

From Data to Code

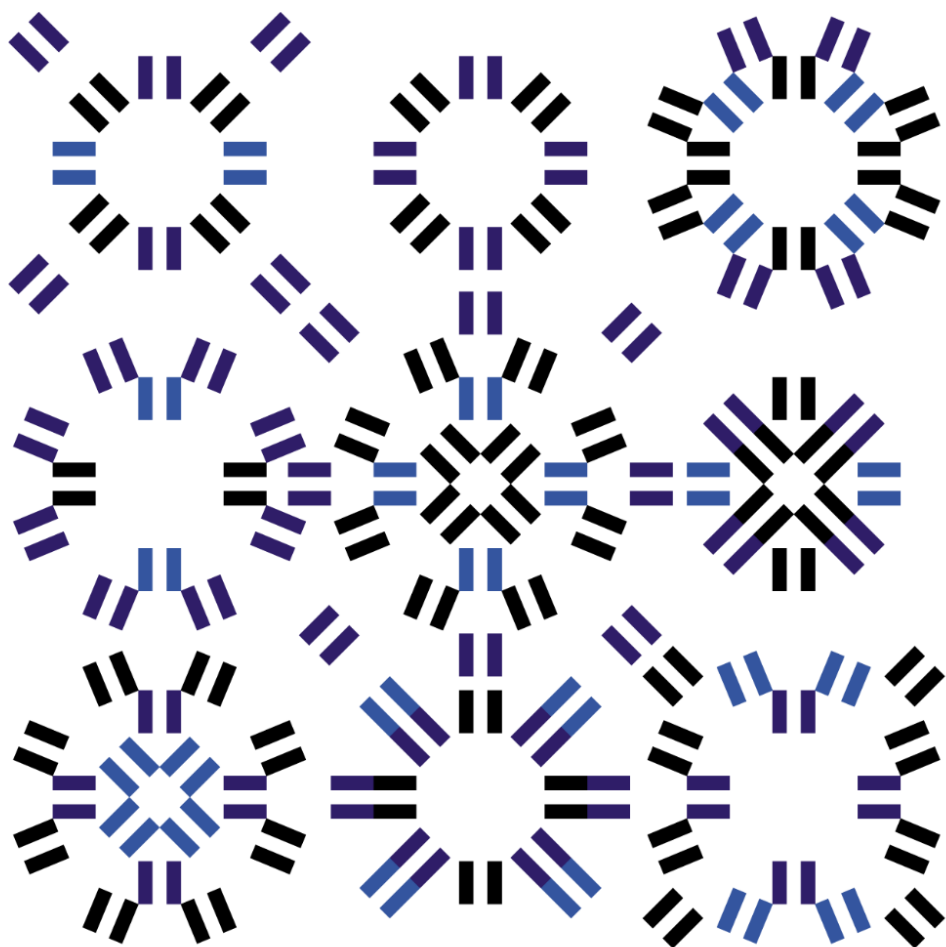
How Libraries advance Open Science practices on source code
Through Data Management Clusters?

Damien Belvèze

damien.belveze@univ-rennes.fr

Université de Rennes

2026-03-25



Activities on Source Code within Data Clusters



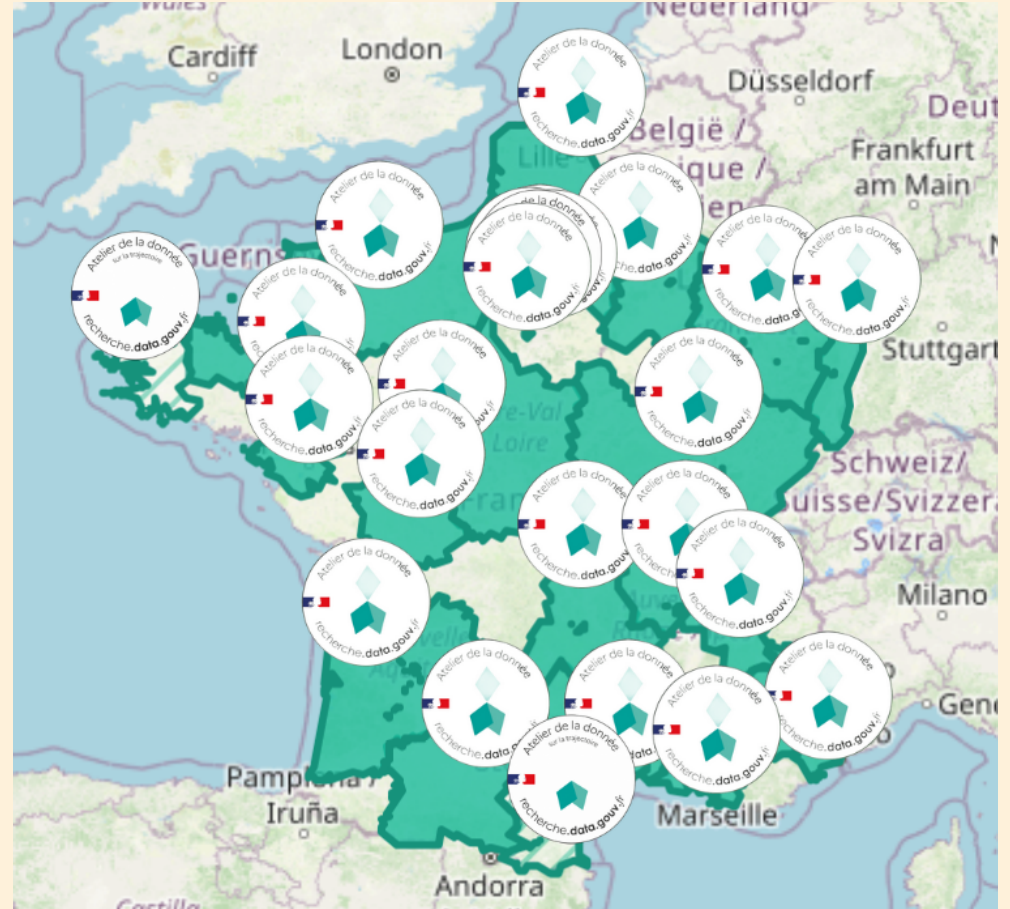
(to display speaker notes, strike “S” key)

what are Data clusters?

Recherche Data Gouv Ecosystem

- local competence centers dedicated to supporting researchers in FAIR policy implementation
- different profiles (Librarians, Archivists, Research Engineers, Data Protection Officer)
- different research institutions operating on the same territory
- collaborate with other Data clusters on specific topics (work group 5 : research source code)
- national Repository is central in a Data Cluster's activity

👉 [Data Cluster lists](#)



Data clusters are part of the global ecosystem framed by the National Data Repository 'Recherche Data Gouv' and its catalogue (still in beta version). Data is central in our activity but to some extent, because data are processed by software and we need both of them to reproduce experiences, source code is also within the scope of the data clusters ; there is an ongoing debate to know up to which point source code should be covered by data curators, since both objects are very different : data are static objects, after the publication of the results, data won't change much, while source code is more dynamic and may involve other developers or even communities long after the publications are made. Besides, some say that skills needed for data curation are not the same as skills required by source code maintenance, promotion and reproducibility

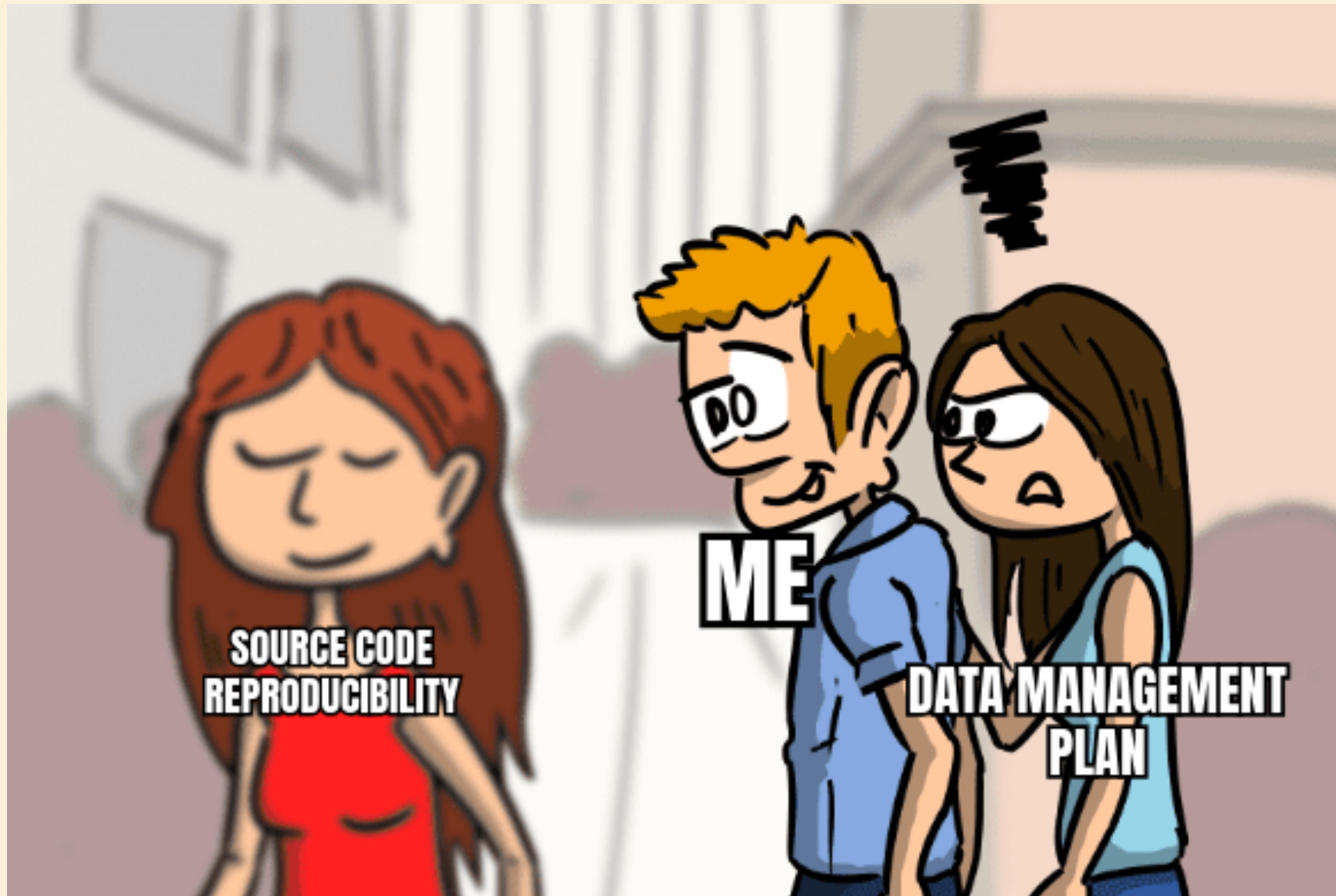
ARDoISE Datahub

- committee : 16 members + corresponding experts
1. centralizes all matters related to the code through [guichet-ardoise](#)
 2. helps researchers to write their DMP
 3. trains them to FAIR data and dissemination of research data and source code
 4. helps them to use data repositories, especially Research Data Gouv
 5. advises them on data preservation and safety
 6. advises them on legal and ethical aspects of data management
 7. **Provides them with information on source code dissemination and reproducibility**
 8. guides them in the use of lab notebooks
 9. receives them in dedicated places in University Rennes 2's library



ARDoISE
Atelier rennais
de la donnée

Sometimes, I feel like...



University policies in Open Source Software

Path Three :

Opening up and promoting source code produced by research

7

Recognize and support the dissemination under an open source license of software produced by publicly funded research programmes

« The opening of software source code is a major challenge for the **reproducibility** of scientific results. »

8

Highlight the production of **source code** from higher education, research and innovation

9

Define and promote an **open source software policy**

« Distribution of software products under **open source licence** will be preferred. »

Source code was neglected in the first version of the French National roadmap toward Open Science but was added later on the 2nd version (2022) as one of the pillars of Open Science close to Publications and Data.

In order to promote Open Source Software in the field of Open Science, our Data cluster has included services on Open Source based on infrastructures : those infrastructures are : - forges - HAL - SoftWare Heritage we try to develop habits of using these platforms to share source code and make them reproducible

We lack a nationwide forge for research source code. In the Social and Human Sciences , academics can rely on GitLab, hosted by Huma-Num , but no equivalent exists for the exact sciences.

working in pair, but with different and multiple connections



Damien Belvèze

- member of Working Group 5 within Research Data Gouv ecosystem
- webinars with INRIA members
- tests Guix package manager



Kilian Heutte

- experienced in R and Rmarkdown
- works with MSHB research engineers
- tests data anonymisation tools

What services do Data Clusters provide to research community regarding research source code?

- communicate on Mastodon / LinkedIn on source code sharing and preserving best practices
- [showcase source code linked to projects they followed up](#)
- write [pages on their websites](#) or blogposts on hypotheses
wordpress websites on source code management
- webinars : how to describe source code and software in a research project (example [here](#))



Bourdaret et al. ([2026](#))

Speaker notes

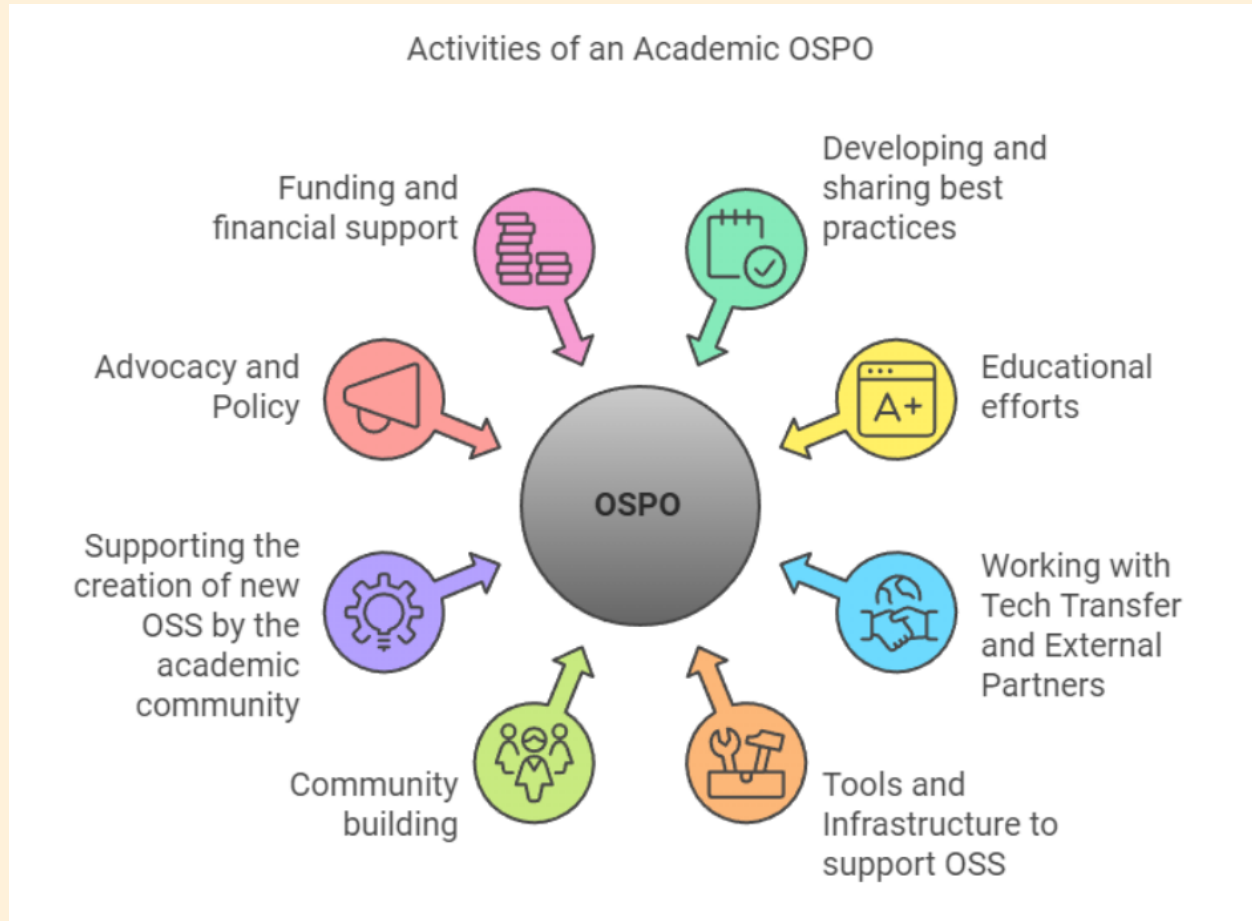
last year, during their internship, Librarians trainees made an inquiry on the behalf of the University of Nanterre to design a set of services that could be offered to this University's research community on research source code development and promotion. I had a pleasant interview with them since one of the first steps of their job was to cover what different stakeholders and among them Data Clusters already do on this topic The work cited here is their. Thank you for their comittment and output.

other stakeholders than Data Clusters take part to source code tracking and promotion:

- groups of research engineers focused on source code management
- University Libraries (cf. University of Nanterre's Library)
- OSPOs (see University of Grenoble's OSPO)
- ARDoISE Data Cluster advocates for becoming part of the future Rennes academic OSPO

project of the University of Nanterre to allocate staff resource to this topic (roadmap to design services that could make easier to support researchers in developing and publishing research open source code)

“University OSPOs are the local execution engines for Open Source Software policies” 📖 Di Cosmo (2025)



(figure) 📖 Young et al. (2024)

Speaker notes

our activity is mainly based on “developing and sharing best practices” and “educational efforts” But we take part to the visibility of research infrastructures such as HAL to link software to publications or Software Heritage Archive to preserve source code.

According to OSPO-Radar project members, our role is also seen as important in community building ; we were mandated by our Vice-Presidents to gather information in order develop a plan to build an OSPO for Rennes research community




Integration in OSPOs

Academic Open Source Programme Offices are built by federating university services concerned with Open Source Programmes :

- Intellectual property office, Technology Transfer Office
- research teams and labs
- libraries (involved in curating HAL and Recherche Data Gouv portals)
- computing centers and experimentation centers
- stakeholders from the local open source industry (including non profit organisations such as OW2)

[Wikipedia](#) : “OW2 is an independent non-profit international consortium dedicated to developing open-source software code infrastructure for middleware information systems. OW2 federates IT vendors and users, universities, and research centers from Europe, Asia, and the Americas, representing IT professionals.”

services adapted to different targets

target	Experience	project's type and dimension
 PhD Students	are learning to code	scripts
 development for individual needs made by researchers	trained users	scripts, add-ons to known software
 collaborative development	state of the art development	comprehensive software



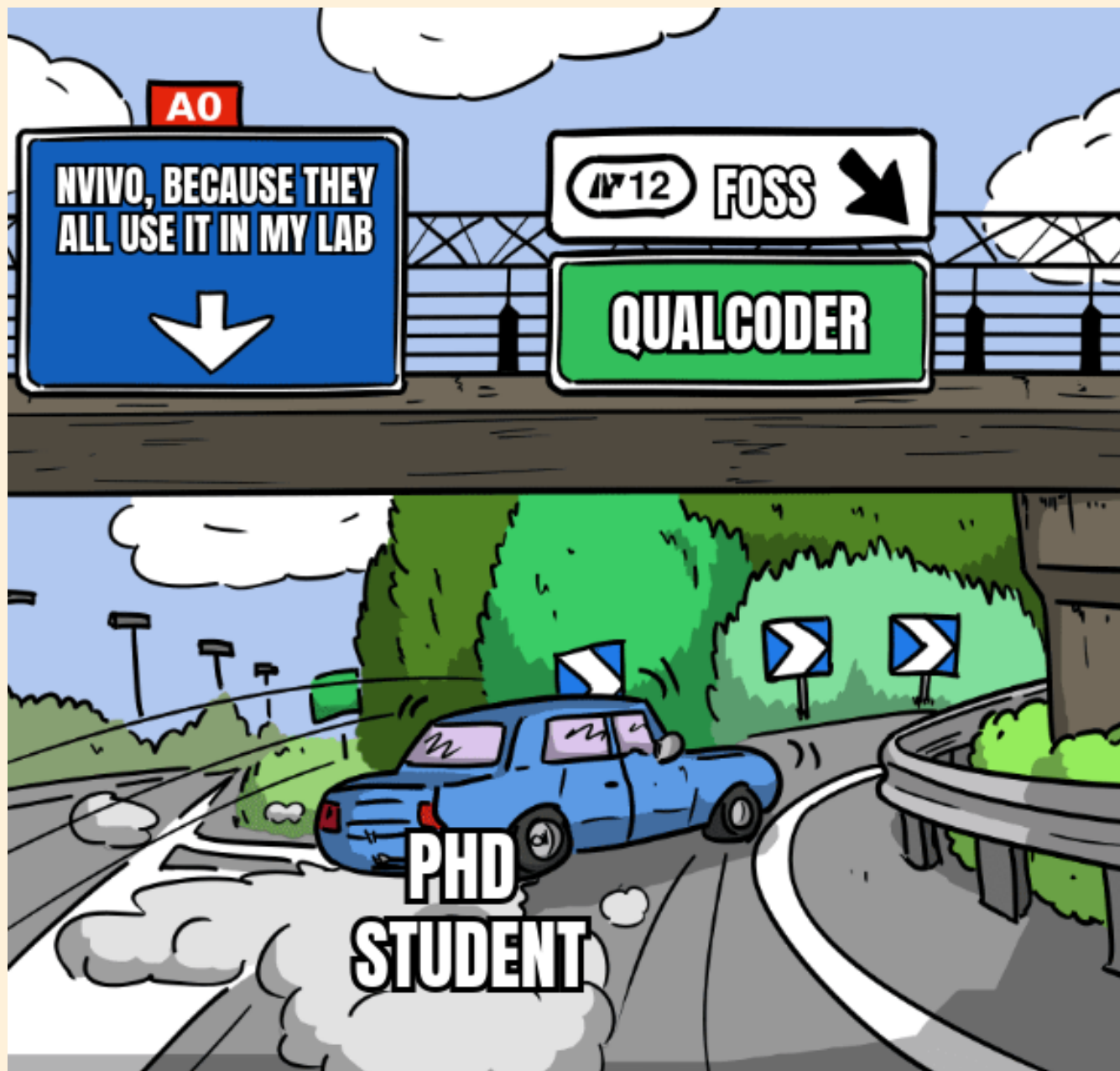
- Seminar organized by ARDoISE on source code reproducibility on december 2024
- survey on the use of research software by Social and Humanity Science Researchers

Support for PhD Students

basic programming skills (R / Python) acquired during their master

- documentation (codemeta, SMP)
- development good practices (virtual env)
- archiving and citing : Software Heritage Archive, SWHID
- promotion of open source programmes to process research data

codemeta (often written in javascript, sometimes in txt) containing the main part of the metadata required to identify a specific software and its computational environment. GT5 tries to make it match with SMP (“Software Management Plan”’s requirements) ; templates to write SMP are provided on the dmp.opidor.fr website along with Data Management Plan templates (more accurately, when source code is involved in the experiment, SMP is a section of the related DMP)



Nvivo is a proprietary tool to conduct qualitative data analysis. There are not enough licences bought to match all the needs. Besides, we are not sure that outputs obtained with Nvivo are aligned to open science standards in the field (rqda), that's why we are trying to replace Nvivo in the habits of young researchers by Qualcoder which comes with its own Graphical User Interface (because the lack of GUI, or poorly developed GUI, is often seen as a problem to make a FOSS software popular)

PhD training

- data visualisation with Quarto ([introduction to Quarto with R and Rstudio](#))
- [initiation to R and Rmarkdown](#)
- [introduction to Jupyter notebooks](#)
- [initiation to Git and Gitlab](#)
- Data Anonymisation tools
- [how to open, share and preserve reproducible source code](#)



individual researchers

they often have already acquired strong development practices since their PhD by collaborating with research engineers

- Questions about licences
- archiving and citing : Software Heritage Archive, SWHID

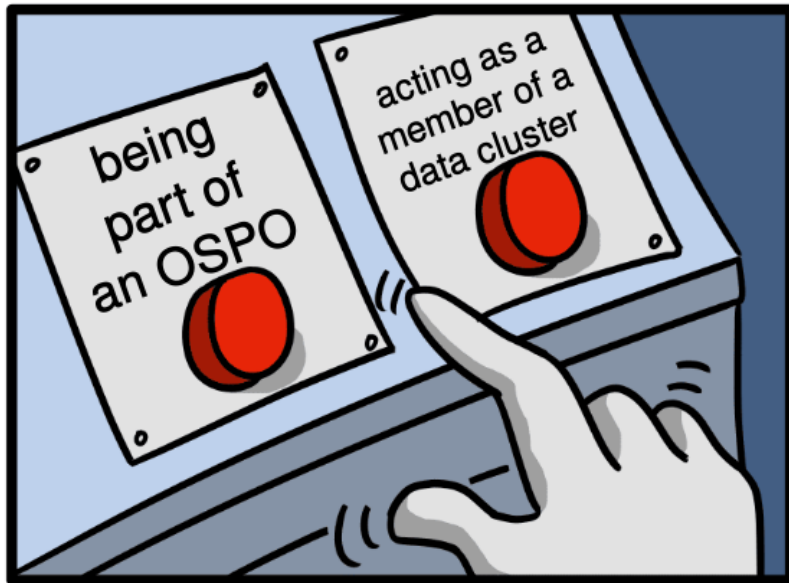
Speaker notes

about software licences, we only give first-grade information and for trickier questions we send people to technology transfer office. We invite these colleagues to participate to webinars we schedule every year

research teams

discussions with software engineers about :

- visibility : HAL, catalogue-esr.gouv.fr
- archiving and citing : Software Heritage Archive, SWHID



GEE, BASED ON JAKE-CLARK.TUMBLR

Looking for the good level:

- Data Cluster? (but more focused on research data)
- Library? (focused on linking publications to underlying source code)
- Library / Data Cluster as member of an OSPO? (what will be our role in this new frame)
- Other?

References

memes are made with [framameme generator](#)

Bourdaret, C., Leclerc, C., Ronat, M., & Sepulveda, B. (2026). *Codes et logiciels pour la recherche* (p. 50) [Gestion de Projet]. ENSSIB.

Di Cosmo, R. (2025). *Academic OSPOs in France - supporting the software pillar of open science*.

Young, Barba, L., Choudhury, S., Flannagan, Lippert, D., & Littauer, R. (2024). *A definition of an Academic OSPO*.