Research Data Management Why and How?

23 July 2019

Yasemin Turkyilmaz-van der Velden

Data Steward @ TU Delft

y.turkyilmaz-vandervelden@tudelft.nl @YaseminTurkyilm

Slides are available: https://doi.org/10.5281/zenodo.3346559

TU Delft

Who am I?

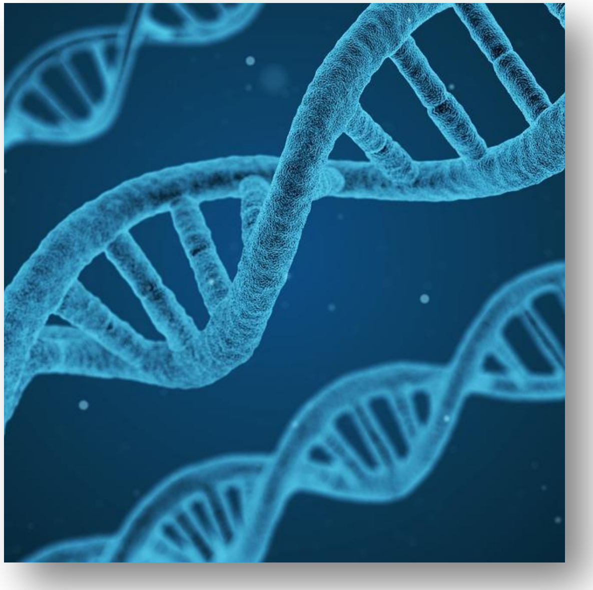
PhD candidate

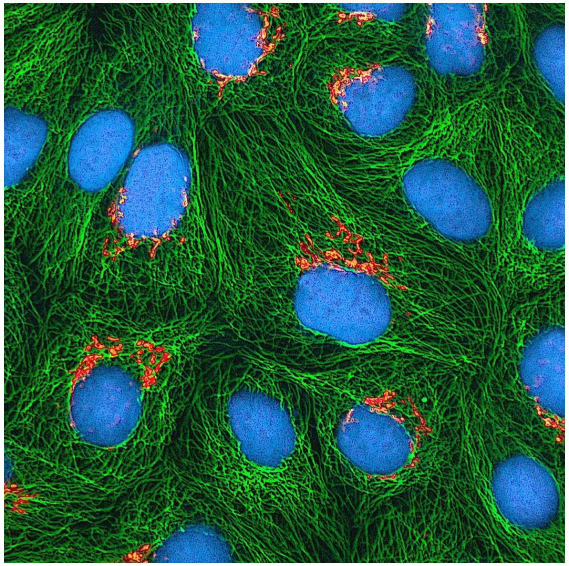
Department of Molecular Genetics

Erasmus MC Rotterdam

UV-induced DNA damage & repair

Proteomics & Microscopy

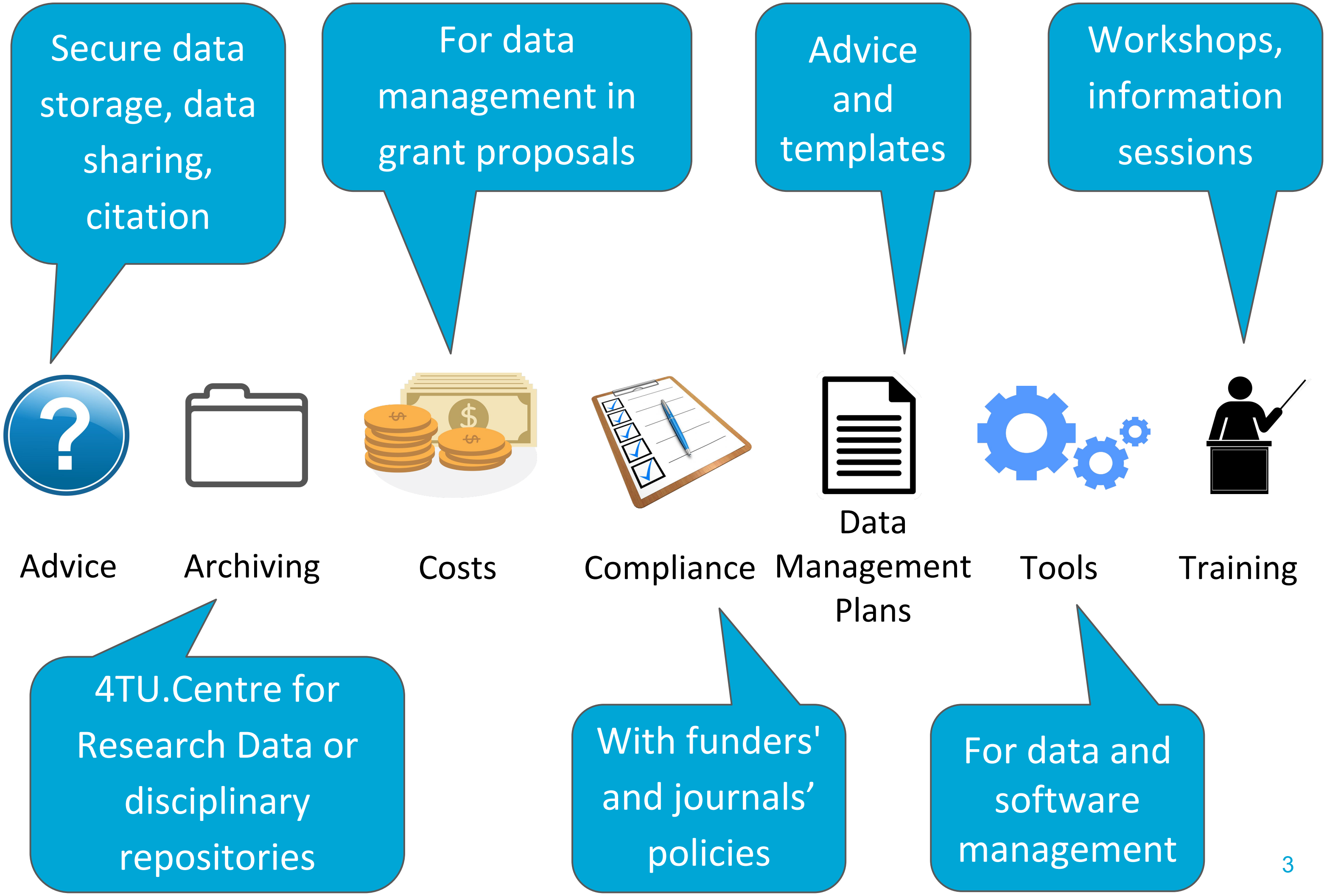




Who am I?

Data Steward @ TU Delft

www.tudelft.nl/library/datastewardship/



Secure data storage, data sharing, citation

For data management in grant proposals

Advice and templates

Workshops, information sessions

Advice

Archiving

Costs

Compliance

Data Management Plans

Tools

Training

4TU.Centre for Research Data or disciplinary repositories

With funders' and journals’ policies

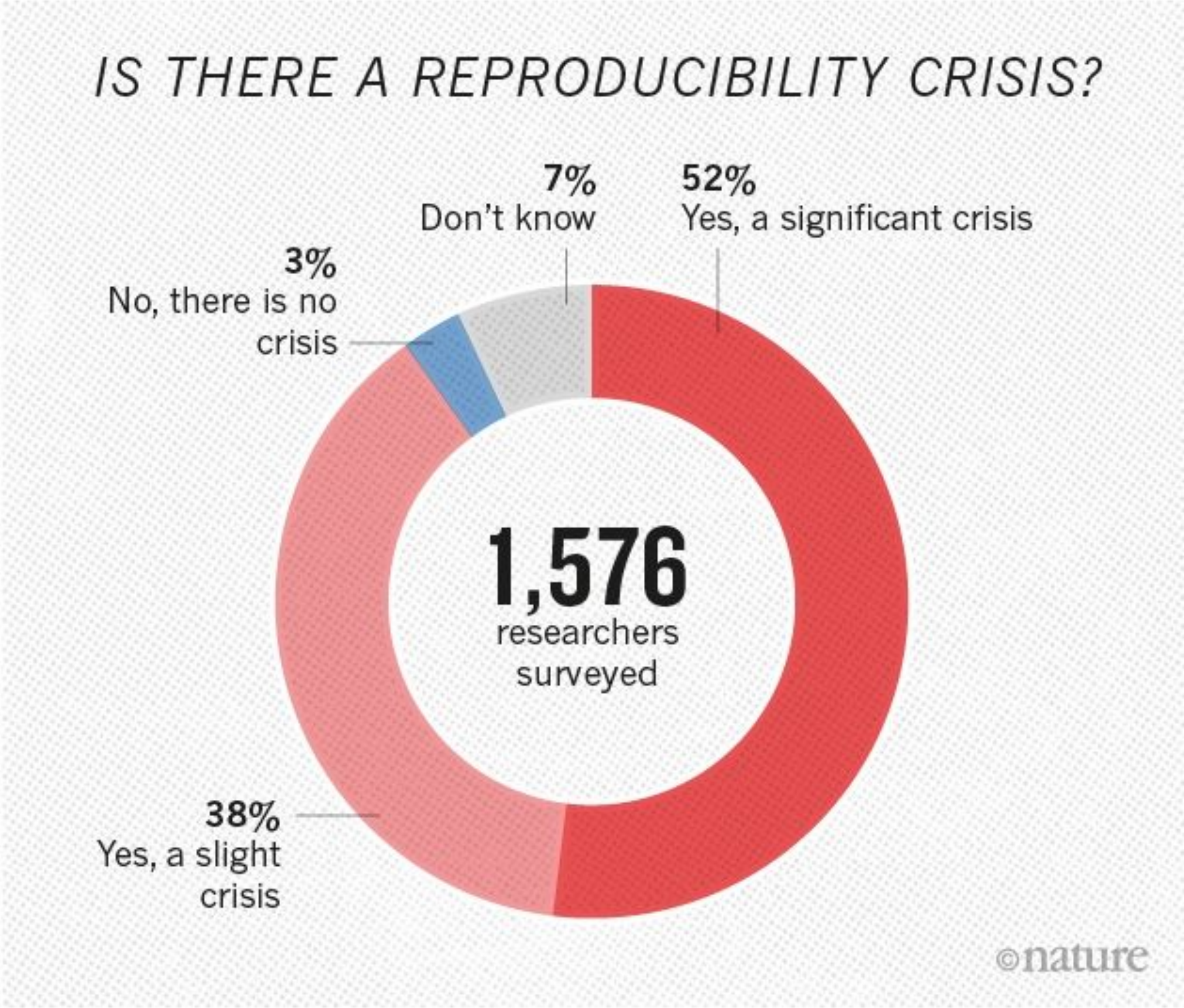
For data and software management

Outline

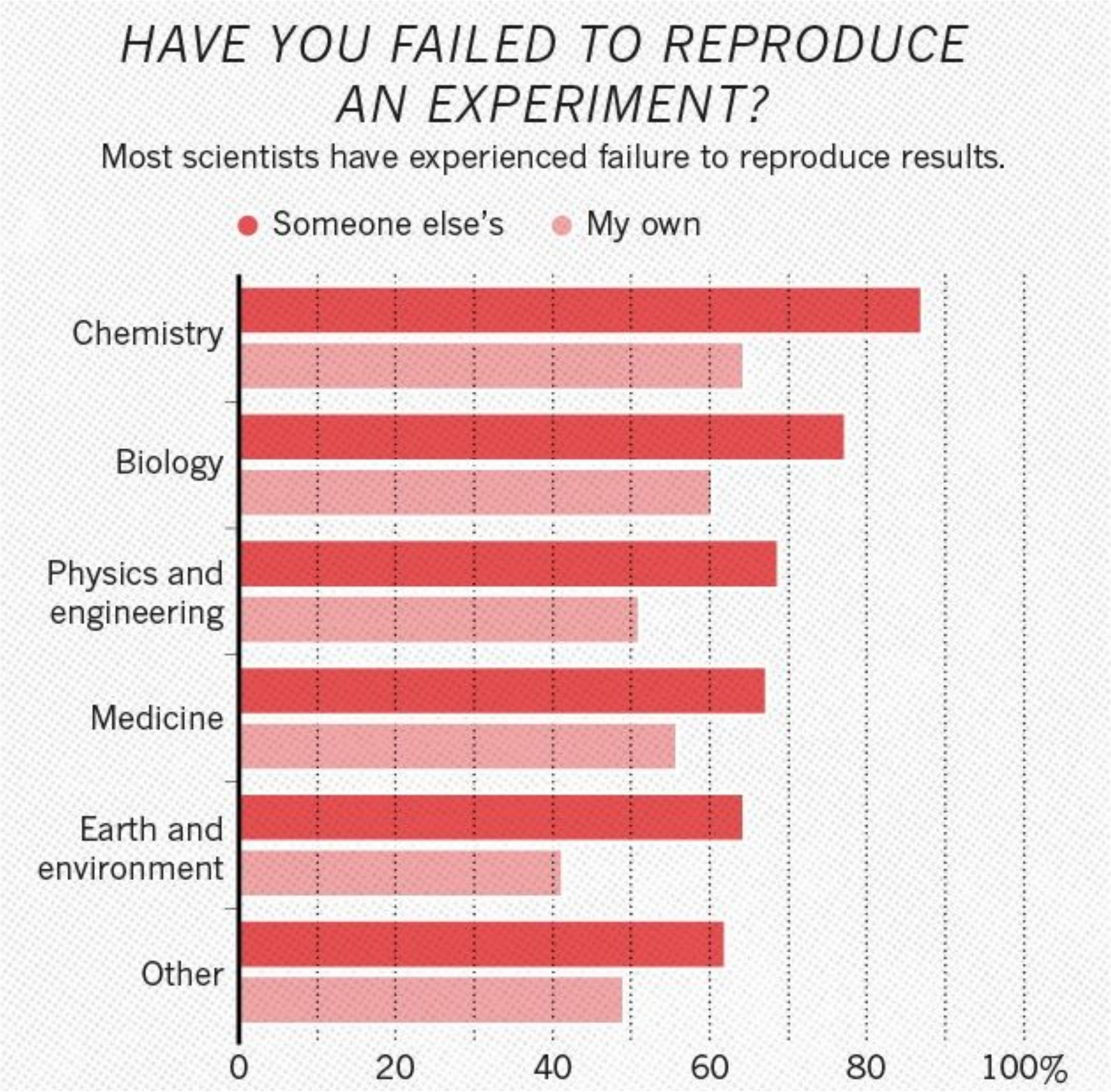
- Research Data management - Why?
  - Reproducibility Crisis
  - Funders’ and Publishers’ Requirements
  - Selfish benefits

- Research Data management - How?
  - Data archiving in repositories
  - Data documentation
  - Secure Data Storage and Backup
  - Data organization
  - Resources and training materials for good practices in scientific computing

Reproducibility crisis

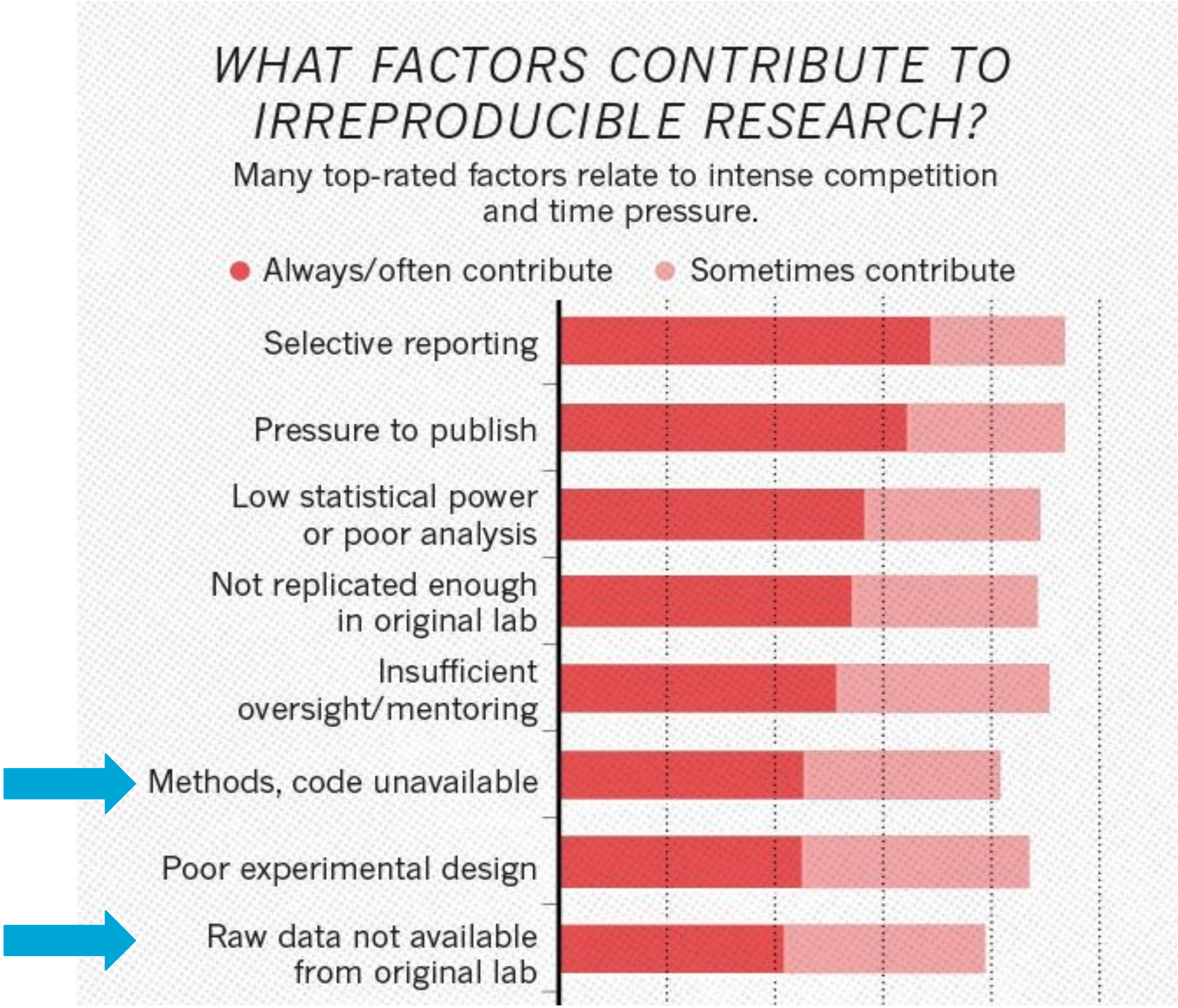


Nature 533, 452–454 (26 May 2016) doi:10.1038/533452a



Nature 533, 452–454 (26 May 2016) doi:10.1038/533452a

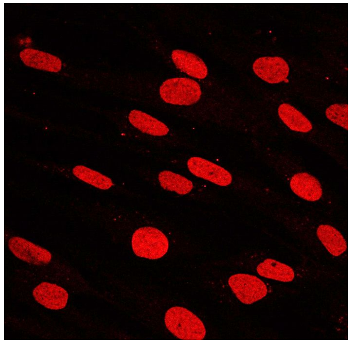
Reproducibility crisis



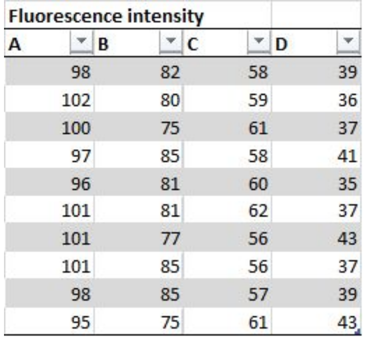
Nature 533, 452–454 (26 May 2016) doi:10.1038/533452a

A close look to the research data life cycle

Raw data



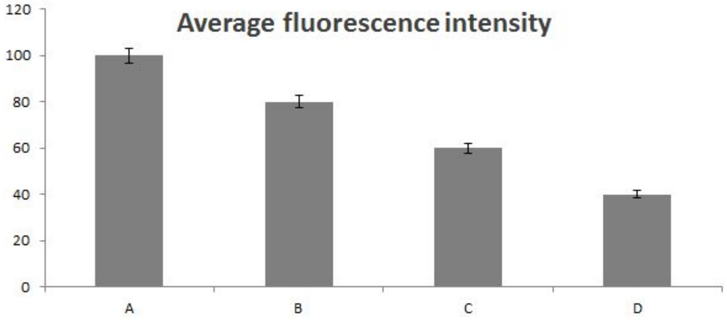
Intermediate data



Fluorescence intensity

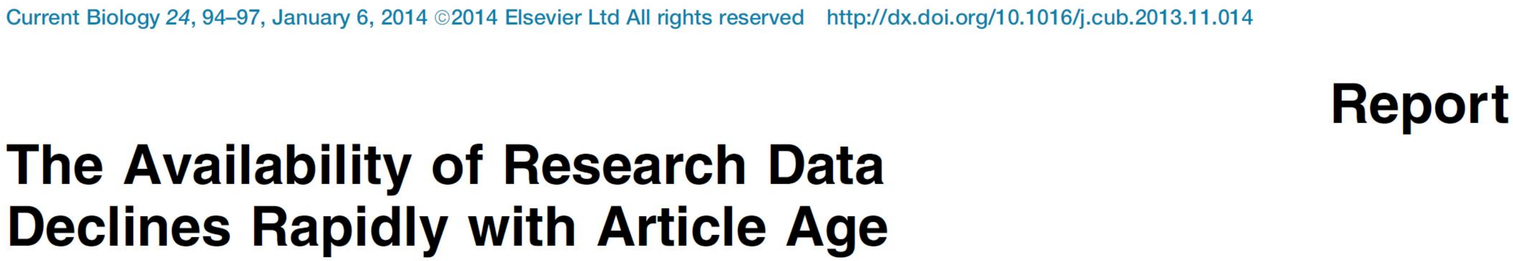
A	B	C	D
98	82	58	39
102	80	59	36
100	75	61	37
97	85	58	41
96	81	60	35
101	81	62	37
101	77	56	43
101	85	56	37
98	85	57	39
95	75	61	43

Final data



- Are the published final data available for validation, reproduction or reuse?
- What about experimental methods and measurement parameters?

Datasets available ‘on request’ are not available



Current Biology *24*, 94–97, January 6, 2014 ©2014 Elsevier Ltd All rights reserved http://dx.doi.org/10.1016/j.cub.2013.11.014

Report

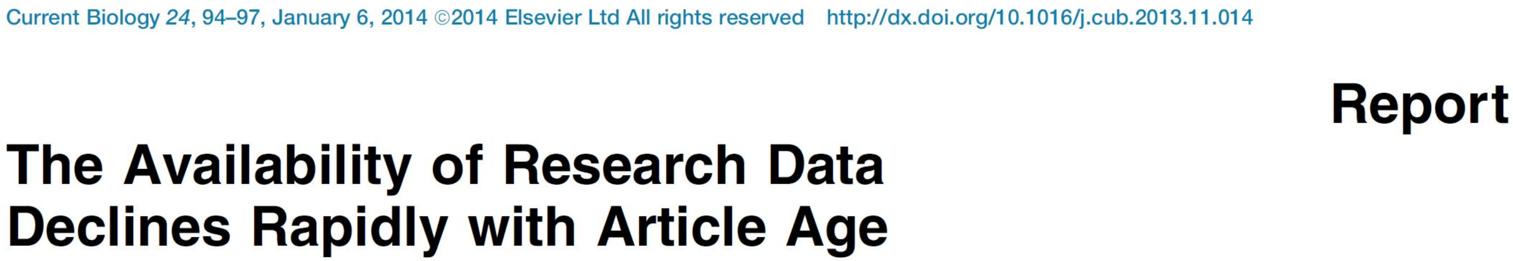
The Availability of Research Data Declines Rapidly with Article Age

- Data availability decreases by 17% per year

- Chance of email address working decreases by 7% per year

http://dx.doi.org/10.1016/j.cub.2013.11.014

Datasets available 'on request' are not available



Current Biology *24*, 94–97, January 6, 2014 ©2014 Elsevier Ltd All rights reserved http://dx.doi.org/10.1016/j.cub.2013.11.014

Report

The Availability of Research Data Declines Rapidly with Article Age

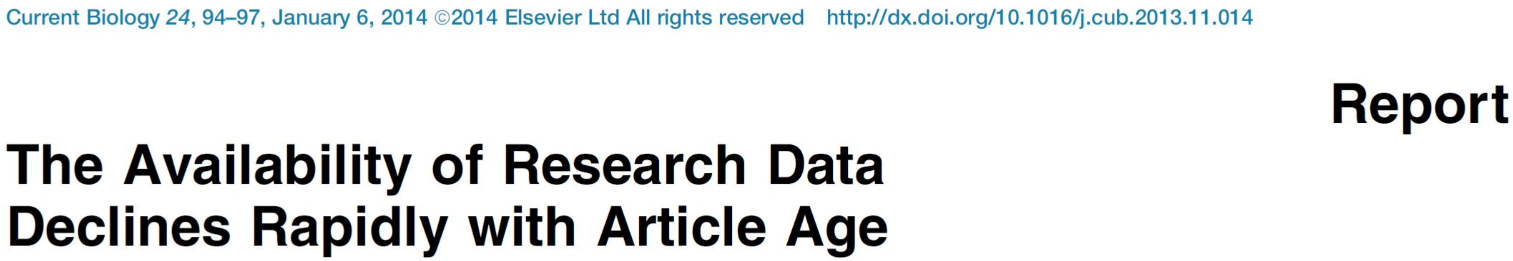
- Data availability decreases by 17% per year
- Chance of email address working decreases by 7% per year

ORCID https://orcid.org/

http://dx.doi.org/10.1016/j.cub.2013.11.014

TU Delft

Datasets available ‘on request’ are not available



Current Biology *24*, 94–97, January 6, 2014 ©2014 Elsevier Ltd All rights reserved http://dx.doi.org/10.1016/j.cub.2013.11.014

Report

The Availability of Research Data Declines Rapidly with Article Age

- Data availability decreases by 17% per year

- Chance of email address working decreases by 7% per year

What’s the alternative to sharing ‘on request’?

http://dx.doi.org/10.1016/j.cub.2013.11.014

Archiving in a repository

A place where things can be stored and shared

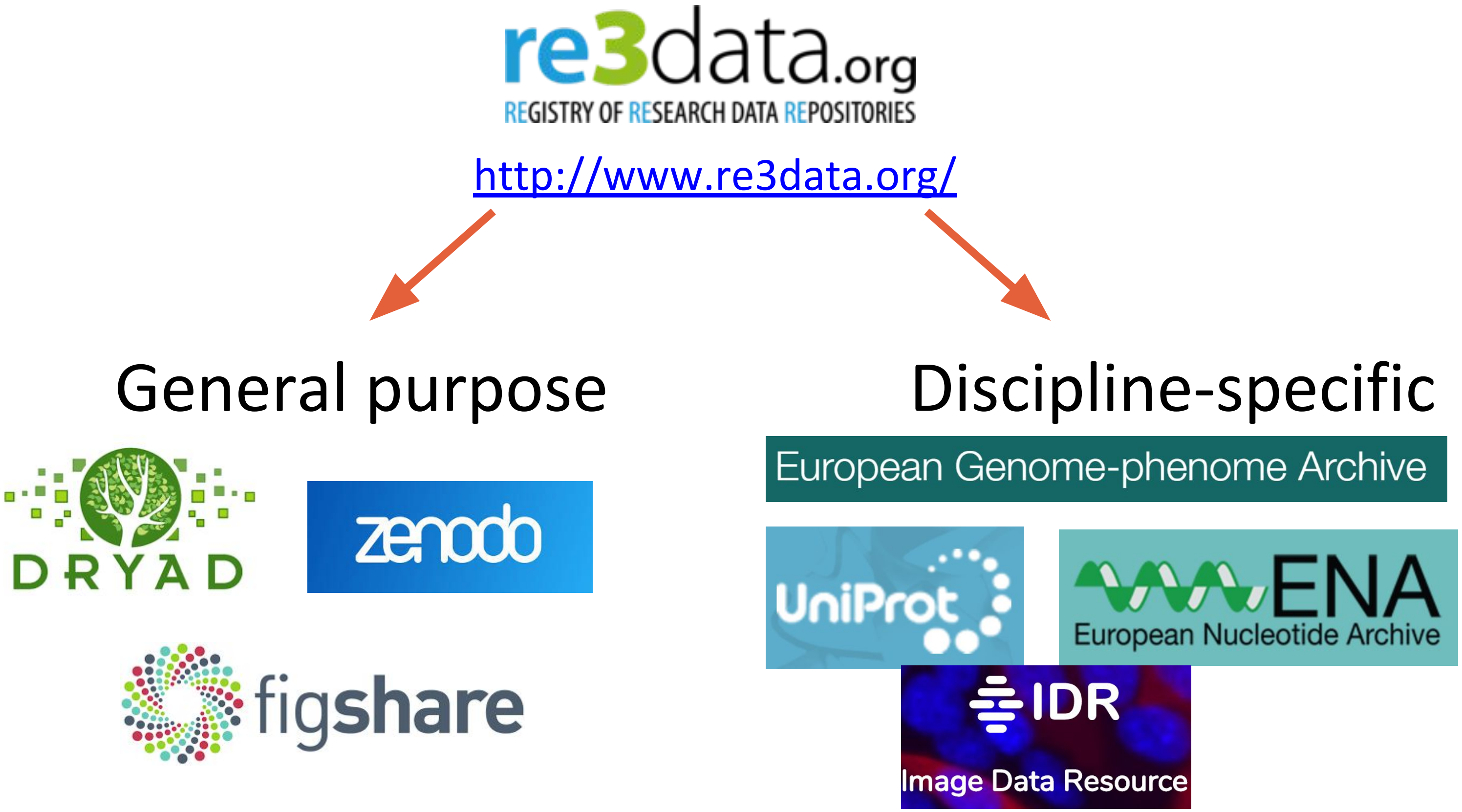
What’s the alternative to sharing ‘on request’?

There are different kinds of repositories:

- for datasets
- for protocols
- for software
- ...

What’s the alternative to sharing ‘on request’?

Repositories for datasets



re3data.org

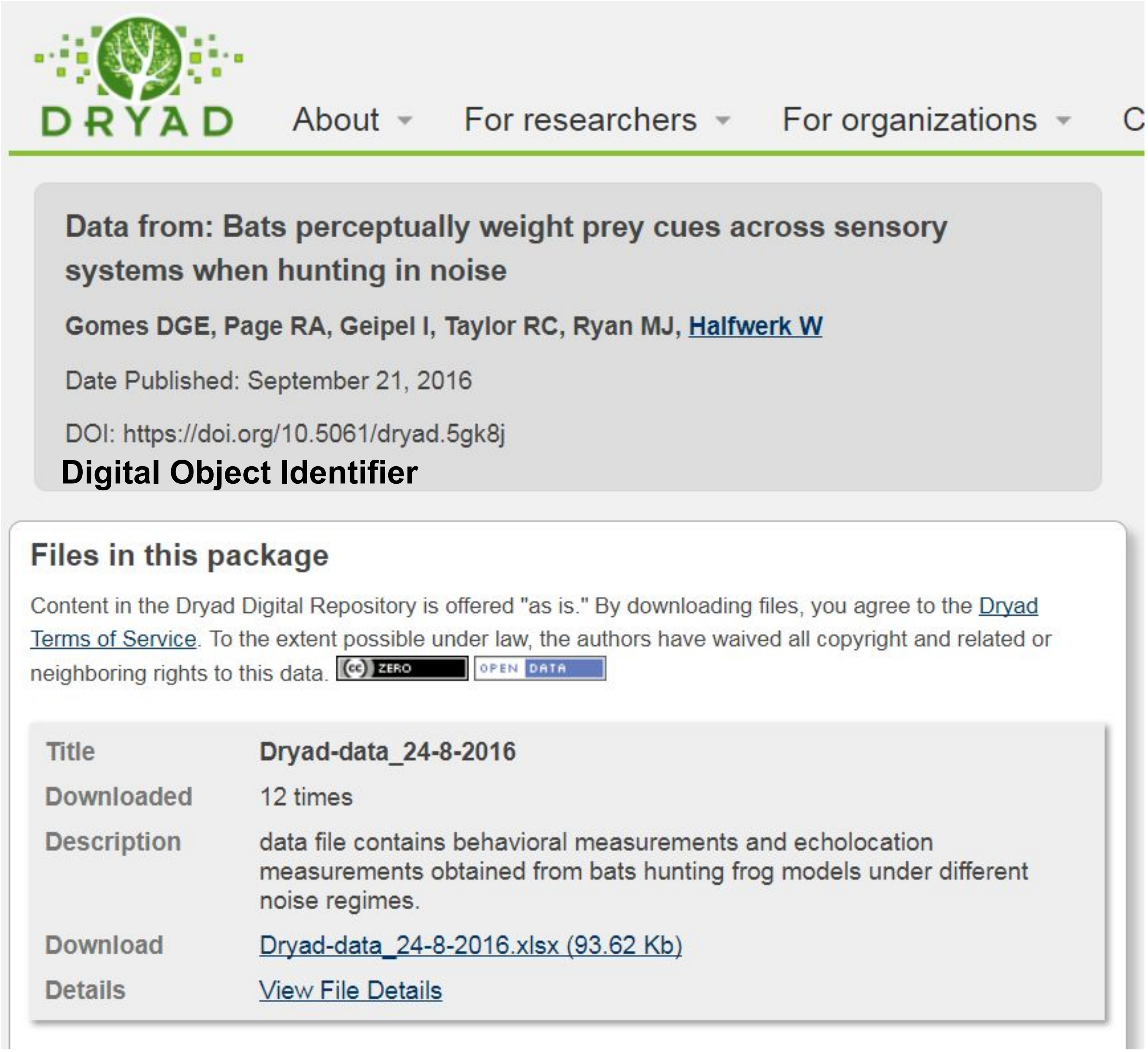
REGISTRY OF RESEARCH DATA REPOSITORIES

http://www.re3data.org/

General purpose

Discipline-specific

Benefits of sharing data via a repository



Data from: Bats perceptually weight prey cues across sensory systems when hunting in noise

Gomes DGE, Page RA, Geipel I, Taylor RC, Ryan MJ, Halfwerk W

Date Published: September 21, 2016

DOI: https://doi.org/10.5061/dryad.5gk8j

Digital Object Identifier

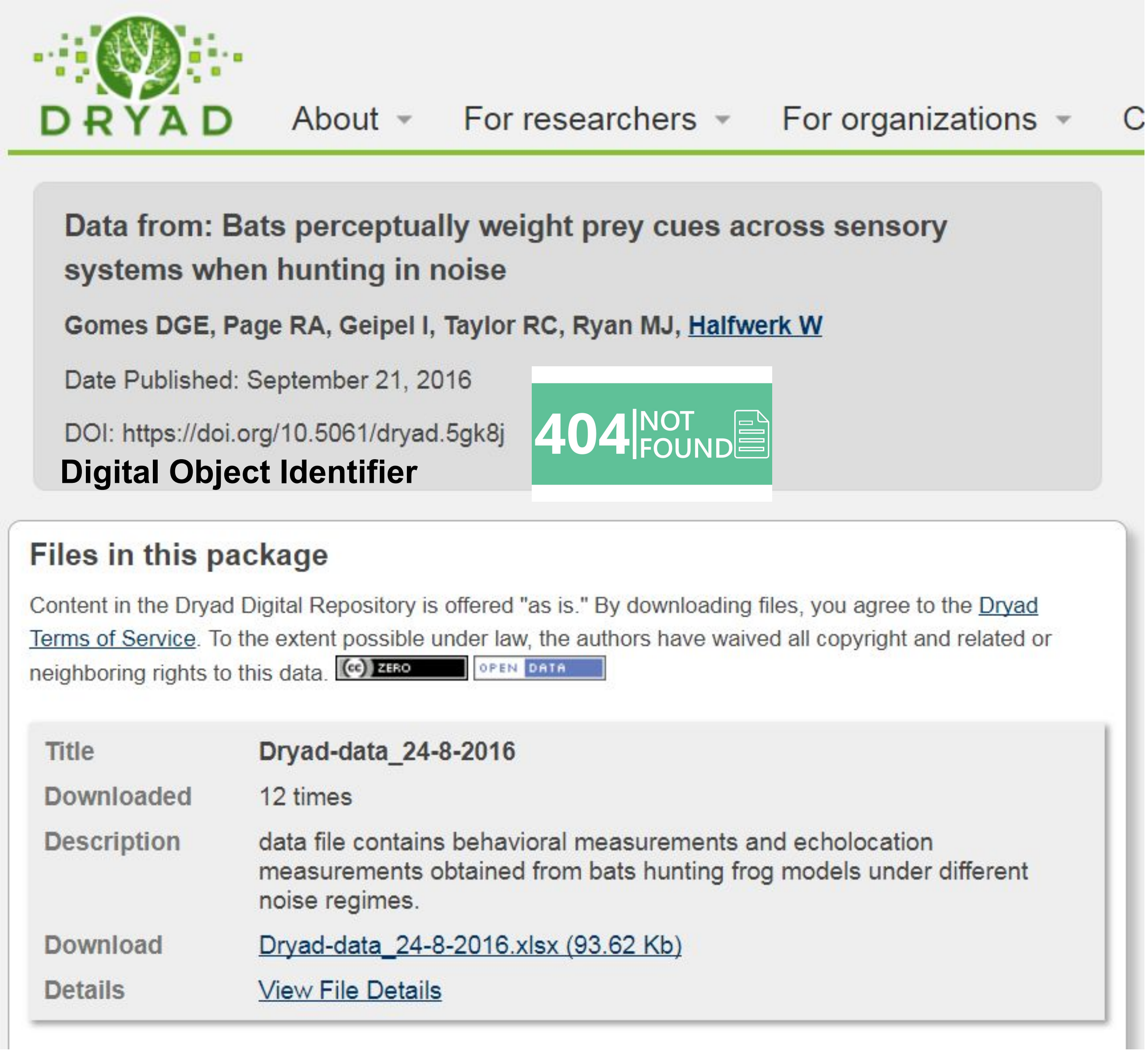
Files in this package

Content in the Dryad Digital Repository is offered "as is." By downloading files, you agree to the Dryad Terms of Service. To the extent possible under law, the authors have waived all copyright and related or neighboring rights to this data. CC ZERO OPEN DATA

Title	**Dryad-data_24-8-2016**
Downloaded	12 times
Description	data file contains behavioral measurements and echolocation measurements obtained from bats hunting frog models under different noise regimes.
Download	Dryad-data_24-8-2016.xlsx (93.62 Kb)
Details	View File Details

https://doi.org/10.5061/dryad.5gk8j

Benefits of sharing data via a repository



Data from: Bats perceptually weight prey cues across sensory systems when hunting in noise

Gomes DGE, Page RA, Geipel I, Taylor RC, Ryan MJ, Halfwerk W

Date Published: September 21, 2016

DOI: https://doi.org/10.5061/dryad.5gk8j

Digital Object Identifier

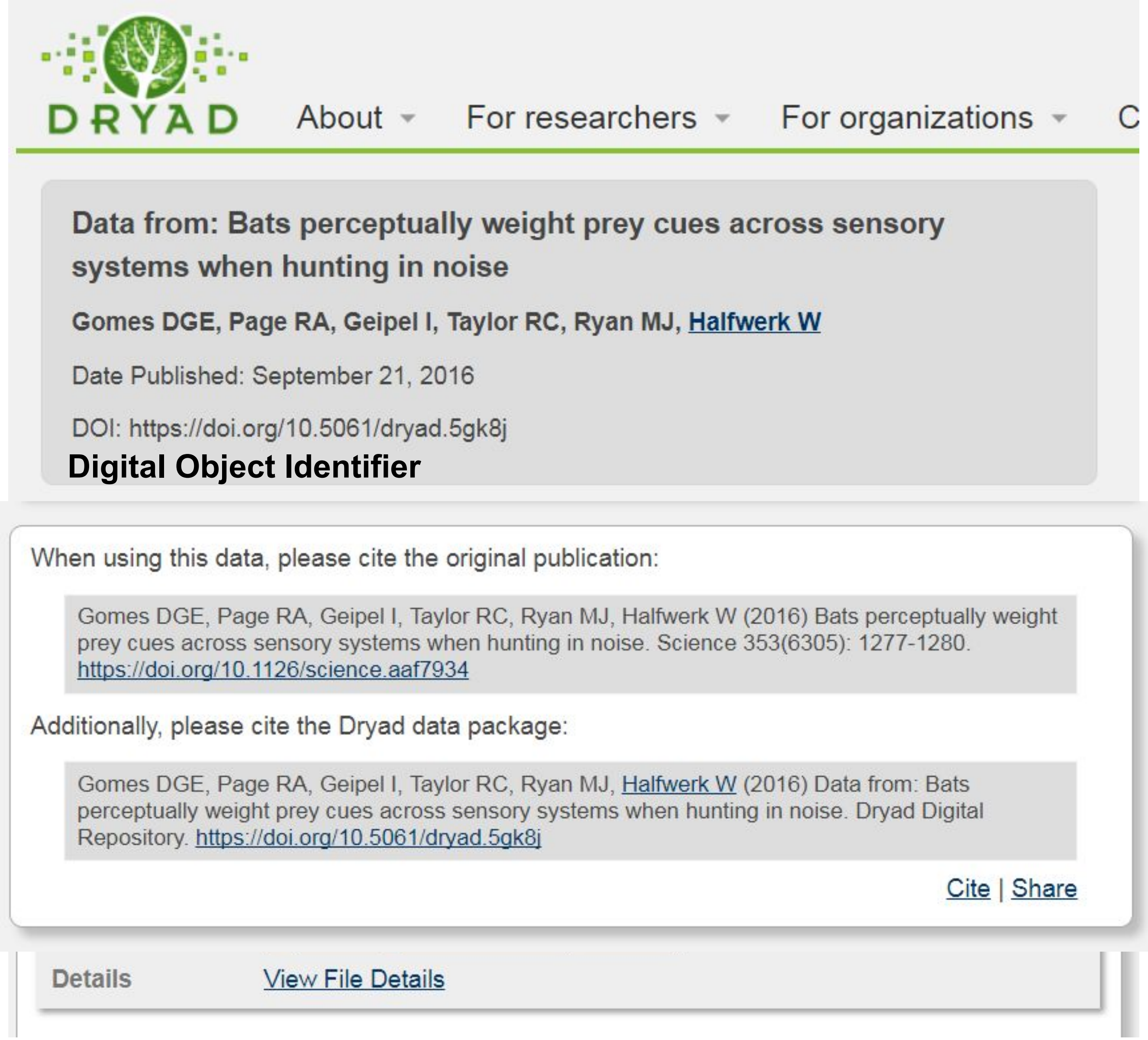
Files in this package

Content in the Dryad Digital Repository is offered "as is." By downloading files, you agree to the Dryad Terms of Service. To the extent possible under law, the authors have waived all copyright and related or neighboring rights to this data. CC ZERO OPEN DATA

Title	Dryad-data_24-8-2016
Downloaded	12 times
Description	data file contains behavioral measurements and echolocation measurements obtained from bats hunting frog models under different noise regimes.
Download	Dryad-data_24-8-2016.xlsx (93.62 Kb)
Details	View File Details

https://doi.org/10.5061/dryad.5gk8j

Benefits of sharing data via a repository



Data from: Bats perceptually weight prey cues across sensory systems when hunting in noise

Gomes DGE, Page RA, Geipel I, Taylor RC, Ryan MJ, Halfwerk W

Date Published: September 21, 2016

DOI: https://doi.org/10.5061/dryad.5gk8j

Digital Object Identifier

When using this data, please cite the original publication:

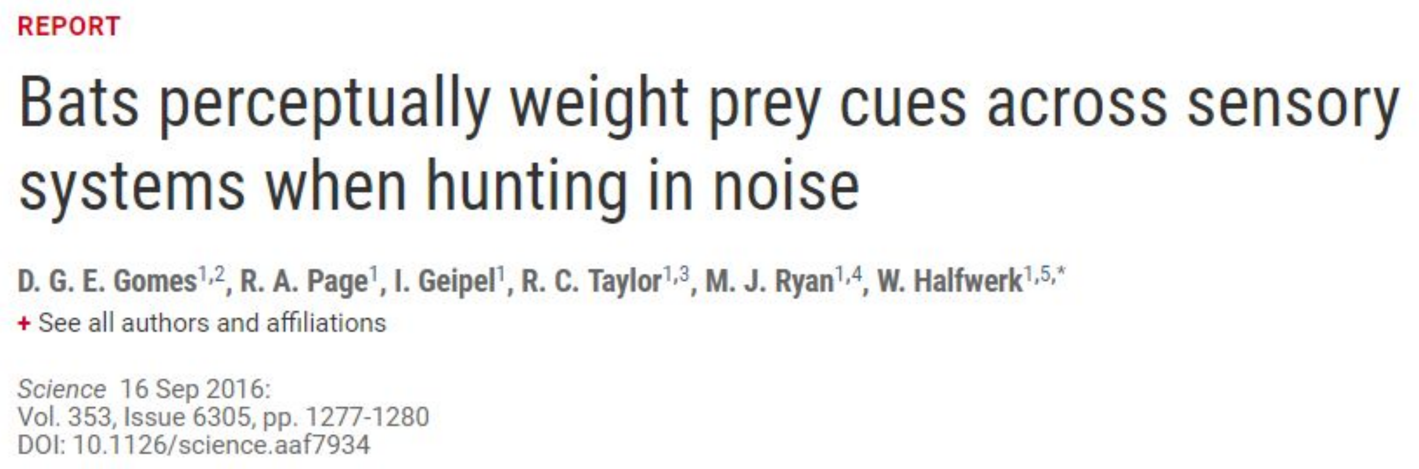
> Gomes DGE, Page RA, Geipel I, Taylor RC, Ryan MJ, Halfwerk W (2016) Bats perceptually weight prey cues across sensory systems when hunting in noise. Science 353(6305): 1277-1280. https://doi.org/10.1126/science.aaf7934

Additionally, please cite the Dryad data package:

> Gomes DGE, Page RA, Geipel I, Taylor RC, Ryan MJ, Halfwerk W (2016) Data from: Bats perceptually weight prey cues across sensory systems when hunting in noise. Dryad Digital Repository. https://doi.org/10.5061/dryad.5gk8j

https://doi.org/10.5061/dryad.5gk8j

Benefits of sharing data via a repository



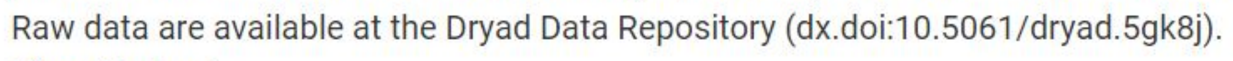
REPORT

Bats perceptually weight prey cues across sensory systems when hunting in noise

D. G. E. Gomes[1,2], R. A. Page[1], I. Geipel[1], R. C. Taylor[1,3], M. J. Ryan[1,4], W. Halfwerk[1,5,*]

+ See all authors and affiliations

Science 16 Sep 2016:
Vol. 353, Issue 6305, pp. 1277-1280
DOI: 10.1126/science.aaf7934



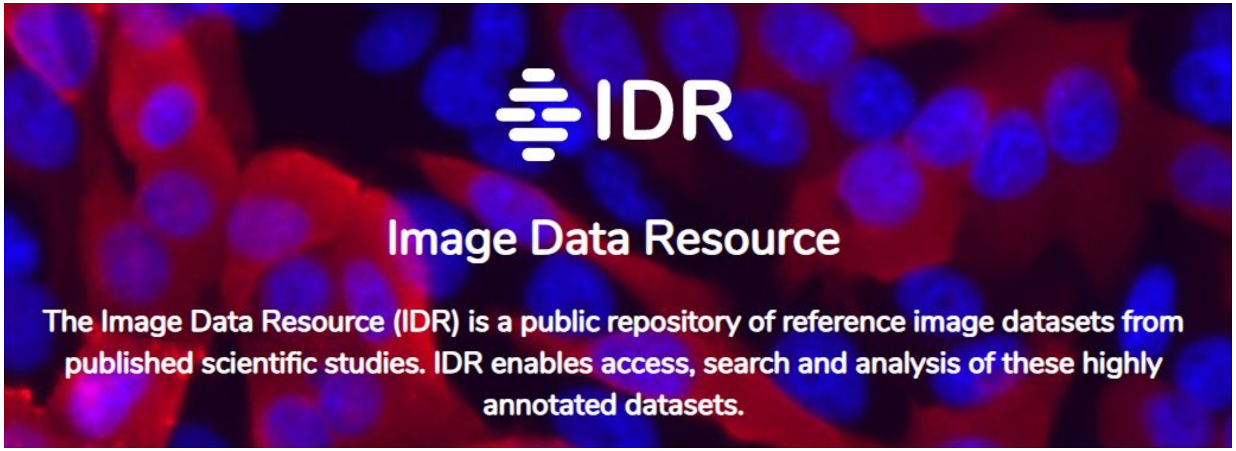
Raw data are available at the Dryad Data Repository (dx.doi:10.5061/dryad.5gk8j).

No emails with requests for data anymore
Citations from not only the papers but also the datasets
Increased visibility and impact

http://science.sciencemag.org/content/353/6305/1277.full

What’s the alternative to sharing ‘on request’?

Repositories for images

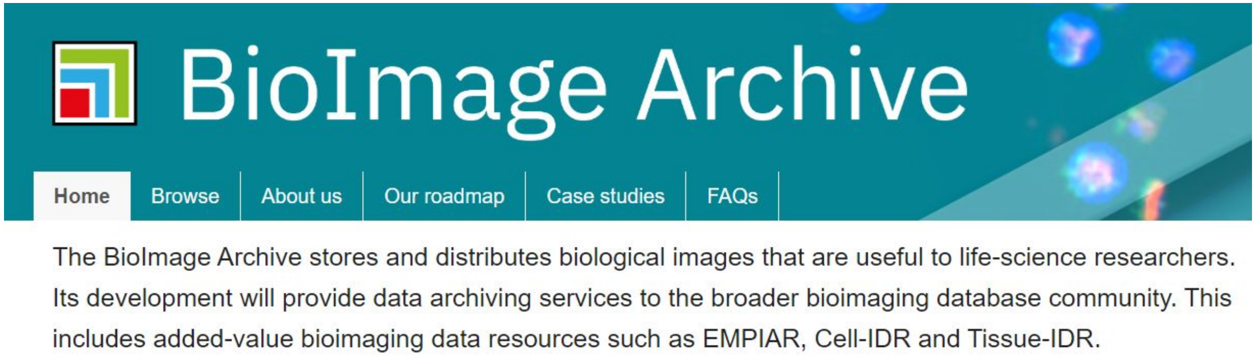


IDR

Image Data Resource

The Image Data Resource (IDR) is a public repository of reference image datasets from published scientific studies. IDR enables access, search and analysis of these highly annotated datasets.

https://idr.openmicroscopy.org/about/about.html

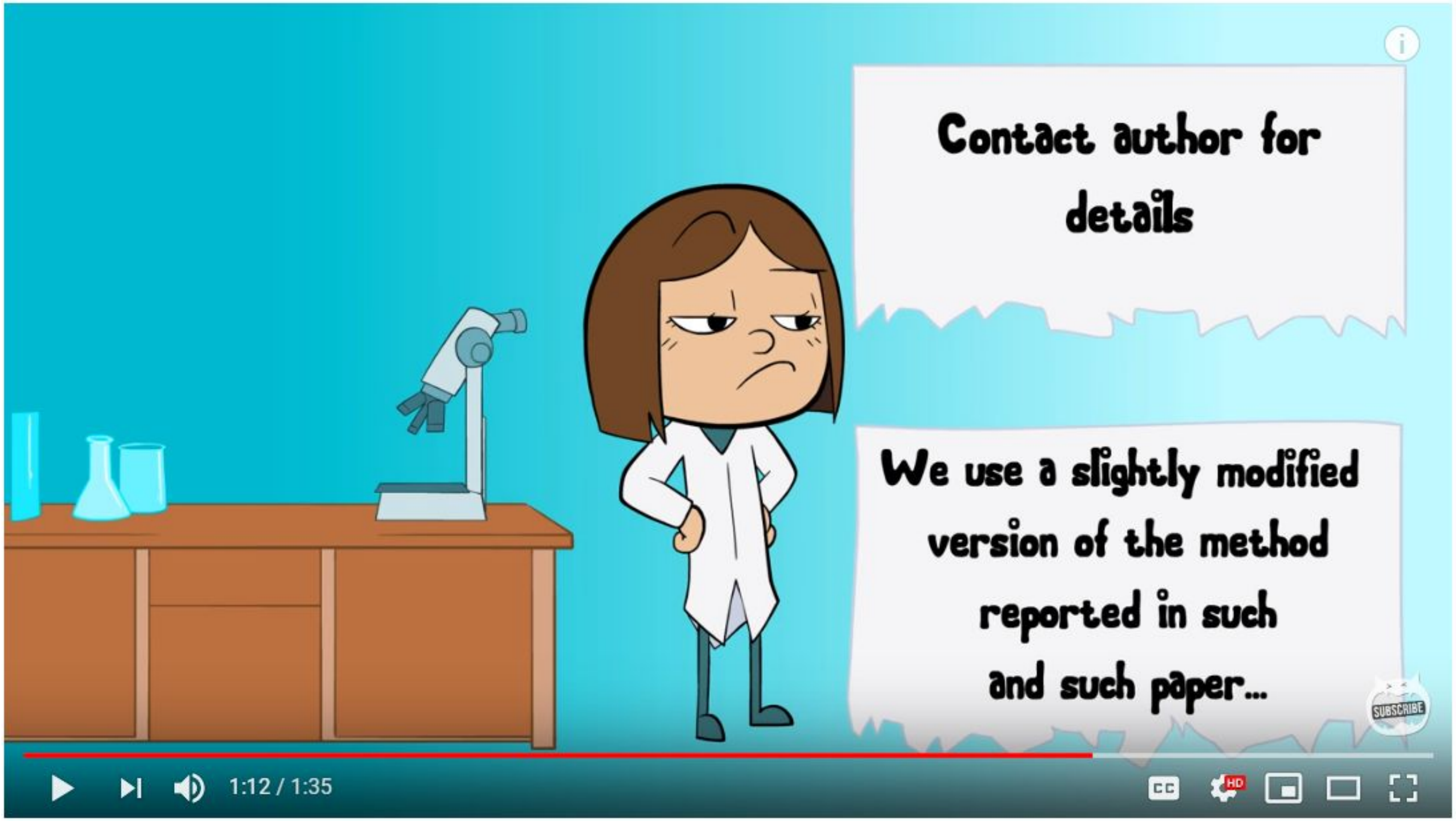


BioImage Archive

The BioImage Archive stores and distributes biological images that are useful to life-science researchers. Its development will provide data archiving services to the broader bioimaging database community. This includes added-value bioimaging data resources such as EMPIAR, Cell-IDR and Tissue-IDR.

https://www.ebi.ac.uk/bioimage-archive/

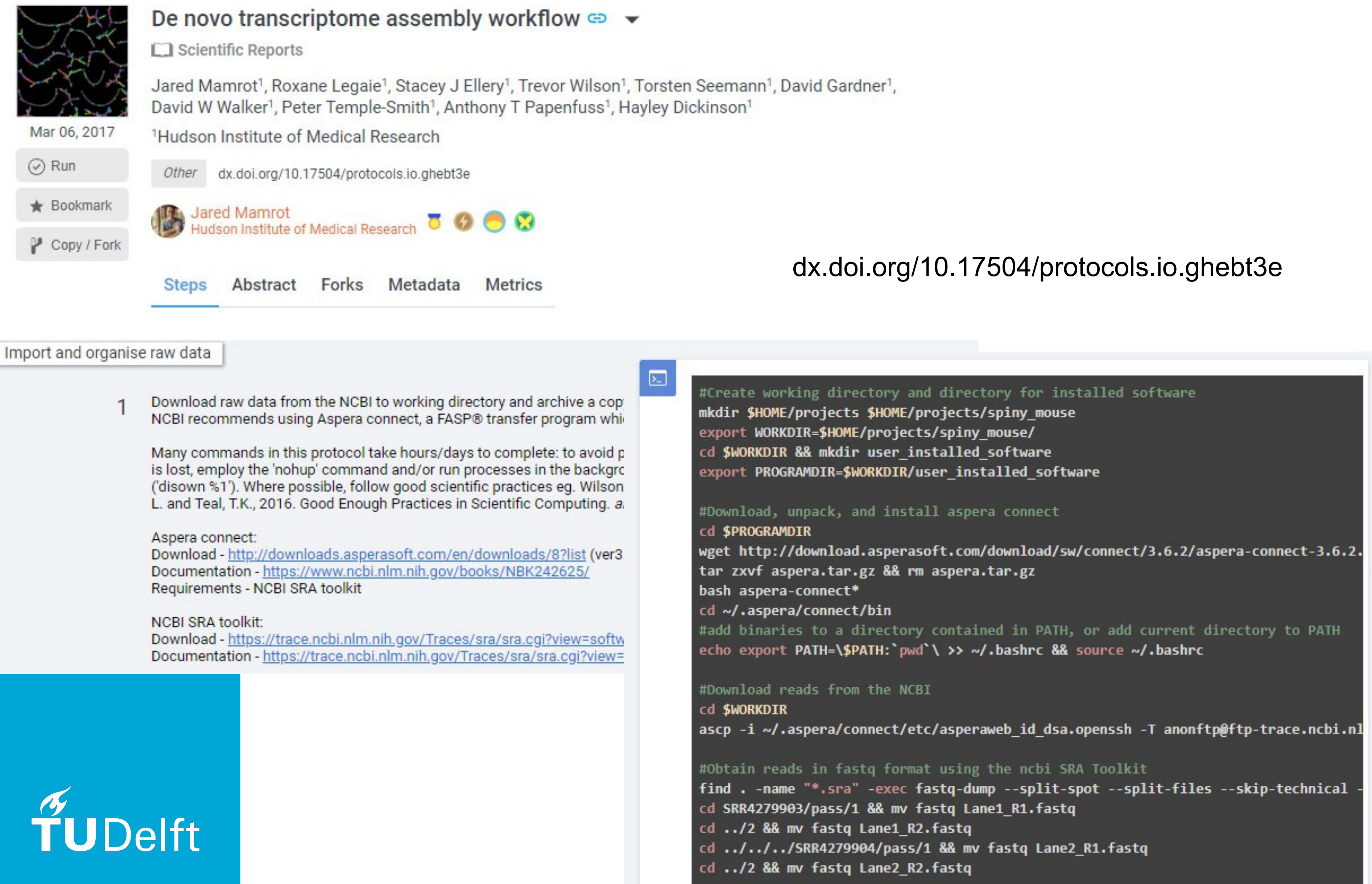
Repositories for protocols


Contact author for details
We use a slightly modified version of the method reported in such and such paper...
SUBSCRIBE
1:12 / 1:35
CC
HD

Protocols.io - Share science protocol knowledge

https://www.youtube.com/watch?v=84B8P6BAOgM
https://www.protocols.io/

What’s the alternative to sharing ‘on request’?



De novo transcriptome assembly workflow

Scientific Reports

Jared Mamrot[1], Roxane Legaie[1], Stacey J Ellery[1], Trevor Wilson[1], Torsten Seemann[1], David Gardner[1], David W Walker[1], Peter Temple-Smith[1], Anthony T Papenfuss[1], Hayley Dickinson[1]

[1]Hudson Institute of Medical Research

Mar 06, 2017

Other dx.doi.org/10.17504/protocols.io.ghebt3e

Jared Mamrot
Hudson Institute of Medical Research

dx.doi.org/10.17504/protocols.io.ghebt3e

Import and organise raw data

1 Download raw data from the NCBI to working directory and archive a cop
NCBI recommends using Aspera connect, a FASP® transfer program whi

Many commands in this protocol take hours/days to complete: to avoid p
is lost, employ the 'nohup' command and/or run processes in the backgrc
('disown %1'). Where possible, follow good scientific practices eg. Wilson
L. and Teal, T.K., 2016. Good Enough Practices in Scientific Computing. a

Aspera connect:
Download - http://downloads.asperasoft.com/en/downloads/8?list (ver3
Documentation - https://www.ncbi.nlm.nih.gov/books/NBK242625/
Requirements - NCBI SRA toolkit

NCBI SRA toolkit:
Download - https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=softw
Documentation - https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=

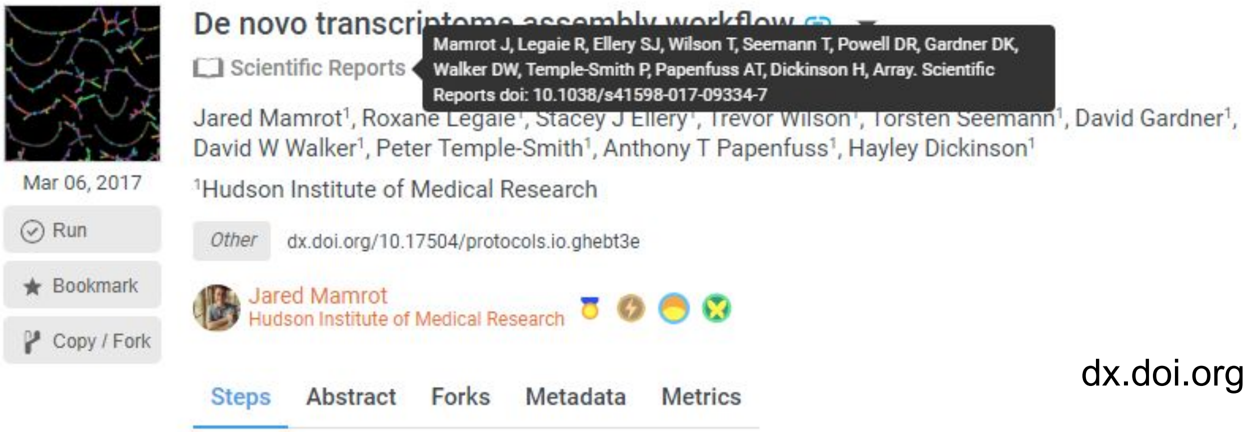
```
#Create working directory and directory for installed software
mkdir $HOME/projects $HOME/projects/spiny_mouse
export WORKDIR=$HOME/projects/spiny_mouse/
cd $WORKDIR && mkdir user_installed_software
export PROGRAMDIR=$WORKDIR/user_installed_software

#Download, unpack, and install aspera connect
cd $PROGRAMDIR
wget http://download.asperasoft.com/download/sw/connect/3.6.2/aspera-connect-3.6.2.
tar zxvf aspera.tar.gz && rm aspera.tar.gz
bash aspera-connect*
cd ~/.aspera/connect/bin
#add binaries to a directory contained in PATH, or add current directory to PATH
echo export PATH=\$PATH:`pwd`\ >> ~/.bashrc && source ~/.bashrc

#Download reads from the NCBI
cd $WORKDIR
ascp -i ~/.aspera/connect/etc/asperaweb_id_dsa.openssh -T anonftp@ftp-trace.ncbi.nl

#Obtain reads in fastq format using the ncbi SRA Toolkit
find . -name "*.sra" -exec fastq-dump --split-spot --split-files --skip-technical -
cd SRR4279903/pass/1 && mv fastq Lane1_R1.fastq
cd ../2 && mv fastq Lane1_R2.fastq
cd ../../../SRR4279904/pass/1 && mv fastq Lane2_R1.fastq
cd ../2 && mv fastq Lane2_R2.fastq
```

What’s the alternative to sharing ‘on request’?



De novo transcriptome assembly workflow

Scientific Reports

Mamrot J, Legaie R, Ellery SJ, Wilson T, Seemann T, Powell DR, Gardner DK, Walker DW, Temple-Smith P, Papenfuss AT, Dickinson H, Array. Scientific Reports doi: 10.1038/s41598-017-09334-7

Jared Mamrot[1], Roxane Legaie[1], Stacey J Ellery[1], Trevor Wilson[1], Torsten Seemann[1], David Gardner[1], David W Walker[1], Peter Temple-Smith[1], Anthony T Papenfuss[1], Hayley Dickinson[1]

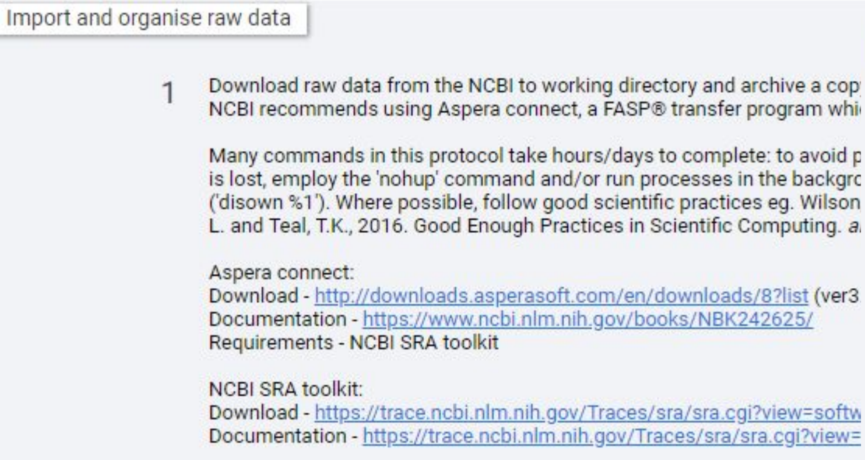
[1]Hudson Institute of Medical Research

Mar 06, 2017

Other dx.doi.org/10.17504/protocols.io.ghebt3e

Jared Mamrot
Hudson Institute of Medical Research

dx.doi.org/10.17504/protocols.io.ghebt3e



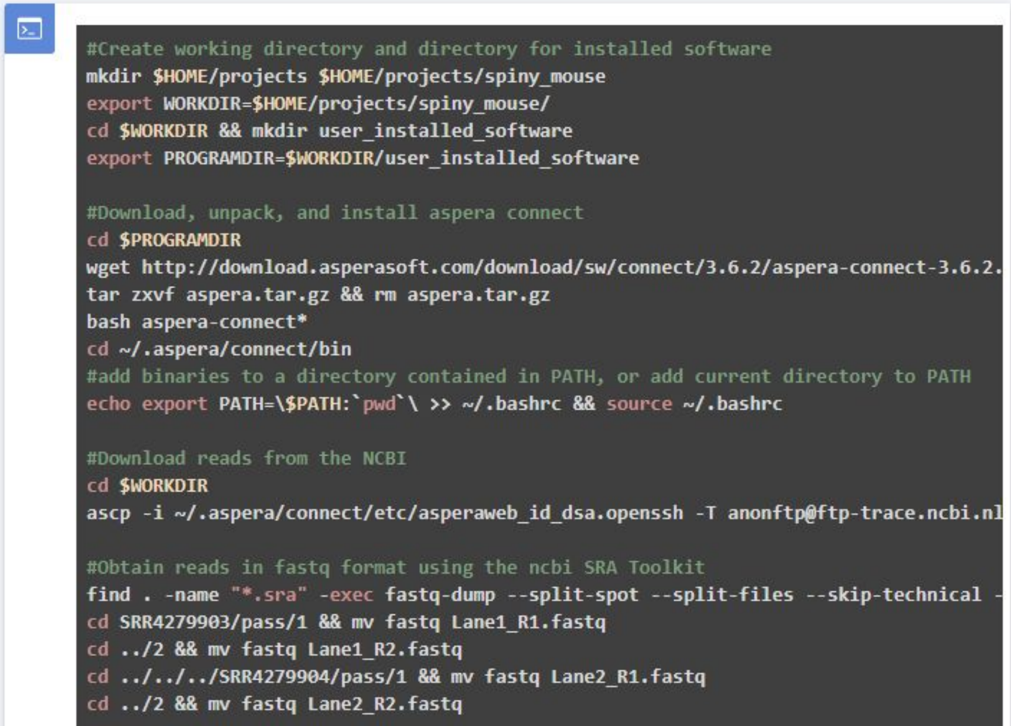
Import and organise raw data

1 Download raw data from the NCBI to working directory and archive a cop
NCBI recommends using Aspera connect, a FASP® transfer program whi

Many commands in this protocol take hours/days to complete: to avoid p
is lost, employ the 'nohup' command and/or run processes in the backgrc
('disown %1'). Where possible, follow good scientific practices eg. Wilson
L. and Teal, T.K., 2016. Good Enough Practices in Scientific Computing. a

Aspera connect:
Download - http://downloads.asperasoft.com/en/downloads/8?list (ver3
Documentation - https://www.ncbi.nlm.nih.gov/books/NBK242625/
Requirements - NCBI SRA toolkit

NCBI SRA toolkit:
Download - https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=softw
Documentation - https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=



```
#Create working directory and directory for installed software
mkdir $HOME/projects $HOME/projects/spiny_mouse
export WORKDIR=$HOME/projects/spiny_mouse/
cd $WORKDIR && mkdir user_installed_software
export PROGRAMDIR=$WORKDIR/user_installed_software

#Download, unpack, and install aspera connect
cd $PROGRAMDIR
wget http://download.asperasoft.com/download/sw/connect/3.6.2/aspera-connect-3.6.2.
tar zxvf aspera.tar.gz && rm aspera.tar.gz
bash aspera-connect*
cd ~/.aspera/connect/bin
#add binaries to a directory contained in PATH, or add current directory to PATH
echo export PATH=\$PATH:`pwd`\ >> ~/.bashrc && source ~/.bashrc

#Download reads from the NCBI
cd $WORKDIR
ascp -i ~/.aspera/connect/etc/asperaweb_id_dsa.openssh -T anonftp@ftp-trace.ncbi.nl

#Obtain reads in fastq format using the ncbi SRA Toolkit
find . -name "*.sra" -exec fastq-dump --split-spot --split-files --skip-technical -
cd SRR4279903/pass/1 && mv fastq Lane1_R1.fastq
cd ../2 && mv fastq Lane2_R2.fastq
cd ../../../SRR4279904/pass/1 && mv fastq Lane2_R1.fastq
cd ../2 && mv fastq Lane2_R2.fastq
```

What’s the alternative to sharing ‘on request’?

GitHub

+

zenodo

Repositories for software

Module 5, Task 2: How to make your code citable using GitHub and Zenodo

277 views

https://www.youtube.com/watch?v=pjsbBQYOOaE&t=1s

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

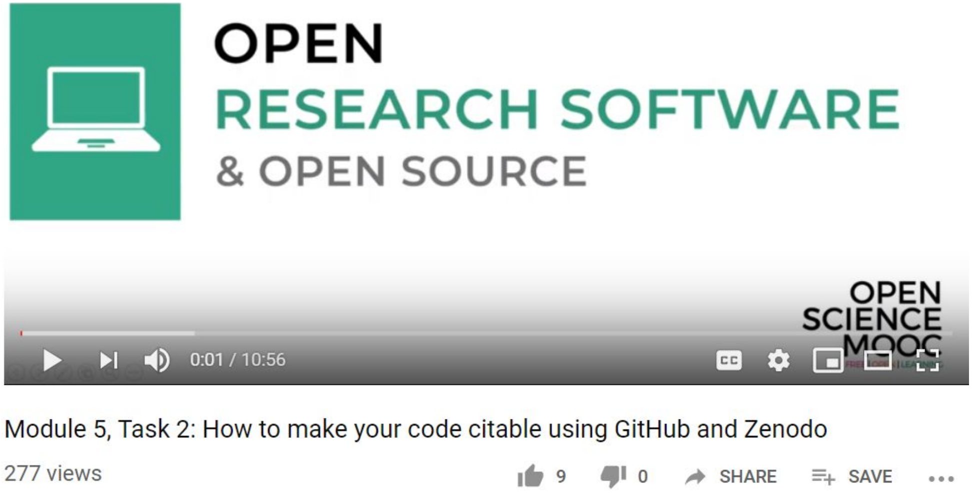
GitHub

+

zenodo

What's the alternative to sharing 'on request'?

Repositories for software

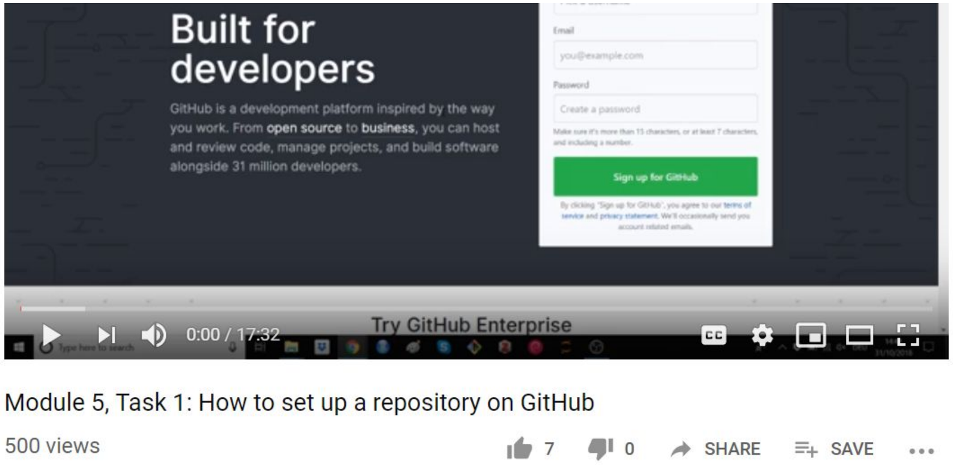


Module 5, Task 2: How to make your code citable using GitHub and Zenodo

277 views

https://www.youtube.com/watch?v=pjsbBQYOOaE&t=1s

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


Built for developers

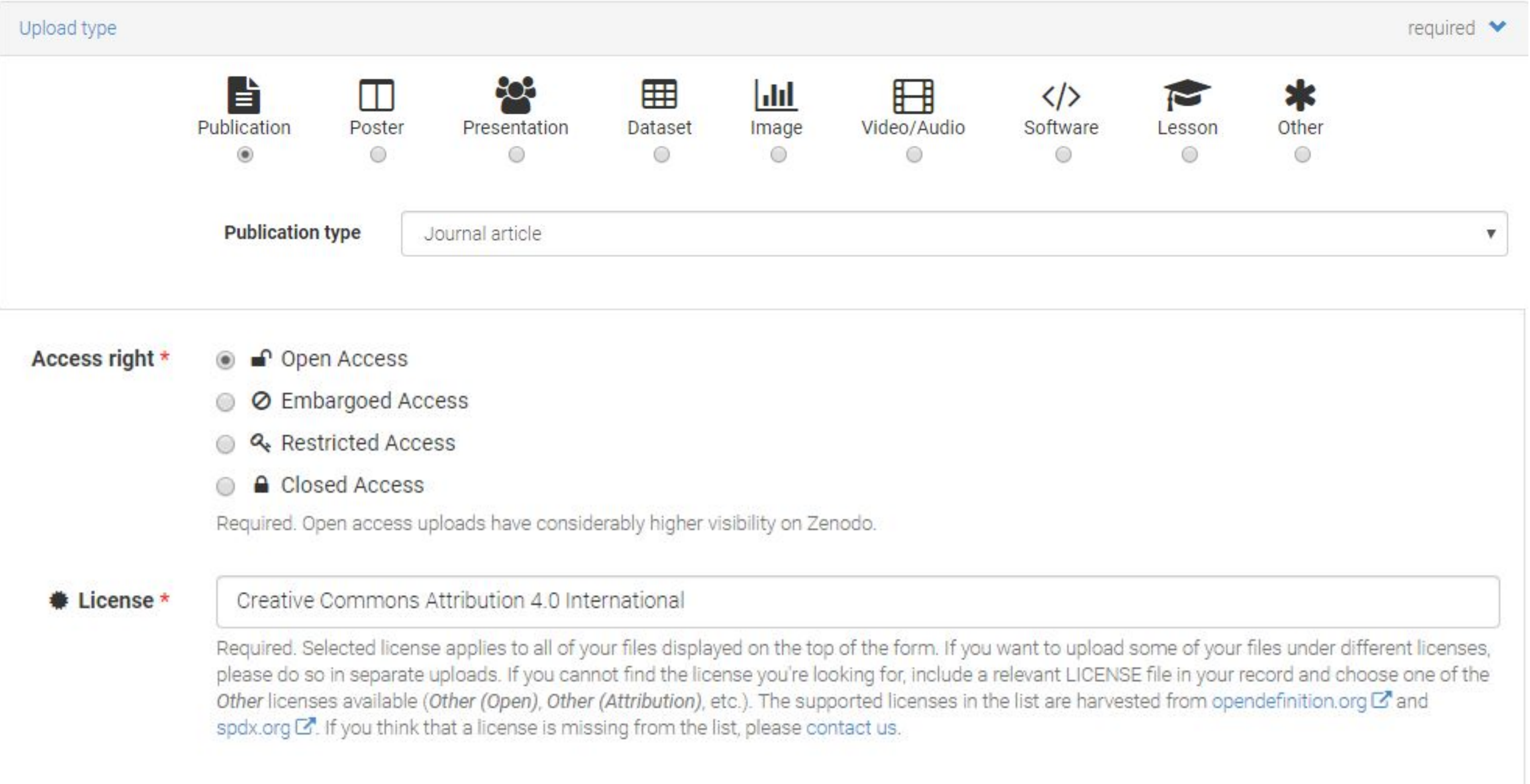
GitHub is a development platform inspired by the way you work. From **open source** to **business**, you can host and review code, manage projects, and build software alongside 31 million developers.

Module 5, Task 1: How to set up a repository on GitHub

500 views

https://www.youtube.com/watch?v=AnftV9HBPSc

TU Delft



Upload type required

● Publication ○ Poster ○ Presentation ○ Dataset ○ Image ○ Video/Audio ○ Software ○ Lesson ○ Other

Publication type Journal article

Access right *

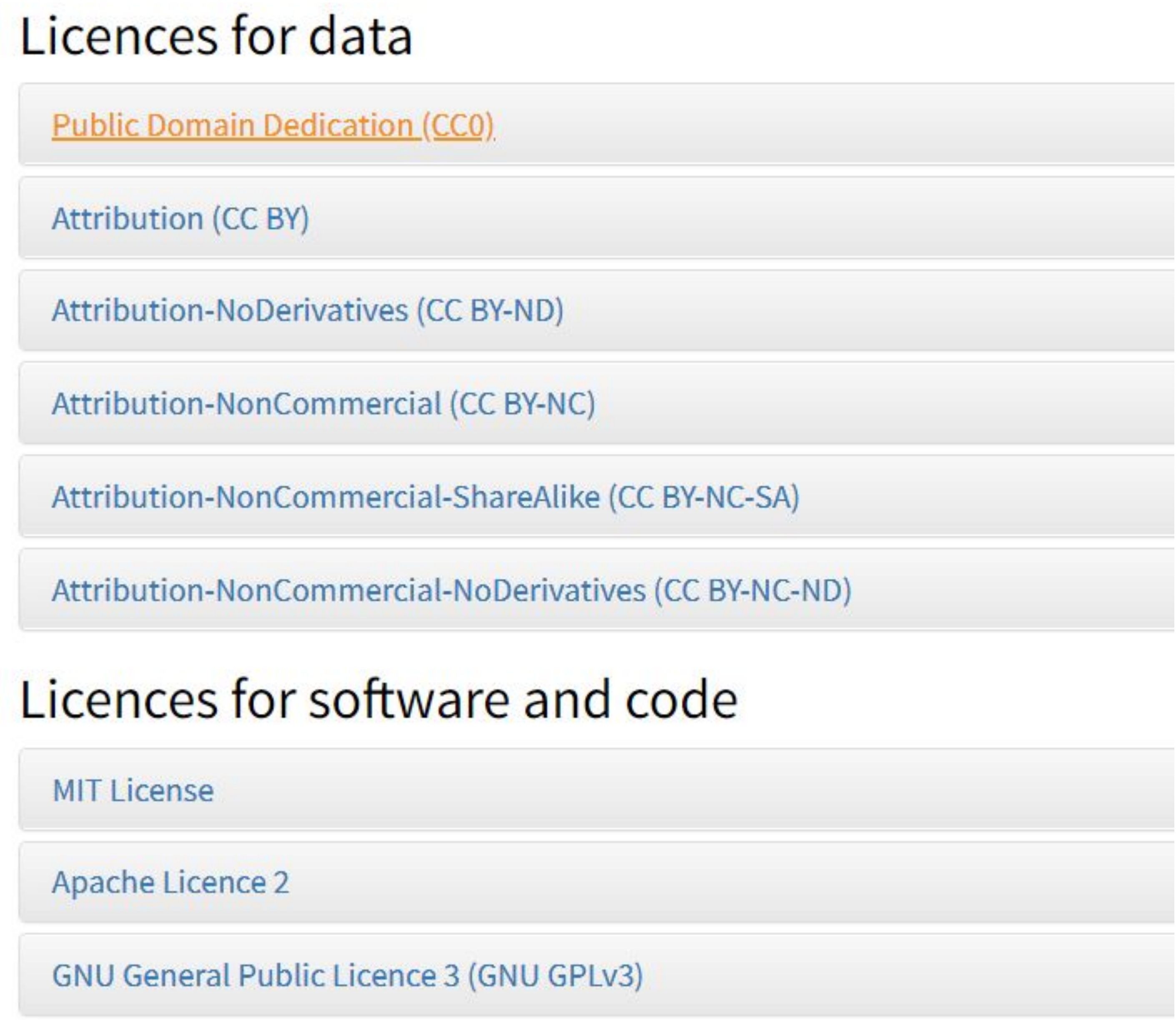
● Open Access
○ Embargoed Access
○ Restricted Access
○ Closed Access

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

License * Creative Commons Attribution 4.0 International

Required. Selected license applies to all of your files displayed on the top of the form. If you want to upload some of your files under different licenses, please do so in separate uploads. If you cannot find the license you're looking for, include a relevant LICENSE file in your record and choose one of the *Other* licenses available (*Other (Open)*, *Other (Attribution)*, etc.). The supported licenses in the list are harvested from opendefinition.org and spdx.org. If you think that a license is missing from the list, please contact us.

https://zenodo.org/



Licences for data

- Public Domain Dedication (CC0)
- Attribution (CC BY)
- Attribution-NoDerivatives (CC BY-ND)
- Attribution-NonCommercial (CC BY-NC)
- Attribution-NonCommercial-ShareAlike (CC BY-NC-SA)
- Attribution-NonCommercial-NoDerivatives (CC BY-NC-ND)

Licences for software and code

- MIT License
- Apache Licence 2
- GNU General Public Licence 3 (GNU GPLv3)

https://researchdata.4tu.nl/en/use-4turesearchdata/archive-research-data/upload-your-data-in-our-data-archive/licencing/

Funders' requirements

FAIR Data Principles

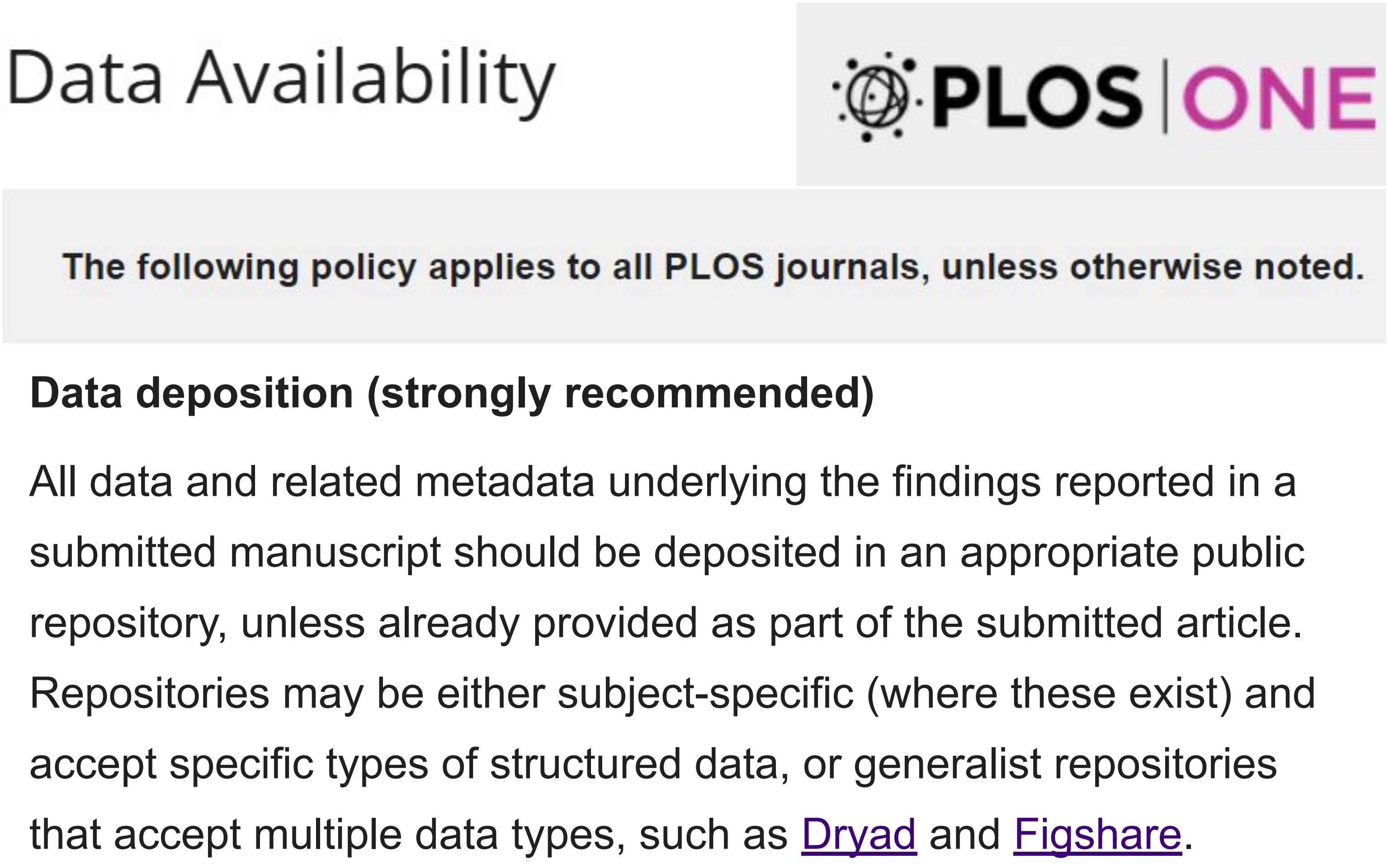
FAIR
DATA!
Accessible
Findable
Interoperable
Re-usable

http://ec.europa.eu/research/press/2016/pdf/opendata-infographic_072016.pdf#view=fit&pagemode=none

- Requirement of increasing number of funders

reach more people, have greater impact
avoid duplication of efforts
preserve data for future researchers

Publishers' requirements



Data Availability

PLOS | ONE

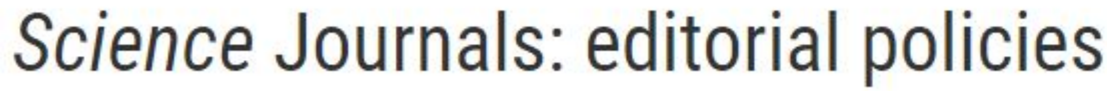
The following policy applies to all PLOS journals, unless otherwise noted.

Data deposition (strongly recommended)

All data and related metadata underlying the findings reported in a submitted manuscript should be deposited in an appropriate public repository, unless already provided as part of the submitted article. Repositories may be either subject-specific (where these exist) and accept specific types of structured data, or generalist repositories that accept multiple data types, such as Dryad and Figshare.

https://journals.plos.org/plosone/s/data-availability

Publishers' requirements



Science Journals: editorial policies

Science

Data Deposition

The *Science* Journals support the efforts of databases that aggregate published data for the use of the scientific community. Therefore, before publication, large data sets (including microarray data, protein or DNA sequences, atomic coordinates or electron microscopy maps for molecular and macromolecular structures, and climate data) must be deposited in an approved database and an accession number or a specific access address must be included in the published paper.

http://www.sciencemag.org/authors/science-journals-editorial-policies

Data Documentation

Human readable

- Readme files with info about:
  - Methods used for data collection and analysis
  - Data-specific information (parameters, variables, column headings, symbols used, etc.)

https://cornell.app.box.com/v/ReadmeTemplate

TUDelft

Data Documentation

Human readable

- Readme files with info about:
  - Methods used for data collection and analysis
  - Data-specific information (parameters, variables, column headings, symbols used, etc.)

Machine readable

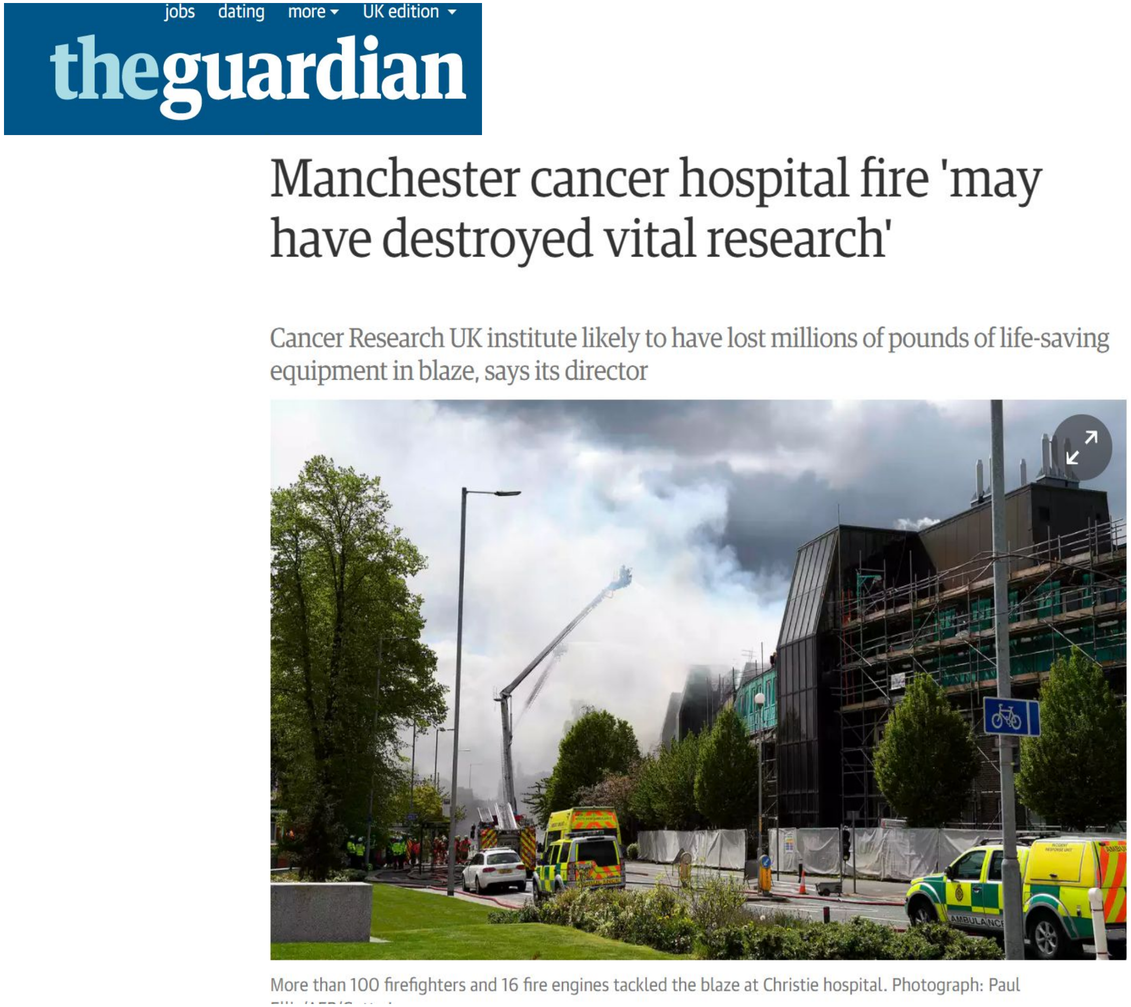
- Metadata with defined fields:
  - Title, date, creator(s), keywords..
  - Disciplinary standards if possible

https://cornell.app.box.com/v/ReadmeTemplate
https://fairsharing.org/standards/
https://rd-alliance.github.io/metadata-directory/

Day-to-day management of data

- Secure Data Storage and Backup
- Data organization
  - File & folder organisation structure
  - File naming
  - Version control
  - Experimental notes

Data loss? It actually happens



Manchester cancer hospital fire 'may have destroyed vital research'

Cancer Research UK institute likely to have lost millions of pounds of life-saving equipment in blaze, says its director

More than 100 firefighters and 16 fire engines tackled the blaze at Christie hospital. Photograph: Paul

https://www.theguardian.com/uk-news/2017/apr/28/manchester-christie-cancer-hospital-fire-research-equipment-destroyed?CMP=Share_iOSApp_Other

To avoid data loss:

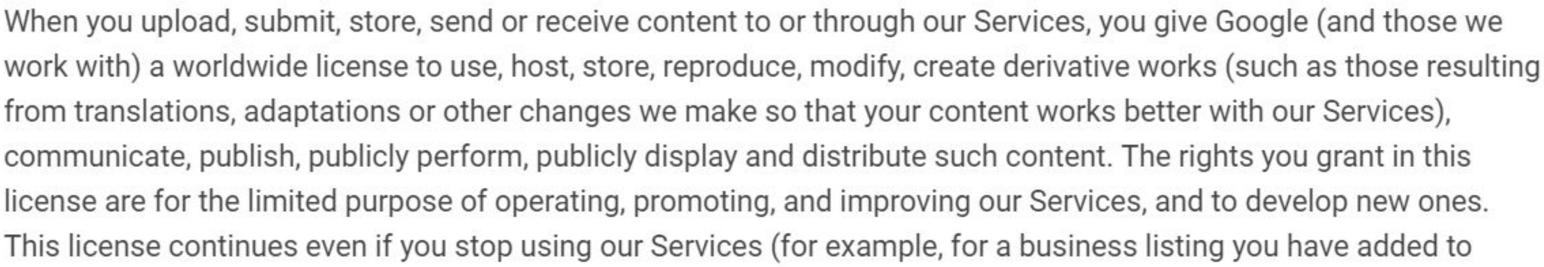
- Backup your data **regularly and preferably automatically**
- Create, at a **minimum, 2 copies of your data**
- Store data at **multiple trusted locations**
- Use **reliable backup solutions**

- Avoid data storage on hard disks, USB's, and personal computers without backup

https://www.dataone.org/best-practices/backup-your-data

Always read the small print…

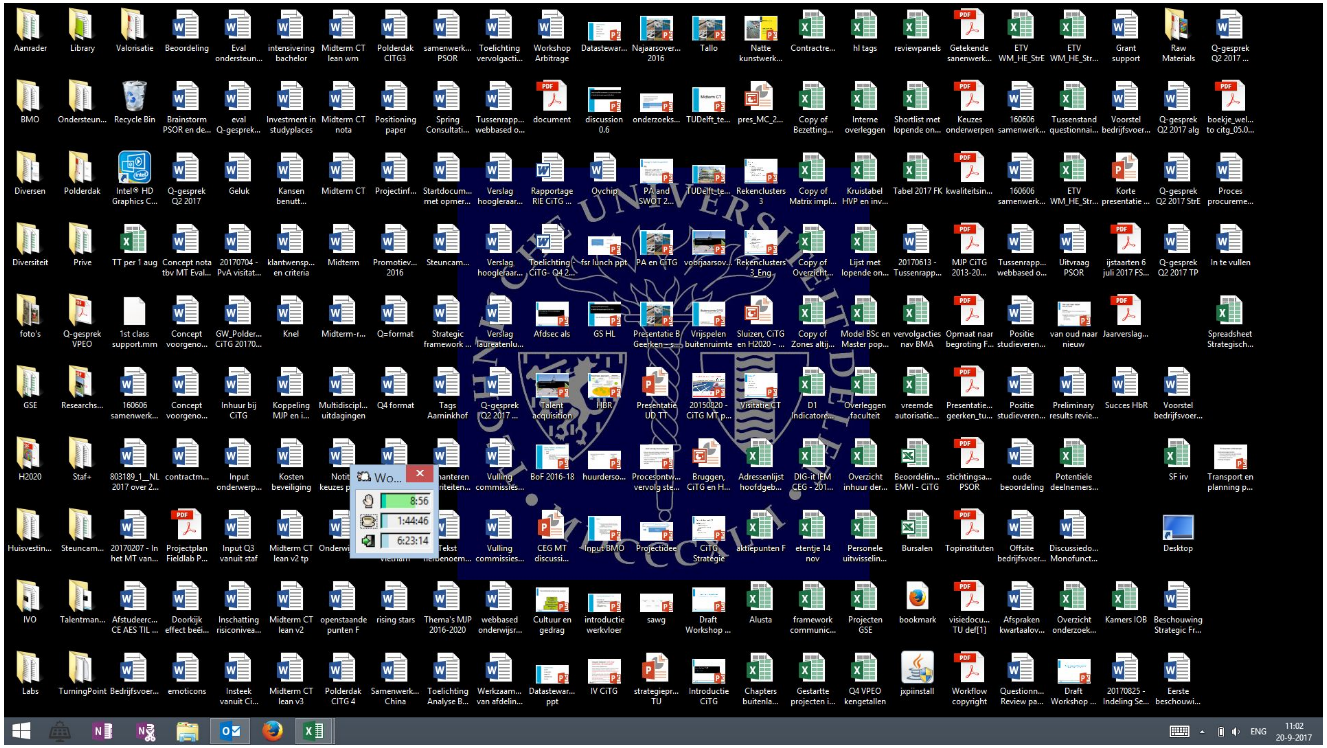
Google services Terms of Use:



When you upload, submit, store, send or receive content to or through our Services, you give Google (and those we work with) a worldwide license to use, host, store, reproduce, modify, create derivative works (such as those resulting from translations, adaptations or other changes we make so that your content works better with our Services), communicate, publish, publicly perform, publicly display and distribute such content. The rights you grant in this license are for the limited purpose of operating, promoting, and improving our Services, and to develop new ones. This license continues even if you stop using our Services (for example, for a business listing you have added to

https://www.google.com/intl/en/policies/terms/

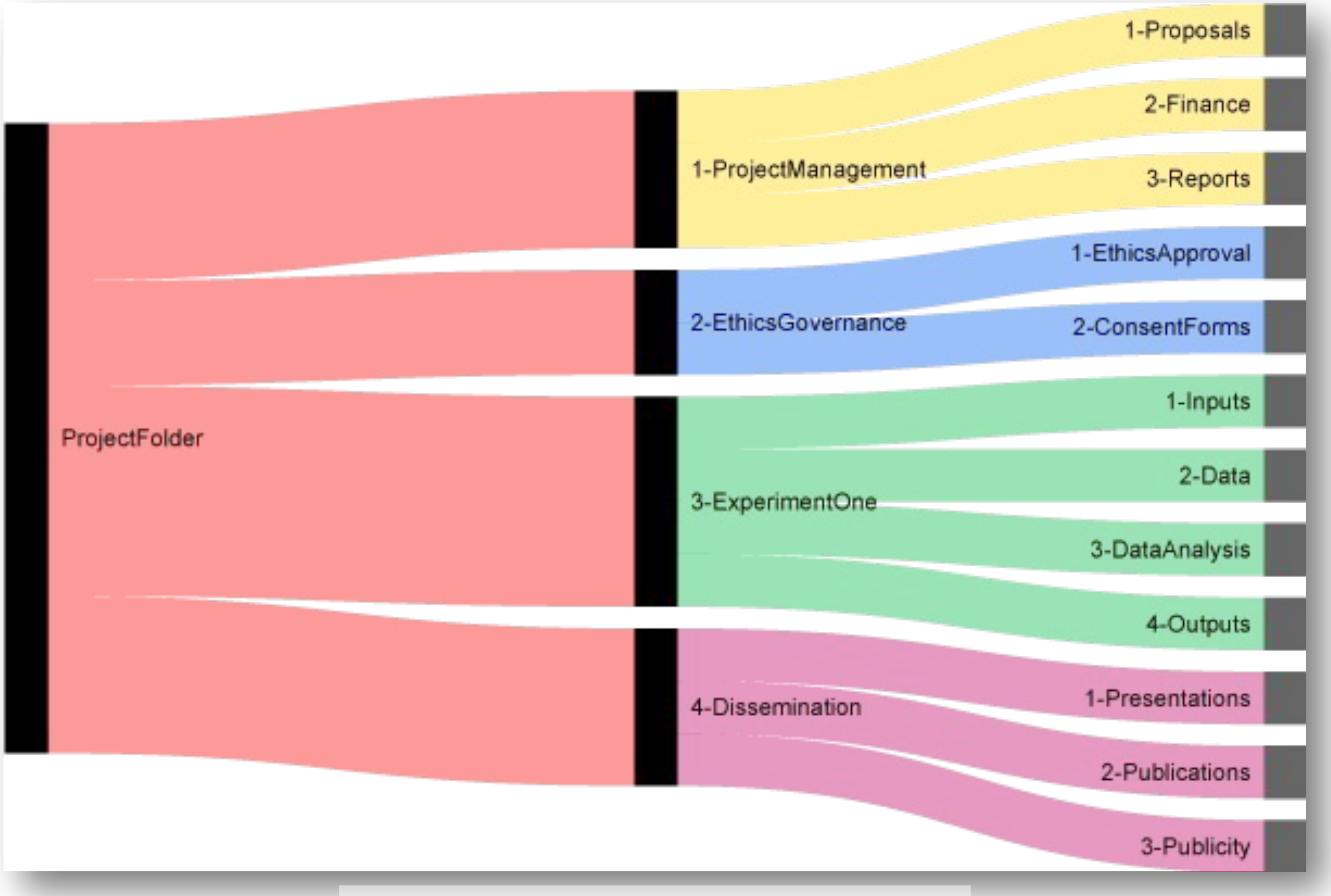
How do you organise your data?



- Consistent
- Meaningful to you and your colleagues
- Allow you to find files easily
- [Project] / [Experiment] / [Instrument or Type of file] / [Date]

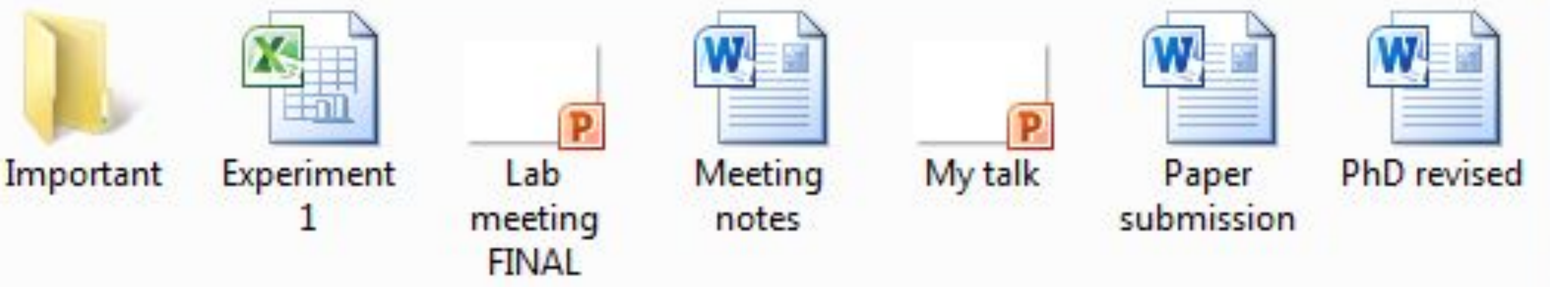
https://libraries.mit.edu/data-management/store/organize/

Example of folder organisation structure



Copyright: http://nikola.me/folder_structure.html

File naming



Important

Experiment 1

Lab meeting FINAL

Meeting notes

My talk

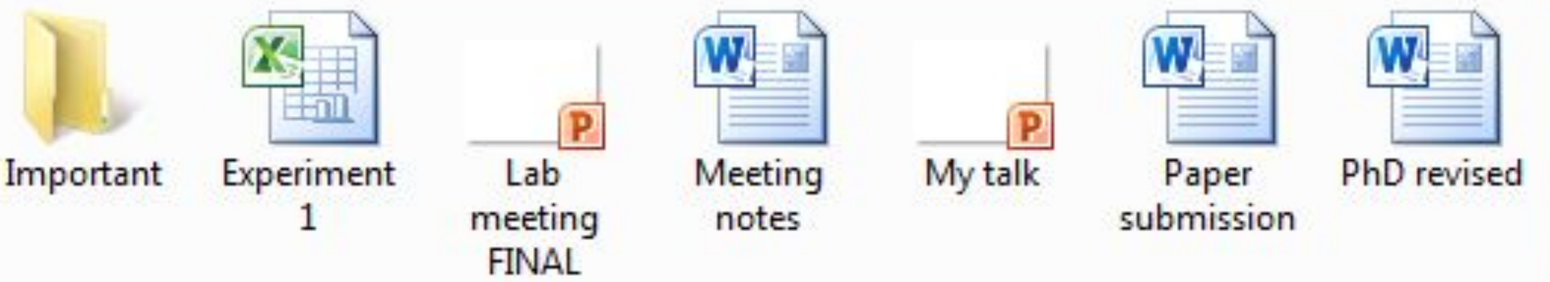
Paper submission

PhD revised

In 3 years time would you know what these are?

https://library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-naming

File naming



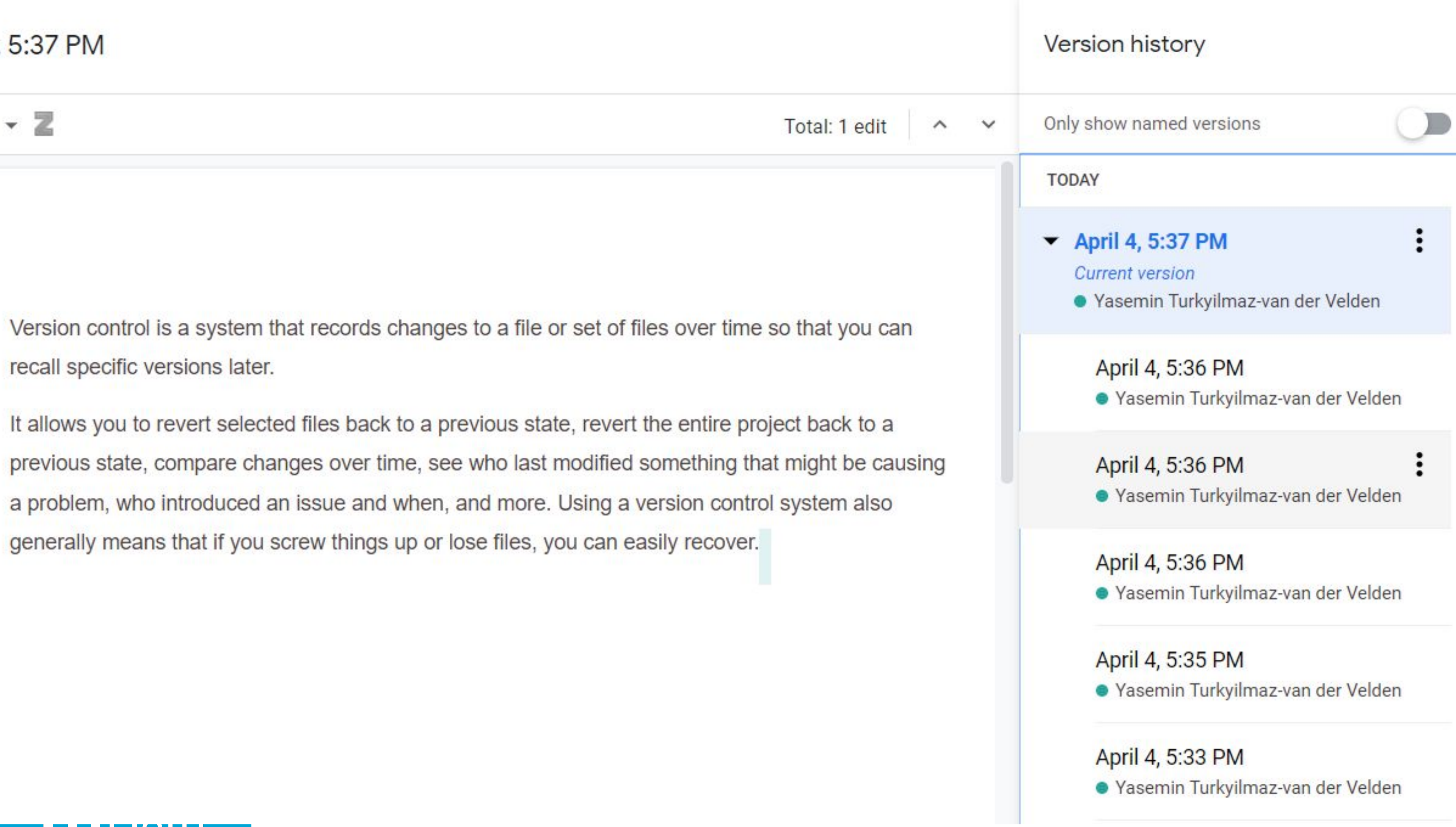
In 3 years time would you know what these are?

- Date or date range of experiment: YYYYMMDD
- File type
- Researcher name/initials
- Version number of file
- 20190723_RSG_Webinar_Presentation_YT2

- Don’t make file names too long
- Avoid special characters and spaces
- Include a README.txt file to explain the naming convention

https://library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-naming

Version Control



5:37 PM

Total: 1 edit

Version control is a system that records changes to a file or set of files over time so that you can recall specific versions later.

It allows you to revert selected files back to a previous state, revert the entire project back to a previous state, compare changes over time, see who last modified something that might be causing a problem, who introduced an issue and when, and more. Using a version control system also generally means that if you screw things up or lose files, you can easily recover.

Version history

Only show named versions

TODAY

April 4, 5:37 PM
Current version
● Yasemin Turkyilmaz-van der Velden

April 4, 5:36 PM
● Yasemin Turkyilmaz-van der Velden

April 4, 5:36 PM
● Yasemin Turkyilmaz-van der Velden

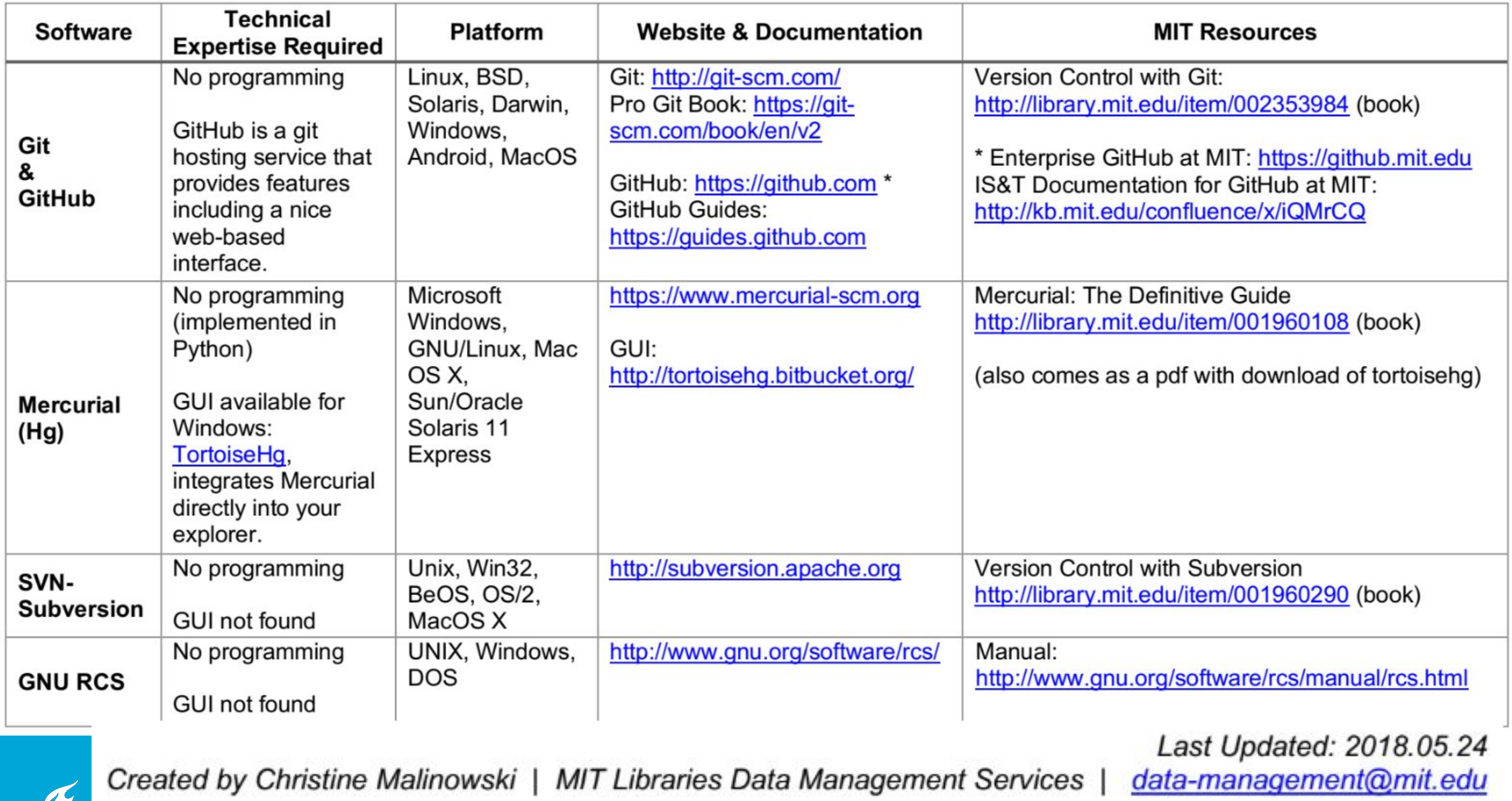
April 4, 5:36 PM
● Yasemin Turkyilmaz-van der Velden

April 4, 5:35 PM
● Yasemin Turkyilmaz-van der Velden

April 4, 5:33 PM
● Yasemin Turkyilmaz-van der Velden

https://git-scm.com/book/en/v2/Getting-Started-About-Version-Control

Version Control



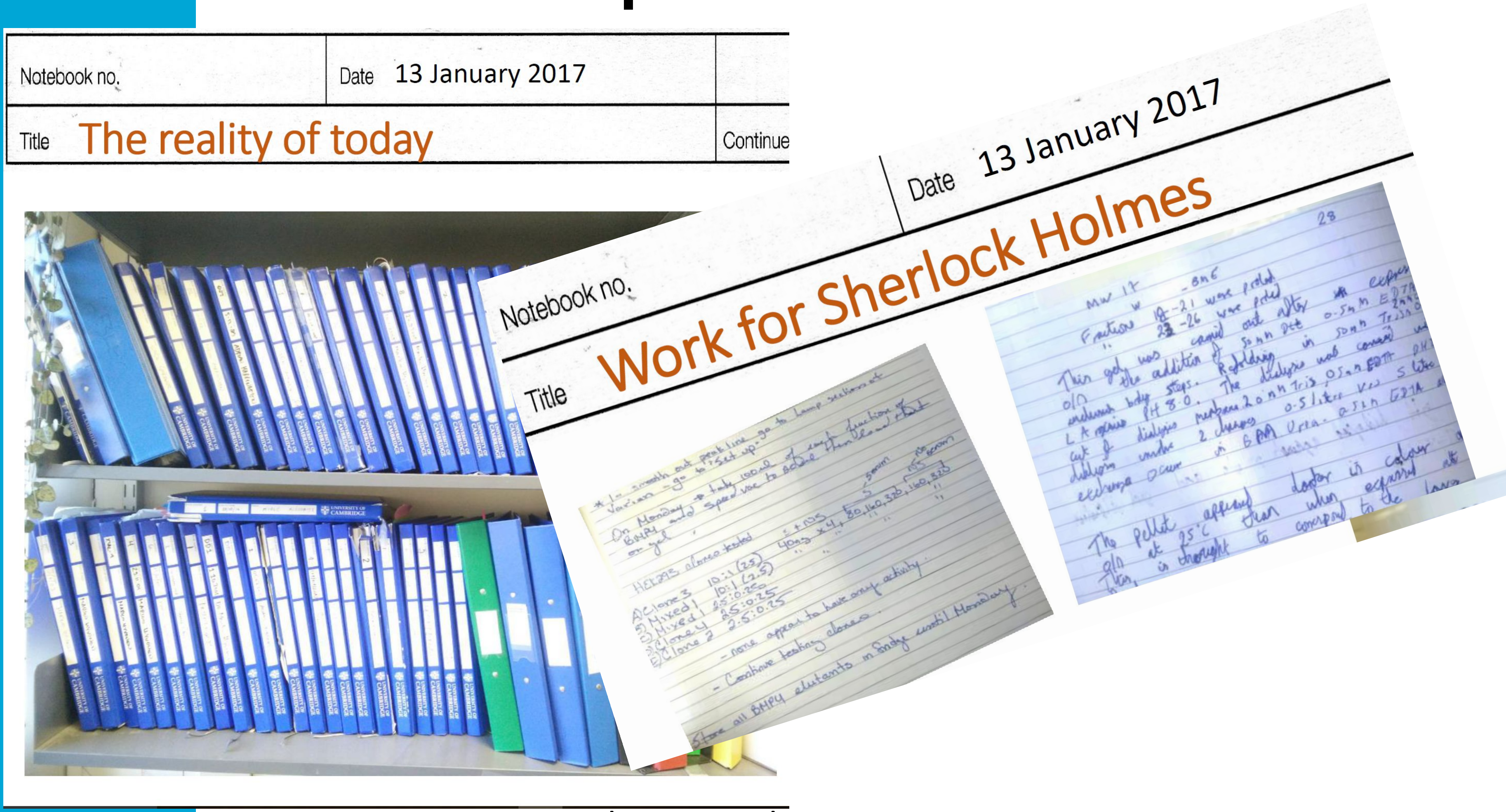
Software	Technical Expertise Required	Platform	Website & Documentation	MIT Resources
Git & GitHub	No programming GitHub is a git hosting service that provides features including a nice web-based interface.	Linux, BSD, Solaris, Darwin, Windows, Android, MacOS	Git: http://git-scm.com/ Pro Git Book: https://git-scm.com/book/en/v2 GitHub: https://github.com * GitHub Guides: https://guides.github.com	Version Control with Git: http://library.mit.edu/item/002353984 (book) * Enterprise GitHub at MIT: https://github.mit.edu IS&T Documentation for GitHub at MIT: http://kb.mit.edu/confluence/x/iQMrCQ
Mercurial (Hg)	No programming (implemented in Python) GUI available for Windows: TortoiseHg, integrates Mercurial directly into your explorer.	Microsoft Windows, GNU/Linux, Mac OS X, Sun/Oracle Solaris 11 Express	https://www.mercurial-scm.org GUI: http://tortoisehg.bitbucket.org/	Mercurial: The Definitive Guide http://library.mit.edu/item/001960108 (book) (also comes as a pdf with download of tortoisehg)
SVN-Subversion	No programming GUI not found	Unix, Win32, BeOS, OS/2, MacOS X	http://subversion.apache.org	Version Control with Subversion http://library.mit.edu/item/001960290 (book)
GNU RCS	No programming GUI not found	UNIX, Windows, DOS	http://www.gnu.org/software/rcs/	Manual: http://www.gnu.org/software/rcs/manual/rcs.html

Last Updated: 2018.05.24

Created by Christine Malinowski | MIT Libraries Data Management Services | data-management@mit.edu

https://www.dropbox.com/s/nfopvc8y7bmmx0v/Handout_versionControl.pdf?dl=0

Experimental notes



Notebook no. Date 13 January 2017

Title The reality of today Continue

Notebook no. Date 13 January 2017

Title Work for Sherlock Holmes

* 1- smooth out peak line go to Lamp sections at Varian - go to ?Set up-

On Monday → take 100ul of each fraction of BMPY and speed vac to adjust then load that on gel

HEK293 clones tested S + NS 40ng x4, 80,160,320, 160,320

A) Clone 3 10:1 (2.5)
B) Mixed 1 10:1 (2.5)
C) Mixed 1 2.5:0.250
D) Clone 4 2.5:0.25
E) Clone 2 2.5:0.25

- none appear to have any activity.

- Continue testing clones.

Store all BMPY elutants in fridge until Monday.

28

MW 17 w -8m6
Fractions 19-21 were pooled
" 23-26 were pooled
This gel was carried out after [illegible] o/n the addition of 50mM DTT in 0.5mM EDTA [illegible] Refolding in 50mM Tris [illegible] pH 8.0. The dialysis was [illegible] dialysis membrane 20mM Tris, 0.5mM EDTA [illegible] 2 changes 0.5 litres [illegible] in 6M Urea. 0.5mM EDTA [illegible] exchange occur

The pellet appeared darker in colour o/n at 25°C than when expressed at [illegible] is thought to correspond to the [illegible]

Presentation by Dr Marko Hyvonen

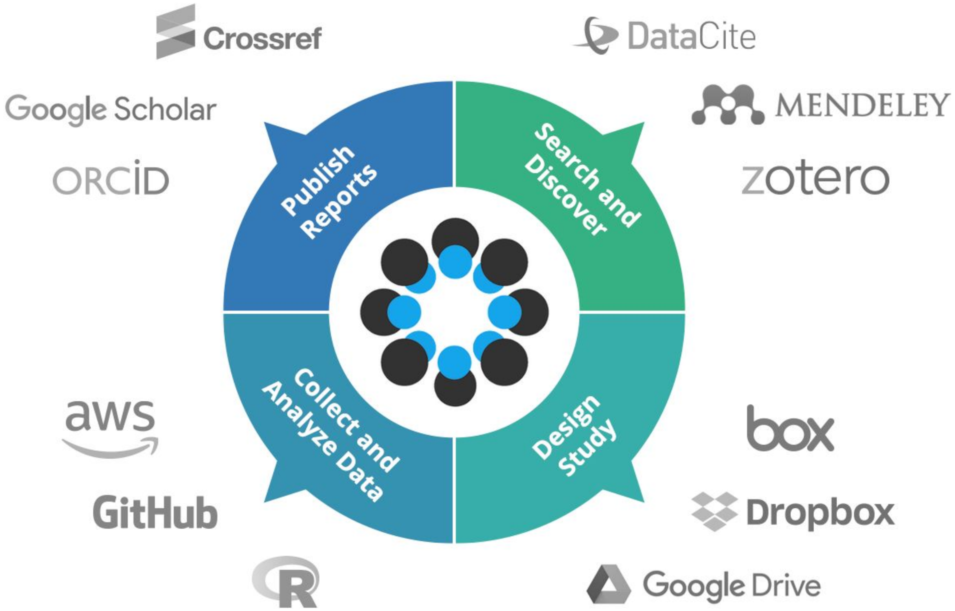
https://doi.org/10.17863/CAM.7217

Electronic Lab Notebooks

- Digital documentation, categorization and linking of
  - Raw, intermediate and final data
  - Experimental and measurement parameters
  - Samples
- Searchable
- Traceable (version control) & fraud-proof

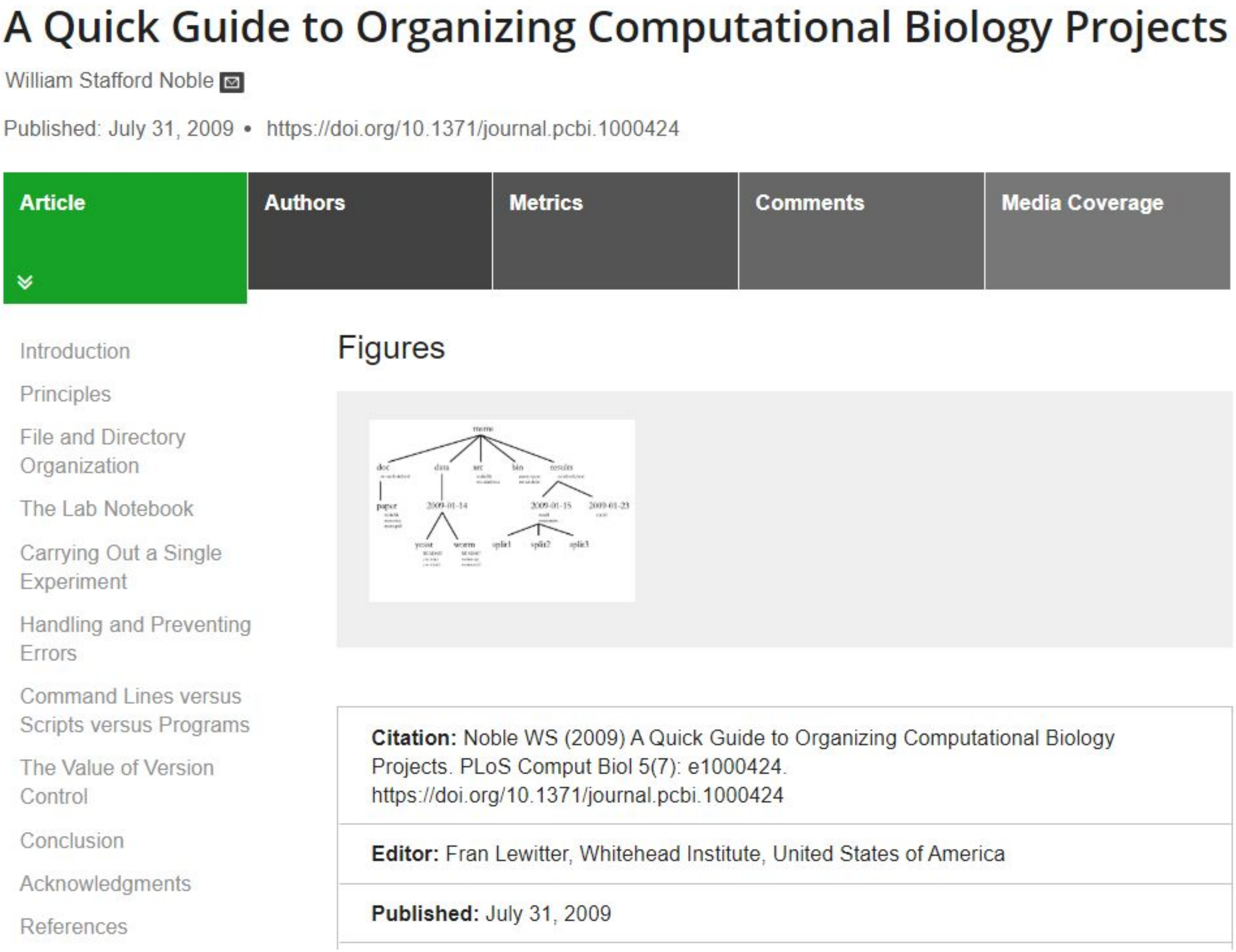
Report: https://doi.org/10.17605/OSF.IO/JR9U2
Talk on youtube:https://bit.ly/2Hlm41X

OSF: Open Science Framework



- **Free and open platform** for project workspaces
- **Collaborative** - share data within and beyond research groups
- **Version control**
- **Access control** at both project and file levels
- **Persistent identifiers**
- **Add-ons** such as Dropbox, GitHub, AWS, Google Drive, and Dataverse
- **Preregistration** of your research plans
- **Preprints**

https://osf.io/



A Quick Guide to Organizing Computational Biology Projects

William Stafford Noble

Published: July 31, 2009 • https://doi.org/10.1371/journal.pcbi.1000424

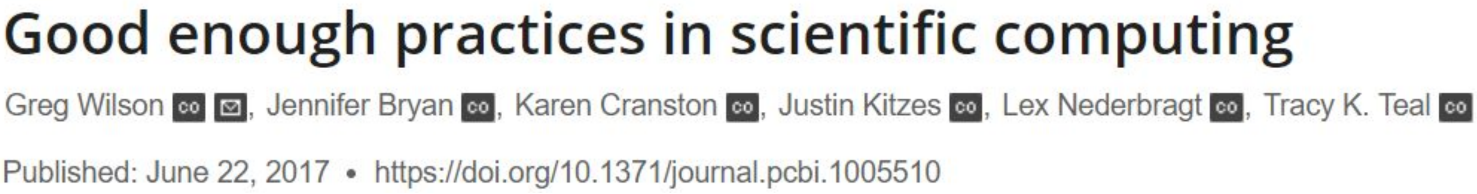
Figures

Citation: Noble WS (2009) A Quick Guide to Organizing Computational Biology Projects. PLoS Comput Biol 5(7): e1000424. https://doi.org/10.1371/journal.pcbi.1000424

Editor: Fran Lewitter, Whitehead Institute, United States of America

Published: July 31, 2009

https://doi.org/10.1371/journal.pcbi.1000424



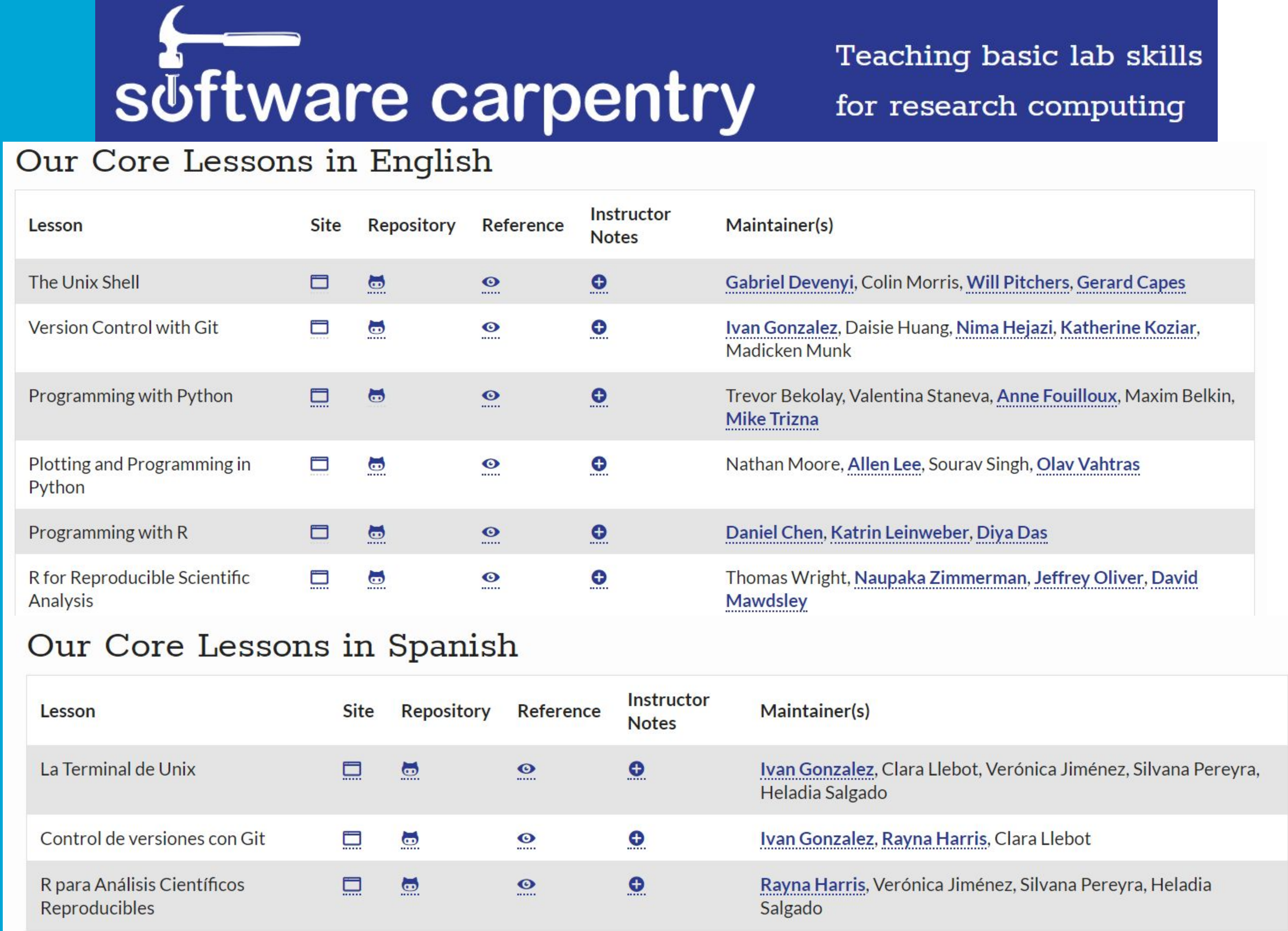
Good enough practices in scientific computing

Greg Wilson, Jennifer Bryan, Karen Cranston, Justin Kitzes, Lex Nederbragt, Tracy K. Teal

Published: June 22, 2017 • https://doi.org/10.1371/journal.pcbi.1005510

1. **Data management:**
   - saving both raw and intermediate forms, documenting all steps, creating tidy data amenable to analysis.
2. **Software:**
   - writing, organizing, and sharing scripts and programs used in an analysis.
3. **Collaboration:**
   - making it easy for existing and new collaborators to understand and contribute to a project.
4. **Project organization:**
   - organizing the digital artifacts of a project to ease discovery and understanding.
5. **Tracking changes:**
   - recording how various components of your project change over time.
6. **Manuscripts:**
   - writing manuscripts in a way that leaves an audit trail and minimizes manual merging of conflicts.

https://doi.org/10.1371/journal.pcbi.1005510



software carpentry

Teaching basic lab skills for research computing

Our Core Lessons in English

Lesson	Site	Repository	Reference	Instructor Notes	Maintainer(s)
The Unix Shell	▭	🐱	👁	⊕	Gabriel Devenyi, Colin Morris, Will Pitchers, Gerard Capes
Version Control with Git	▭	🐱	👁	⊕	Ivan Gonzalez, Daisie Huang, Nima Hejazi, Katherine Koziar, Madicken Munk
Programming with Python	▭	🐱	👁	⊕	Trevor Bekolay, Valentina Staneva, Anne Fouilloux, Maxim Belkin, Mike Trizna
Plotting and Programming in Python	▭	🐱	👁	⊕	Nathan Moore, Allen Lee, Sourav Singh, Olav Vahtras
Programming with R	▭	🐱	👁	⊕	Daniel Chen, Katrin Leinweber, Diya Das
R for Reproducible Scientific Analysis	▭	🐱	👁	⊕	Thomas Wright, Naupaka Zimmerman, Jeffrey Oliver, David Mawdsley

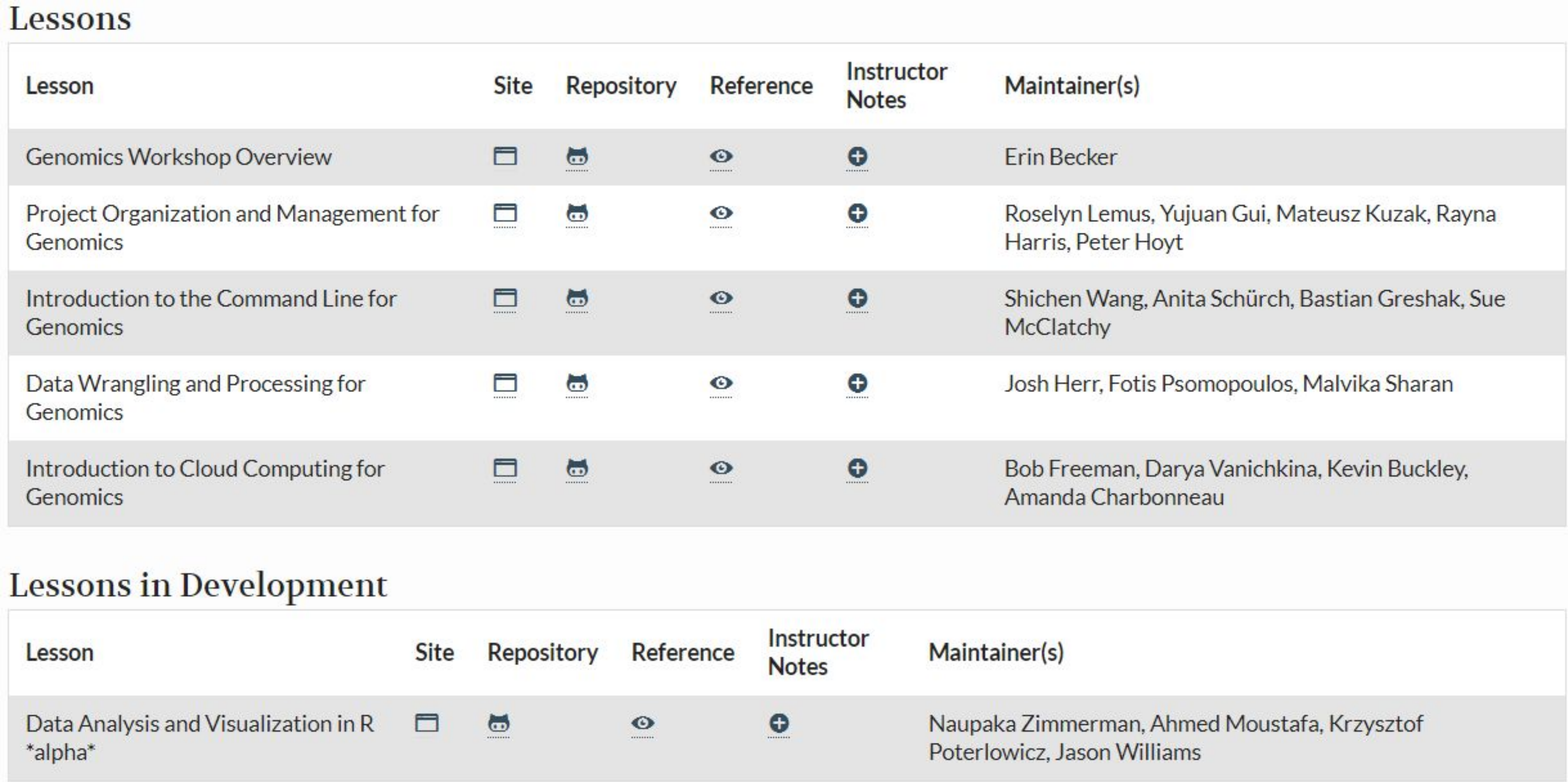
Our Core Lessons in Spanish

Lesson	Site	Repository	Reference	Instructor Notes	Maintainer(s)
La Terminal de Unix	▭	🐱	👁	⊕	Ivan Gonzalez, Clara Llebot, Verónica Jiménez, Silvana Pereyra, Heladia Salgado
Control de versiones con Git	▭	🐱	👁	⊕	Ivan Gonzalez, Rayna Harris, Clara Llebot
R para Análisis Científicos Reproducibles	▭	🐱	👁	⊕	Rayna Harris, Verónica Jiménez, Silvana Pereyra, Heladia Salgado

https://software-carpentry.org/lessons/index.html

DATA CARPENTRY

BUILDING COMMUNITIES TEACHING UNIVERSAL DATA LITERACY



Lessons

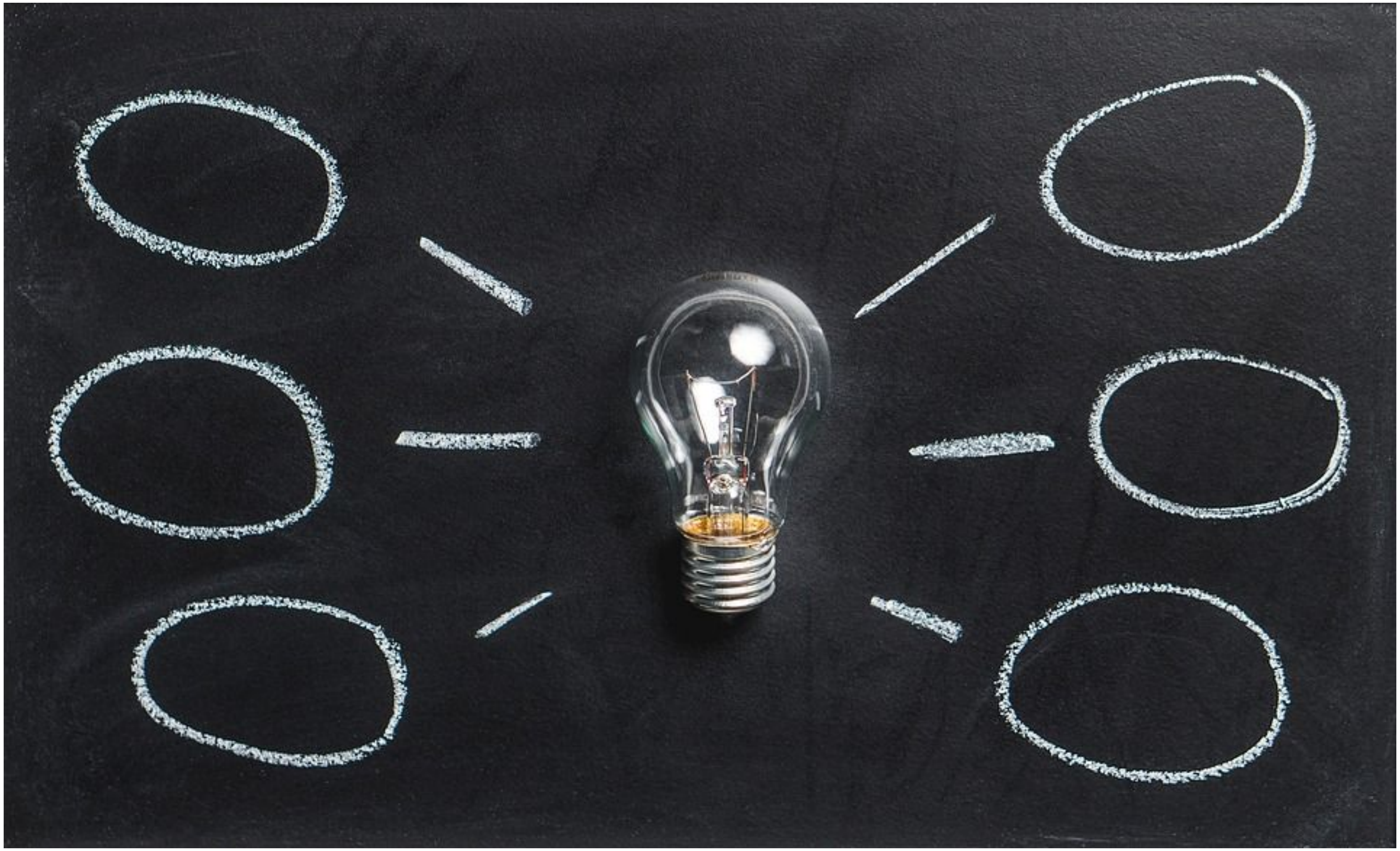
Lesson	Site	Repository	Reference	Instructor Notes	Maintainer(s)
Genomics Workshop Overview	▭	🐱	👁	⊕	Erin Becker
Project Organization and Management for Genomics	▭	🐱	👁	⊕	Roselyn Lemus, Yujuan Gui, Mateusz Kuzak, Rayna Harris, Peter Hoyt
Introduction to the Command Line for Genomics	▭	🐱	👁	⊕	Shichen Wang, Anita Schürch, Bastian Greshak, Sue McClatchy
Data Wrangling and Processing for Genomics	▭	🐱	👁	⊕	Josh Herr, Fotis Psomopoulos, Malvika Sharan
Introduction to Cloud Computing for Genomics	▭	🐱	👁	⊕	Bob Freeman, Darya Vanichkina, Kevin Buckley, Amanda Charbonneau

Lessons in Development

Lesson	Site	Repository	Reference	Instructor Notes	Maintainer(s)
Data Analysis and Visualization in R *alpha*	▭	🐱	👁	⊕	Naupaka Zimmerman, Ahmed Moustafa, Krzysztof Poterlowicz, Jason Williams

https://datacarpentry.org/lessons/#genomics-workshop

Thank you
Questions?



Slides are available:
https://doi.org/10.5281/zenodo.3346559