

Software Heritage

CURI OSS

Bastien Guerry
Head of partnerships
bastien.guerry@inria.fr

February 5th, 2026



Software Heritage
THE GREAT LIBRARY OF SOURCE CODE

1 Software Heritage's mission

2 A new phase: products and services for all stakeholders

3 OSPO Radar

4 Strategic Software Insights Report

Unveiled in 2016



Software Heritage

THE GREAT LIBRARY OF SOURCE CODE

Collect, preserve and share *all* software source code

Preserving our heritage, enabling better software and better science for all

Unveiled in 2016



Software Heritage

THE GREAT LIBRARY OF SOURCE CODE

Collect, preserve and share *all* software source code

Preserving our heritage, enabling better software and better science for all

Reference catalog



**find and reference all
software source code**

Unveiled in 2016



Software Heritage

THE GREAT LIBRARY OF SOURCE CODE

Collect, preserve and share *all* software source code

Preserving our heritage, enabling better software and better science for all

Reference catalog



Universal archive



find and reference all
software source code

preserve and share all
software source code

Unveiled in 2016



Software Heritage

THE GREAT LIBRARY OF SOURCE CODE

Collect, preserve and share *all* software source code

Preserving our heritage, enabling better software and better science for all

Reference catalog



find and **reference** all
software source code

Universal archive



preserve and **share** all
software source code

Research infrastructure



enable analysis of all
software source code

Sharing the vision



United Nations
Educational, Scientific and
Cultural Organization



And many more ...

www.softwareheritage.org/support/testimonials

An international, non profit initiative

for the long term

Sharing the vision



United Nations
Educational, Scientific and
Cultural Organization



Members and sponsors

Inria



And many more ...

www.softwareheritage.org/support/testimonials

A universal software archive, as a shared infrastructure

One infrastructure
open and shared

Cultural Heritage Industry Research Public Administration



Inria  unesco

Software Heritage

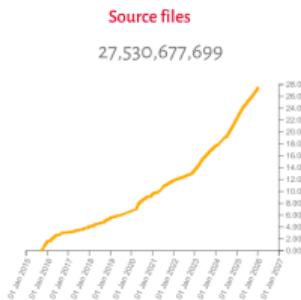
A universal software archive, as a shared infrastructure

One infrastructure
open and shared



inria unesco

The largest archive ever built



A universal software archive, as a shared infrastructure

Cultural Heritage

Industry

Research

Public Administration

One infrastructure
open and shared



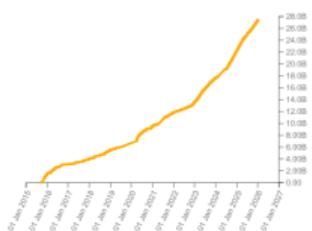
inria unesco

Software Heritage

The largest archive ever built

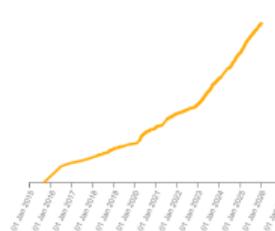
Source files

27,530,677,699



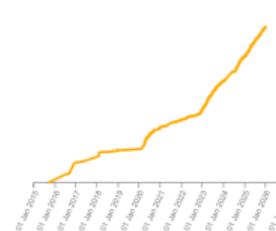
Commits ↘

5,790,477,790



Projects

422,797,163



Directories

21,551,272,560

Authors

103,062,151

Releases

135,015,027

Bitbucket

2,979,416 origins



30,710 origins



300,899,172 origins



3,973 origins



75,284 origins



664,326 origins



4,739,410 origins



72,459 origins



56,983 origins



25,242 origins



494 origins



354 origins



520,999 origins



73,325 origins



384,120 origins



198 origins



73,902 origins

git

33,507 origins



1,391 origins



2,326,857 origins



73,325 origins



384,120 origins



198 origins



73,902 origins

figures as of January 8 2026

The Software Heritage European Network: A Distributed Infrastructure for Global Source Code Preservation



Standards and policy: Software Hash Identifier (SWHID)



intrinsic (no need for a registry), cryptographically strong (integrity)

50+ billion artifacts in the Archive

Check the above SWHID [here](#)

1 Software Heritage's mission

2 A new phase: products and services for all stakeholders

3 OSPO Radar

4 Strategic Software Insights Report

Institutional portal



Strategic Insights



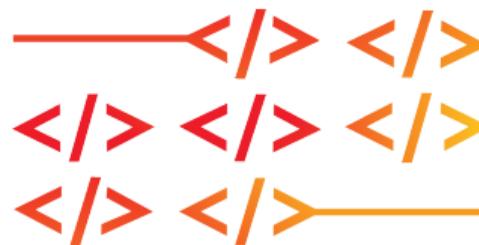
Software Heritage

Strategic Software
Insights Report

Vulnerability graph



Dataset Factory



1 Software Heritage's mission

2 A new phase: products and services for all stakeholders

3 OSPO Radar

4 Strategic Software Insights Report



Morane Gruenpeter
Director of Scholarly
Ecosystems

Can you imagine science without
Software?

A nighttime satellite view of Earth from space, showing city lights and auroras.

Photo by NASA on Unsplash.

10 years of driving a **culture shift** in Academia

To recognize **Source Code** as a first class output



Photo by NASA on Unsplash.

10 years of driving a **culture shift** in Academia To recognize **Source Code** as a first class output

We see more SWH adopters

- 26 Archive and Libraries Members
- 22 Ambassadors in Academia
- 15 Infrastructures connected and using SWH
 - ◆ 4 Depositing code
 - ◆ 2 Distributions use SWH as fallback



Photo by NASA on Unsplash.

10 years of driving a **culture shift** in Academia To recognize **Source Code** as a first class output

We see more SWH adopters

- 26 Archive and Libraries Members
- 22 Ambassadors in Academia
- 15 Infrastructures connected and using SWH
 - ◆ 4 Depositing code
 - ◆ 2 Distributions use SWH as fallback



We are actively leading working groups and networks influencing policies



Photo by NASA on Unsplash.

Preserving knowledge: Two **voices** from Academia

*"For the library, it's important to encourage software authors to **enrich the source code** description by completing a codemeta.json file and submitting the software to HAL."*



Cédric Mercier, ADAC Administrator,
Paris-Saclay University

*"Software Heritage services strongly align with the university's strategic priorities. Via its "**Springboard**" project (...), Paris-Saclay University has strategically dedicated both financial and human resources **to support open-source code**"*



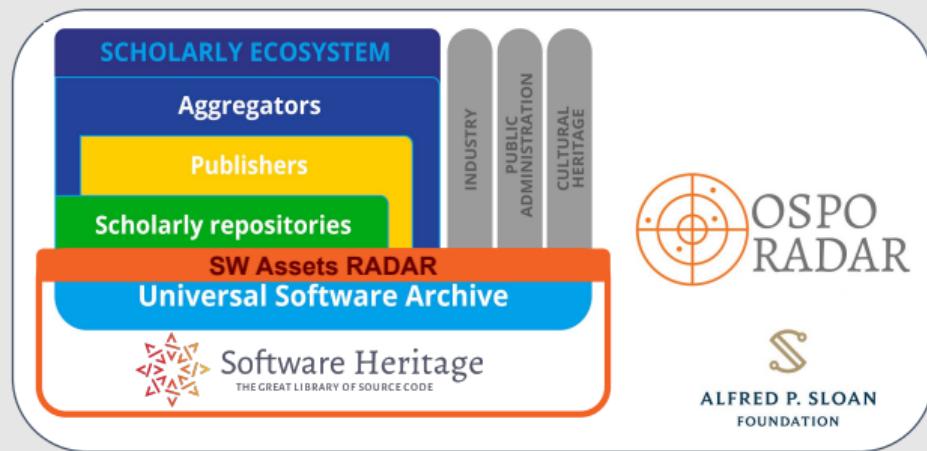
Fanny Sébire, Data librarian,
Institut Pasteur

The [full interviews series](#)

The perception is fundamentally a cultural shift, not a technical adjustment

An Assets dashboard: Addressing OSPOs' needs

- Manage institutional **software assets**;
- Enrich metadata with **annotations**;
- Support proper software **citation and attribution**;
- Track contributions, **funding**, and reuse;
- Interconnect with **existing infrastructures and services**.



"Software Heritage's OSPO-RADAR project will make visible that influence by giving research institutions a set of tools to better manage, document, and share their open-source work." **Joshua M. Greenberg**, Program Director, [Sloan Foundation](#)

Designing a product with the OSPOs Community

The MVP capabilities

- Create an **institution collection**
 - **Batch import;**
 - **Signal code repo**
- **Management dashboard:** to view/filter/curate;
- A public **front-shop** that can be injected in institutional website through an **API**.

Full specifications are public



Software Heritage as a Metadata-broker

Leveraging the CodeMeta vocabulary

- A subset of **schema.org**
- Community governed project

The CodeMeta Project

Software Heritage has the capacity to

- Ingest metadata
- Extract metadata
- Deliver consolidated views to stakeholders

Interested in joining the
OSPO-RADAR journey?

→ Join swh-ospo-radar@inria.fr

Thanks for your support making
software a **first-class research output!**



Eduardo Campagnoli Unsplash

1 Software Heritage's mission

2 A new phase: products and services for all stakeholders

3 OSPO Radar

4 Strategic Software Insights Report

Software is a strategic asset for your business, but

- ... tracking and analysing your software assets is **difficult**
- ... assessing your open source presence is **hard**
- ... knowing how you compare to other is **complex**

Software Heritage helps you gain deep insight into your software assets

A preview of the Strategic Software Insights report

Software Heritage



Report on the public
software collected,
preserved and
referenced for
Acad-Hal-Fr

Version 1.0 – 2025-12-29

Contents

1	Introduction	3
1.1	Bird's-Eye View: the key numbers	3
1.2	Technical Information	4
2	Analysis of Contributions	5
2.1	Density of contributions per project	5
2.2	Distribution of first contributions over time	7
2.3	Timespan of institutional contributions	9
2.4	Distribution of project lifetime (based on Software Heritage information)	11
2.5	Distribution of project lifetime (GitHub-based)	13
2.6	Distribution of Citation Files	15
2.6.1	Projects with Citation Files	15
2.6.2	Projects with Citation Files Only in Software Heritage	17
2.7	Top Projects by Repository Host	19
3	Analysis of the Projects to which Acad-Hal-Fr contributes	34
3.1	Distribution by number of stars	34
3.2	Main organizations	36
3.3	Top 40 projects by number of contributions	37
3.4	Top 40 projects by GitHub stars	39
3.5	Top 40 projects by repository age	41
4	Programming Languages	53
5	License Analysis	55
5.1	License Distribution	55
5.2	License Status Overview	55
5.3	License Statistics	55
5.4	Top License Types	55
6	Platform Distribution Analysis	57
6.1	Platform Distribution Overview	57
6.2	Platform Distribution (Log Scale)	57
6.3	Platform Statistics	57
6.4	Top Platforms	57
6.5	Summary	58
7	Tentative Classification in Research Areas/Domains	59
7.1	Distribution of domains of contributors (all projects)	60
7.2	Distribution of domains of contributors (selected projects with more than 10 contributions)	61
8	Methodology: From Source Graph to Analysis-Ready Dataset	61
9	Conclusion	63

A peek inside the French Academic Software report

**3531626**

Contributions

**46533**

Software Projects

Active last 12 months: 3425

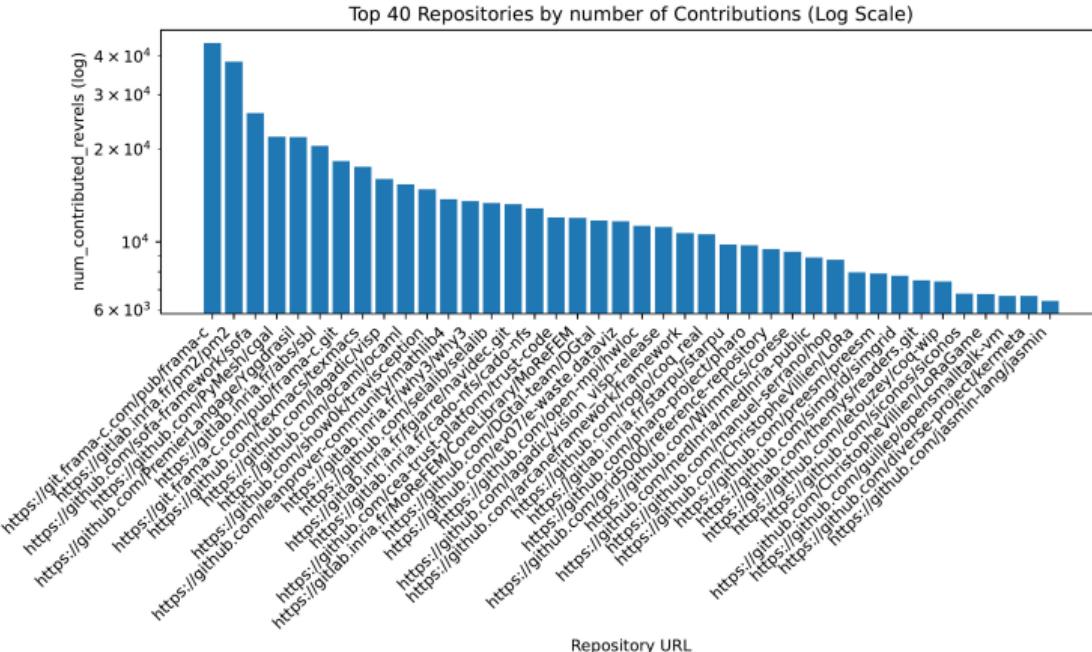
Still crawled: 44874

**30101**

Contributors

**1992-07-08
2025-05-17**

Time Span



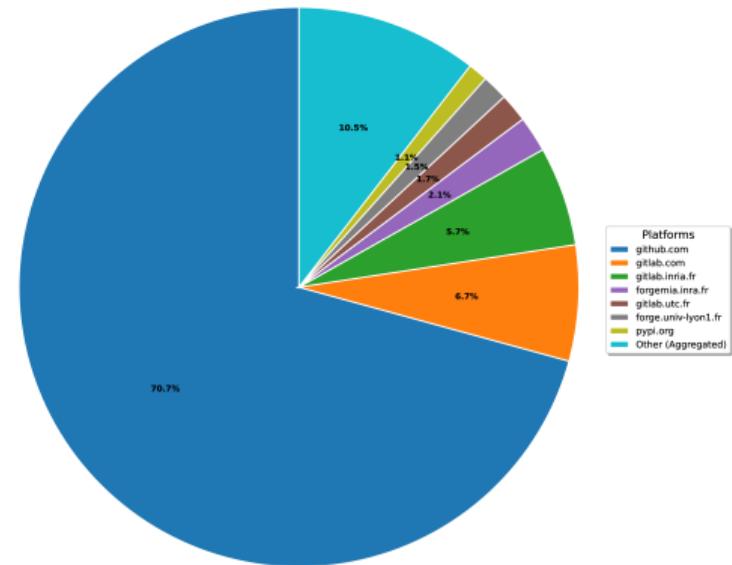
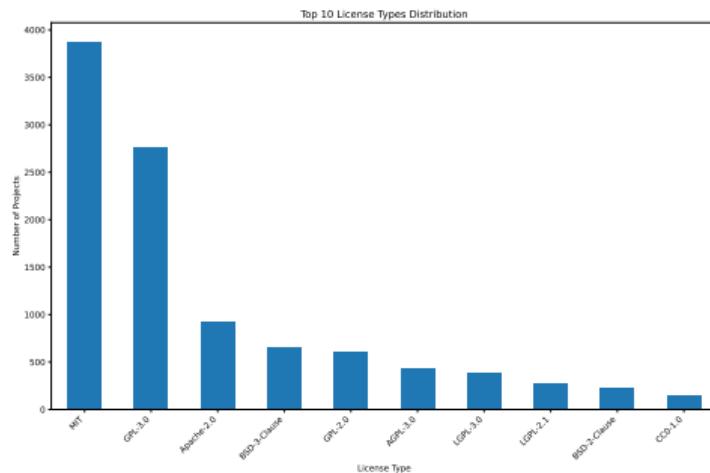
domain	count
inria.fr	3833
univ-lyon1.fr	2806
univ-orleans.fr	1555
uca.fr	1373
ucl.fr	1247
utc.fr	1216
cea.fr	1187
insa-toulouse.fr	747
ens-lyon.fr	706
univ-artois.fr	696
inrae.fr	690
univ-lemans.fr	688
cnes.fr	682
ens.fr	675
inra.fr	644
univ-tcomte.fr	634
unilim.fr	541
insa-rennes.fr	491
univ-avignon.fr	489
univ-pau.fr	483
ut2.fr	472
polytechnique.fr	437
uha.fr	422
pasteur.fr	420
telecom-paris.fr	412
univ-cotedazur.fr	393
univ-smb.fr	352
univ-tours.fr	350
u-pca.fr	349
enpc.fr	328
sciencespo.fr	327
universite-paris-saclay.fr	323
univ-poitiers.fr	312
u-paris.fr	289
univ-angers.fr	267
univ-paris8.fr	264
inserm.fr	211
centralesupelec.fr	192
ird.fr	171
imt-atlantique.fr	159
onera.fr	158
univ-tlse2.fr	131
instea.fr	130
enac.fr	126
univ-eiffel.fr	125
agroparistech.fr	105
cnam.fr	104
mnhn.fr	95
uphf.fr	94
univ-tln.fr	89

Note: Showing top 40 of 312 total domains

A peek inside the French Academic Software report, cont'd

Where is the software developed?

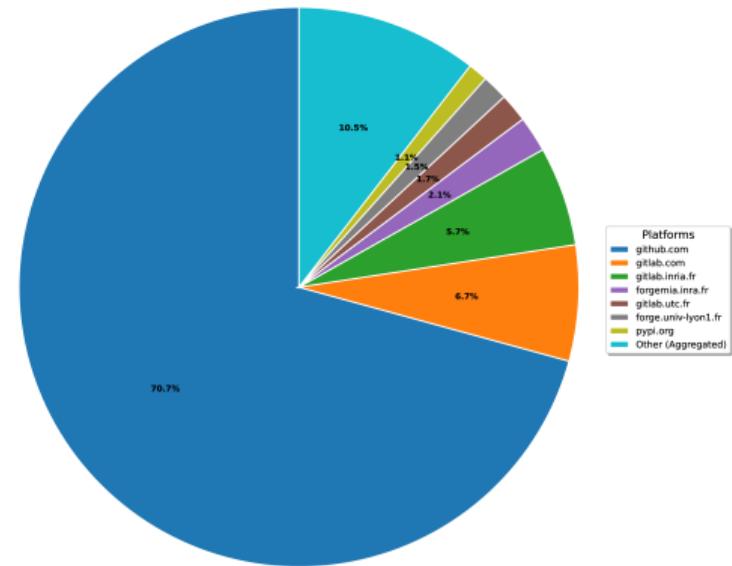
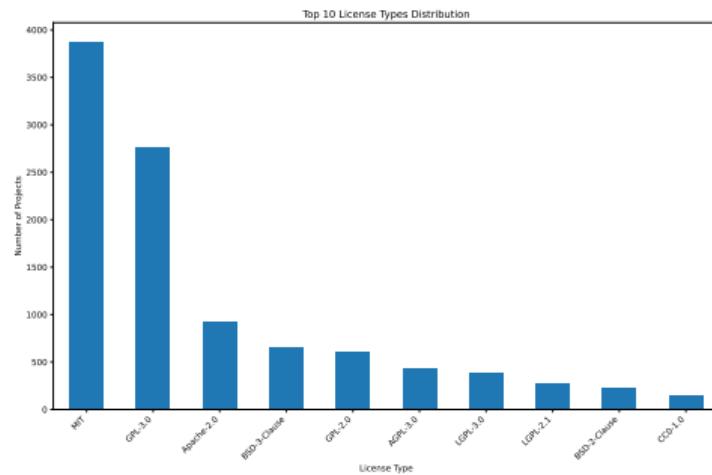
What are the top licences used?



A peek inside the French Academic Software report, cont'd

Where is the software developed?

What are the top licences used?



what is interesting? your input is needed

