

Software Heritage

Building the Universal Software Archive

Roberto Di Cosmo

`roberto@dicosmo.org`

October 18th, 2016



Software Heritage

(Free and Open Source) Software is everywhere



Our source code is spread all around

Debian
Sourceforge
Maven
Bitbucket
GoogleCode
CRAN
Gitorious
Inria
Git Hub
BerliOs
Gitlab
Adullact
CRAN



A word cloud of terms related to source code fragility, set against a background of a world map. The words are arranged in various sizes and orientations. The largest word is 'damage' in a dark red font. Other prominent words include 'disaster' in purple, 'deletion' in blue, 'malicious' in brown, 'obsolete' in purple, 'attack' in blue, 'format' in green, 'reference' in blue, 'storage' in brown, 'dependencies' in blue, 'corruption' in brown, 'encryption' in blue, 'dangling' in brown, 'wear' in brown, 'media' in brown, 'aging' in blue, and 'tear' in purple. The background features a faint world map and a decorative pattern of red and orange triangles on the right side.

damage
disaster
deletion
malicious
obsolete
attack
reference
storage
dependencies
corruption
encryption
dangling
wear
media
aging
tear
format



Software Heritage



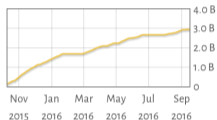
Our mission

Collect, preserve and share the *source code* of *all the software* that is publicly available

The archive (we collect!)

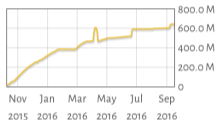
Source files

2,970,266,880



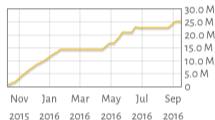
Commits

644,628,800



Projects

25,258,776



- GitHub
- Debian, GNU
- Gitorious, Google Code
- ...

Supporters and sponsors



ACM, Bell Labs, Creative Commons, DANS, Eclipse, Engineering, FSF, OSI, GitHub, GitLab, Huawei, IEEE, Informatics Europe, Microsoft, OIN, OW2, SIF, SFC, SFLC, The Document Foundation, The Linux Foundation, ...

Cultural Heritage



Industry



Research



Education



Software Heritage

Open approach

- open source
- transparency

In for the long haul

- non profit
- replication

Come in, we're open!



Software Heritage

www.softwareheritage.org

@swheritage

Everybody can participate

sponsoring / partnership

working groups, leads

our own code

sponsorship.softwareheritage.org

wiki.softwareheritage.org

forge.softwareheritage.org