Issue #14, April 19th, 2024

Software Heritage: 2024 Vision & Milestones Newsletter

"Empowerment of individuals is a key part of what makes open source work, since in the end, innovations tend to come from small groups, not from large, structured efforts."

- Susan Wojcicki

Welcome to the first newsletter of 2024! We're delighted to share the latest updates and insights from Software Heritage. In this edition, we'll cover significant developments and initiatives that have shaped our journey.

Firstly, we're excited to present our 2023 Annual Report, offering a comprehensive overview of the progress and achievements we've made over the past year. From tackling complex technical challenges to making strides in scalability and sustainability, and expanding the community, it's been an eventful period for Software Heritage.

Next, it's a pleasure to announce that all materials are now accessible online from the 2024 Software Heritage Symposium. From session recordings to detailed slide decks, you can catch up on the discussions and insights shared during this high profile annual event.

The Software Heritage 2024 Roadmap marks a pivotal moment in our approach to building a resilient and scalable infrastructure. With a renewed focus on project-oriented development and community-driven features, we're paving the way for future innovation and growth.

The recent 2024 Software Heritage Community Workshop, a prelude to the 2024 Symposium, was a chance for people from various backgrounds – developers, curators, librarians, open source

advocates, and researchers from around the world – to share ideas and work together towards a common goal: ensure the legacy of software for generations to come. The *"Towards Preserving Digital Culture,"* an interview series by Camille Françoise, one of our ambassadors, explores the importance of software preservation from various perspectives. Through insightful conversations with experts, we gain a valuable understanding of the global challenges and opportunities in this field.

Software Heritage's mission is to collect, preserve and share all publicly available software source code, particularly emphasizing Free and Open Source Software (FOSS) as a digital commons embodying decades of human ingenuity. Together with safeguarding this resource for future generations comes the need to enhance its utility for science and software development. The advent of Large Language Models (LLMs) brings significant challenges and opportunities for the software commons: they hold the promise of making it more accessible and reusable, but we are worried to see so many closed models developed in a fully opaque way. We believe LLM development should abide to founding principles that respect and protect the software commons. We published a first version of these principles in October 2023, and we are excited to witness the success of StarCoder2, the first AI model for code aligned with our principles: it shows that *one can build a world class LLM for code while being fully open, transparent on the training data and pipeline, and supporting author opt-out*. Like for software development, there will always be errors, but *transparency makes it easy to find them, and makes it possible to fix them*, which is much more difficult when working with models built behind closed doors.

Lastly, we're proud to announce our participation in the FAIR-IMPACT Open Call, where we're partnering to implement Research Software Metadata Guidelines. Through this initiative, we're working to promote the recognition and preservation of research software, furthering our mission to safeguard software knowledge for future generations.

Stay tuned for more updates and opportunities as we continue our mission to preserve, share, and advance Software Heritage!

2023 ANNUAL REPORT

Software Heritage in 2023: a perspective

We are pleased to announce the release of our annual report for 2023. This document provides a comprehensive overview of Software Heritage's mission, progress, and achievements. It highlights the expansion of our source code collection, user-facing improvements, and community engagement efforts, and expresses gratitude to our sponsors and supporters. We invite you to access and share the report as we continue working together to preserve humankind's software heritage.

Thank you for your continued support!

Read more and download the 2023 annual report

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.
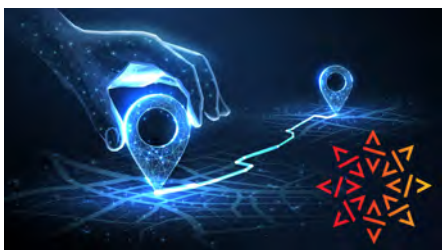

Software Heritage will provide solid, common foundations to serve the different needs of heritage preservation, science, and industry.

Donate

IT'S NEW!

April 9, 2024

The 2024 Roadmap Unveiled: A Vision for Software Heritage

Since the beginning, we have faced complex technical challenges to build a robust, scalable and resilient infrastructure. 2023 has been a decisive year to make significant progress towards scalability and sustainability for the long term.
While we continue our long-term effort on the core infrastructure, 2024 marks a turning point in the way we are building the technical roadmap: *from this year, we are adding a project-oriented layer to guide development and operations, in order to focus on products and features driven by key user communities.*

Read more

March 26, 2024

"Towards Preserving Digital Culture": an interview series by Camille Françoise

Our Software Heritage ambassador [Camille Françoise](#) published the "Towards Preserving Digital Culture" interview series, highlighting experts with different backgrounds and perspectives, to collectively advocate for software preservation. The interview series is published in [Medium](#).

⚠️ 👀 🫣 *Spoiler alert: this addictive series contains also reading recommendations from Camille Françoise's guests.*

Read more

March 21, 2024

The 2024 Software Heritage Community Workshop

On January 31st, 2024, members of the Software Heritage community came together at the Inria research center in Paris, France. This gathering, a prelude to the 2024 Symposium, was a chance for people from various backgrounds – developers, curators, librarians, open source advocates, and researchers from around the world – to share ideas and work together towards a common goal, like an orchestra: where every instrument contributes to the melody. In this blog post, we'll give you a brief overview of what happened at the workshop and highlight some of the key insights that were shared. The 2024 workshop marks its third edition and introduces a full day of sessions for the first time, including several in an unconference style.
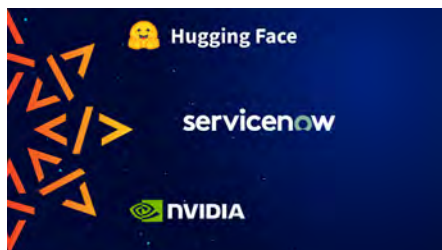
Read more

March 12, 2024

Highlights of the 2024 Software Heritage Symposium

The Software Heritage Symposium 2024 materials are now accessible online: those who couldn't join us in person for our annual event on February 1st, 2024 will now find video recordings of all sessions and the detailed slide decks from our speakers on the event webpage.

Read more

February 28, 2024

Pioneering the Future of Code Preservation and AI with StarCoder2

We recently unveiled our guiding principles for the ethical use of Large Language Models (LLMs) for code, emphasizing openness, transparency, and respect for authors. Today, we celebrate the launch of StarCoder2, the first AI model for code developed in collaboration with Hugging Face, ServiceNow, and NVIDIA. This milestone signifies a step towards lowering barriers to software innovation while upholding ethical AI development principles. It underscores our commitment to advancing technology for the greater good and highlights the ongoing importance of our mission to collect, preserve, and share the wealth of human knowledge embedded in software.

Read more

February 15, 2024

Partnering in the FAIR-IMPACT Open Call to implement the Research Software Metadata Guidelines

We're happy to share with you a new opportunity emerging from the FAIR-IMPACT European project: the Research Software support offer. Software Heritage is a partner in the FAIR-IMPACT project to increase recognition of research software and improve software curation and metadata standardization in the scholarly ecosystem. The current support offer is open to researchers, practitioners, scholarly infrastructures, institutions and anyone interested in research software.

Read more

Join the 2024 Software Heritage Symposium at UNESCO

At the beginning of this year, we announced the Software Heritage Annual Symposium on February 1st, 2024, hosted at UNESCO Headquarters. Now in its third iteration, this exclusive gathering focused on the preservation and sharing of software source code, addressing its impact on digital development, scientific innovation, and cultural heritage.

Read more

AMBASSADORS

We invite you to join our Ambassadors Program**, where we are seeking enthusiastic organizations and individuals to volunteer and play a vital role in expanding the Software Heritage community!**

Become a Software Heritage ambassador to support this effort in your area by:

- presenting Software Heritage, the services it provides, and its long-term mission to key persons in your area;
- sharing information with relevant mailing lists, venues and contacts;
- posting Software Heritage news on blogs and social media;
- identifying code hosting platforms in your area that it is urgent and important to archive;
- updating the Software Heritage team with your community's feedback.

We are welcoming to the ambassadors program:

Wendy Hagenmaier

Baptiste Mélès

We are delighted to introduce our 10th ambassador in 2023, Wendy Hagenmaier! Wendy Hagenmaier is the Software Preservation Program Manager at Yale University Library, where she leads the Emulation-as-a-Service Infrastructure (EaaSI) program to empower the widespread use of emulation for interaction with software, computer systems, collections, and data. Wendy participated in the 2022 SWHAP Days event and 2023 SWHAP Workshop and is happy to become a Software Heritage ambassador. **She believes software preservation and curation are global challenges that require international cooperation, relationship-building, and understanding.**

We are delighted to introduce our 1st ambassador in 2024.
Baptiste Mélès is a researcher at the Henri-Poincaré Archives, in a unit dedicated to philosophy and research on science and technology. Therefore, a significant part of his job is to study source code, using a philosophical, linguistic and historical approach.
If you want to know more about how Software Heritage products can be used in a daily research workflow, then you should ask Baptiste who heavily relies on the archived source code and also chases gems to be preserved.

👉 Read more.

👉 Read more.

Interested in becoming a Software Heritage ambassador?
Please tell us a bit about yourself and your interest in the mission of Software Heritage: info@softwareheritage.org

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, and to add features to the archive to better access it, Software Heritage needs expert contributors.

Open calls for cascading grants may be available to fund your contributions: look at the NLnet

foundation and the recently announced NGI Search, and contact us for a letter of support if you think you can apply with a proposal that contributes to the long term mission of Software Heritage.

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.

Become 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

Host a mirror

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!

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

🎞️ **Discover five years of Software Heritage in just five minutes!** 🎦

July 29, 2021

Software for Open Science: a leap forward

In 2018, the French national plan for Open Science recognized Software Heritage as a key initiative for supporting research software, leading to the creation of a dedicated working group. **Three years later, software is now fully recognized as a pillar of Open Science in the second national plan unveiled by the French Ministry of Research in July 2021.**

👉 Read more.

October 10, 2019

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

We were delighted to make a step forward about an important and urgent challenge in October 2019: collecting and curating landmark legacy source code written by pioneers of the digital age, many of which were still around, and willing to contribute their knowledge, but only for a limited time.

👉 Read more.

February 26, 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

February 26, 2019

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

In April 2017, Inria entered into a framework agreement with UNESCO to uphold our mission of collecting, preserving, and sharing all available source code.
In June 2018, at UNESCO Headquarters, we launched the Software Heritage Archive to the public.

to recall why Software Heritage is building an infrastructure that plays a key role for Open Science.

👉 Read more.

As part of our ongoing partnership, we helped bring together a group of experts at the UNESCO Headquarters and it resulted in the creation of the Paris Call, which now accessible online.

👉 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.

You can read news from Software Heritage on our blog, follow us on 🐘 Mastodon, X, 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!