Alvis Brazma. The article was posted on February 11, 2022, and is now published in the Journal of Molecular Biology with DOI 10.1101/2022.167505. The copyright notice states: "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." The page also includes links for download, print/save options, revision summary, citation tools, and social sharing.

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](#). 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](#). 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.

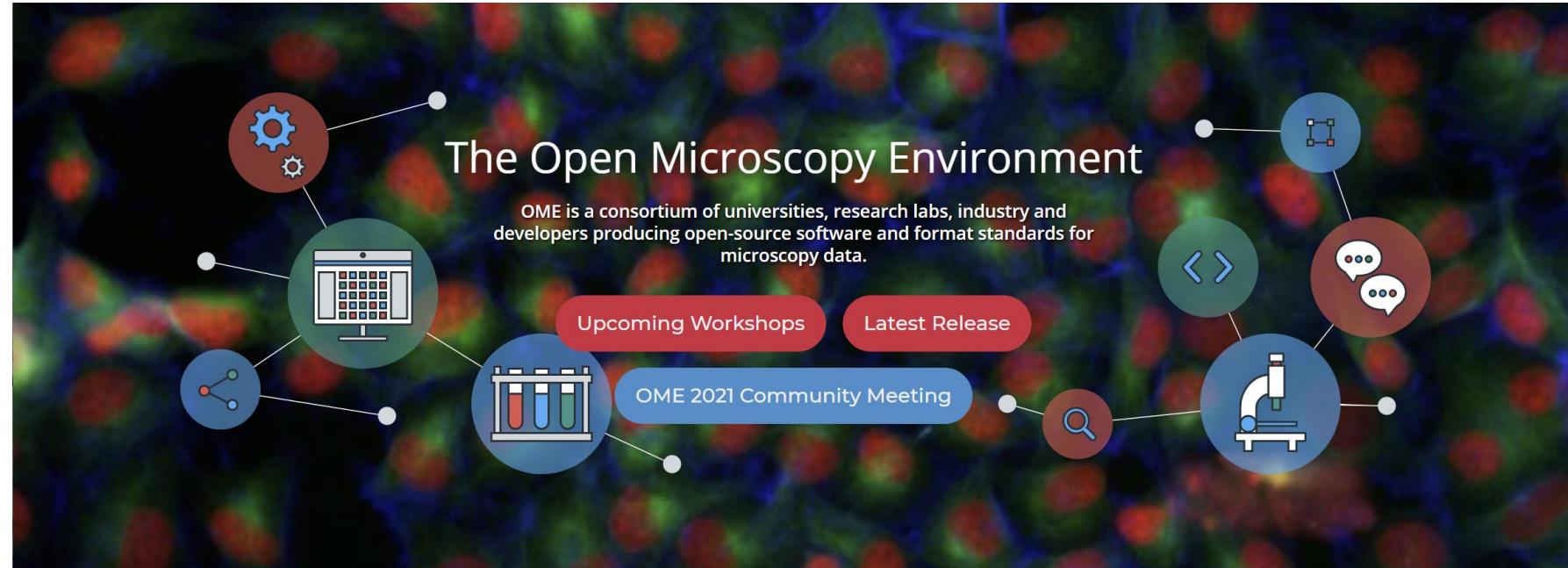
CC BY includes the following elements:

BY - Credit must be given to the creator

Content on this slide was adapted from <https://creativecommons.org/about/cclicenses/> which is licensed under a Creative Commons Attribution 4.0 International license. 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](#)

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:



- Credit must be given to the creator



- Adaptations must be shared under the same terms

“Restrictive”
licensing

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

Licensing: Creative Commons (CC)



to the creator.

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

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/cclicenses/> which is licensed under a [Creative Commons Attribution 4.0 International license](#). Icons by The Noun Project.

Quiz

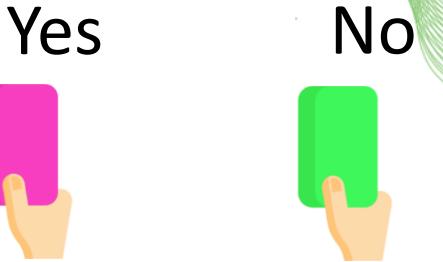
- May I use one of the Figures from this preprint?
- May I download and redistribute this preprint to students of a course for free?

This is a Preprint and has not been peer reviewed. The published version of this Preprint is available: <https://doi.org/10.1016/j.rsase.2025.101538>. This is version 1 of this Preprint.

Downloads
[Download Preprint](#)

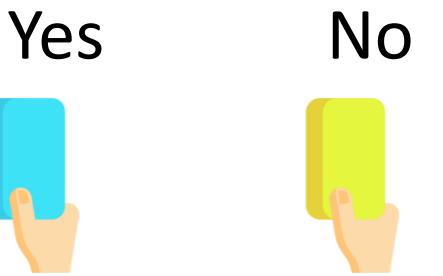
Authors
Sepideh Jalayer, Samira Alkaee Taleghan, Rafael Pires de Lima, Behzad Vahedi, Nick Hughes, Farnoush Banaei-Kashani, Morteza Karimzadeh

ABSTRACT
Accurate mapping of sea ice is crucial for marine navigation and monitoring climate change. Automating sea ice mapping remains challenging due to remotely-sensed signal ambiguity, the dynamic nature of sea ice, and limited field measurements.



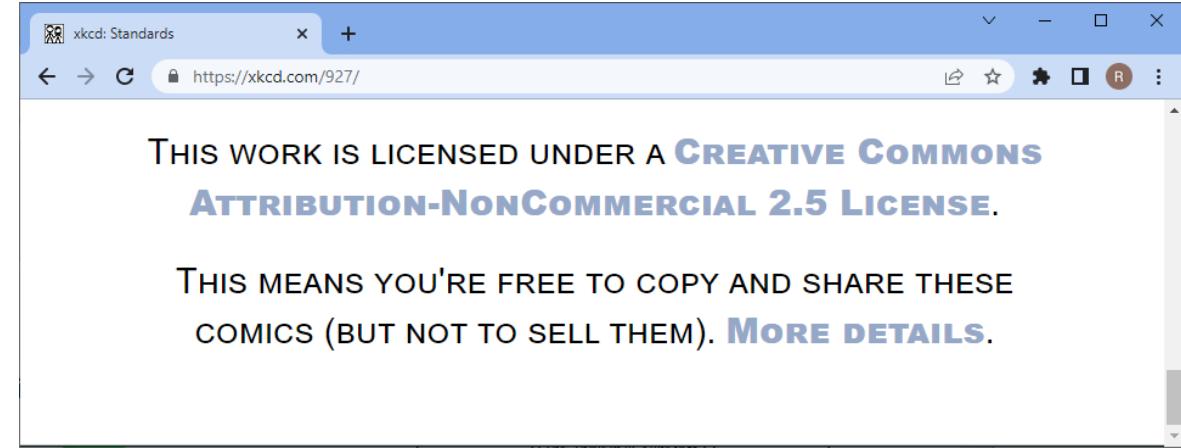
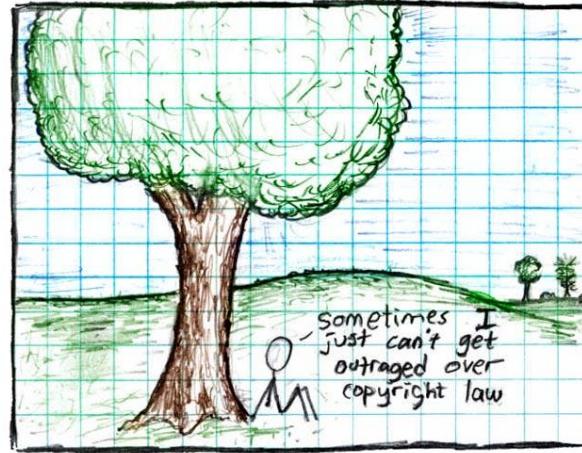
...

License
CC-By Attribution-NonCommercial-NoDerivatives
4.0 International



Quiz

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



Mention
author



Ask the
authors



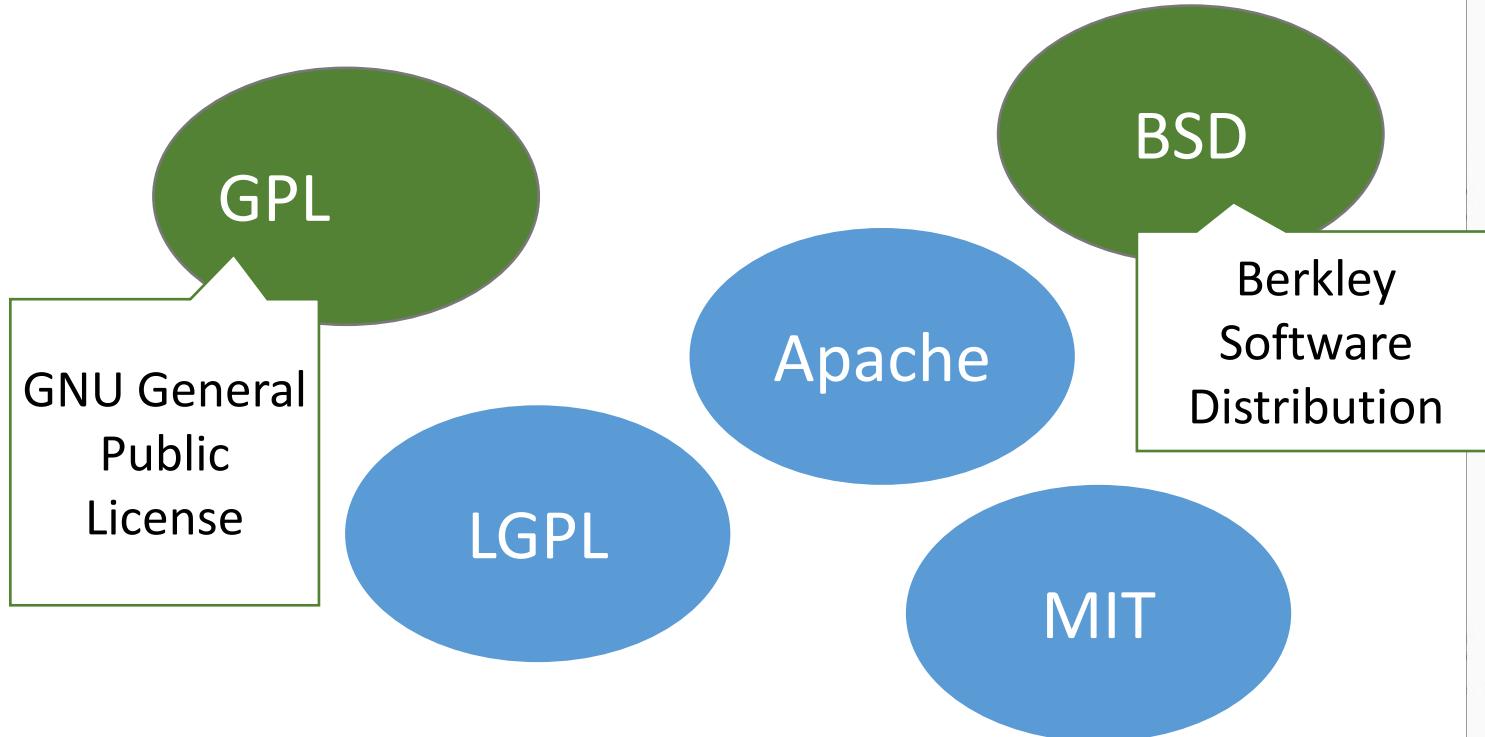
Link to the
license



Copy the
copyright
statement

Licensing Software

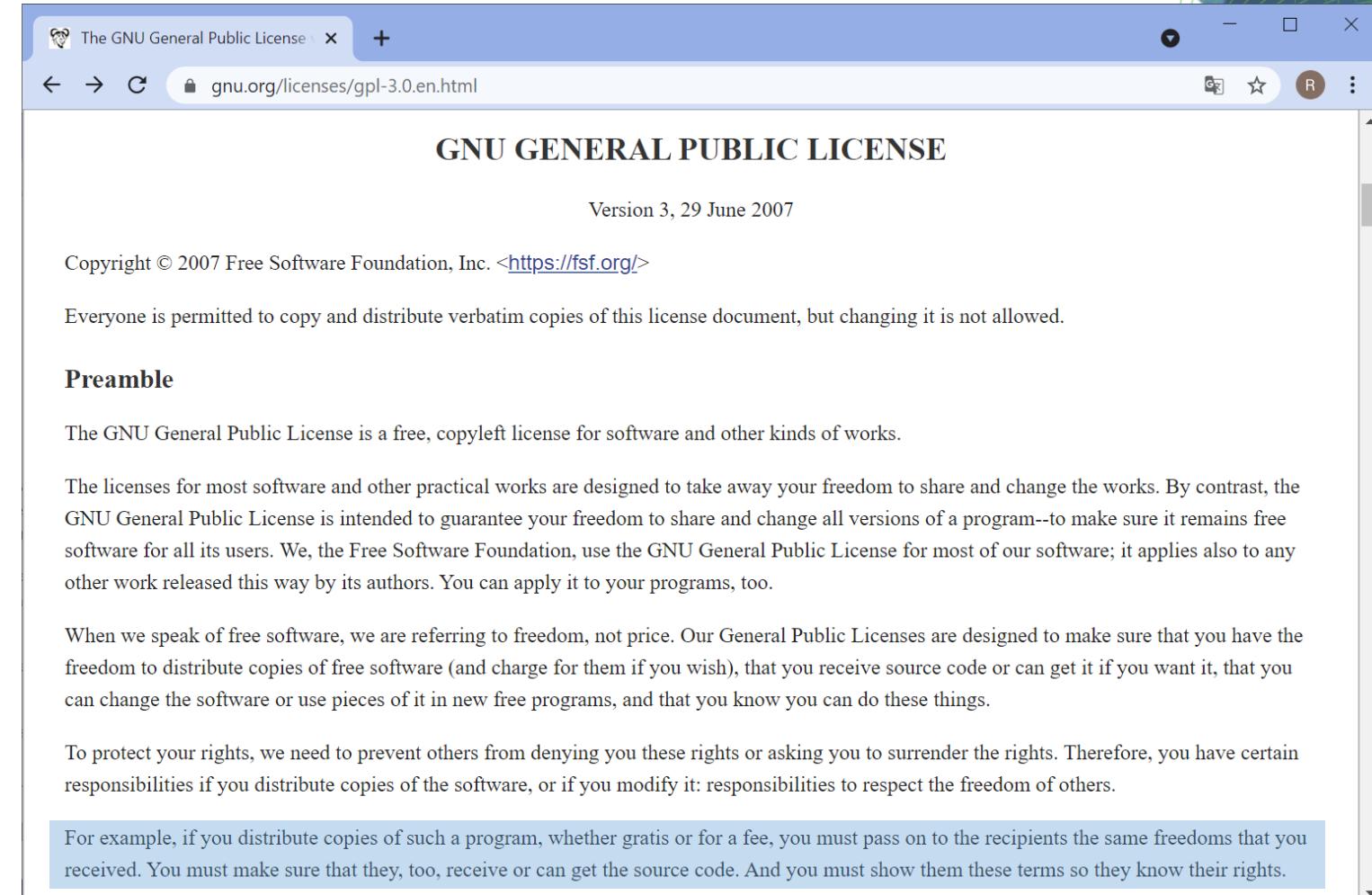
In the software world, other licenses are more popular, historically grown.



Licensing Software: GPL

GPL-derivatives must also
be GPL-licensed

“Restrictive”
licensing



The screenshot shows a web browser window with the title "The GNU General Public License" and the URL "gnu.org/licenses/gpl-3.0.en.html". The page content is the GNU GENERAL PUBLIC LICENSE, Version 3, dated 29 June 2007. It includes the FSF logo, copyright information, a preamble, and sections about freedom to redistribute and modify the software. A note at the bottom states that recipients must pass on the same freedoms.

The GNU GENERAL PUBLIC LICENSE
Version 3, 29 June 2007

Copyright © 2007 Free Software Foundation, Inc. <<https://fsf.org/>>

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

Preamble

The GNU General Public License is a free, copyleft license for software and other kinds of works.

The licenses for most software and other practical works are designed to take away your freedom to share and change the works. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change all versions of a program--to make sure it remains free software for all its users. We, the Free Software Foundation, use the GNU General Public License for most of our software; it applies also to any other work released this way by its authors. You can apply it to your programs, too.

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for them if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs, and that you know you can do these things.

To protect your rights, we need to prevent others from denying you these rights or asking you to surrender the rights. Therefore, you have certain responsibilities if you distribute copies of the software, or if you modify it: responsibilities to respect the freedom of others.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must pass on to the recipients the same freedoms that you received. You must make sure that they, too, receive or can get the source code. And you must show them these terms so they know their rights.

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.

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

} Similar to
CC-BY

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

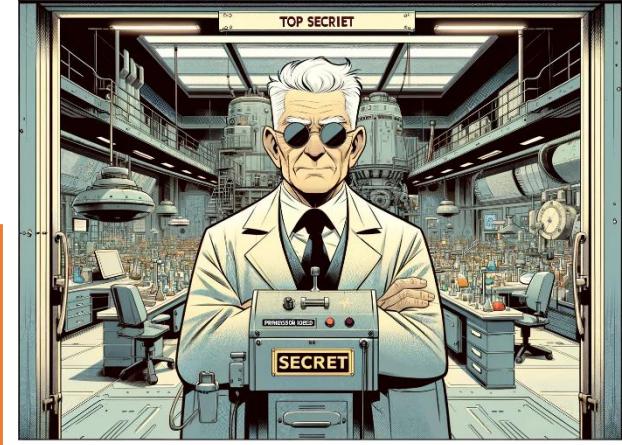
} Similar to
CC-BY

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**

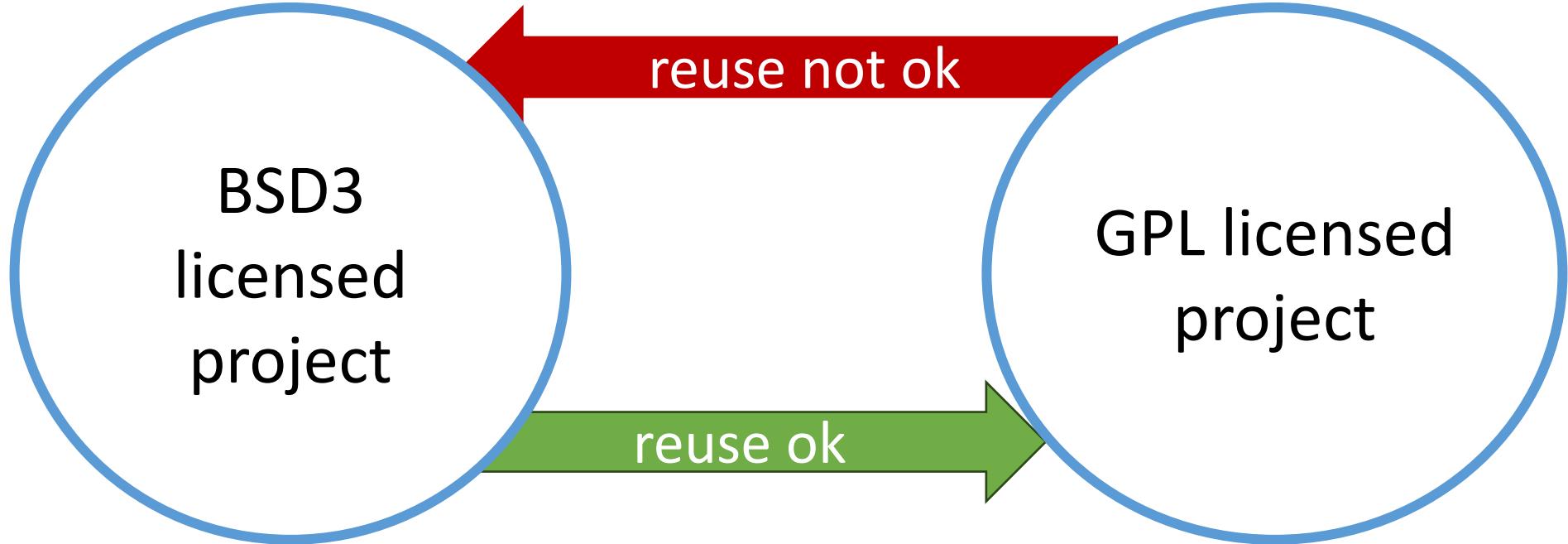
I conclude,
these are
less open in
a sense



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



The screenshot shows a GitHub repository page for `pyrocko/pyrocko`. The repository is described as an official read-only mirror of <https://git.pyrocko.org/pyrocko/pyrocko>. It is a seismology toolkit for Python. The 'About' section includes links to pyrocko.org, tags for `python`, `seismology`, `geophysics`, `earthquake`, and `geophysical-inversions`, and sections for `Readme`, `GPL-3.0 license`, `Cite this repository`, `Activity`, and `Custom properties`. A large orange arrow points from the 'No' button on the left towards the 'About' section on the right.

Author	Commit Message	Date
emolch	gnss_unr_campaign example: update url	deee904 · 27 days ago
	.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

Quiz

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

Yes



No



The screenshot shows a GitHub repository page for `pyrocko/pyrocko`. The repository is described as an official read-only mirror of <https://git.pyrocko.org/pyrocko/pyrocko>, a seismology toolkit for Python. The page lists several recent commits by various contributors, including updates to examples, dockerfiles, and documentation. On the right side, there are links for the repository's website (`pyrocko.org`), tags like `python`, `seismology`, and `geophysics`, and other repository details. An orange arrow points from the 'No' hand icon to the URL <https://github.com/pyrocko/pyrocko> at the bottom of the page.

Author	Commit Message	Date
emolch	gnss_unr_campaign example: update url	deee904 · 27 days ago
	.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

Quiz

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

Yes



No



A screenshot of a web browser displaying the GitHub repository page for `fatiando/verde`. The repository is public and has 38 issues, 6 pull requests, and 333 commits. The commit history shows various updates to the .github, data, doc, env, paper, and verde directories, along with Sphinx and NumPy replacements, acknowledgements, and documentation fixes. On the right side of the screen, there is a sidebar with repository details: "About" (Processing and gridding spatial data, machine-learning style), a link to the website (www.fatiando.org/verde), and a list of tags: python, machine-learning, interpolation, geospatial, geoscience, python3, geophysics, scipy, earth-science, scipy-stack, and fatiando-a-terra. Below the sidebar, there are links for Readme, BSD-3-Clause license, Code of conduct, Activity, and Custom properties. An orange arrow points from the "Activity" link towards the "No" hand icon.

Commit	Message	Date
leouieda	Add a link to the Fatiando F...	4a961f6 · 2 months ago
.github	Bump pypa/gh-action-pypi-public...	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 numpy...	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

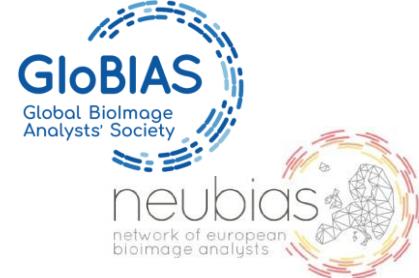
Communities &
platforms



NFDI4
BIOIMAGE



image.sc



BIDS team UL



Lea Gihlein, Lea Kabjesz, Mara Lampert

Funding



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.

Chan
Zuckerberg
Initiative



Deutsche
Forschungsgemeinschaft

Some Figures were
generated using



DALL-E