El Software en la investigación

Les logiciels dans la recherche

Roberto Di Cosmo

Director, Software Heritage
Inria y Université Paris Cité

29 de octubre de 2026
IV Encuentro Francia - América Latina sobre Ciencia Abierta
Montevideo, Uruguay



Software Heritage

THE GREAT LIBRARY OF SOURCE CODE

Contenido

1. Apertura / Ouverture
2. El código fuente, memoria de la ciencia / Le code source, mémoire de la science
3. Encuentro con Software Heritage / Voici Software Heritage
4. Archivar, Referenciar, Describir, Citar / Archiver, Referencer, Décrire, Citer
5. Describir su software / Décrire son logiciel
6. Herramientas prácticas / Des outils pratiques
7. Las exigencias van subiendo / Des obligations qui montent
8. El software de la región / Le logiciel de la région
9. En la política de Ciencia Abierta / Dans la politique de Science Ouverte
10. Lo que queda abierto: cita mañana / Ce qui reste ouvert : rendez-vous demain
11. Conclusión / Conclusion

Profesor de informática en París, hoy en Inria

- *más de 35 años* de investigación (informática teórica, programación, ingeniería del software)
- *más de 25 años* de software libre
- *más de 15 años* creando y dirigiendo estructuras al servicio del bien común



1999 *DemoLinux*, la primera distribución GNU/Linux *live*

2007 *Grupo temático Software Libre* (polo Systematic, París)

2010 *IRILL*

2015 *Software Heritage*, en Inria

2018 *Comité Nacional para la Ciencia Abierta*, Francia

2021 *Task Force EOSC sobre infraestructuras para el software*, Unión Europea



COMMENT | 07 October 2025

Stop treating code like an afterthought: record, share and value it

Scientists, research institutions, funders, libraries and publishers must all improve software practices.

By Roberto Di Cosmo, Sabrina Granger, Konrad Hinsen, Nicolas Jullien, Daniel Le Berre, Violaine Louvet, Camille Maumet, Clémentine Maurice, Raphaël Monat & Nicolas P. Rougier

Illustration: Phil Wheeler

Un llamado de *Nature*

Comentario en *Nature*, octubre de 2025

doi:10.1038/d41586-025-03196-0

Esta conferencia es:

- la motivación profunda para dar al software un papel central
- una perspectiva del punto de vista de politica de ciencia (abierta)
- el *cómo* actuar en practica para :
  - archivar, referenciar, describir y citar el software de investigación, concretamente y desde hoy
  - adaptar la politicas científicas para agregar este pilár software

El código fuente es conocimiento

Le code source est connaissance

Harold Abelson, Structure and Interpretation of Computer Programs (1st ed.) 1985

"Programs must be written for people to read, and only incidentally for machines to execute."

Apollo 11 source code (excerpt)

```
P63SPOT3        CA      BIT6            # IS THE LR ANTENNA IN POSITION 1 YET
                EXTEND
                RAND    CHAN33
                EXTEND
                BZF     P63SPOT4        # BRANCH IF ANTENNA ALREADY IN POSITION 1

                CAF     CODE500         # ASTRONAUT:    PLEASE CRANK THE
                TC      BANKCALL        #               SILLY THING AROUND
                CADR    GOPERF1
                TCF     GOTOPOOH        # TERMINATE
                TCF     P63SPOT3        # PROCEED       SEE IF HE'S LYING

P63SPOT4        TC      BANKCALL        # ENTER         INITIALIZE LANDING RADAR
                CADR    SETPOS1

                TC      POSTJUMP        # OFF TO SEE THE WIZARD ...
                CADR    BURNBABY
```

Covid Sim (excerpt)

```
/**
 * @brief The basic unit of the simulation and is associated to a geographical location.
 *
 * Interventions (e.g., school closures) are tracked at this level. It contains a list of its
 * members (people), places (schools, universities, workplaces etc.), road networks, links to
 * airports etc.
 */
struct Microcell
{
        /* Note use of short int here limits max run time to USHRT_MAX*ModelTimeStep - e.g. 65536*0.25=16384 days=44 yrs.
           Global search and replace of 'unsigned short int' with 'int' would remove this limit, but use more memory.
        */

        int n; // Number of people in microcell
        int adunit; // admin unit microcell belongs to
        int* members; // array of members/hosts of microcell

        int* places[MAX_NUM_PLACE_TYPES]; // list of places (of various place types) within microcell
        unsigned short int NumPlacesByType[MAX_NUM_PLACE_TYPES]; // number of places (of various place types) within mircocell
        unsigned short int keyworkerproph, move_trig, place_trig, socdist_trig, keyworkerproph_trig;
        unsigned short int move_start_time, move_end_time;
        unsigned short int place_end_time, socdist_end_time, keyworkerproph_end_time;
        TreatStat moverest, treat, vacc, socdist, placeclose;
        unsigned short int treat_trig, vacc_trig;
        unsigned short int treat_start_time, treat_end_time;
        unsigned short int vacc_start_time;
        IndexList* Airportlist;
};
```

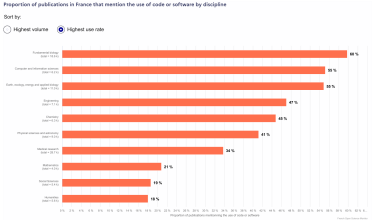
Len Shustek, Computer History Museum 2006

"Source code provides a view into the mind of the designer."

Contenido

1. Apertura / Ouverture
2. El código fuente, memoria de la ciencia / Le code source, mémoire de la science
3. Encuentro con Software Heritage / Voici Software Heritage
4. Archivar, Referenciar, Describir, Citar / Archiver, Referencer, Décrire, Citer
5. Describir su software / Décrire son logiciel
6. Herramientas prácticas / Des outils pratiques
7. Las exigencias van subiendo / Des obligations qui montent
8. El software de la región / Le logiciel de la région
9. En la política de Ciencia Abierta / Dans la politique de Science Ouverte
10. Lo que queda abierto: cita mañana / Ce qui reste ouvert : rendez-vous demain
11. Conclusión / Conclusion

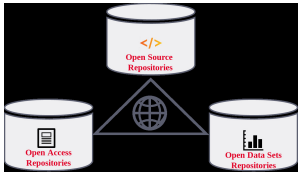
El software sostiene la investigación moderna



Más del 20 % de los artículos que usan software, en todas las disciplinas, lo comparten

Barómetro francés de la ciencia abierta, 2024

Tres pilares



El software libre en la Recomendación de la UNESCO sobre la Ciencia Abierta (2021), §7 d)

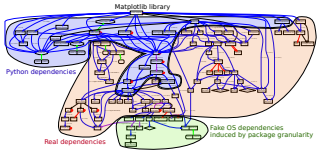
« programas informáticos cuyo código fuente está disponible públicamente, de manera oportuna, en un formato fácil de utilizar, legible y modificable por personas y máquinas, mediante una licencia abierta [...] »

El software *evoluciona*

- proyectos que duran décadas
- la *historia del desarrollo* es clave para *entenderlo*

Complejidad

- *millones* de líneas de código
- una gran *red de dependencias*
  - fácil de romper, difícil de mantener
  - el software de investigación es solo la capa superior
- *comunidades de desarrollo* sofisticadas



El lado humano

diseño, algoritmo, código, pruebas, documentación, comunidad, financiación

los principios FAIR, pensados para los datos, no bastan

de menor a mayor dificultad

Archivar

El software de investigación debe estar archivado

para poder *recuperarlo* (*reproducibilidad*)

Referenciar

El software de investigación debe estar referenciado

para poder *identificarlo* (*reproducibilidad*)

Describir

El software de investigación debe estar descrito

para poder *descubrirlo* y *reutilizarlo* (*visibilidad*)

Citar y dar crédito

El software de investigación debe ser citado (*no es lo mismo que referenciarlo*)

para dar *crédito* a sus autores (*evaluación*)

Contenido

1. Apertura / Ouverture
2. El código fuente, memoria de la ciencia / Le code source, mémoire de la science
3. Encuentro con Software Heritage / Voici Software Heritage
4. Archivar, Referenciar, Describir, Citar / Archiver, Referencer, Décrire, Citer
5. Describir su software / Décrire son logiciel
6. Herramientas prácticas / Des outils pratiques
7. Las exigencias van subiendo / Des obligations qui montent
8. El software de la región / Le logiciel de la région
9. En la política de Ciencia Abierta / Dans la politique de Science Ouverte
10. Lo que queda abierto: cita mañana / Ce qui reste ouvert : rendez-vous demain
11. Conclusión / Conclusion

Software Heritage: la misión

Software Heritage
THE GREAT LIBRARY OF SOURCE CODE

Recolectar, preservar y compartir *todo* el código fuente del software

preservar nuestro patrimonio, para un mejor software y una mejor ciencia para todos

Catálogo de referencia



encontrar y referenciar
todo el código fuente

Archivo universal



preservar y compartir
todo el código fuente

Infraestructura de investigación



habilitar el análisis de
todo el código fuente

acuerdo con la UNESCO desde 2017: el código fuente como patrimonio de la humanidad

Software Heritage

THE GREAT LIBRARY OF SOURCE CODE

Inria with unesco



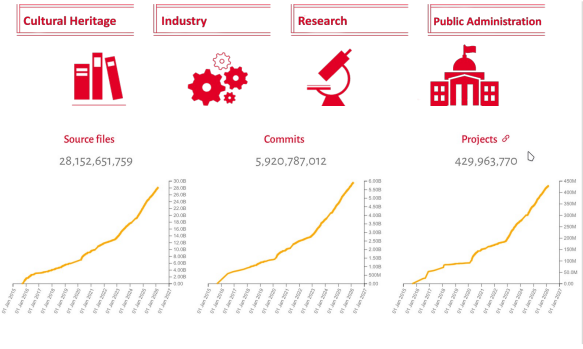
Software Heritage
THE GREAT LIBRARY OF SOURCE CODE

Inria with unesco



The largest open source code archive: one infrastructure, open, shared, non profit

Unique digital common good *built in France since 2015*



Software Heritage
THE GREAT LIBRARY OF SOURCE CODE

Inria with unesco



The largest open source code archive: one infrastructure, open, shared, non profit

Unique digital common good *built in France since 2015*

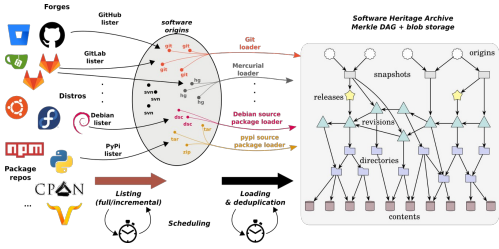
Cultural Heritage | **Industry** | **Research** | **Public Administration**

Source files
28,152,651,759

Commits
5,920,787,012

Projects
429,963,770





5000+ platforms

All versions, all history development in a single graph

Software Heritage
THE GREAT LIBRARY OF SOURCE CODE

Inria with unesco



The largest open source code archive: one infrastructure, open, shared, non profit

Unique digital common good *built in France since 2015*

Cultural Heritage | **Industry** | **Research** | **Public Administration**

Source files
28,152,651,759

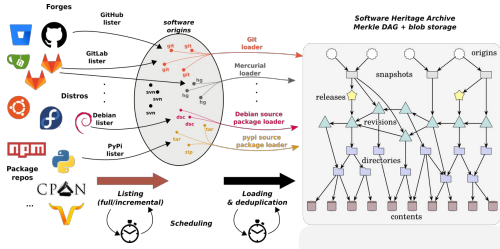


Commits
5,920,787,012



Projects
429,963,770





5000+ platforms

All versions, all history development in a single graph

- **50 × 10⁹** nodes
- **1000 × 10⁹** edges
- ~ **3 PB** of storage

Software Heritage
THE GREAT LIBRARY OF SOURCE CODE

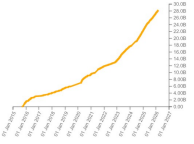
Inria with unesco

The largest open source code archive: one infrastructure, open, shared, non profit

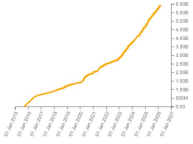
Unique digital common good *built in France since 2015*

Cultural Heritage | **Industry** | **Research** | **Public Administration**

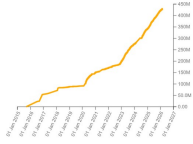
Source files
28,152,651,759

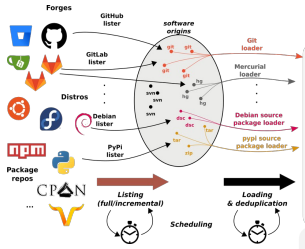


Commits
5,920,787,012



Projects
429,963,770





5000+ platforms

All versions, all history development in a single graph

- **50 × 10⁹** nodes
- **1000 × 10⁹** edges
~ **3 PB** of storage

A revolutionary **infrastructure** ensures **availability** guarantees **integrity** enables **traceability**

Software Heritage

THE GREAT LIBRARY OF SOURCE CODE

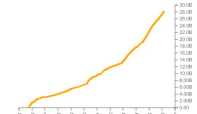
Inria with unesco

The largest open source code archive: one infrastructure, open, shared, non profit

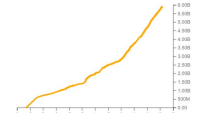
Unique digital common good *built in France since 2015*

Cultural Heritage | **Industry** | **Research** | **Public Administration**

Source files
28,152,651,759

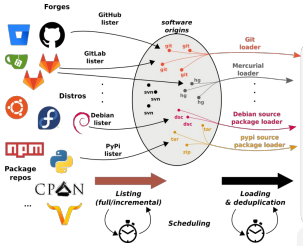


Commits
5,920,787,012



Projects
429,963,770





5000+ platforms

All versions, all history development in a single graph

- **50 × 10⁹** nodes
- **1000 × 10⁹** edges
- ~ **3 PB** of storage

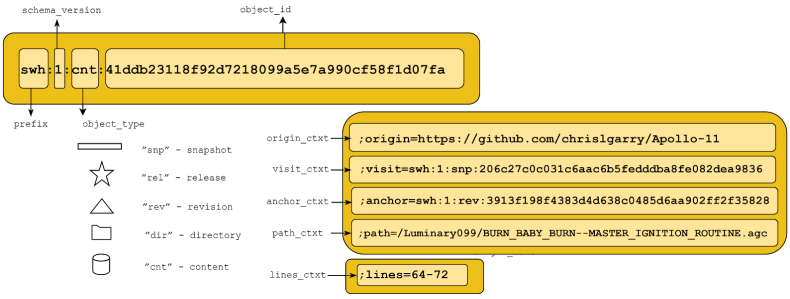
A revolutionary **infrastructure** ensures **availability** guarantees **integrity** enables **traceability**

IBM Microsoft TII Technology Innovation Institute HUAWEI cnrs Infosys cea Ministère de l'Enseignement Supérieur et de la Recherche aws Google Université Paris Cité openinventionnetwork

AdaCore GitHub Université di Pisa Sorbonne Université Université de Lorraine Sant'Anna NWO université PARIS-SACLAY

SoftWare Hash IDentifiers (SWHID)

swhid.org

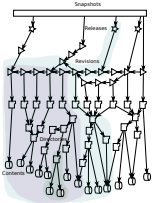


Referencias completas al código fuente: trazabilidad, integridad, reproducibilidad

- intrínseco, descentralizado, criptográfico: cualquiera puede recalcularlo
- granularidad: instantánea, versión, directorio, archivo, *líneas*
- norma internacional ISO/IEC 18670

guía: el HOWTO

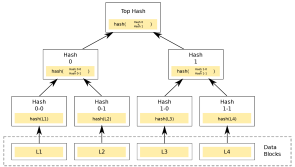
El *grafo* del desarrollo de software público



Todo el desarrollo de software en un solo grafo …

- trazabilidad

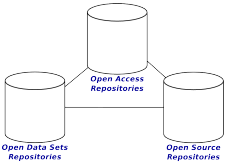
La *blockchain* del código público



… un grafo de Merkle

- integridad

Un *pilar* de la ciencia abierta

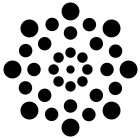


Archivo de referencia del software de investigación

- reproducibilidad
- referencia

Plataforma de referencia para el *Big Code*

estructura de datos uniforme



- estudios a gran escala
- ciberseguridad
- aprendizaje automático, IA, …

Contenido

1. Apertura / Ouverture
2. El código fuente, memoria de la ciencia / Le code source, mémoire de la science
3. Encuentro con Software Heritage / Voici Software Heritage
4. Archivar, Referenciar, Describir, Citar / Archiver, Referencer, Décrire, Citer
5. Describir su software / Décrire son logiciel
6. Herramientas prácticas / Des outils pratiques
7. Las exigencias van subiendo / Des obligations qui montent
8. El software de la región / Le logiciel de la région
9. En la política de Ciencia Abierta / Dans la politique de Science Ouverte
10. Lo que queda abierto: cita mañana / Ce qui reste ouvert : rendez-vous demain
11. Conclusión / Conclusion

Cinco pasos, en vivo, con un código de esta ciudad

1. Buscar: ¿su código ya está en el archivo?
2. Navegar y referenciar: un SWHID para un archivo, o para unas líneas
3. Archivar: con un clic
4. Verificar: recalcular el SWHID, sin conexión
5. Describir y citar: metadatos y botón de cita

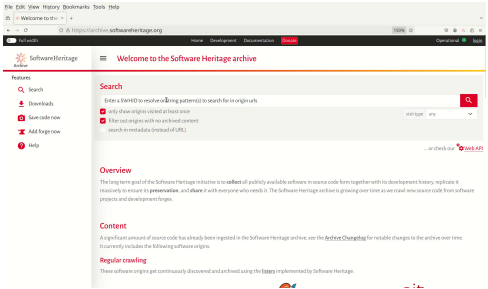
si la red falla, cada paso tiene su video

¿Su código ya está en el archivo?

- abrir archive.softwareheritage.org y escribir el nombre de un proyecto, de un grupo, de una institución
- en vivo: ONSAS, luego udelar
- la recolección: GitHub, GitLab.com, Bitbucket y muchas otras plataformas las recolectamos, mas detalles mas abajo

Pruébenlo ahora, con su propio proyecto

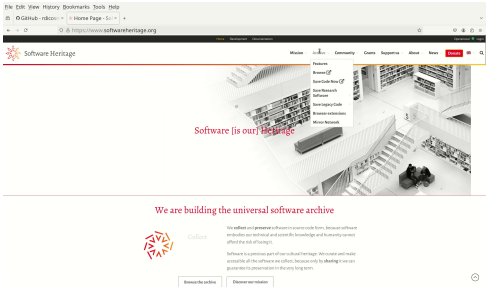
- navegar como en una plataforma de desarrollo: ONSAS en el archivo
- elegir un archivo, seleccionar líneas, pestaña *Permalinks*: el SWHID con su contexto
- el clásico: Apollo 11, líneas 245 a 261



video de respaldo

Paso 3: archivar con un clic

- Save Code Now: pegar la URL de un repositorio, y listo
- extensión UpdateSWH (Firefox, Chrome, Edge): un botón en la página del repositorio
- gris: no archivado; amarillo: desactualizado; verde: al día



vídeo de respaldo

El SWHID se calcula a partir del *contenido*: cualquiera lo recalcula, sin confiar en nadie.

```
$ D=swh:1:dir:1f3dcbf064df1bd365f0f903f291c6c1af9cd4aa   # ONSAS
$ A=https://archive.softwareheritage.org/api/1
# 1. pedir a la bóveda que prepare ese directorio, y esperar "done"
$ curl -X POST $A/vault/flat/$D/
# 2. descargar y descomprimir (-L: la descarga es una redirección)
$ curl -L -o onsas.tar.gz $A/vault/flat/$D/raw/
$ tar xzf onsas.tar.gz
# 3. recalcular el SWHID localmente
$ swh identify --type directory $D/ | cut -f1
swh:1:dir:1f3dcbf064df1bd365f0f903f291c6c1af9cd4aa
```

idéntico: el código es exactamente el citado, byte por byte

video de respaldo

- ONSAS trae un archivo de cita en su repositorio
- el archivo lo lee: botón *Cite* en la página del código
- ¿sin metadatos todavía? el generador CodeMeta: un formulario, y se descarga el `codemeta.json`

video de respaldo

Lo que genera el archivo para ONSAS

```
@softwareversion{swh-rev-20ff794,
 author = "Pérez Zerpa, J. M. and ...",
 title = .ºNSAS: an Open Nonlinear
   Structural Analysis Solver ...",
 version = "0.4.4",
 swhid = "swh:1:rev:20ff794...;
   origin=https://github.com/ONSAS/ONSAS;
   visit=swh:1:snp:02139aa..."}
```

formato biblatex-software, con el SWHID

Lo que un zip con DOI no puede hacer

Ce qu'un zip avec DOI ne peut pas faire

Un DOI que apunta a un *zip*

- lleva a una copia en la que hay que *confiar*: sin verificación de integridad
- un bloque opaco: sin *historia*, sin *diferencias*, nada direccionable *dentro*
- la granularidad se detiene en .[el] depósito entero"

Un *SWHID*

- intrínseco: el identificador *es* el hash criptográfico, se verifica solo
- designa *cualquier* nivel: instantánea, versión, directorio, archivo, líneas
- funciona con *cualquier* sistema de versiones (git, hg, svn) y con archivos tar

un DOI responde « *¿qué depósito está registrado?* »; un SWHID responde « *exactamente qué bytes, y lo demuestro* »

Contenido

1. Apertura / Ouverture
2. El código fuente, memoria de la ciencia / Le code source, mémoire de la science
3. Encuentro con Software Heritage / Voici Software Heritage
4. Archivar, Referenciar, Describir, Citar / Archiver, Referencer, Décrire, Citer
5. Describir su software / Décrire son logiciel
6. Herramientas prácticas / Des outils pratiques
7. Las exigencias van subiendo / Des obligations qui montent
8. El software de la región / Le logiciel de la région
9. En la política de Ciencia Abierta / Dans la politique de Science Ouverte
10. Lo que queda abierto: cita mañana / Ce qui reste ouvert : rendez-vous demain
11. Conclusión / Conclusion

El software tiene derecho de autor

Le logiciel relève du droit d'auteur

Convenio de Berna: el derecho de autor nace *automáticamente* con la obra, sin registro ni mención.

Por defecto: « todos los derechos reservados ». Sin una *licencia* explícita, nadie puede reutilizar legalmente su código.

- el menú práctico:
  - permisivas (MIT, BSD, Apache-2.0): favorecen la reutilización, sin obligación de devolver
  - copyleft (GPL, LGPL, AGPL): las versiones modificadas siguen abiertas
- declararla en forma *legible por máquina*, con un identificador SPDX en cada archivo:
  `# SPDX-License-Identifier: MIT`
- elegir con choosealicense.com, verificar en la lista OSI

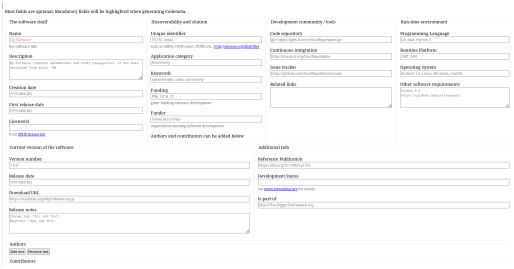
Atención: las licencias Creative Commons no están pensadas para el software

Metadatos legibles por máquina, que viajan *con* el código

- `codemeta.json`: esquema JSON-LD para el software (autores, licencia, referencias, financiamiento)
- `CITATION.cff`: el primo ligero, generable desde CodeMeta
- Software Heritage usa ambos para su botón de cita
- sin escribir JSON: un formulario web y se descarga un `codemeta.json` listo

  codemeta.github.io/codemeta-generator

El generador CodeMeta



Cuatro gestos

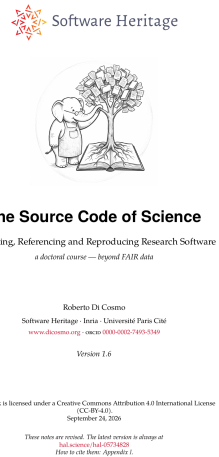
- archivar el código exacto detrás de cada resultado (un clic, o un webhook)
- referenciar con un SWHID, al nivel justo: versión, directorio, archivo, líneas
- describir: un README, una licencia, un `codemeta.json`
- citar el software en la bibliografía, como cualquier otro resultado de investigación

Como evaluador o editor

- pedir a los autores una *prueba de archivo* (un SWHID), no una URL frágil
- un hábito, aplicado cada vez, cambia las normas de una comunidad

Contenido

1. Apertura / Ouverture
2. El código fuente, memoria de la ciencia / Le code source, mémoire de la science
3. Encuentro con Software Heritage / Voici Software Heritage
4. Archivar, Referenciar, Describir, Citar / Archiver, Referencer, Décrire, Citer
5. Describir su software / Décrire son logiciel
6. Herramientas prácticas / Des outils pratiques
7. Las exigencias van subiendo / Des obligations qui montent
8. El software de la región / Le logiciel de la région
9. En la política de Ciencia Abierta / Dans la politique de Science Ouverte
10. Lo que queda abierto: cita mañana / Ce qui reste ouvert : rendez-vous demain
11. Conclusión / Conclusion



The Source Code of Science

Archiving, Referencing and Reproducing Research Software: un curso doctoral, de los identificadores a la evaluación de artefactos, con recetario y glosario.

CC-BY 4.0: para reutilizar, traducir y enseñar. ¿una versión en español?

hal.science/hal-05734828

Contenido

1. Apertura / Ouverture
2. El código fuente, memoria de la ciencia / Le code source, mémoire de la science
3. Encuentro con Software Heritage / Voici Software Heritage
4. Archivar, Referenciar, Describir, Citar / Archiver, Referencer, Décrire, Citer
5. Describir su software / Décrire son logiciel
6. Herramientas prácticas / Des outils pratiques
7. Las exigencias van subiendo / Des obligations qui montent
8. El software de la región / Le logiciel de la région
9. En la política de Ciencia Abierta / Dans la politique de Science Ouverte
10. Lo que queda abierto: cita mañana / Ce qui reste ouvert : rendez-vous demain
11. Conclusión / Conclusion

Lo que todavía se acepta con frecuencia

- « código disponible a pedido »
- una URL hacia una forja, que puede desaparecer
- un zip con DOI: una copia, sin historia, sin granularidad

Lo que habría que pedir

- el código archivado en un archivo dedicado, con su historia
- un identificador intrínseco (SWHID), al nivel justo, en el artículo

Ejemplos afuera de Ciencia de la Computación:

- JTCAM recomienda Software Heritage en sus instrucciones; *eLife* archiva automáticamente el código de sus artículos desde 2020
- Una revista de ciencias de la vida: la preservación del código forma parte *del proceso de publicación*, no de la buena voluntad de los autores.
  - *desde noviembre de 2020*: el código de cada artículo se archiva en Software Heritage
  - el SWHID se integra en el XML del artículo y aparece en su página
  - una validación verifica, antes de publicar, que el SWHID corresponde al repositorio

anuncio: eLife, 12 de noviembre de 2020

Contenido

1. Apertura / Ouverture
2. El código fuente, memoria de la ciencia / Le code source, mémoire de la science
3. Encuentro con Software Heritage / Voici Software Heritage
4. Archivar, Referenciar, Describir, Citar / Archiver, Referencer, Décrire, Citer
5. Describir su software / Décrire son logiciel
6. Herramientas prácticas / Des outils pratiques
7. Las exigencias van subiendo / Des obligations qui montent
8. El software de la región / Le logiciel de la région
9. En la política de Ciencia Abierta / Dans la politique de Science Ouverte
10. Lo que queda abierto: cita mañana / Ce qui reste ouvert : rendez-vous demain
11. Conclusión / Conclusion

From the **Software Sky** to **Institutional Insight:** Anatomy of an Observational Pipeline

Observation parameters (config file)

Stage 0: The Software Sky

The process begins with a complete map of the software universe, built from the Software Heritage archive, containing all observable software 'stars'. This stage is configured by observation parameters and a spectral filter definition (email domain selection).

Phase 1: Target Selection

Identify the stars associated with a specific institution by selecting contributors using institutional identifiers, which isolates the relevant region of the software sky for focused observation.

Quality Gate: Signal Strength Check

Before proceeding, the pipeline verifies that the selected region contains enough observable signals (software contributions) to produce statistically meaningful results.

Egeait to Nous Kaba

Phase 2: Gravitational Cleaning (Deforking)

The process renvents gravitational artifects such as mitrers and fures from the data to retain only the physically meaningful celestial objects (original software projects).

Phase 4: Spectral Enrichment

Each observed object is euriched with spectral data—such as activits, domains, and licenses—collected from multiple observation instruments to provide a deeper understanding.

Activity

Domains

Licenses

Adeptive Observation Scheduling

This fecture ensures the automatic relation of observation tokens to bypass iassument (DPQ limits, ensuring continuous data collection.

Observation Cache

A oerWixeen observation log is maintained, preventing redandant measurements and speeding up subsequent analyzes.

Phase 5: Astrophysical Analysis

The entiched observations are analysed to compute metrics, classily objects into scientific domains, and generate a synthetic, comprehensive view of the observed system.

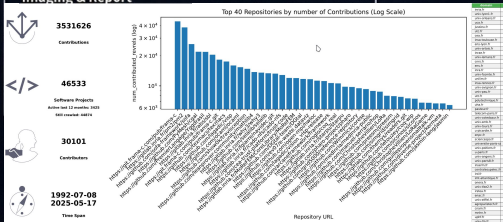
Phase 6: High-Resolution Imaging & Report

3531626
Contributions

46533
Software Projects
Active last 12 months: 3025
Still-created: 16874

30101
Contributors

1992-07-08
2025-05-17
Time Span



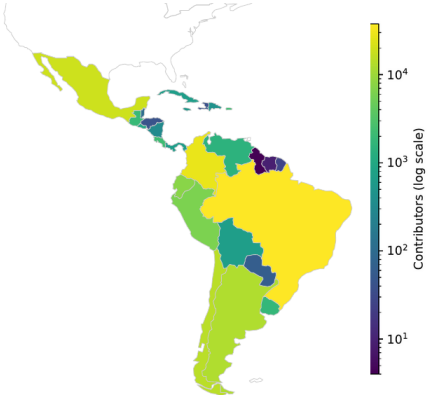
- **620 095** proyectos con contribuciones firmadas desde instituciones de investigación de la región, en **26** países
- **139 221** personas firmaron esas contribuciones
- **50 605** proyectos ya no están donde se publicaron: existen hoy solo porque fueron archivados
- **3 580** recibieron una contribución en los últimos doce meses
- **743** (0,1 %) declaran cómo quieren ser citados



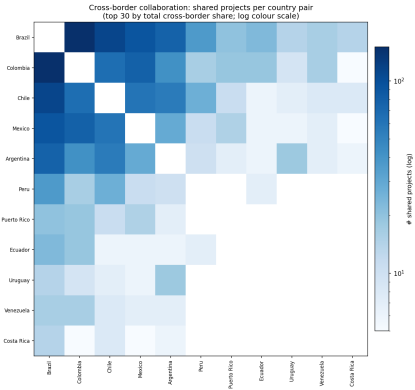
el informe completo, para los delegados

Quién construye, y con quién

Qui construit, et avec qui



139 221 personas, según el país de su dominio institucional (escala logarítmica)



proyectos compartidos por cada par de países (pares con 5 o más)

la región ya construye software *como región*

Los 40 proyectos con más contribuciones

Les 40 projets les plus contribués

1 wazuh	11 RocketPy	21 igac	31 sits
2 terralib	12 agda	22 OpenROAD	32 COVID-19
3 metamask-mobile	13 rainbow	23 MOSAIC	33 jaziku
4 Tesis_Maestria	14 NumCosmo	24 nclcomposer	34 RIOT
5 lua	15 salva	25 TCGAbiolinks	35 wrf-baq-0.5km
6 web	16 spyder	26 CardioLux	36 ONSAS
7 techminer-api	17 inland	27 keystone	37 radiation-benchmarks
8 rsat-code	18 openbus-core	28 python-orm	38 NEOs.jl
9 meran	19 carmen_lcad	29 NgenBundle	39 firecrown
10 ginga	20 terrame	30 spyro	40 verigraph

por contribuciones (revisiones y lanzamientos) de las instituciones de la región, entre los 620 095 proyectos del informe

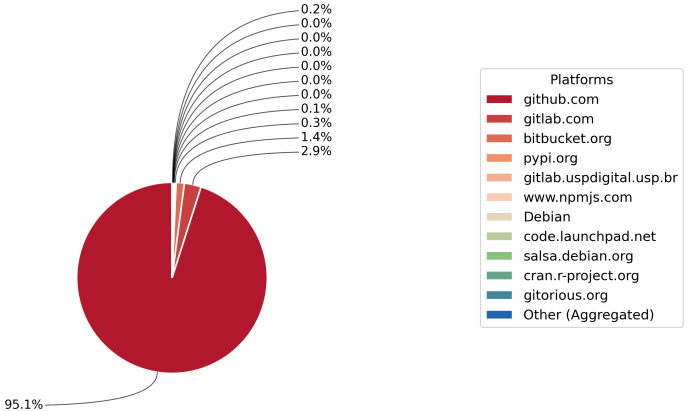
ONSAS: Open Nonlinear Structural Analysis Solver

Facultad de Ingeniería, Universidad de la República: Instituto de Estructuras y Transporte, e Instituto de Ingeniería Mecánica y Producción Industrial; con coautores en Porto Alegre (UFRGS) y Montreal

- nacido en 2017 como apoyo de un curso de la Facultad
- código abierto, licencia GPL-3.0, archivo de cita en el repositorio
- artículo en el *Journal of Open Source Software*, 1 de abril de 2026: 10.21105/joss.09604
- archivado en Software Heritage: `swh:1:dir:1f3dcbf064df1bd365f0f903f291c6c1af9cd4aa`

¿Puede su CVUy valorarlo como corresponde?

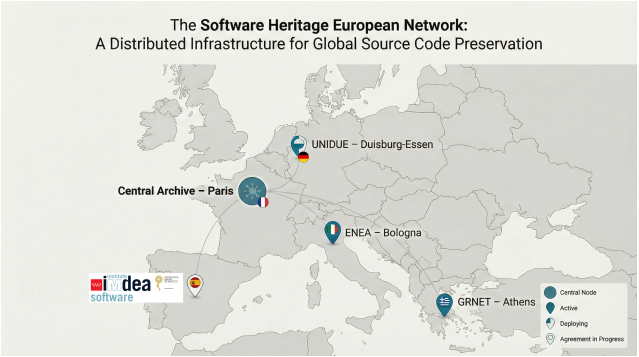
El 95 % en una sola plataforma



dónde están los 620 095 proyectos del informe

- el 95,1 % del software académico de la región está en una sola plataforma, de una sola empresa, en una sola jurisdicción
- detrás, las forjas de las propias instituciones: se preservan solo si alguien pide archivarlas

Concentración global = riesgo global



cuatro espejos, todos en Europa

Una copia

- bajo el control de *una* organización
- protege contra las fallas

Un espejo

- bajo control *independiente*, en su propia jurisdicción
- protege contra las decisiones

Cada país puede alojarlo.

Ningún país lo controla.

Contenido

1. Apertura / Ouverture
2. El código fuente, memoria de la ciencia / Le code source, mémoire de la science
3. Encuentro con Software Heritage / Voici Software Heritage
4. Archivar, Referenciar, Describir, Citar / Archiver, Referencer, Décrire, Citer
5. Describir su software / Décrire son logiciel
6. Herramientas prácticas / Des outils pratiques
7. Las exigencias van subiendo / Des obligations qui montent
8. El software de la región / Le logiciel de la région
9. En la política de Ciencia Abierta / Dans la politique de Science Ouverte
10. Lo que queda abierto: cita mañana / Ce qui reste ouvert : rendez-vous demain
11. Conclusión / Conclusion

América Latina ya registra la investigación

L'Amérique latine enregistre déjà la recherche

País	Sistema nacional de CV o registro	¿Lugar para declarar software?
Argentina	CVar (MINCyT), SIGEVA (CONICET)	?
Bolivia	no encontrado en nuestra búsqueda	?
Brasil	Currículo Lattes (CNPq)	sí, según guías universitarias
Chile	Portal del Investigador (ANID)	?
Colombia	CvLAC (MinCiencias)	**sí**: « Producción técnica y tecnológica »
Costa Rica	Hipatia, Talento Costa Rica (MICITT)	?
Cuba	no encontrado en nuestra búsqueda	?
Ecuador	Registro de investigadores (SENESCYT)	?
México	CVU (Secihti, antes Conacyt)	**sí**: « Desarrollo de Software »
Panamá	Directorio Nacional de Investigadores (SENACYT)	?
Paraguay	CVPy (CONACYT)	?
Perú	CTI Vitae (CONCYTEC)	?
Uruguay	CVUy (ANII)	?
Venezuela	ReNII (ONCTI)	?

Colombia: preguntas frecuentes CvLAC, MinCiencias (2014). México: manual del CVU, Conacyt (2009). "?": no lo hemos podido confirmar, ¿ustedes lo saben?

la región decidió hace años que el software se declara; ¿alguno de estos sistemas *apunta* al software mismo?

Quince acuerdos básicos para la evaluación de la investigación académica en Uruguay (2025)

II: « armonizar la evaluación [...] se requiere en especial en lo que tiene que ver con la valoración de la formación de grado [...], de la producción tecnológica y de la dimensión de construcción institucional »

XIV: « CVUy debería poder contener cualquier mérito pertinente para una evaluación abarcadora »

Participaron: ANII, Comisión Honoraria del SNI, Udelar, UCU, ORT, UTEC, INIA, Institut Pasteur de Montevideo, PEDECIBA, IIBCE, ANCIU, AUPP, CFE (ANEP), Investiga.uy. (texto)

Ya declararon software en su CVUy.
¿Quién lo lee puede llegar al software?

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR, DE LA RECHERCHE ET DE L'ESPACE
Liberté
Égalité
Fraternité

Ouvrir la science !



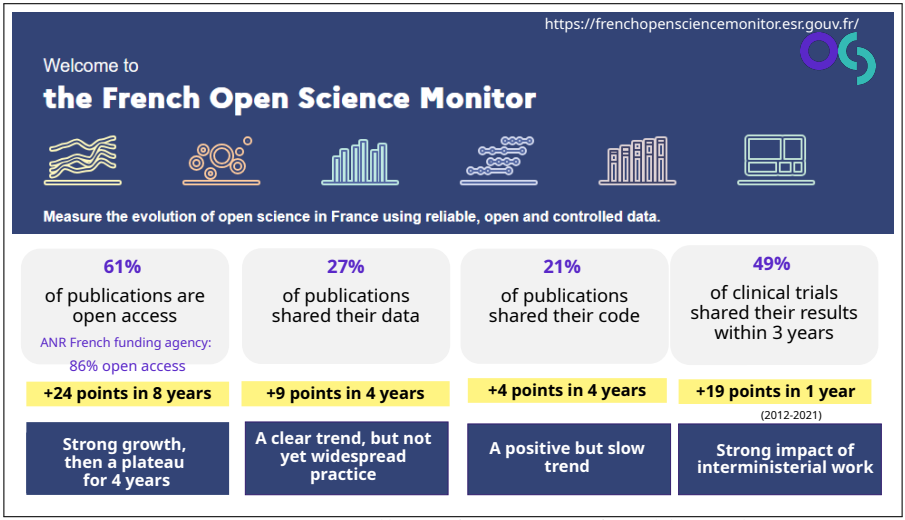
- planes nacionales de ciencia abierta: julio de 2018, y 6 de julio de 2021
- Estrategia Nacional de Ciencia Abierta 2026-2033, publicada el 16 de septiembre de 2026: un objetivo dedicado al software, que pide una forja soberana y el archivo en Software Heritage; acción OS7.A4, « Soutenir Software Heritage »
- Software Heritage: infraestructura de investigación de la hoja de ruta nacional, con apoyo plurianual del Estado

hechos, no recetas: cada país elige su camino

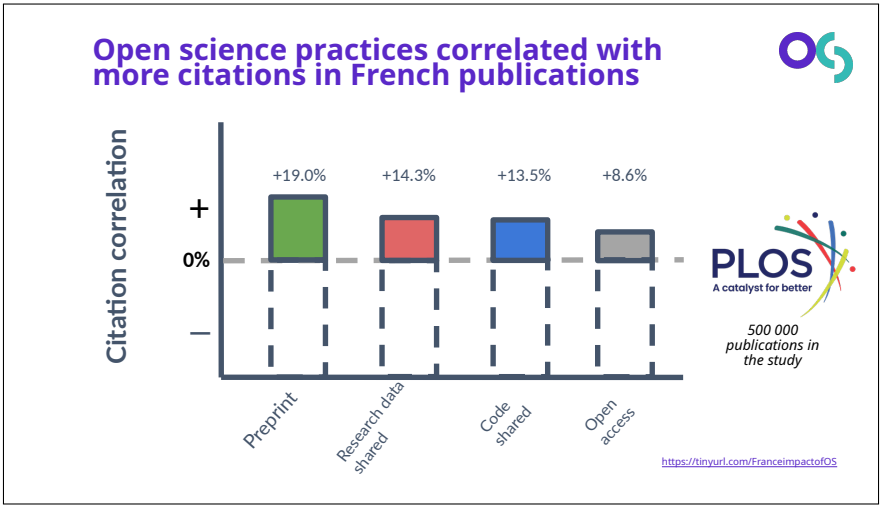
la Estrategia: doi.org/10.52949/93

Francia mide: el 21 % comparte su código

La France mesure : 21 % partagent leur code



publicaciones francesas, Barómetro francés de la ciencia abierta; Marin Dacos, octubre de 2026



correlación, no causalidad; 500 000 publicaciones francesas

Colavizza et al., *QSS*, 2026, doi:10.1162/QSS.a.487



Un pasaporte para el código fuente

conceptos clave para estudiantes, docentes e investigadores



Informe sobre las forjas académicas

necesidades, opciones, limitaciones



Premios nacionales de ciencia abierta

software y datos; tesis desde 2024

Eduarda Centeno: premio de tesis 2025

medicina, biología y salud

Ministerio de Educación Superior e Investigación

Focus en el premio software

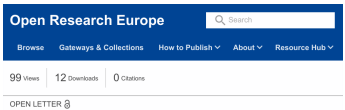
2022	Coq · scikit-learn · Faust · Gammapy	+ 6 accessits
2023	Smilei · OCaml · Brian · Hyphe	+ 4 espoirs
2024	Ichthyop · PARI/GP · scikit-rf · SOFA	+ 4 espoirs
2025	GNU MPFR · Aladin Lite · Pharo · HyperSpy	+ 4 espoirs

34 laureates,
open data

categories: scientific and technical · community · documentation · jury / coup de cœur

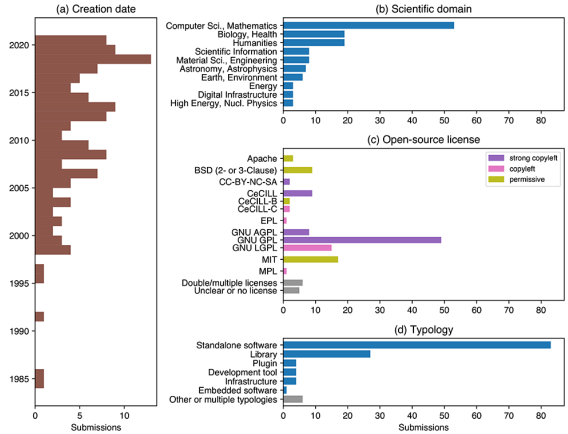
próxima entrega: noviembre de 2026, Montpellier (Assises nationales des données de la recherche)

Blueprints and analysis available



- goals, design decisions
- challenges and solutions
- lessons learned
- detailed data

A look at the data



Contenido

1. Apertura / Ouverture
2. El código fuente, memoria de la ciencia / Le code source, mémoire de la science
3. Encuentro con Software Heritage / Voici Software Heritage
4. Archivar, Referenciar, Describir, Citar / Archiver, Referencer, Décrire, Citer
5. Describir su software / Décrire son logiciel
6. Herramientas prácticas / Des outils pratiques
7. Las exigencias van subiendo / Des obligations qui montent
8. El software de la región / Le logiciel de la région
9. En la política de Ciencia Abierta / Dans la politique de Science Ouverte
10. Lo que queda abierto: cita mañana / Ce qui reste ouvert : rendez-vous demain
11. Conclusión / Conclusion

INSTITUTIONS

• **Support development**	**mandatory**
• Promote software	recommended
• Build and maintain institutional forges	optional

FUNDERS

• **Provide grants for long-term support**	**mandatory**
• Promote reproducibility	recommended
• Facilitate collaborations	optional

LIBRARIES

• **Prepare software archival plans**	**mandatory**
• **Create and curate software metadata**	**mandatory**
• Catalog software	recommended

PUBLISHERS

• **Enforce open source**	**mandatory**
• Link publications and codes	recommended
• Review software	optional

- el investigador no puede solo: cuatro actores, cada uno con su parte *obligatoria*
- instituciones: apoyar el desarrollo
- financiadores: financiar el mantenimiento a largo plazo
- bibliotecas: planes de archivo y metadatos curados
- editoriales: exigir el código abierto

responde a las necesidades de cada nivel

Una oficina dedicada en cada institución

- interfaz entre investigadores y servicios: transferencia, jurídico, ciencia abierta
- una visión del software que produce la institución
- asesoría y formación: archivar y referenciar, describir y citar, licencias, forjas

Francia, 2026

- red de centros de experticia en software certificados en SSO: objetivo cinco OSPO certificados
- mañana, viernes 11:00: Violaine Louvet: OSPO de la Université Grenoble Alpes, y Eduarda Centeno: qué piden las grillas de evaluación

Brasil, desde 2008

- el CCSL de la USP, Centro de Competência em Software Livre
- nacido del proyecto europeo QualiPSo: una idea de casi veinte años



Proyecto de Software Heritage financiado por la Fundación Alfred P. Sloan

Localizar, seguir y entender su impacto en código abierto

1. Definir: la lista de URL del perímetro de software de la institución
2. Recolectar: los metadatos de forjas, registros y plataformas
3. Preservar: el código se archiva en Software Heritage, con SWHID
4. Gestionar: curar, anotar y exponer la colección en un tablero y una API pública

Complemento: un informe de análisis, del dominio de correo institucional a la lista de URL

Se buscan instituciones piloto y socios de co-diseño `swh-ospo-radar@inria.fr`

fuente: M. Gruenpeter, *Building a window into the Software Heritage Cathedral*, 24/06/2026, CC-BY-4.0

Ejemplos de preguntas claves

- ¿Quién mantiene el software de investigación cuando termina el proyecto, y quién paga?
- ¿Cómo cuenta el código cuando se evalúa a un investigador? (la mesa redonda de esta tarde)
- ¿Cómo aprende un doctorando a escribir código que dure?

Viernes 30 de octubre, 11:00 a 13:00, Edificio Los Ombúes

Mesa redonda: hacer del software un pilar efectivo de la ciencia abierta (infraestructura, reconocimiento y gobernanza), con Violaine Louvet (CNRS) y Eduarda Centeno (Rede Brasileira de Reprodutibilidade)

Viernes 30 de octubre, 14:30 a 16:30, en esta misma sala

Taller: oportunidades de cooperación en software libre y de código abierto, en español, y a las 16:30 la conclusión de los talleres: cooperación y proyectos futuros

vengan con su propuesta

Contenido

1. Apertura / Ouverture
2. El código fuente, memoria de la ciencia / Le code source, mémoire de la science
3. Encuentro con Software Heritage / Voici Software Heritage
4. Archivar, Referenciar, Describir, Citar / Archiver, Referencer, Décrire, Citer
5. Describir su software / Décrire son logiciel
6. Herramientas prácticas / Des outils pratiques
7. Las exigencias van subiendo / Des obligations qui montent
8. El software de la región / Le logiciel de la région
9. En la política de Ciencia Abierta / Dans la politique de Science Ouverte
10. Lo que queda abierto: cita mañana / Ce qui reste ouvert : rendez-vous demain
11. Conclusión / Conclusion

no traten su código como algo secundario: *regístrenlo, compártanlo, valórenlo*

Roberto Di Cosmo roberto@dicosmo.org

- el archivo archive.softwareheritage.org
- archivar un repositorio archive.softwareheritage.org/save
- el curso hal.science/hal-05734828
- el informe anual
  annex.softwareheritage.org/public/annual-reports/2025/
- el informe regional annex.softwareheritage.org/public/reports/2026/

¿Quieren saber más, o sumarse? Escriban a Morane Gruenpeter, morane@softwareheritage.org

estas diapositivas

softwareheritage.org

¿Preguntas?

Estas diapositivas son CC-BY 4.0.