Issue #6, July 1st, 2021

5 years collecting, preserving and sharing humankind's software heritage

"The best way to communicate from one human being to another is through story"

Donald Knuth

On June 30, 2016, after almost two years of preparatory work, we opened up to the world the Software Heritage website unveiling to the world our long-term mission to collect, preserve and make easily available the source code of all software publicly available. Today, we are celebrating five years of continuous and passionate dedication to this mission, and we would like to start by thanking the many partners that have made this possible: sponsors that have followed up on Inria's initial full support, UNESCO for the precious collaboration on the vision and the policy issues, private foundations that have made it possible to fund key expert contributors, many organizations and individuals that share the vision, and a growing number of contributors and donors.

In the occasion of this fifth anniversary, we have prepared for you a special blog post that revisits, through the annual reports, the long road traveled with passion and dedication in pursuit of our mission to collect, preserve and share the source code of all software publicly available: we hope you will enjoy reading it as much as we enjoyed writing it.

Building an open, non-profit, multi-stakeholder, long-term and common infrastructure to collect, preserve and share all the source code is a humbling undertaking. We know well that 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. 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.

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.

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

Donate

IT'S NEW!

June 30, 2021

Celebrating Five Years of Building Software Heritage

A detailed account of the long road traveled can be found in the annual reports that have been published for each of the past years, and this fifth anniversary is an occasion to review some of the many milestones reached.

Read more

June 17, 2021

Introducing Vicky Rampin, our fourth ambassador

We are delighted to introduce our fourth ambassador, Vicky Rampin, librarian for research data management and reproducibility at New York University, and an advocate for open scholarship.

Read more

International Archives Day

Launched on June 9th, 2007, this day turned into a whole week in 2019 and is meant to remind all of us of the importance of archives for keeping the memories of our society.
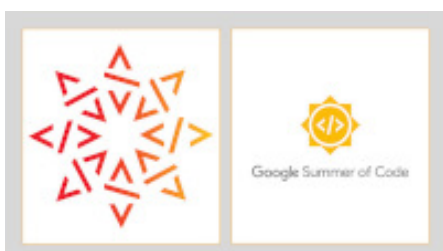
Read more

Introducing the 3rd Software Heritage's ambassador

We are delighted to introduce our 3rd ambassador, Malin Sandström, the Community Engagement officer for the International Neuroinformatics Coordinating Facility (INCF).

Read more

Software Heritage welcomes Google Summer of Code students… again!

The Software Heritage team is delighted to welcome two new students for the Google Summer of Code 2021: Kumar Shivendu et Daniele Serafini.

Read more

Meet the second Software Heritage ambassador

We are delighted to introduce Bruno Khélifi, the second Software Heritage ambassador in academia.
Bruno is a physicist working at APC in the experimental area of very-high-energy gamma-ray astrophysics, as a member of the H.E.S.S. and CTA collaborations.

Read more

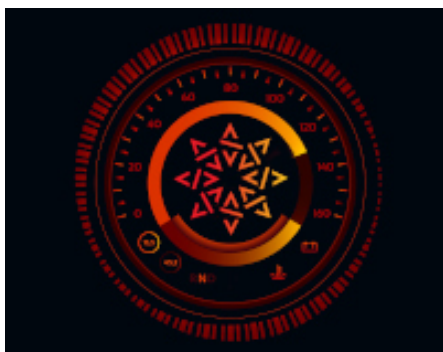## Know your software: contributing to cybersecurity through traceability

A recent executive order from the President of the United States focused on responding to cybersecurity threats has a full section dedicated to "Enhancing Software Supply Chain Security". It announces forthcoming standards, procedures, or criteria regarding in particular means for:

(x) ensuring and attesting, to the extent practicable, to the integrity and provenance of open source software used within any portion of a product.

Read more

Next Generation counters are up and running!

Software Heritage is archiving hundreds of millions of repositories, produced by tens of millions of developers. This amount to billions of distinct software artefacts (files, directories, revisions, releases) that need to be properly ingested in the gigantic Merkle graph that lies at the core of the archive.

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 Software Heritage.

Discover our sponsorship program!

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 participating in our research activities.
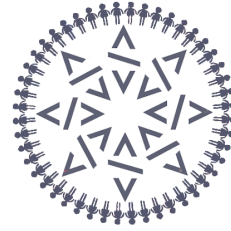
Join our research community!

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!

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

June 30, 2016

Software Heritage is now open, please come in!
Source code is special: it's executable knowledge that is meant to be human readable, by design.
Software Heritage is an ambitious initiative that aims at collecting, organizing, preserving and sharing all the source code publicly available in the world. We decided to undertake this ambitious task because software is our heritage, because software is essential for science because software is the cornerstone of our industry, our society and our own lives.
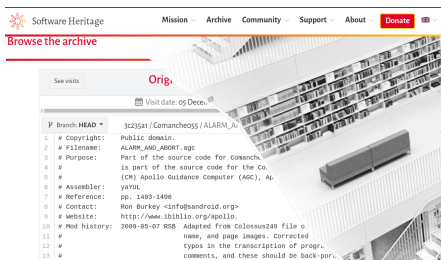
Read more

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.
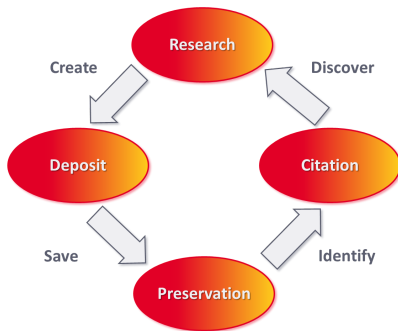
Read more

June 7, 2018

Opening the doors of the Software Heritage archive

Today, June 7th 2018, we are proud to be back at Unesco headquarters to unveil a major milestone in our roadmap: the grand opening of the doors of the Software Heritage archive to the public (the slides of the presentation are online). You can now look at what we archived, exploring the largest collection of software source code in the world: you can explore the archive right away, via your web browser. If you want to know more, an upcoming post will guide you through all the features that are provided and the internals backing them.

Read more

Depositing scientific software into Software Heritage

Thanks to a collaboration between Software Heritage (SWH), HAL-Inria and the CCSD, HAL is opening its doors to a new type of scientific deposit: software. Researchers now have the ability to deposit source code while contributing to Software Heritage, the Library of Alexandria of Software.

Read more

You can "save code now"!

Features

Software Heritage is now open for your source code submissions!
In order to push code to the Archive, go
to archive.softwareheritage.org and click on "Save code now" in the menu on the left side of the page.

Read more

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

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

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 recensions des dépôts, nous faisions 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.
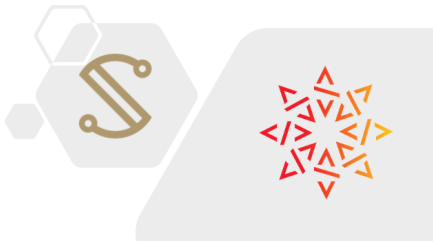
Read more

Software Heritage sponsors meet at UNESCO's headquarters

At a meeting organized on UNESCO's headquarters on February 4th 2020, we have been delighted to bring together over 30 participants, representing the expanding network that supports our mission, to learn about the progress made so far on Software Heritage's implementation and contribute to the discussion on the next steps and strategic directions for the next years.

Read more

Welcoming Sloan support to expand the Software Heritage archive

We are deligthed 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.
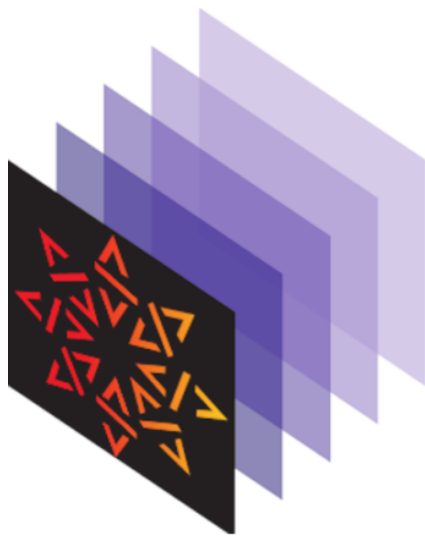
Read more

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

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

November 12, 2020

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.

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.
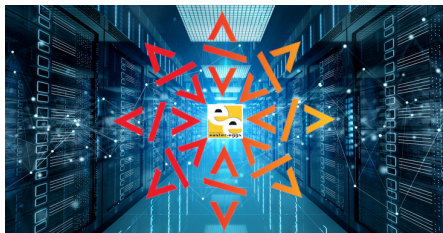
Read more

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.
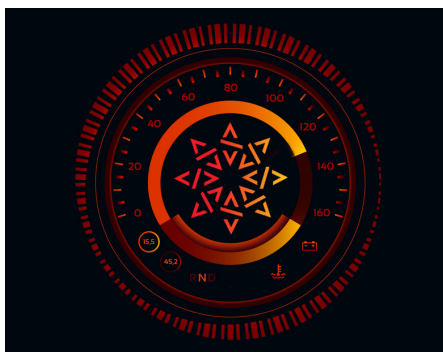
Read more

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!