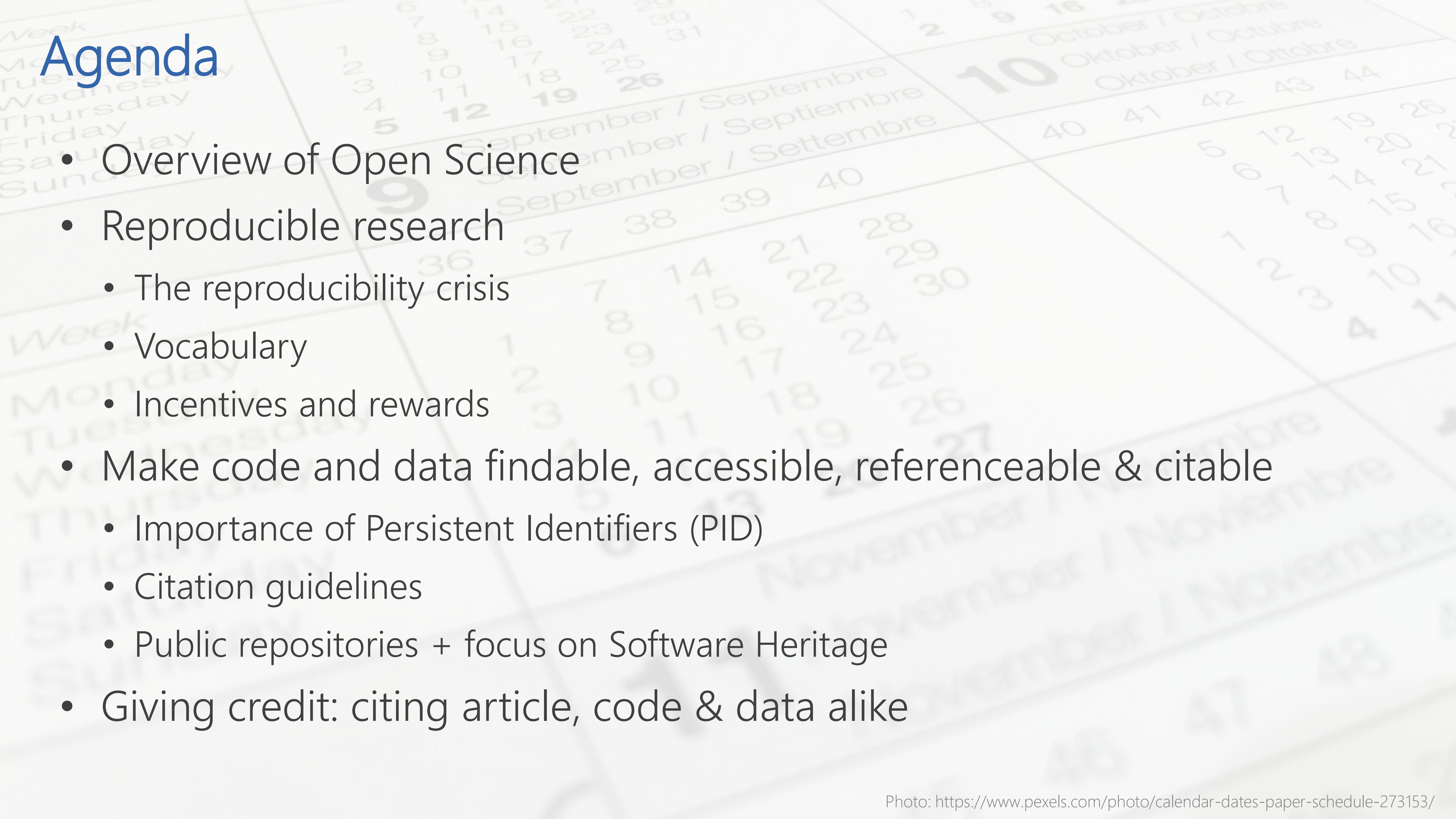
[article+code+data]:

A virtuous tryptic towards reproducible research

Research Publish Reproduce

Franck MICHEL

Agenda



- Overview of Open Science
- Reproducible research
  - The reproducibility crisis
  - Vocabulary
  - Incentives and rewards
- Make code and data findable, accessible, referenceable & citable
  - Importance of Persistent Identifiers (PID)
  - Citation guidelines
  - Public repositories + focus on Software Heritage
- Giving credit: citing article, code & data alike

Photo: https://www.pexels.com/photo/calendar-dates-paper-schedule-273153/

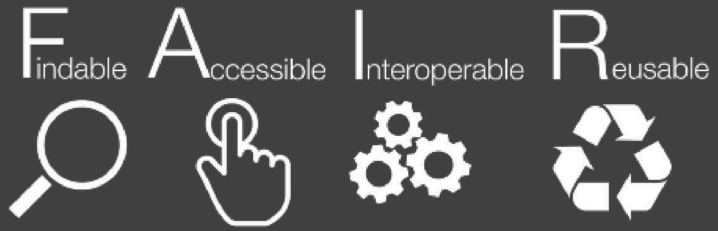
Science

Photo: https://www.pexels.com/photo/photo-of-open-signage-2763246/

Open Science

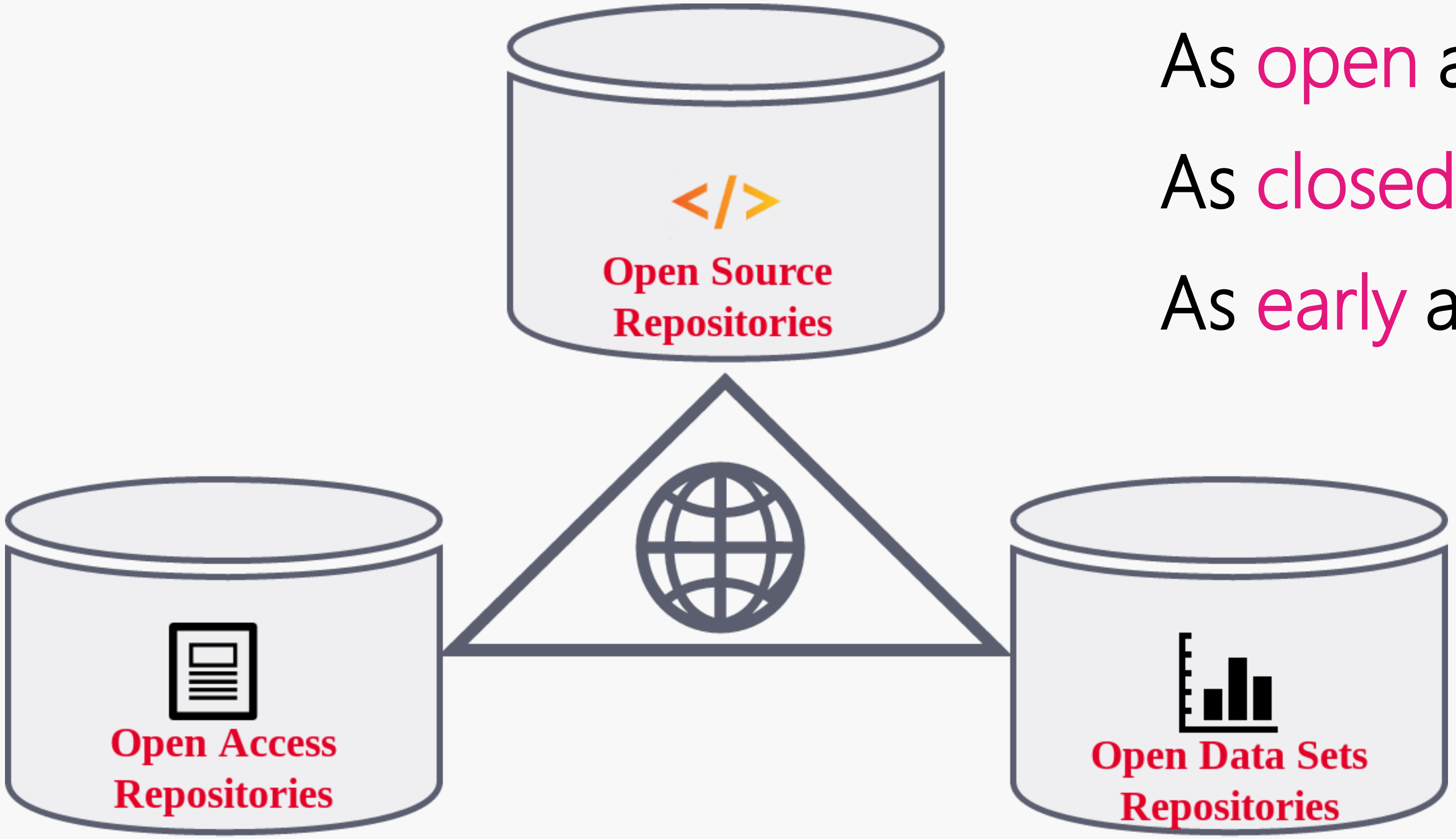
"is the unhindered spreading of the results, methods and products of scientific research.

*It is based on the (...) **open access to publications** and, as much as possible, to **data**, **source code** and **research methods**."**



*: *translated from https://www.ouvrirlascience.fr/deuxieme-plan-national-pour-la-science-ouverte/*

The Three Pillars of Open Science



As open as possible

As closed as necessary

As early as possible

Source: Software Heritage, 2019. https://www.softwareheritage.org/save-and-reference-research-software/

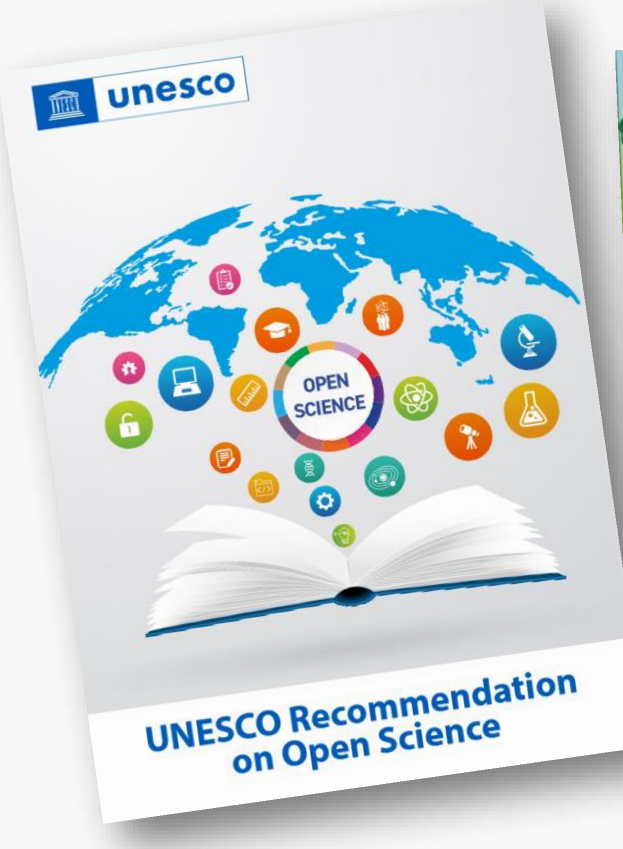
What benefits are expected from Open Science?

- Foster reproducible research:
  - A cornerstone of the scientific method: transparency and sharing of methods/code/data allow other researchers to reproduce experiments, verify results
  - **Non reproducible research is not science!**
- Make **scientific knowledge** accessible **to everyone**, regardless of location, institution or financial resources
- Increase scientific integrity
- Foster more effective collaboration of researchers across disciplines, institutions…
- Increase creativity through collective intelligence
- Accelerate scientific discovery and innovation: easier to build upon others' work, reuse vs. redo
- Increase public trust in science by making scientific research more accessible and understandable to non-experts.

How to verify or measure these claims?

S. Friesike, B. Fecher, & G.G. Wagner. **Open science: One term, five schools of thought**. In Opening science (pp. 17-47). Springer International Publishing (2014). DOI: 10.1007/978-3-319-00026-8_2

Open Science: a widely shared concern


unesco

OPEN SCIENCE

UNESCO Recommendation on Open Science

EUROPEAN UNION

HORIZON EUROPE

#HorizonEU

June 2021

"Horizon Europe will set a new standard for dissemination of knowledge and new skills across European societies. With clear and immediate open access requirements for beneficiaries, the Open Research Europe publishing platform and a strengthened European Open Science Cloud, we are well underway in making truly open science a reality".

Mariya Gabriel Commissioner for Innovation, Research, Culture, Education and Youth

OPEN SCIENCE

EARLY KNOWLEDGE AND DATA SHARING, AND OPEN COLLABORATION

The challenge is for Europe **to embrace open science as the modus operandi for all researchers**. Open science consists in the sharing of knowledge, data and tools as early as possible in the Research and Innovation (R&I) process, in open collaboration with all relevant knowledge actors, including academia, industry, public authorities, end users, citizens and society at large. Open science has the potential to increase the quality, efficiency and impact of R&I, lead to greater responsiveness to societal challenges, and increase trust of society in the science system.

What are open science practices?

- Open access to research outputs such as publications, data, software, models, algorithms, and workflows;
- Early and open sharing of research, for example through preregistration, registered reports, pre-prints, and crowd-sourcing of solutions to a specific problem;
- Use of open research infrastructures for knowledge and data sharing;
- Participation in open peer-review;
- Measures to ensure reproducibility of results; and

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR, DE LA RECHERCHE ET DE L'INNOVATION

Liberté Égalité Fraternité

Deuxième Plan national pour la science ouverte

GÉNÉRALISER LA SCIENCE OUVERTE EN FRANCE 2021-2024

The National Academies of SCIENCES · ENGINEERING · MEDICINE

CONSENSUS STUDY REPORT

OPEN SCIENCE BY DESIGN

Realizing a Vision for 21st Century Research

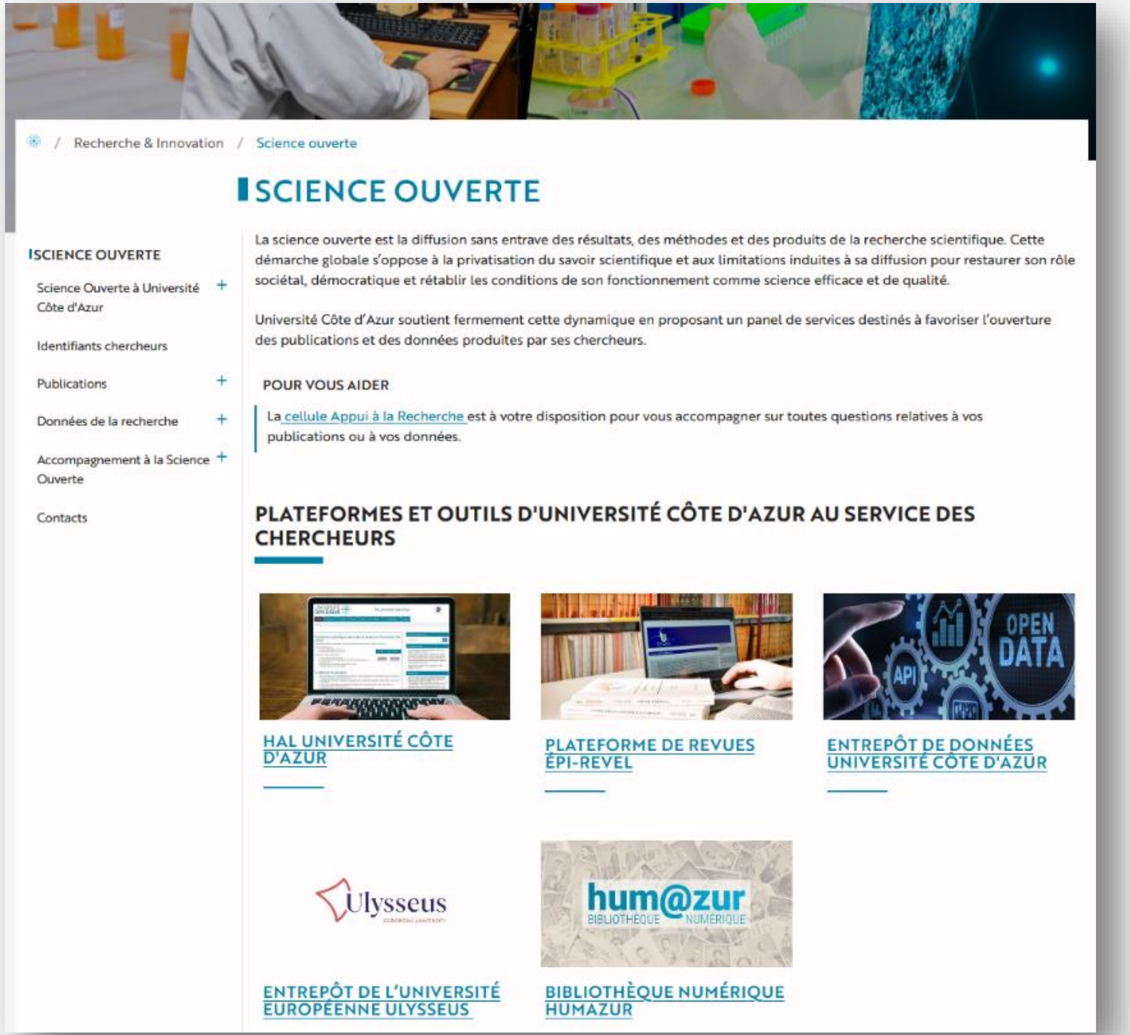
UNESCO. **UNESCO Recommendation on Open Science** (2021). https://www.unesco.org/en/legal-affairs/recommendation-open-science

European Commission, Directorate-General for Research and Innovation. **Horizon Europe, open science : early knowledge and data sharing, and open collaboration.** *Publications Office of the European Union* (2021). https://data.europa.eu/doi/10.2777/18252

Ministère de l'ES, la Recherche et l'Innovation. **Deuxième Plan national pour la science ouverte** (2021). https://www.ouvrirlascience.fr/deuxieme-plan-national-pour-la-science-ouverte/

National Academies of Sciences, Engineering, and Medicine, Policy and Global Affairs, Board on Research Data and Information, Committee on Toward an Open Science Enterprise. **Open Science by Design: Realizing a Vision for 21st Century Research** (2018).

Open Science @ UniCA



/ Recherche & Innovation / Science ouverte

SCIENCE OUVERTE

SCIENCE OUVERTE

- Science Ouverte à Université Côte d'Azur +
- Identifiants chercheurs
- Publications +
- Données de la recherche +
- Accompagnement à la Science Ouverte +
- Contacts

La science ouverte est la diffusion sans entrave des résultats, des méthodes et des produits de la recherche scientifique. Cette démarche globale s'oppose à la privatisation du savoir scientifique et aux limitations induites à sa diffusion pour restaurer son rôle sociétal, démocratique et rétablir les conditions de son fonctionnement comme science efficace et de qualité.

Université Côte d'Azur soutient fermement cette dynamique en proposant un panel de services destinés à favoriser l'ouverture des publications et des données produites par ses chercheurs.

POUR VOUS AIDER

La cellule Appui à la Recherche est à votre disposition pour vous accompagner sur toutes questions relatives à vos publications ou à vos données.

PLATEFORMES ET OUTILS D'UNIVERSITÉ CÔTE D'AZUR AU SERVICE DES CHERCHEURS

HAL UNIVERSITÉ CÔTE D'AZUR

PLATEFORME DE REVUES ÉPI-REVEL

ENTREPÔT DE DONNÉES UNIVERSITÉ CÔTE D'AZUR

ENTREPÔT DE L'UNIVERSITÉ EUROPÉENNE ULYSSEUS

BIBLIOTHÈQUE NUMÉRIQUE HUMAZUR

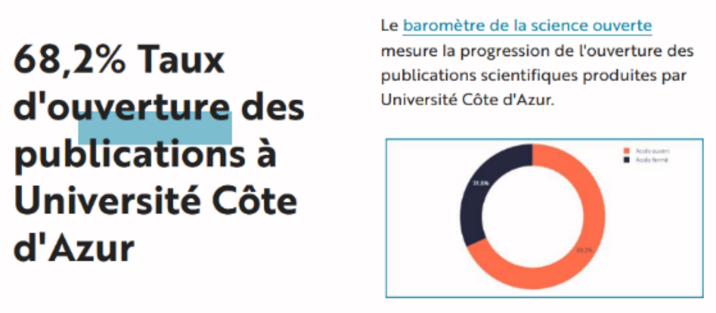
https://univ-cotedazur.fr/recherche-innovation/science-ouverte

Trainings & masterclasses for PhD/master students and researchers

(https://univ-cotedazur.fr/recherche-innovation/science-ouverte/accompagnement-a-la-science-ouverte/formations-a-la-science-ouverte)

Multiple resources & guides

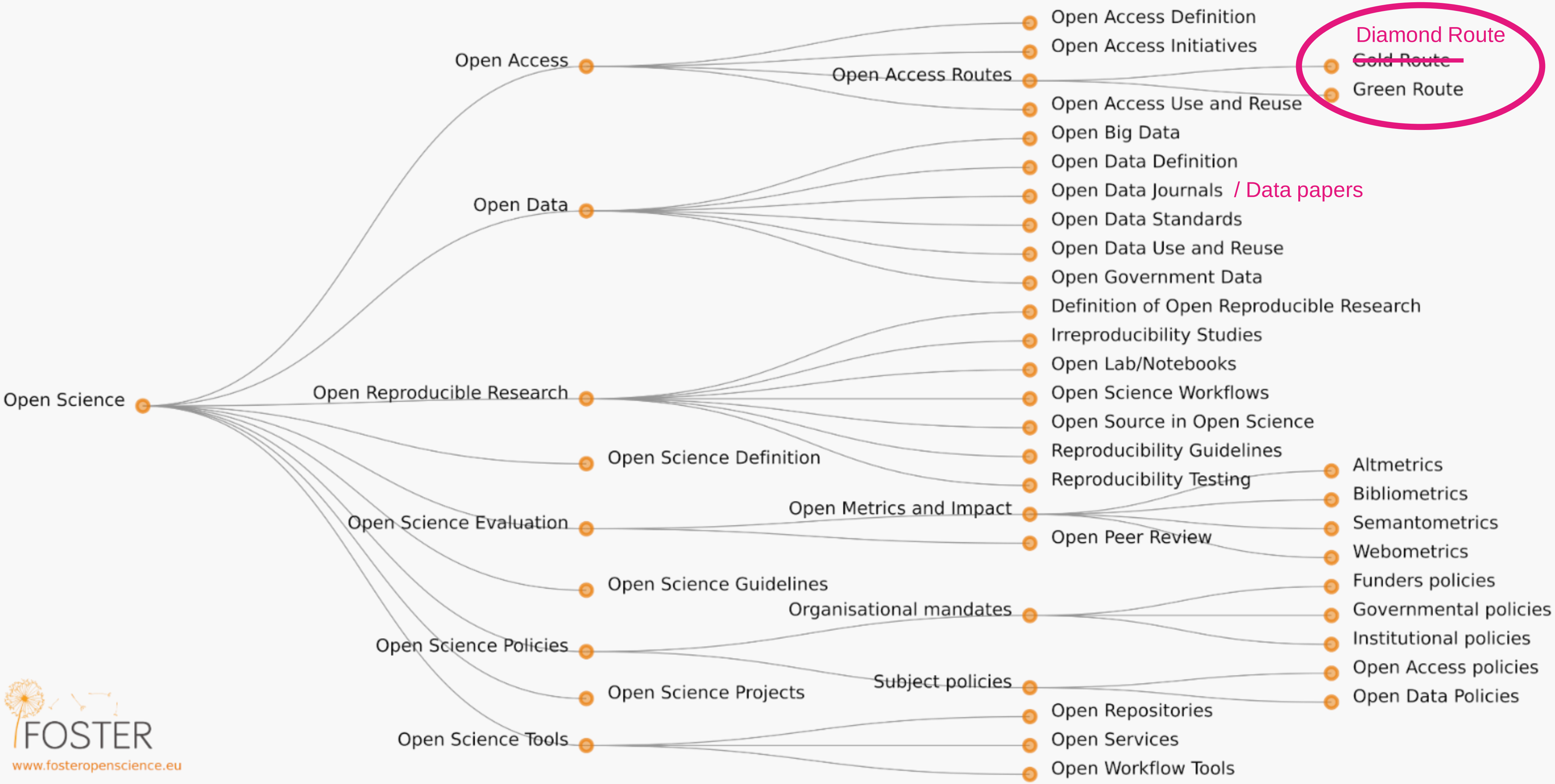
- Publishing articles/manuscripts in open access
- Licenses
- Where to store data, how to write a DMP
- DOIs, researcher ids…
- Open Science Barometer
  (https://apps-scd.univ-cotedazur.fr/barometre-science-ouverte/dashboard-publications)



68,2% Taux d'ouverture des publications à Université Côte d'Azur

Le baromètre de la science ouverte mesure la progression de l'ouverture des publications scientifiques produites par Université Côte d'Azur.

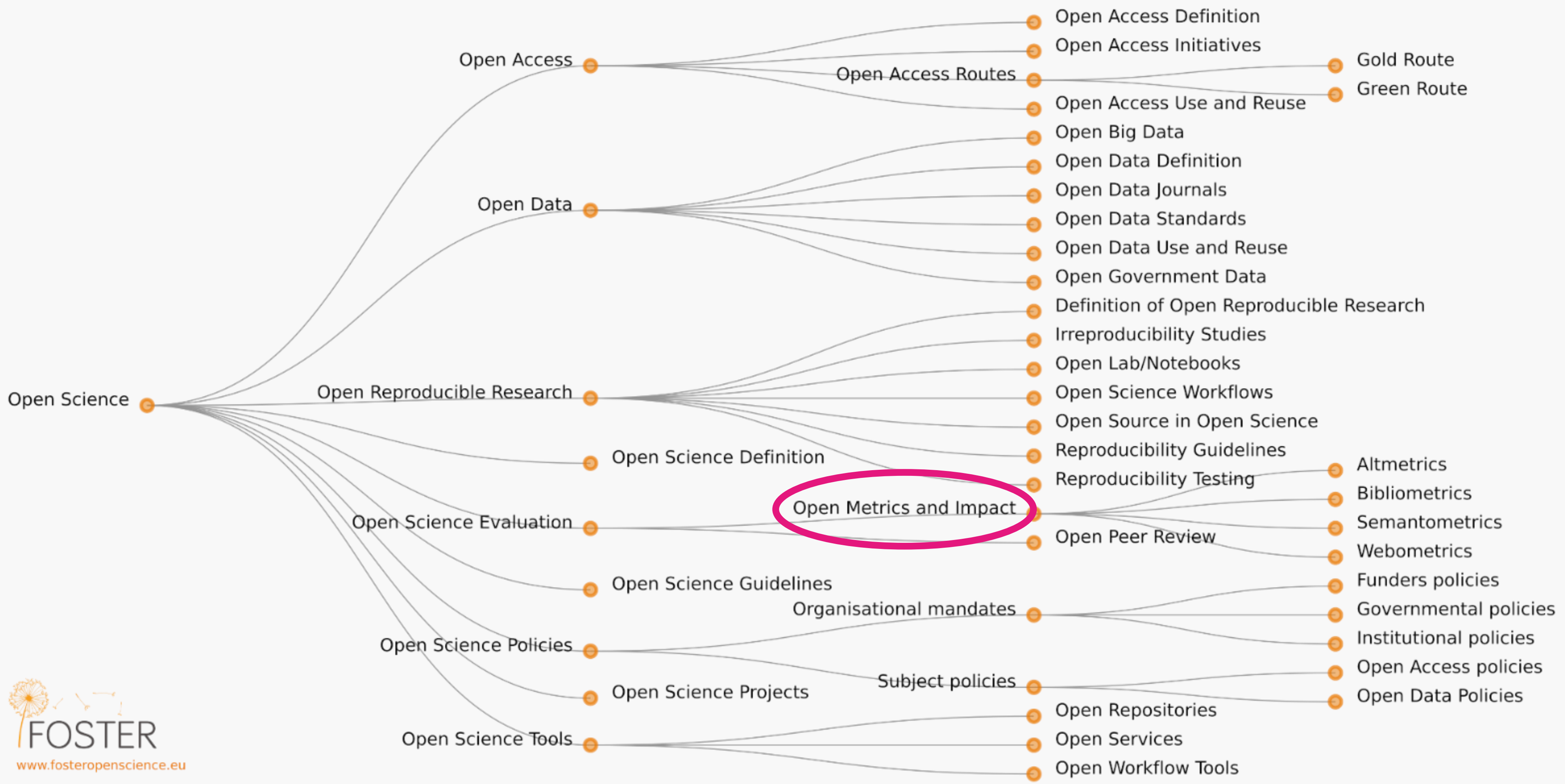
How can I do my part? Check the OS taxonomy!



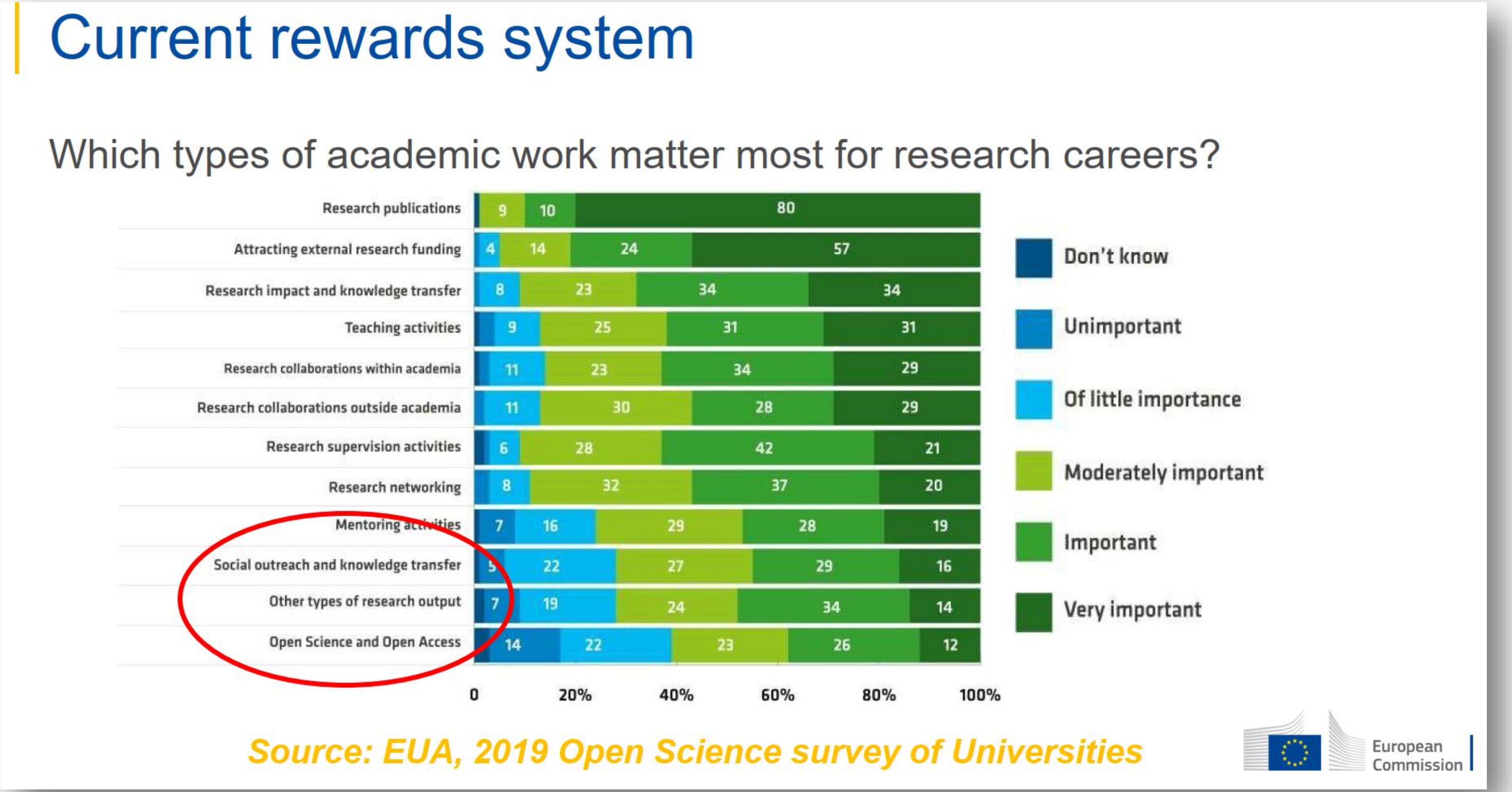
- Open Science
  - Open Access
    - Open Access Definition
    - Open Access Initiatives
    - Open Access Routes
      - Diamond Route
      - ~~Gold Route~~
      - Green Route
    - Open Access Use and Reuse
  - Open Data
    - Open Big Data
    - Open Data Definition
    - Open Data Journals / Data papers
    - Open Data Standards
    - Open Data Use and Reuse
    - Open Government Data
  - Open Reproducible Research
    - Definition of Open Reproducible Research
    - Irreproducibility Studies
    - Open Lab/Notebooks
    - Open Science Workflows
    - Open Source in Open Science
    - Reproducibility Guidelines
    - Reproducibility Testing
  - Open Science Definition
  - Open Science Evaluation
    - Open Metrics and Impact
      - Altmetrics
      - Bibliometrics
      - Semantometrics
      - Webometrics
    - Open Peer Review
  - Open Science Guidelines
  - Open Science Policies
    - Organisational mandates
      - Funders policies
      - Governmental policies
      - Institutional policies
    - Subject policies
      - Open Access policies
      - Open Data Policies
  - Open Science Projects
  - Open Science Tools
    - Open Repositories
    - Open Services
    - Open Workflow Tools

Source: https://www.fosteropenscience.eu/taxonomy/term/134

How can I do my part? Check the OS taxonomy!



Declaration on Research Assessment (DORA) https://sfdora.org/read/
Coalition for Advancing Research Assessment (CoARA) https://coara.eu/



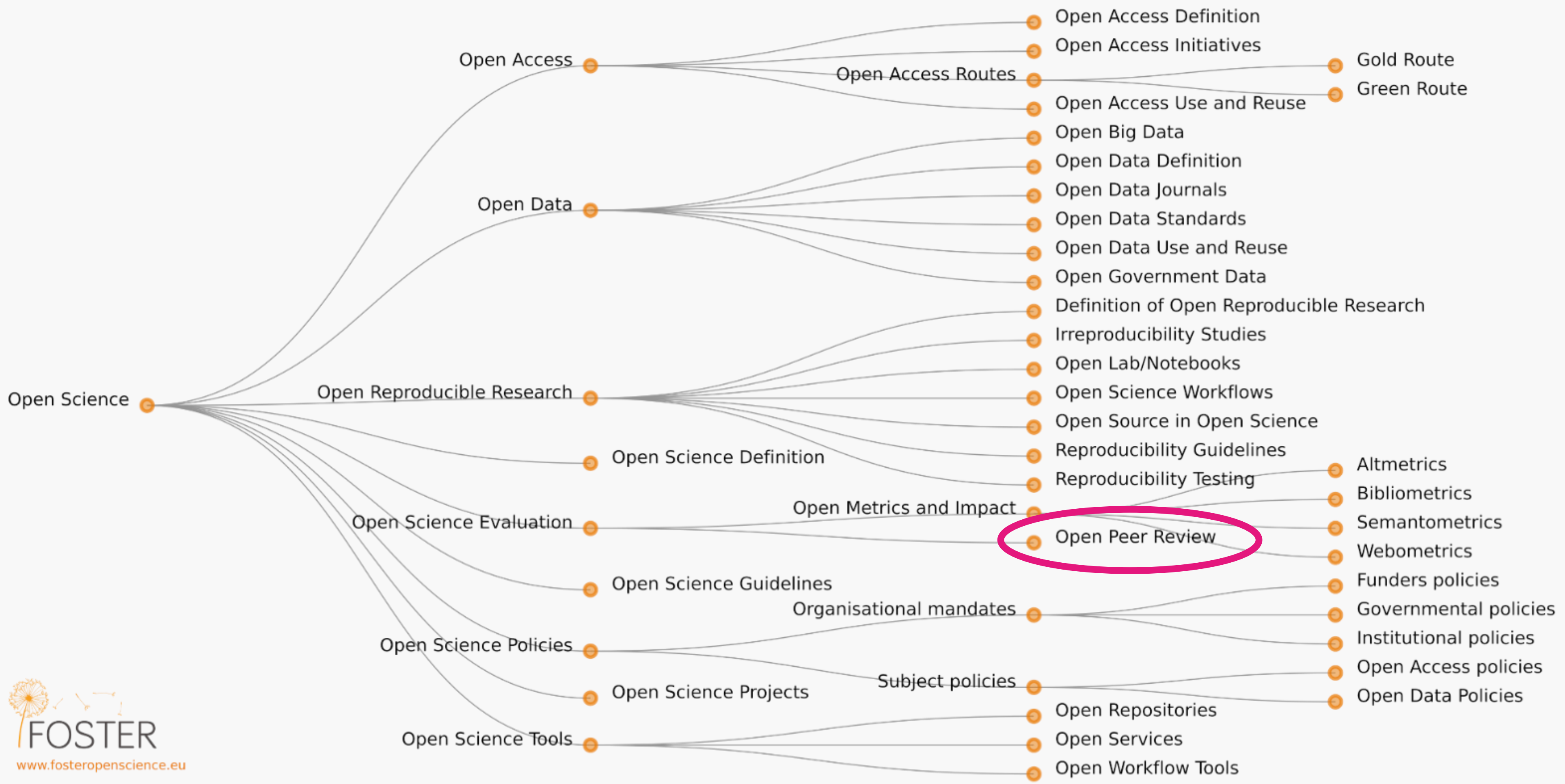
Current rewards system

Which types of academic work matter most for research careers?

Source: EUA, 2019 Open Science survey of Universities

Slide by Kostas Glinos. Challenges for the 21st century science.
Open Science Seminar, University Côte d'Azur, 2023-05-05.
https://drive.google.com/drive/folders/1ARjE-G8vWeSPmygLcMZiiVWmXKh1SDAk

How can I do my part? Check the OS taxonomy!



- Open Science
  - Open Access
    - Open Access Definition
    - Open Access Initiatives
    - Open Access Routes
      - Gold Route
      - Green Route
    - Open Access Use and Reuse
  - Open Data
    - Open Big Data
    - Open Data Definition
    - Open Data Journals
    - Open Data Standards
    - Open Data Use and Reuse
    - Open Government Data
  - Open Reproducible Research
    - Definition of Open Reproducible Research
    - Irreproducibility Studies
    - Open Lab/Notebooks
    - Open Science Workflows
    - Open Source in Open Science
    - Reproducibility Guidelines
    - Reproducibility Testing
  - Open Science Definition
  - Open Science Evaluation
    - Open Metrics and Impact
      - Altmetrics
      - Bibliometrics
      - Semantometrics
      - Webometrics
    - Open Peer Review
  - Open Science Guidelines
  - Open Science Policies
    - Organisational mandates
      - Funders policies
      - Governmental policies
      - Institutional policies
    - Subject policies
      - Open Access policies
      - Open Data Policies
  - Open Science Projects
  - Open Science Tools
    - Open Repositories
    - Open Services
    - Open Workflow Tools

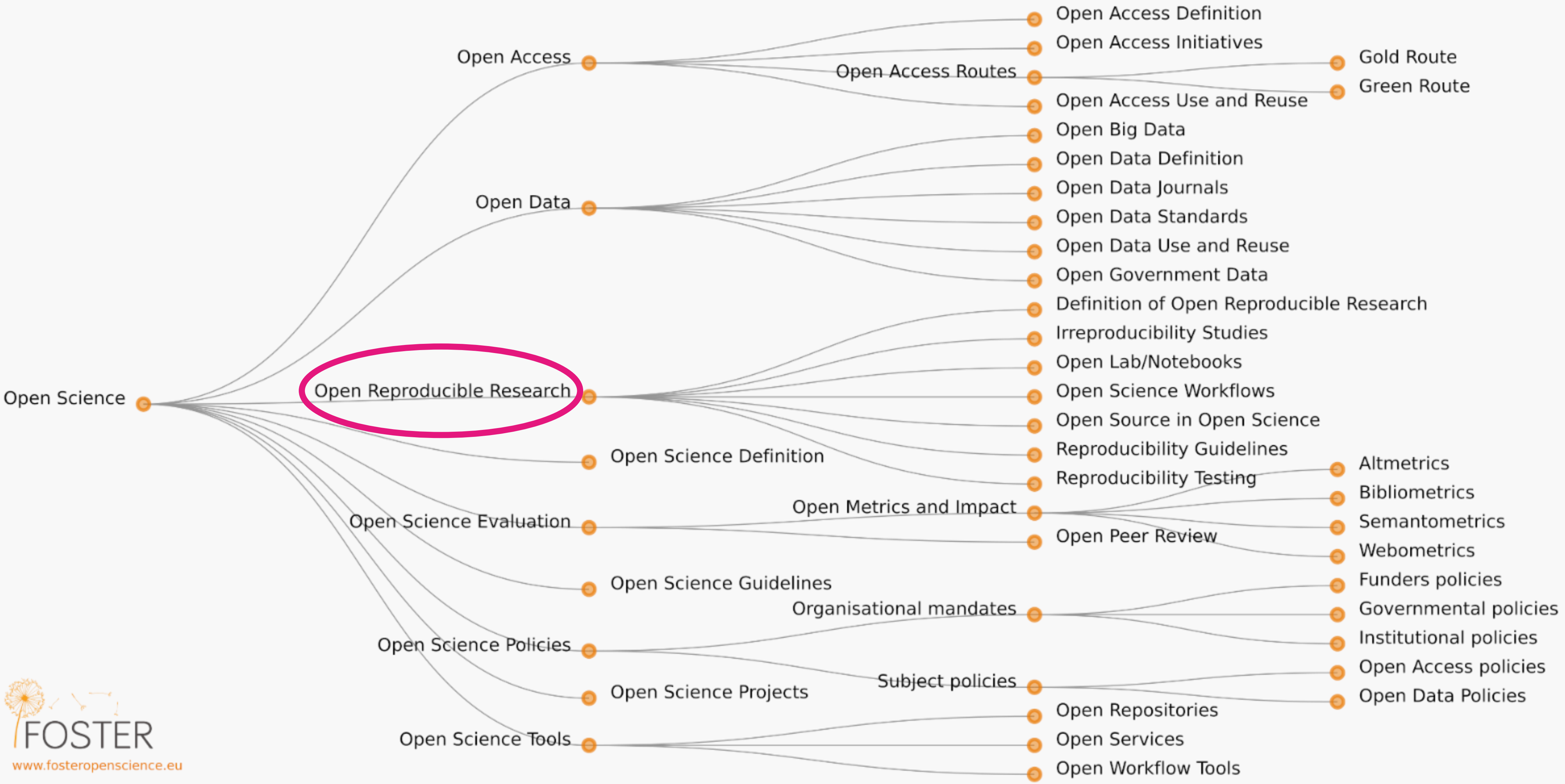
FOSTER

www.fosteropenscience.eu

"20% of the researchers performed 69% to 94% of the reviews. Among researchers actually contributing to peer review, 70% dedicated 1% or less of their research work-time to peer review while 5% dedicated 13% or more of it".

M. Kovanis, R. Porcher, P. Ravaud, L. Trinquart. The Global Burden of Journal Peer Review in the Biomedical Literature: Strong Imbalance in the Collective Enterprise. PLOSOne, 2016. https://doi.org/10.1371/journal.pone.0166387

How can I do my part? Check the OS taxonomy!



Source: https://www.fosteropenscience.eu/taxonomy/term/134

Agenda

- Overview of Open Science
- Reproducible research
  - The reproducibility crisis
  - Vocabulary
  - Incentives and rewards
- Make code and data findable, accessible, referenceable & citable
  - Importance of Persistent Identifiers (PID)
  - Citation guidelines
  - Public repositories + focus on Software Heritage
- Giving credit: citing article, code & data alike

Photo: https://www.pexels.com/photo/calendar-dates-paper-schedule-273153/

The “crisis” of reproducibility

Photo: https://www.flickr.com/photos/franckmichel/44548013464/

Repeat
Replicate
Reproduce
Reuse

?

Same words used differently in different contexts.

L. A. Barba. **Terminologies for Reproducible Research.** ArXiv preprint. 2018, https://doi.org/10.48550/arXiv.1802.03311.

Repeat

Same experiment

Same setup

Same lab

S. Cohen-Boulakia, K. Belhajjame, O. Collin, J. Chopard, C. Froidevaux, A. Gaignard, K. Hinsen, P. Larmande, Y. Le Bras, F. Lemoine, F. Mareuil, H. Ménager, C. Pradal, C. Blanchet. Scientific workflows for computational reproducibility in the life sciences: Status, challenges and opportunities. Future Generation Computer Systems, Volume 75, 2017, https://doi.org/10.1016/j.future.2017.01.012 .

Repeat > Replicate

Same experiment	Same experiment
Same setup	Same setup
Same lab	~~Same lab~~

S. Cohen-Boulakia, K. Belhajjame, O. Collin, J. Chopard, C. Froidevaux, A. Gaignard, K. Hinsen, P. Larmande, Y. Le Bras, F. Lemoine, F. Mareuil, H. Ménager, C. Pradal, C. Blanchet. Scientific workflows for computational reproducibility in the life sciences: Status, challenges and opportunities. Future Generation Computer Systems, Volume 75, 2017, https://doi.org/10.1016/j.future.2017.01.012 .

Repeat > Replicate > Reproduce

Same experiment	Same experiment	Same experiment
Same setup	Same setup	~~Same setup~~
Same lab	~~Same lab~~	~~Same lab~~

S. Cohen-Boulakia, K. Belhajjame, O. Collin, J. Chopard, C. Froidevaux, A. Gaignard, K. Hinsen, P. Larmande, Y. Le Bras, F. Lemoine, F. Mareuil, H. Ménager, C. Pradal, C. Blanchet. Scientific workflows for computational reproducibility in the life sciences: Status, challenges and opportunities. Future Generation Computer Systems, Volume 75, 2017, https://doi.org/10.1016/j.future.2017.01.012 .

Repeat > Replicate > Reproduce > Reuse

Same experiment	Same experiment	Same experiment	New ideas,
Same setup	Same setup	~~Same setup~~	new experiment,
Same lab	~~Same lab~~	~~Same lab~~	new data

S. Cohen-Boulakia, K. Belhajjame, O. Collin, J. Chopard, C. Froidevaux, A. Gaignard, K. Hinsen, P. Larmande, Y. Le Bras, F. Lemoine, F. Mareuil, H. Ménager, C. Pradal, C. Blanchet. Scientific workflows for computational reproducibility in the life sciences: Status, challenges and opportunities. Future Generation Computer Systems, Volume 75, 2017, https://doi.org/10.1016/j.future.2017.01.012 .

Repeat > Replicate > Reproduce > Reuse

continuum

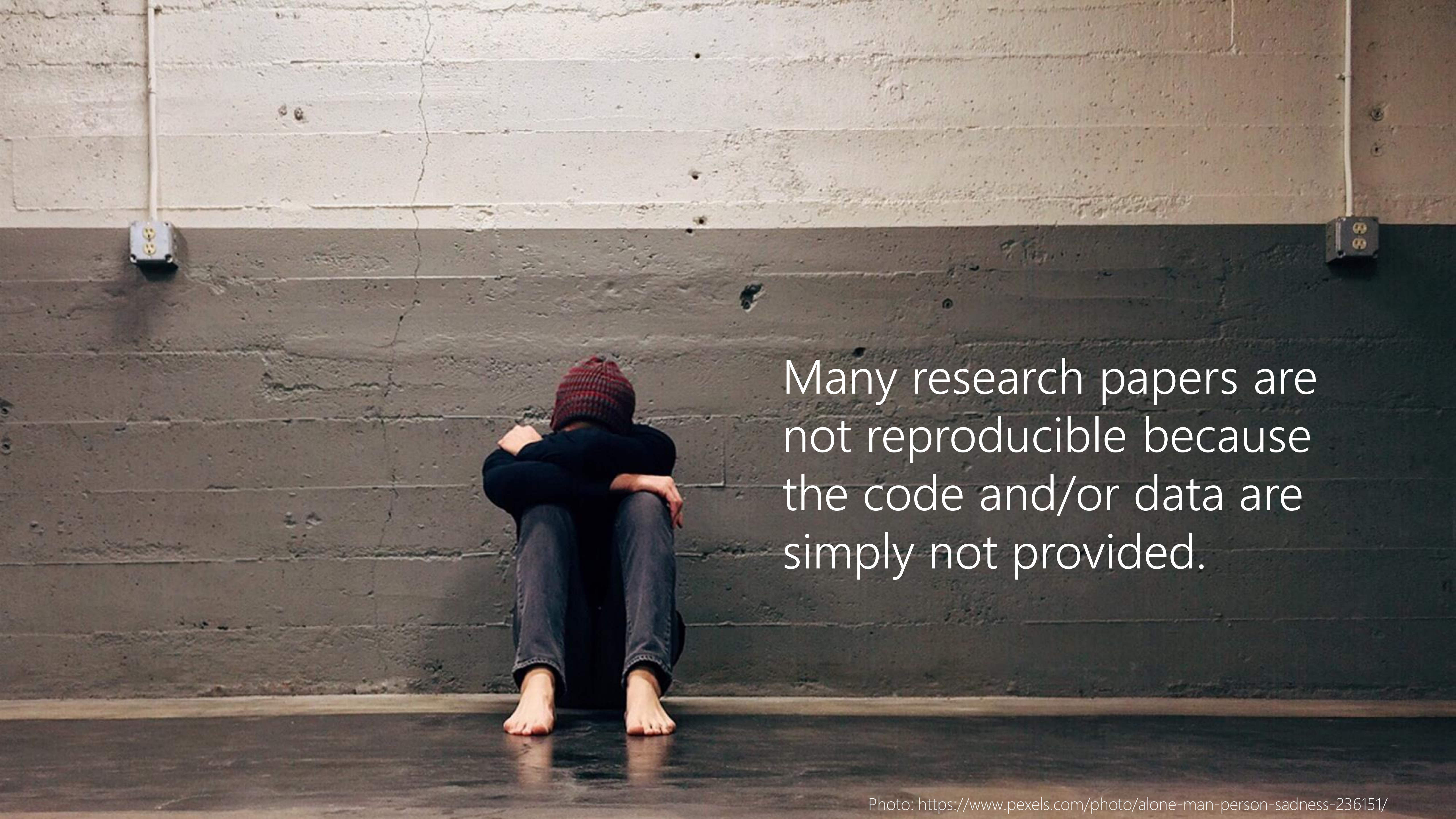
Same experiment	Same experiment	Same experiment	New ideas,
Same setup	Same setup	~~Same setup~~	new experiment,
Same lab	~~Same lab~~	~~Same lab~~	new data

S. Cohen-Boulakia, K. Belhajjame, O. Collin, J. Chopard, C. Froidevaux, A. Gaignard, K. Hinsen, P. Larmande, Y. Le Bras, F. Lemoine, F. Mareuil, H. Ménager, C. Pradal, C. Blanchet. **Scientific workflows for computational reproducibility in the life sciences: Status, challenges and opportunities.** Future Generation Computer Systems, Volume 75, 2017, https://doi.org/10.1016/j.future.2017.01.012 .

Repeat > Replicate > Reproduce > Reuse

Same experiment	Same experiment	Same experiment	New ideas,
Same setup	Same setup	~~Same setup~~	new experiment,
Same lab	~~Same lab~~	~~Same lab~~	new data

S. Cohen-Boulakia, K. Belhajjame, O. Collin, J. Chopard, C. Froidevaux, A. Gaignard, K. Hinsen, P. Larmande, Y. Le Bras, F. Lemoine, F. Mareuil, H. Ménager, C. Pradal, C. Blanchet. Scientific workflows for computational reproducibility in the life sciences: Status, challenges and opportunities. Future Generation Computer Systems, Volume 75, 2017, https://doi.org/10.1016/j.future.2017.01.012 .



Many research papers are not reproducible because the code and/or data are simply not provided.

Photo: https://www.pexels.com/photo/alone-man-person-sadness-236151/

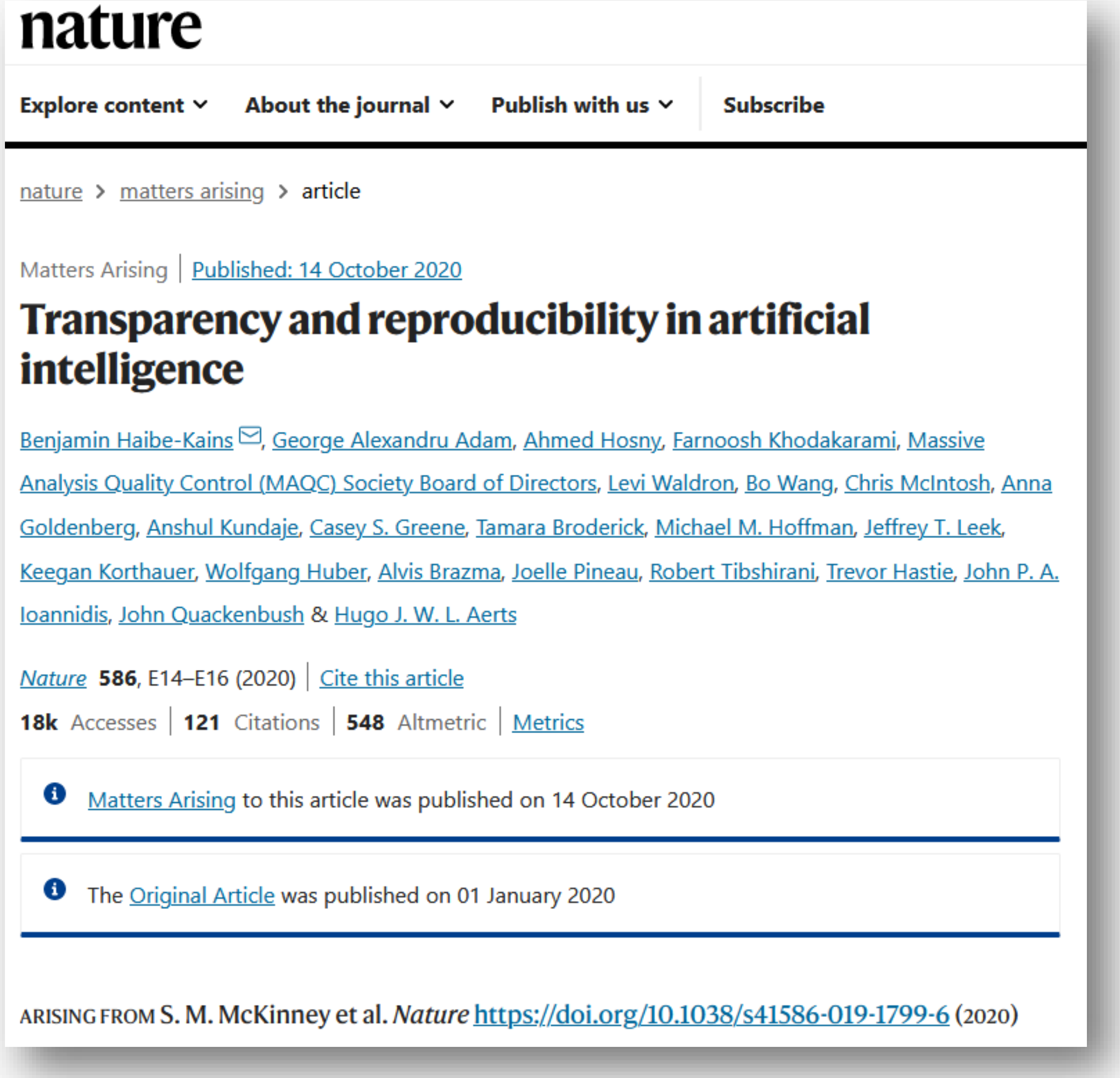
What is the problem?

- Some important choices may only be in the code (e.g. architecture, hyperparams, protocols, …)
- The method works with the authors' data but not with yours. Beyond applicability scope? Flaw?
- Hardly possible to verify results, and therefore build upon the original work

Munafò, M., Nosek, B., Bishop, D. *et al.* **A manifesto for reproducible science**. *Nat Hum Behav* **1**, 0021 (2017). https://doi.org/10.1038/s41562-016-0021

EC, Directorate-General for Research and Innovation, Baker, L., Cristea, I., Errington, T., et al., **Reproducibility of scientific results in the EU: scoping report**, Lusoli, W. (editor), *Publications Office*, 2020, https://data.europa.eu/doi/10.2777/341654

What is the problem?



Matters Arising | Published: 14 October 2020

Transparency and reproducibility in artificial intelligence

Benjamin Haibe-Kains, George Alexandru Adam, Ahmed Hosny, Farnoosh Khodakarami, Massive Analysis Quality Control (MAQC) Society Board of Directors, Levi Waldron, Bo Wang, Chris McIntosh, Anna Goldenberg, Anshul Kundaje, Casey S. Greene, Tamara Broderick, Michael M. Hoffman, Jeffrey T. Leek, Keegan Korthauer, Wolfgang Huber, Alvis Brazma, Joelle Pineau, Robert Tibshirani, Trevor Hastie, John P. A. Ioannidis, John Quackenbush & Hugo J. W. L. Aerts

Nature **586**, E14–E16 (2020) | Cite this article

18k Accesses | **121** Citations | **548** Altmetric | Metrics

Matters Arising to this article was published on 14 October 2020

The Original Article was published on 01 January 2020

ARISING FROM S. M. McKinney et al. *Nature* https://doi.org/10.1038/s41586-019-1799-6 (2020)

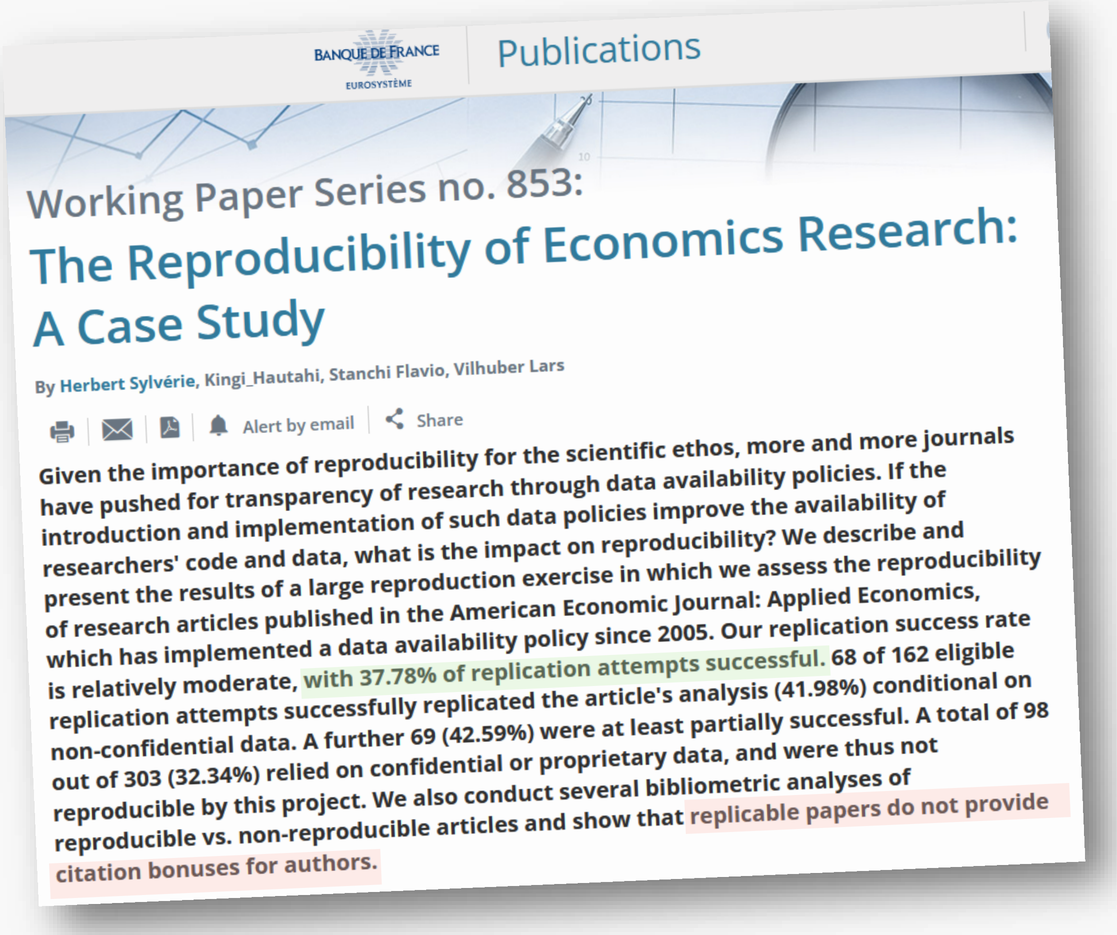
https://www.nature.com/articles/s41586-020-2766-y (2020)

"The lack of access to code and data in prominent scientific publications may lead to **unwarranted and even potentially harmful clinical trials**. (...)

Making one's methods reproducible may **surface biases or shortcomings** to authors before publication.

Preventing external validation of a model will likely **reduce its impact**, as it also prevents other researchers from using and building upon it in future studies."

What is the problem?



BANQUE DE FRANCE EUROSYSTÈME | Publications

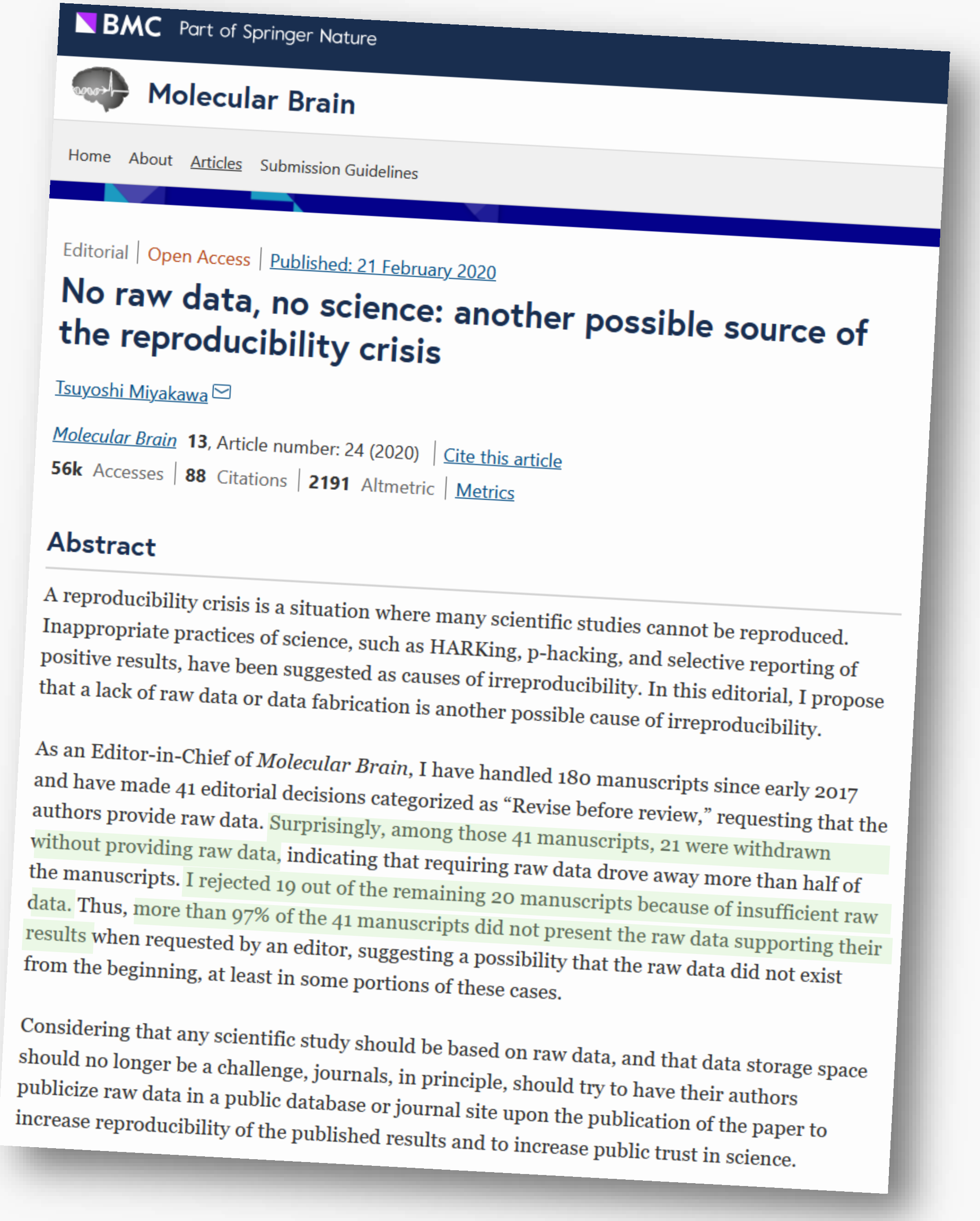
Working Paper Series no. 853:

The Reproducibility of Economics Research: A Case Study

By Herbert Sylvérie, Kingi_Hautahi, Stanchi Flavio, Vilhuber Lars

Given the importance of reproducibility for the scientific ethos, more and more journals have pushed for transparency of research through data availability policies. If the introduction and implementation of such data policies improve the availability of researchers' code and data, what is the impact on reproducibility? We describe and present the results of a large reproduction exercise in which we assess the reproducibility of research articles published in the American Economic Journal: Applied Economics, which has implemented a data availability policy since 2005. Our replication success rate is relatively moderate, with 37.78% of replication attempts successful. 68 of 162 eligible replication attempts successfully replicated the article's analysis (41.98%) conditional on non-confidential data. A further 69 (42.59%) were at least partially successful. A total of 98 out of 303 (32.34%) relied on confidential or proprietary data, and were thus not reproducible by this project. We also conduct several bibliometric analyses of reproducible vs. non-reproducible articles and show that replicable papers do not provide citation bonuses for authors.

https://publications.banque-france.fr/en/reproducibility-economics-research-case-study (2021)



BMC Part of Springer Nature

Molecular Brain

Editorial | Open Access | Published: 21 February 2020

No raw data, no science: another possible source of the reproducibility crisis

Tsuyoshi Miyakawa

Molecular Brain **13**, Article number: 24 (2020) | Cite this article

56k Accesses | **88** Citations | **2191** Altmetric | Metrics

Abstract

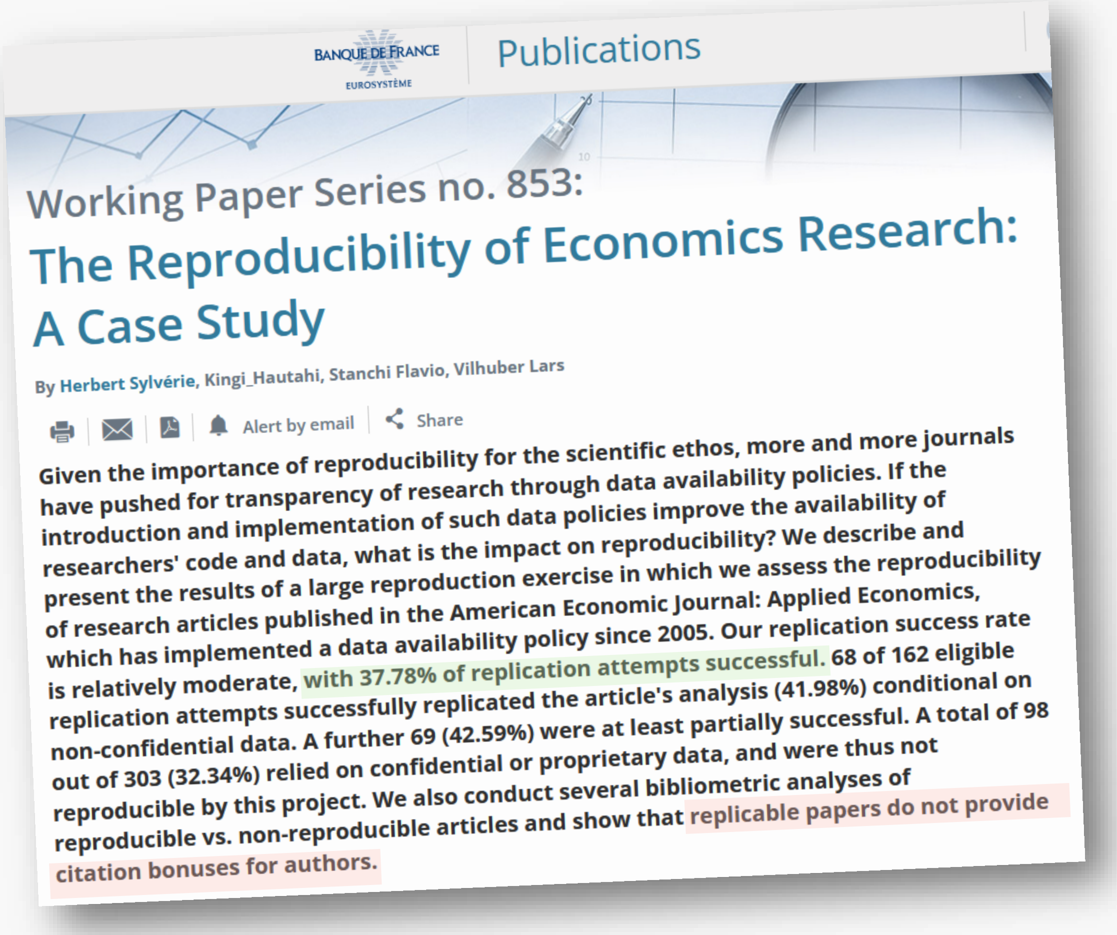
A reproducibility crisis is a situation where many scientific studies cannot be reproduced. Inappropriate practices of science, such as HARKing, p-hacking, and selective reporting of positive results, have been suggested as causes of irreproducibility. In this editorial, I propose that a lack of raw data or data fabrication is another possible cause of irreproducibility.

As an Editor-in-Chief of *Molecular Brain*, I have handled 180 manuscripts since early 2017 and have made 41 editorial decisions categorized as "Revise before review," requesting that the authors provide raw data. Surprisingly, among those 41 manuscripts, 21 were withdrawn without providing raw data, indicating that requiring raw data drove away more than half of the manuscripts. I rejected 19 out of the remaining 20 manuscripts because of insufficient raw data. Thus, more than 97% of the 41 manuscripts did not present the raw data supporting their results when requested by an editor, suggesting a possibility that the raw data did not exist from the beginning, at least in some portions of these cases.

Considering that any scientific study should be based on raw data, and that data storage space should no longer be a challenge, journals, in principle, should try to have their authors publicize raw data in a public database or journal site upon the publication of the paper to increase reproducibility of the published results and to increase public trust in science.

https://molecularbrain.biomedcentral.com/articles/10.1186/s13041-020-0552-2 (2020)

What is the problem?



BANQUE DE FRANCE
EUROSYSTÈME

Publications

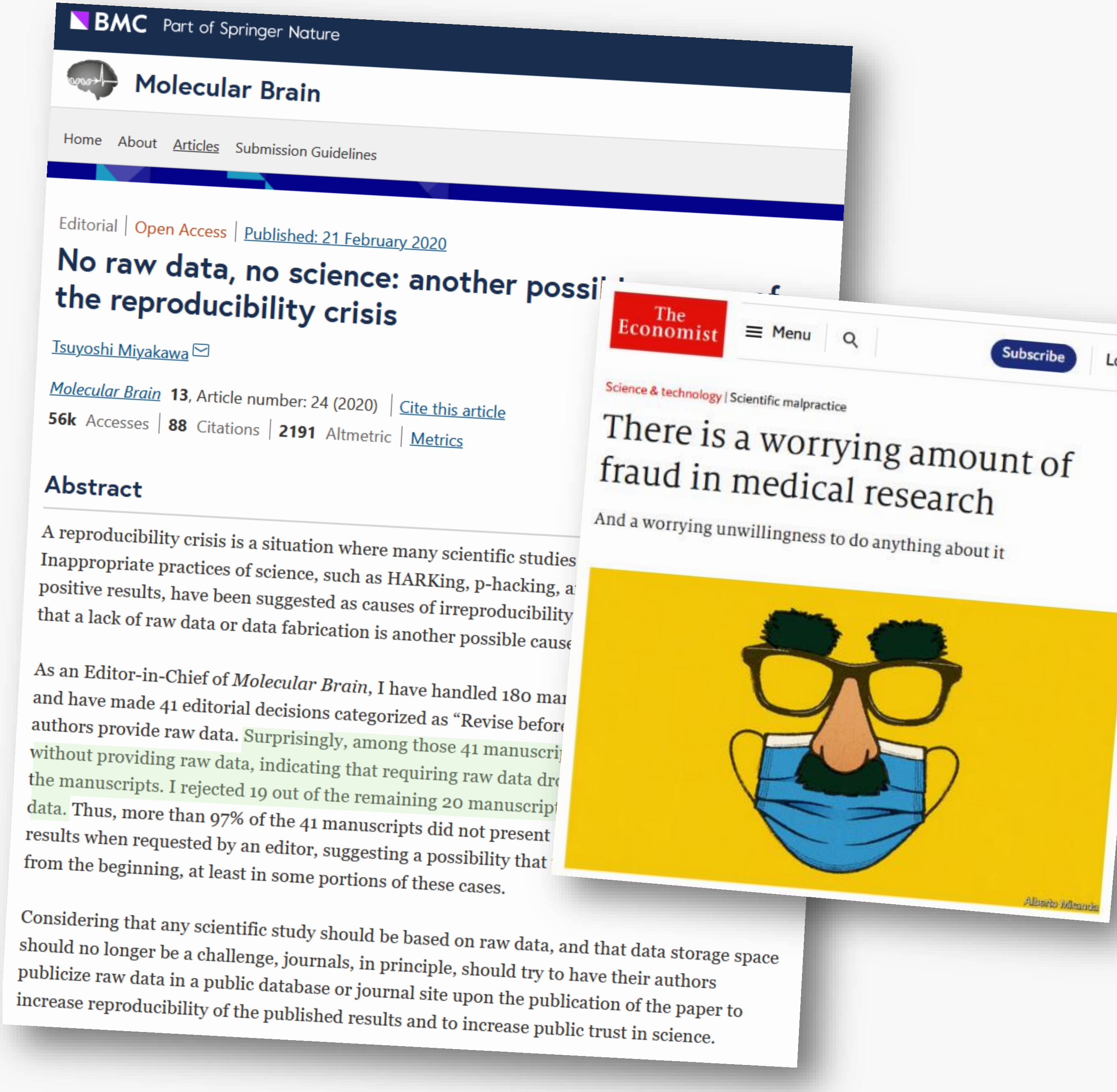
Working Paper Series no. 853:

The Reproducibility of Economics Research: A Case Study

By Herbert Sylvérie, Kingi_Hautahi, Stanchi Flavio, Vilhuber Lars

Given the importance of reproducibility for the scientific ethos, more and more journals
have pushed for transparency of research through data availability policies. If the
introduction and implementation of such data policies improve the availability of
researchers' code and data, what is the impact on reproducibility? We describe and
present the results of a large reproduction exercise in which we assess the reproducibility
of research articles published in the American Economic Journal: Applied Economics,
which has implemented a data availability policy since 2005. Our replication success rate
is relatively moderate, with 37.78% of replication attempts successful. 68 of 162 eligible
replication attempts successfully replicated the article's analysis (41.98%) conditional on
non-confidential data. A further 69 (42.59%) were at least partially successful. A total of 98
out of 303 (32.34%) relied on confidential or proprietary data, and were thus not
reproducible by this project. We also conduct several bibliometric analyses of
reproducible vs. non-reproducible articles and show that replicable papers do not provide
citation bonuses for authors.

https://publications.banque-france.fr/en/reproducibility-economics-research-case-study (2021)



BMC Part of Springer Nature

Molecular Brain

Editorial | Open Access | Published: 21 February 2020

**No raw data, no science: another possi
the reproducibility crisis**

Tsuyoshi Miyakawa

Molecular Brain 13, Article number: 24 (2020) | Cite this article

56k Accesses | 88 Citations | 2191 Altmetric | Metrics

Abstract

A reproducibility crisis is a situation where many scientific studies
Inappropriate practices of science, such as HARKing, p-hacking, a
positive results, have been suggested as causes of irreproducibility
that a lack of raw data or data fabrication is another possible cause

As an Editor-in-Chief of *Molecular Brain*, I have handled 180 ma
and have made 41 editorial decisions categorized as "Revise befor
authors provide raw data. Surprisingly, among those 41 manuscri
without providing raw data, indicating that requiring raw data dr
the manuscripts. I rejected 19 out of the remaining 20 manuscrip
data. Thus, more than 97% of the 41 manuscripts did not present
results when requested by an editor, suggesting a possibility that
from the beginning, at least in some portions of these cases.

Considering that any scientific study should be based on raw data, and that data storage space should no longer be a challenge, journals, in principle, should try to have their authors publicize raw data in a public database or journal site upon the publication of the paper to increase reproducibility of the published results and to increase public trust in science.

The Economist

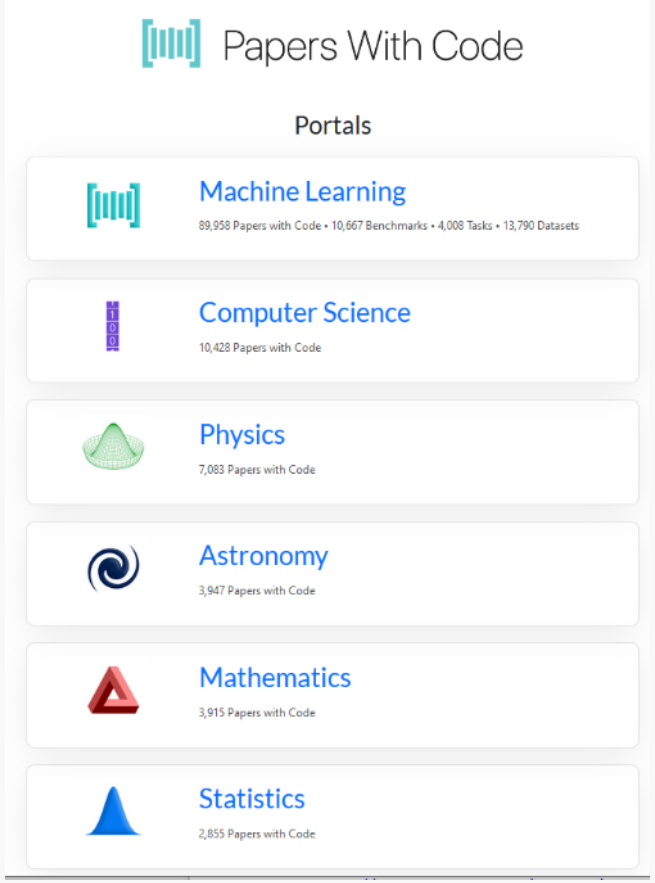
Science & technology | Scientific malpractice

There is a worrying amount of fraud in medical research

And a worrying unwillingness to do anything about it

https://molecularbrain.biomedcentral.com/articles/10.1186/s13041-020-0552-2 (2020)

What is the problem?



Papers With Code

Portals

Machine Learning
89,958 Papers with Code • 10,667 Benchmarks • 4,008 Tasks • 13,790 Datasets

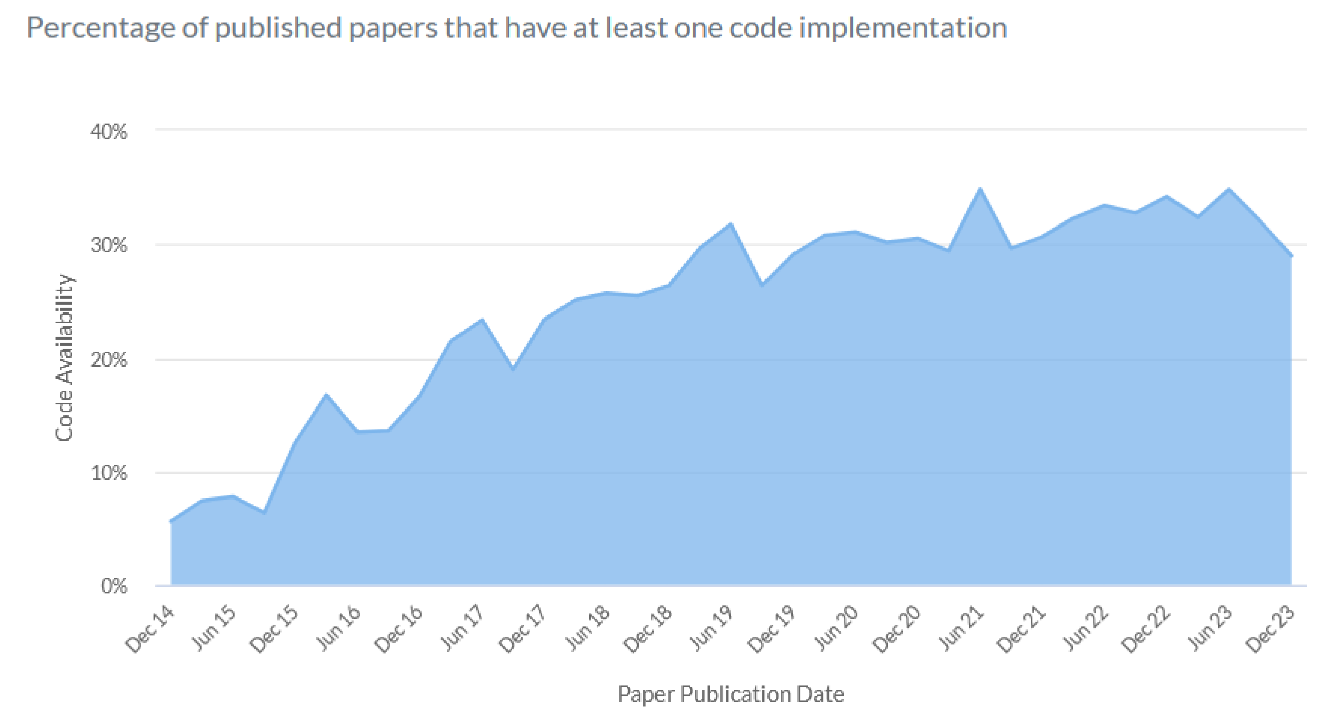
Computer Science
10,428 Papers with Code

Physics
7,083 Papers with Code

Astronomy
3,947 Papers with Code

Mathematics
3,915 Papers with Code

Statistics
2,855 Papers with Code



“Code Availability: For every open access ML paper, we check whether a code implementation is available on GitHub. (…) We include both official and community implementations.”

https://paperswithcode.com/trends

Why publish my [article+code+data] and link them?

- My research is funded by **public grants**, code/data have to be public too
- It **convinces** reviewers that the results in my paper can be trusted
- People can **use** my algorithm and **compare** it with others without having to re-implement it
- Tends to increase the number of **citations** of my paper*
- It improves **quality**: forces me to organize, document, clean my code & data

*Colavizza C., Hrynaszkiewicz I., Staden I., Whitaker K., McGillivray B. The citation advantage of linking publications to research data. *PLOS ONE*, 2020, https://doi.org/10.1371/journal.pone.0230416

> *"In 2018, 94% of 21,793 PLOS articles and 88% of 31,956 BMC had data availability statements. In 2017 and 2018, 21% of PLOS and 12% of BMC publications provided **data availability statements containing a link to data** in a repository. (…) association between articles that include statements that link to data in a repository and **up to 25.36% higher citation**"*

Reproducibility **incentives**?

Availability of code+data not yet a requirement from journals and conferences, but visible uptake.

Photo: https://www.pexels.com/photo/side-view-of-giving-treats-to-a-dog-8473725/

Uptake of reproducibility requirement?

- Reproducibility checklist for reviewers

  IJCAI, AAAI, NAACL, MICCAI

- Conference Reproducibility Track

  ISWC, ECIR, SIGIR, ACM Multimedia

- Reproducibility of Results in the ACM Digital Library

  Prototypes of *active digital curation platforms*, close to “executable paper”

  https://www.acm.org/publications/reproducibility

- Replicability Stamp for papers published in some journals

  https://www.acm.org/publications/policies/artifact-review-and-badging-current

acm
acm
acm
acm
acm

available, functional, reusable, reproduced, replicated

http://www.replicabilitystamp.org/

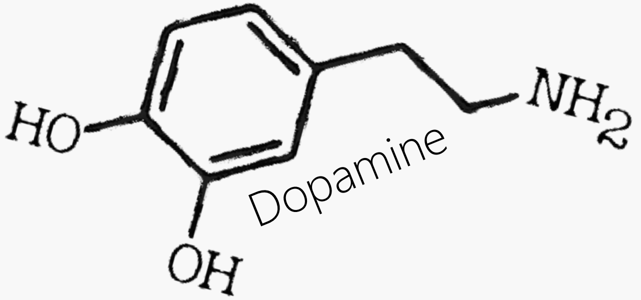
(graphics community effort)

Data availability incentives?

Some journals implement a data availability policy ranging from encouragement to requirement:

- Springer data policies: 4 policies to apply to different journals https://www.springernature.com/gp/authors/research-data-policy/research-data-policy-types
  - Springer Scientific Data - *Data sharing, evidence of data sharing and peer review of data* ***required***
  - Springer Humanities and Social Science Communications: *Data sharing* ***encouraged*** *and statements of data availability required*
- Cell Discovery, Nature
  - *Data sharing encouraged and statements of data availability required*
- International Economic Review https://economics.sas.upenn.edu/ier/submissions/data-availability-policy
- IOP Publishing data availability policy https://publishingsupport.iopscience.iop.org/iop-publishing-data-availability-policy/
- Set of Data Availability policies: https://www.lib.uiowa.edu/data/cite-data-and-code/#availability

How do I get rewarded?



Data: Data Paper (data journal)

Code: Software Paper (research software journal)

Data or code: Resource Track Paper (conference)

Advantage: it's a regular citable paper

Fits in existing citation fw, accounted for in common citation indicators

Examples of Data Journals:

- Biomedical Data Journal
- Biodiversity Data Journal
- Elvesier Data in Brief
- Nature Scientific Data

Examples of Software Journals:

- Journal of Open Source Software
- Open Research Software
- ScienceDirect SoftwareX
- BMC Neuroscience

But we want to cite data/code, not (*only*) data/software papers!

Agenda

- Overview of Open Science
- Reproducible research
  - The reproducibility crisis
  - Vocabulary
  - Incentives and rewards
- Make code and data findable, accessible, referenceable & citable
  - Importance of Persistent Identifiers (PID)
  - Citation guidelines
  - Public repositories + focus on Software Heritage
- Giving credit: citing article, code & data alike

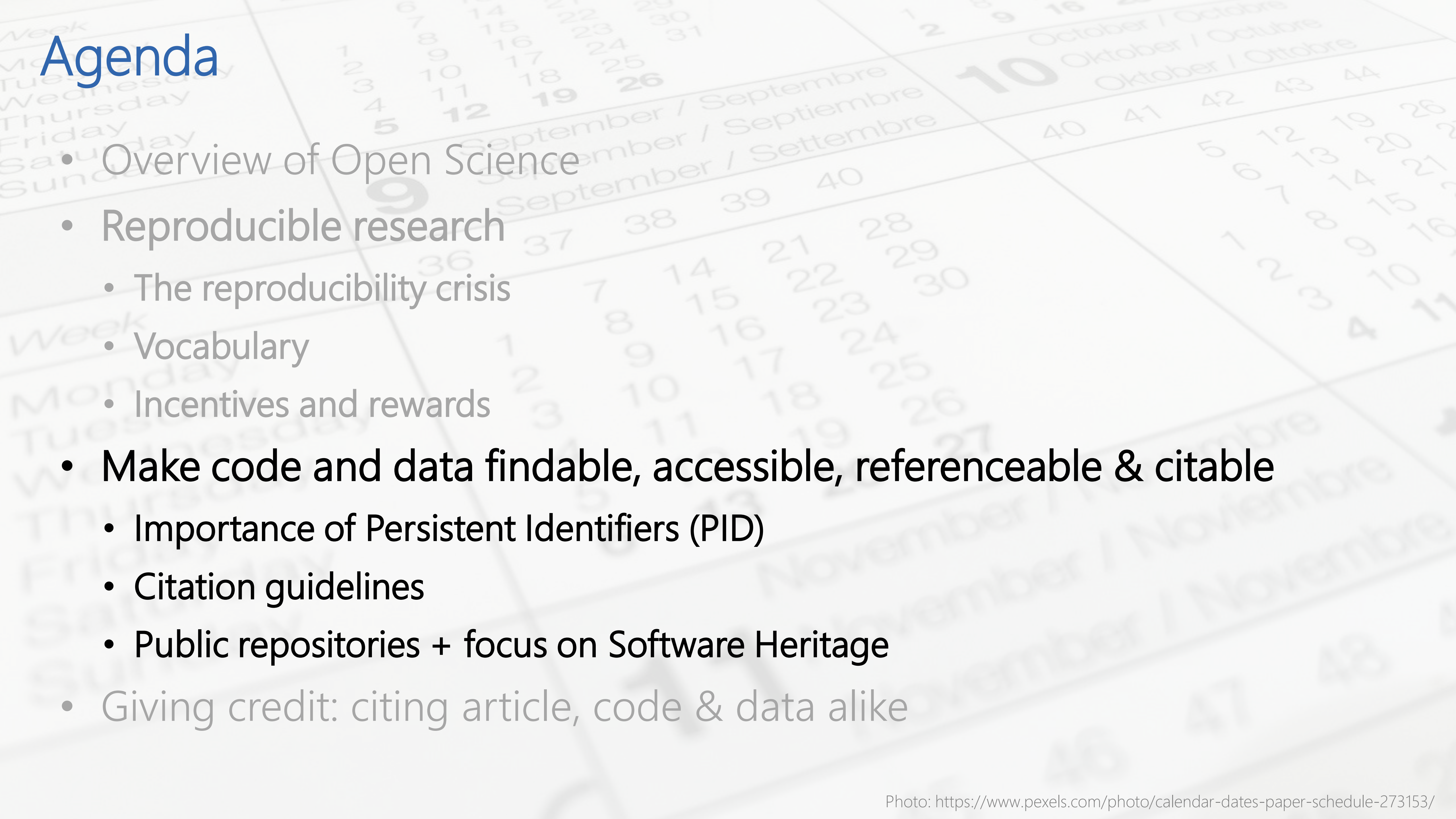


Photo: https://www.pexels.com/photo/calendar-dates-paper-schedule-273153/

Reproducibility requires to make code and data FARC*

* Findable, Accessible, Referenceable, Citable

FARC: Findable Accessible Referenceable Citable

- Findable: publish rich metadata
  - Title, Authors/publishers
  - Dates (first publication, release),
  - Version
  - License
  - Provenance
  - **Persistent identifier (PID)**
  - …
- Accessible: published on sustainable public repository
- Citable: give credit, attribution
- Referenceable: exact version of code/data for reproducibility

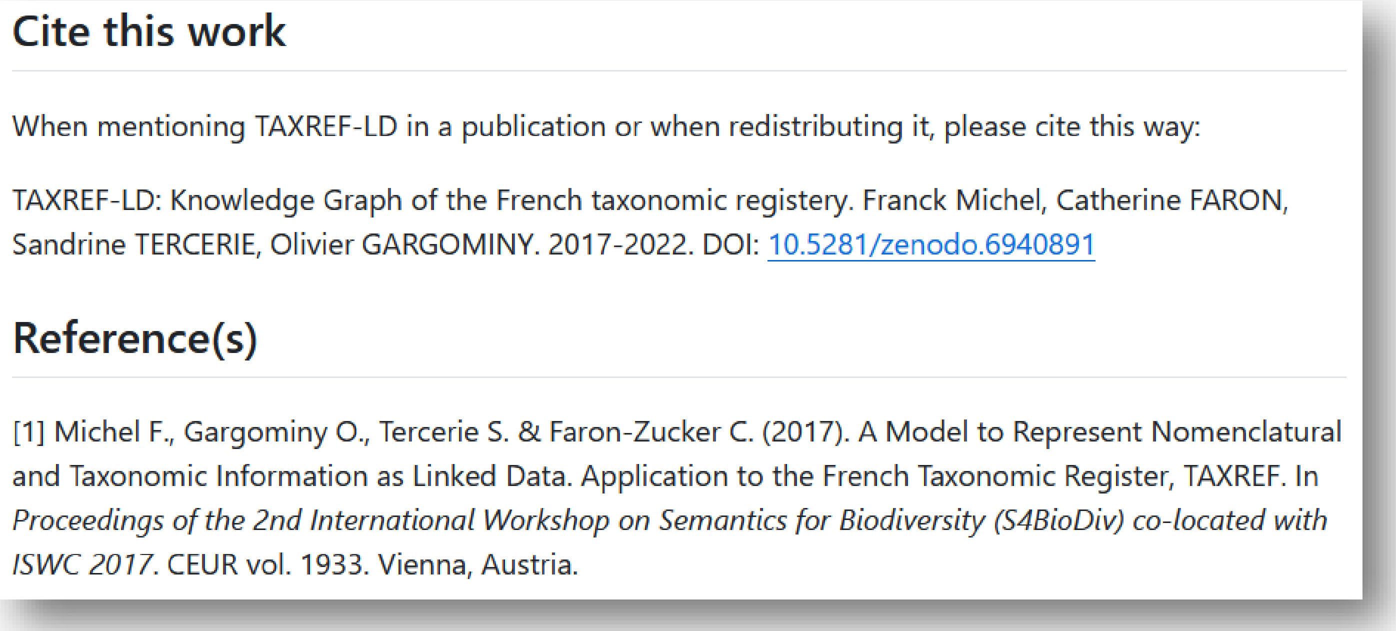
PID "is all you need!"

- Name a resource unambiguously
- Associate metadata to a resource
- Cite a resource
- Reference a version of a resource

Add *(human-readable)* **citation guidelines** for your code and data

README(.*) of a repository, or citation field in the metadata etc.

https://github.com/frmichel/taxref-ld/



Cite this work

When mentioning TAXREF-LD in a publication or when redistributing it, please cite this way:

TAXREF-LD: Knowledge Graph of the French taxonomic registery. Franck Michel, Catherine FARON, Sandrine TERCERIE, Olivier GARGOMINY. 2017-2022. DOI: 10.5281/zenodo.6940891

Reference(s)

[1] Michel F., Gargominy O., Tercerie S. & Faron-Zucker C. (2017). A Model to Represent Nomenclatural and Taxonomic Information as Linked Data. Application to the French Taxonomic Register, TAXREF. In *Proceedings of the 2nd International Workshop on Semantics for Biodiversity (S4BioDiv) co-located with ISWC 2017*. CEUR vol. 1933. Vienna, Austria.

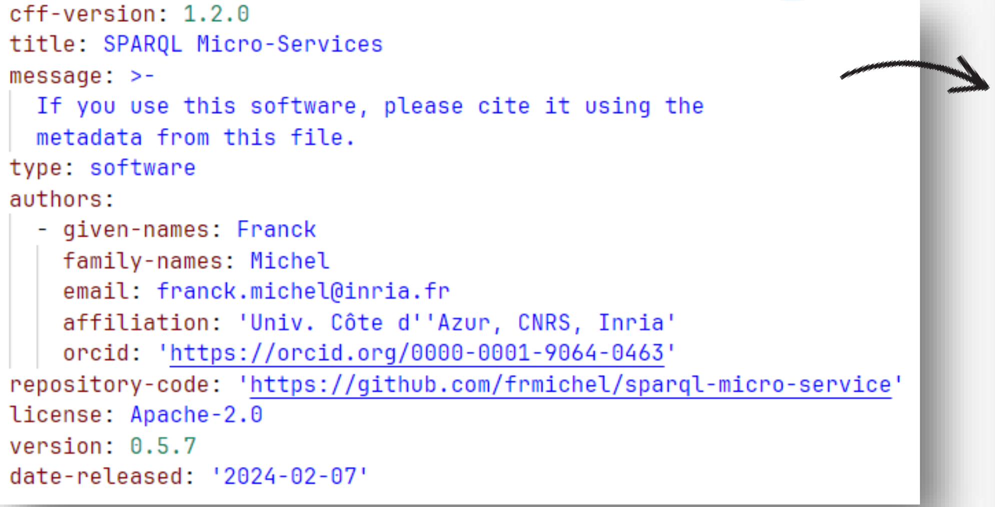
Drawback: human-readable only

Add *(machine-readable)* **citation guidelines** for your code and data

Citation File Format (CFF)

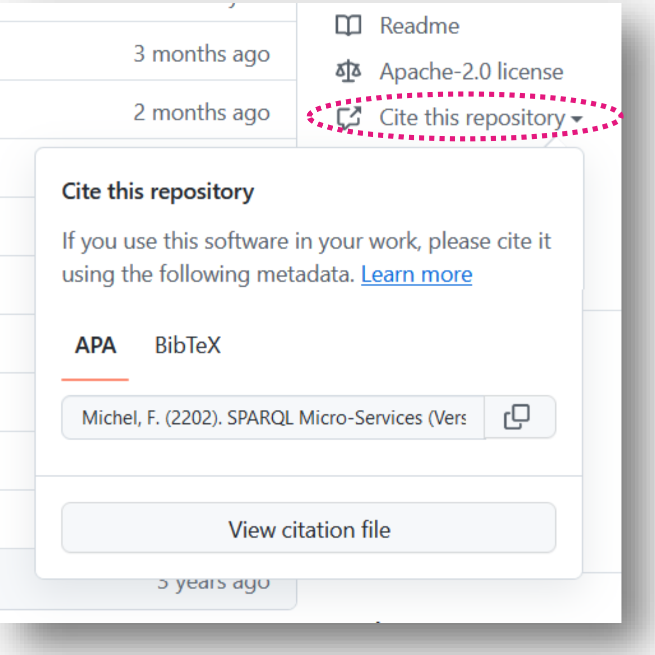
CITATION.cff: plain text file with **human- and machine-readable citation information** for software and datasets to **let others know how to correctly cite them**.
https://citation-file-format.github.io/

https://github.com/frmichel/sparql-micro-service/



```
cff-version: 1.2.0
title: SPARQL Micro-Services
message: >-
  If you use this software, please cite it using the
  metadata from this file.
type: software
authors:
  - given-names: Franck
    family-names: Michel
    email: franck.michel@inria.fr
    affiliation: 'Univ. Côte d''Azur, CNRS, Inria'
    orcid: 'https://orcid.org/0000-0001-9064-0463'
repository-code: 'https://github.com/frmichel/sparql-micro-service'
license: Apache-2.0
version: 0.5.7
date-released: '2024-02-07'
```

APA and Bibtex @*software* citation



Cite this repository

If you use this software in your work, please cite it using the following metadata. Learn more

Michel, F. (2202). SPARQL Micro-Services (Vers

Generation form: https://citation-file-format.github.io/cff-initializer-javascript/#/start

Supported by Github, Zenodo, Zotero's browser plugin.

Make DATA findable, accessible, referenceable

Where do I publish my data?

French p/f:

- https://www.data.gouv.fr/
- https://recherche.data.gouv.fr/ (PNSO2)
- Spaces and collections per institution, laboratory…
  - https://entrepot.recherche.data.gouv.fr/dataverse/[univ-cotedazur|inria|cnrs|I3S]

"To be used when there is no well-adopted domain or community specific repository."

Research or general-purpose repositories, e.g.:

- **Zenodo** (doi): by OpenAire (European infrastructure that supports Open Sc.)
- **Figshare** (doi): hosted by Digital Science, a subsidiary of Springer Nature
- **Internet Archive** (ark)

Make DATA findable, accessible, referenceable

Where do I publish my data? (*cont.*)

- Institutional or community data repositories
- Find a repo on **OpenDOAR**: Directory of Open Access Repositories
  - Faceted search by country, domain, type of research object… (https://v2.sherpa.ac.uk/opendoar/)

Fenner, M., Crosas, M., Grethe, J.S. *et al.* **A data citation roadmap for scholarly data repositories**. *Sci Data* **6**, 28 (2019). https://doi.org/10.1038/s41597-019-0031-8

Force11 Data Citation. **Data Citations: A Primer**. (2016). Retrieved December 22, 2016. From http://force11.github.io/data-citation-primer/

Task Group on Data Citation Standards and Practices, C.-I., 2013. **Out of Cite, Out of Mind: The Current State of Practice, Policy, and Technology for the Citation of Data**. *Data Science Journal*, 12, pp.CIDCR1–CIDCR7. http://doi.org/10.2481/dsj.OSOM13-043

Make CODE findable, accessible, referenceable, citable

Where do I publish my code?

- Public/institutional CVS: Github, Gitlab, Bitbucket, Redmine.
  - No guarantee of sustainability/long term preservation.
  - e.g. Google Code 2006-2016 R.I.P.
- Research or general-purpose repositories: Zenodo, Figshare, Internet Archive
- Software Heritage
- HAL

Smith AM, Katz DS, Niemeyer KE. **FORCE11 Software Citation Working Group**. 2016. **Software citation principles**. *PeerJ Computer Science* 2:e86. DOI: 10.7717/peerj-cs.86

P. Alliez, R. Di Cosmo, B. Guedj, A. Girault, MS. Hacid, et al.. **Attributing and Referencing (Research) Software: Best Practices and Outlook from Inria**. 2019. ⟨hal-02135891v1⟩

Automatic deposit of code from Github

- Github to Zenodo

  Automated on release action (snapshot), gets a **DOI**

  https://docs.github.com/en/repositories/archiving-a-github-repository/referencing-and-citing-content

- Github to Figshare

  Need to add a GitHub workflow action in the repository

  Updates on every change, new version on release action, gets a **DOI**

  20GB limit.

  https://help.figshare.com/article/how-to-connect-figshare-with-your-github-account

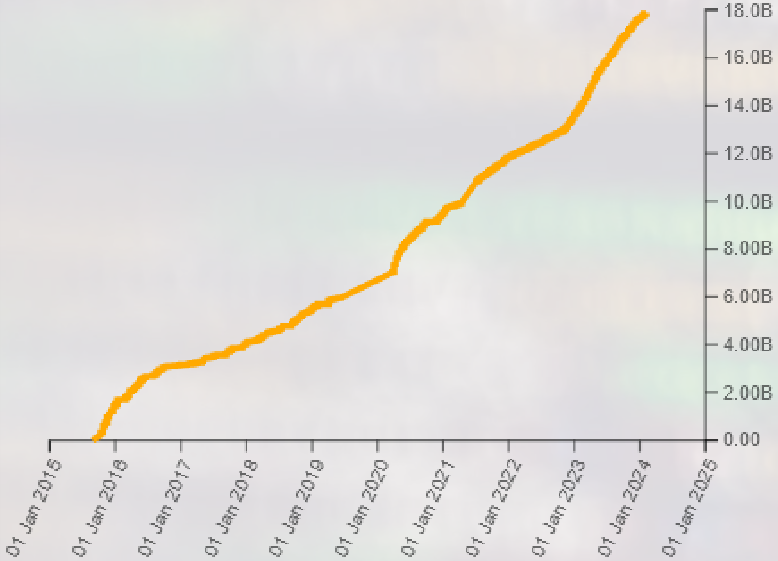
Software [is our] Heritage

"collect, preserve, share
all software publicly available
with full dev. history, in source code form"

Software Heritage

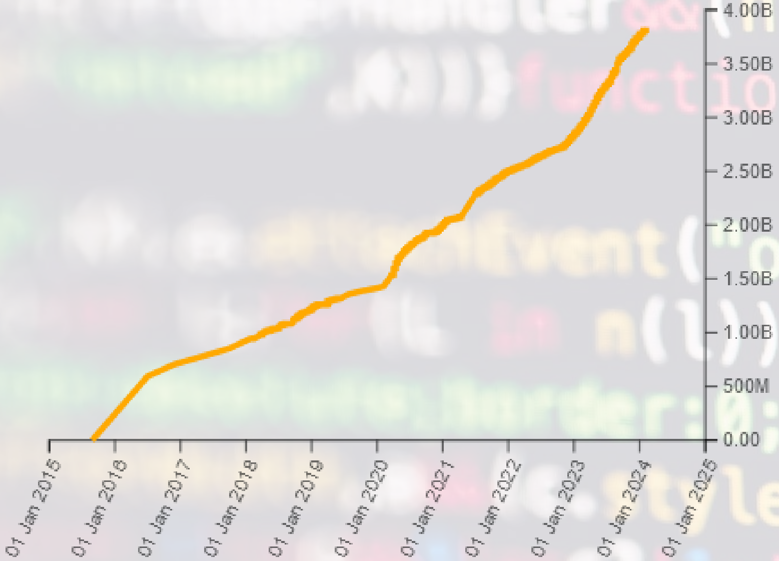
Source files

18 470 899 783



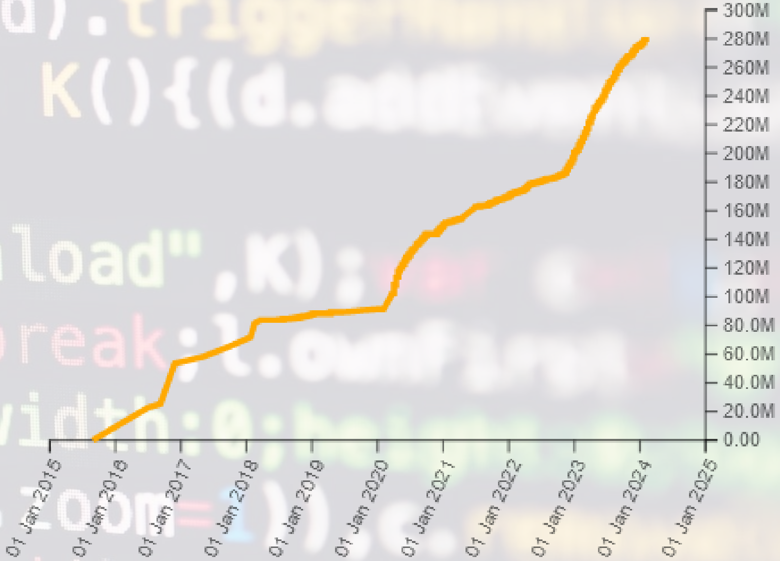
Commits

3 967 081 515



Projects

289 085 734



(2024-03-28)

https://archive.softwareheritage.org/

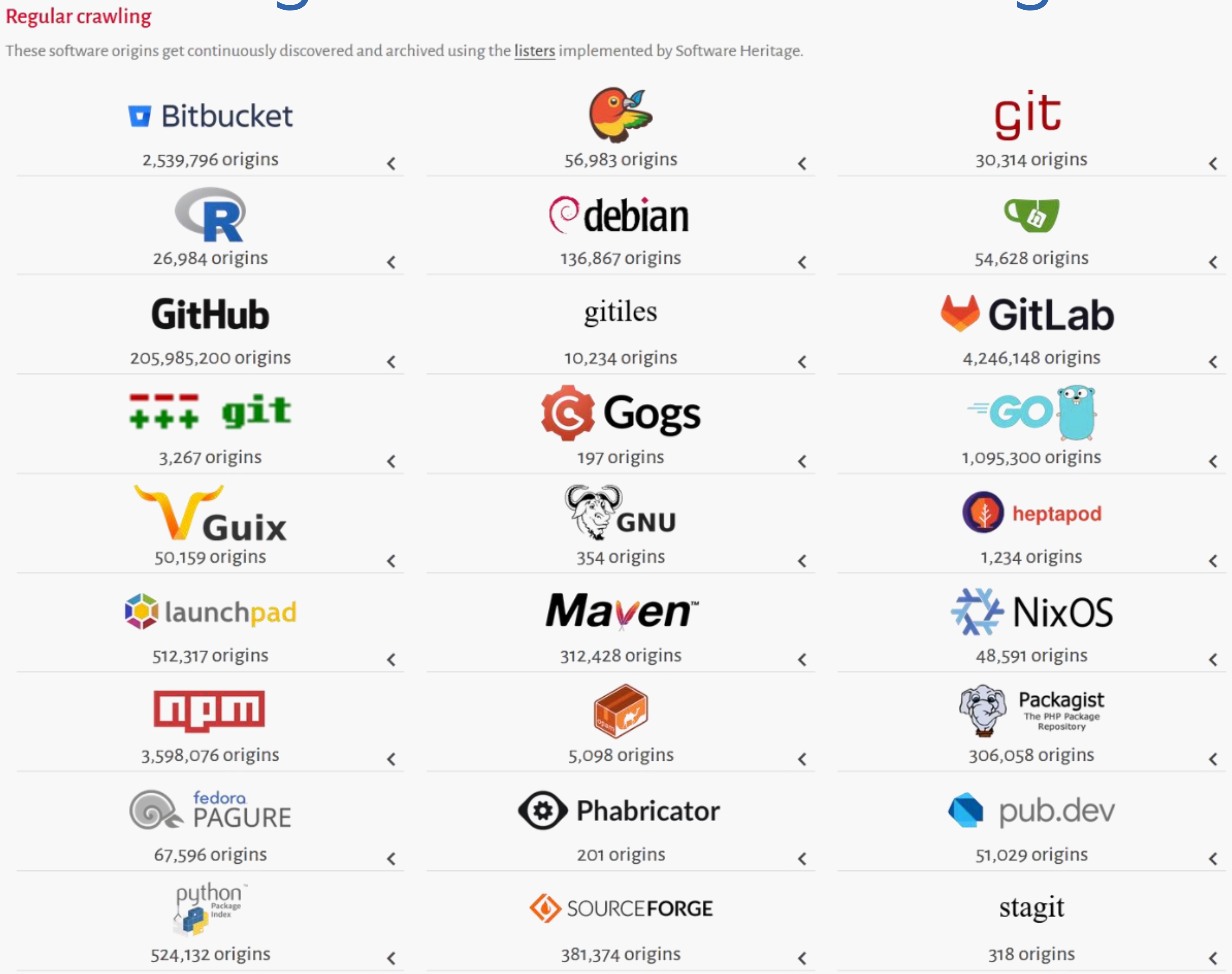
Software Heritage persistent Identifiers: SWHIDs

- Intrinsic identifiers: no need for external register
- Do not depend on external resolvers that can be compromised/discontinued
- Resolvable: prepend with https://archive.softwareheritage.org/
- Reference a precise point in the sw dev history, independently of releases
- Flexible granularity of referenced content, from project down to code lines

R. D. Cosmo, M. Gruenpeter and S. Zacchiroli. **Referencing Source Code Artifacts: A Separate Concern in Software Citation**. *Computing in Science & Engineering*, vol. 22, no. 2, pp. 33-43, 2020, doi: 10.1109/MCSE.2019.2963148.

Research Data Alliance/FORCE11 Software Source Code Identification WG, Allen, A., Bandrowski, A., Chan, P., di Cosmo, R., Fenner, M., Garcia, L., Gruenpeter, M., Jones, C. M., Katz, D. S., Kunze, J., Schubotz, M., & Todorov, I. T. **Software Source Code Identification Use cases and identifier schemes for persistent software source code identification** (1.1). 2020. DOI: 10.15497/RDA00053

Software Heritage automatic harvesting



Regular crawling

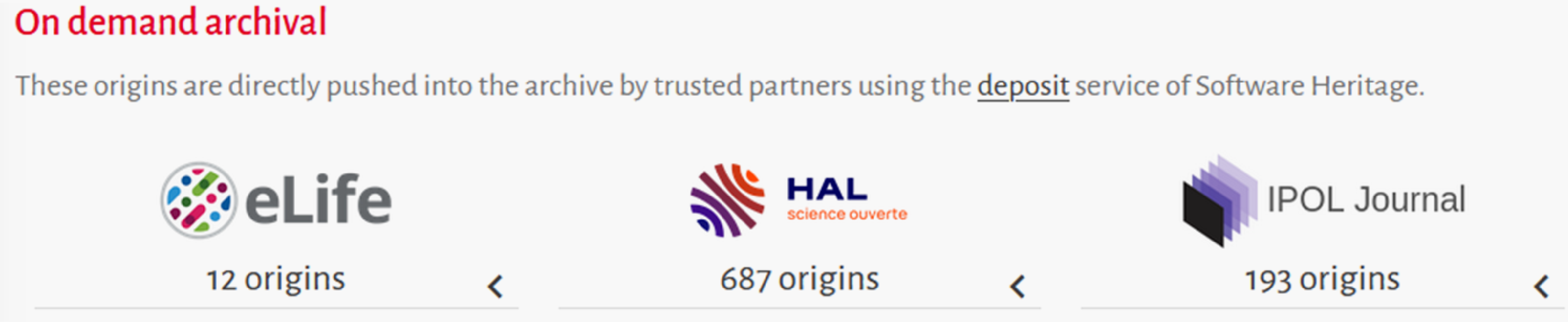
These software origins get continuously discovered and archived using the listers implemented by Software Heritage.

Bitbucket 2,539,796 origins	56,983 origins	git 30,314 origins
R 26,984 origins	debian 136,867 origins	54,628 origins
GitHub 205,985,200 origins	gitiles 10,234 origins	GitLab 4,246,148 origins
git 3,267 origins	Gogs 197 origins	GO 1,095,300 origins
Guix 50,159 origins	GNU 354 origins	heptapod 1,234 origins
launchpad 512,317 origins	Maven 312,428 origins	NixOS 48,591 origins
npm 3,598,076 origins	5,098 origins	Packagist The PHP Package Repository 306,058 origins
fedora PAGURE 67,596 origins	Phabricator 201 origins	pub.dev 51,029 origins
python Package Index 524,132 origins	SOURCEFORGE 381,374 origins	stagit 318 origins

Software Heritage on-demand harvesting

Github: https://github.com/marketplace/actions/save-to-software-heritage

A Gitlab/Github/Bitbucket endpoint using the API:
https://archive.softwareheritage.org/api/1/origin/save/doc/

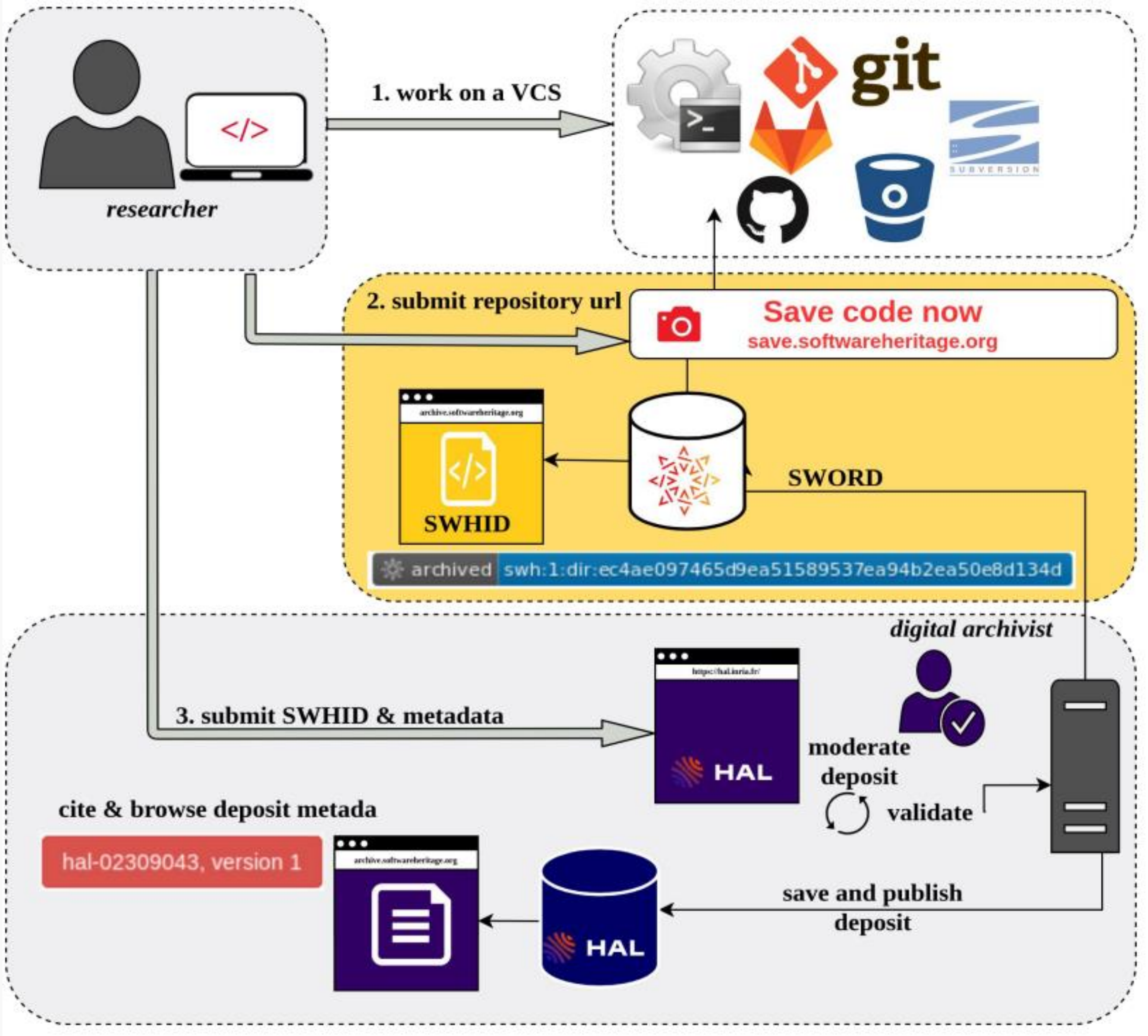


On demand archival

These origins are directly pushed into the archive by trusted partners using the deposit service of Software Heritage.

HAL https://www.softwareheritage.org/2018/09/28/depositing-scientific-software-into-software-heritage/

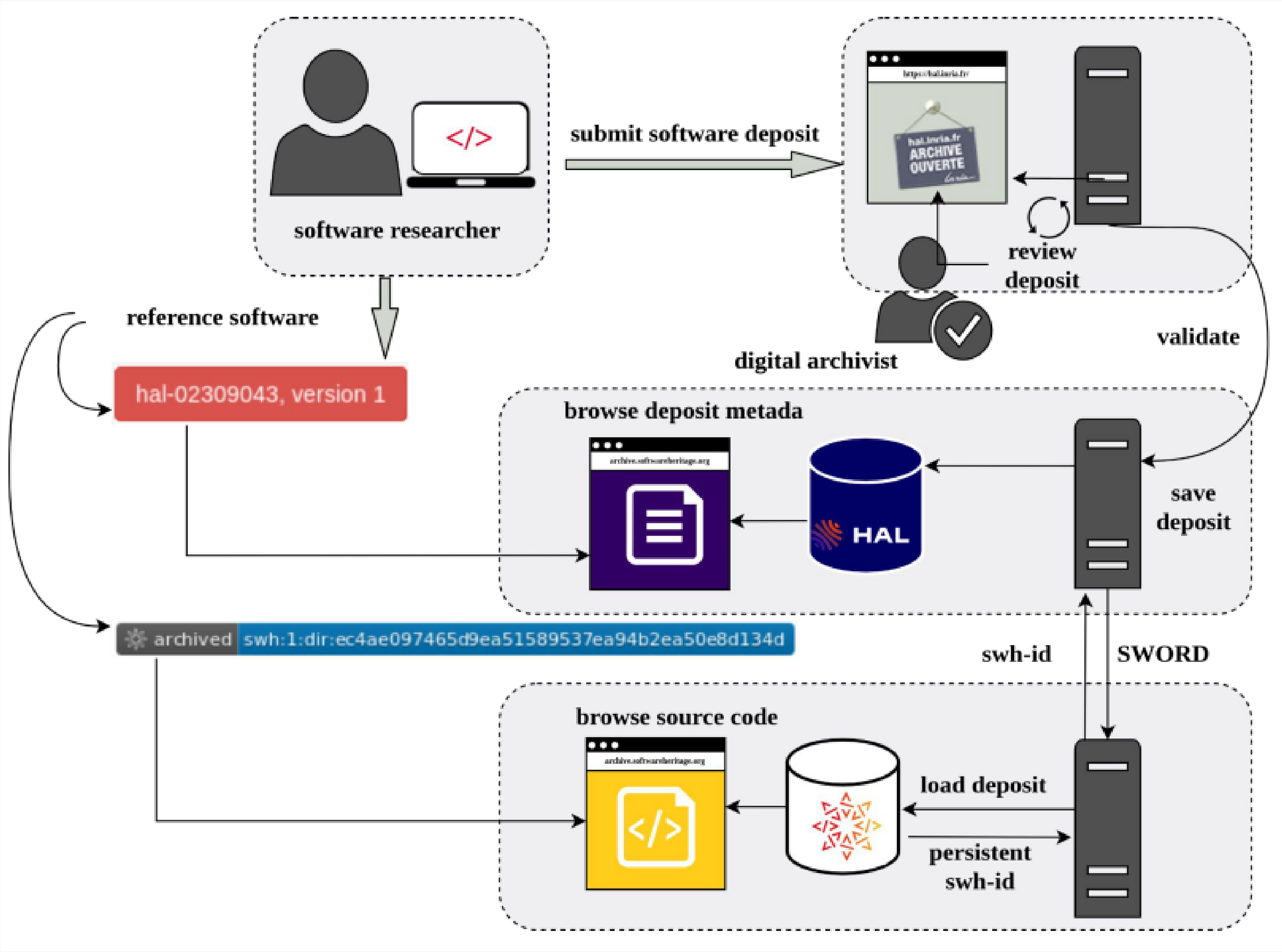
CVS -> Software Heritage -> HAL



Source:

Morane Gruenpeter, Jozefina Sadowska, Estelle Nivault, Alain Monteil. **Create software deposit in HAL: User guide and best practices**. [Technical Report] Inria; CCSD; Software Heritage. 2022. ⟨hal-01872189v2⟩

HAL -> Software Heritage



Source:

Morane Gruenpeter, Jozefina Sadowska, Estelle Nivault, Alain Monteil. **Create software deposit in HAL: User guide and best practices**. [Technical Report] Inria; CCSD; Software Heritage. 2022. ⟨hal-01872189v2⟩

Archive and reference with Software Heritage

1. Prepare the repository

   README

   AUTHORS

   LICENSE

   codemeta.json

Archive and reference with Software Heritage

1. Prepare the repository

README, README.md, README.txt

-> see HAL documentation (https://hal.inria.fr/hal-01872189v2)

Archive and reference with Software Heritage

1. Prepare the repository

README

AUTHORS (.md, .rst), CONTRIBUTORS, CREDITS, CITATION, CITATION.cff

Recommended to use Inria's taxonomy of contributors (design, architecture, coding, testing…) but other options e.g. Contributor Role Ontology.

Pierre Alliez, Roberto Di Cosmo, Benjamin Guedj, Alain Girault, Mohand-Said Hacid, et al.. **Attributing and Referencing (Research) Software: est Practices and Outlook from Inria**. 2019. ⟨hal-02135891v1⟩

No recommended file format but common practices:

```
John Smith
John Smith <john.smith@domain.org>
John Smith <john.smith@domain.org> (https://homepage.me/johnsmith)
John Smith - author and maintainer
    <john.smith@domain.org>
    https://homepage.me/johnsmith
```

Archive and reference with Software Heritage

1. Prepare the repository

   README

   AUTHORS

   LICENSE

   Most cases: use the SPDX license list (https://spdx.org/licenses/)
   SPDX: open standard for communicating license and copyright information used in free/open sw, data, hw or documentation => Find a license, copy its content to your LICENSE file

   Mix of licenses, or material with different licenses: use the REUSE guidelines (https://reuse.software/). More cumbersome.

Archive and reference with Software Heritage

1. Prepare the repository

README

AUTHORS

LICENSE

codemeta.json

CodeMeta initiative (https://codemeta.github.io/): std for sharing metadata about sw across repositories.
Addresses citation (authors), reproducibility (dependencies/env), discovery (keywords/description)

Jones, M. B., Boettiger, C., Mayes, A. C., Arfon Smith, Slaughter, P., Niemeyer, K., Gil, Y., Fenner, M., Nowak, K., Hahnel, M., Coy, L., Allen, A., Crosas, M., Sands, A., Hong, N. C., Cruse, P., Katz, D., & Goble, C. (2017). **CodeMeta: an exchange schema for software metadata**. KNB Data Repository. DOI: 10.5063/schema/codemeta-2.0

Supported by Github, DataCite, Figshare, Zenodo, NSF.

Format: JSON-LD, mostly relies on schema.org + few extensions.

Issue: does not address contributors roles, only author and contributor.

CodeMeta generator (https://codemeta.github.io/codemeta-generator/)+ various tools
Example: https://github.com/frmichel/sparql-micro-service/blob/master/codemeta.json

Archive and reference with Software Heritage

1. Prepare the repository

   README: for humans

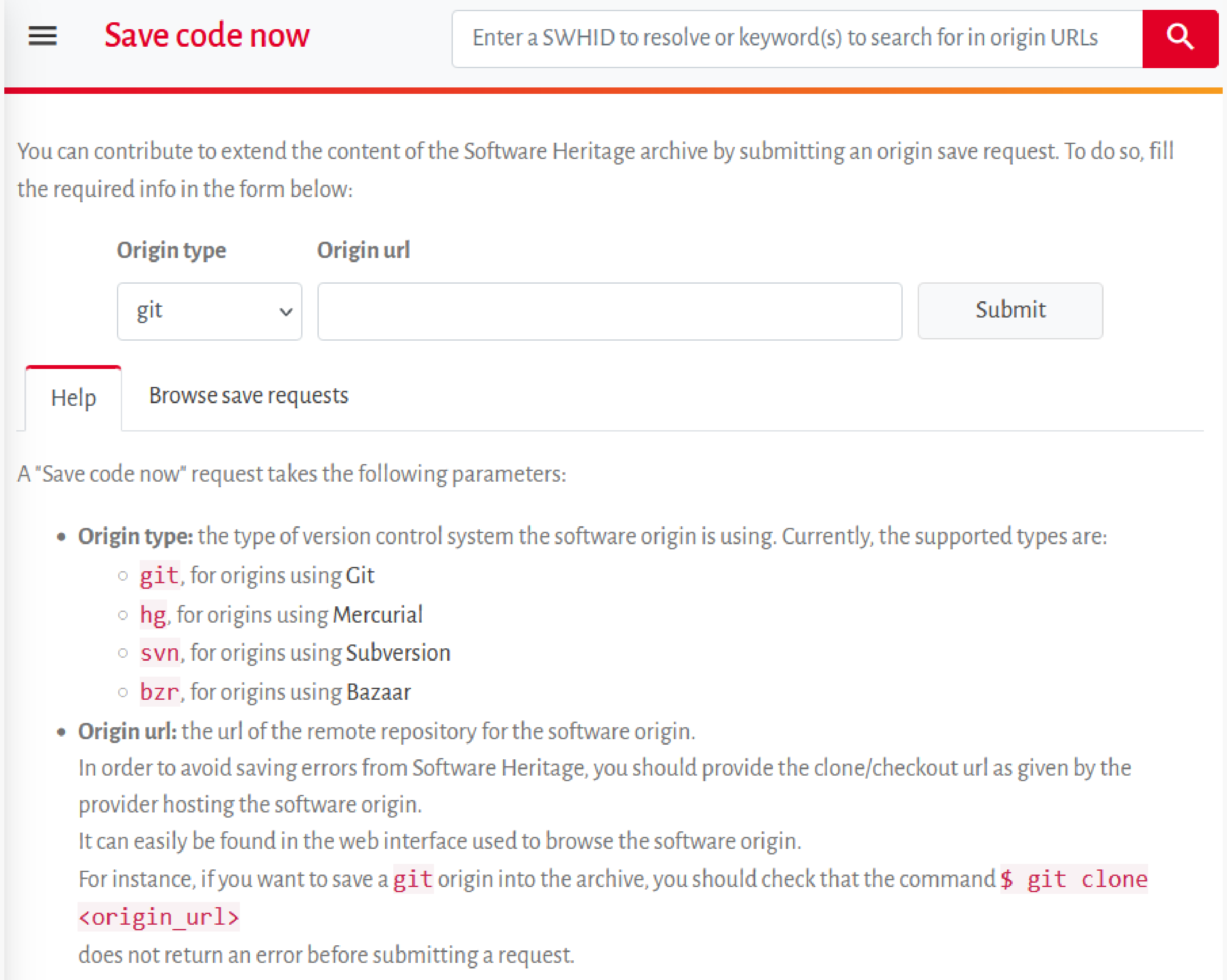
   AUTHORS for humans, DATACITE.cff for machines

   LICENSE: for humans (& machines)

   codemeta.json: for machines

Archive and reference with Software Heritage

1. Prepare the repository
2. Save the code
   https://archive.softwareheritage.org/save/



☰
Save code now
Enter a SWHID to resolve or keyword(s) to search for in origin URLs
🔍

You can contribute to extend the content of the Software Heritage archive by submitting an origin save request. To do so, fill the required info in the form below:

Origin type	Origin url	
git		Submit

Help
Browse save requests

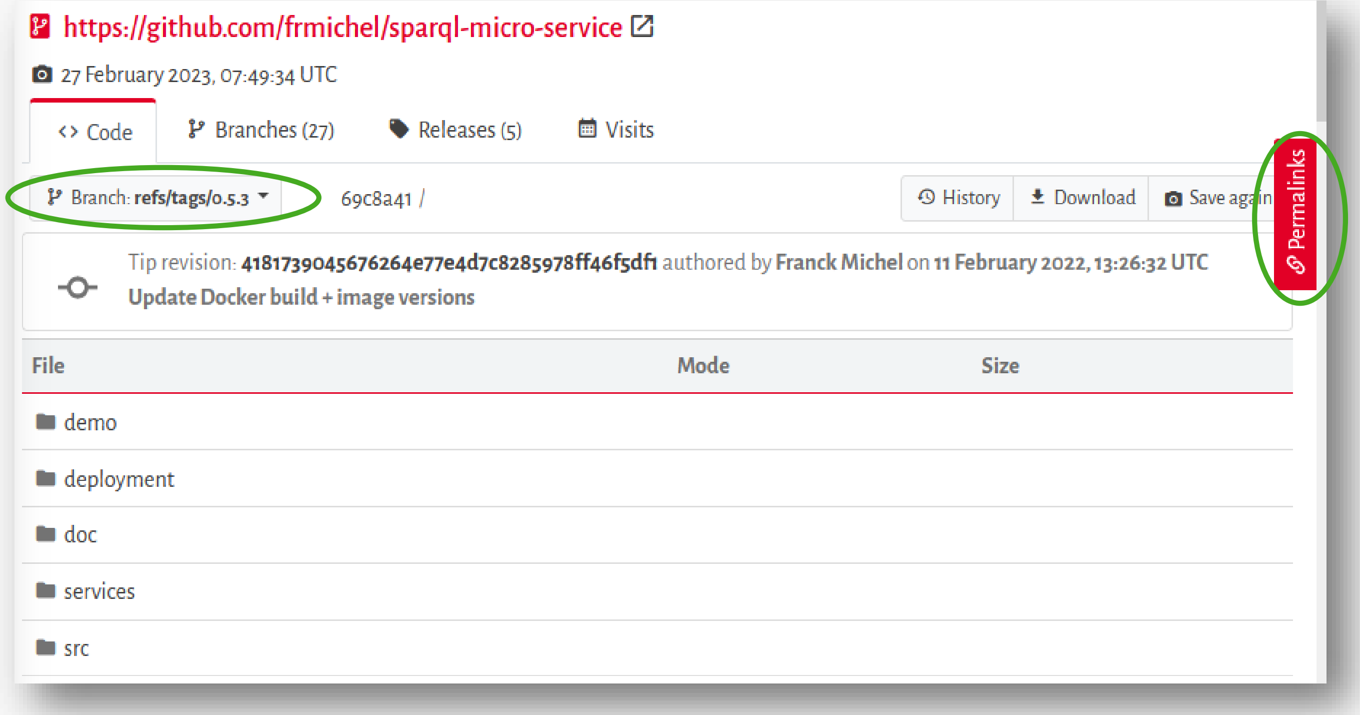
A "Save code now" request takes the following parameters:

- **Origin type:** the type of version control system the software origin is using. Currently, the supported types are:
  - `git`, for origins using Git
  - `hg`, for origins using Mercurial
  - `svn`, for origins using Subversion
  - `bzr`, for origins using Bazaar
- **Origin url:** the url of the remote repository for the software origin.
  In order to avoid saving errors from Software Heritage, you should provide the clone/checkout url as given by the provider hosting the software origin.
  It can easily be found in the web interface used to browse the software origin.
  For instance, if you want to save a `git` origin into the archive, you should check that the command `$ git clone <origin_url>`
  does not return an error before submitting a request.

Archive and reference with Software Heritage

3. Cite and reference https://archive.softwareheritage.org/browse/origin/directory/?origin_url=<original_uri>

Link to a version of the software project



https://github.com/frmichel/sparql-micro-service

27 February 2023, 07:49:34 UTC

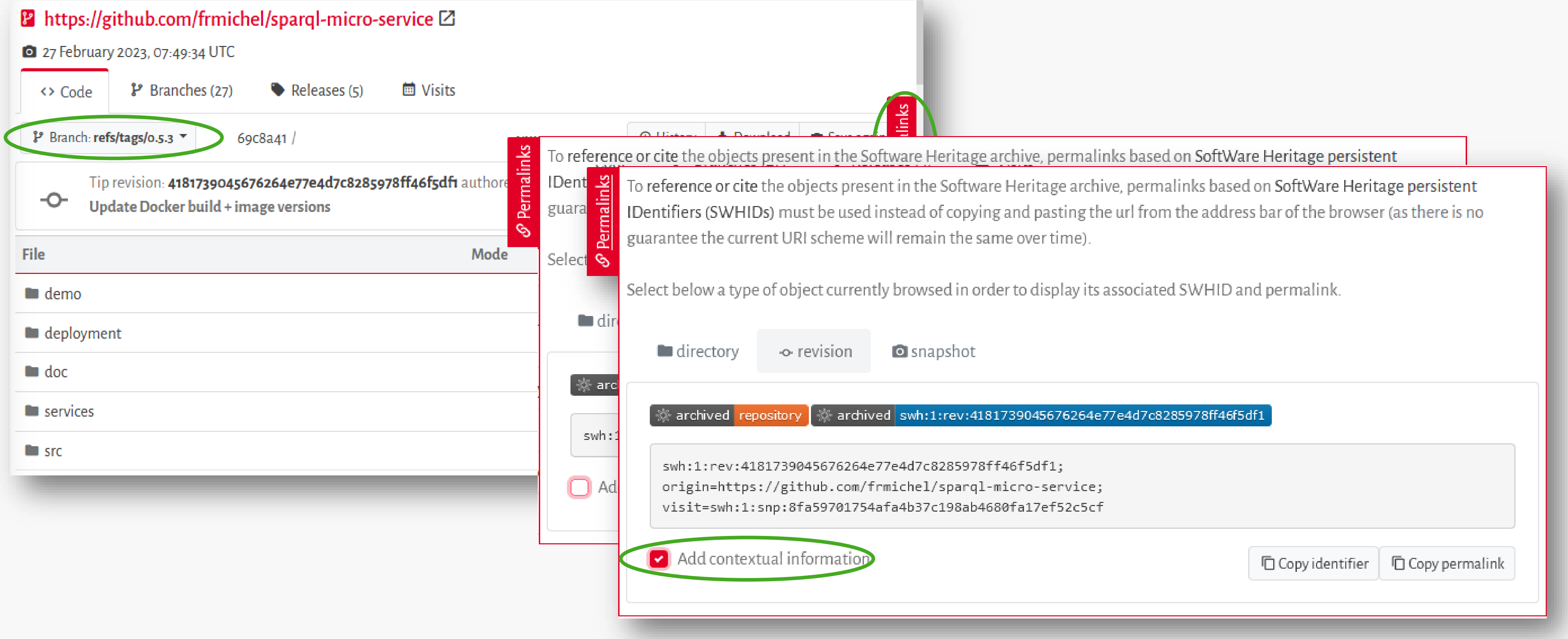
Tip revision: 4181739045676264e77e4d7c8285978ff46f5df1 authored by Franck Michel on 11 February 2022, 13:26:32 UTC
Update Docker build + image versions

File	Mode	Size
demo		
deployment		
doc		
services		
src		

Archive and reference with Software Heritage

3. Cite and reference https://archive.softwareheritage.org/browse/origin/directory/?origin_url=<original_uri>

   Link to a version of the software project



https://github.com/frmichel/sparql-micro-service

27 February 2023, 07:49:34 UTC

Branch: **refs/tags/0.5.3** 69c8a41 /

Tip revision: **4181739045676264e77e4d7c8285978ff46f5df1** authore

Update Docker build + image versions

File	Mode
demo	
deployment	
doc	
services	
src	

To reference or cite the objects present in the Software Heritage archive, permalinks based on SoftWare Heritage persistent IDentifiers (SWHIDs) must be used instead of copying and pasting the url from the address bar of the browser (as there is no guarantee the current URI scheme will remain the same over time).

Select below a type of object currently browsed in order to display its associated SWHID and permalink.

archived repository archived swh:1:rev:4181739045676264e77e4d7c8285978ff46f5df1

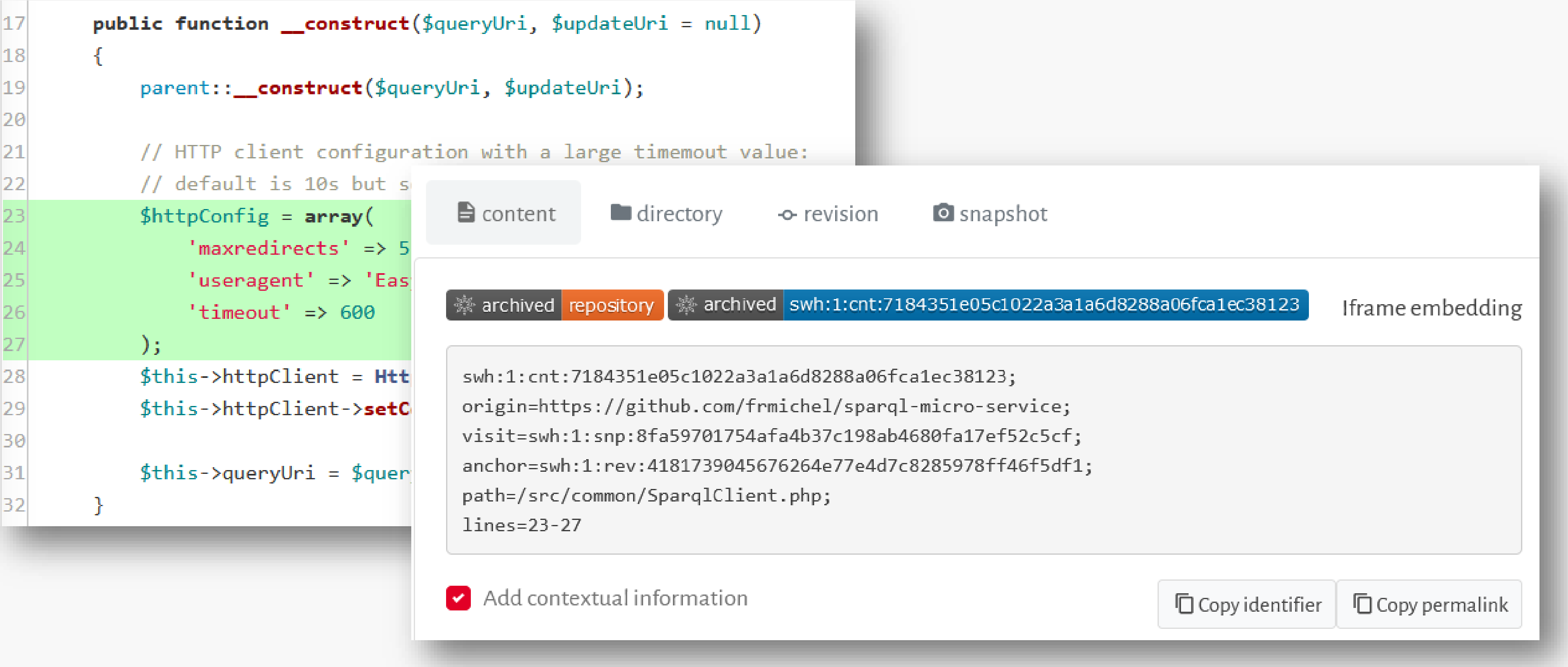
```
swh:1:rev:4181739045676264e77e4d7c8285978ff46f5df1;
origin=https://github.com/frmichel/sparql-micro-service;
visit=swh:1:snp:8fa59701754afa4b37c198ab4680fa17ef52c5cf
```

☑ Add contextual information

Archive and reference with Software Heritage

3. Cite and reference

   Link to a version of a source file, down to the line of code

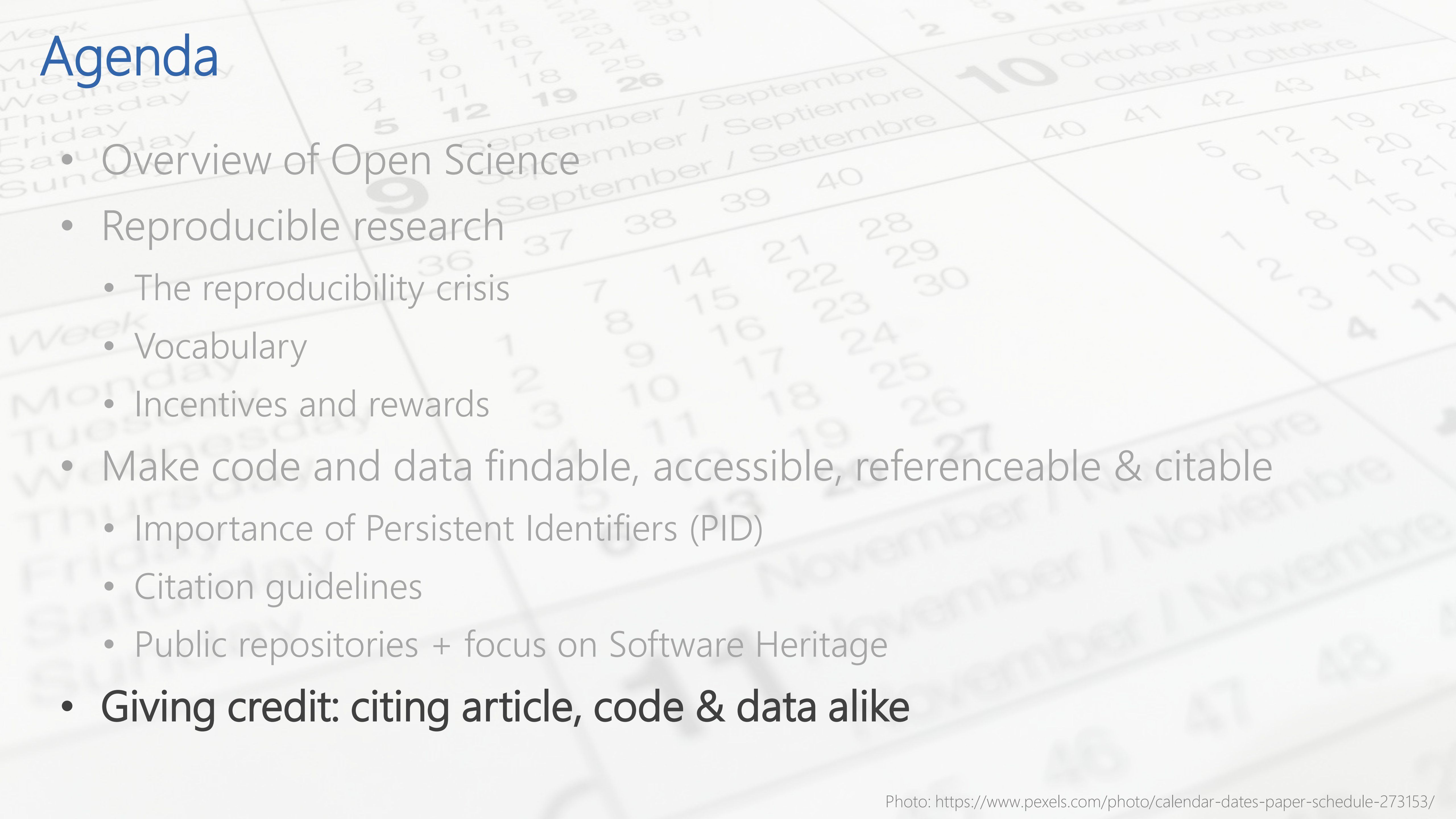


```
17      public function __construct($queryUri, $updateUri = null)
18      {
19          parent::__construct($queryUri, $updateUri);
20
21          // HTTP client configuration with a large timeout value:
22          // default is 10s but s
23          $httpConfig = array(
24              'maxredirects' => 5
25              'useragent' => 'Eas
26              'timeout' => 600
27          );
28          $this->httpClient = Htt
29          $this->httpClient->setC
30
31          $this->queryUri = $quer
32      }
```

```
swh:1:cnt:7184351e05c1022a3a1a6d8288a06fca1ec38123;
origin=https://github.com/frmichel/sparql-micro-service;
visit=swh:1:snp:8fa59701754afa4b37c198ab4680fa17ef52c5cf;
anchor=swh:1:rev:4181739045676264e77e4d7c8285978ff46f5df1;
path=/src/common/SparqlClient.php;
lines=23-27
```

☑ Add contextual information

Agenda



- Overview of Open Science
- Reproducible research
  - The reproducibility crisis
  - Vocabulary
  - Incentives and rewards
- Make code and data findable, accessible, referenceable & citable
  - Importance of Persistent Identifiers (PID)
  - Citation guidelines
  - Public repositories + focus on Software Heritage
- **Giving credit: citing article, code & data alike**

Photo: https://www.pexels.com/photo/calendar-dates-paper-schedule-273153/

Giving credit:
citing article, code & data alike

Cite others' works

Have a look at DataCite FIND, ACCESS, AND REUSE DATA

"A global community of organizations and researchers identifying and citing research outputs and resources."

Multiple services:

DOI registration, metadata management, discovery,
citation tracking, **citation formatter**, bibliometrics…

Non-profit association (German right)

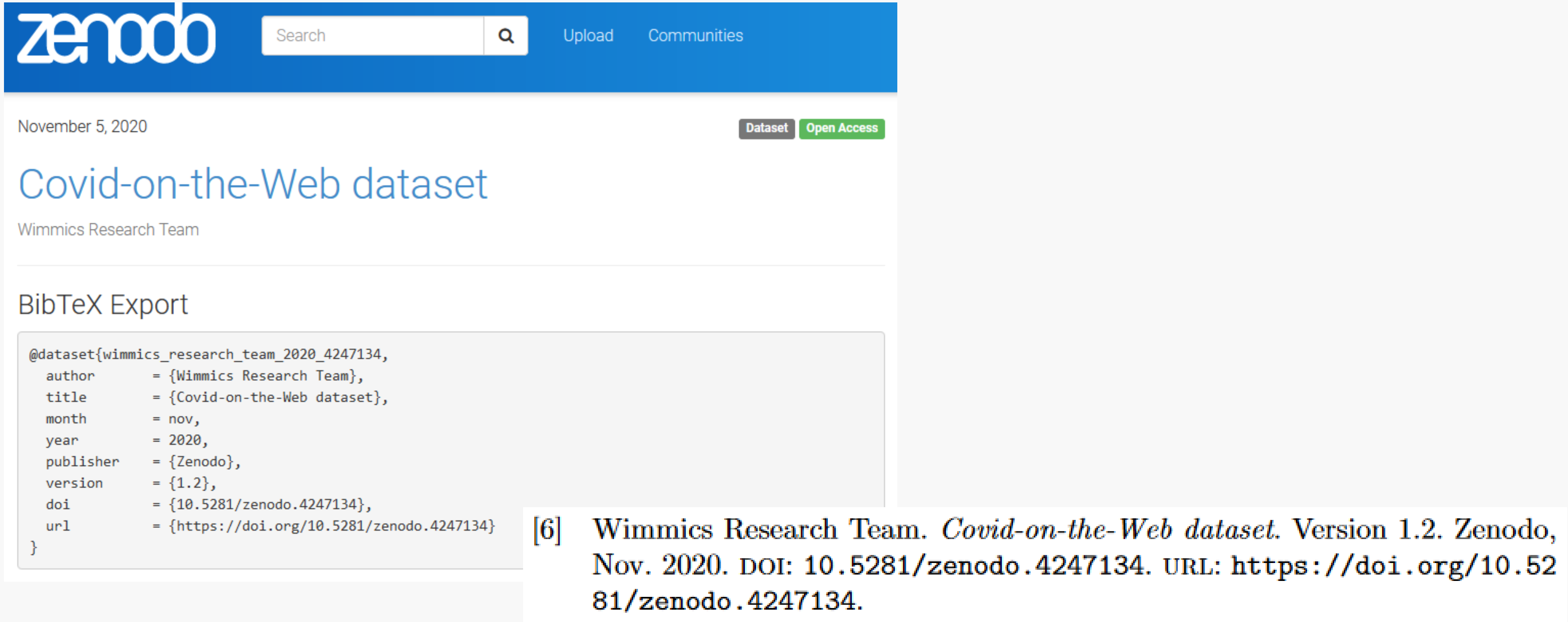
Federation of organizations allowing a single point of entry: institutions, universities, libraries, archives… (https://datacite.org/members/)

Not just data: code, articles and any research output.

Harvest existing repositories e.g. Zenodo, Crossref

Cite data in LaTex

@dataset entry type in BibLaTex (not in BibTex)



November 5, 2020

Covid-on-the-Web dataset

Wimmics Research Team

BibTeX Export

```
@dataset{wimmics_research_team_2020_4247134,
  author       = {Wimmics Research Team},
  title        = {Covid-on-the-Web dataset},
  month        = nov,
  year         = 2020,
  publisher    = {Zenodo},
  version      = {1.2},
  doi          = {10.5281/zenodo.4247134},
  url          = {https://doi.org/10.5281/zenodo.4247134}
}
```

[6] Wimmics Research Team. *Covid-on-the-Web dataset*. Version 1.2. Zenodo, Nov. 2020. DOI: 10.5281/zenodo.4247134. URL: https://doi.org/10.5281/zenodo.4247134.

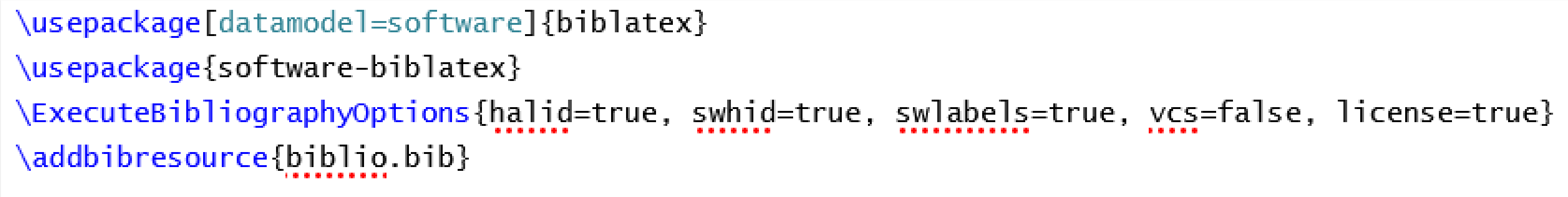
My experience: worked with LNCS, not with many other templates e.g. with ACM SIG-ALTERNATE.

Cite code in LaTex

software-biblatex: package that makes full use of SWHIDs, HAL ids, DOIs.
(https://www.ctan.org/tex-archive/macros/latex/contrib/biblatex-contrib/biblatex-software)

Extensions: `@software, @softwarerevision, @softwaremodule, @codefragment`

.tex file:

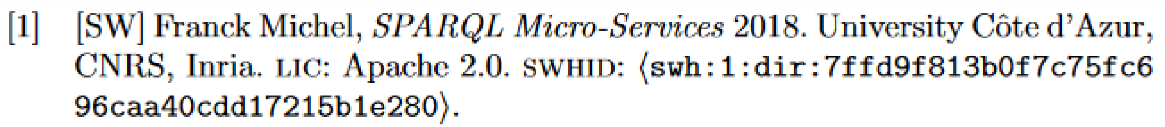


```
\usepackage[datamodel=software]{biblatex}
\usepackage{software-biblatex}
\ExecuteBibliographyOptions{halid=true, swhid=true, swlabels=true, vcs=false, license=true}
\addbibresource{biblio.bib}
```

My experience: worked with LNCS, not with many other templates e.g. with ACM SIG-ALTERNATE.

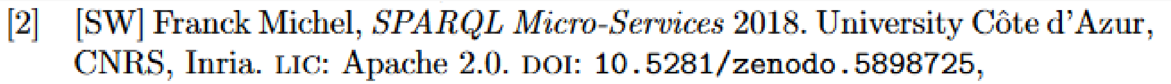
Cite code in LaTex

```
@software{sparql-micro-services,
    title = {SPARQL Micro-Services},
    author = {Michel, Franck},
    date = {2018},
    institution = {University Côte d'Azur, CNRS, Inria},
    license = {Apache 2.0},
    repository= {https://github.com/frmichel/sparql-micro-service/},
    swhid = {swh:1:dir:7ffd9f813b0f7c75fc696caa40cdd17215b1e280}
}
```



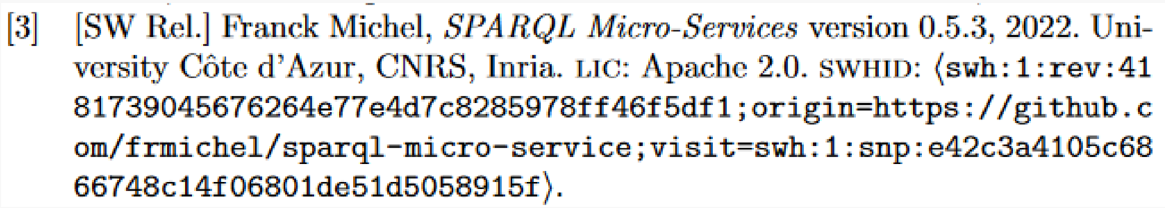
[1] [SW] Franck Michel, *SPARQL Micro-Services* 2018. University Côte d'Azur, CNRS, Inria. LIC: Apache 2.0. SWHID: ⟨swh:1:dir:7ffd9f813b0f7c75fc696caa40cdd17215b1e280⟩.

```
@software{sparql-micro-services-doi,
    title = {SPARQL Micro-Services},
    author = {Michel, Franck},
    date = {2018},
    institution = {University Côte d'Azur, CNRS, Inria},
    license = {Apache 2.0},
    doi = {10.5281/zenodo.5898725},
    repository= {https://github.com/frmichel/sparql-micro-service/}
}
```



[2] [SW] Franck Michel, *SPARQL Micro-Services* 2018. University Côte d'Azur, CNRS, Inria. LIC: Apache 2.0. DOI: 10.5281/zenodo.5898725,

```
@softwareversion{sparql-micro-services-0.5.3,
    crossref = {sparql-micro-services}
    version = {0.5.3},
    date = {2022},
    swhid = {swh:1:rev:4181739045676264e77e4d7c8285978ff46f5df1;
      origin=https://github.com/frmichel/sparql-micro-service;
      visit=swh:1:snp:e42c3a4105c6866748c14f06801de51d5058915f}
}
```



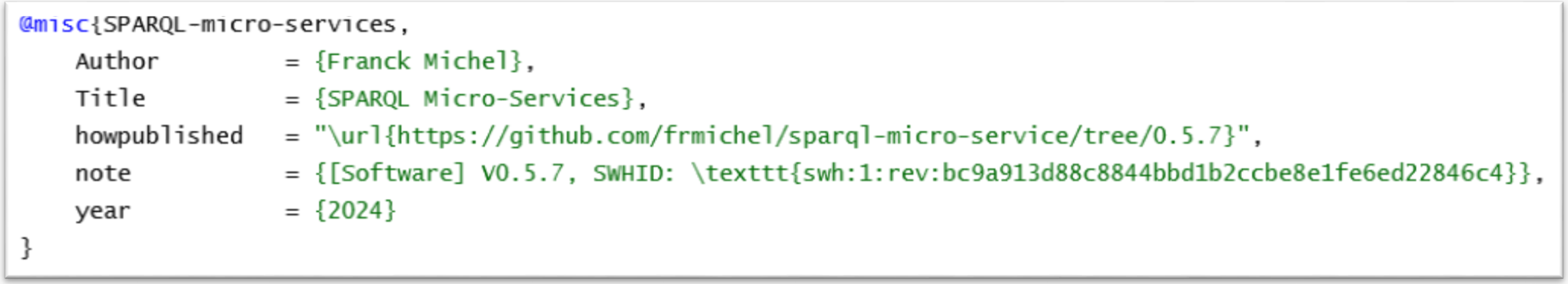
[3] [SW Rel.] Franck Michel, *SPARQL Micro-Services* version 0.5.3, 2022. University Côte d'Azur, CNRS, Inria. LIC: Apache 2.0. SWHID: ⟨swh:1:rev:4181739045676264e77e4d7c8285978ff46f5df1;origin=https://github.com/frmichel/sparql-micro-service;visit=swh:1:snp:e42c3a4105c6866748c14f06801de51d5058915f⟩.

What if only BibTex can be used?

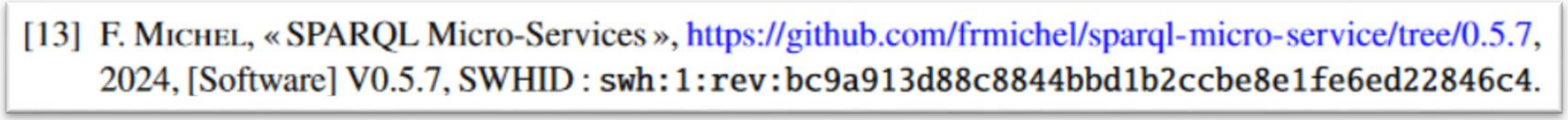
Cite data paper/software paper/resource track paper, if any

Use BibTex entry type @misc and field ‘note’ to add version, license, SWHID, HAL id, DOI…

https://www.bibtex.com/e/entry-types/#misc



```
@misc{SPARQL-micro-services,
    Author          = {Franck Michel},
    Title           = {SPARQL Micro-Services},
    howpublished    = "\url{https://github.com/frmichel/sparql-micro-service/tree/0.5.7}",
    note            = {[Software] V0.5.7, SWHID: \texttt{swh:1:rev:bc9a913d88c8844bbd1b2ccbe8e1fe6ed22846c4}},
    year            = {2024}
}
```



[13] F. MICHEL, « SPARQL Micro-Services », https://github.com/frmichel/sparql-micro-service/tree/0.5.7, 2024, [Software] V0.5.7, SWHID : swh:1:rev:bc9a913d88c8844bbd1b2ccbe8e1fe6ed22846c4.

But not standard so likely not machine-processable ☹

TAKE
AWAY

Good research is reproducible research

"Open Science is science done right"

Open Access, Open Data, Open Source

Cultural shift to openness and collaboration

Make code & data "FARC"

Data on **recherche.data.gouv.fr** or Zenodo or…

Code on **SWH and HAL**, or Zenodo and HAL

Prepare repositories, rich metadata

PID is *(almost)* all you need

Name a resource, Cite it, Reference a specific version, Associate metadata to it

Still a long way to go…

Editors must support code/data citation

- Update the Latex templates!
- Lengthy PIDs require unlimited pages for references

Track article/code/data relationships

- Metadata (on most portals) often not ready. HAL doing it but progress still needed

Give credit & reward:

- Change the metrics to reward impactful code/data
- Reward reviewing work…
- Change the mentalities…

More: **DORA, CoARA**

Thank you!

Links:
https://www.ouvrirlascience.fr/passeport-pour-la-science-ouverte-guide-pratique-a-lusage-des-doctorants/
https://www.ouvrirlascience.fr/science-ouverte-codes-et-logiciels/
https://www.ouvrirlascience.fr/partager-les-donnees-liees-aux-publications-scientifiques-guide-pour-les-chercheurs/