



Software Heritage, le logiciel comme patrimoine de l'humanité

Annoncé il y a un an, le partenariat entre [Inria](#) et l'Unesco pour préserver et rendre disponible l'intégralité des logiciels existants a été officiellement signé en présence du président François Hollande. 3 milliards de codes sources sont déjà archivés.



A l'Unesco, le 3 avril 2017, Antoine Petit, PDG d'[Inria](#), et Irina Bokova, directrice générale de l'Unesco, signaient la convention de partenariat entre leurs deux institutions pour le projet Software Heritage en présence de François Hollande.

Hamilton-Pool/SIPA



[Visualiser l'article](#)

Au cours de son mandat, le président de la République François Hollande est venu à l'Unesco en juin 2013 pour y recevoir le prix pour la paix Houphouët-Boigny. Puis après le 13 novembre 2015 pour une intervention devant le Forum des dirigeants à la suite des attentats parisiens. La raison de sa troisième visite, lundi 3 avril, était plus inattendue. Il est venu assister à la signature officielle d'un partenariat entre l'organisation internationale et d'[Inria](#) ([Institut de recherche en informatique](#)). L'occasion pour le président de glisser dans son discours quelques allusions à la campagne présidentielle en cours et à l'usage d'Internet pour faire circuler mensonges et rumeurs. L'occasion surtout pour les deux partenaires de célébrer en grande pompe un travail annoncé il y a un an : la création d'une « *bibliothèque d'Alexandrie du logiciel* » selon la formule d'Alexandre Petit, PDG d'[Inria](#), pour préserver le patrimoine logiciel de l'humanité au titre du patrimoine immatériel.

Appelé Software Heritage, ce projet consiste concrètement à retrouver, collecter puis stocker sur des serveurs tous les codes sources produits par l'humanité, et les rendre accessible en ligne (Microsoft participe au projet avec son service Azure) depuis un site Internet en plusieurs langues ouvert ce même 3 avril. « *L'idée est de pouvoir conserver une archive du code source* », continue le PDG d'[Inria](#). *Certaines entreprises ne possèdent plus qu'un fichier .exe de leurs logiciels. Outre la préservation, ce sera aussi un outil de vérification, pour les travaux d'étudiants, et de référencement.* »

D'abord les logiciels libres, à terme les logiciels propriétaires

Actuellement, le projet en est à 55 millions de logiciels pour 3 milliards de fichiers sources. Car c'est le code source « *la partie importante, celle qui contient la véritable connaissance* », affirme Roberto Di Cosmo, directeur de Software Heritage. Pour l'instant, le projet se concentre sur les logiciels libres, qui seront consultables par tout le monde. Mais les logiciels propriétaires sont tout autant concernés, même si eux ne seront accessibles qu'à des personnes autorisées et sous condition.

Pour Indrajit Banerjee, directeur de la Division de la société de l'information à l'Unesco, cette entreprise participe pleinement du mandat de l'institution, censée travailler à un « *accès universel à l'information et au savoir. Or, sans logiciel aujourd'hui, on n'a accès à rien.* » Bien sûr, il existe déjà des plateformes en ligne proposant des logiciels en accès libre (Debian, Github, GNU, SourceForge). Mais aucune entrée unique informant de manière uniforme quel code a été développé par qui. Il faut donc explorer ces bibliothèques, réapprendre à chaque fois un langage technique différent. Sans compter que certaines disparaissent purement et simplement (Gitorious en mars 2015, Google Code début 2016). « *Qui sait ce que deviendront ces plateformes dans dix ans* », continue Roberto Di Cosmo. *Le logiciel est un objet fragile : il se casse, peut être corrompu, piraté. Il existe donc un vrai danger de perte.* » C'est pourquoi, même si Software Heritage a commencé à rendre possible le téléchargement, à l'usage des développeurs, la priorité actuelle reste à la récupération des codes sources.