

## Des nouvelles de Software Heritage

Le Paris Open Source Summit a été l'occasion de faire le point sur les avancées du projet Software Heritage.



Présenté au public en juin 2016 pour la première fois - nous vous en avons parlé -, Software Heritage, mené par l'Inria et l'Irill, fait un point sur ses avancées. Ce projet très ambitieux vise à collecter et archiver tous les codes sources libres pour les préserver. Autant dire qu'il y a du travail. Mais l'archivage avance et la base de données contient aujourd'hui 200 To de données, 45 millions de projets en ligne et 3 milliards de fichiers. Software Heritage a repris tous les codes hébergés sur Github, mais aussi le projet GNU, les dépôts Debian et à terme également les codes sur Google Code.

Chaque fichier d'un projet reçoit un hash afin de pouvoir l'identifier et le retrouver. Toutes les versions d'un même logiciel sont stockées. Software Heritage possède deux copies de la base plus une troisième stockée sur le cloud de Microsoft, qui est un des principaux partenaires.

### Liens :

- ▶ Le site de Software Heritage <https://www.softwareheritage.org/>
- ▶ La page d'accès aux archives <https://www.softwareheritage.org/ar...>