



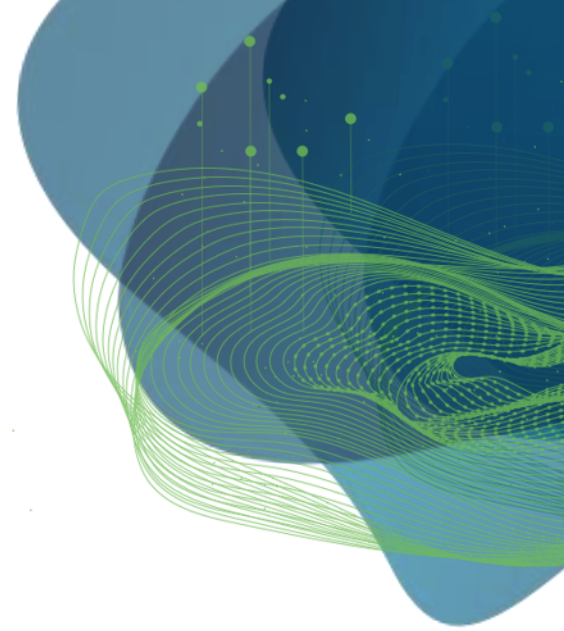
ScaDS.AI

DRESDEN LEIPZIG

CENTER FOR SCALABLE DATA ANALYTICS
AND ARTIFICIAL INTELLIGENCE

Research Data 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.

Recap quiz

- We write good documentation to enabling others to do an experiment. This is good for ...

Repeatability



Reproducibility



Replicability



Reliability



Recap quiz

- “Resolution” in microscopy imaging describes

Camera
pixel size



Screen
pixel size



Size of
differentiable
objects



Objective
magnification



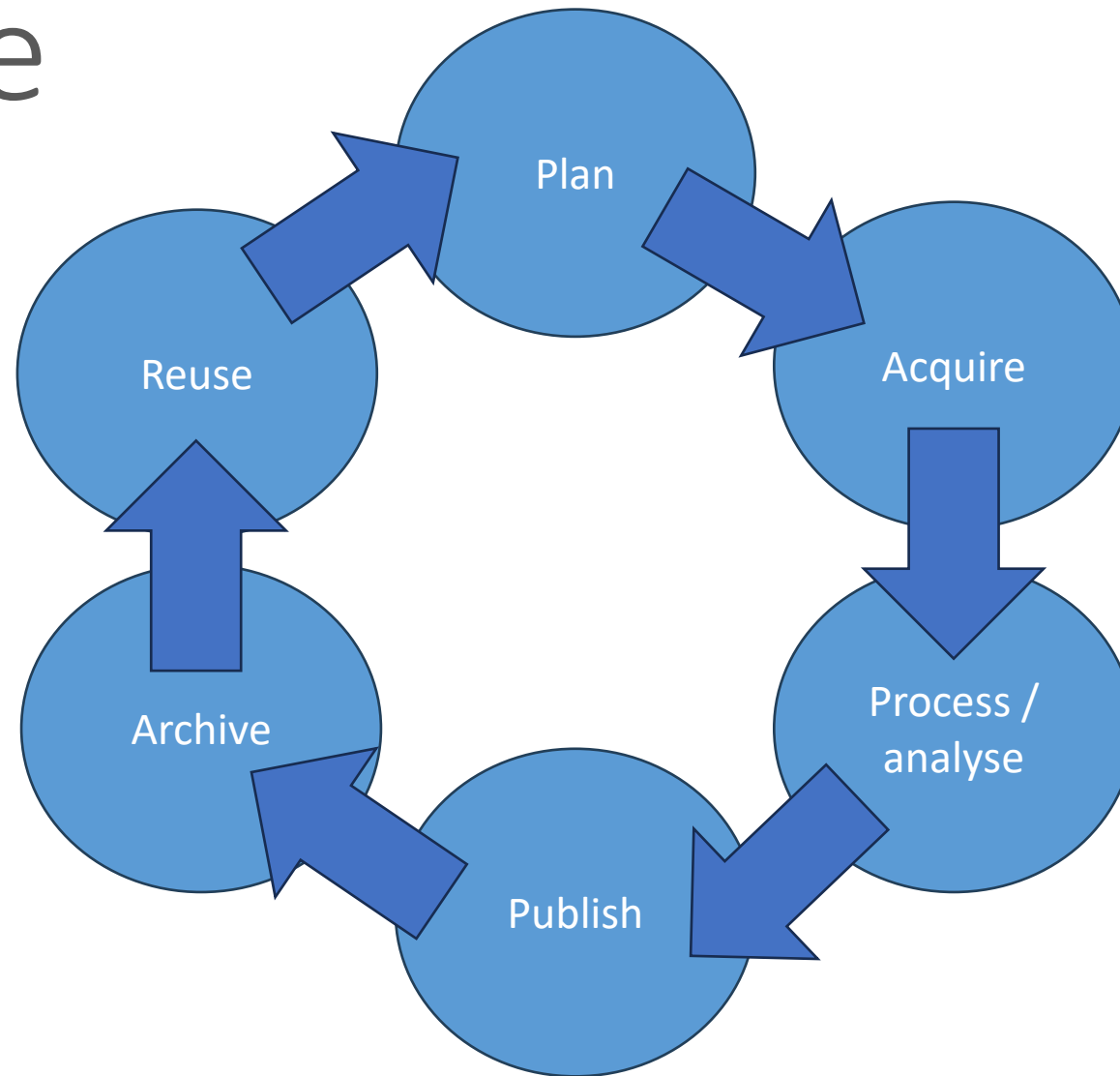
Research Data Management (RDM)

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



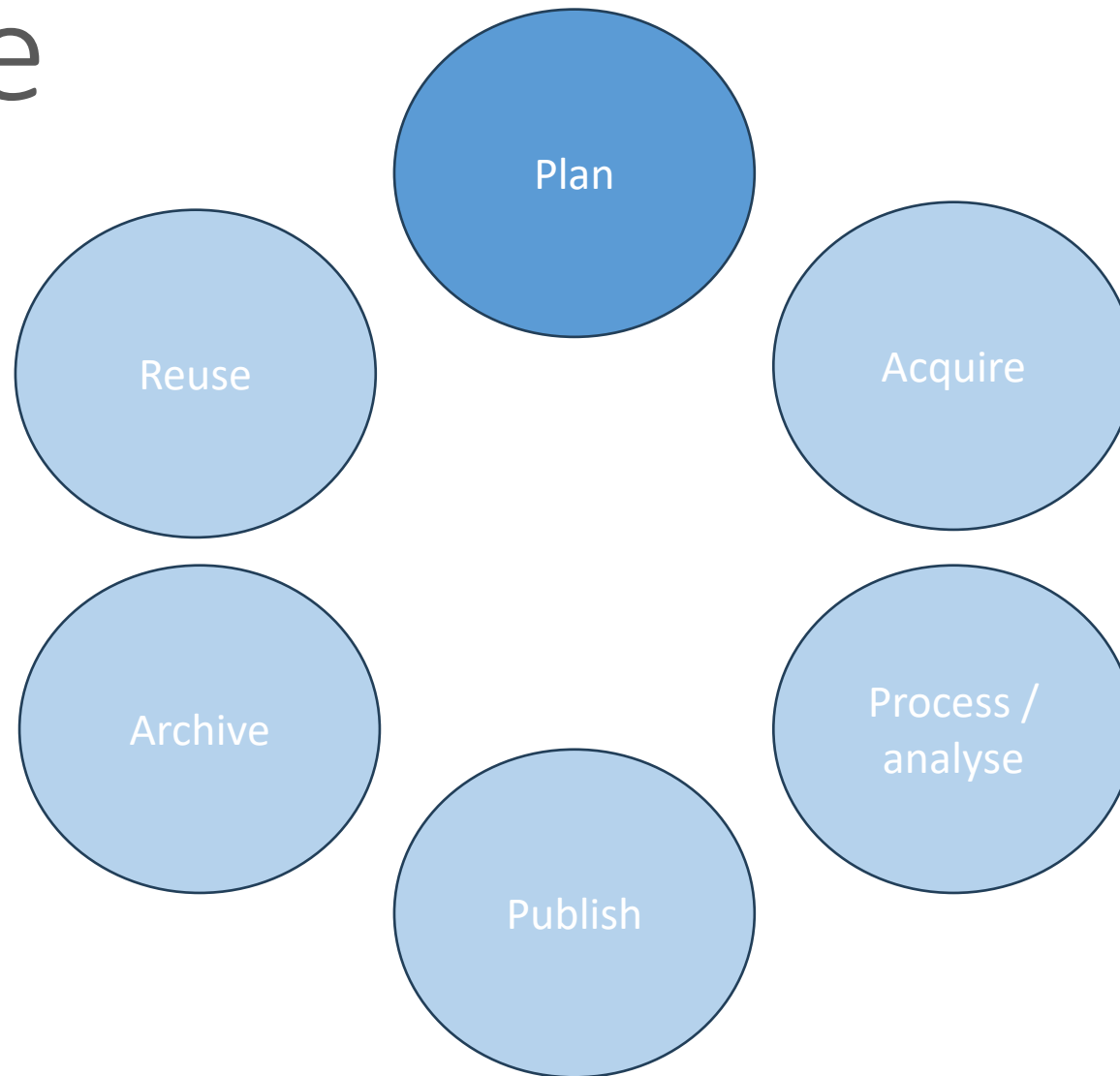
RDM Life Cycle

- Processes are ideally cyclic



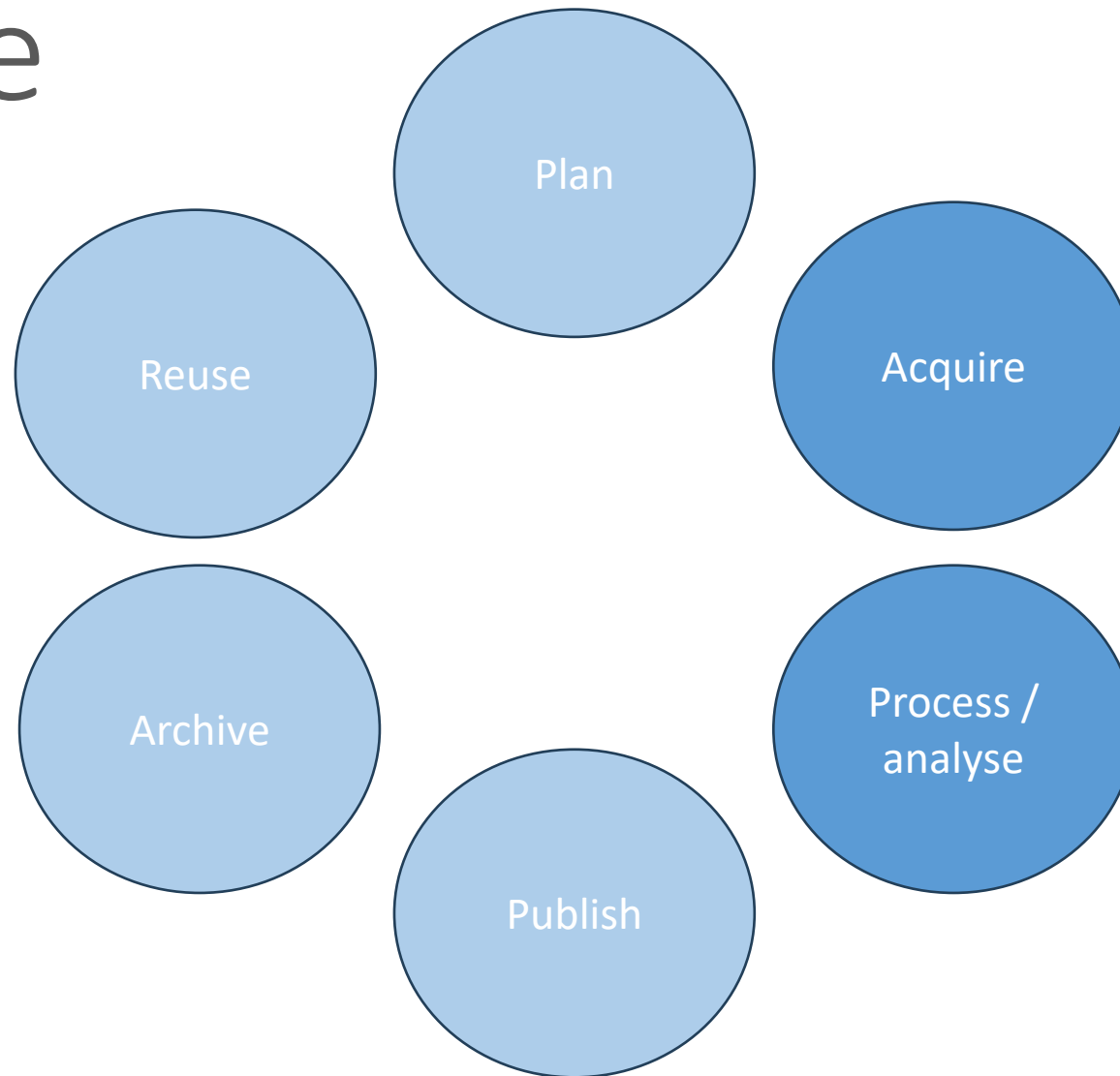
RDM Life Cycle

- Cost
- Benefit
- Quality
- Strategic decisions



RDM Life Cycle

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



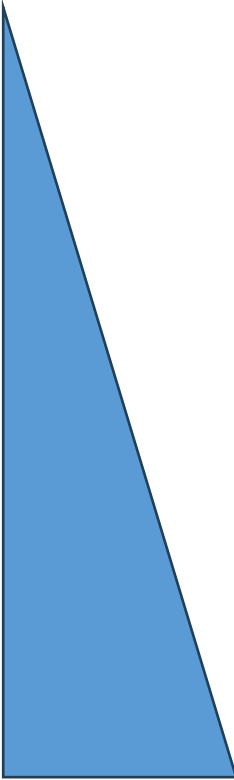
Types of data

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



Types of data

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

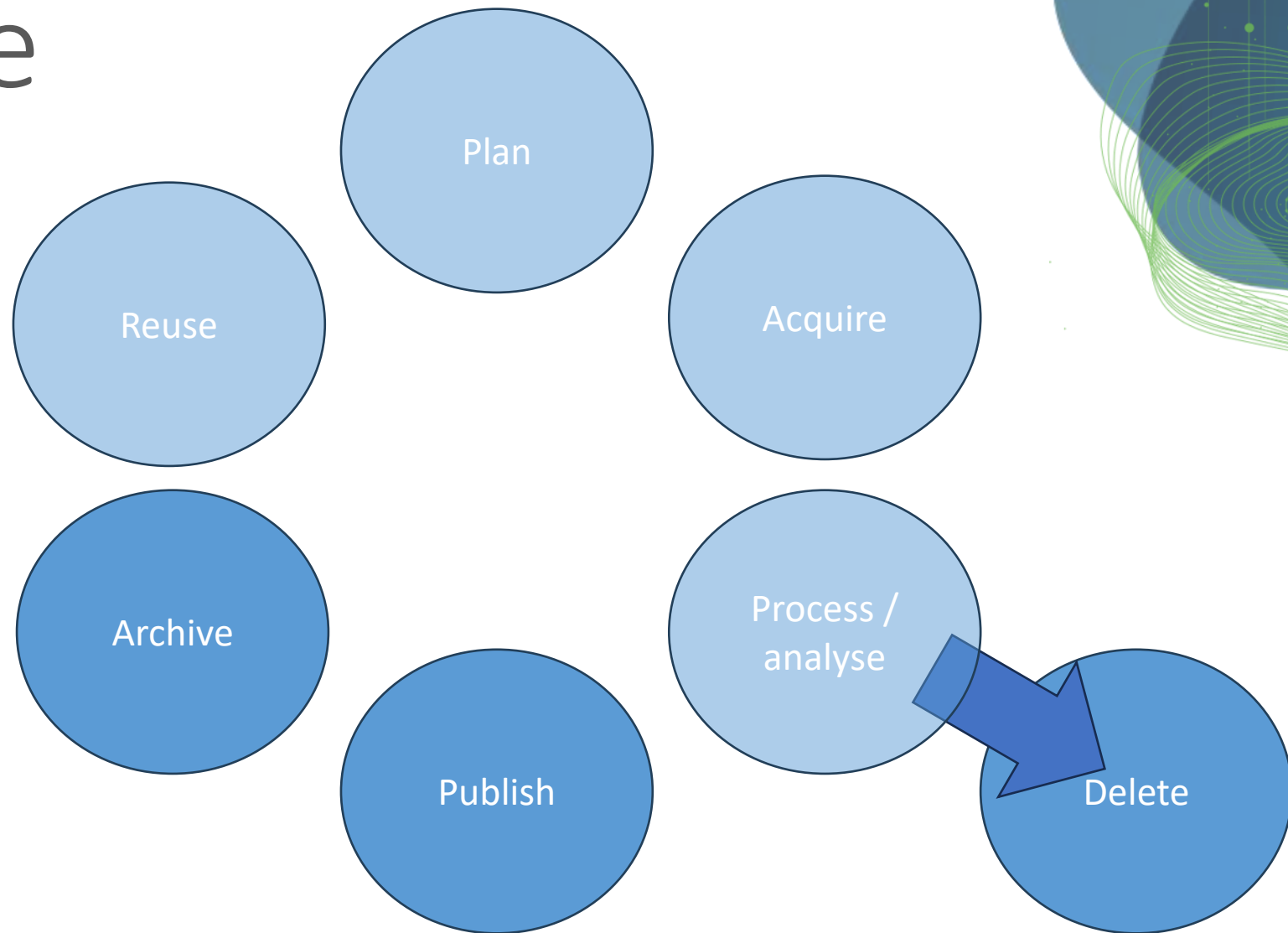


In need of
protection
(schutzbedürftig)



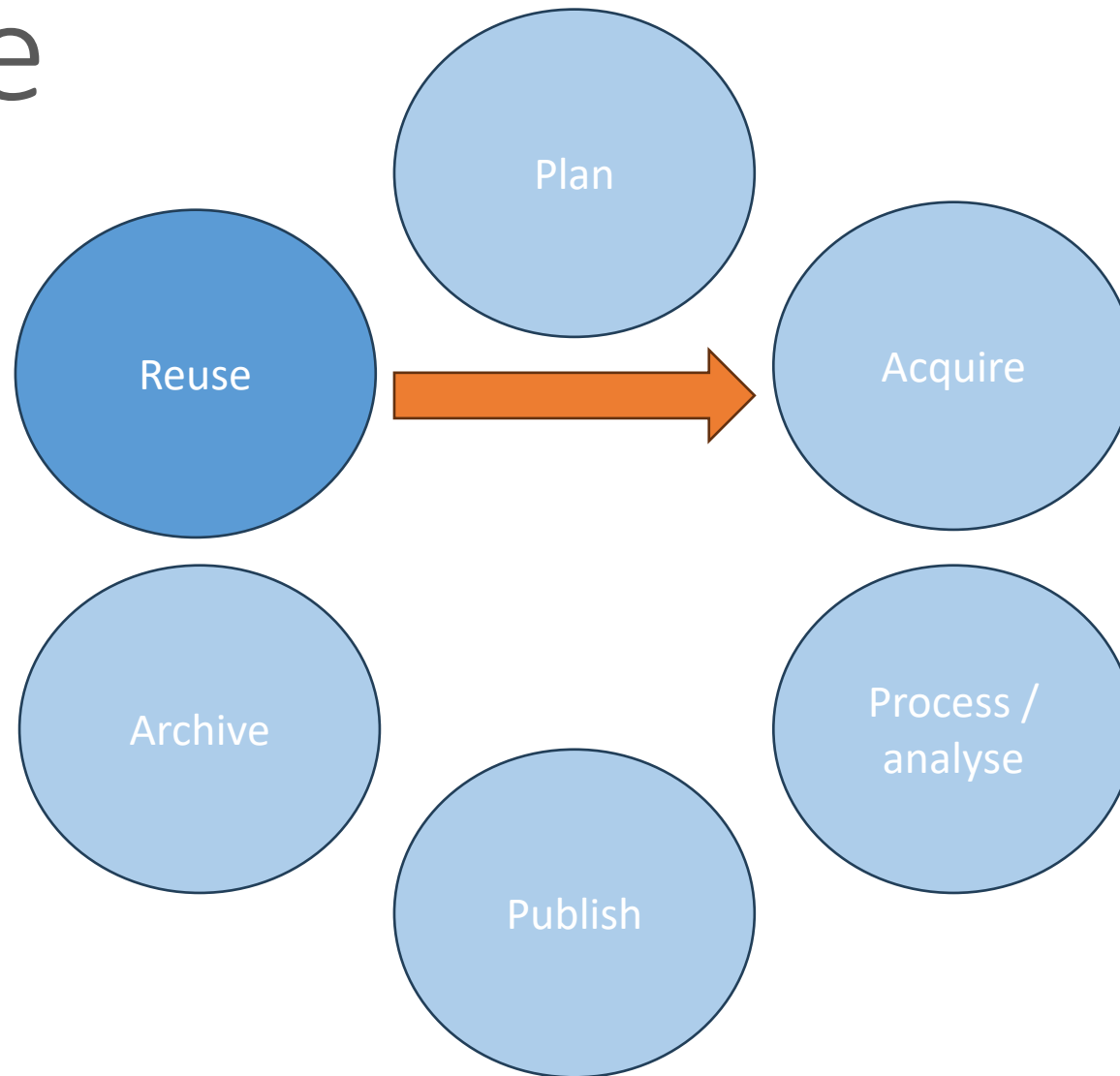
RDM Life Cycle

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



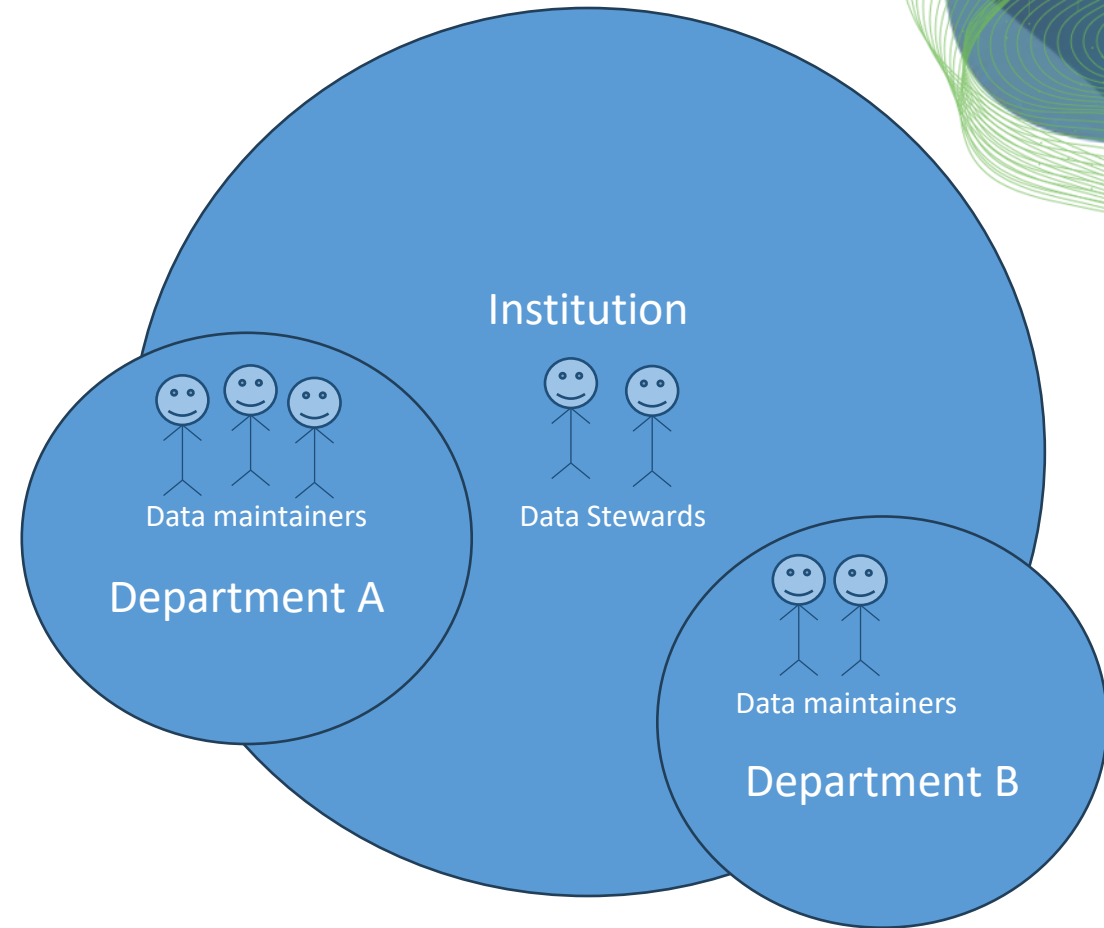
RDM Life Cycle

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



What is **good** RDM?

- Clearly defined responsibilities and processes (Governance)
 - Data Management Plan (DMP)
- Communication of goals, metrics, responsibilities, processes
- Dedicated personnel
 - “Data maintainers”
 - IT infrastructure maintainers
- Expert consultants
 - “Data stewards”



Roles != Job profiles

Domain specialist

- Focuses on scientific question, often related to the physical world
- Requires sound insights and sustainable solutions
- Examples:
 - Biologist
 - Geologist
 - City planner

Data analyst

- Focuses on methods for data processing / visualization
- Gains sound insights
- Examples:
 - Statistician
 - Bioinformatician
 - Data Scientist

IT specialist

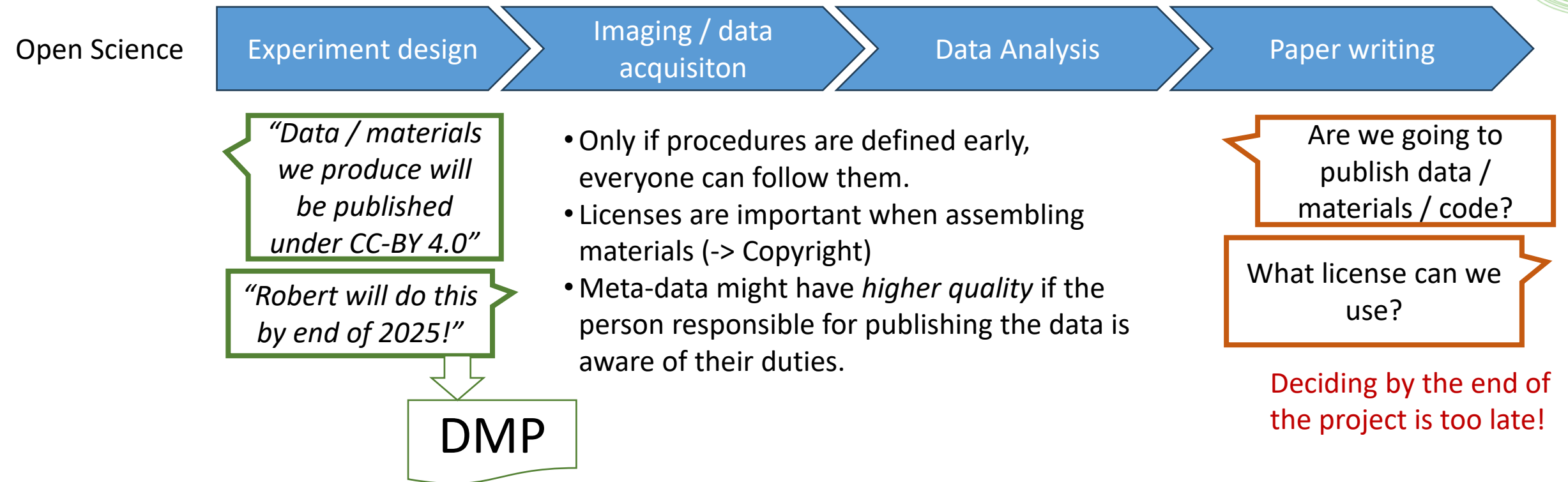
- Focuses on IT infrastructure
 - Hardware
 - Software
- Builds sustainable solutions
- Examples:
 - Computer scientist
 - IT specialist

Data Management Plans (DMPs)

- Describes the IS-state of a data environment
 - Which data is acquired / processed?
(content, format, amount)
 - What meta-data is collected?
 - Which quality standards are targeted?
 - How is data saved, archived, backed-up, shared, published...?
 - Who is responsible for what?
 - Roles, job-profiles
 - What does this all cost?
(IT infrastructure + human resources)

Data Management Plans (DMPs)

- Define responsibilities and procedures early!



Quiz

- Regularly copying files to a remote place is ...

Archiving



Backup



Quiz

- Data Scientists is a ...

Role



Job profile



Quiz

- Data Steward is a ...

Role



Job profile



Sharing & licensing

Robert Haase

Code

Slides

Text

Data

...

GEFÖRDERT VOM



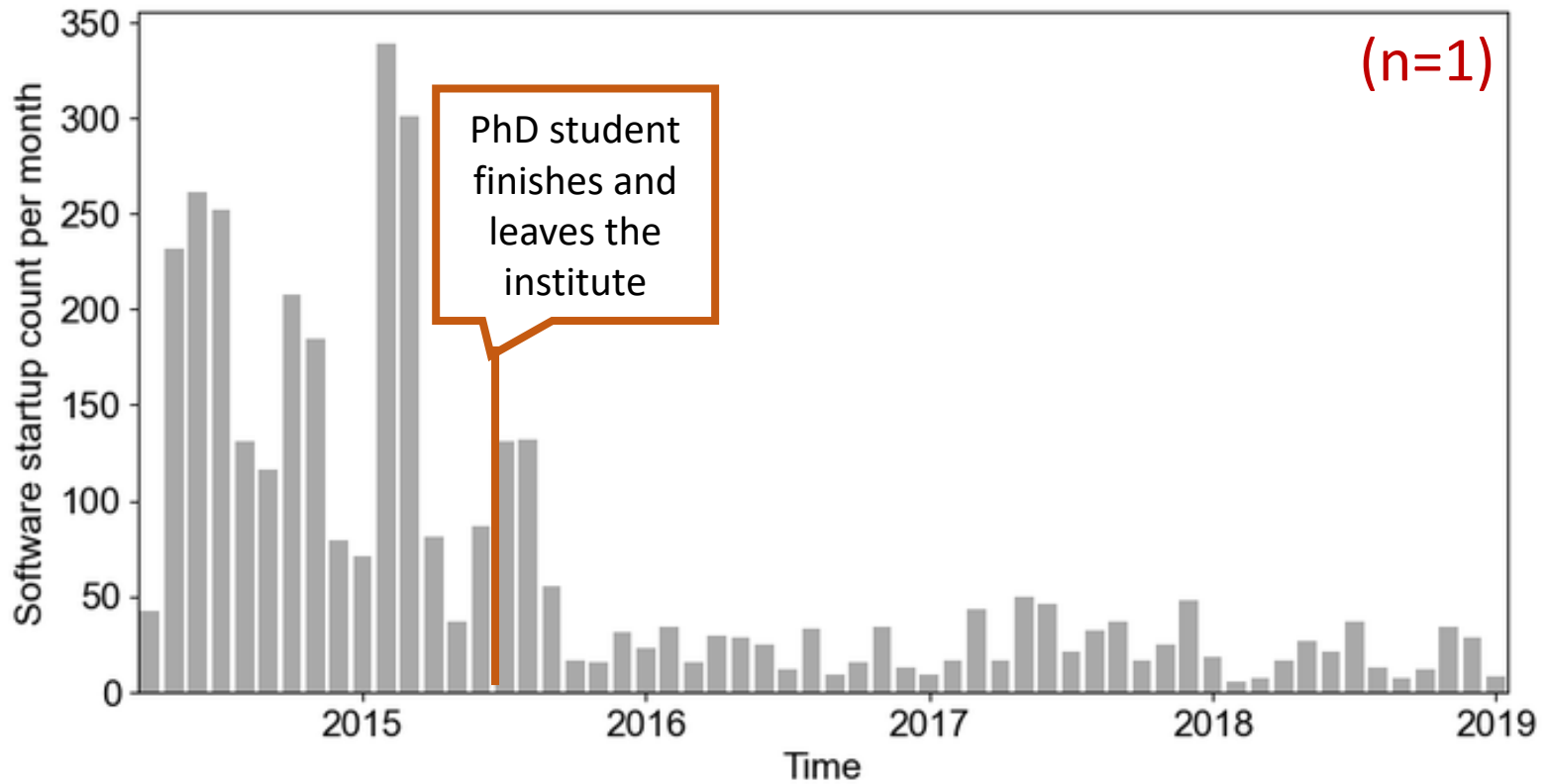
Bundesministerium
für Bildung
und Forschung



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.

Sustainability of *my* contribution to science

- What happens to research software once the PhD student leaves the institute / field?



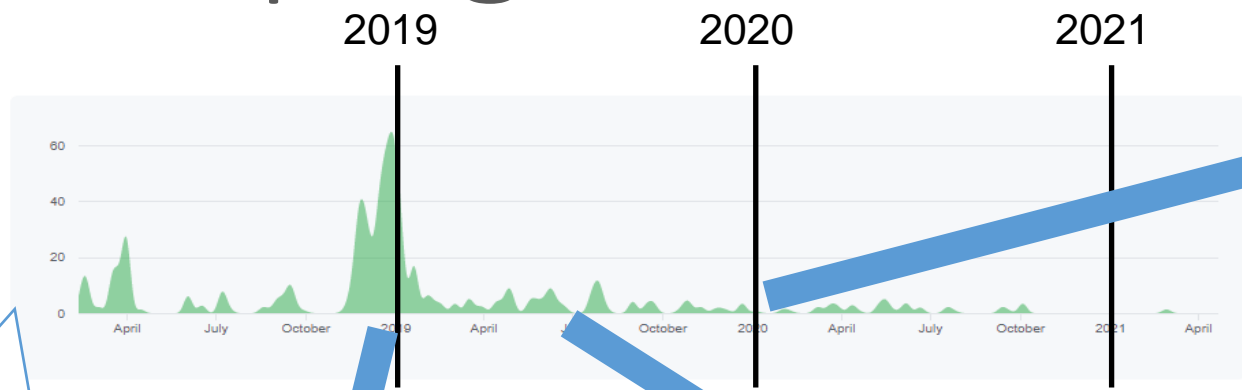
Developing software in the open



Nov. 2017: I took over microscopy control software from Loic and "found" some GPU-accelerated image processing in there



Loic A. Royer (CZ Biohub) @loicaroyer



forum.image.sc/t/opencl-gpu-based-image-processing-in-imagej-macro/21286

Community Partners

OpenCL/GPU based image processing in ImageJ macro

Announcements imagej, fiji, macro, clij, gpu

Robert Haase haesleinhuepf clij maintainer 2 Dec '18

Dear friends of GPU based image processing, dear early adopters,

I recently put some efforts into making GPU-based image processing in ImageJ macro run. It has the potential to save us massive amounts of processing time. GPU programming in macro looks like this:

bioRxiv THE PREPRINT SERVER FOR BIOLOGY

CLIJ: Enabling GPU-accelerated image processing in Fiji

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

nature > nature methods > correspondence > article

nature methods

Correspondence | Published: 18 November 2019

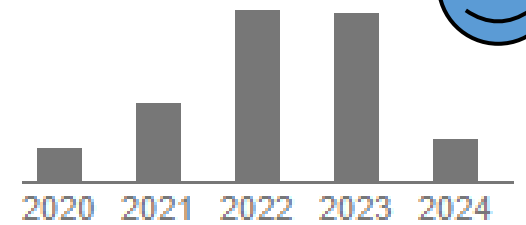
CLIJ: GPU-accelerated image processing for everyone

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

Today: 134 citations (Google scholar, 2024-03-18)

Zitiert von: 134



Scientific culture

Public access to research results -> Reusability



Guidelines for Safeguarding Good Research Practice

Code of Conduct

“Kodex”

Guideline 13: Providing public access to research results

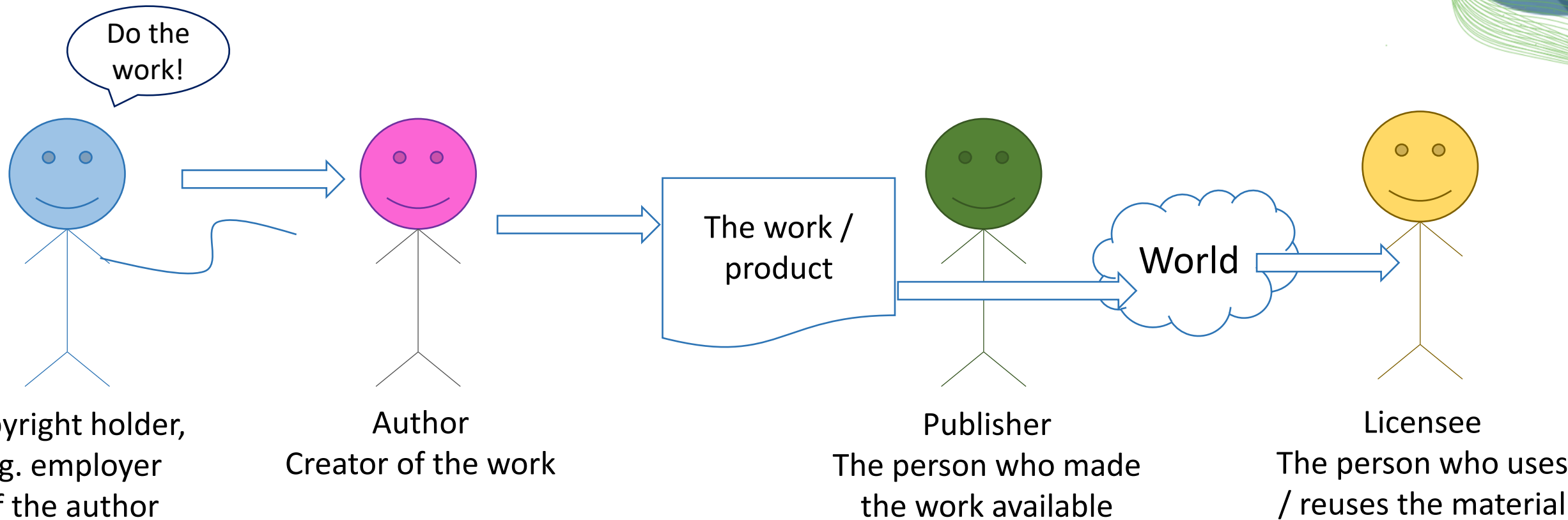
► As a rule, researchers **make all results available** as part of scientific/academic discourse. In specific cases, however, there may be reasons not to make results publicly available (in the narrower sense of publication, but also in a broader sense through other communication channels); this decision 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

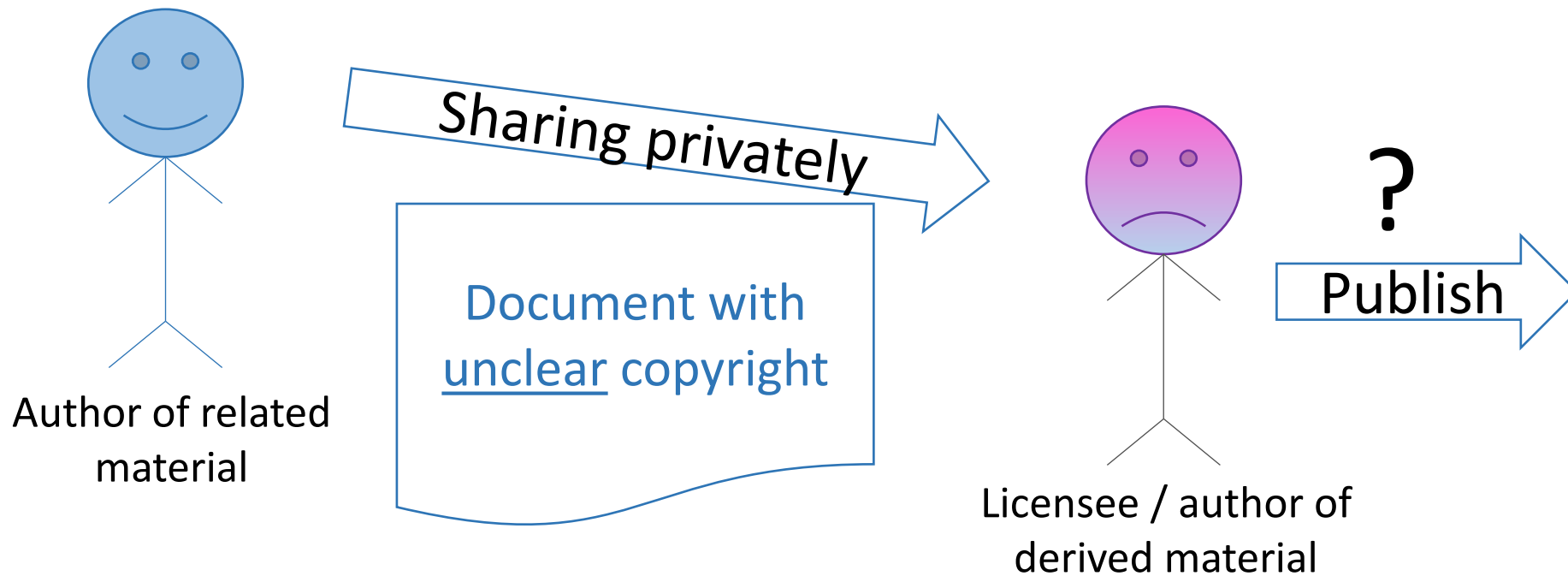
Am I allowed to publish my stuff?

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



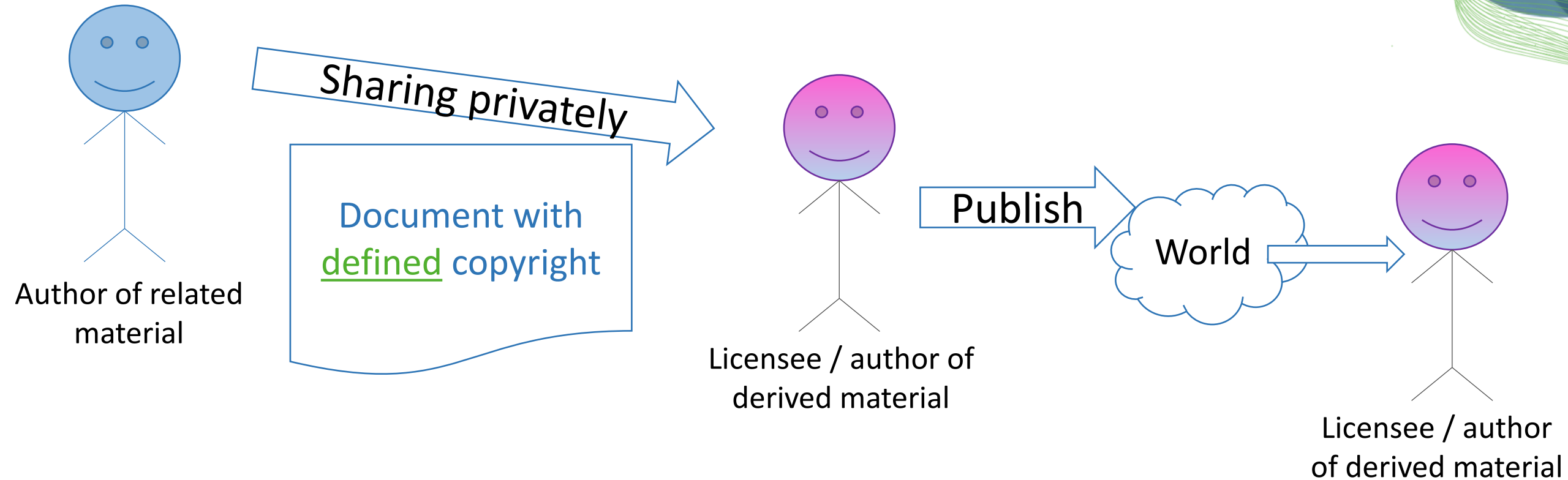
Am I allowed to publish my stuff?

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



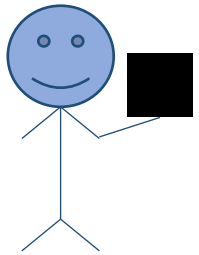
Am I allowed to publish my stuff?

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



Openness of software / projects

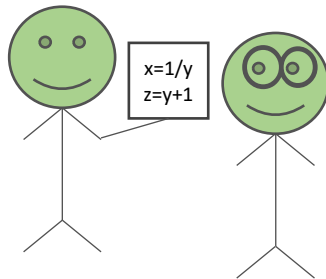
Closed source



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

Hardware device drivers

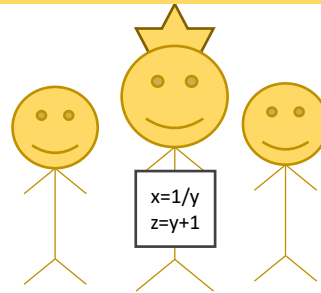
Open source



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

Custom image analysis scripts

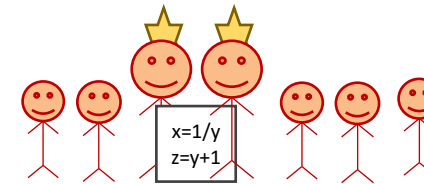
Benevolent dictatorship



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

TrackMate, SNT, MorpholibJ, CLIJ

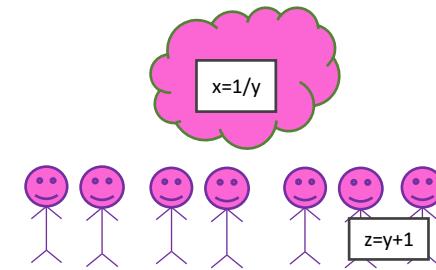
Community driven



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

scikit-image, scipy, OpenCL

Openly extensible



- Openly extensible; without maintainers involved
- Partially community driven

ImageJ, Python, numpy

Quiz

- What is the role of Github in the context of publishing open-source code?

Copyright holder



Author



Publisher



Licensee



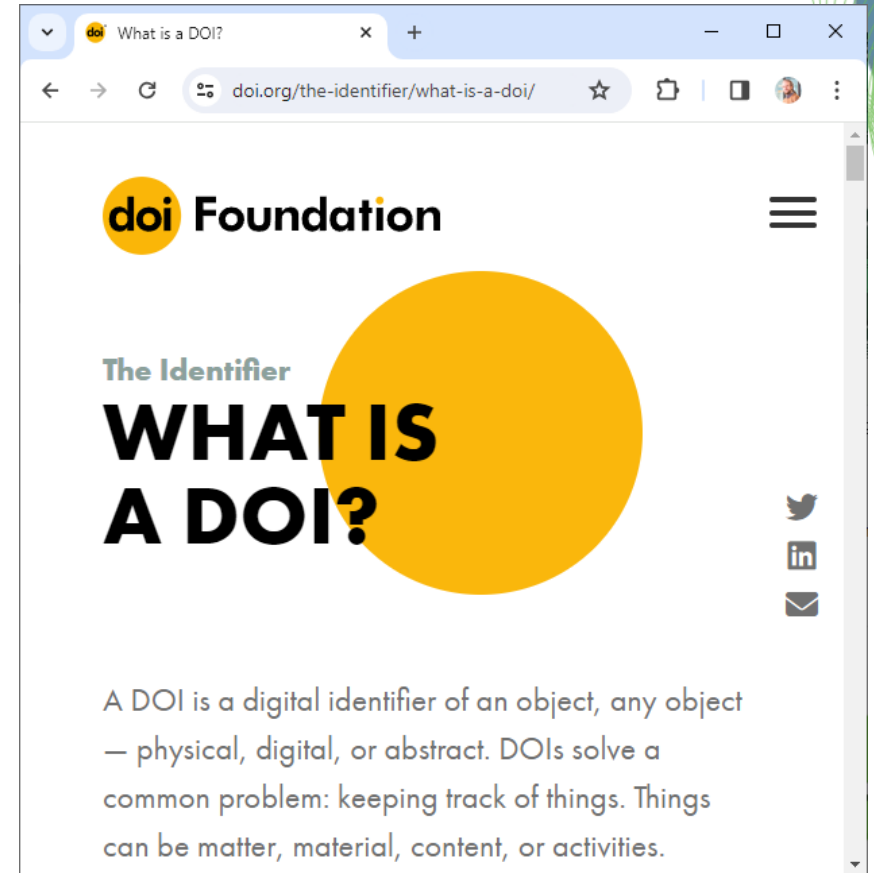
Standard for sharing: The FAIR-principles

- Findable
- Accessible
- Interoperable
- Reusable



The FAIR-principles

- Findable
- F1. (Meta)data are assigned a globally unique and persistent identifier
 - Universal Resource Identifier (URI)
 - Digital Object Identifier (DOI)
- F2. Data are described with rich metadata (defined by R1 below)
- F3. Metadata clearly and explicitly include the identifier of the data they describe
- F4. (Meta)data are registered or indexed in a searchable resource



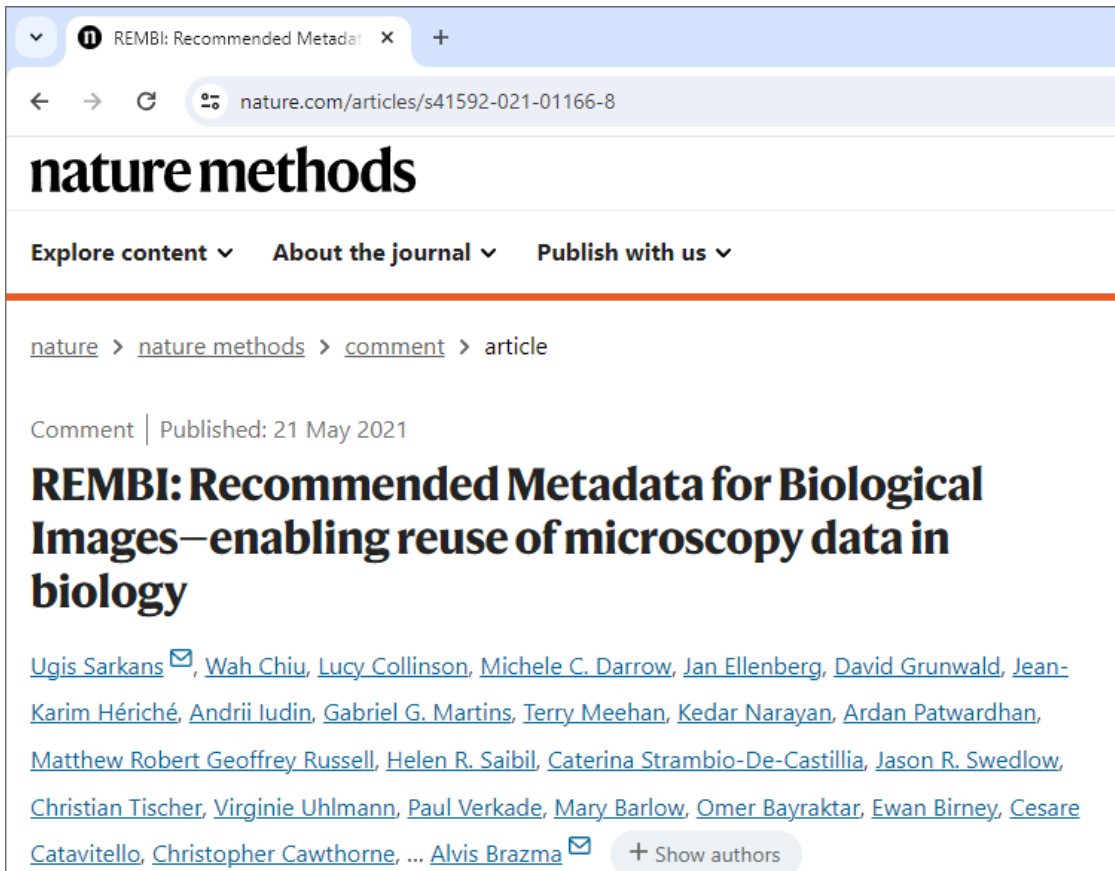
Meta data

- Generic
 - Author
 - Usage license
 - Creation date
- Field-specific (microscopy)
 - Exposure time
 - Wavelength (colour)
 - Microscope type/vendor

Goals:
FAIR+
Reproducibility

REMBI: Recommended Metadata for Biological Images—enabling reuse of microscopy data in biology

• Read more:



REMBI: Recommended Metadata for Biological Images—enabling reuse of microscopy data in biology

nature.com/articles/s41592-021-01166-8



nature methods

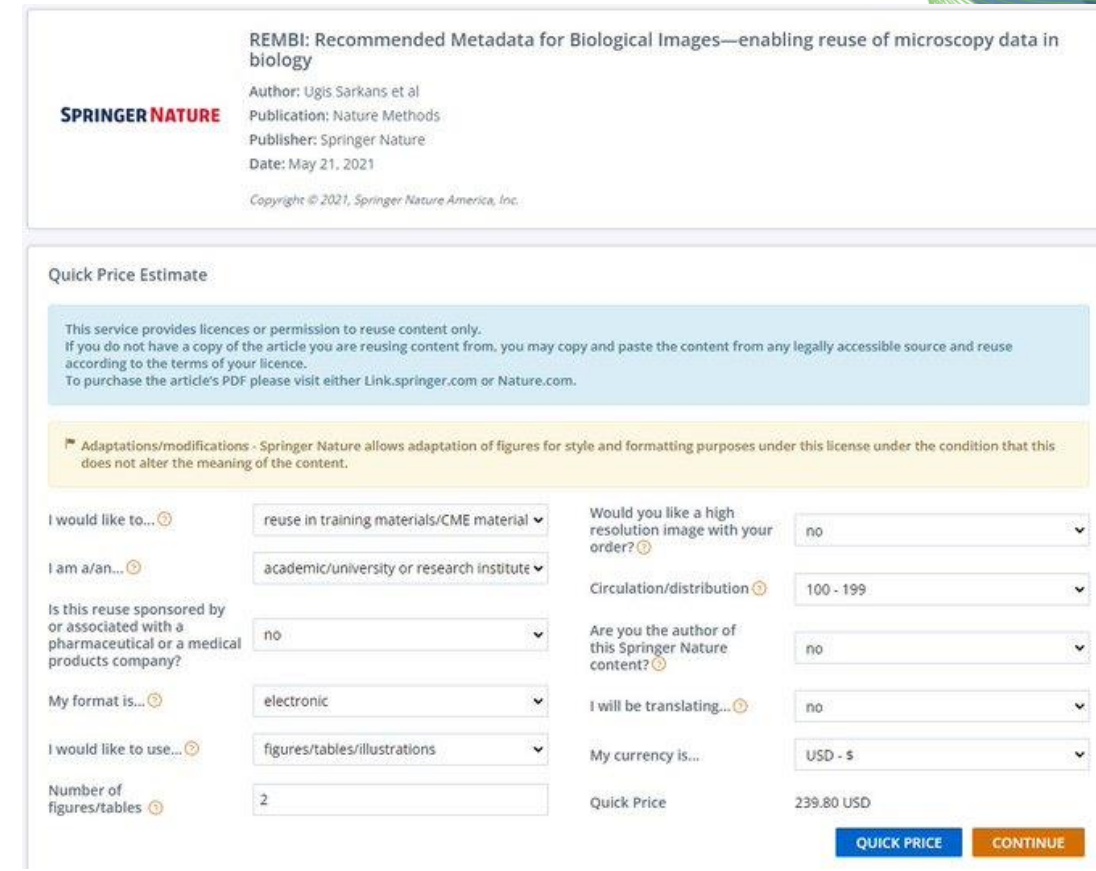
Explore content ▾ About the journal ▾ Publish with us ▾

nature > nature methods > comment > article

Comment | Published: 21 May 2021

REMBI: Recommended Metadata for Biological Images—enabling reuse of microscopy data in biology

[Ugis Sarkans](#) , [Wah Chiu](#), [Lucy Collinson](#), [Michele C. Darrow](#), [Jan Ellenberg](#), [David Grunwald](#), [Jean-Karim Hériché](#), [Andrii Iudin](#), [Gabriel G. Martins](#), [Terry Meehan](#), [Kedar Narayan](#), [Ardan Patwardhan](#), [Matthew Robert Geoffrey Russell](#), [Helen R. Saibil](#), [Caterina Strambio-De-Castillia](#), [Jason R. Swedlow](#), [Christian Tischer](#), [Virginie Uhlmann](#), [Paul Verkade](#), [Mary Barlow](#), [Omer Bayraktar](#), [Ewan Birney](#), [Cesare Catavittello](#), [Christopher Cawthorne](#), ... [Alvis Brazma](#)  [+ Show authors](#)








REMBI: Recommended Metadata for Biological Images—enabling reuse of microscopy data in biology

SPRINGER NATURE Author: Ugis Sarkans et al
Publication: Nature Methods
Publisher: Springer Nature
Date: May 21, 2021
Copyright © 2021, Springer Nature America, Inc.

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 ▾	Would you like a high resolution image with your order? 	no ▾
I am a/an...	academic/university or research institute ▾	Circulation/distribution 	100 - 199 ▾
Is this reuse sponsored by or associated with a pharmaceutical or a medical products company?	no ▾	Are you the author of this Springer Nature content? 	no ▾
My format is...	electronic ▾	I will be translating... 	no ▾
I would like to use...	figures/tables/illustrations ▾	My currency is...	USD - \$ ▾
Number of figures/tables 	2	Quick Price	239.80 USD

[QUICK PRICE](#) [CONTINUE](#)

Digital Object Identifiers (DOI)

- DOIs / URIs always point at the same data
- DOIs are centrally registers, URIs not
- Unified Resource Locators (URLs) may point at different things

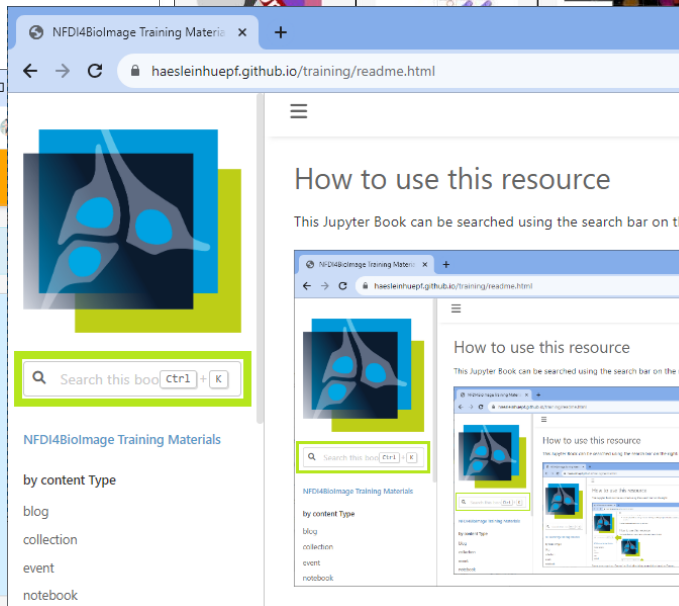
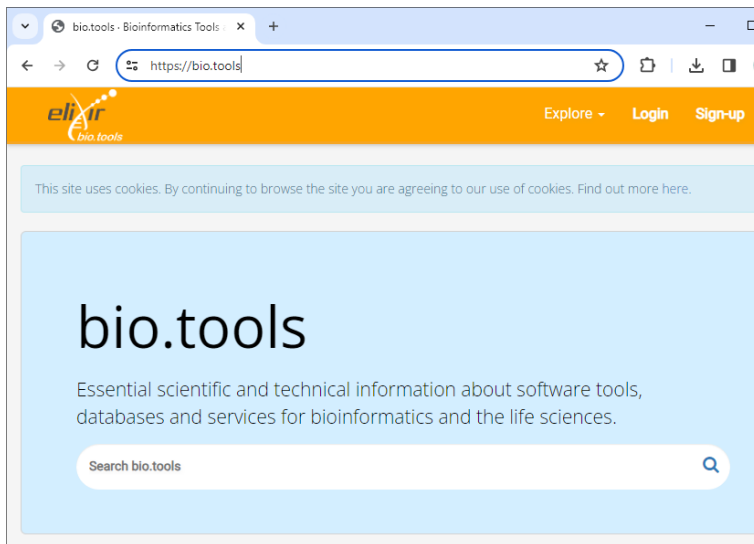
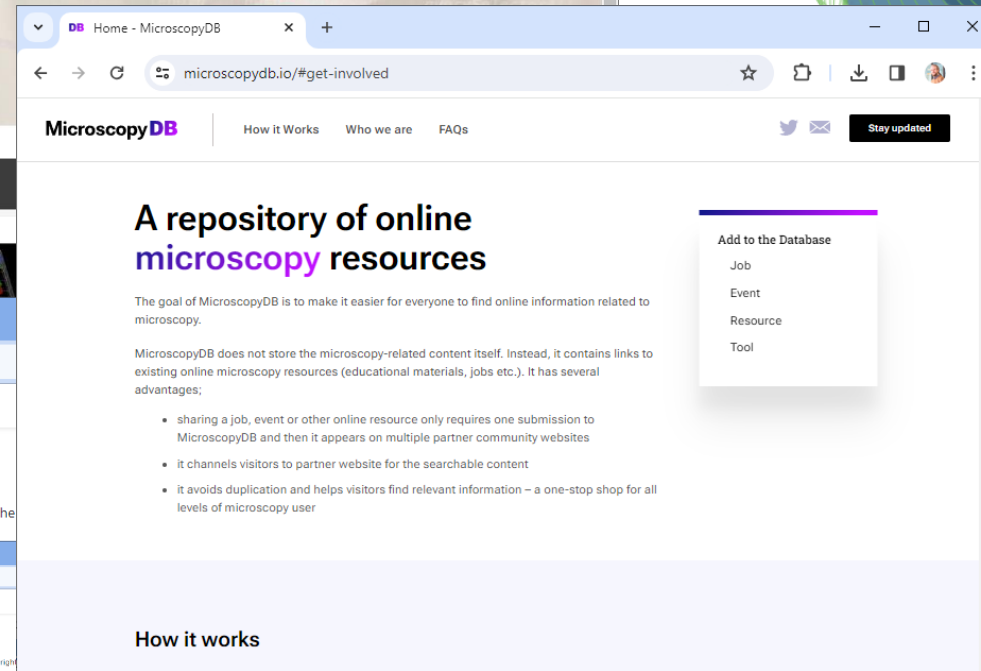
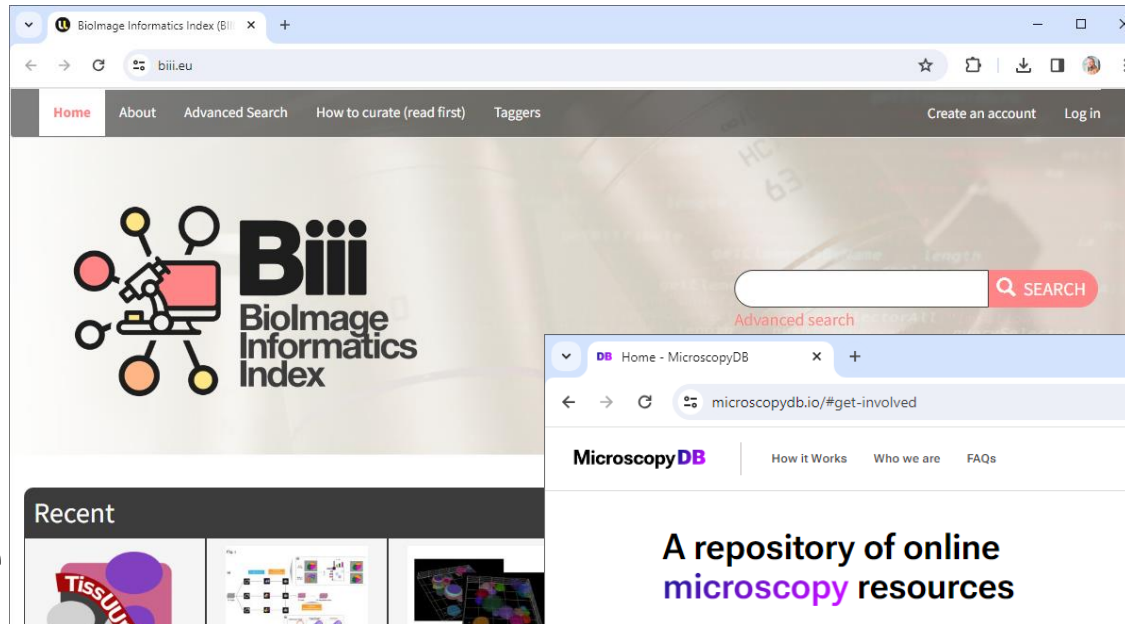
Feld	Wert
Ansprechpartner	Verkehrs- und Tiefbauamt, Stadt Leipzig
E-Mail	vta@leipzig.de
Verwaltungsebene	kommunale Ebene
Gemeindename	Leipzig, Stadt
Ausgestellt	2021-08-20
Aktualisiert	2024-01-17

Gemeindename	Leipzig, Stadt
Ausgestellt	2021-08-20
Aktualisiert	2024-01-17

This no DOI, no URI, it's a URL

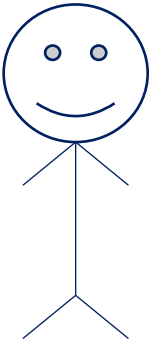
Indexing

- Make sure your materials are listed in public search indices
- Do not trust google to make your stuff findable



Incentives: Findability

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

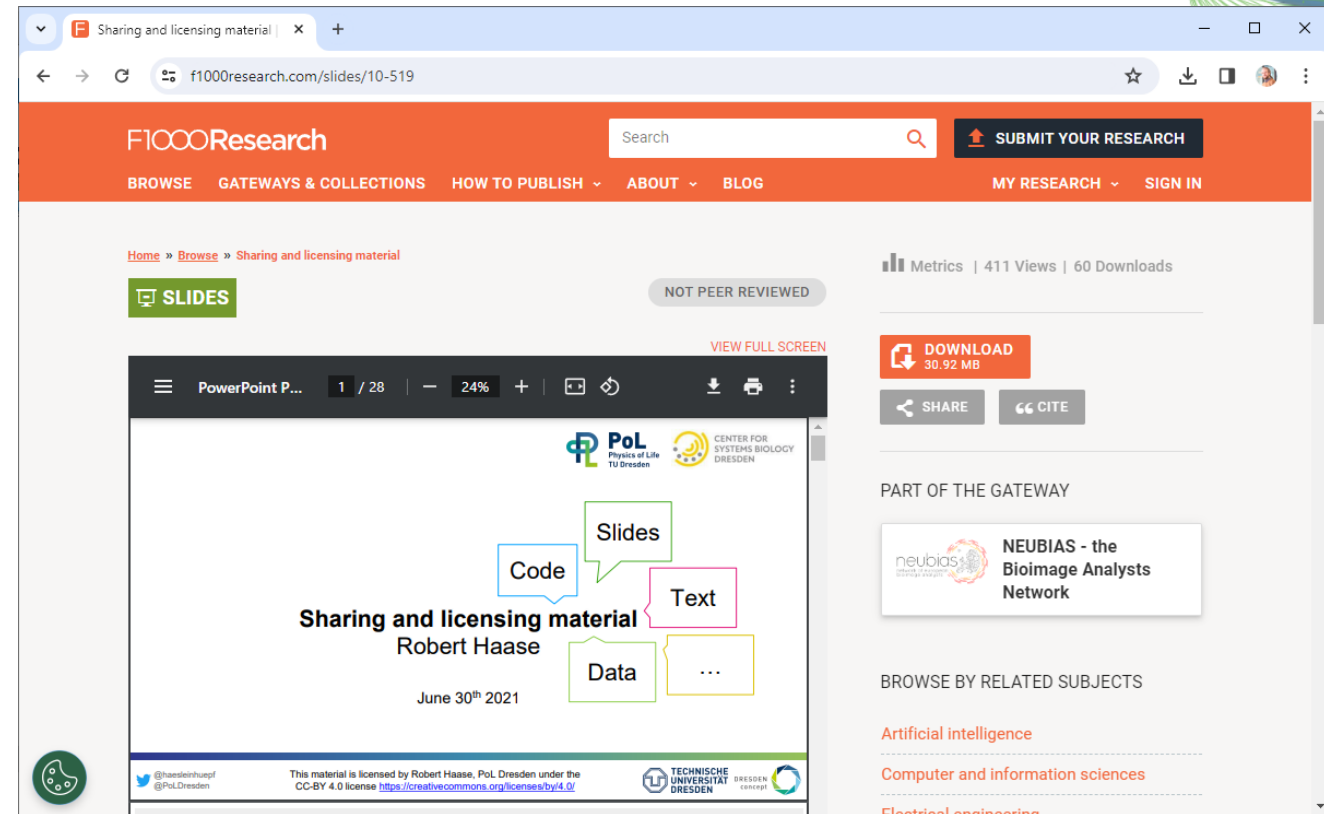


You remember that talk you gave in 2021?

Where are the slides?



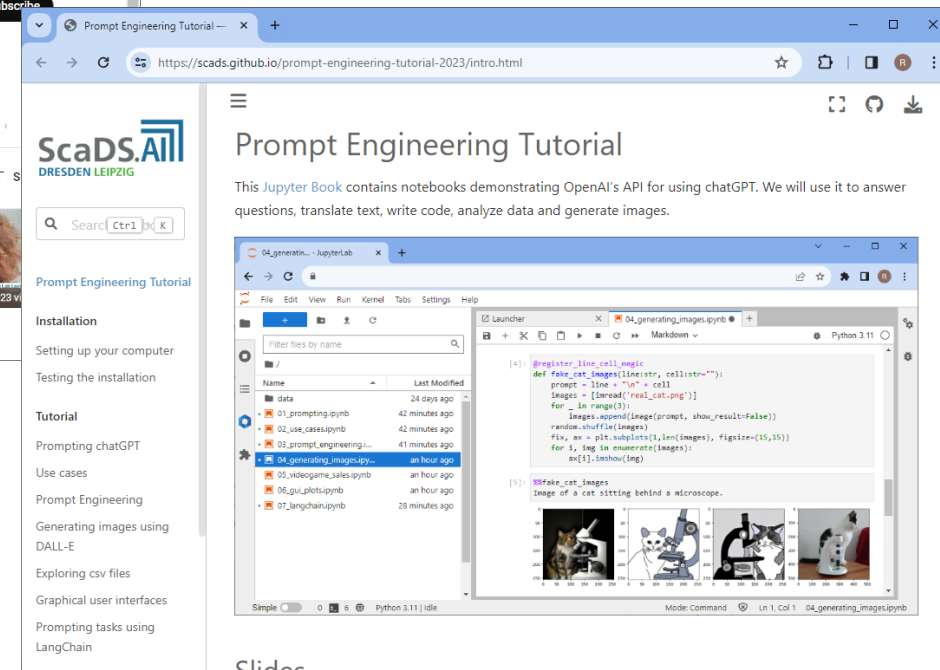
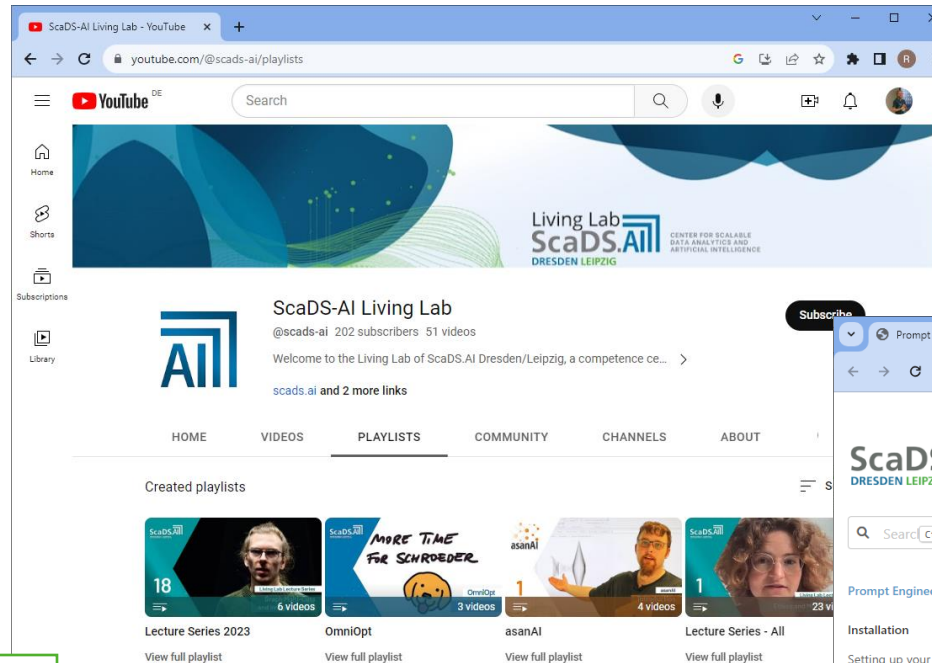
Online, open access!



The screenshot shows a web browser displaying a slide on the f1000research platform. The slide title is "Sharing and licensing material" by Robert Haase, dated June 30th 2021. The slide content includes a diagram with boxes for "Code", "Slides", "Text", and "Data". The website interface includes a search bar, navigation menu, and a "DOWNLOAD" button (30.92 MB). The slide footer contains a license notice: "This material is licensed by Robert Haase, PoL Dresden under the CC-BY 4.0 license https://creativecommons.org/licenses/by/4.0/".

Incentives: Findability -> Visibility

- YouTube
- Github

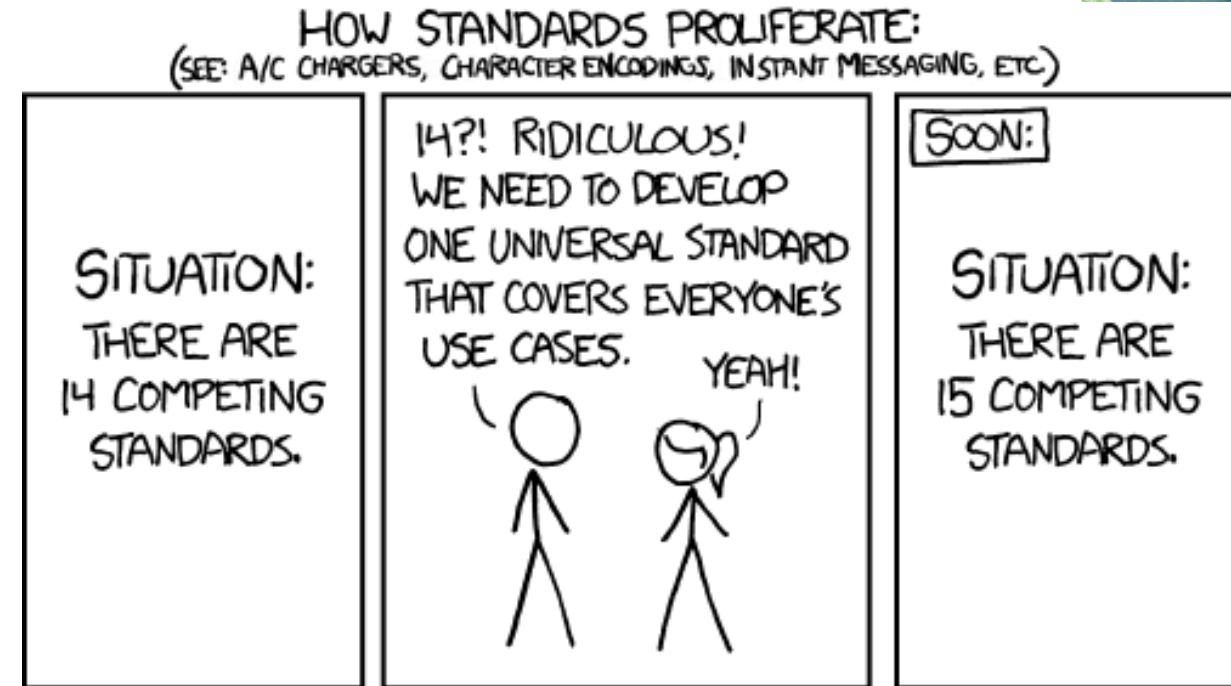


Open & FAIR sharing
is a PR instrument

- ... leading to
- more software users
 - new collaborations

The FAIR-principles

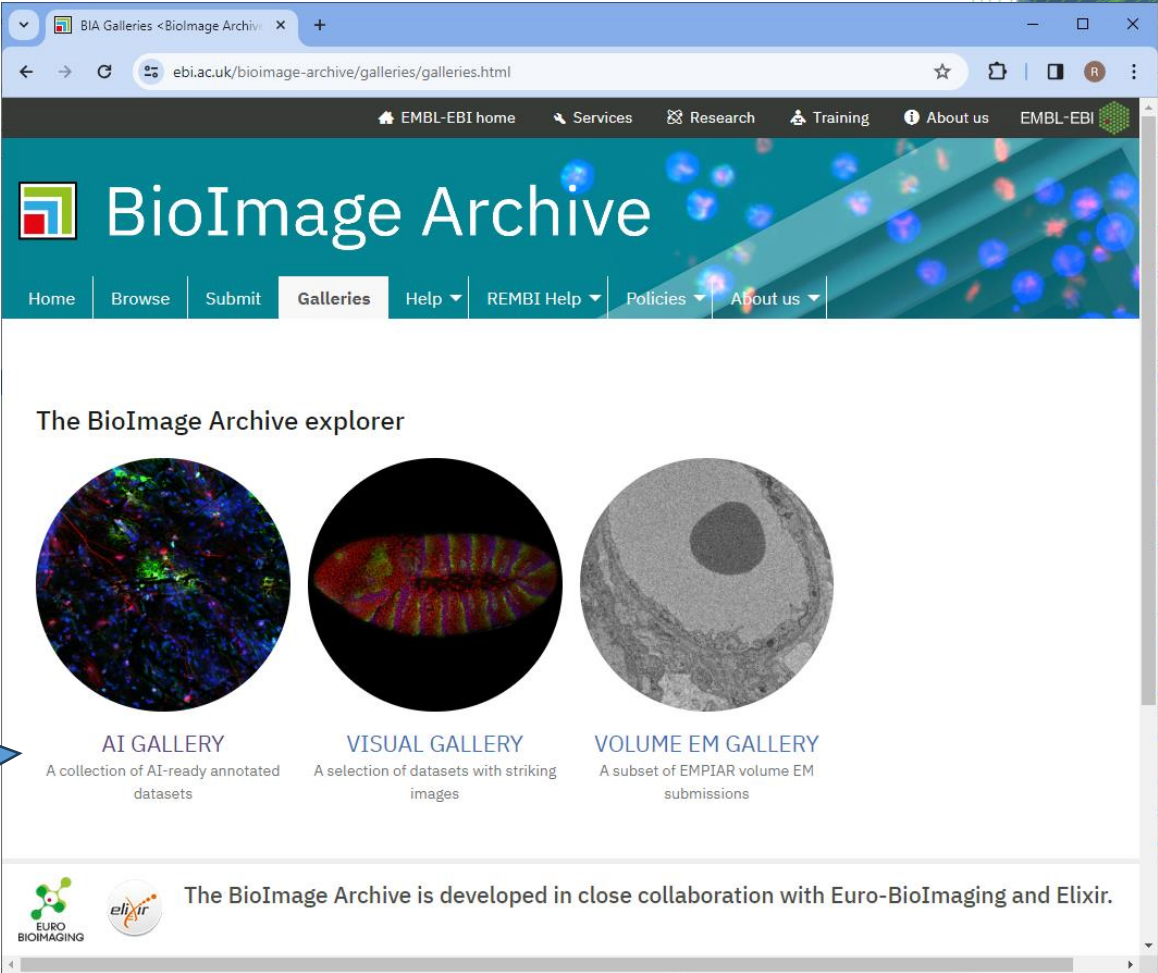
- Accessible
- A1. (Meta)data are retrievable by their identifier using a standardised communications protocol
 - A1.1 The protocol is open, free, and universally implementable
 - A1.2 The protocol allows for an authentication and authorisation procedure, where necessary
- A2. Metadata are accessible, even when the data are no longer available



Accessibility

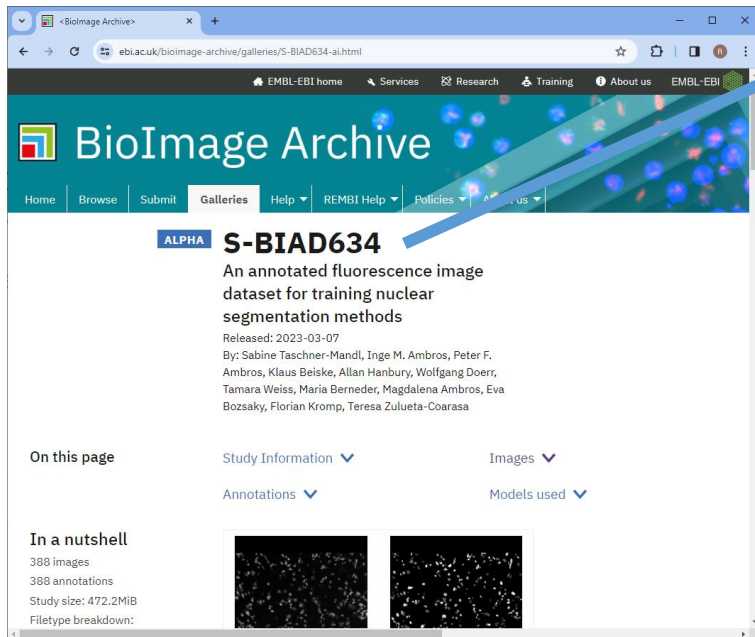
- The ability to download data, for humans and computers

Essential for AI developers =-)



Accessibility

- The ability to download data, for humans and computers



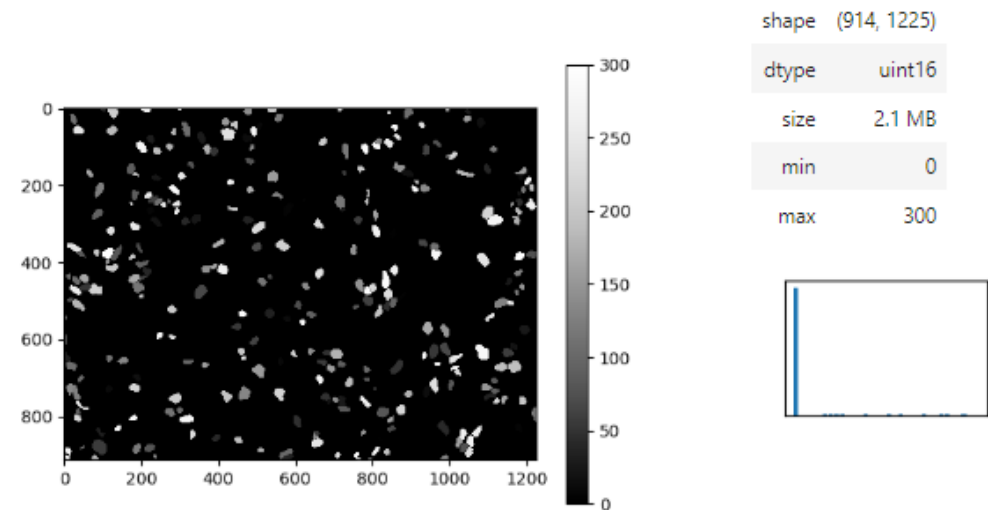
```
[1]: from bia_explorer import io, biostudies
      from skimage.io import imread
      import stackview

      accession = 'S-BIAD634'
      study = io.load_bia_study(accession)
      image = study.images[0]
```

Displaying images using stackview

```
[2]: uri = image.uri.replace("\\", "/")
      image_data = imread(uri)
      stackview.insight(image_data)
```

[2]:



Restricted Access

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

The screenshot shows the Zenodo record page for 'blobs.tif' (DOI: 10.5281/zenodo.10829230). The page is marked as 'Restricted' with a red lock icon. A blue arrow points to this icon. The record is published on March 18, 2024, and has 0 views and 0 downloads. The description states that the dataset is public-domain but files are restricted. The 'Files' section shows a red 'Restricted' label with the text: 'The record is publicly accessible, but files are restricted to users with access.'

This screenshot shows the same Zenodo record page for 'blobs.tif', but with a file preview visible. The 'Files' section shows a preview of the 'blobs.tif' image, which displays a blurry, grayscale pattern of circular shapes. The 'Restricted' status is still indicated by a red lock icon. The page also shows options for 'Edit', 'New version', and 'Share'.

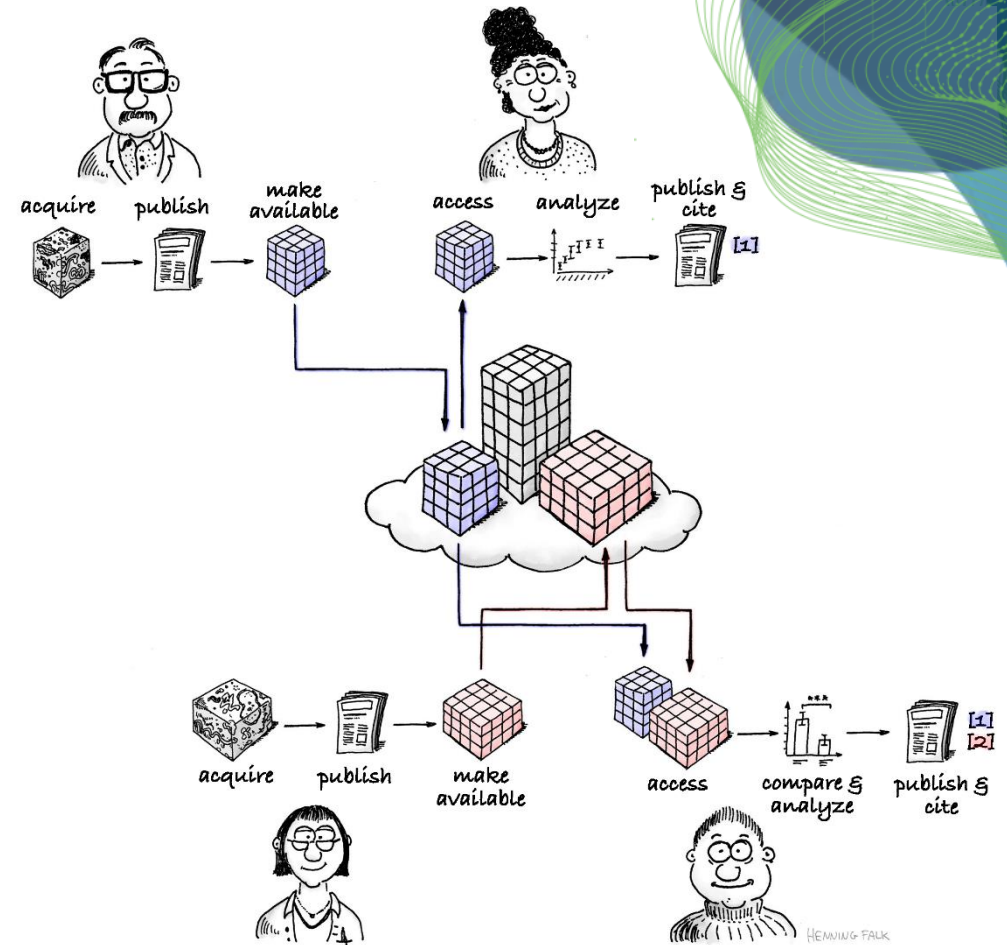
The FAIR-principles

- Interoperable
 - 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



The FAIR-principles

- Reusable
 - 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 Access -> Others teach how to use your tools & methods

Interactive image data flow graph x +
f1000research.com/slides/10-201

F1000Research Search SUBMIT YOUR RESEARCH

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

Home » Browse » Interactive image data flow graphs and GPU-accelerated image processing...

SLIDES NOT PEER REVIEWED

VIEW FULL SCREEN

PowerPoint ... 1 / 17 24% +

Pol Physics of Life TU Dresden CENTER FOR SYSTEMS BIOLOGY DRESDEN

Interactive Image Data Flow Graphs and GPU-accelerated image processing for everyone
Robert Haase

ABRF Annual Meeting March 10th 2021

@haesleinhuepf @PoLDresden

TECHNISCHE UNIVERSITÄT DRESDEN

Metrics | 329 Views | 32 Downloads

DOWNLOAD 53.58 MB

SHARE CITE

PART OF THE GATEWAY

neubias NEUBIAS - the Bioimage Analysts Network

BROWSE BY RELATED SUBJECTS

Artificial intelligence
Computer and information sciences

Interactive image data flow graph x +
f1000research.com/slides/11-1175

F1000Research Search SUBMIT YOUR RESEARCH

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

Home » Browse » Interactive image data flow graphs and reproducible GPU-accelerated...

SLIDES NOT PEER REVIEWED

VIEW FULL SCREEN

f1000resear... 1 / 33 24% +

TECHNISCHE UNIVERSITÄT DRESDEN DRESDEN concept Pol Physics of Life TU Dresden CENTER FOR SYSTEMS BIOLOGY DRESDEN

Interactive Image Data Flow Graphs and reproducible GPU-accelerated image processing
Martin Schätz

Adapted from Robert Haase, PoL, TU Dresden

@haesleinhuepf @SchatzCZ @PoLDresden

TECHNISCHE UNIVERSITÄT DRESDEN

Metrics | 20 Views | 5 Downloads

DOWNLOAD 43.20 MB

SHARE CITE

PART OF THE GATEWAY

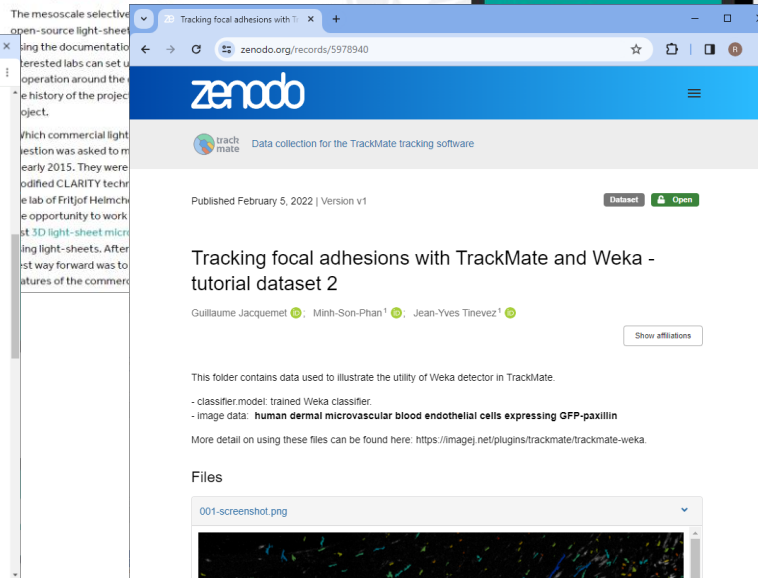
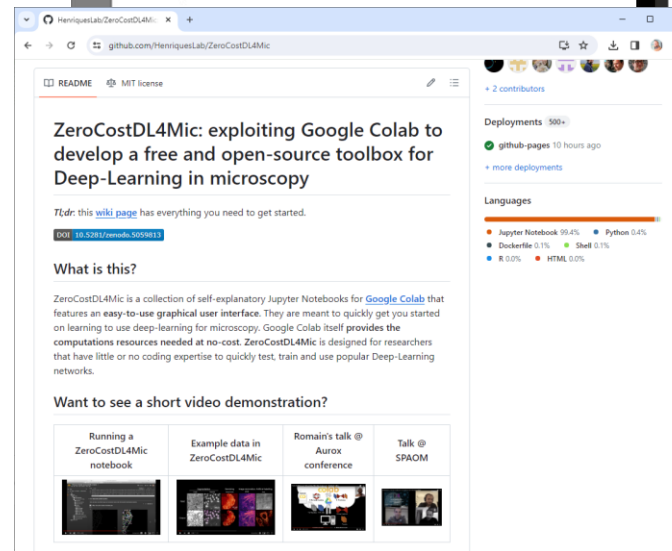
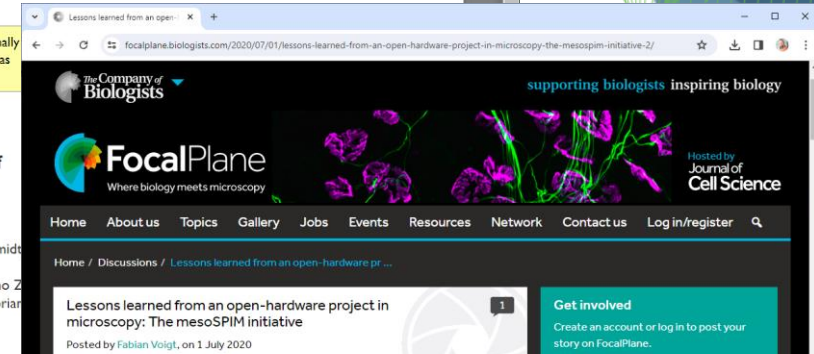
neubias NEUBIAS - the Bioimage Analysts Network

BROWSE BY RELATED SUBJECTS

Artificial intelligence
Computer and information sciences

Where to share?

- Open science related content
 - bioRxiv (manuscripts, no reviews)
 - Figshare
 - F1000
 - Bioimage Archive (data)
 - Github (code)
 - Zenodo
 - Focalplane
 - Institutional servers (if there is no alternative)



Quiz

- Where might open source code be most *visible*?

Git server of the
university



Zenodo.org



Github.com



Group / institute /
personal website



A short excursion...

- Published 2 weeks ago:

Claim to
outperform current
state of the art.

nature methods View all journals Search Log in

Explore content ▾ About the journal ▾ Publish with us ▾ Sign up for alerts RSS feed

nature > nature methods > analyses > article

Analysis | Published: 26 March 2024

The multimodality cell segmentation challenge: toward universal solutions

Jun Ma, Ronald Xie, Shamini Ayyadhury, Cheng Ge, Anubha Gupta, Ritu Gupta, Song Gu, Yao Zhang, Gihun Lee, Joonkee Kim, Wei Lou, Haofeng Li, Eric Upschulte, Timo Dickscheid, José Guilherme de Almeida, Yixin Wang, Lin Han, Xin Yang, Marco Labagnara, Vojislav Gligorovski, Maxime Scheder, Sahand Jamal Rahi, Carly Kempster, Alice Pollitt, ... Bo Wang + Show authors

Nature Methods (2024) | [Cite this article](#)

7645 Accesses | 19 Altmetric | [Metrics](#)

You have full access to this article via **Universitätsbibliothek Leipzig**

[Download PDF](#)

Associated content

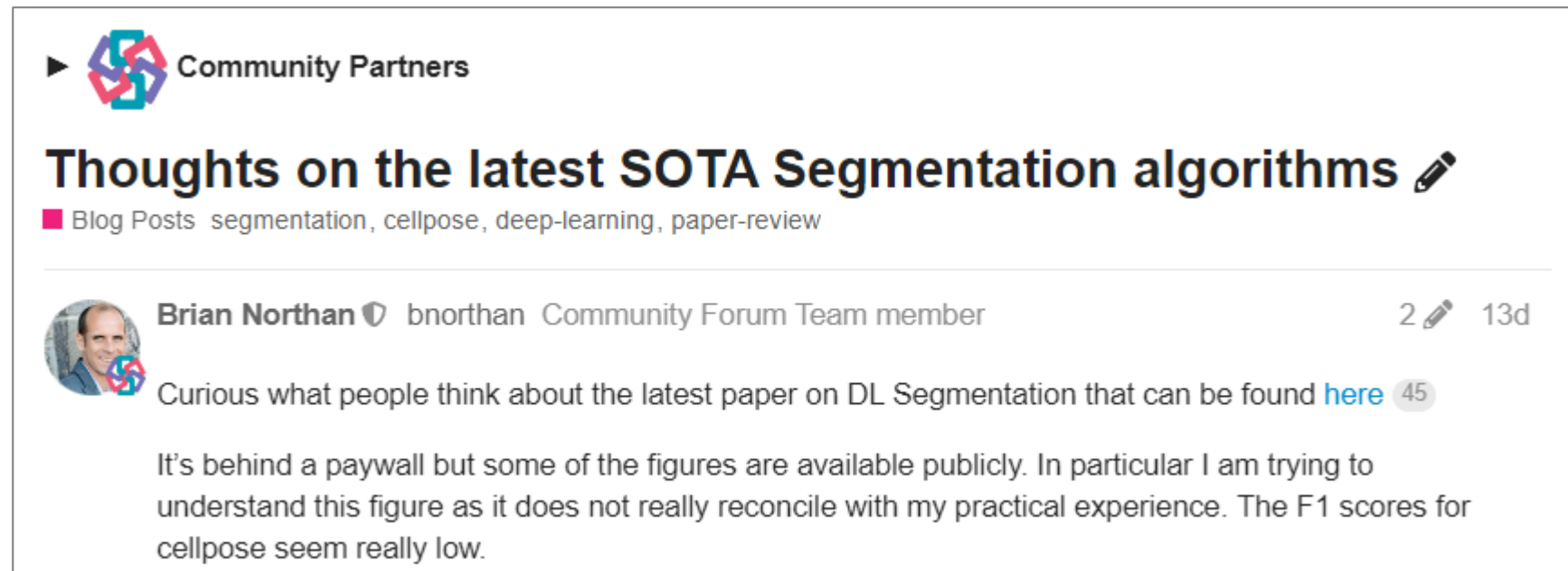
[Creating a universal cell segmentation algorithm](#)

Nature Methods | Research Briefing | 01 Apr 2024

“The T1 algorithms achieved a median F1 score of 89.7% (IQR 36.7–82.4%), surpassing the KIT-GE, Cellpose-pretrain, Cellpose-scratch, Omnipose-pretrain and Omnipose-scratch by 49.9%, 24.4%, 35.4%, 58.9% and 48.7%, respectively.”

A short excursion...

- “However my initial testing of the winning entry [mediar 8](#) reveals errors (which are typical without tuning) when applied to recent data posted [here 6](#) for which cellpose produced a (subjectively) good result.”



The screenshot shows a forum post from the 'Community Partners' section. The post title is 'Thoughts on the latest SOTA Segmentation algorithms' with a pencil icon. Below the title are tags: 'Blog Posts', 'segmentation', 'cellpose', 'deep-learning', and 'paper-review'. The author is 'Brian Northan', a 'Community Forum Team member', with a profile picture and a small icon. The post content reads: 'Curious what people think about the latest paper on DL Segmentation that can be found [here](#) 45'. Below this is a paragraph: 'It's behind a paywall but some of the figures are available publicly. In particular I am trying to understand this figure as it does not really reconcile with my practical experience. The F1 scores for cellpose seem really low.'

A short excursion...

- Published on Sunday:
(by the CellPose authors)
- “In fact, Mediar directly copied our codebase for implementing the Cellpose framework without modification”

New Results

Follow this preprint

Transformers do not outperform Cellpose

Carsen Stringer, Marius Pachitariu

doi: <https://doi.org/10.1101/2024.04.06.587952>

This article is a preprint and has not been certified by peer review [what does this mean?].

0 0 0 1 0 0 126

Abstract

Info/History Metrics

Preview PDF

Abstract

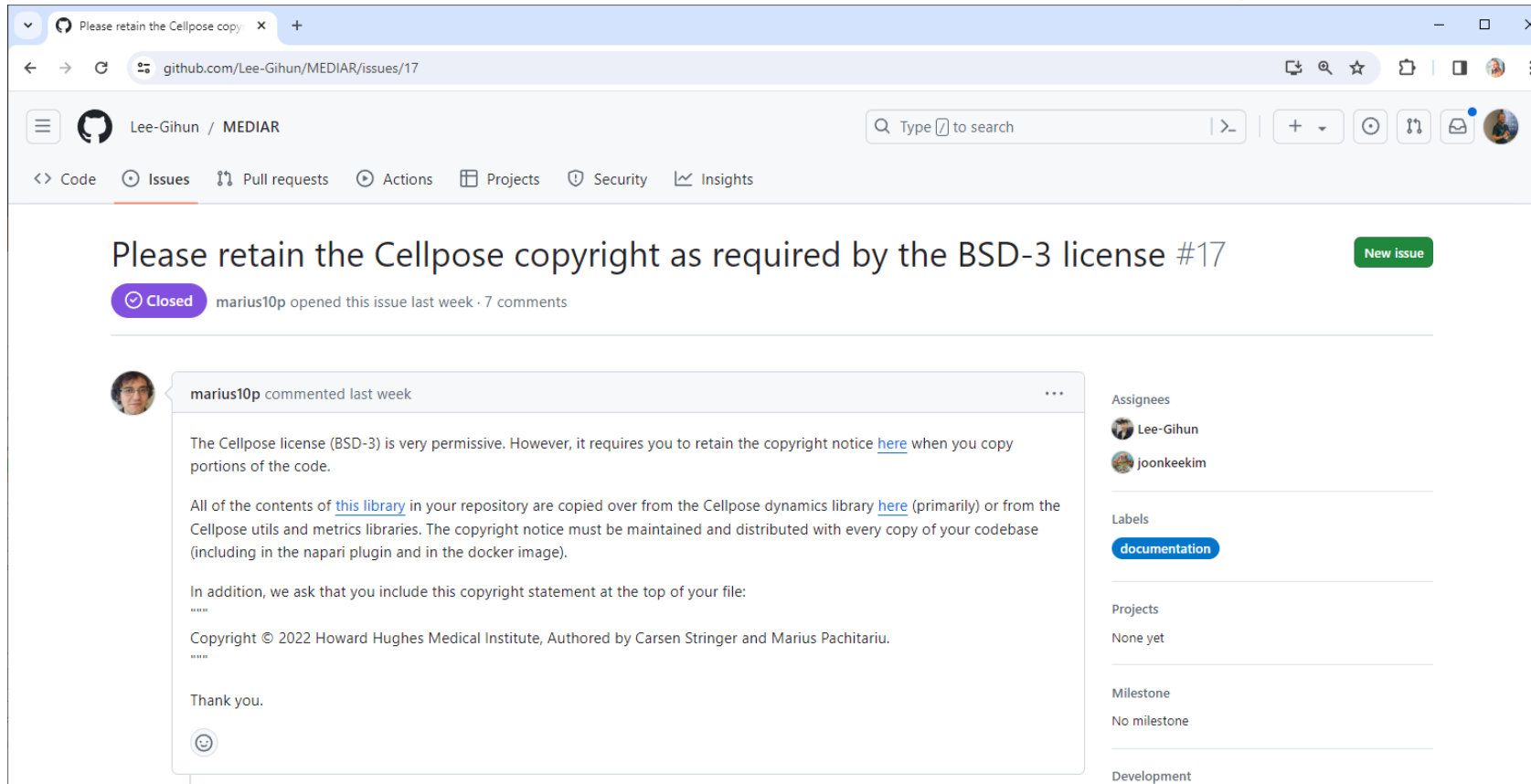
In a recent publication, Ma et al (2024) claim that a transformer-based cellular segmentation method called Mediar - which won a Neurips challenge - outperforms Cellpose (0.897 vs 0.543 median F1 score). Here we show that this result was obtained by artificially impairing Cellpose in multiple ways. When we removed these impairments, Cellpose outperformed Mediar (0.861 vs 0.826 median F1 score on the updated test set). To further investigate the performance of transformers for cellular segmentation, we replaced the Cellpose backbone with a transformer. The transformer-Cellpose model also did not outperform the standard Cellpose (0.848 median F1 test score). Our results suggest that transformers do not advance the state-of-the-art in cellular segmentation.

Competing Interest Statement

The authors have declared no competing interest.

A short excursion...

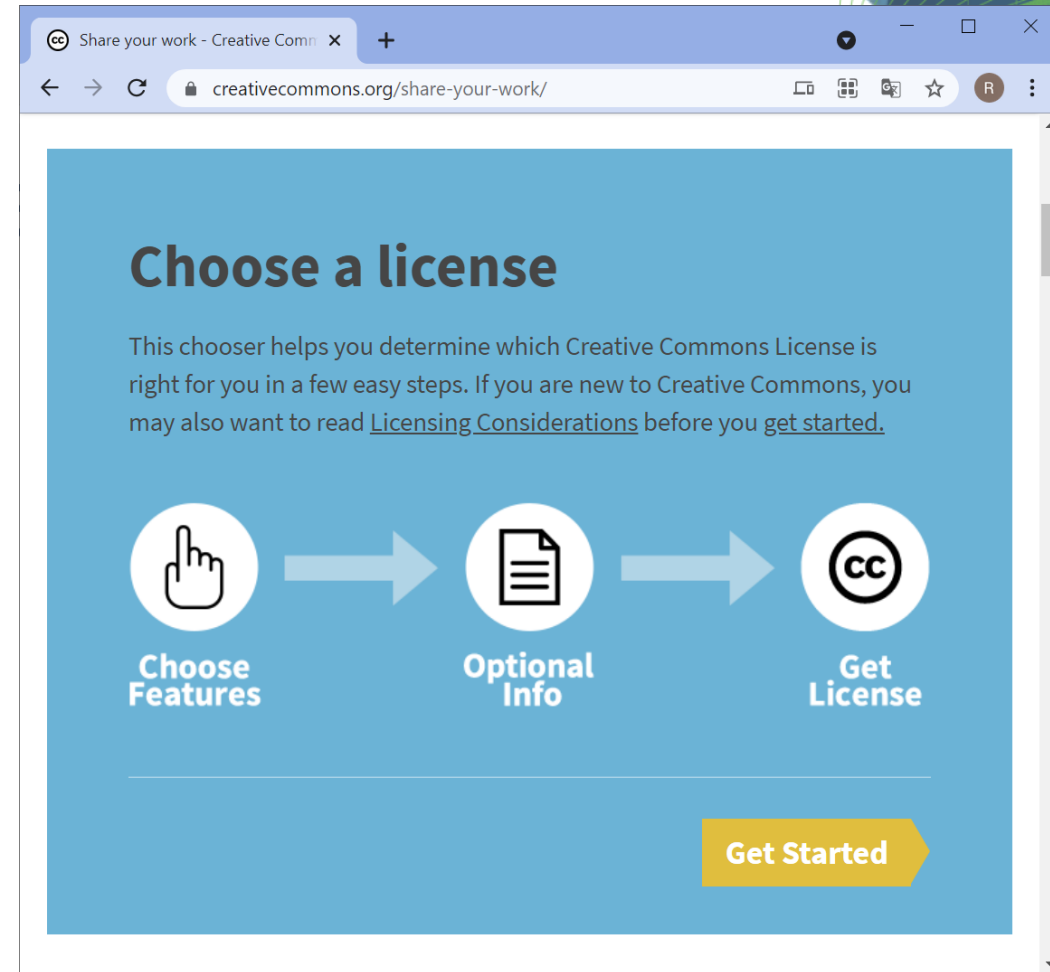
- Professional discussion about Copyright:



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/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)



“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.

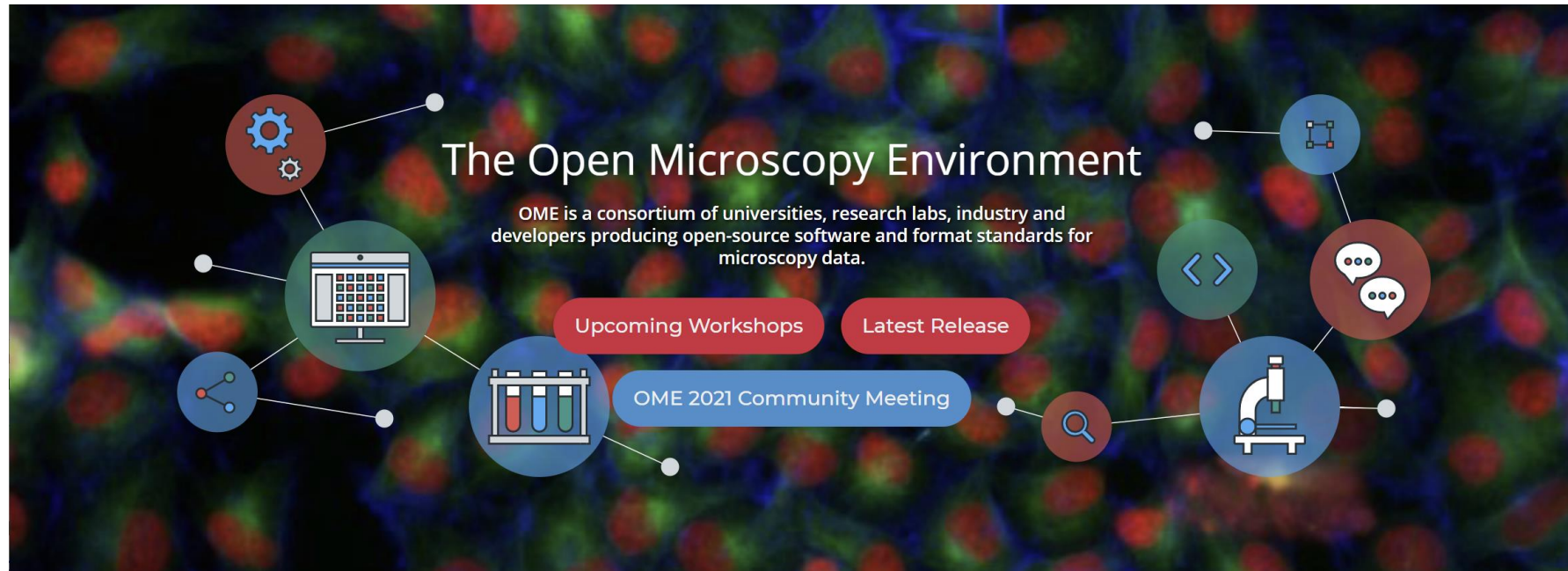
CC BY includes the following elements:

BY  – Credit must be given to the creator

Content on this slide was adapted from <https://creativecommons.org/licenses/by/4.0/>, 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?
- May I download and redistribute this preprint to students of a course for free?

 Cold Spring Harbor Laboratory **bioRxiv**
THE PREPRINT SERVER FOR BIOLOGY

bioRxiv posts many COVID19-related papers. A reminder: they have not been formally peer-reviewed and should not guide health-related behavior or be reported in the press as conclusive.








New Results [Follow this preprint](#)

Content-Aware Image Restoration: Pushing the Limits of Fluorescence Microscopy

 Martin Weigert,  Uwe Schmidt,  Tobias Boothe,  Andreas Müller,  Alexandr Dibrov,  Akanksha Jain,  Benjamin Wilhelm,  Deborah Schmidt,  Coleman Broaddus,  Siân Culley,  Mauricio Rocha-Martins,  Fabián Segovia-Miranda,  Caren Norden,  Ricardo Henriques,  Marino Zerial,  Michele Solimena,  Jochen Rink,  Pavel Tomancak,  Loic Royer,  Florian Jug,  Eugene W. Myers

doi: <https://doi.org/10.1101/236463>

Now published in *Nature Methods* doi: [10.1038/s41592-018-0216-7](https://doi.org/10.1038/s41592-018-0216-7)

 1  0  2  0  2  0  618

Abstract Full Text **Info/History** Metrics [Preview PDF](#)

ARTICLE INFORMATION

doi <https://doi.org/10.1101/236463>
History July 3, 2018.

ARTICLE VERSIONS

Version 1 (December 19, 2017 - 23:34).
Version 2 (December 21, 2017 - 14:29).
Version 3 (January 23, 2018 - 12:52).
Version 4 (January 23, 2018 - 20:15).
You are viewing Version 5, the most recent version of this article.

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-NC-ND 4.0 International license](#).

Yes



No



Yes

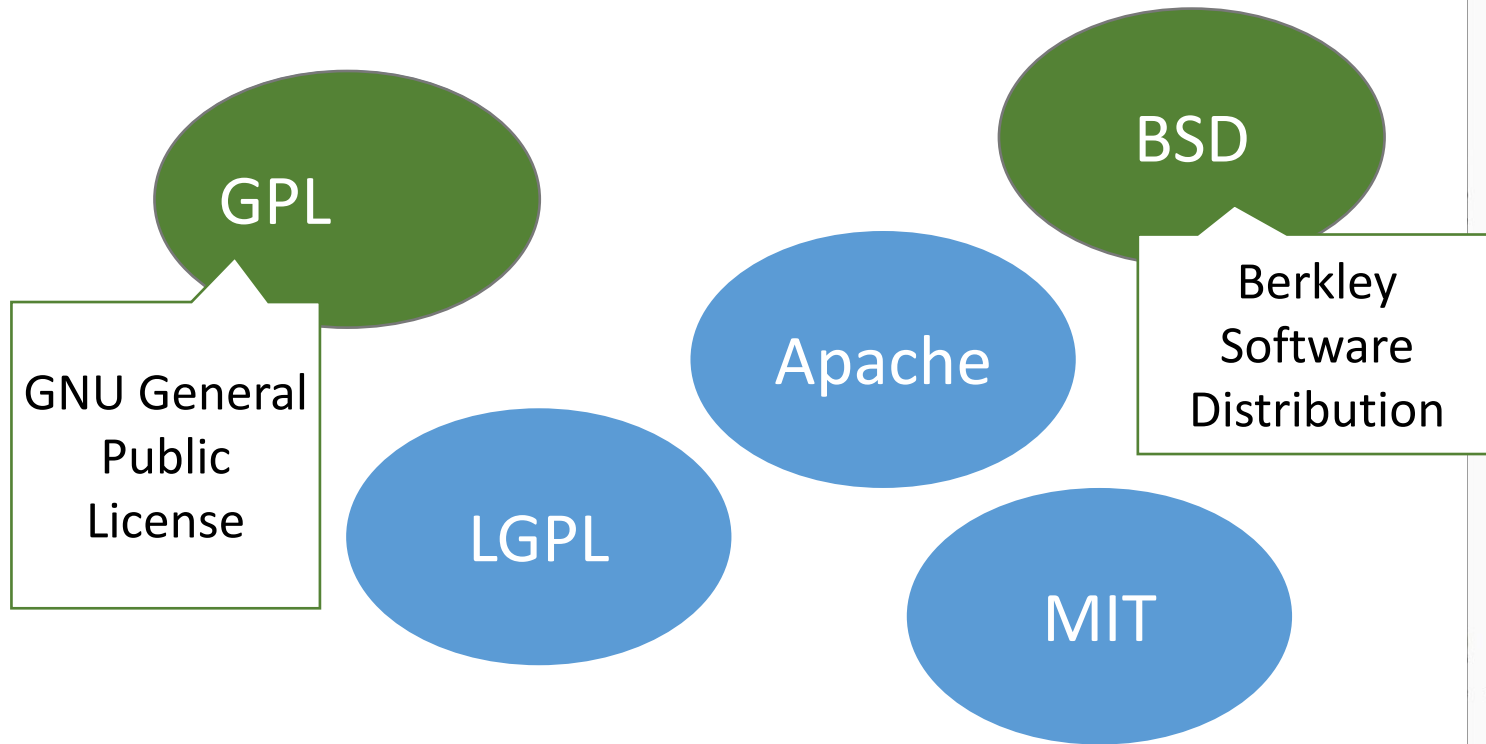


No



Licensing Software

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



The screenshot shows the homepage of choosealicense.com. The main heading is "Choose an open source license". Below it, a question asks: "Which of the following best describes your situation?". Three options are presented:

- I need to work in a community.** (Icon: three people) - Description: "Use the license preferred by the community you're contributing to or depending on. Your project will fit right in. If you have a dependency that doesn't have a license, ask its maintainers to add a license." Example: "Babel, .NET Core, and Rails use the MIT License."
- I want it simple and permissive.** (Icon: double-headed arrow) - Description: "The MIT License is short and to the point. It lets people do almost anything they want with your project, like making and distributing closed source versions." Example: "Ansible, Bash, and GIMP use the GNU GPLv3."
- I care about sharing improvements.** (Icon: circular arrow) - Description: "The GNU GPLv3 also lets people do almost anything they want with your project, except distributing closed source versions." Example: "Ansible, Bash, and GIMP use the GNU GPLv3."

Below this, another question asks: "What if none of these work for me?". Three options are presented:

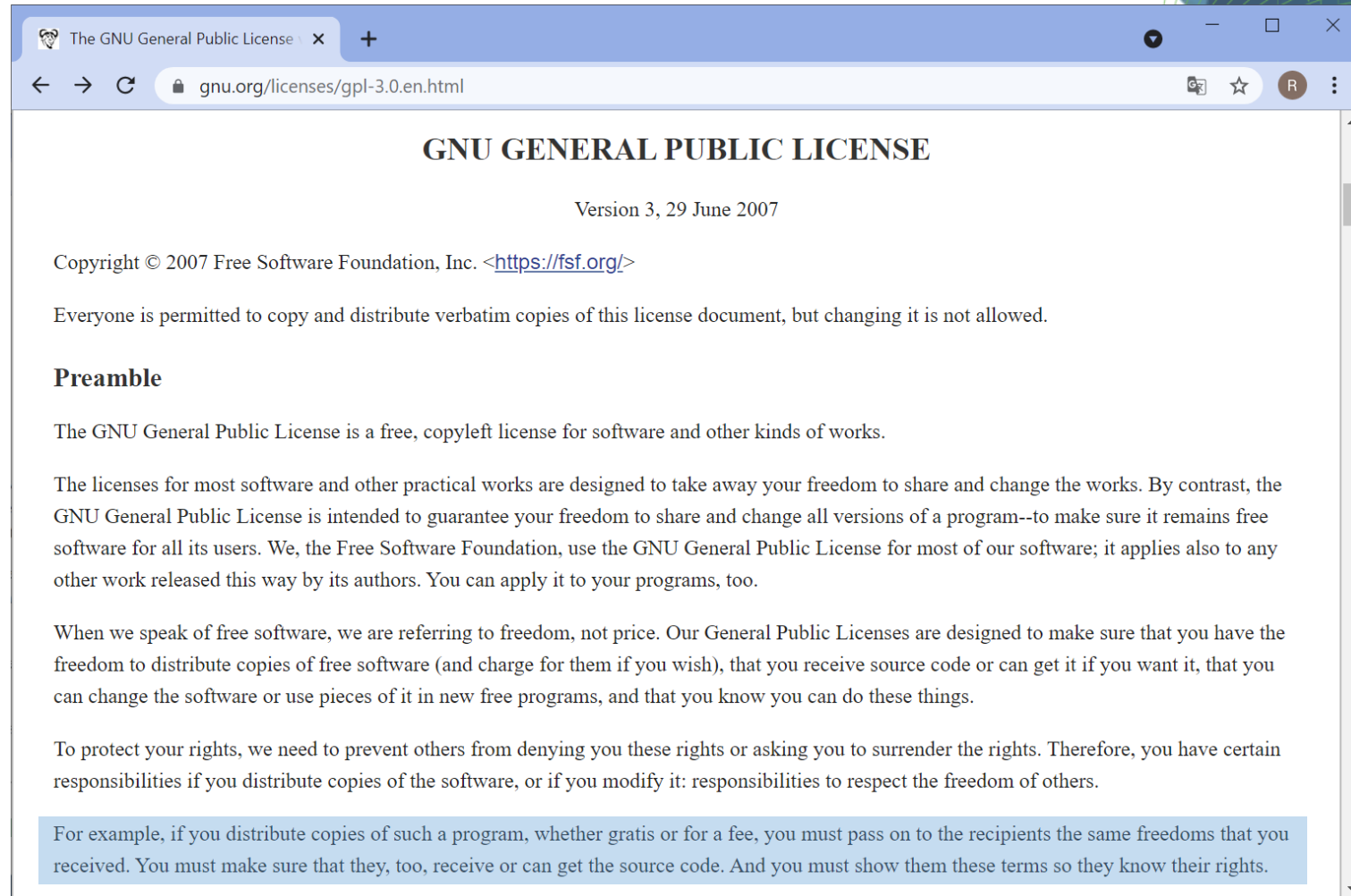
- My project isn't software.** - Description: "There are licenses for that."
- I want more choices.** - Description: "More licenses are available."
- I don't want to choose a license.** - Description: "Here's what happens if you don't."

At the bottom, there is a footer with the Creative Commons Attribution 3.0 Unported License and links for "About", "Terms of Service", and "Help improve this page".

Licensing Software: GPL

GPL-derivatives must also be GPL-licensed

“Restrictive”
licensing



The screenshot shows a web browser window displaying the GNU General Public License (GPL) page. The browser's address bar shows the URL gnu.org/licenses/gpl-3.0.en.html. The page title is "GNU GENERAL PUBLIC LICENSE" and the version is "Version 3, 29 June 2007". The copyright notice is "Copyright © 2007 Free Software Foundation, Inc. <<https://fsf.org/>>". The text states: "Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed." The section "Preamble" begins with: "The GNU General Public License is a free, copyleft license for software and other kinds of works." It explains that the license is intended to guarantee freedom to share and change all versions of a program. A highlighted blue box at the bottom of the page contains the text: "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 openly 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.

} Similar to CC0
(public domain)

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

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

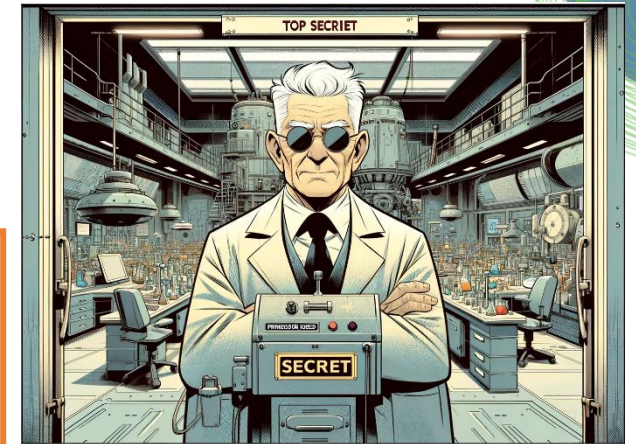
Similar to
CC-BY

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



Quiz

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

Yes



No



cnr-isti-vclab/meshlab: The open source mesh processing system

github.com/cnr-isti-vclab/meshlab

Search or jump to... Pull requests Issues Marketplace Explore

cnr-isti-vclab / meshlab Public

Watch 150 Fork 686 Star 3.3k

Code Issues 124 Pull requests 2 Discussions Actions Projects Wiki Security Insights

main 3 branches 40 tags Go to file Add file Code

alemuntoni bugfix while loading exif info ✓ 2672c14 22 days ago 10,254 commits

.github	fix missing windeployqt in windows workflows	5 months ago
docs	Fix various typos	8 months ago
resources	icon on windows folder	5 months ago
sample	add gltf samples	16 months ago
scripts	deploy all plugins in macos script	4 months ago
src	bugfix while loading exif info	22 days ago
textures	moved textures folder outside distrib	3 years ago
unsupported	remove "vertex color noise" filter	12 months ago
.gitignore	move build and install dirs on macos outside src	5 months ago

About

The open source mesh processing system

www.meshlab.net

point-cloud mesh mesh-generation

3d-printing 3d-scanning 3d

3d-reconstruction 3d-models

mesh-processing mesh-editing

mesh-simplification triangle-mesh

Readme

GPL-3.0 license

3.3k stars

150 watching

686 forks

Quiz

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

Yes



No



The screenshot shows the GitHub repository page for napari/napari. The repository is public and has 1.5k stars, 45 watchers, and 328 forks. It contains 6 branches and 172 tags. The main branch is selected. The repository description is "napari: a fast, interactive, multi-dimensional image viewer for python". The repository includes a README, BSD-3-Clause license, and 114 releases. The file list shows folders like .devcontainer, .github, binder, examples, napari, napari_builtins, resources, tools, and .env_sample, along with their respective commit messages and dates.



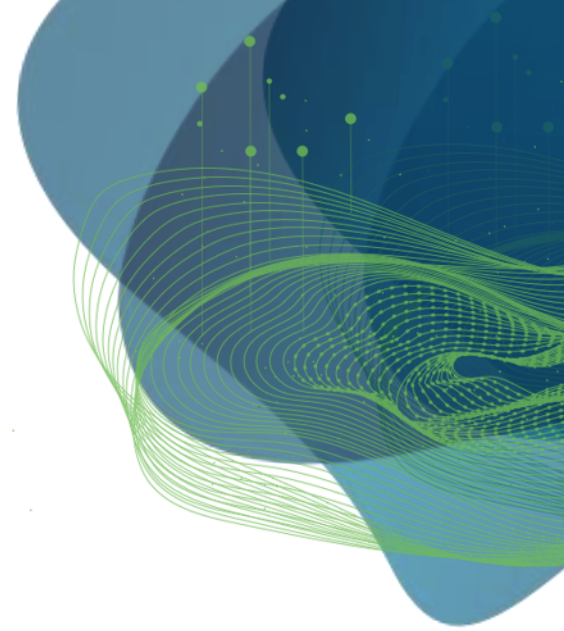
ScaDS.AI

DRESDEN LEIPZIG

CENTER FOR SCALABLE DATA ANALYTICS
AND ARTIFICIAL INTELLIGENCE

Software environments

Robert Haase



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

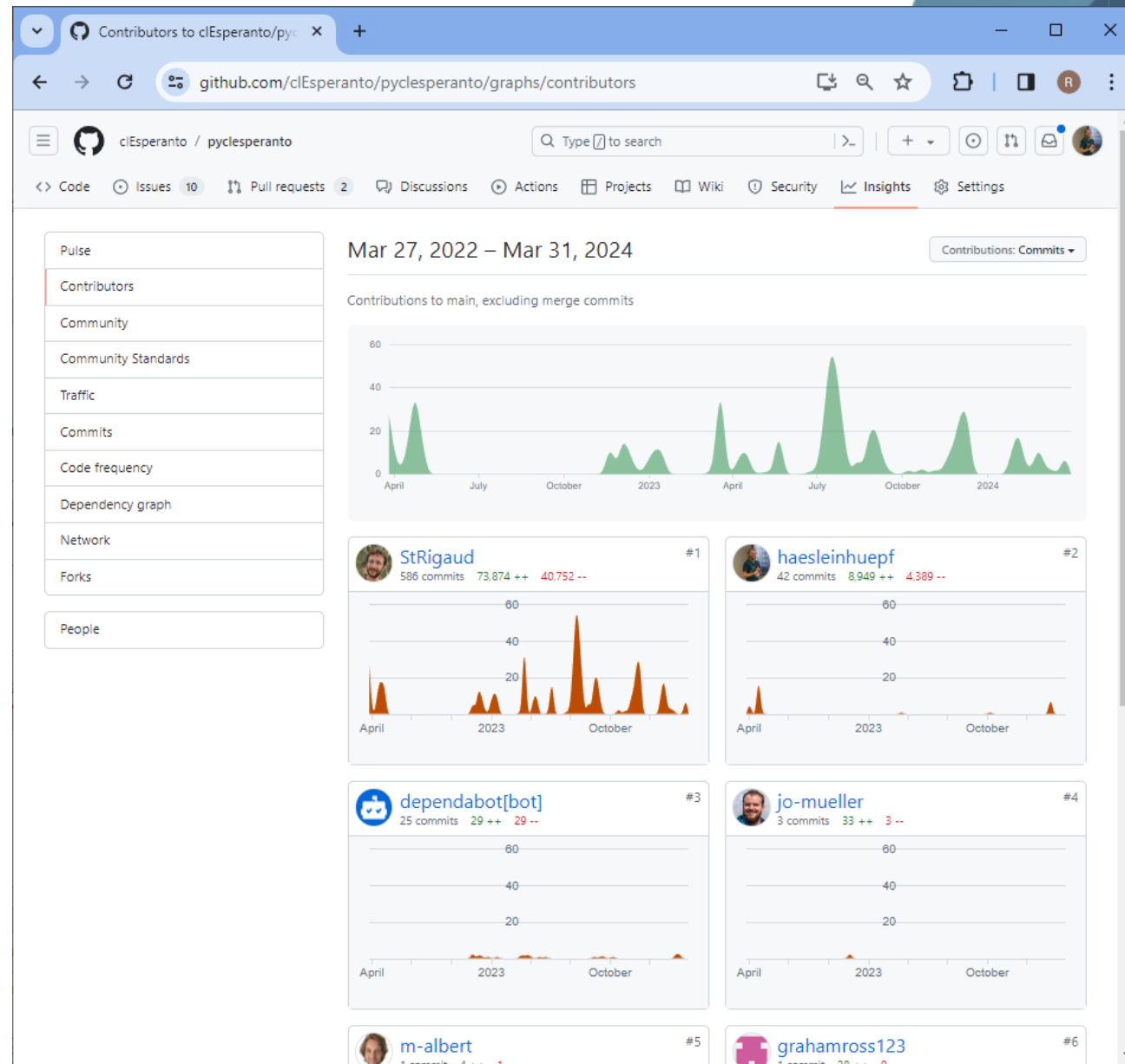


SACHSEN

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.

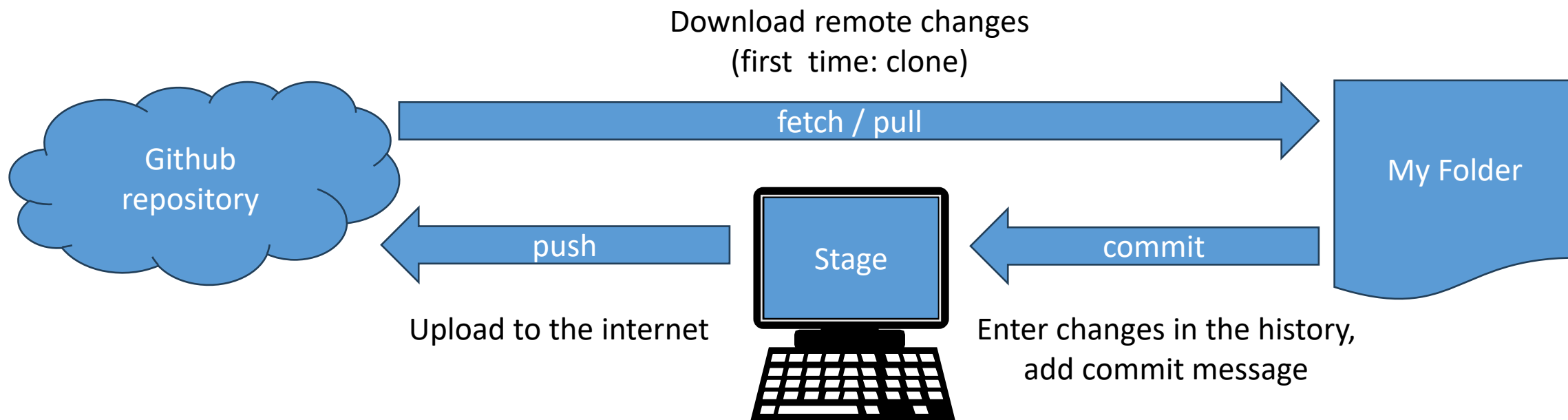
git

- Version control is key element of data scientist's toolbox
- Distributed file system with sophisticated logging mechanisms
- Control about what becomes part of a repository and what not



git

- Git makes file modifications a more active / involved process (making people think about)



git

- Who wrote this code
- when and
- why?

The image shows two browser windows from GitHub. The left window displays the commit history for the repository `haesleinhuepf/example_image_analysis_script`. The commit list includes:

- Merge pull request #1 from BiAPoL/main
- bugfix: threshold_otsu** (highlighted with an orange arrow)
- Add minimal working example
- add gitignore

The right window shows the diff for the commit `bugfix: threshold_otsu`. It displays changes to the file `my_library.py`. The diff shows a function `segment_image` with several lines of code. Lines 9 and 10 are highlighted in green, indicating additions:

```
6 6 blurred = gaussian(image, sigma=2)
7 7
8 8 # binarize the image
9 + binary = threshold_otsu(blurred)
10 + threshold = threshold_otsu(blurred)
11
12 # label connected components
13 result = label(binary)
```

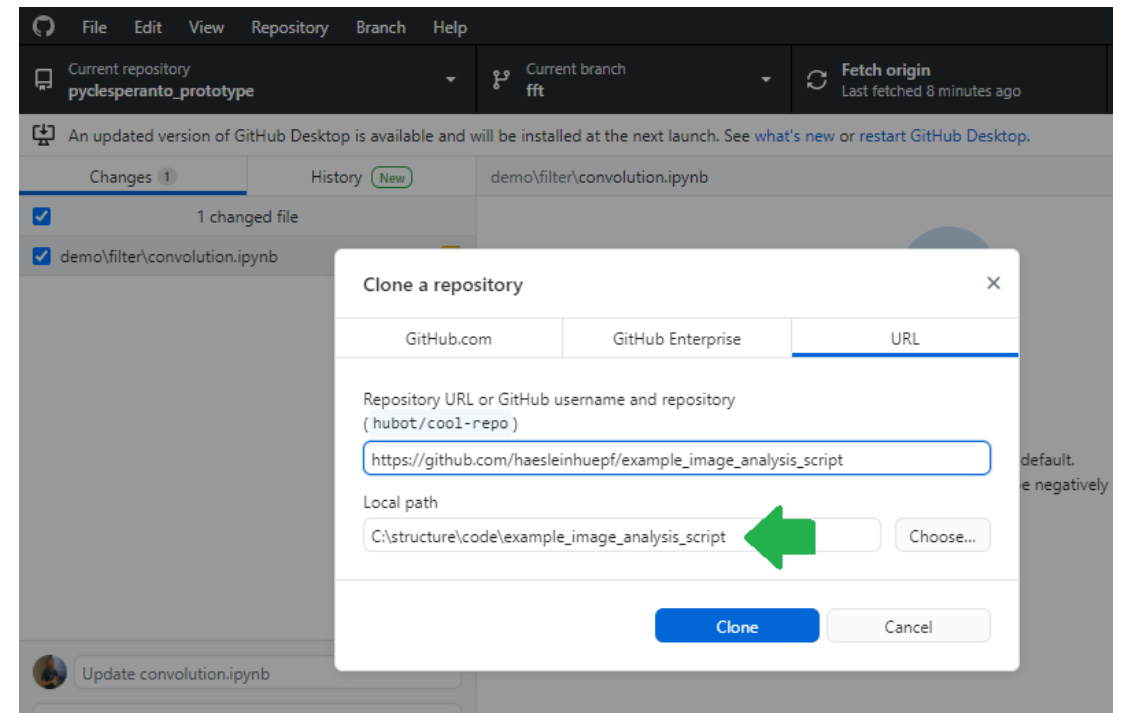
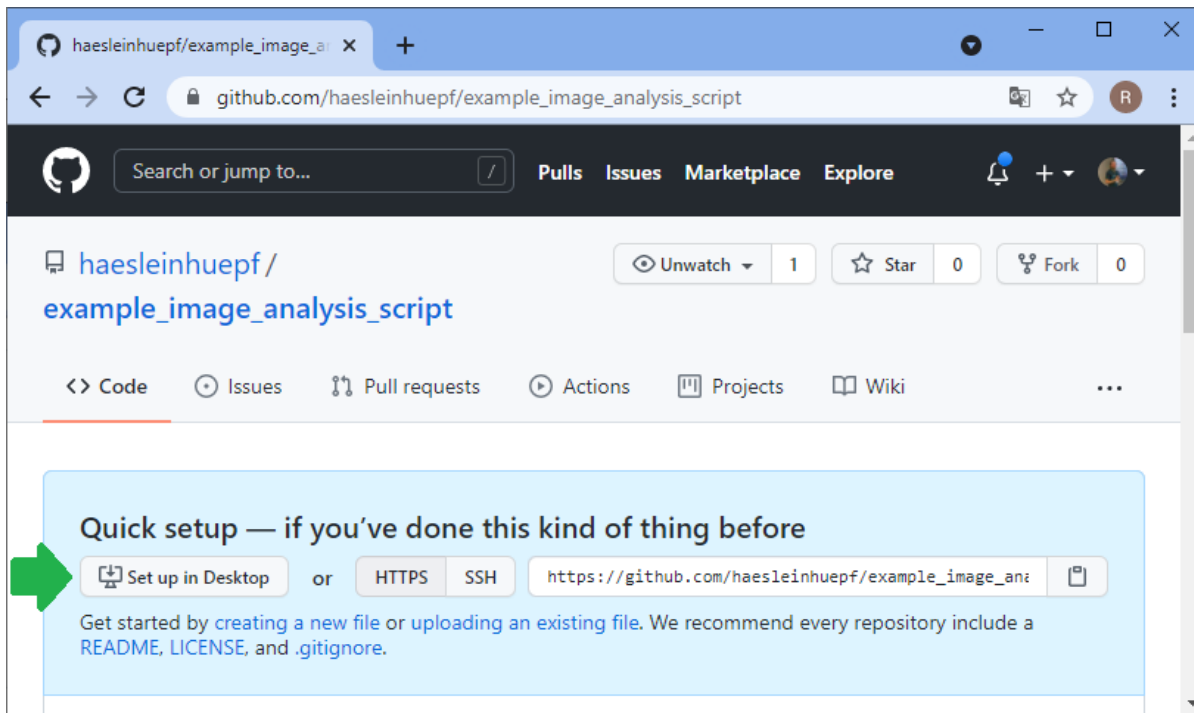
github – creating repositories

- Add a new, empty repository

The image consists of three overlapping screenshots of the GitHub website, illustrating the process of creating a new repository. The top-left screenshot shows the user's profile page for Robert Haase (haesleinhuepf), with a green arrow pointing to the 'Repositories' tab. The middle screenshot shows the 'Create a New Repository' page, with a green arrow pointing to the 'Repository name' field where 'example_image_analysis_script' is entered and validated. The bottom-right screenshot shows the same 'Create a New Repository' page, with a green arrow pointing to the 'Create repository' button. The interface includes navigation menus, search bars, and various repository options.

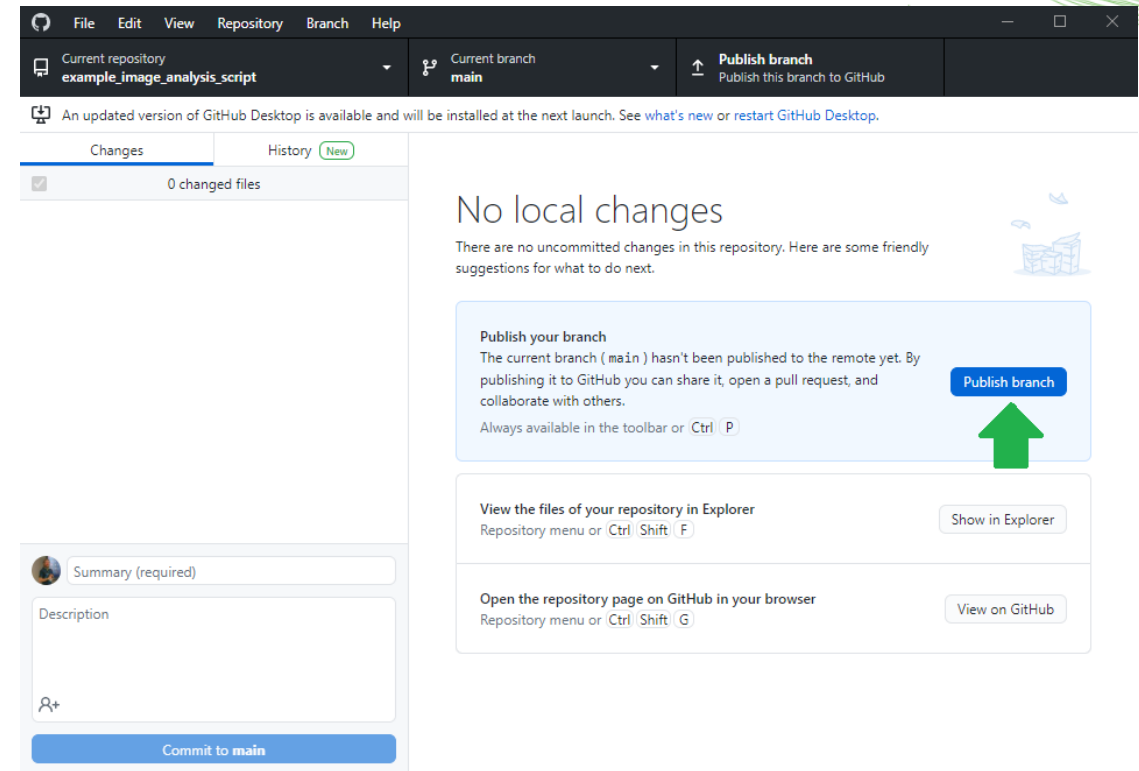
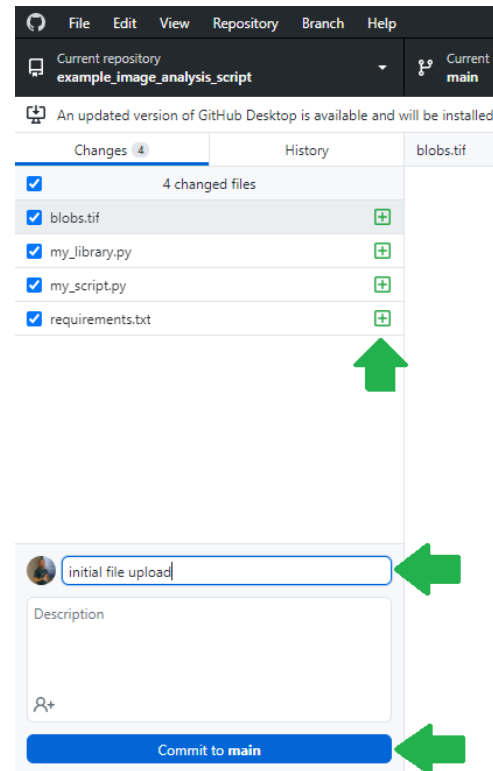
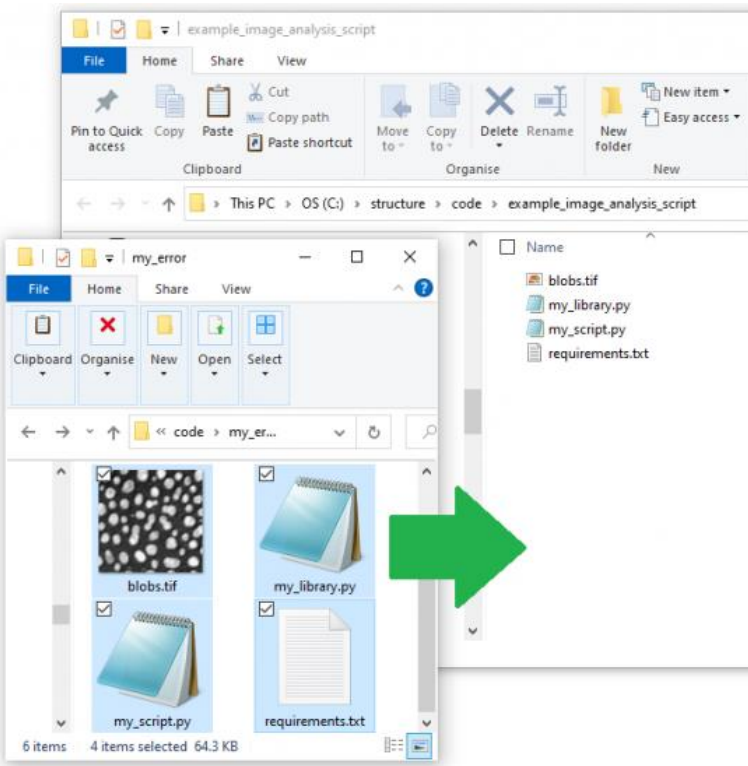
github – clone repositories

- `git clone https://github.com/organization/repository`
- Or: Use the Github Desktop app



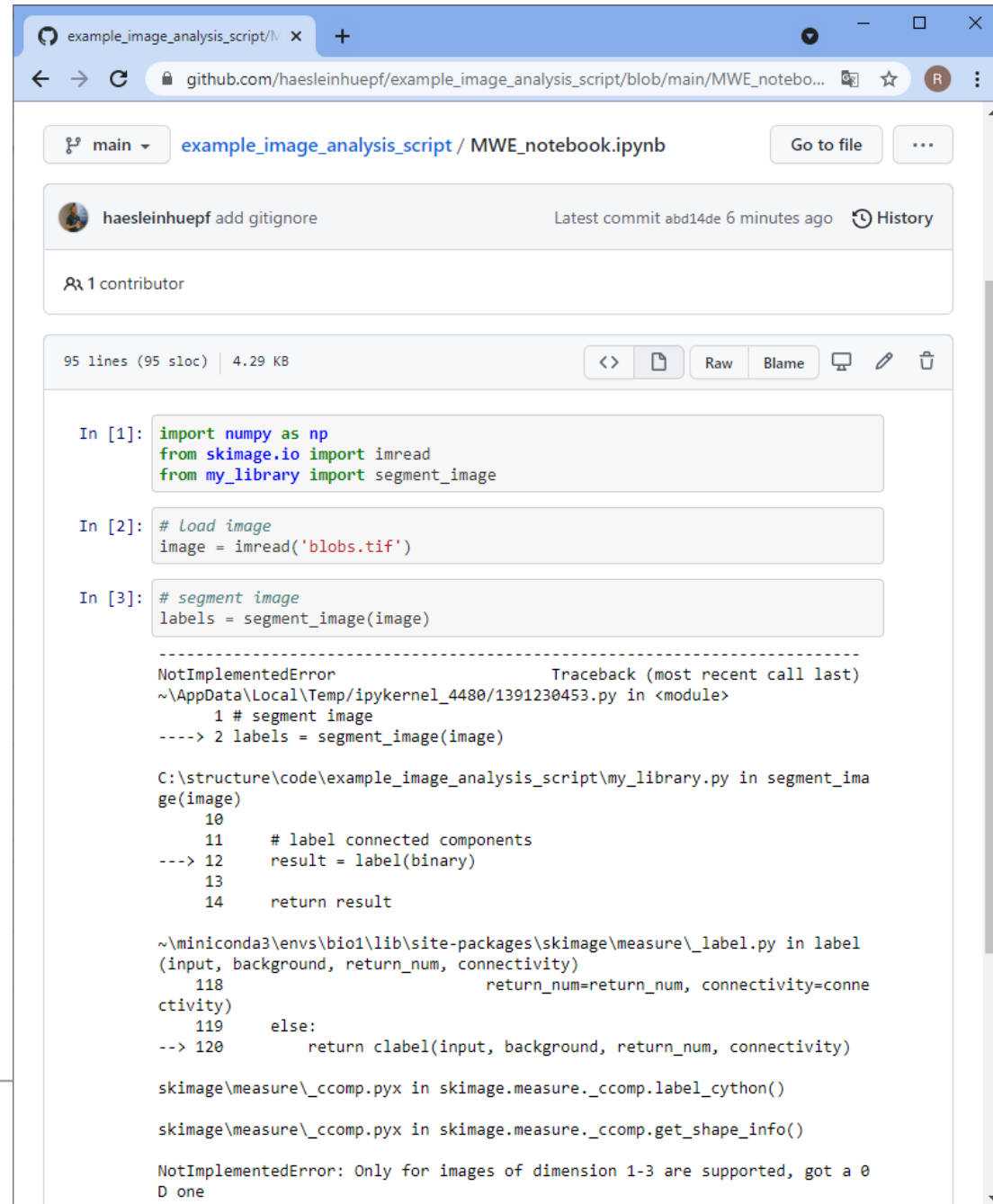
github - uploading

- git [add], commit, push



github

- Ease of reading notebooks online
- No need to download and execute code



The screenshot shows a GitHub repository page for 'example_image_analysis_script' by user 'haesleinhuepf'. The file 'MWE_notebook.ipynb' is displayed, containing three code cells. The first cell imports 'numpy', 'skimage.io', and 'my_library'. The second cell loads an image from 'blobs.tif'. The third cell calls 'segment_image' on the loaded image. Below the code, a traceback error is shown: 'NotImplementedError: Only for images of dimension 1-3 are supported, got a 0 D one'. The error originates from 'my_library.py' in the 'segment_image' function, which calls 'label' from 'skimage.measure._label.py', which in turn calls 'label_cython' from 'skimage.measure._ccomp.pyx'.

```
In [1]: import numpy as np
        from skimage.io import imread
        from my_library import segment_image

In [2]: # Load image
        image = imread('blobs.tif')

In [3]: # segment image
        labels = segment_image(image)

-----
NotImplementedError                                Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_4480\1391230453.py in <module>
      1 # segment image
----> 2 labels = segment_image(image)

C:\structure\code\example_image_analysis_script\my_library.py in segment_image(image)
     10
     11 # label connected components
----> 12 result = label(binary)
     13
     14 return result

~\miniconda3\envs\bio1\lib\site-packages\skimage\measure\_label.py in label(input, background, return_num, connectivity)
    118         return_num=return_num, connectivity=connectivity)
    119     else:
--> 120         return clabel(input, background, return_num, connectivity)

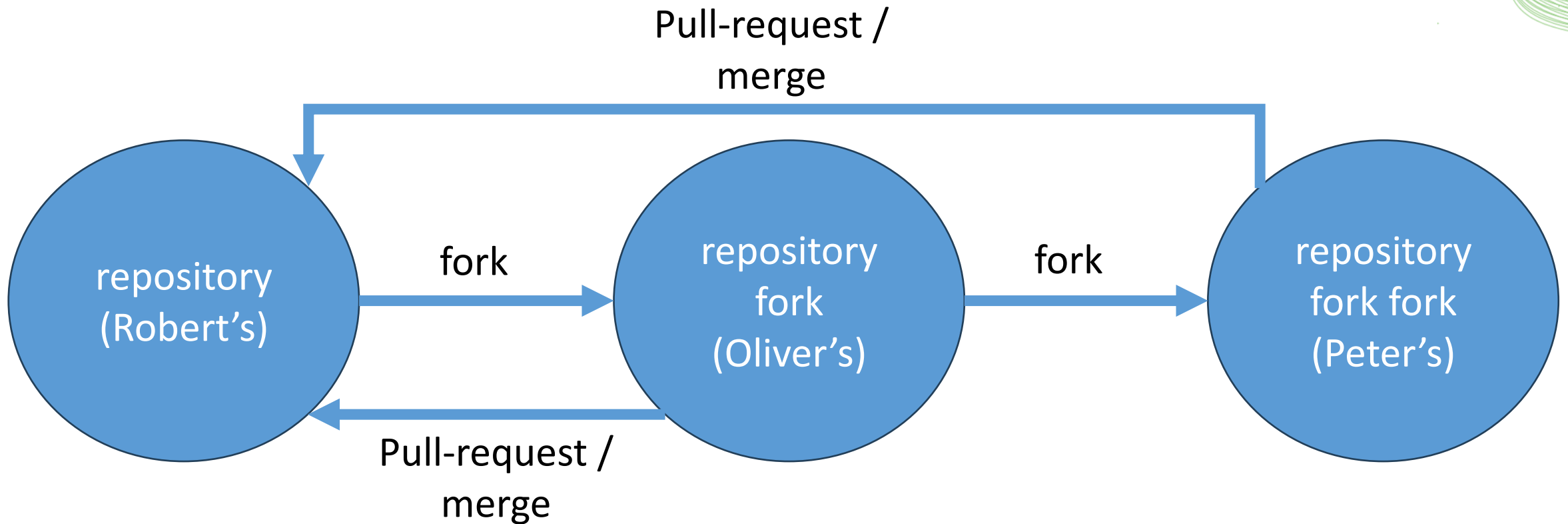
skimage\measure\_ccomp.pyx in skimage.measure._ccomp.label_cython()

skimage\measure\_ccomp.pyx in skimage.measure._ccomp.get_shape_info()

NotImplementedError: Only for images of dimension 1-3 are supported, got a 0
D one
```

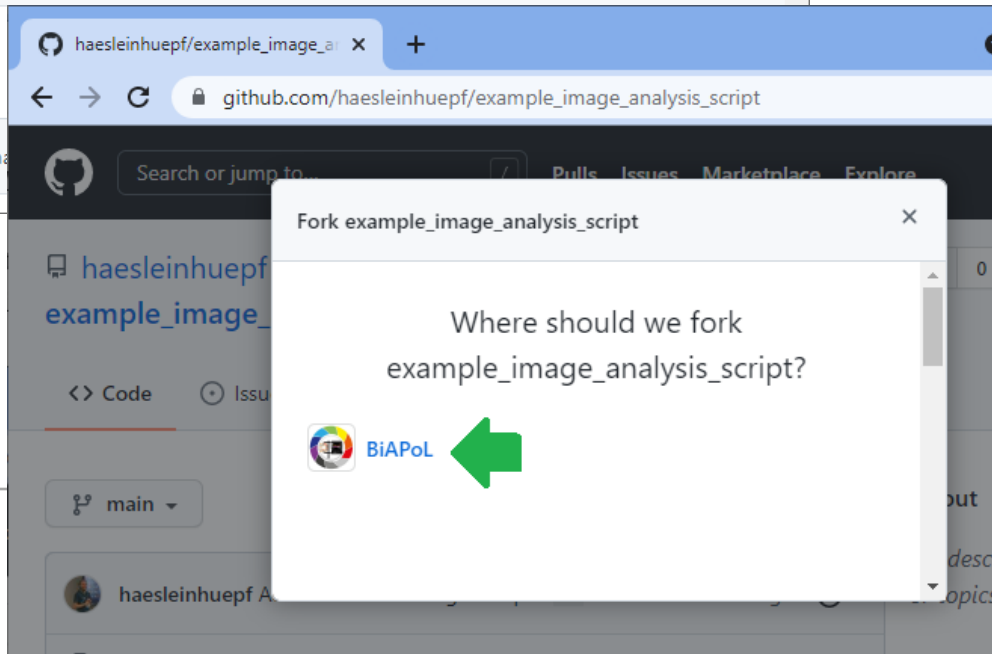
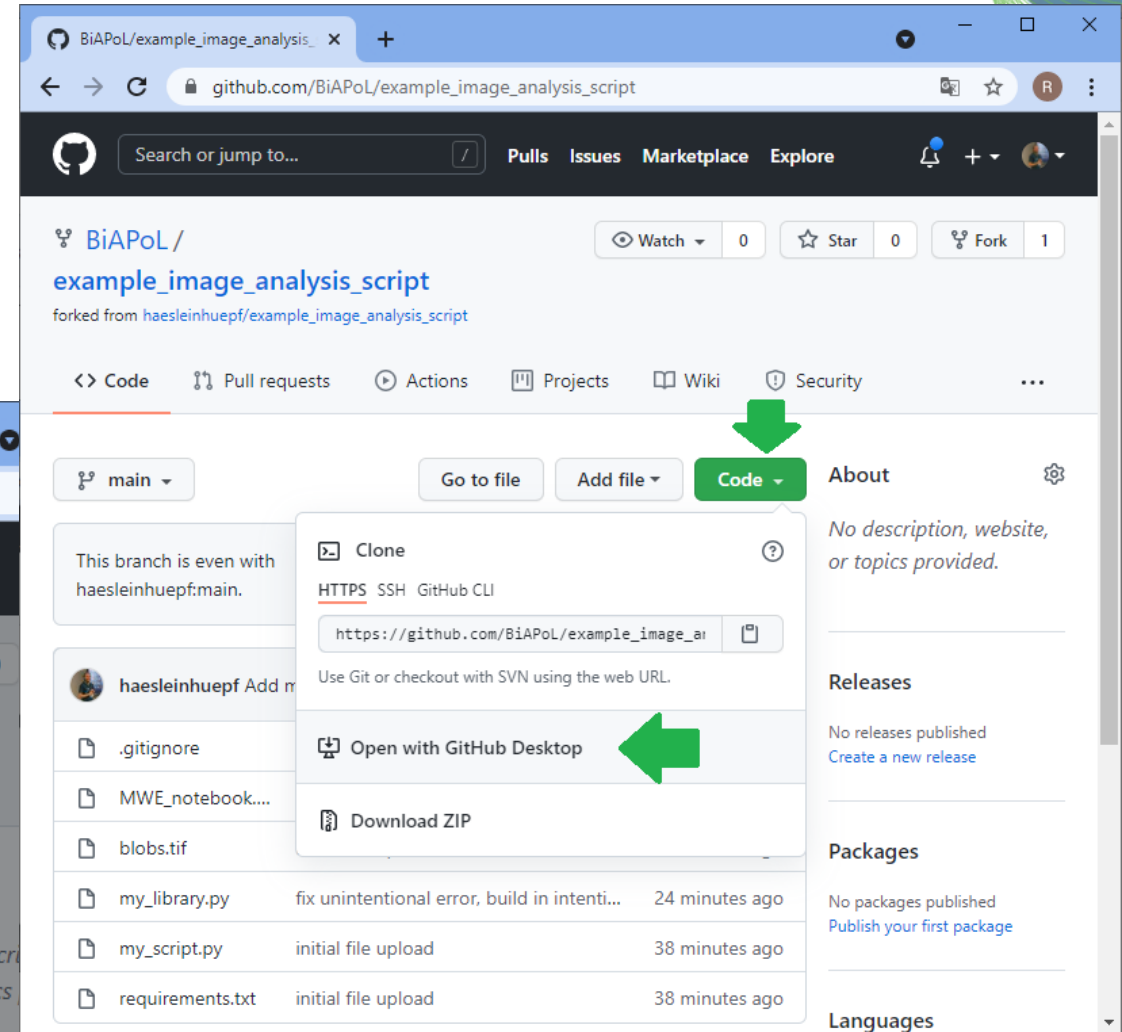
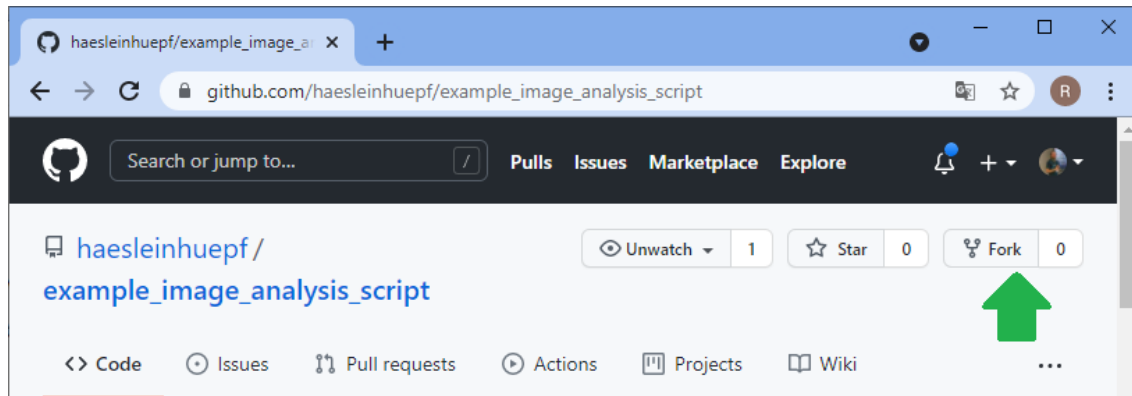
git - forking

- Making a copy where we have edit rights



github - forking

- Making a copy where we have edit rights



github – uploading (again)

- After fixing a bug, we upload the changes to our fork

File Edit View Repository Branch Help

Current repository: example_image_analysis_script | Current branch: main | Fetch origin: Never fetched

Changes: 2 changed files

- MWE_notebook.ipynb
- my_library.py

```
@@ -6,7 +6,8 @@ def segment_image(image):
6 6     blurred = gaussian(image, sigma=2)
7 7
8 8     # binarize the image
9 9     - binary = threshold_otsu(blurred)
10 10    + threshold = threshold_otsu(blurred)
11 11    + binary = blurred > threshold
12 12
13 13     # label connected components
14 14     result = label(binary)
@@ @@
```

Commit message: bugfix: threshold_otsu

threshold_otsu delivers a number (the threshold), not a binary image. For thresholding the image, an additional step is necessary.

Commit to main

File Edit View Repository Branch Help

Current repository: example_image_analysis_script | Current branch: main | Push origin: Last fetched just now

Changes: 0 changed files

No local changes

There are no uncommitted changes in this repository. Here are some friendly suggestions for what to do next.

Push commits to the origin remote

You have 1 local commit waiting to be pushed to GitHub.

Always available in the toolbar when there are local commits waiting to be pushed or Ctrl | P

Push origin

View the files of your repository in Explorer

Repository menu or Ctrl | Shift | F

Show in Explorer

Open the repository page on GitHub in your browser

Repository menu or Ctrl | Shift | G

View on GitHub

Summary (required)

Description

Commit to main

Committed just now

bugfix: threshold_otsu

Undo

Github – pull requests

- Contribute to open-source projects

The screenshot shows the GitHub repository page for 'BiAPoL/example_image_analysis_script', which is a fork of 'haesleinhuepf/example_image_analysis_script'. The repository is currently on the 'main' branch. A green arrow points to the 'Contribute' button, which has opened a dropdown menu. Another green arrow points to the 'Open pull request' button in the dropdown menu. The repository page includes a file list with columns for file name, description, and time since last update.

File Name	Description	Time
blobs.tif	initial file upload	1 hour ago
my_library.py	bugfix: threshold_otsu	1 minute ago
my_script.py	initial file upload	1 hour ago
requirements.txt	initial file upload	1 hour ago

The screenshot shows the 'Open a pull request' form in GitHub. The form is titled 'bugfix: threshold_otsu'. The base repository is 'haesleinhuepf/example_image_analysis_script' and the base branch is 'main'. The head repository is 'BiAPoL/example_image_analysis_script' and the compare branch is 'main'. A green checkmark indicates that the pull request is 'Able to merge'. The form includes a text area for the pull request description, which contains the following text: 'Dear Robert, here comes a bug fix for your image segmentation function. threshold_otsu delivers a number (the threshold), not a binary image. For thresholding the image, an additional step is necessary.' The form also includes a 'Create pull request' button.

Github – pull requests

- Reviewer perspective

github.com/haesleinhuepf/example_image_analysis_script/pull/1/files

haesleinhuepf / example_image_analysis_script

bugfix: threshold_otsu #1

haesleinhuepf wants to merge 1 commit into haesleinhuepf:main from BIAPOL:main

Files changed 2 +13 -20

```
@@ -6,7 +6,8 @@ def segment_image(image):
6 6     blurred = gaussian(image, sigma=2)
7 7
8 8     # binarize the image
9 -     binary = threshold_otsu(blurred)
9 +     threshold = threshold_otsu(blurred)
10 +     binary = blurred > threshold
11
11 12     # label connected components
12 13     result = label(binary)
```

bugfix: threshold_otsu #1

haesleinhuepf wants to merge 1 commit into haesleinhuepf:main from BIAPOL:main

Continuous integration has been enabled for this repository. This will automatically run GitHub Actions and several other checks on every push to this branch to help catch bugs and enforce style.

This branch has no conflicts with the branch you're merging. Merging can be performed automatically.

Merge pull request

Dear Robert,

here comes a pull request that delivers a number of improvements to the image, an additional feature, and a bugfix.

Best,
Robert

Thank you Robert! That's great 🐱

bugfix: threshold_otsu #1

haesleinhuepf wants to merge 1 commit into haesleinhuepf:main from BIAPOL:main

Merged

Dear Robert,

here comes a pull request that delivers a number of improvements to the image, an additional feature, and a bugfix.

Best,
Robert

example_image_analysis_script / MWE_notebook

haesleinhuepf bugfix: threshold_otsu ...

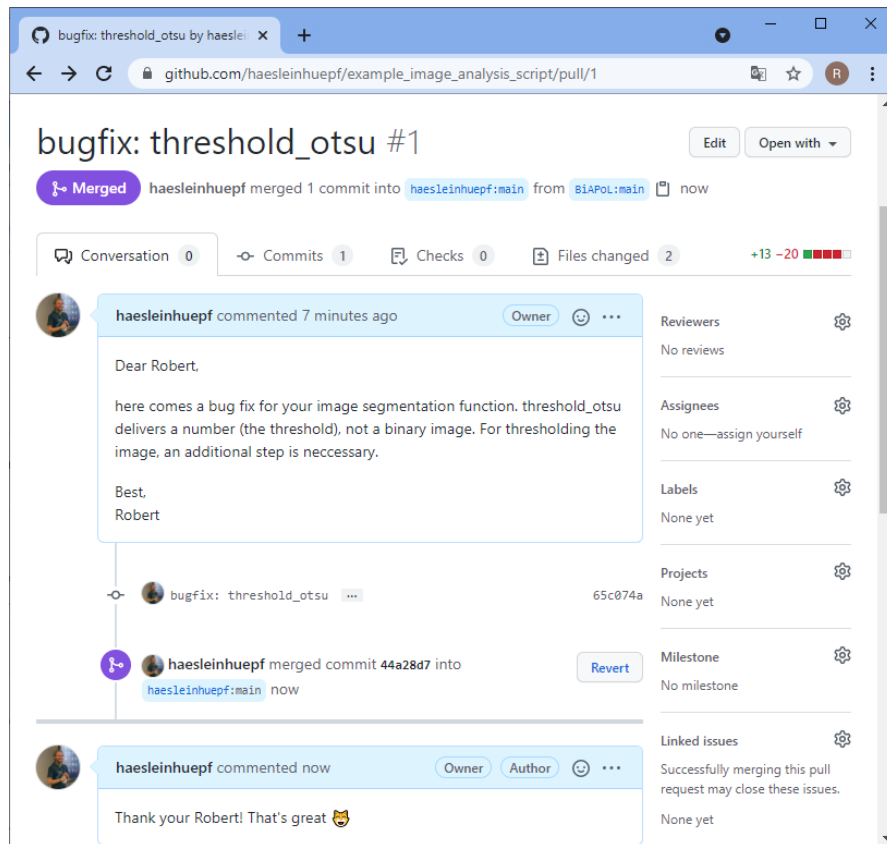
1 contributor

87 lines (87 sloc) | 1.56 KB

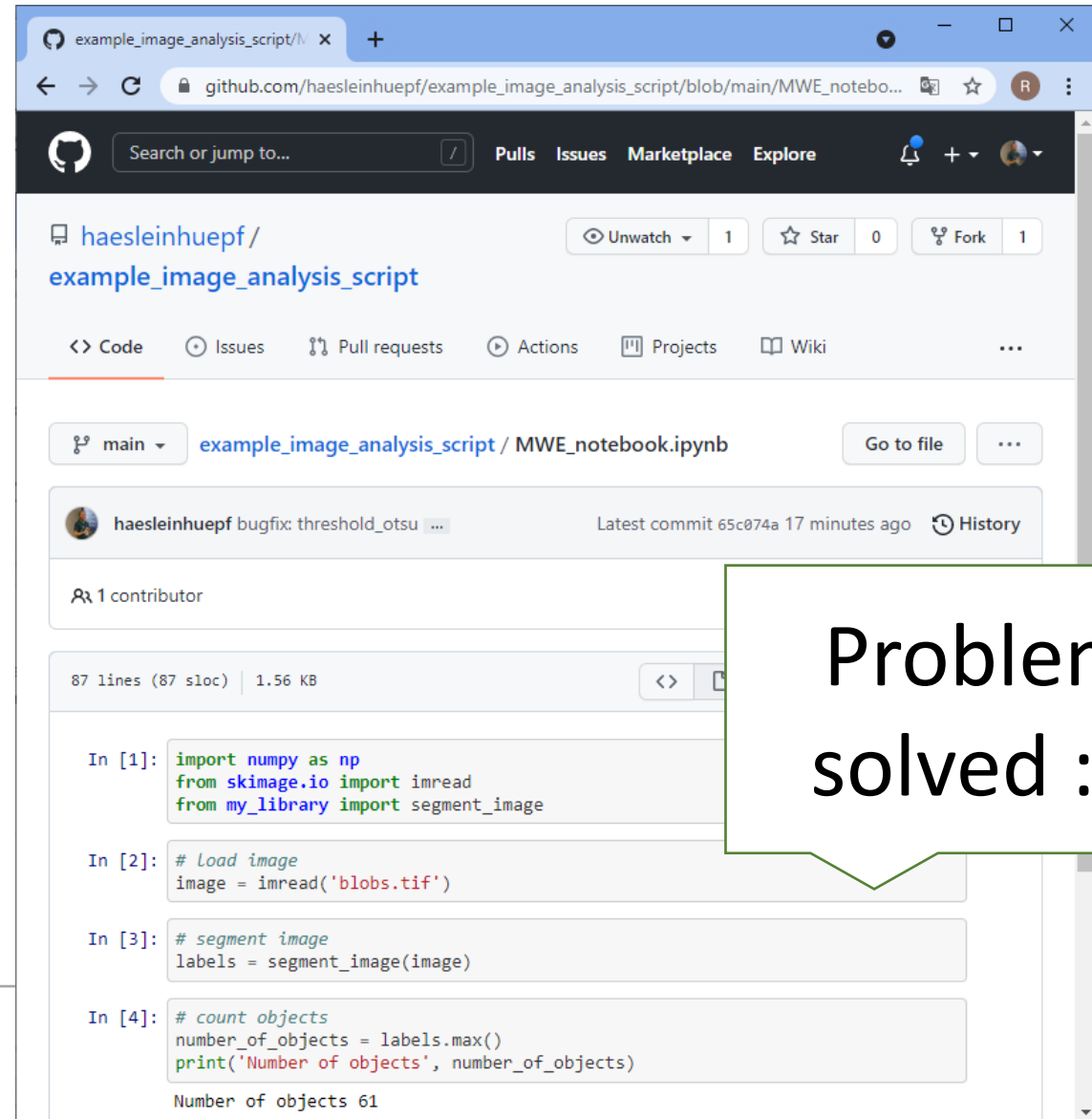
```
In [1]: import numpy as np
```

Github – pull requests

- Reviewer perspective



The screenshot shows a pull request titled "bugfix: threshold_otsu #1" in a repository named "example_image_analysis_script". The pull request is marked as "Merged" and shows that the author, haesleinhuepf, merged 1 commit into the main branch. The reviewer, Robert, has commented: "Dear Robert. here comes a bug fix for your image segmentation function. threshold_otsu delivers a number (the threshold), not a binary image. For thresholding the image, an additional step is necessary. Best, Robert". The pull request details show 1 commit, 2 files changed, and 13 additions and 20 deletions. The reviewer's comment is highlighted in blue.



The screenshot shows the same pull request from the author's perspective. The repository name is "haesleinhuepf / example_image_analysis_script". The pull request title is "bugfix: threshold_otsu" and it shows the latest commit 65c074a from 17 minutes ago. The pull request is for the file "MWE_notebook.ipynb" in the "main" branch. The code snippet shows the following Python code:

```
In [1]: import numpy as np
        from skimage.io import imread
        from my_library import segment_image

In [2]: # Load image
        image = imread('blobs.tif')

In [3]: # segment image
        labels = segment_image(image)

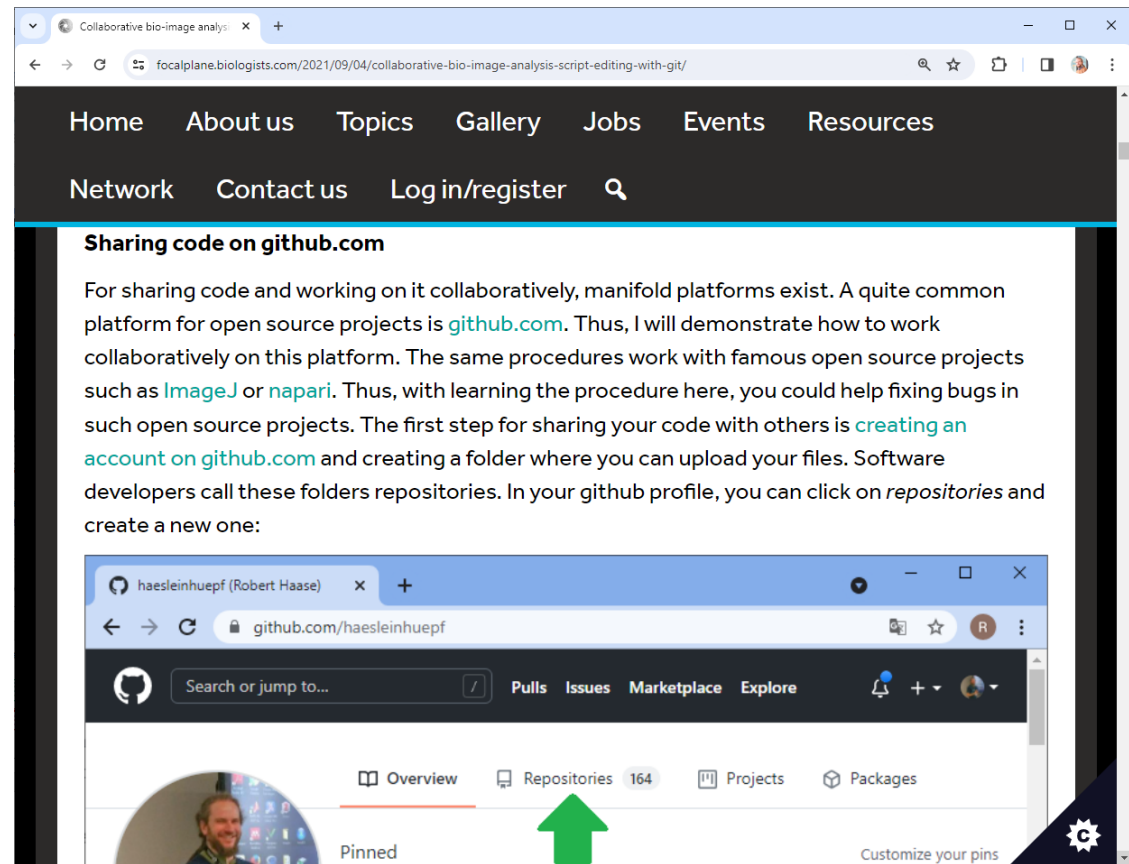
In [4]: # count objects
        number_of_objects = labels.max()
        print('Number of objects', number_of_objects)

Number of objects 61
```

A green callout box with the text "Problem solved :-)" is overlaid on the right side of the code snippet.

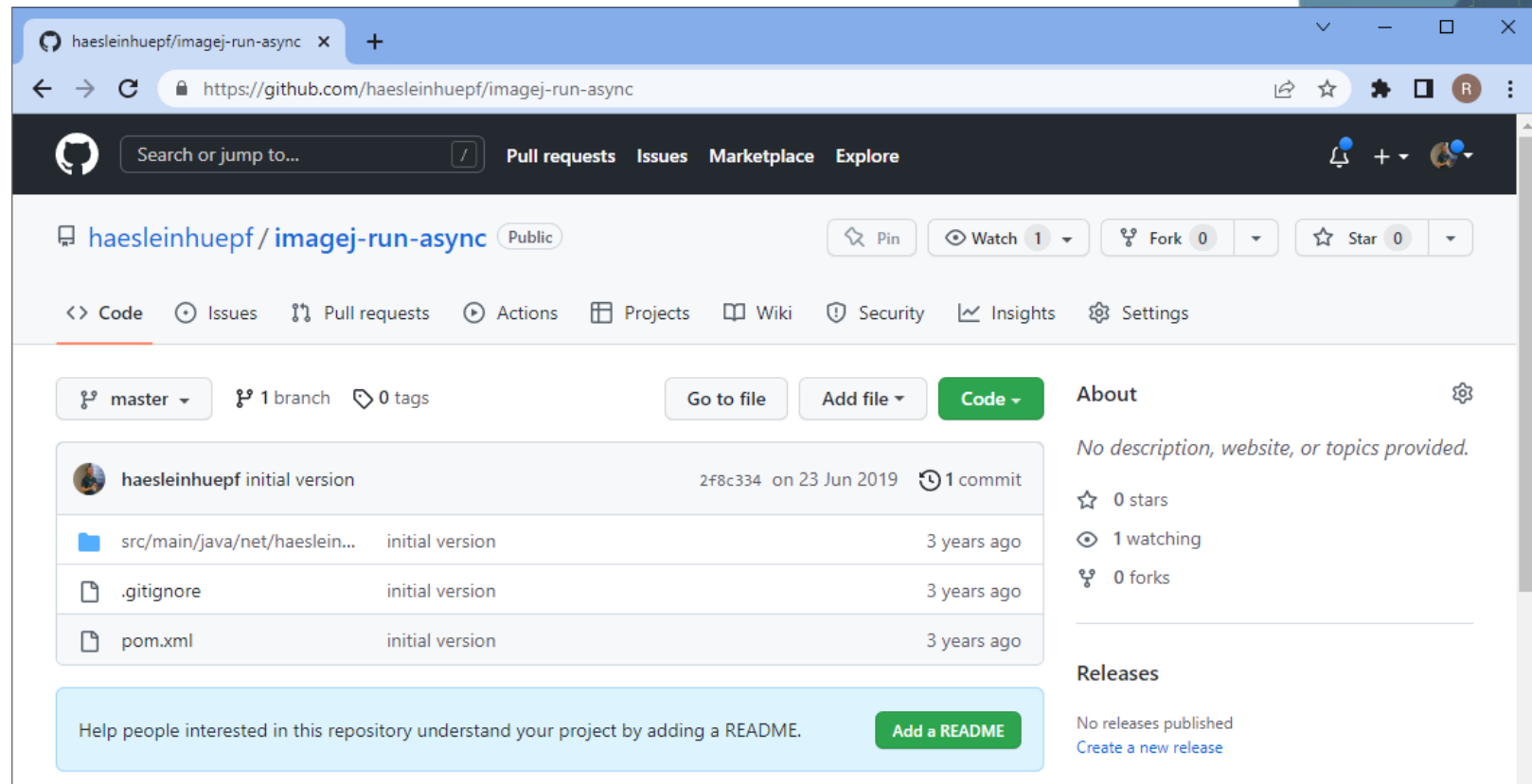
Github

- If this was too fast...



Quiz

It's ok to reuse this code if ...



Mention
author



Ask the
authors



Link to the
license



Copy the
copyright
statement



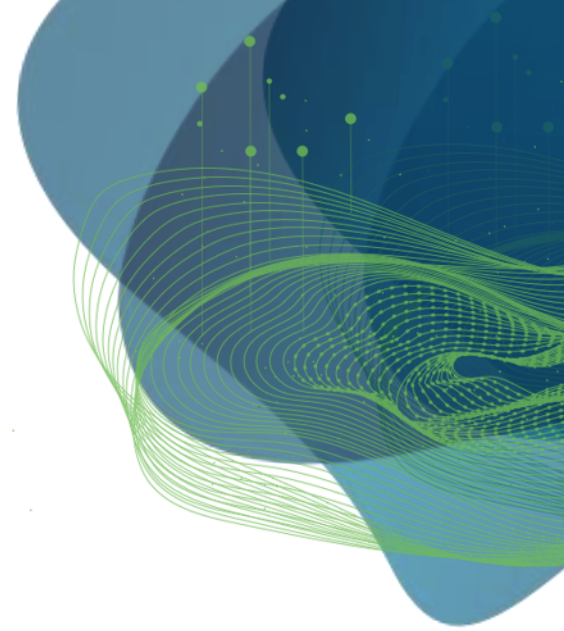
ScaDS.AI

DRESDEN LEIPZIG

CENTER FOR SCALABLE DATA ANALYTICS
AND ARTIFICIAL INTELLIGENCE

Exercises

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.

Exercises

ScaDS / BIDS-lecture-2024

Training resources for Students at Uni Leipzig who want to dive into bio-image Python. The material will develop between April and July 2024

CC-BY-4.0 license

1 star 0 forks 2 watching 1 Branch 0 Tags Activity

Public repository

main

haesleinhuepf added first exercise

- 01a_setting_up_local_environment added first exercise
- 01b_setting_up_sc_ulei_environment added first exercise
- 01c_testing_environment added first exercise
- .gitignore added first exercise
- LICENSE-CC-BY The Beginning
- README.md added first exercise

Bio-image Data Science

This repository contains training resources for Students at Uni Leipzig who want to dive into bio-image data science with Python. The material will develop between April and July 2024 and shared here in this github repository.

Teaching Goal

Students learn the full workflow of common bio-image data science projects to a degree that they can execute a scientific data analysis project in this context on their own. They will be familiar with common bio-image analysis algorithms and workflows, how to choose them according to a scientific goal, and how to measure quality of derived results. Attending the lecture and executing the practicals qualifies the students to work as bio-image data scientist in the pharmaceutical industry or basic biological research.

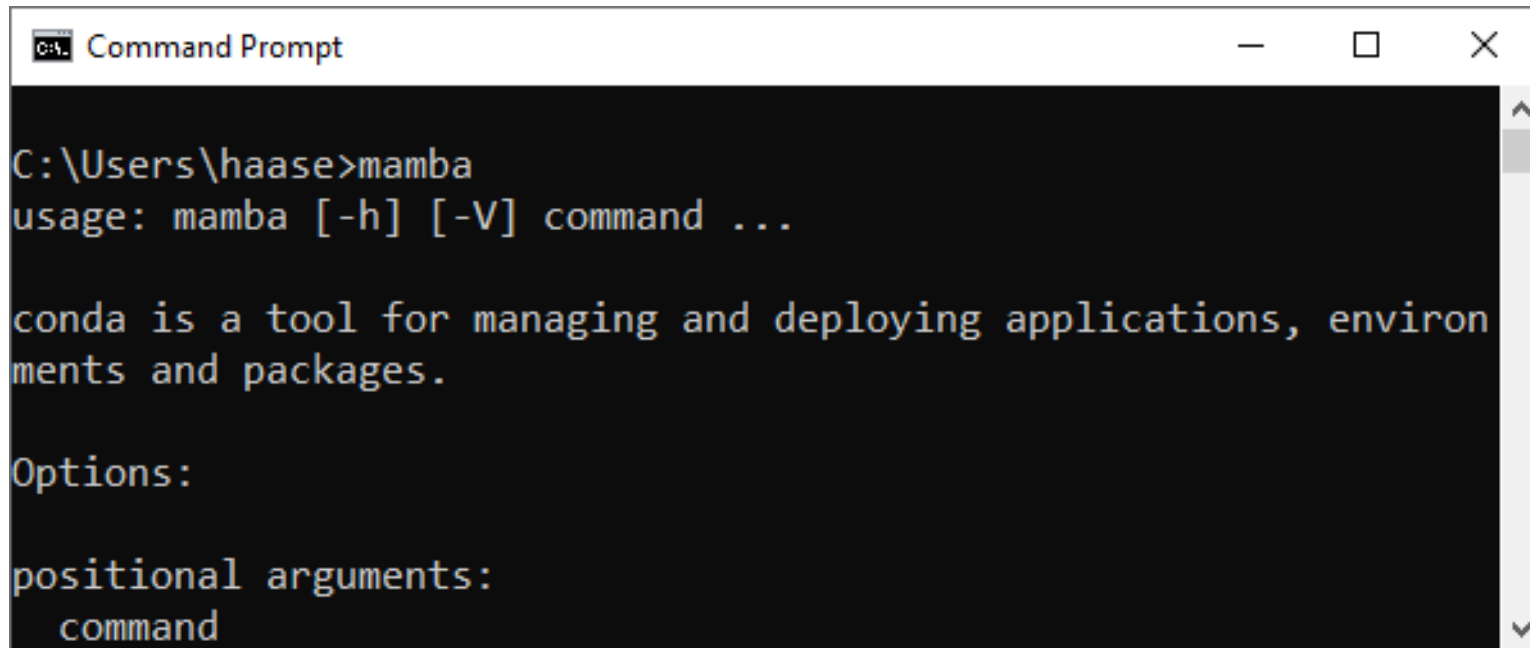
Course contents

- Introduction to Bio-image Data Science (Apr 2nd 2024)
 - Basics of microscopy
 - Introduction to Bio-image Analysis
 - Exercises:
 - Setting up a local environment
 - Setting up Jupyter Hub at Scientific Computing / Leipzig University
 - Execute the trailer notebook



Exercise (recap)

- Make sure `mamba` is installed on your computer (see instructions from last week)



```
Command Prompt
C:\Users\haase>mamba
usage: mamba [-h] [-V] command ...

conda is a tool for managing and deploying applications, environ-
ments and packages.

Options:

positional arguments:
  command
```

Exercise (BioImage Archive)

- Download a dataset from the BioImage Archive

The image displays three overlapping screenshots of a JupyterLab interface, illustrating the workflow for downloading a dataset from the BioImage Archive.

- Left Screenshot: Exploring the BioImage Archive**
 - Text: "In this notebook we use the `bia-explorer` project to explore the `S-BIAD634` dataset in the BioImage Archive. We will download some images and store them in a local directory."
 - Code cell [1]:

```
from bia_explorer import io, biostudies
from skimage.io import imread, imsave
from IPython.display import display, Markdown
import stackview
```
 - Section: "Accessing meta-data"
 - Text: "First we access the meta data of the datasets. Here we can for example see what the data under which license it can be used."
 - Code cell [2]:

```
accession = 'S-BIAD634'
submission = biostudies.load_submission(accession)
```
- Middle Screenshot: Visualizing images**
 - Text: "A single image can be loaded and shown like this (See also)."
 - Code cell [4]:

```
image = study.images[0]
image
```
 - Code cell [4]:

```
BIAImage(uri='https://www.ebi.ac.uk/biostudies/files/S-BIAD634/dataset/ganglioneuroblastoma_0.tif', size=2239668, fpath=WindowsPath('dataset/groundtruth/Ganglioneuroblastoma_0.tif'))
```
 - Code cell [5]:

```
uri = image.uri.replace("\\", "/")
image_data = imread(uri)
stackview.insight(image_data)
```
 - Code cell [5]:

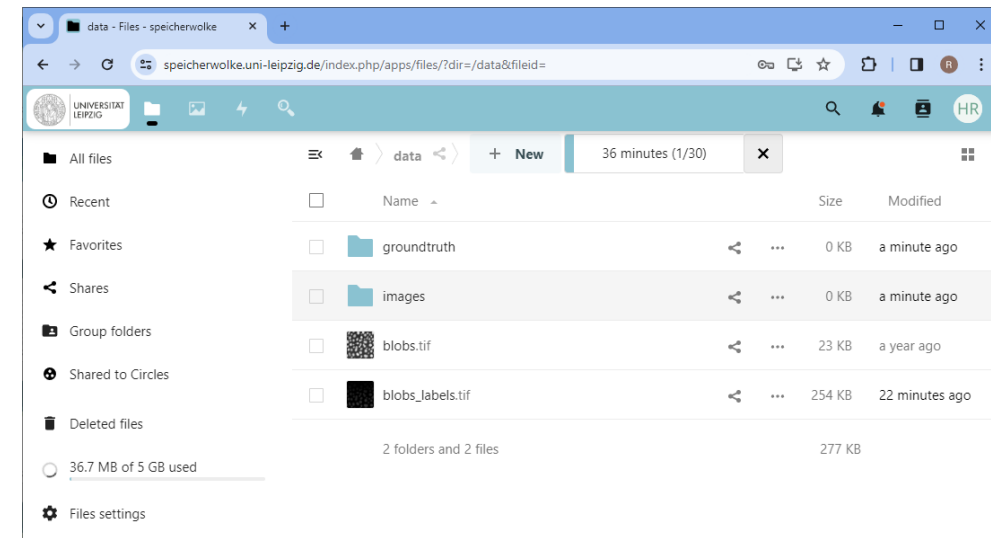
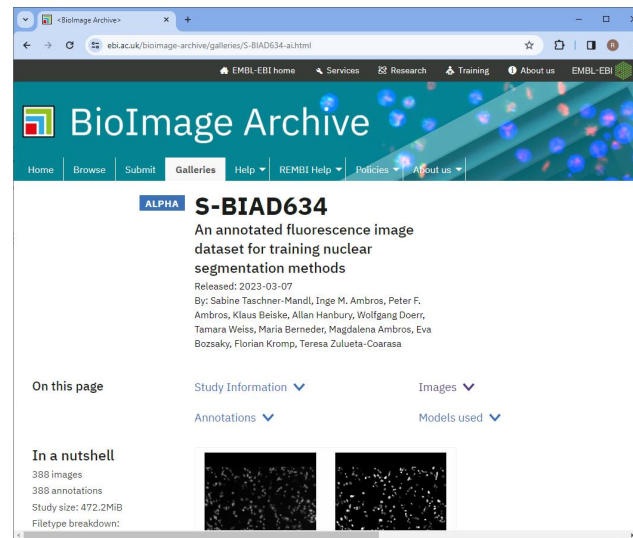
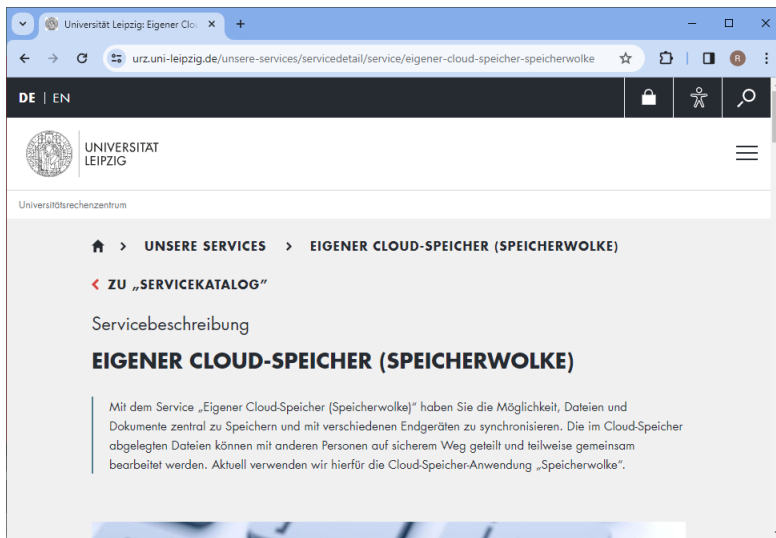
```
shape (300, 300, 3)
dtype uint8
size 270000
min 0
```
 - Visual output: A grayscale image showing a dense field of small white spots on a black background, representing a histological section of neuroblastoma tissue.
- Right Screenshot: Exercise**
 - Code cell [10]:

```
for f in os.listdir(groundtruth_folder):
    print(f)
```
 - Output: A list of files:

```
Ganglioneuroblastoma_0.tif
Ganglioneuroblastoma_1.tif
Ganglioneuroblastoma_10.tif
Ganglioneuroblastoma_2.tif
Ganglioneuroblastoma_3.tif
Ganglioneuroblastoma_4.tif
Ganglioneuroblastoma_6.tif
Ganglioneuroblastoma_7.tif
Ganglioneuroblastoma_8.tif
Ganglioneuroblastoma_9.tif
```
 - Section: "Exercise"
 - Text: "Download all images with 'Neuroblastoma' in their name and upload them to a folder in the owncloud. Do not download and upload files which already exist."
 - Code cell []: An empty code cell for the user to write their solution.

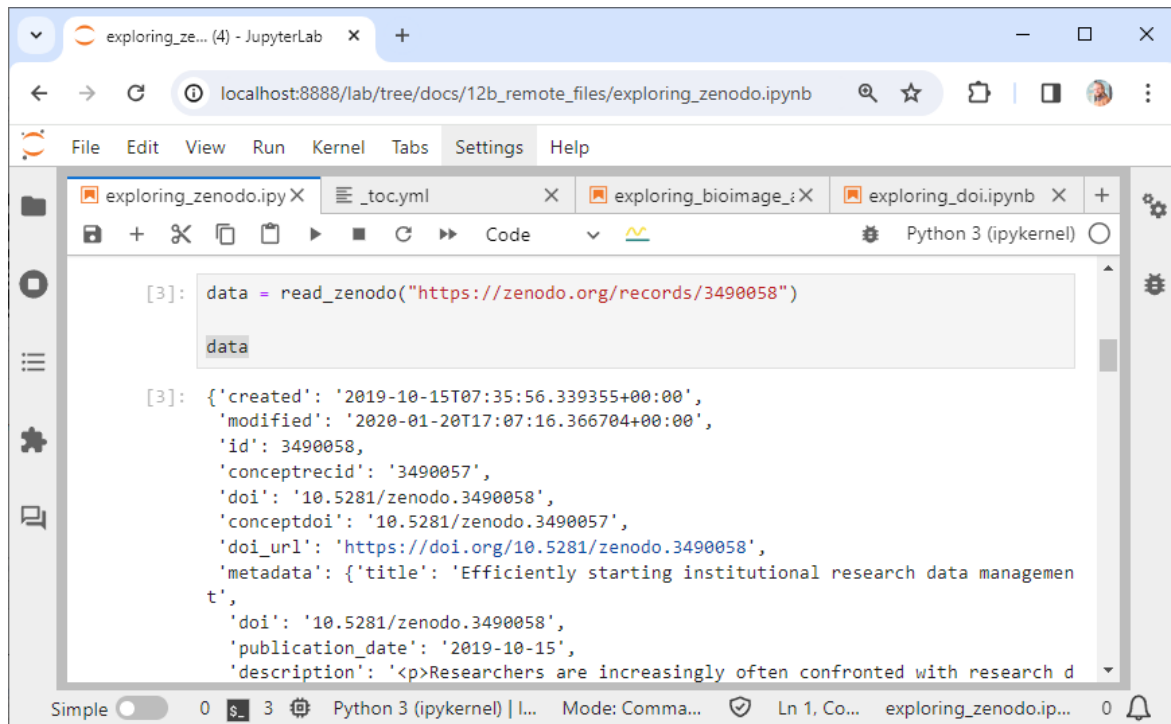
Exercise (nextcloud)

- Register at Speicherwolke @ Uni Leipzig,
- Upload the images from the BioImage Archive to a folder in the Speicherwolke.



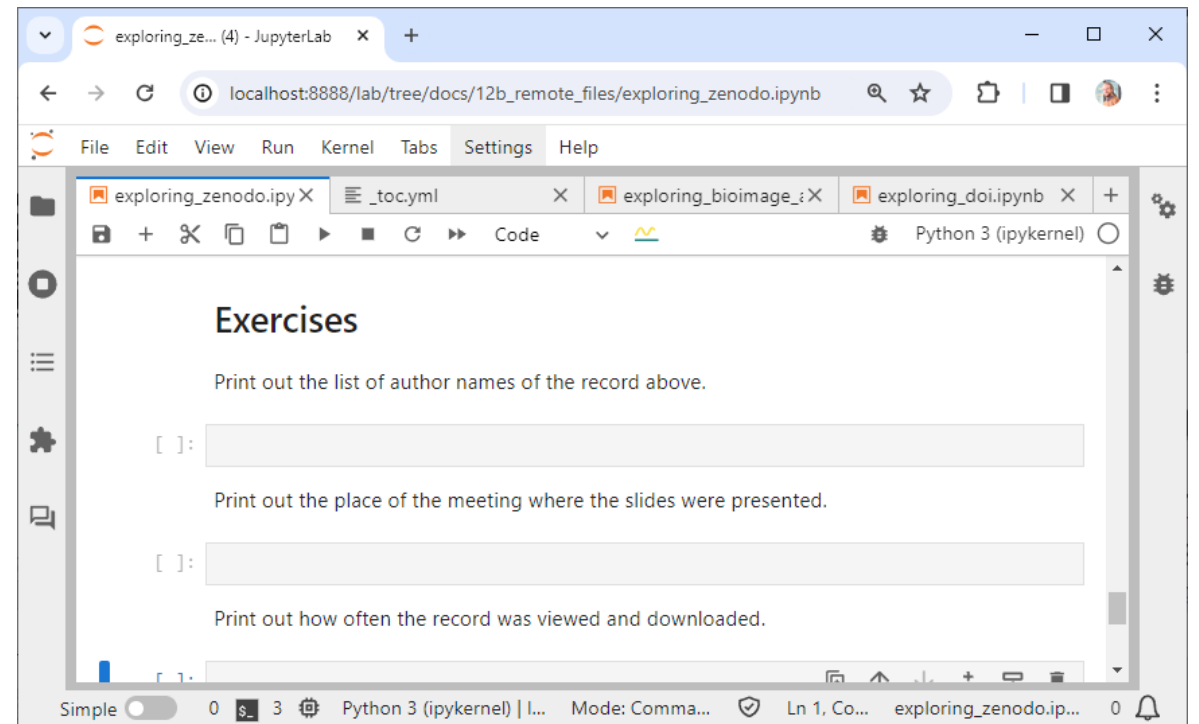
Exercise (Zendo and DOI)

- Explore the DOI and Zenodo APIs to find out the author of online records



```
[3]: data = read_zenodo("https://zenodo.org/records/3490058")
data

[3]: {'created': '2019-10-15T07:35:56.339355+00:00',
      'modified': '2020-01-20T17:07:16.366704+00:00',
      'id': 3490058,
      'conceptrecid': '3490057',
      'doi': '10.5281/zenodo.3490058',
      'conceptdoi': '10.5281/zenodo.3490057',
      'doi_url': 'https://doi.org/10.5281/zenodo.3490058',
      'metadata': {'title': 'Efficiently starting institutional research data management',
                  'doi': '10.5281/zenodo.3490058',
                  'publication_date': '2019-10-15',
                  'description': '<p>Researchers are increasingly often confronted with research d
```



Exercises

Print out the list of author names of the record above.

[]:

Print out the place of the meeting where the slides were presented.

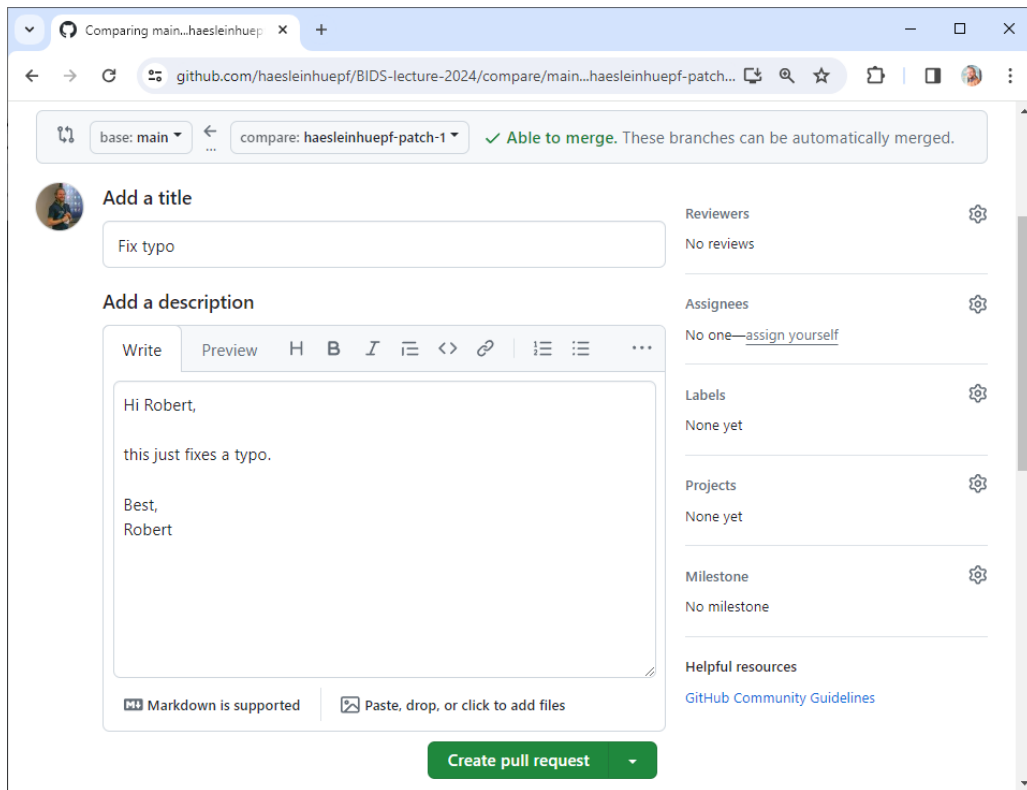
[]:

Print out how often the record was viewed and downloaded.

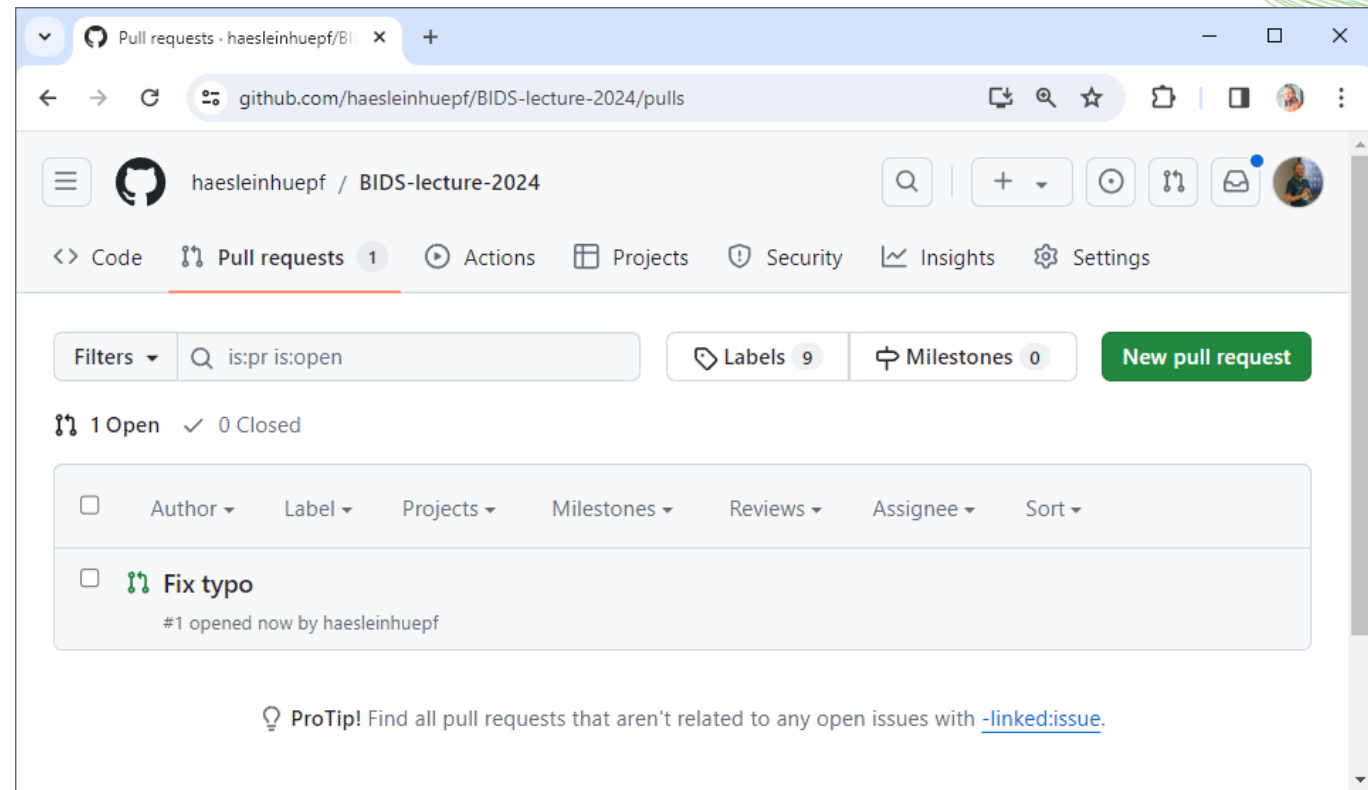
[]:

Exercise

- Clone the training materials repository
- Fix the typo on this page, send a pull-request



A screenshot of a web browser showing the GitHub interface for creating a pull request. The browser tab is titled "Comparing main...haesleinhuepf". The address bar shows the URL: github.com/haesleinhuepf/BIDS-lecture-2024/compare/main...haesleinhuepf-patch-1. The interface includes a "base: main" dropdown, a "compare: haesleinhuepf-patch-1" dropdown, and a green checkmark indicating "Able to merge. These branches can be automatically merged." Below this, there is a section for "Add a title" with a text input field containing "Fix typo". To the right of the title field are "Reviewers" (No reviews) and "Assignees" (No one—assign yourself) sections. The "Add a description" section has a rich text editor with the text: "Hi Robert, this just fixes a typo. Best, Robert". At the bottom right, there is a "Create pull request" button.



A screenshot of a web browser showing the GitHub interface for the pull requests page of the repository "haesleinhuepf / BIDS-lecture-2024". The browser tab is titled "Pull requests · haesleinhuepf/BIDS-lecture-2024". The address bar shows the URL: github.com/haesleinhuepf/BIDS-lecture-2024/pulls. The page shows a navigation bar with "Code", "Pull requests 1", "Actions", "Projects", "Security", "Insights", and "Settings". Below the navigation bar, there are filters for "is:pr is:open", "Labels 9", and "Milestones 0", along with a "New pull request" button. The main content area shows "1 Open" and "0 Closed" pull requests. A list of pull requests is displayed with a table header: "Author", "Label", "Projects", "Milestones", "Reviews", "Assignee", and "Sort". The first pull request is titled "Fix typo" and is marked as "#1 opened now by haesleinhuepf". At the bottom, there is a "ProTip!" section with the text: "Find all pull requests that aren't related to any open issues with [-linked:issue](#)."