



Issue #7, December 22nd, 2021



You can help building the universal archive of software source code

"Open science infrastructures should be organized and financed upon an essentially not-for-profit and long-term vision"

UNESCO recommendation on Open Science, November 2021

Software is the fabric that binds together all aspects of modern society, and we now need a **global infrastructure** that enables us to preserve, access, study and improve the large corpus of technical, scientific and organisational knowledge that is embodied in software source code. We decided to answer these needs by building an **open, non-profit, multi-stakeholder, long-term and shared infrastructure**. It is a humbling undertaking, and we will only succeed with the contribution of a large community, from [funders](#) to users and adopters, from industries to public bodies and individual contributors.

This November 30, 2021, [a special event took place at UNESCO's headquarters](#) to celebrate the [five years of Software Heritage](#), bringing together our growing community, and [welcoming CEA as our first diamond sponsor](#). It was the opportunity to take stock of the achievements and status of Software Heritage, and to highlight the relevance of building a universal software source code archive in the context of today's dynamic digital innovation landscape.

The [video included below](#) recaps, in just five minutes, the key milestones of these first five years.

Looking at what has been accomplished in these past years and the growing number of [contributions of all kinds to our work](#), and the growing number of [ambassadors volunteers](#), gives us renewed energy to continue, with passion and dedication, working on Software Heritage's long-term mission.

This year, we decided to [launch our first open fundraising campaign](#), with the goal to broaden the community of individuals involved in our work. **You can make a difference**, by [becoming a donor today](#), and telling the people you know why this is important.

♥ Yes, I want to become a donor now! ♥



OUR MISSION

We collect and preserve software in source code form, because software embodies our technical and scientific knowledge and humanity cannot afford the risk of losing it. Software is a precious part of our cultural heritage. We curate and make accessible all the software we collect, because only by **sharing** it we can guarantee its preservation in the very long term.

Collect



Preserve



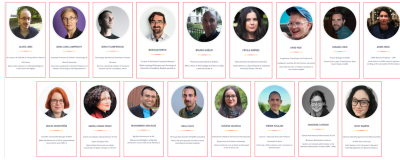
Share



Software Heritage will provide solid, common foundations to serve the different needs of heritage, preservation, science, and industry.

Donate

IT'S NEW!



Seventeen Software Heritage Ambassadors in 2021
We are delighted to see that in this first year 17 ambassadors from around the world and with different backgrounds have come on board: you can find out more about them on the dedicated ambassador program page.

[Read more](#)

December 20, 2021



the University of Pisa.

Software Stories: a novel approach to present landmark legacy code

We have unveiled Software Stories, a novel approach to present the history of landmark software projects, making it accessible to a wide range of software enthusiasts without any technical background. The Software Stories project is a joint collaboration with the sciencestories.io team and

[Read more](#)

December 16, 2021



Meeting at UNESCO to celebrate five years of Software Heritage

This year 2021 we celebrate five years of continuous and passionate dedication to the Software Heritage mission, which we announced to the world on June 30th 2016: to collect, preserve and share all the available

source code, for the benefit of society as a whole.

[Read more](#)

December 7, 2021



Join the donor community, and help preserve software

For the first time in our history, we are launching an end-of-year fundraising campaign, with the initial goal of reaching 100 donations by December 31st, 2021. Every donation counts, by any amount.

[Read more](#)

October 14, 2021



A new effort to expand the Software Heritage archive

We are excited to announce that a third grant managed by NLNet has been awarded to Octobus to help expand the coverage of the Software Heritage archive, by developing new listers and loaders (...)

[Read more](#)

September 24, 2021



Building the next generation object storage for Software Heritage

Building the Software Heritage infrastructure is challenging. This is why we partnered with funders around the world to provide grants for experts that are willing to engage with the long term mission of Software Heritage.

[Read more](#)

August 12, 2021



Archiving SourceForge

We are delighted to announce that we have started archiving SourceForge, thanks to a collaboration with Octobus, funded by a Sloan grant.

[Read more](#)

July 29, 2021



Software for Open Science: a leap forward

We are excited to see software fully recognized as a key pillar of Open Science in the second national plan for Open Science, unveiled by the French Ministry of Research on July 6 2021.

[Read more](#)



Archiving the Maven ecosystem

We are delighted to announce that a fifth subgrant has been awarded to Castalia Solutions.

Castalia Solutions will be developing the Maven Repositories connector to archive the Maven ecosystem.

[Read more](#)

GRANTS



Helping you to contribute to the mission

Source code is spread around a broad spectrum of code hosting platforms and websites, and developed using a variety of version control systems. In order to collect and archive it properly, Software Heritage needs adaptors for each of the platforms and version control systems out there.

We partnered with the [Alfred P. Sloan Foundation](#) and the [NLnet foundation](#) to provide grants for experts that are willing to engage with the long term mission of Software Heritage.

[For more information on how to apply for the programs that are still open.](#)

YOU CAN HELP!



The Software Heritage archive will serve the needs of the many, from [cultural institutions](#) to [scientists](#) and [industries](#).

Everyone can help us achieving these ambitious goals and there are several ways to help.



Becoming a sponsor

Pursuing our roadmap for the archive requires significant resources. We welcome companies, institutions, and individuals who would like to join our sponsorship program and sustain the



Tackle scientific challenges

Building, maintaining, and exploiting the universal source code archive poses relevant scientific challenges. We welcome scientists who would like to contribute to this mission by

Software Heritage.

[Discover our sponsorship program!](#)



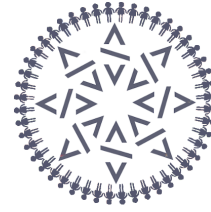
Code with us

All the software we develop ourselves is open source. We welcome contributors that are willing to delve into it and help us building the many components that are needed to make the archive progress towards the next milestones.

[Dive into the code!](#)

participating in our research activities.

[Join our research community!](#)



Host a mirror

Institutions and companies from all around the world are welcome to join our mirror program. This is essential to prevent information loss, and will greatly simplify access to humankind's software heritage.

[Contact us to host a mirror!](#)



Discover the growing number of [institutions, organizations and companies that sponsor Software Heritage](#), and those that express support for the [software preservation mission](#). [Take a look at their testimonials!](#)

Individual [donations](#) are also most welcome, in any amount, just click on the button below to give to the mission.

Donate

SELECTED MILESTONES

April 4, 2017



Raising awareness for Software Source Code preservation and access

At Unesco's headquarters in Paris, a formal partnership agreement has been signed between Unesco and Inria to establish a framework for collaboration on preserving the knowledge embedded in software source code, and making it

widely available.

Read more

*Part of our Heritage, Pillar
of our Present, Enabler
of our Future*

The Paris Call is here: calling for action to sustain Software Source Code!

A shared mission and a continuing collaboration with UNESCO: a call to action on software, and software source code

[Read more](#)

April 17, 2019

Saving software development from the European copyright reform

Today I have the meager satisfaction to see that one year and half of a hard and exhausting work of explanations and mobilizations, at national and European level, to which I have actively contributed, has led to a clear exclusion formulated in Article 2(6) for

"open source developing and sharing platforms"

[Read more](#)

June 6, 2019



Announcing the Software Heritage Graph Dataset

We are delighted to announce that the first public dataset encompassing the entire Software Heritage graph is now available.

The Software Heritage graph dataset is a representation of the fully-deduplicated Merkle DAG of the Software Heritage archive.

It links together file content identifiers, source code directories, commits in version control systems (VCS) tracking evolution over time, up to the full states of

VCS repositories as observed by Software Heritage during periodic crawls.

[Read more](#)

September 5, 2019

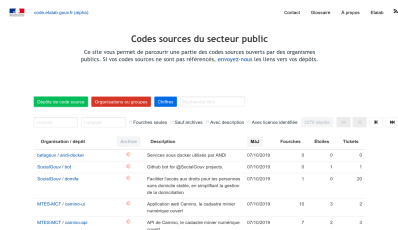


Software Heritage for Open Science

On July 4th 2018 was unveiled the French national plan for Open Science, that identifies Software Heritage as one key initiative to support. Today we are delighted to welcome the French Ministry of Research as our first gold sponsor, contributing effectively to our efforts. On this occasion, we are happy to recall why Software Heritage is building an infrastructure that plays a key role for Open Science.

[Read more](#)

October 9, 2019



Organisme / projet	Description	MU	Fonctions	Statut	Titulaire
Stratégie nationale	Services aux citoyens offerts par ARS	07102019	8	0	0
Stratégie 2018	Offre de services numériques	07102019	4	1	1
Stratégie 2018	Faciliter l'accès aux droits pour les personnes dans le cadre d'un accompagnement global de la santé	07102019	1	0	0
SPESMCT / Centre de données	Aggrégation des données de santé issues de multiples sources	07102019	10	3	2
SPESMCT / Centre de données	API de données de santé issues de multiples sources	07102019	7	2	3

Comment trouver les codes sources aujourd'hui ouverts par l'administration ?

Vous remarquez une icône dans la colonne « Archive » : cet acronyme désigne le projet Software Heritage, dont la mission est de sauvegarder tout le code source existant. À mesure que nous recensons des dépôts, nous faisons aussi

en sorte de demander leur archivage : c'est ainsi que nous avons aujourd'hui atteint les 99% de dépôts de codes sources archivés sur Software Heritage, une garantie supplémentaire pour la pérennité de ces informations publiques.

[Read more](#)

October 3, 2019



Software Heritage has a long-term strategy to ensure that its contents will be safeguarded. An important part of this strategy is the development of an international network of mirrors, implemented using a variety of storage technologies, running in various administrative domains, controlled by different institutions, and located in different jurisdictions. The mirror network will provide much higher resilience than multiple copies under our own administrative control.

[Read more](#)

October 10, 2019



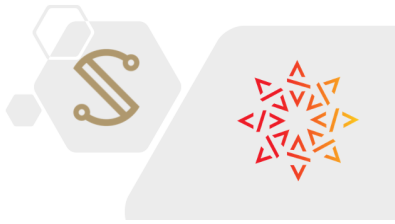
Curating and archiving landmark legacy source code: you can help!

We are delighted to make a step forward about an important and urgent challenge: collecting and curating landmark legacy source code written by pioneers of the digital age, many of which are still around, and willing to contribute their knowledge, but only for a limited time. They are spread all around the world, and the amount of knowledge at risk of

being lost is huge, so we'll only succeed if a broad international community steps up to tackle this noble task.

[Read more](#)

May 20, 2020



Welcoming Sloan support to expand the Software Heritage archive

We are delighted to announce a grant from the Alfred P. Sloan Foundation to foster the emergence of a community of expert contributors to increase the coverage of the Software Heritage archive.

[Read more](#)

May 26, 2020



Citing software with style

*Treating software as a first-class citizen...
... and citing it with the biblatex-software style*

Software, and software source code in particular, is a pillar of modern research, and a stepping stone for Open Science. It must be properly archived, referenced, described, and cited in order to build a stable and long-lasting corpus of scientific knowledge.

[Read more](#)

June 11, 2020



IPOL archives research software in Software Heritage

Supporting Open Science and reproducibility of research is part of our mission, and we are excited to announce that the Image Processing On Line journal (IPOL) has decided to deposit systematically in the Software Heritage archive all the software artifacts associated to the articles it publishes.

[Read more](#)



eLife archives research software in Software Heritage

We are excited to announce that the journal eLife has decided to systematically save source code repositories in the Software Heritage archive containing all software artifacts associated with the published article.

artifacts associated with the published article.

[Read more](#)

December 7, 2020



The Task Force on Scholarly Infrastructures of Research Software, as part of the Architecture WG of the European Open Science Cloud (EOSC) Executive Board, has established a set of recommendations to allow EOSC to include software, next to other research outputs like publications and data, in the realm of its research artifacts.

This work is built upon a survey and documentation of

a representative panel of current operational infrastructures across Europe, comparing their scopes and approaches.

[Read more](#)

December 22, 2020

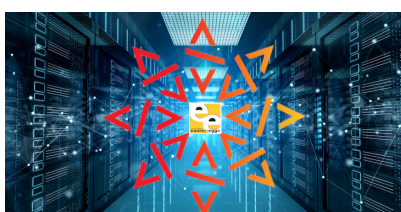


Introducing the Software Heritage Ambassador programme

We are delighted to announce the Software Heritage ambassador programme, and look forward to welcome enthusiastic organizations and individuals that want to help spread the word about Software Heritage and the services it provides to society as a whole.

[Read more](#)

March 11, 2021



Towards a next-generation object storage for Software Heritage

We are delighted to announce a collaboration with Easter-eggs to address another key issue we face: building an object storage that can scale to tens of billions of source code files, which can be quite small compared to the usual workload for

which object storage systems are designed.



Read more

May 11, 2021



Next Generation counters are up and running!

It is quite natural to want to know how many artifacts are in the archive, from how many projects, and involving how many contributors: to answer these questions, we show a variety of counters on the main archive.

Counting billions of objects...

... using probabilistic algorithms...

... is now deployed

Read more

2020 ANNUAL REPORT



The Software Heritage annual report is out!

Despite the crisis and the difficulties, 2020 has seen great progress at Software Heritage, that now preserves over 2 billion commits from more than 150 million projects!

Learn what has been achieved at the service of industry, research, cultural heritage, government transparency and Open Science.

Discover how Software Heritage is now not only grant seeking, but grant making. Explore the new possibilities that are now available to get involved and contribute to the long term mission that Software Heritage is undertaking. Read the news and watch the videos on our channel. Promote the use of Software Heritage by becoming an Ambassador.

Read more



You can read news from Software Heritage [on our blog](#), follow us on [Twitter](#), [LinkedIn](#) and also our [RSS flux](#)!

You can watch Software Heritage's presentations on our YouTube [video channel](#)!

And if you missed our newsletter's first issues, consult the [archive](#)!

Let's spread the word!