Academic OSPOs in France

Supporting the Software Pillar of Open Science

Roberto Di Cosmo

Director, Software Heritage
Inria and Université Paris Cité

September 23rd 2025

Software Heritage

THE GREAT LIBRARY OF SOURCE CODE

Outline

Short Bio: Roberto Di Cosmo

Computer Science professor in Paris, now working at INRIA

- *35+ years* of research (Theor. CS, Programming, Software Engineering, Erdos #: 3)
- *25+ years* of Free and Open Source Software
- *15+ years* building and directing structures for the common good

1999 *DemoLinux* – first live GNU/Linux distro

2007 *Free Software Thematic Group*
150 members 40 projects 200Me

2008 *Mancoosi project* www.mancoosi.org

2010 *IRILL* www.irill.org

2015 *Software Heritage* at INRIA

2018 *National Committee for Open Science*, France

2021 *EOSC Task Force on Infrastructures for Software*, European Union

Open Source, and Open Source Program Offices (OSPOs)

"Free Software", "Open Source": a reminder [1]

- Freedom #0, to run the program, for any purpose
- Freedom #1, to study how the program works, and change it
- Freedom #2, to redistribute copies
- Freedom #3, to improve the program, and release improvements

Open Source Program Office (OSPO) in industry – the traditional mandate

- Competence center: strategy, policy, governance. [2, 3]
- Core functions: license compliance, IP hygiene; security; contribution guidelines; upstream and community engagement; enablement/training.
- Many large orgs operate an OSPO (recent LF research). [4]

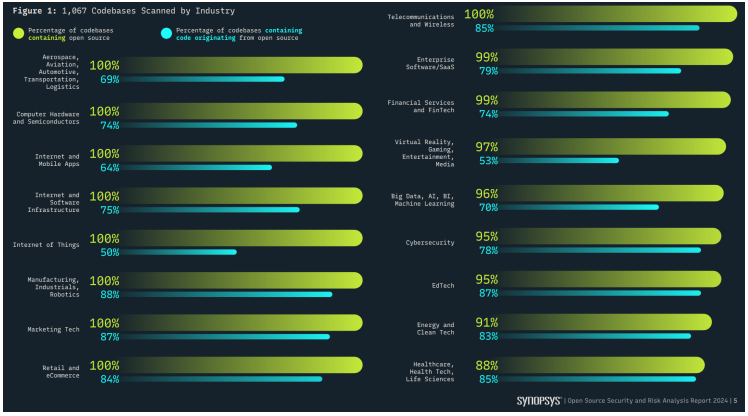
[1] FSF Free Software Definition; OSI Open Source Definition

[2] TODO Group OSPO definition. ospodefinition.org

[3] OSI blog: What is an OSPO and why you should have one

[4] Linux Foundation, "State of OSPOs & Open Source Management 2024". LF OSPO 2024

The Open Source tidal wave in audited commercial codebases



In a nutshell

- 96–97% Synopsys audited projects include OSS (~77% is OSS) [OSSRA 2024/2025] [5]

[5] Synopsys, "Open Source Security and Risk Analysis" OSSRA 2024 report and OSSRA 2025

Software Heritage: a *universal* software archive, as a shared infrastructure

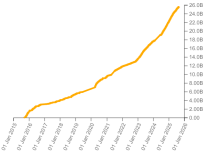
One infrastructure
non profit
open and shared

Unveiled in 2016

The largest archive ever built

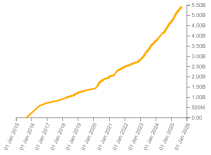
Source files

25,681,084,548



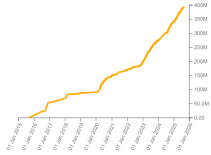
Commits

5,427,907,921



Projects

394,217,408



Diamond sponsors

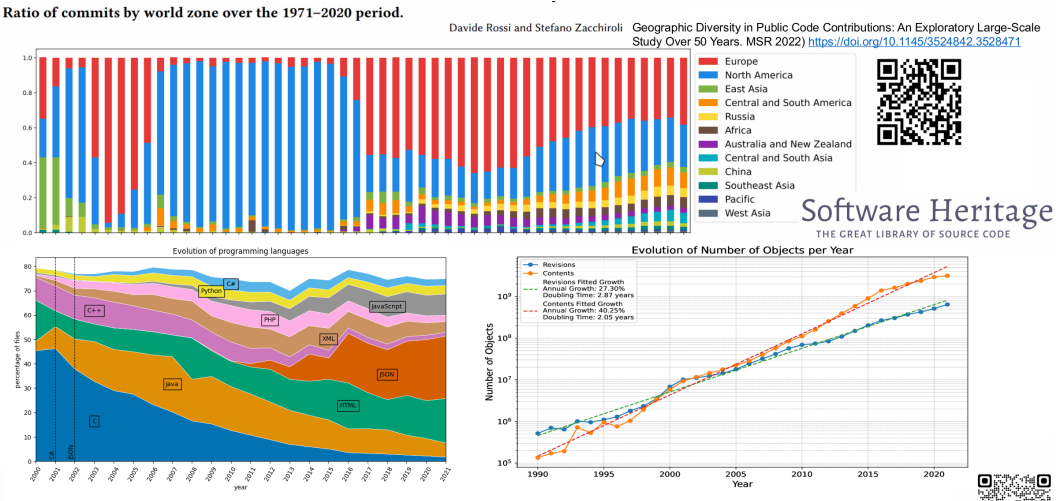
Platinum sponsors

Gold sponsors

Silver sponsors

Bronze sponsors

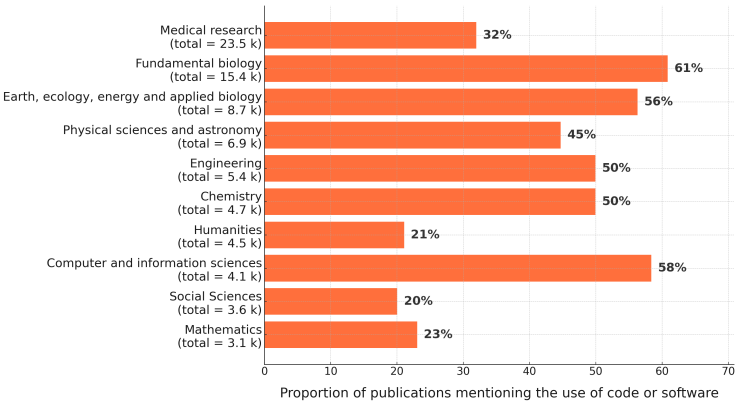
Software sky (image: Software Heritage very large telescope)



Outline

(Open) Source Code is a pillar of Open Science

Software powers modern research



French Open Science Monitor 2025

From national to global: Open Science Monitoring Framework Initiative

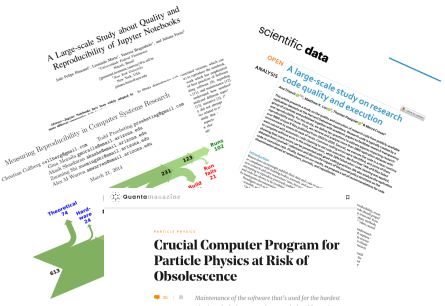
MINISTÈRE CHARGÉ DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE

Liberté
Égalité
Fraternité

unesco

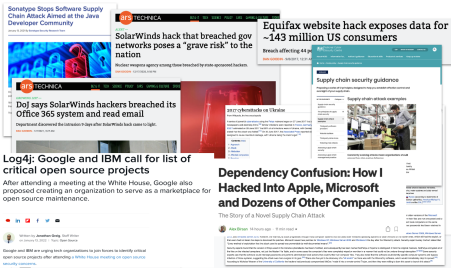
How are we managing our software ?

Reproducibility, maintenance in Academia



(articles: here, here, here and here)

Security, integrity, traceability in Industry



Can they track the software that they

- ship, use, acquire
- has that bug or vulnerability

awareness is raising at the level of public policy

What is at stake for research software

ARDC

- Archive for retrieval (*reproducibility*)
- Reference for identification (*reproducibility*)
- Describe for discovery and reuse
- Cite/Credit for credit and evaluation

Before ARDC

- Development practices and tools (VCS, build system, test suites, CI, code quality, …)
- Opening up towards a community (documentation, organization, communication)

Need training, tooling, infrastructures, best practices

Beyond ARDC

- Policies (dissemination, reuse, careers, …)
- Sustainability (legal, financial, etc.)
- Technology transfer
- Advanced technologies and tools (quality, traceability, etc.)

a humbling challenge, and a complex one (we are not in a vacuum)

French National plan for Open Science, 2021-2024

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

Liberté
Égalité
Fraternité

SECOND FRENCH PLAN FOR OPEN SCIENCE

Generalising open science in France 2021-2024

1

Launch on 6 July 2021 by Frédérique Vidal, Minister for Higher Education, Research and Innovation

- Multiplying the **levers for change** in order to **generalise open science practices**
- Structuring the **policy for opening up or sharing research data**
- New commitments to the **opening of source code** produced by research
- **European and international inclusion** in the context of the French Presidency of the European Union
- **Disciplinary and thematic variations**: open science policies must be adapted to disciplinary specificities

2

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

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

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

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

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

3

Define and promote an open source software policy

- Produce a **National Charter for Open Source Software** coming from higher education, research and innovation
- Develop the **link between data and software** through a network of **Chief Data Officers** in the various universities and research performing organisations.
- Develop the **economic models of open source software** and make them known within commercialization services
- **Support Software Heritage** and recommend it for the archiving and referencing of source code

Recognise source code as a contribution to research

- Create an **open source research software prize**
- **Provide greater recognition** for software production in the career of researchers, research support staff

Build an ecosystem that connects code, data and publications

- Develop **proper coordination** between software forges, open publication archives, data repositories and the scientific publishing sector.

4

Plan national pour la science ouverte — Path Three

Key points from the plan

- Recognize open-source dissemination of research software as the norm.
- Catalog & monitor research software at national level
- Connect code $\leftrightarrow$ data $\leftrightarrow$ publications.
- Archive & reference source code and version history via Software Heritage and HAL.
- Enable proper citation and career recognition.
- Adress sustainability issues (economic model) and proper infrastructure support

National coordination: the Software College in the CoSO

Five action lines (see details online) — Since April 2022

- Identifying and highlighting research software production
- Technical and social tools and best practices
- Valorization and sustainability
- Liaison and animation at national, European, and international levels
- Recognition and careers

Source Code primer

key concepts

- for students
- for teachers
- for researchers

Report on software forges

in academia (FR):

- needs
- options
- limitations

Annual award

Establishing a national research software award. Open Research Europe 2023

A broad membership

See the official Software College web page

Outline

French academic software (image: Software Heritage very large telescope)

3531626

Contributions

46533

Software Projects

Active last 12 months: 3425

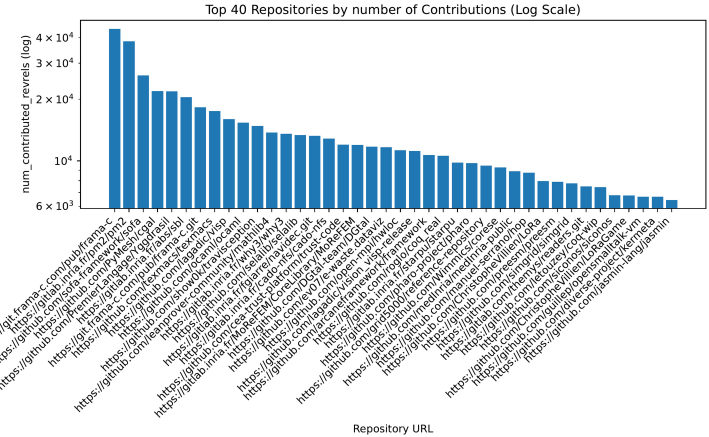
Still crawled: 44874

30101

Contributors

1992-07-08
2025-05-17

Time Span

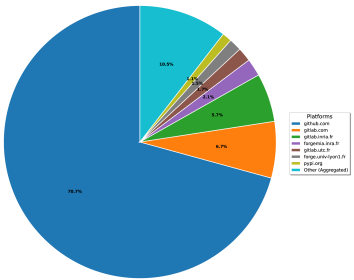


domain	count
inria.fr	3833
univ-lyon1.fr	2806
univ-orleans.fr	1555
uca.fr	1371
jussieu.fr	1247
utc.fr	1216
cea.fr	1187
insa-toulouse.fr	747
ens-lyon.fr	706
univ-artois.fr	696
inrae.fr	690
univ-lemans.fr	688
cnrs.fr	682
ens.fr	675
inra.fr	644
univ-fcomte.fr	634
unilim.fr	541
insa-rennes.fr	491
univ-avignon.fr	489
univ-pau.fr	483
utt.fr	472
polytechnique.fr	437
uha.fr	422
pasteur.fr	420
telecom-paris.fr	412
univ-cotedazur.fr	363
univ-smb.fr	352
univ-tours.fr	350
u-picardie.fr	349
enpc.fr	328
sciencesconf.fr	327
universite-paris-saclay.fr	323
univ-poitiers.fr	312
u-paris.fr	289
univ-angers.fr	267
univ-parisdl.fr	264
insa-rouen.fr	211
centralesupelec.fr	192
ird.fr	171
imt-atlantique.fr	159
onera.fr	158
univ-ubse2.fr	131
inria.fr	130
enac.fr	126
univ-eiffel.fr	125
agroparistech.fr	105
cnam.fr	104
mines.fr	95
uphf.fr	94
univ-tln.fr	89

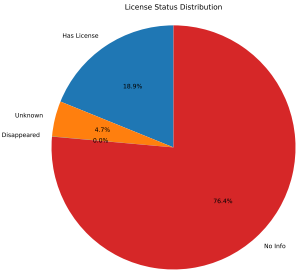
Note: Showing top 50 of 51 total domains

French academic software: where is it, how is it licensed?

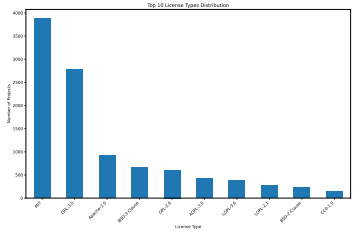
Main code hosting platforms



Licence Status



Top 10 licences



Outline

The (ideal) Role of an Academic OSPO

University OSPOs are the local execution engines for these policies

- interface with key University bodies: tech transfer, legal, open science, etc.
- practical advice, training to researchers/engineers/PhD students on:
  - Archive & Reference: in Software Heritage and HAL
  - Describe & Cite: codemeta.json, software citation, link to outputs and evaluation
  - Compliance & License: open by default; decision flow; legal support; compatibility
  - Development & Dissemination: forges, CI/testing, packaging, onboarding communities
- follow closely national and local policy evolution, leverage resources:
  - Infrastructures: Software Heritage+HAL; national catalogs/monitors
  - Training: SW Eng basics; reproducible research practices
  - Communities: CURIOSS, RSE/RDA/ReSA; national working groups (DevLog, Calcul, etc.)
  - Internal partners: libraries/data stewards, transfer, legal, etc.
- create a view of the software production for their University/Institution

Analogies and differences with Industry OSPOs

- Same core:
  - policy, compliance, security, community, education, standards, metrics, developmnent best practices
- Different levers:
  - industry optimizes product & risk;
  - academia optimizes openness, credit, and continuity despite short cycles.
- Practical setup in academia:
  - separate from campus IT, OSPO prioritizes
    - policy, training, archival (e.g., deposits, PIDs), and
    - upstream bridge-building over operating product pipelines.

UGA's OSPO as a model case

Long standing collaborations

- Software College: Violaine Louvet working on the catalog
- Open Science service: Joanna Janik
- fluid interaction with technology transfer and legal offices

More recently

- Software Heritage: one research engineer embedded in GRICAD

Outline

UGA's software through Software Heritage's very large telescope

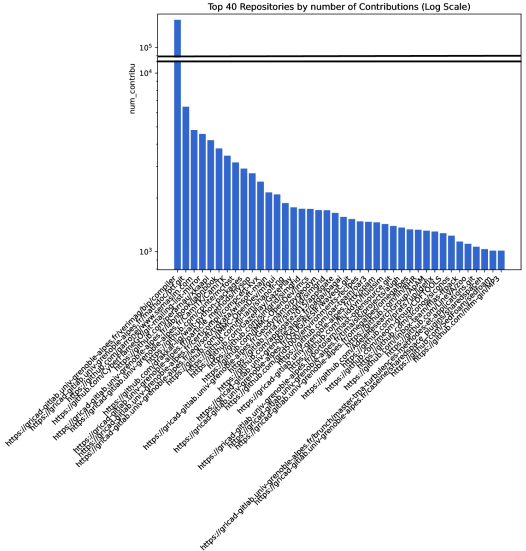
Key figures

481644
Contributions

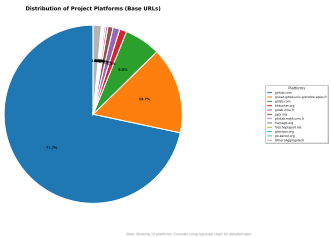
9314
Software Projects
Active last 12 months: 404
Still crawled: 8748

3583
Contributors

1993-10-19
2025-05-16
Time Span

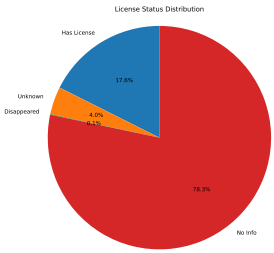


Main code hosting platforms

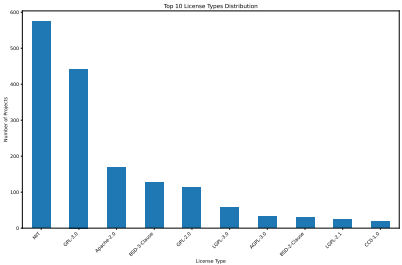


UGA's software through Software Heritage's very large telescope

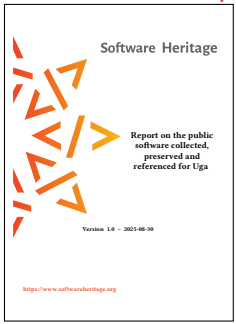
Licence Status



Top 10 licences



More in the full report



A lot to do… Policy, people, tools are here: let's move forward!