



# Logiciels et codes sources en science ouverte : Du politique à l'opérationnel

## Parlons Science Ouverte #10

12.12.2024

Lucie Albaret, Université Grenoble Alpes

Jaime Arias, Université Sorbonne Paris Nord

Karim Boualem, Université Sorbonne Paris Nord

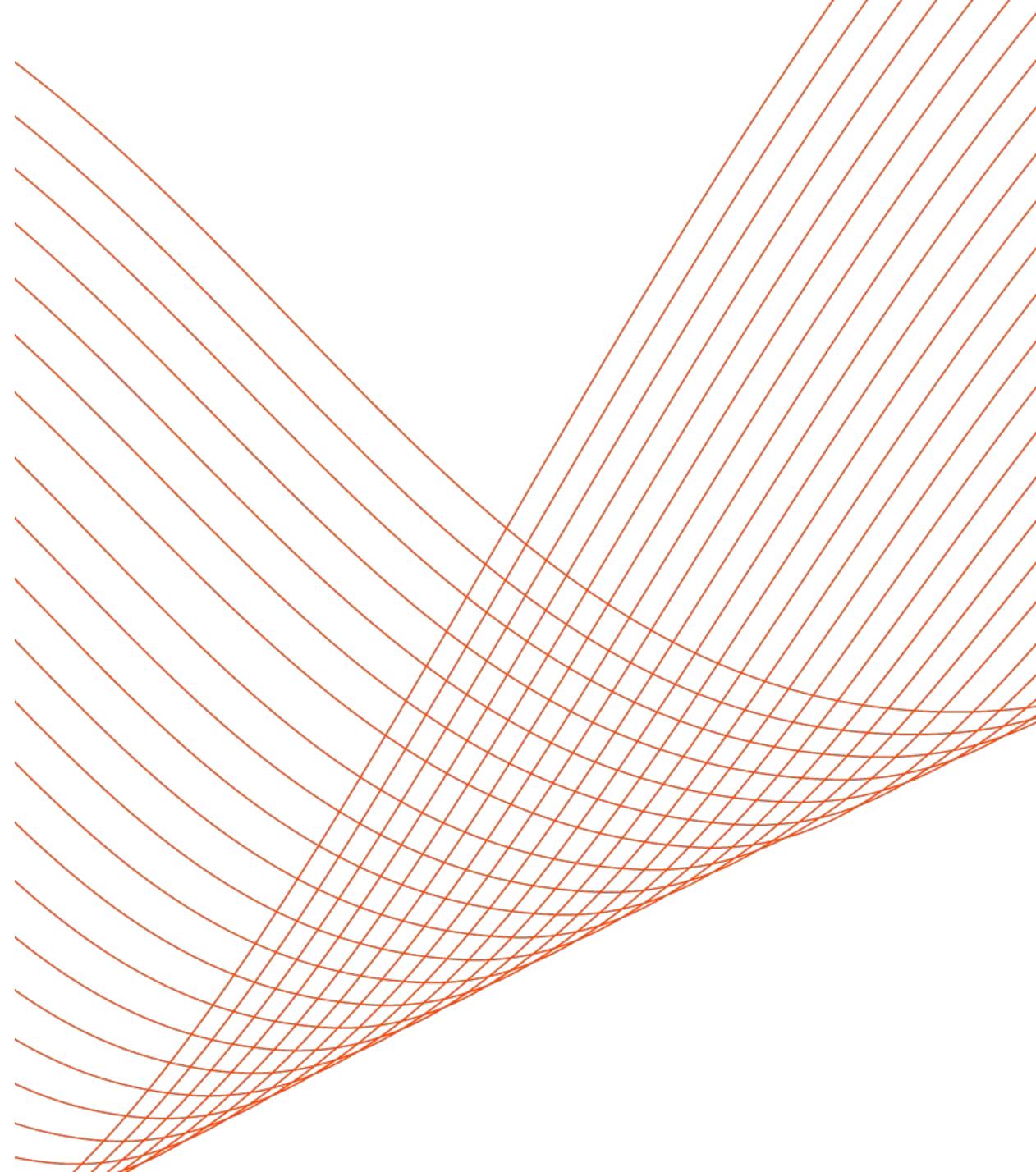
Violaine Louvet, Université Grenoble Alpes



# 01

---

## **Logiciels et codes sources dans la science ouverte : cadre national**





# Logiciels et codes sources dans la science ouverte : cadre national

Violaine Louvet

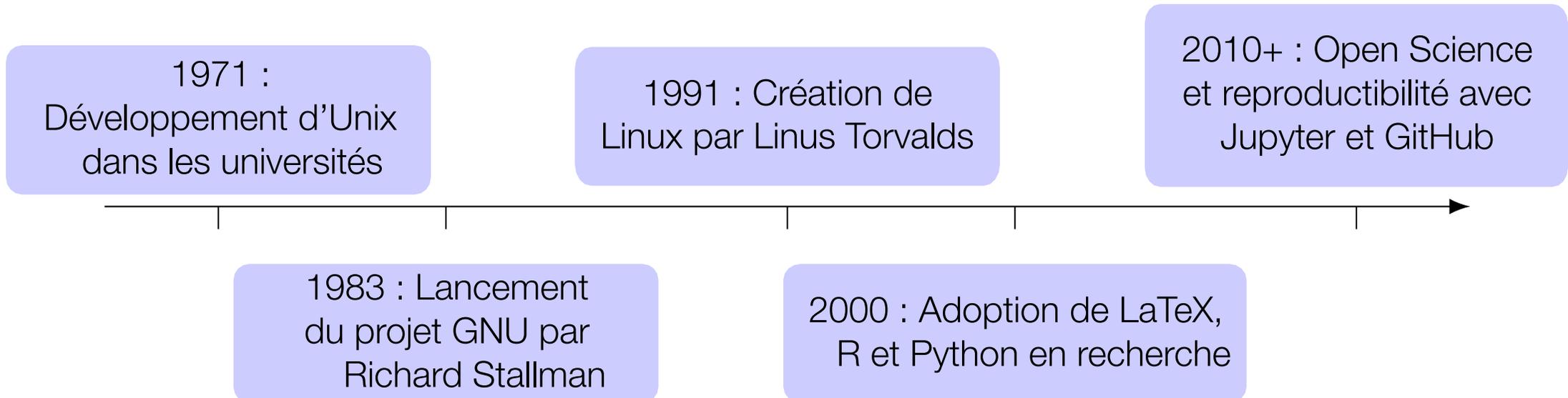
CNRS, laboratoire Jean Kuntzmann

Collège Codes Sources et Logiciels

Parlons Science Ouverte, 12 décembre 2024

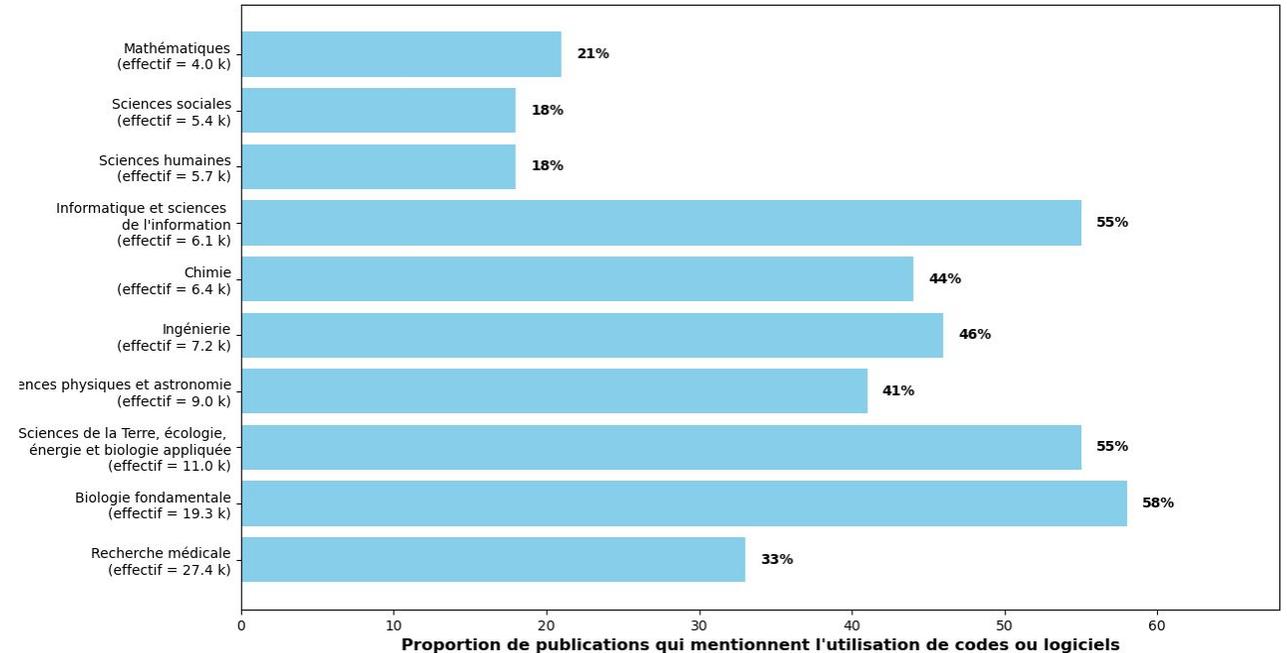
# Logiciel libre et recherche académique

- ▶ Le logiciel libre est **né dans les universités** : développement d'Unix dans les années 1970, fondation de GNU dans les années 1980 avec la création de la licence GPL ...
- ▶ Des **outils libres clés pour la recherche** comme LaTeX, R, Python, git ... sont devenus depuis les années 2000 des standards incontournables dans toutes les disciplines
- ▶ La recherche académique **s'appuie et développe massivement depuis toujours** des logiciels libres et s'inscrit historiquement dans ce contexte.



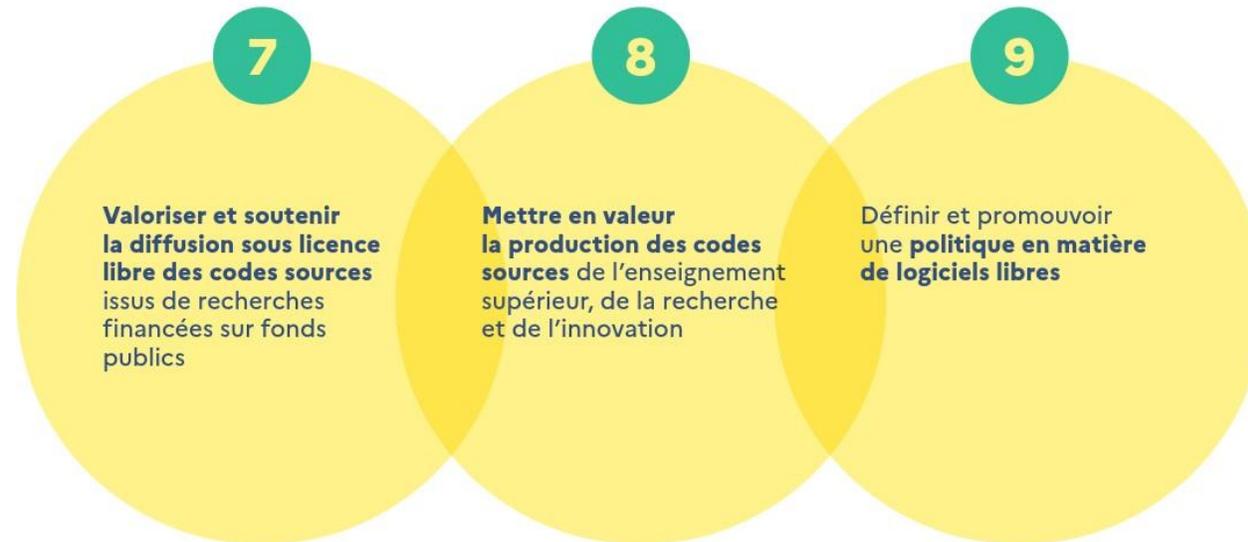
# Genèse du collège

- ▶ **2018** : Plan National Science Ouverte v1 (sans mention du logiciel)
- ▶ **2018** : création du Groupe Projet Logiciels Libres et Ouverts au sein du COSO
- ▶ **2021** : Plan National Science Ouverte v2 incluant un volet dédié au logiciel
- ▶ **2021** : premier prix science ouverte du logiciel libre de recherche
- ▶ **2022** : création du collège codes sources et logiciels
- ▶ **2023** : Intégration des codes et logiciels dans le baromètre science ouverte



# Le 3e axe du plan national science ouverte v2

- ▶ Définir et promouvoir une **politique** en matière de logiciels libres
- ▶ Reconnaître le code source comme une **contribution à la recherche** en particulier :
  - ▶ Créer un prix du logiciel libre pour la recherche
  - ▶ Construire un catalogue des logiciels issus de la recherche
- ▶ Coordonner les **communautés de pratiques** sur les codes sources et le logiciel libre en particulier :
  - ▶ Créer un Collège des codes sources et des logiciels
- ▶ Construire un **écosystème** reliant codes, données et publications



# Missions et composition du collège

- 1. Accompagner** l'essor des modes de développement logiciels et matériels libres et ouverts dans les différentes communautés scientifiques et de support
  - ▶ Contribuer à la production et à la dissémination de **méthodologies de référence et de bonnes pratiques** relatives à la production et à la gouvernance des projets, y compris en ce qui concerne leur référencement, leur pérennisation, leur valorisation et leur préservation patrimoniale
  - ▶ Proposer des **plans d'action permettant aux ministères et aux établissements** de l'enseignement supérieur et de la recherche de promouvoir les modes de développement libres et ouverts dans les actions de soutien à la recherche (appels à projets, etc.)

# Missions et composition du collège

- 1. Accompagner** l'essor des modes de développement logiciels et matériels libres et ouverts dans les différentes communautés scientifiques et de support
- 2.** Proposer des **métriques de production scientifique** pour les logiciels et matériels, tenant compte des modes de développement libres et ouverts ; dans ce cadre, étudier les mécanismes pour le référencement des logiciels et matériels, et la production d'indicateurs permettant de mesurer la pénétration des divers modes de développement au sein de la communauté scientifique

# Missions et composition du collège

- 1. Accompagner** l'essor des modes de développement logiciels et matériels libres et ouverts dans les différentes communautés scientifiques et de support
- Proposer des **métriques de production scientifique** pour les logiciels et matériels, tenant compte des modes de développement libres et ouverts ; dans ce cadre, étudier les mécanismes pour le référencement des logiciels et matériels, et la production d'indicateurs permettant de mesurer la pénétration des divers modes de développement au sein de la communauté scientifique
- Effectuer **une veille sur les initiatives relatives au développement libre et ouvert** aux niveaux national et international, en particulier dans le domaine de la recherche scientifique

# Missions et composition du collège

- 1. Accompagner** l'essor des modes de développement logiciels et matériels libres et ouverts dans les différentes communautés scientifiques et de support
- Proposer des **métriques de production scientifique** pour les logiciels et matériels, tenant compte des modes de développement libres et ouverts ; dans ce cadre, étudier les mécanismes pour le référencement des logiciels et matériels, et la production d'indicateurs permettant de mesurer la pénétration des divers modes de développement au sein de la communauté scientifique
- Effectuer **une veille sur les initiatives relatives au développement libre et ouvert** aux niveaux national et international, en particulier dans le domaine de la recherche scientifique
- Aider le Comité pour la science ouverte** à répondre aux questions issues de la communauté scientifique

# Missions et composition du collège

- 1. Accompagner** l'essor des modes de développement logiciels et matériels libres et ouverts dans les différentes communautés scientifiques et de support
- Proposer des **métriques de production scientifique** pour les logiciels et matériels, tenant compte des modes de développement libres et ouverts ; dans ce cadre, étudier les mécanismes pour le référencement des logiciels et matériels, et la production d'indicateurs permettant de mesurer la pénétration des divers modes de développement au sein de la communauté scientifique
- Effectuer **une veille sur les initiatives relatives au développement libre et ouvert** aux niveaux national et international, en particulier dans le domaine de la recherche scientifique
- Aider le Comité pour la science ouverte** à répondre aux questions issues de la communauté scientifique
- Établir des liens** avec les autres collèges, la DINUM, et l'écosystème national et international

# Composition du collège

- ▶ 41 personnes
- ▶ Des profils variés :
  - ▶ Recherche
  - ▶ Technique
  - ▶ Information scientifique et technique
  - ▶ Valorisation

## Pilotes



## Roberto Di Cosmo François Pellegrini

Professeur d'Informatique et  
Directeur de Software Heritage -  
Inria et Université Paris Cité

Professeur des universités en  
informatique et membre du  
collège de la Commission  
nationale de l'informatique et des  
libertés

## Membres

- Jaime ARIAS (CNRS, LIPN, Université Sorbonne Paris Nord)
- Thomas ARSOUZE (Cirad, AMAP)
- Céline BLITZ-FRAYRET (Cirad)
- Pierre Antoine BOUTTIER (CNRS-GRICAD)
- David CHAMONT (Université Paris-Saclay, CNRS-IN2P3, IJCLab)
- Florent CHUFFART (Université Grenoble Alpes)
- Mélanie CLÉMENT-FONTAINE (Université Paris-Saclay, Versailles Saint-Quentin)
- Laurent COSTA (UMR7041 ArScAn)
- Ludovic COURTÈS (Inria)
- Philip DAVIS (CNRS, Laboratoire de Physique Corpusculaire de Caen)
- Andrew DAVISON (CNRS, Université Paris-Saclay)
- Laurent DEMAGISTRI (IRD)
- Sébastien GÉRARD (Université Paris-Saclay, CEA, List)
- Mathieu GIRAUD (CNRS, Université de Lille) - *Co-pilote GT5*
- Timothée GIRAUD (CNRS)
- Loïc GOUARIN (CNRS, CMAP, Ecole polytechnique)
- Sabrina GRANGER (Software Heritage, Inria) - *Co-pilote GT1*
- Sonia GUERIN-HAMDI (CNRS)
- Konrad HINSEN (CNRS, UPR4301 Centre de Biophysique Moléculaire)
- Marie-Christine JACQUEMOT (Inist-CNRS)
- Jean-Yves JEANNAS (Université de Lille, AFUL) - *Co-pilote GT2*
- Nicolas JULLIEN (IMT Atlantique)
- Daniel LE BERRE (Université d'Artois, CNRS) - *Co-pilote GT2*
- Sébastien LE ROUX (CNRS, Institut de Physique et Chimie des Matériaux de Strasbourg)
- Guillaume LEVRIER (Sciences Po (CEVIPOF), Bibliothèque Nationale de France)
- Violaine LOUVET (CNRS - Laboratoire Jean Kuntzmann - Université Grenoble Alpes) - *Co-pilote GT1*
- Chiara MARMO (Université Paris Saclay) - *Co-pilote GT5*
- Camille MAUMET (Inria, Université de Rennes, CNRS, Inserm)
- Clémentine MAURICE (CNRS)
- Raphaël MONAT (Inria, Université de Lille) - *Co-pilote GT4*
- Patrick MOREAU (CNRS)
- Rémi OSSANT (Université de Tours, MSH Val-de-Loire, C-VaLo)
- Jeanne ROBINEAU (IRD)
- Nicolas ROUGIER (Inria, Université de Bordeaux, CNRS) - *Co-pilote GT4*
- François SABOT (IRD - Mission Science Ouverte)
- Kévin SCHMIDMAYER (Inria Bordeaux, Université de Pau et des Pays de l'Adour, LMAP)
- Jean-Christophe SOUPLLET (CNRS)
- Samuel THIBAUT (Université de Bordeaux)
- Anne-Laure THOMAS DEREPAIS (CNRS)
- Sylvie TONDA-GOLDSTEIN (Inria) - *Co-pilote GT3*
- Claire WAJEMAN (Université Grenoble Alpes - UMR5217 Laboratoire d'Informatique de Grenoble)

# Organisation

- ▶ **Thème 1** : Identification et mise en avant de la production logicielle de l'ESR
  - ▶ Construction d'un catalogue/référentiel national
  - ▶ [Étude d'impact de la production logicielle de l'ESR](#)
  - ▶ Métadonnées standardisées pour l'ensemble des usages (de la citation au transfert)
  - ▶ Archivage, référencement et citation du logiciel

# Organisation

- ▶ **Thème 1** : Identification et mise en avant de la production logicielle de l'ESR
- ▶ **Thème 2** : Outils et bonnes pratiques techniques et sociales
  - ▶ Forges : analyse des besoins, identification des limitations actuelles, propositions d'action
  - ▶ Modèles de gouvernance pour les logiciels de recherche
  - ▶ Bonnes pratiques pour incuber et gérer une communauté de contribution

# Organisation

- ▶ **Thème 1** : Identification et mise en avant de la production logicielle de l'ESR
- ▶ **Thème 2** : Outils et bonnes pratiques techniques et sociales
- ▶ **Thème 3** : Valorisation et durabilité
  - ▶ Aspects juridiques (licences, ayants droit, etc.)
  - ▶ Modèles économiques
  - ▶ Cycle de vie du logiciel
  - ▶ Financement du développement

# Organisation

- ▶ **Thème 1** : Identification et mise en avant de la production logicielle de l'ESR
- ▶ **Thème 2** : Outils et bonnes pratiques techniques et sociales
- ▶ **Thème 3** : Valorisation et durabilité
- ▶ **Thème 4** : Liaison et animation nationale, Europe, international
  - ▶ Proposition de charte pour le logiciel de recherche
  - ▶ Constitution d'un réseau de compétences national
  - ▶ Représentation dans des groupes de travail Europe/international

# Organisation

- ▶ **Thème 1** : Identification et mise en avant de la production logicielle de l'ESR
- ▶ **Thème 2** : Outils et bonnes pratiques techniques et sociales
- ▶ **Thème 3** : Valorisation et durabilité
- ▶ **Thème 4** : Liaison et animation nationale, Europe, international
- ▶ **Thème 5** : Reconnaissance et carrières
  - ▶ Pérennisation et organisation du prix du logiciel libre de recherche
  - ▶ Évaluation de la contribution logicielle pour la carrière

# Organisation

- ▶ **Thème 1** : Identification et mise en avant de la production logicielle de l'ESR
- ▶ **Thème 2** : Outils et bonnes pratiques techniques et sociales
- ▶ **Thème 3** : Valorisation et durabilité
- ▶ **Thème 4** : Liaison et animation nationale, Europe, international
- ▶ **Thème 5** : Reconnaissance et carrières
- ▶ **Nouveau thème 6** : Traitements algorithmiques pour la production logicielle dans la recherche
  - ▶ Usage des traitements algorithmiques dans la production de logiciels (de recherche)
  - ▶ Usage des codes sources pour la création de modèles

# Liens externes

Le collège entretient des **liens forts au niveau national et international** avec l'ensemble de l'écosystème

- ▶ Autres collèges du COSO
- ▶ RDA France
- ▶ Software Heritage
- ▶ CCSD
- ▶ Ateliers de la donnée
- ▶ EOSC, à travers certains projets, task forces, ...
- ▶ Nombreuses autres initiatives à l'international (ReSa, CURIOS, divers organismes de recherche, ...)

# Quelques éléments sur les travaux actuels du collège

- ▶ Le **catalogue des logiciels** développés dans les laboratoires de recherche
  - ▶ Un premier document [état des lieux](#) (infrastructures, catalogues existants, métadonnées, accompagnement, attentes des communautés)
  - ▶ Une identification des points essentiels :
    - ▶ Répondre aux **attentes des communautés** scientifiques ;
    - ▶ S'appuyer sur **l'existant** ;
    - ▶ S'intégrer dans **l'écosystème du chercheur** ;
    - ▶ **Automatiser** au maximum les processus
  - ▶ Des choix structuraux pour le développement d'un prototype :
    - ▶ **HAL** comme base de données principale, connu et maîtrisé par les communautés, modéré, s'appuyant sur des référentiels ;
    - ▶ **Software Heritage**, disposant de données à jour, et l'accès à des informations complémentaires ;
    - ▶ **SILL** (Socle Interministériel des Logiciels Libres), pour un portail homogène côté MESR, et qu'on peut enrichir de fonctionnalités complémentaires

# Quelques éléments sur les travaux actuels du collège

- ▶ Le **catalogue des logiciels** développés dans les laboratoires de recherche
- ▶ L'état des lieux sur les **forges logicielles** avec un [rapport produit en 2023](#) et enrichi en permanence
  - ▶ Daniel Le Berre désigné comme **réfèrent national pour les forges logicielles de l'ESR** en mai 2024
  - ▶ Création d'un **réseau d'administrateurs de forges académiques** (<https://groupes.renater.fr/sympa/info/forum-national-forges>)

# Quelques éléments sur les travaux actuels du collège

- ▶ Le **catalogue des logiciels** développés dans les laboratoires de recherche
- ▶ L'état des lieux sur les **forges logicielles** avec un [rapport produit en 2023](#) et enrichi en permanence
- ▶ Le **prix science ouverte** du logiciel libre de recherche
  - ▶ Remis depuis 2022 : <https://www.ouvrirlascience.fr/category/prix-logiciel-libre/>

# Quelques éléments sur les travaux actuels du collège

- ▶ Le **catalogue des logiciels** développés dans les laboratoires de recherche
- ▶ L'état des lieux sur les **forges logicielles** avec un [rapport produit en 2023](#) et enrichi en permanence
- ▶ Le **prix science ouverte** du logiciel libre de recherche
- ▶ Les **politiques institutionnelles**
  - ▶ [Atelier RDA France](#) co-organisé sur le sujet

# Quelques éléments sur les travaux actuels du collège

- ▶ Le **catalogue des logiciels** développés dans les laboratoires de recherche
- ▶ L'état des lieux sur les **forges logicielles** avec un [rapport produit en 2023](#) et enrichi en permanence
- ▶ Le **prix science ouverte** du logiciel libre de recherche
- ▶ Les **politiques institutionnelles**
- ▶ La **sensibilisation aux enjeux** autour du logiciel
  - ▶ Nombreuses **interventions** devant différents publics en France et à l'international

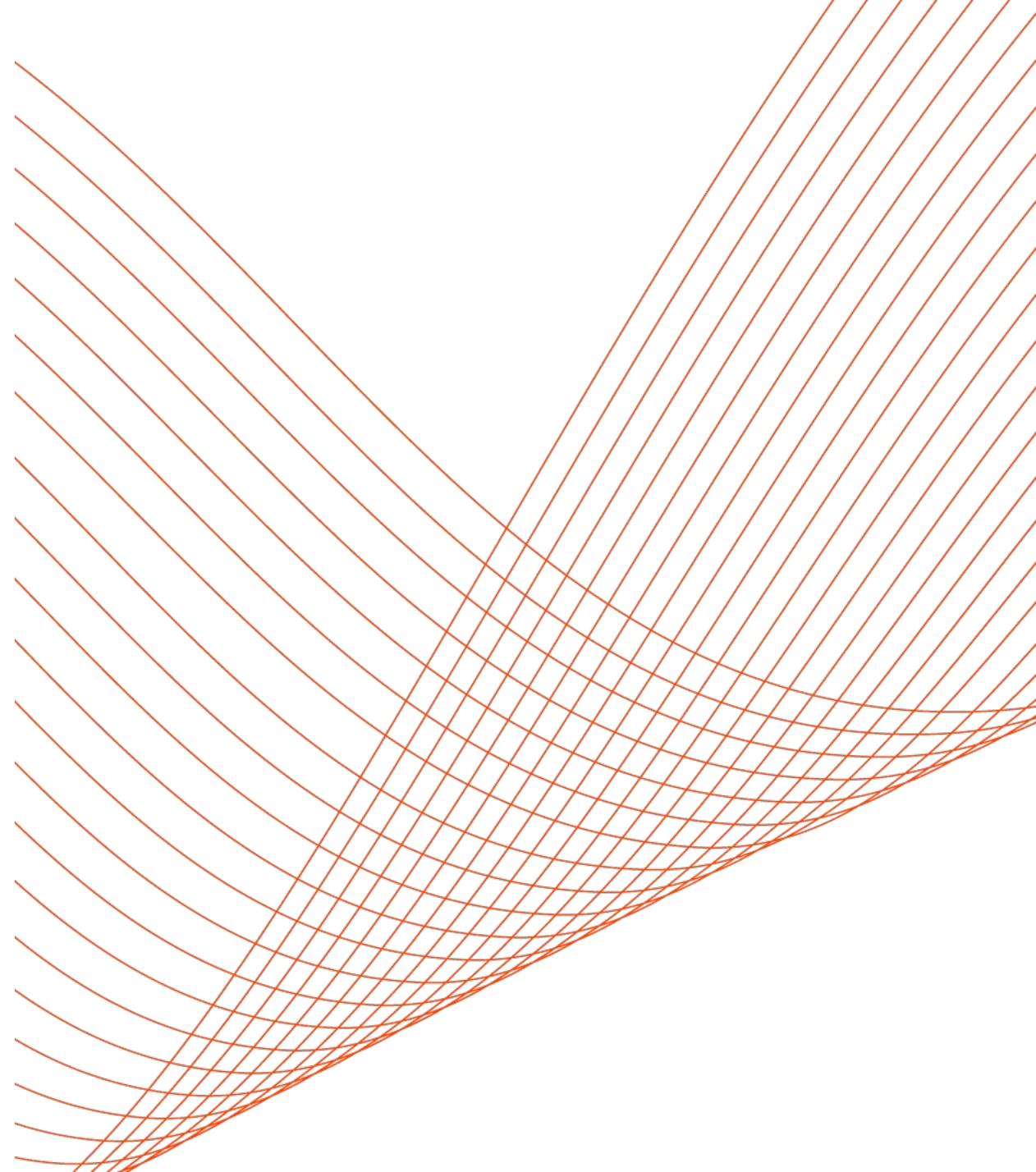
# Références

- ▶ [Plan national science ouverte 2021-2024](#)
- ▶ [Collège Codes sources et logiciels](#)
- ▶ [Productions du collège](#)
- ▶ [Article de blog sur le logiciel de recherche](#)
- ▶ [Prix science ouverte du logiciel libre de recherche 2023](#)
- ▶ [SPOS 2023, Highlights of the Software Pillar of Open Science workshop.](#)
- ▶ [Atelier RDA France sur les politique des logiciels des entités de recherche Forges de IESR Définition, usages, limitations rencontrées et analyse des besoins](#)

# 02

---

## Présentation de Software Heritage

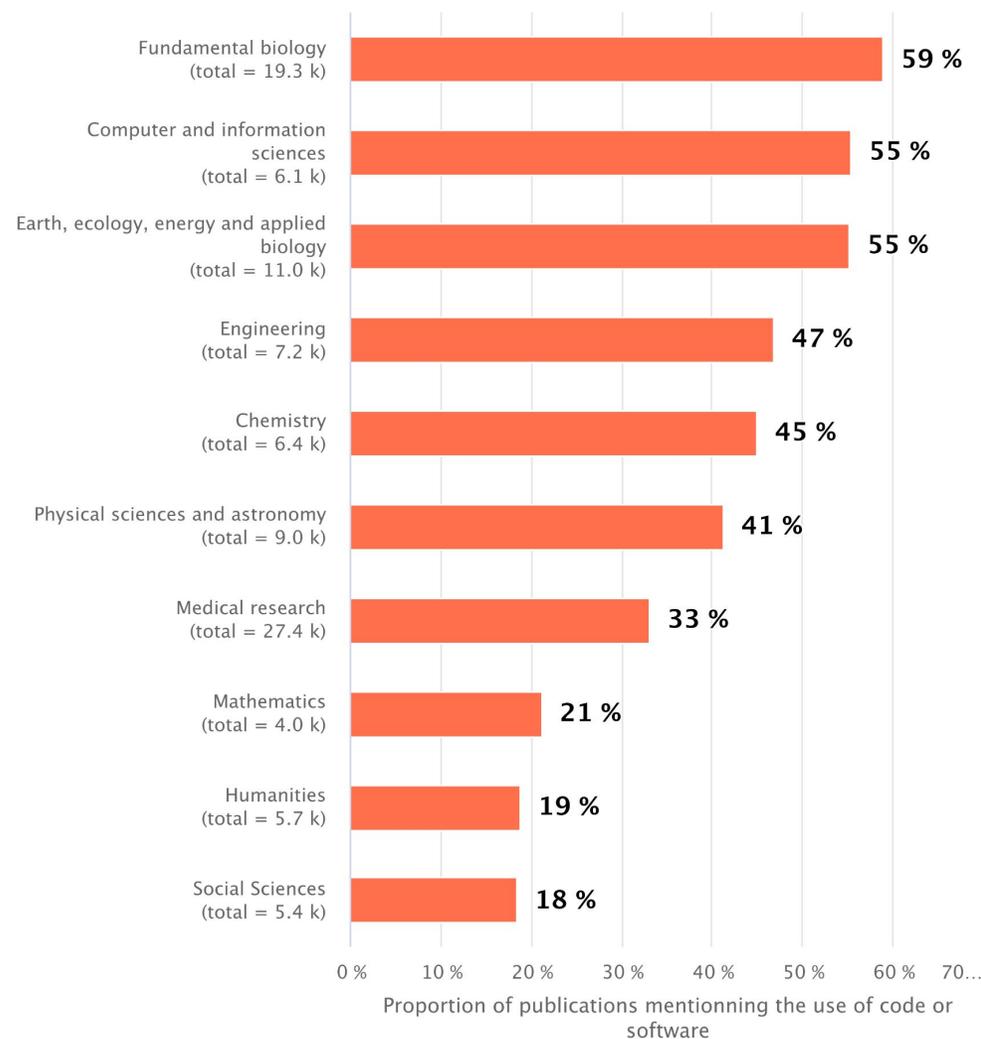




# Software Heritage

THE GREAT LIBRARY OF SOURCE CODE

### Proportion of publications in France published in 2022 that mention the use of code or software by discipline



French Open Science Monitor, Sources : MESR,

## 2024 le baromètre français de la Science Ouverte

### Sonatype Survey (2017)

80 à 90 % d'une nouvelle application logicielle est ... juste ré-utiliser !

### Art. L. 112-2 du Code de la Propriété Intellectuelle (1994)

“Sont considérés notamment comme œuvres de l'esprit au sens du présent code: . . . 13o «Les logiciels, y compris le matériel de conception préparatoire»; . . .”



ScienceDirect Journals & Books

[View PDF](#)  
Download full issue

 **Electronic Notes in Theoretical Computer Science**  
Volume 312, 24 April 2015, Pages 161-177

## A Symbolic Model for Timed Concurrent Constraint Programming

Jaime Arias , **Author 2** , **Author 3** 

Show more 

 Outline |  Share  Cite

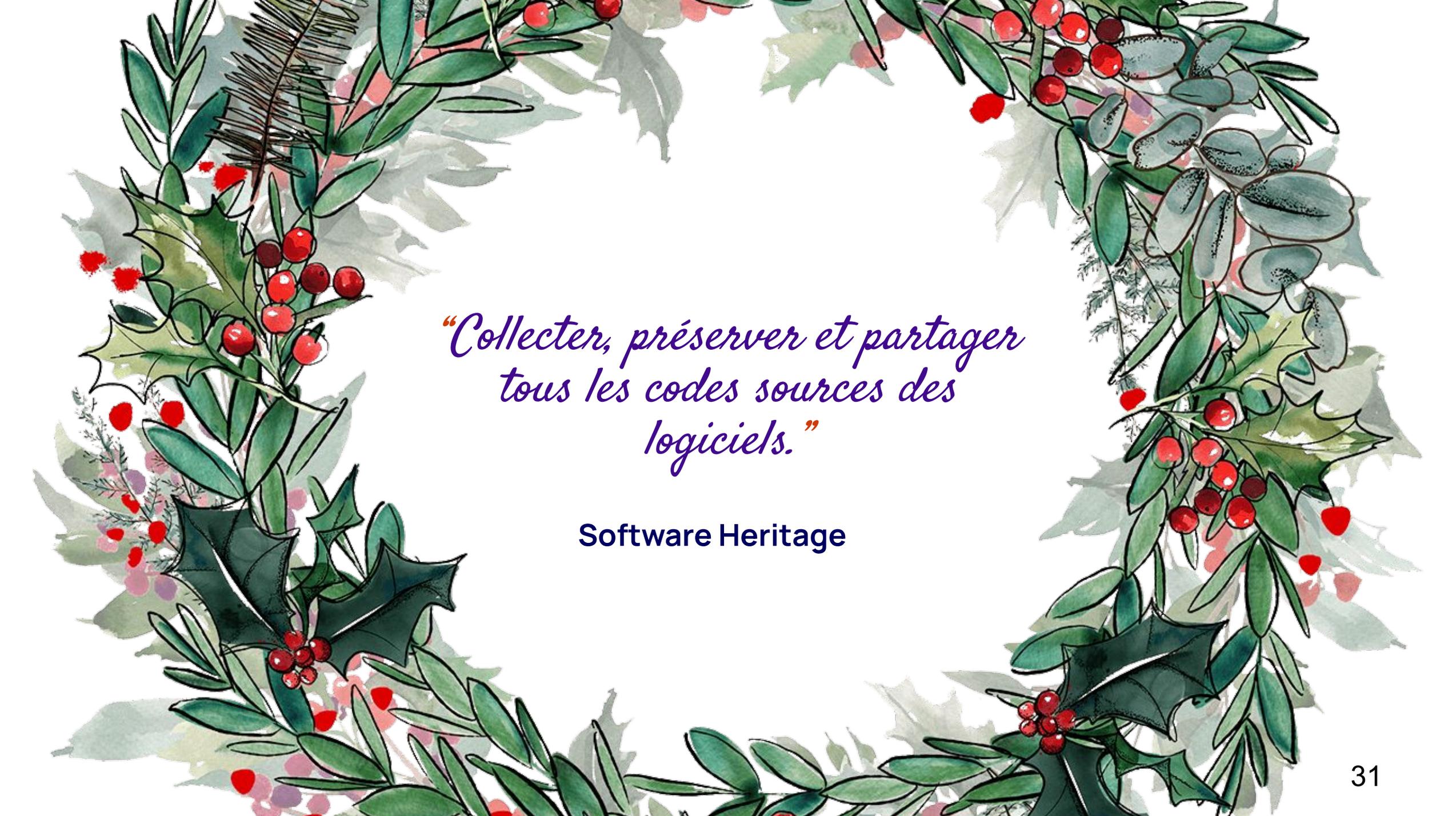
<https://doi.org/10.1016/j.entcs.2015.04.010>  

[Under a Creative Commons license](#) 



We do not describe the implementation of our tool in depth here in order to give a higher priority to the technical aspects of our approach. The reader can find the details of the implementation as well as the execution of the examples described in this paper at <http://www.labri.fr/perso/jarias/symbolicMC>.



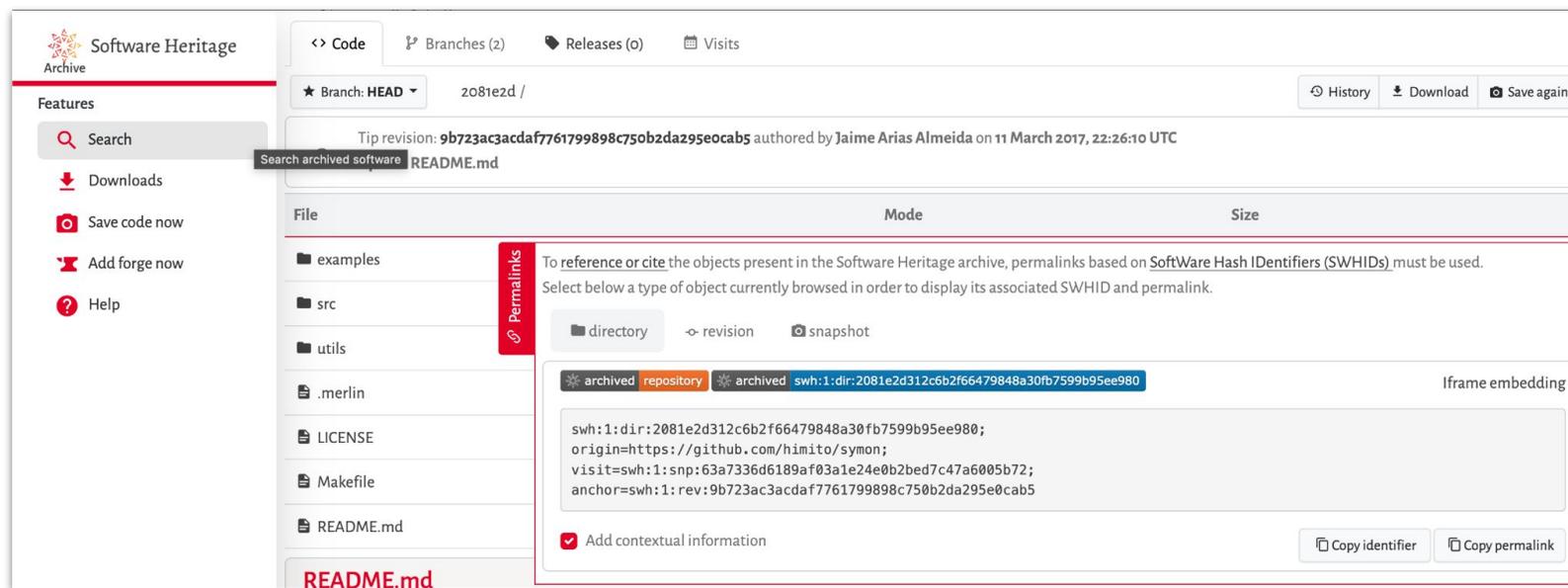


*“Collecter, préserver et partager  
tous les codes sources des  
logiciels.”*

**Software Heritage**

## Software Heritage en quelques mots

Catalogue de référence : Trouver et référencer tous les codes sources des logiciels



Software Heritage Archive

<> Code Branches (2) Releases (0) Visits

★ Branch: HEAD 2081e2d / History Download Save again

Tip revision: 9b723ac3acdaf7761799898c750b2da295e0cab5 authored by Jaime Arias Almeida on 11 March 2017, 22:26:10 UTC

Search archived software README.md

| File      | Mode | Size |
|-----------|------|------|
| examples  |      |      |
| src       |      |      |
| utils     |      |      |
| .merlin   |      |      |
| LICENSE   |      |      |
| Makefile  |      |      |
| README.md |      |      |

Permalinks

To reference or cite the objects present in the Software Heritage archive, permalinks based on SoftWare Hash Identifiers (SWHIDs) must be used. Select below a type of object currently browsed in order to display its associated SWHID and permalink.

directory ↔ revision snapshot

archived repository archived swh:1:dir:2081e2d312c6b2f66479848a30fb7599b95ee980 Iframe embedding

```
swh:1:dir:2081e2d312c6b2f66479848a30fb7599b95ee980;
origin=https://github.com/himito/symon;
visit=swh:1:snp:63a7336d6189af03a1e24e0b2bed7c47a6005b72;
anchor=swh:1:rev:9b723ac3acdaf7761799898c750b2da295e0cab5
```

Add contextual information Copy identifier Copy permalink

README.md

## Software Heritage en quelques mots

### Catalogue de référence : Trouver et référencer tous les codes sources des logiciels

International Conference on Computational Science, ICCS 2012

A “minimal disruption” skeleton experiment:  
seamless map & reduce embedding in OCaml<sup>☆</sup>

M. Danelutto<sup>a</sup>, R. Di Cosmo<sup>b</sup>

- [7] [Software excerpt] R. Di Cosmo and M. Danelutto, “Core mapping routine”, from *The Parmap library* version 1.1.1, 2020. Inria, University of Paris, and University of Pisa. VCS: <https://github.com/rdicosmo/parmap>, SWHID: `{swh:1:cnt:43a6b232768017b03da934ba22d9cc3f2726a6c5;origin=https://github.com/rdicosmo/parmap;visit=swh:1:snp:2a6c348c53eb77d458f24c9cbcecaf92e3c45615;anchor=swh:1:rel:373e2604d96de4ab1d505190b654c5c4045db773;path=/src/parmap.ml;lines=192-228}`.

```
al combine =  
  
-> tempfd () in  
  
on the chunk *)  
  in-1  
  ze-1 in  
  = compute al lo hi opid in  
  l fdarr.(i) v;  
  )  
  th "Fork error"  
  
children *)  
res-1 do ignore(Unix.wait()) done;  
22 (* read in all data *)  
23 let res = ref [] in  
24 (* accumulate the results in the right order *)  
25 for i = 0 to ncores-1 do  
26   res := ((unmarshal fdarr.((ncores-1)-i)): 'd)::!res;  
27 done;  
28 (* combine all results *)  
29 combine !res;;
```

Figure 1: Simple implementation of the distribution, fork, and recollection phases in Parmap (slightly simplified from the actual code in the version of Parmap used for this article)

## Software Heritage en quelques mots

Catalogue de référence : Trouver et référencer tous les codes sources des logiciels



HAL  
science ouverte

The screenshot shows the HAL Software Heritage interface for the project ADT2AMAS. It includes a navigation bar with 'Consulter sur Software Heritage' and 'Logiciel Année : 2023'. The main content is divided into several sections: 'Dates et versions' (hal-04152510, version 1), 'Identifiants' (HAL Id, SWHID, and various repository URLs), and 'Métadonnées' (version 2.2.0, GNU General Public License v3.0, C++ programming language, and Linux/MacOS/Windows platforms). The right side features a 'Résumé' section with a description of the tool for transforming Attack-Defense Trees (ADTrees) into Asynchronous Multi-Agent Systems (AMAS), and a 'Mots clés' section with tags for 'Attaque-Defense Trees', 'Multi-Agent Systems', and 'Security'. A 'Liste complète des métadonnées' button is also present.

Consulter sur  
Software Heritage

Logiciel Année : 2023

**ADT2AMAS: Managing agents in attack-defence scenarios**  
Jaime Arias (1, 2, 3), Wojciech Penczek (4), Laure Petrucci (2, 3), Teofil Sidoruk (4), Gianni Karlo Aguirre-Samboni

Afficher plus de détails

1 CNRS - Centre National de la Recherche Scientifique  
2 LIPN - Laboratoire d'Informatique de Paris-Nord  
3 Université Sorbonne Paris Nord  
4 PAN - Polska Akademia Nauk = Polish Academy of Sciences = Académie polonaise des sciences

Résumé en

Tool for transforming Attack-Defense Trees (ADTrees) into Asynchronous Multi-Agent Systems (AMAS).

Mots clés en

Attaque-Defense Trees Multi-Agent Systems  
Security

Domaines

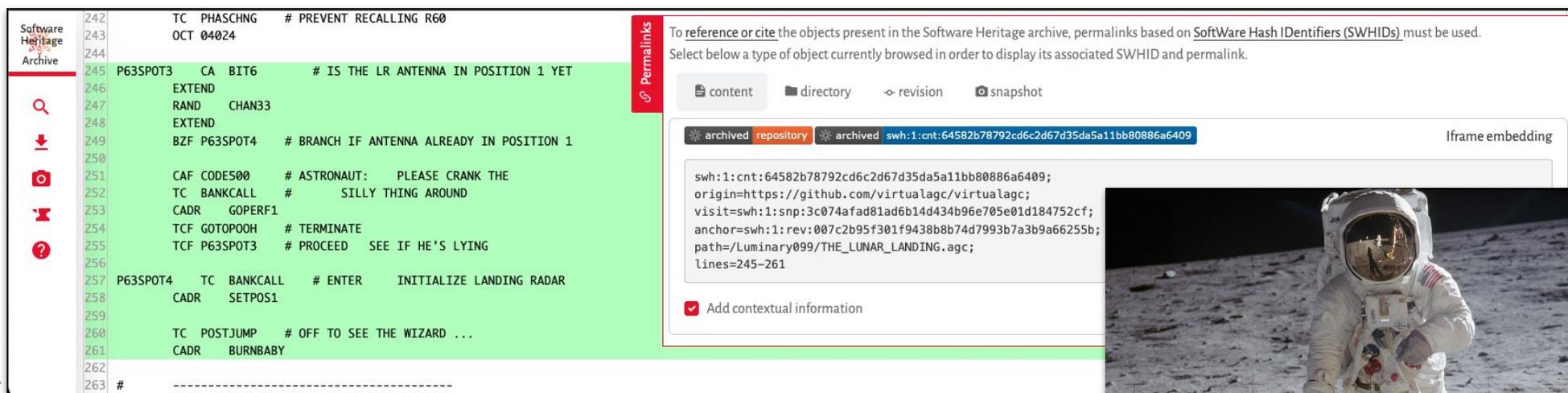
Informatique [cs]

Liste complète des métadonnées

Jaime ARIAS · Connectez-vous pour contacter le contributeur  
<https://hal.science/hal-04152510>  
Soumis le : lundi 10 juillet 2023, 14:13:35

## Software Heritage en quelques mots

Archive universelle : Conserver et partager tous les codes sources des logiciels



The screenshot displays the Software Heritage interface. On the left, a sidebar contains navigation icons: a magnifying glass for search, a download arrow, a camera for sharing, a question mark for help, and a red circle with a white question mark. The main content area shows a list of code objects with their identifiers and descriptions. A specific object is highlighted in green, showing its SWID and a permalink. To the right, a 'Permalinks' section explains that SWHIDs must be used to reference objects and provides options to view content, directory, revision, or snapshot. Below this, a text box contains the SWID and a detailed permalink for the highlighted object, along with an 'iframe embedding' option and a checkbox for 'Add contextual information'.

```
242 TC PHASCHNG # PREVENT RECALLING R60
243 OCT 04024
244
245 P63SPOT3 CA BIT6 # IS THE LR ANTENNA IN POSITION 1 YET
246 EXTEND
247 RAND CHAN33
248 EXTEND
249 BZF P63SPOT4 # BRANCH IF ANTENNA ALREADY IN POSITION 1
250
251 CAF CODE500 # ASTRONAUT: PLEASE CRANK THE
252 TC BANKCALL # SILLY THING AROUND
253 CADR GOPERF1
254 TCF GOTOPOOH # TERMINATE
255 TCF P63SPOT3 # PROCEED SEE IF HE'S LYING
256
257 P63SPOT4 TC BANKCALL # ENTER INITIALIZE LANDING RADAR
258 CADR SETPOS1
259
260 TC POSTJUMP # OFF TO SEE THE WIZARD ...
261 CADR BURNBABY
262
263 # -----
```

To reference or cite the objects present in the Software Heritage archive, permalinks based on SoftWare Hash IDentifiers (SWHIDs) must be used. Select below a type of object currently browsed in order to display its associated SWHID and permalink.

content directory revision snapshot

archived repository archived swh:1:cnt:64582b78792cd6c2d67d35da5a11bb80886a6409 iframe embedding

```
swh:1:cnt:64582b78792cd6c2d67d35da5a11bb80886a6409;
origin=https://github.com/virtuallgc/virtuallgc;
visit=swh:1:snp:3c074afad81ad6b14d434b96e705e01d184752cf;
anchor=swh:1:rev:007c2b95f301f9438b8b74d7993b7a3b9a66255b;
path=/Luminary099/THE_LUNAR_LANDING.agc;
lines=245-261
```

Add contextual information



Astronaut Edwin E. Aldrin Jr., NASA  
<https://www.nasa.gov/mission/apollo-11/>

# Software Heritage en quelques mots

## Infrastructure de recherche : Permettre l'analyse de tous les codes sources des logiciels



**Sonatype Stops Software Supply Chain Attack Aimed at the Java Developer Community**  
January 16, 2020 By Sonatype Security Research Team

**SolarWinds hack that breached gov networks poses a "grave risk" to the nation**  
Nuclear weapons agency among those breached by state-sponsored hackers.

**Equifax website hack exposes data for ~143 million US consumers**  
Breach affecting 44 pe

**DoJ says SolarWinds hackers breached its Office 365 system and read email**  
Department discovered the intrusion 9 days after SolarWinds hack came to light.

**Log4j: Google and IBM call for list of critical open source projects**  
After attending a meeting at the White House, Google also proposed creating an organization to serve as a marketplace for open source maintenance.

**Dependency Confusion: How I Hacked Into Apple, Microsoft and Dozens of Other Companies**  
The Story of a Novel Supply Chain Attack

**2017 cyberattacks on Ukraine**  
From malware to botnets

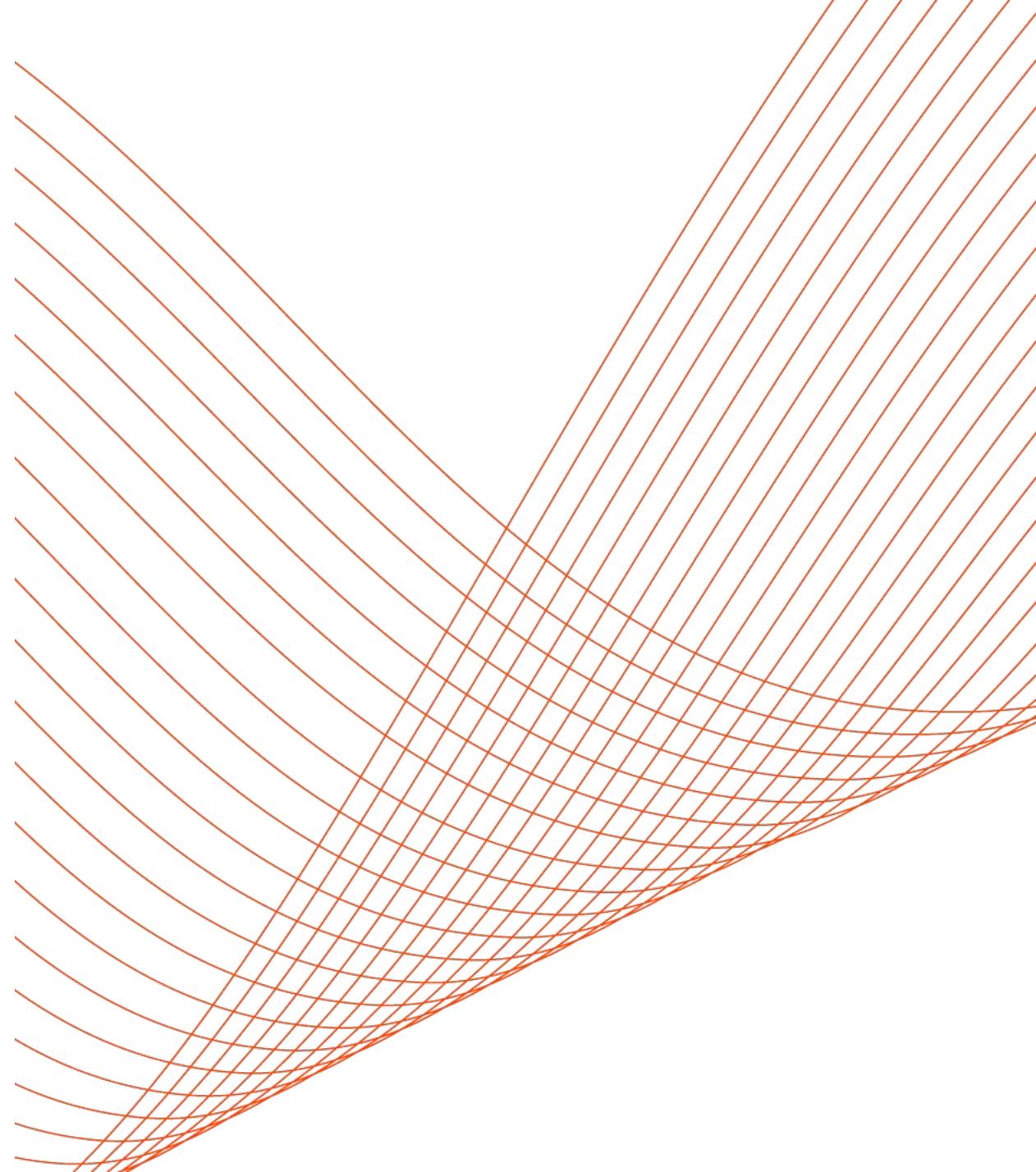
**Supply chain security guidance**  
Preparing a series of 17 principles, designed to help you establish effective control and oversight of your supply chain.

**Supply chain attack examples**

# 03

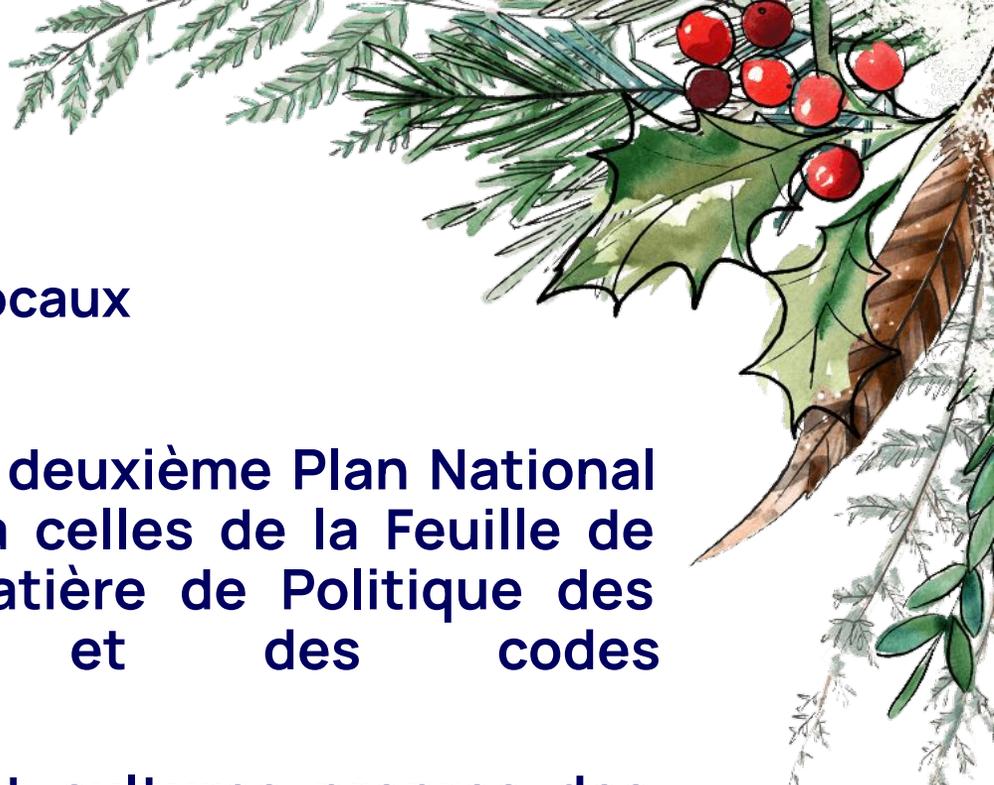
---

## Le rôle de la bibliothèque dans l'appui à la recherche





# **Le rôle de la bibliothèque dans l'appui à la recherche**



## Stratégie nationale et analyse des besoins locaux

Se conformer aux préconisations du deuxième Plan National pour la Science Ouverte (2021) et à celles de la Feuille de route (2021-2024) du MESRI en matière de Politique des données, des algorithmes et des codes

L'analyse des besoins particuliers et cultures propres des communautés, des spécificités disciplinaires et des différents contextes de recherche (Université pluridisciplinaire)



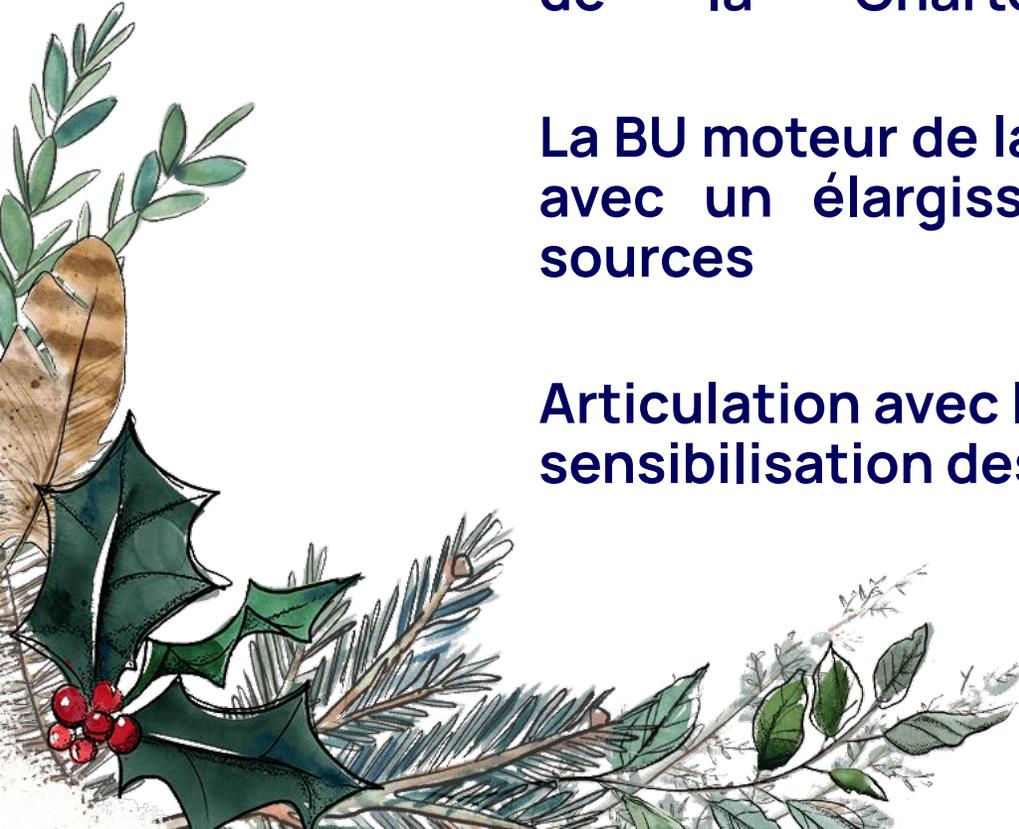


## **Le rôle de la bibliothèque dans l'appui à la recherche**

**Le positionnement central des BU dans la promotion de la Science Ouverte (HAL) est un facteur de légitimité (rédaction de la Charte Science Ouverte par la BU)**

**La BU moteur de la structuration d'un réseau Science Ouverte avec un élargissement prévu pour les logiciels et code sources**

**Articulation avec les instances politiques et décisionnelles (la sensibilisation des décideurs)**





## **Le rôle de la bibliothèque dans l'appui à la recherche**

**Formations des personnes-ressources (juridiques, techniques, etc.)**

**Le rôle de l'ambassadeur Software Héritage (mettre en place des actions de sensibilisation et de formation des doctorants et des chercheurs)**

**Les possibilités de mutualisation avec d'autres établissements**





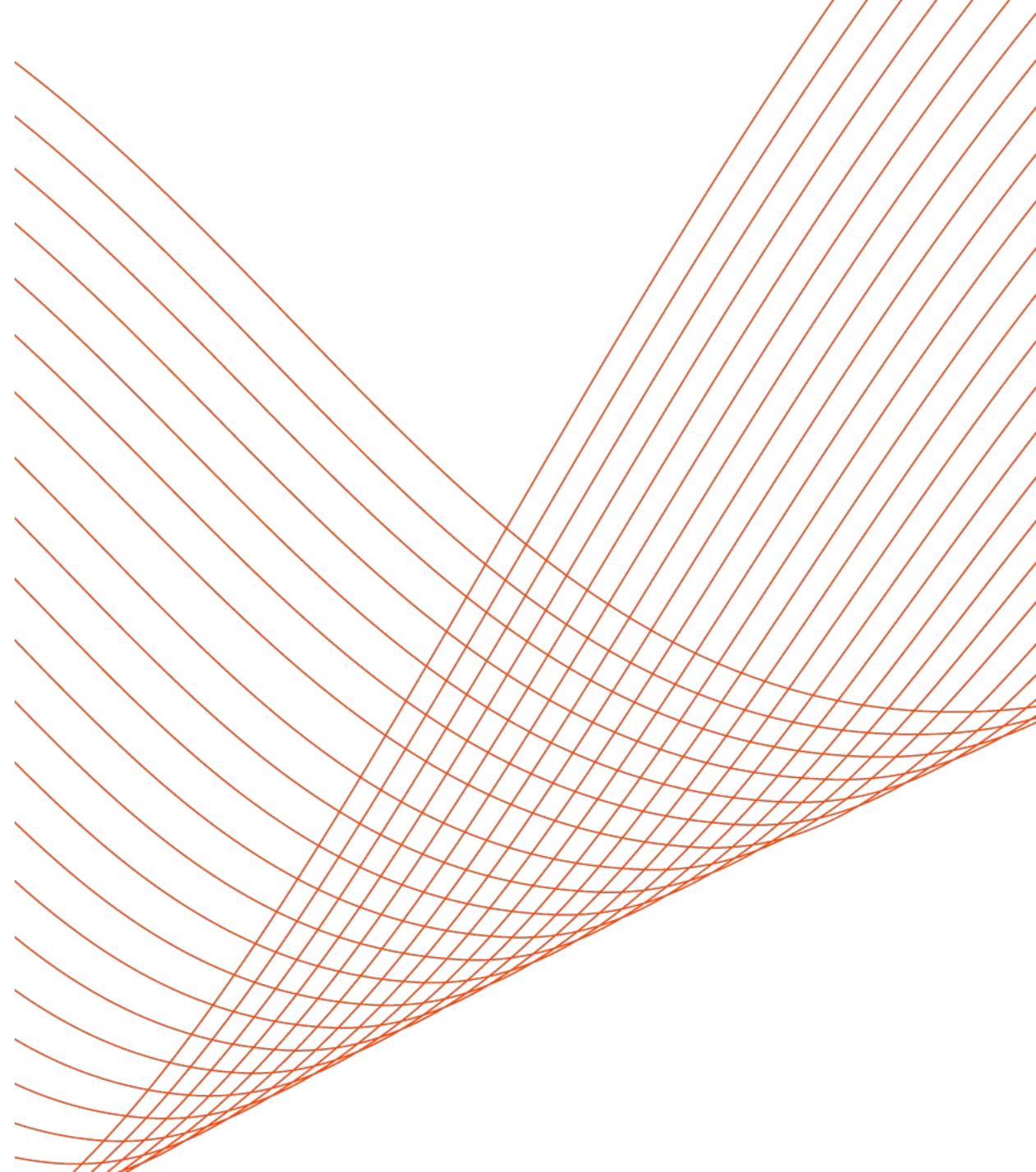
**Merci de votre attention**



# 04

---

## Les codes de recherche à l'Université Grenoble Alpes





## Les codes de recherche à l'Université Grenoble Alpes

Parlons science ouverte  
12 décembre 2024



## Contexte

- *Engagements UGA*
- *Les acteurs*

## Quelques actions concrètes

- *Former - sensibiliser*
- *Modérer*
- *Comprendre les pratiques et les besoins*
- *Participer aux actions nationales*

## Quelques perspectives

- *Impliquer*
- *Identifier les codes et logiciels à l'UGA*

## Conclusion



- Université Grenoble Alpes
  - UGA, CNRS, INRIA, CEA, INSERM, INRAE, IRD, CHUGA
  - > 90 unités de recherche
  - > 2900 doctorants
- Engagements de l'UGA sur la science ouverte
  - [Charte science ouverte](#) (oct. 2022)
  - [Schéma directeur de la science ouverte](#) (déc.2022)
- Codes : recensement, formation, développement bonnes pratiques, développement reproductibilité, utilisation de forges logicielles, utilisation de notebooks, diffusion des codes dans HAL via SWH ...



- **Les acteurs**
  - [Cellule Data Grenoble Alpes](#) (UAR Gricad - DGD BAPSO)
    - Référents données
  - [Equipe HAL UGA](#) (DGD BAPSO)
    - Correspondants HAL
  - [Plateforme Universitaire de Données Grenoble Alpes](#)
  - [UAR GRICAD](#) (Mésocentre)
  - [Gates](#)
- Atouts : - complémentarité des métiers
  - implication du mésocentre GRICAD et de l'INRIA
- Les enjeux : décroisonner !
  - Codes = le 3<sup>e</sup> pilier de la SO >>> traités avec les données
  - Faire le lien publications - données - codes



# OFFRE DE SERVICES CELLULE DATA GRENOBLE ALPES

UGA  
Université  
Grenoble Alpes



## DIFFUSER

Diffusion des données

Diffusion de codes

Archivage numérique

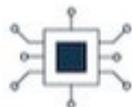
## PLANIFIER



Plan de Gestion  
de Données (PGD)

Aspects juridiques  
sur les données et les codes

Mise en place de politiques  
des données de laboratoire

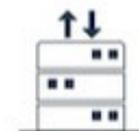


## TRANSFORMER

Conseil pour l'analyse de données,  
le calcul et l'IA

Accompagnement à la gestion de codes

## ACQUÉRIR



Recherche de données

Stockage

Cahiers de laboratoire



## TRANSVERSE

Formations

Sensibilisation aux impacts  
environnementaux

<https://scienceouverte.univ-grenoble-alpes.fr>  
[sos-data@univ-grenoble-alpes.fr](mailto:sos-data@univ-grenoble-alpes.fr)  
[sos-codes-recherche@univ-grenoble-alpes.fr](mailto:sos-codes-recherche@univ-grenoble-alpes.fr)



- **Sensibilisation- Formations**
- Formation des personnels de la bibliothèque
- [Webinaires](#) - séminaires
  - Diffusion des codes à travers Software Heritage et HAL (2 x)
  - Présentation de l'enquête nationale du MESR sur les logiciels issus des travaux de recherche
  - Atelier sur la politique logiciel des établissements RDA-France
  - Formations autour du développement scientifique (git/gitlab)
- Sensibilisation des correspondants HAL
- Interventions dans les labos
  - Séminaire des ingénieurs du centre Inria de l'Université Grenoble Alpes
  - Atelier sur le référencement logiciel aux référents données INRAE



- **Sensibilisation - Formations**
- Doctorants : « Codes et logiciels : bonnes pratiques depuis leur création pour la gestion, la diffusion, l'archivage »
  - Durée : 3h
  - [Contenu](#) : cycle de vie, forges, cadre juridique, métadonnées, diffusion,
- En projet : recherche reproductible (2025)
- Atout : soutien du collège doctoral

- [Science ouverte UGA](#)
  - Une [rubrique](#) dédiée
  - Une [fiche pratique](#)
  - Des actualités
  - Une [FAQ](#)

### J'ouvre mon code en 4 étapes



1

#### J'intègre 4 fichiers

|               |   |   |
|---------------|---|---|
| README.md     | → | Description et liens vers les documentations  |
| AUTHORS.md    | → | Les auteurs et les contributeurs du logiciel  |
| LICENSE.txt   | → | Je veux un <b>copyleft</b> = GNU GPL<br>Je veux une <b>licence permissive</b> = MIT license |
| CODEMETA.json | → | À générer automatiquement<br>avec le <b>CodeMetagenerator</b> de github                     |

2

#### Quand je code

Je **documente et commente** mon code pour moi et pour les autres  
J'utilise une **forge** comme le Gitlab de Gricad

3

#### J'archive mon code sur Software Heritage



4

#### Je signale mon code sur HAL grâce au SWHID



### Mon code est ouvert

#### Contacts

Pour plus d'informations sur la diffusion des codes, contacter la Cellule Data Université Grenoble Alpes :

[sos-codes-recherche@univ-grenoble-alpes.fr](mailto:sos-codes-recherche@univ-grenoble-alpes.fr)





## Outils et bonnes pratiques de développement

Forges logicielles

Notebooks

Documentation

## Questions juridiques

Licences

Logiciels libres

Dépôt APP

En pratique

## Reproductibilité

## Référencement, diffusion et archivage

SWH / HAL

Métadonnées

Référencement à l'UGA

## Contexte national et international



- [Dataacc'](#)
  - Plateforme d'aide à la gestion des données en physique et en chimie
  - Mise en œuvre par le SCD Lyon 1 et la BAPSO UGA (Financement Collex-Persée)
  - Une nouvelle rubrique sur les codes et logiciels

BONNES PRATIQUES

## Ouvrir ses codes et logiciels

— Accueil • Bonnes pratiques • Ouvrir ses codes et logiciels

### Licences et propriété intellectuelle

Propriété intellectuelle applicable aux codes et logiciels  
Licences d'usage des logiciels  
Focus sur les licences libres  
Exemple de la licence [...]

### Bonnes pratiques de développement

Introduction Cycle de vie du logiciel Plan de gestion des logiciels Forges logicielles  
Documentation Quelques exemples d'outils en physique et [...]

### Valoriser ses codes et logiciels

Introduction Définition Bonnes pratiques  
Archivage et diffusion des codes via Software Heritage et HAL  
Comment faire? Quelques ressources et tutoriels [...]



- **Modération HAL**

- Une équipe de 10 personnes formées
- Des points d'attention :
  - Licences
  - Readme
  - Auteurs

**Demande de modification** - Bonjour, Après avoir examiné votre dépôt, nous avons constaté qu'il manquait dans votre archive quelques fichiers pour pouvoir effectuer la publication. ( Pour en savoir plus, consulter la documentation <https://doc.archives-ouvertes.fr/deposer/deposer-le-code-source>) - le fichier README qui décrit le logiciel que vous déposez. Pour vous aider à le constituer, vous pouvez consulter par exemple le site suivant : <https://www.makeareadme.com/> . - le fichier AUTHORS qui décrit les auteurs du logiciel. Vérifier que les mêmes auteurs sont renseignés dans les métadonnées, avec leurs affiliations. -le fichier LICENCE qui décrit les droits d'usage du logiciel. Vérifier que cette même licence est renseignée dans les métadonnées, plutôt que "Other (open)". Nous vous remercions de re-déposer une archive zip avec ces fichiers. Pour plus de renseignements, vous pouvez nous contacter à cette adresse : [sos-codes-recherche@univ-grenoble-alpes.fr](mailto:sos-codes-recherche@univ-grenoble-alpes.fr)



- **Modération Recherche Data Gouv**
  - Champ : « Publications associées »
  - Champ « Workflow de calcul »
- **Liens publications données codes**
  - Un exemple :
    - Sur Recherche Data Gouv
    - Fernandez, Simon, 2024, "Replication Data for: WHOIS Right? An Analisis of WHOIS and RDAP Consistency", <https://doi.org/10.57745/RJX9XH>, Recherche Data Gouv, V1,
    - Sur HAL
  - Simon Fernandez. WHOIS Right? - Dataset Analysis. 2023, <https://github.com/drakkar-lig/whois-right-dataset;visit=swh:1:snp:47ce58d56c4f4339670f652de2c22db15266adff;anchor=swh:1:rev:d44ca350efd7f06f821600cbe9215aee4331f0a4>. [\(hal-04418015\)](#)



- **Adresses support** : [sos-codes-recherche@univ-grenoble-alpes.fr](mailto:sos-codes-recherche@univ-grenoble-alpes.fr) et [sos-data@univ-grenoble-alpes.fr](mailto:sos-data@univ-grenoble-alpes.fr)
  - Questions juridiques
  - Développement des codes (forges)
  - Diffusion des codes
- **Mise à disposition de services**
  - Forge UGA : service d'informatique scientifique le plus utilisé
  - [Jupyter Notebooks](#)



- **Implication des référents données**
  - Sept 2024 : échange sur les pratiques autour des logiciels
  - Des pratiques hétérogènes :
    - Informatique, physique, sciences de l'environnement, linguistique, littérature
  - Des besoins :
    - Sensibilisation aux enjeux d'ouverture
    - Formations aux bonnes pratiques (PGL, gestion de versions, ...)
      - Rôle du collège doctoral à renforcer
    - Aide au choix des outils pertinents
    - Proposer des services autour de l'IA
    - Aide juridique (licences, recommandations)



- Focus [Gates](#)
  - 5 ingénieurs pour la gestion des données de recherche en SHS
    - Implique la création de codes
    - Compétences SHS mais aussi langages de programmation (R, Python), bases de données, visualisation, dashboard, site web ...
  - Quelques exemples :
    - [Translatoscope](#)
    - [Jade](#)
  - Un [blog](#)



- Un outil de suivi : l'[Open Research Software Monitor](#) (juillet 2024)  
Créateurs : Maxence Larrieu, Joseph Beau-Reder, Lilas Lacoste, Elias Chetouane  
  
Mesure : interrogation hebdomadaire de la forge UGA (gricad-gitlab)  
8000 logiciels  
2400 disponibles publiquement.
- Conçu à la suite de l'[Open Research Data Monitor](#)  
Créateurs Maxence Larrieu, Elias Chetouane, Louis Maritaud





## Do repositories contain a `codemeta.json` ?

---

We currently count 5 repositories that contains a `codemeta.json` file.

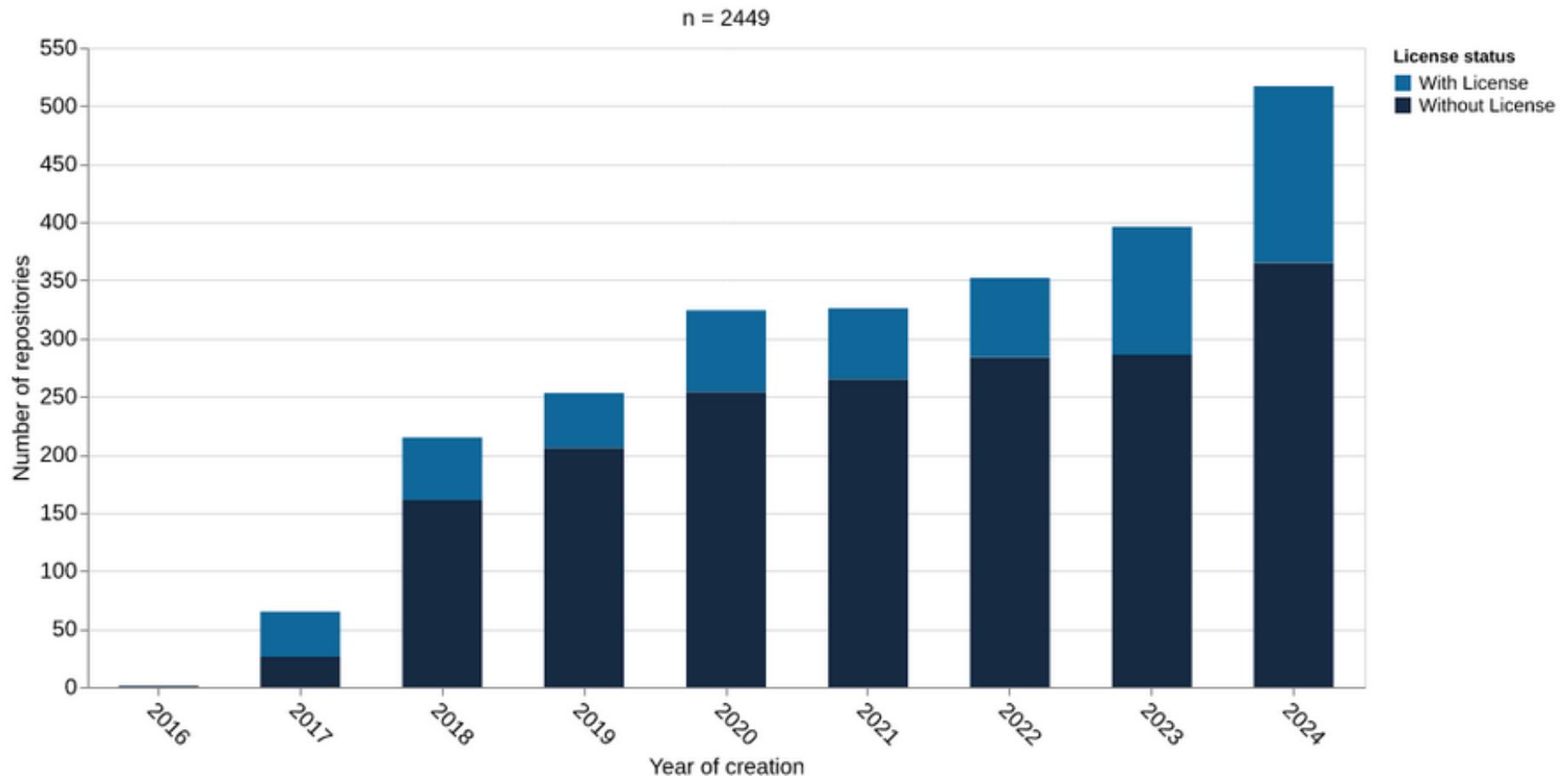
## Are repositories referenced in HAL ?

---

We currently count 9 repositories that are referenced in HAL.



## How many public repositories by year ?





- **Participations aux actions nationales**
- **Collège Codes et logiciels du Comité Science Ouverte**
  - Participation à la conception du catalogue, basé sur HAL
- **GT5 des ateliers de la donnée**
  - Plan de gestion des logiciels (PGL)
  - Description des logiciels
  - Forges et logiciels
  - Mapping entre le PGL et codemeta.json
  - Bonnes pratiques en faveur de la répliquabilité



- Renforcer la coopération entre CDGA et équipe HAL
- Renforcer la coopération avec les référents données
- Impliquer d'autres acteurs
  - Les services juridiques (licences, déclarations APP)
  - Le service des thèses (BAPSO), le collège doctoral et les ED
- Communiquer (chercheurs, doctorants, directeurs de thèse, DU, équipes d'accompagnement)
- Renforcer la formation (doctorants, chercheurs, services juridiques ...)
  - Identifier les besoins
- Mettre en place des indicateurs spécifiques
  - Besoin de leviers (HCERES, ANR ?)



- Identifier les codes pour le dépôt dans HAL
  - Questions :
    - Quels codes ?
    - Quelle méthode ?
  - Etapes :
    - Arriver à une preuve de concept
    - Moissonner la forge pour aller vers l'automatisation de la création des notices HAL ?
    - Prendre en compte la mention des codes dans les thèses ?
- Systématiser les liens publications/données/codes/thèses
  - Rôle du projet HALiance !
  - Evolution de Recherche Data Gouv (catalogue)



- Rôle des personnels de documentation ?
  - Expérience sur les publications et les données
  - Périmètre d'action :
    - Formation – sensibilisation
    - Animation des réseaux
    - Aide à la description (code meta !)
    - Modération
    - Liens publications – données – codes
  - Assurer le décroisement :
    - Actions en binôme
    - Collaboration indispensable mésocentre/bibliothèques !
  - Enjeu : formation interne et appropriation par les équipes
- Aboutir à un cercle vertueux



Merci pour votre attention !

Des questions ?

