The great library of source code opens its doors

A year ago, Inria signed an agreement with UNESCO - opening the way to joint actions in favour of the preservation and sharing of software source code, in particular through the Software Heritage project which, for this occasion, launched its website in several language.

https://www.softwareheritage.org/

Today, with the collection work continuing and increasing, and new partners having joined the project, UNESCO and Software Heritage are delighted to celebrate a new and crucial stage together: opening the doors of the biggest archive ever built, making the source code of tens of millions of preserved software programs accessible to all and, on this occasion, reaffirming the common good dimension of open source software.

Software Heritage serving society, science and industry
Software is at the crux of all technological development and therefore plays a pivotal - and even critical - role in our daily lives, our industry and our society.

In order to preserve this heritage and address the technological and scientific challenges of tomorrow, it is essential to build, as of today, a universal and sustainable software archive. That is why Inria, the research institute dedicated to computer sciences - convinced of the vital role of software in the development of the digital society - wanted to take on the challenge by initiating the project in 2015.

The objective of Software Heritage is to collect, organise, preserve and make accessible to all the source code of all available software. By building a universal and sustainable software archive, Software Heritage aims to put in place an essential infrastructure for society, science and industry.

The need to build an infrastructure that aims at indexing all publicly available source code, with a long-term commitment, is clear. It’s not just a matter of preserving old source code, but to help developers of new software projects to find, reuse, reference and archive new source codes.

Providing a single and universal archive making software source code readily available facilitates access to the knowledge contained therein, supports programming education, and creates a reference catalogue with all knowledge about this software - a kind of software Wikipedia. Software Heritage is based on a distributed infrastructure, with international partners, in order to ensure long term availability and reliability of the archive.
Today, Software Heritage is unveiling the contents of the archive

To date, Software Heritage has collected more than 83 million software projects, archiving more than four billion unique source files, as well as all related development history - already making it the world’s leading source code archive.

https://www.softwareheritage.org/archive/

For example, the following are archived there:

• the source code of the Apollo 11 guidance system
• the source code of the NCSA Mosaic Web browser, which popularised the Web
• the source code of the Quake III Arena game
• the source code of the extremely popular AngularJS framework

Software Heritage has also forged a partnership with HAL, the multidisciplinary open archive, aimed at the deposit and dissemination of research-level scientific articles - published or unpublished - and theses, from French or foreign teaching and research establishments and public or private laboratories.

The project is backed by numerous international partners, such as Microsoft, DANS, University of Bologna, Société Générale, Huawei, Nokia Bell Labs and Intel, recently joined by GitHub, Google, UQAM, Qwant, and FOSSID.

Roberto Di Cosmo, the project director, adds: "we are very proud of this stage, which represents a significant amount of work; finally, everybody can consult and download the content of the archive that we have building for the last three years. A great deal of work still lies ahead: find and archive all the source codes we do not have, generalise the deposit of scientific codes, and add functionalities to make the archive easy to use whatever the utilisation. For this, we need new partners and the support of everyone”.

“Softwareheritage.org marks a necessary addition to UNESCO’s longstanding heritage work,” says Audray Azoulay, the Director-General of UNESCO. “The project embodies commitments that are at the heart of our mandate: favouring the free flow of information and culture, and fostering intellectual cooperation for peace and development”.

For example, the following are archived there:

• the source code of the Apollo 11 guidance system
• the source code of the NCSA Mosaic Web browser, which popularised the Web
• the source code of the Quake III Arena game
• the source code of the extremely popular AngularJS framework
Software Heritage partners

Software Heritage was unveiled on 30 June 2016 with two early international partners that had already committed their support to the project in order to help it to grow: Microsoft, one of the largest software industries in the world, and DANS, an institution of the Royal Academy of Arts and Sciences and the Netherlands organisation for scientific research, dedicated to preserving and promoting sustained access to digital research data.

Over the months that followed, many new collaborations have been set up, in particular with Société Générale, Intel, Huawei, Nokia Bell Labs, and the University of Bologna. On 3 April 2017 an agreement was signed with UNESCO, opening the way to common actions in favour of the preservation and sharing of software source code. Since then, other partners have joined Software Heritage, notably GitHub, Google, UQAM, Qwant and FOSSID.

All of the project partners' testimonials are available on the website:

https://www.softwareheritage.org/support/sponsors/

However, it is now possible for everyone to support the Software Heritage project, in particular through online donations on:

https://www.softwareheritage.org/donate

Getting involved with Software Heritage is the unique opportunity to take on a major societal challenge, to support a global initiative and to participate in the preservation of the world's software heritage.

PRESS CONTACT
Laurence GOUSSU - 06 81 44 17 33
Laurence.goussu@inria.fr