Preserving our Software Heritage

Roberto Di Cosmo
Software preservation workshop

Director, Software Heritage
Inria and Université de Paris Cité

September 14th 2023
Outline

1. Software as Heritage
2. Software Heritage’s commitment
3. Working together
“Programs must be written for people to read, and only incidentally for machines to execute.”

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

Apollo 11 source code (excerpt)

```c
float Q_rsqrt(float number)
{
    long i;
    float x2, y;
    const float threehalves = 1.5F;

    x2 = number * 0.5F;
    y = number;
    i = * (long *) &y; // evil floating point bit level hacking
    i = 0x5f3759df - (i >> 1); // what the fuck?
    y = * (float *) &i;
    y = y * (threehalves - (x2 * y * y)); // 1st iteration
    // y = y * (threehalves - (x2 * y * y)); // 2nd iteration, this can be removed

    return y;
}
```

Len Shustek, Computer History Museum 2006

“Source code provides a view into the mind of the designer.”
Experts call for greater recognition of software source code as heritage for sustainable development

UNESCO, Inria, Software Heritage invite 40 international experts meet in Paris ...

“We call to] support efforts to gather and preserve the artifacts and narratives of the history of computing, while the earlier creators are still alive”

"Telling historical stories is the best way to teach. It’s much easier to understand something if you know the threads it is connected to."

Let’s Not Dumb Down the History of Computer Science
Donald E. Knuth, Len Shustek
https://doi.org/10.1145/3442377

A unique opportunity
most of the creators are still here: we can talk to them!
but the clock is ticking…
Outline

1. Software as Heritage
2. Software Heritage’s commitment
3. Working together
Universal software archive

One infrastructure
open and shared

Software Heritage

Largest archive

Technology
- transparency and FOSS
- replicas all the way down

Content (billions!)
- intrinsic identifiers
- facts and provenance

Organization
- non-profit
- multi-stakeholder
Paris Call on Software Source Code

“[We call to] support efforts to gather and preserve the artifacts and narratives of the history of computing, while the earlier creators are still alive”

- **Detailed process** for Legacy Software
  - *curation*
    - reconstruction of the development history
    - collecting metadata
  - *archival*
    - in Software Heritage

- **Traceability** of all process phases
  - using modern version control tools
Paris Call on Software Source Code

“[We call to] support efforts to gather and preserve the artifacts and narratives of the history of computing, while the earlier creators are still alive”

- Expand the SWHAP scope to documents, media (videos, pictures, images, etc.), oral history
- Preserve and Present all this material
- Share process and tools (all open source!) with museums, archives and all interested parties

see this live on the Software Stories website, and get the guide and the SWHAP Days hybrid event, 19 and 20 October 2022
1. Software as Heritage
2. Software Heritage’s commitment
3. Working together
A noble and urgent mission

A great challenge
collecting, curating, preserving and telling the history of landmark software

Software Heritage’s commitment
- provide archival and reference infrastructure
- humbly contribute some tools to the vast preservation landscape
  - the SWHAP source code acquisition and curation protocol
  - the Software Stories approach to interconnect historical data with source code
- help bring together the community of software preservation enthusiast

Thank you for being here, our software legacy needs you!
alone we go faster, together we go further
ancient african saying