



L'entreprise française Software Heritage va préserver le logiciel comme patrimoine immatériel



La signature de la convention de partenariat autour de Software Heritage entre [Antoine Petit](#), le PDG d'Inria, et la directrice générale de l'Unesco, Irina Bokova, en présence du président de la République, François Hollande. - Camille Le Hyaric

L'Unesco et l'Institut national de la recherche en informatique ont signé un partenariat autour de la plateforme française d'archivage Software Heritage, « la bibliothèque d'Alexandrie » du logiciel.

Les logiciels sont partout : dans nos ordinateurs, nos smartphones, désormais nos montres connectées ou encore nos voitures. Pourtant, aussi omniprésents soient-ils, « on ne s'aperçoit pas vraiment du patrimoine qu'ils recèlent. Celui de la connaissance humaine qu'il faut pour les créer », assure Roberto Di Cosmo, fondateur et directeur de Software Heritage. Cette plateforme française d'archivage s'est fixée comme objectif, en juin 2016, de recenser et conserver les codes sources de tous les logiciels disponibles sur Internet.

Une initiative soutenue et financée dès son commencement par l'Institut national de la recherche en informatique (Inria) et qui a pris, cette semaine, une nouvelle dimension. L'Unesco a accepté de signer avec l'Inria un partenariat pour soutenir Software Heritage. L'organisation internationale reconnaît ainsi le logiciel libre comme patrimoine immatériel de l'humanité.



« L'Unesco a tout de suite compris la valeur de notre initiative, assure Roberto Di Cosmo. C'est un accord-cadre qui a été signé en présence de François Hollande, et qui va permettre de mener à bien ce projet de conservation et de mise à disposition des logiciels. »

Microsoft pour partenaire

Si l'Unesco a accepté ce partenariat, c'est probablement parce que Roberto di Cosmo n'est pas venu les mains vides. Son équipe et lui, avec l'appui de l'Inria, avaient déjà récolté et archivé pas moins de « 3 milliards de codes sources uniques » et plus de « 58 millions de projets de logiciels ». « Mais notre mission, aujourd'hui, va plus loin que ça », explique le directeur de Software Heritage. Il faut maintenant à la fois continuer à récupérer et à archiver les logiciels mais aussi permettre de les comprendre, les retrouver et s'orienter sur la plateforme. L'objectif est que chacun puisse venir, comme dans une bibliothèque, récupérer les codes sources dont il a besoin pour mener à bien un projet.

C'est aussi ce qui a plu aux partenaires de Software Heritage, dont certains sont de grandes entreprises comme Microsoft, Intel ou encore la Société Générale et Huawei. Mais aussi l'université de Bologne (Italie), le gouvernement français, et l'organisation à but non lucratif Creative Commons... La plateforme fonctionne sur les principes même du logiciel libre : ouverte à tous, elle doit devenir participative.

Quant à la présence de grandes entreprises parmi les partenaires qui pourraient profiter d'un tel projet ? « Cela ne me dérange pas du tout », affirme Roberto di Cosmo. Auteur de plusieurs livres dénonçant le monopole de certaines entreprises sur le monde des logiciels, il reconnaît que celles-ci ont changé. « Aujourd'hui, 10.000 ingénieurs travaillent chez Microsoft uniquement à la création de logiciels libres et ouvert à tous, constate-t-il. C'est le plus gros contributeur au monde. » Il était donc important pour Software Heritage de l'avoir comme partenaire.

Un projet, deux étapes

« Mais nous avons imaginé le projet sur le modèle du World Wide Web Consortium (W3C), afin qu'aucune entreprise ne puisse avoir la main mise sur ce qui doit être archivé ou non », précise Roberto di Cosmo. Le directeur préfère ainsi avoir de nombreux partenaires qui contribuent à hauteur d'un montant maximum plutôt qu'un seul très gros qui donnerait beaucoup et aspirerait à prendre des décisions importantes.

Car, in fine, Software Heritage veut pouvoir s'adresser à quatre types de publics : les entreprises, certes, mais aussi les scientifiques, les conservateurs du patrimoine, ainsi que les professeurs et leurs élèves. Un projet qui se déroulera en deux étapes.

Roberto di Cosmo estime ainsi qu'il faudra encore 3 à 4 ans pour archiver tous les logiciels déjà présents sur Internet et convaincre les plateformes qui hébergent des nouveaux codes sources de les lui transmettre. Une première partie chiffrée à 30 millions d'euros. Puis d'ici 5 à 10 ans, l'objectif sera de rendre l'ensemble « disponible, traçable et uniforme », afin que tout le monde voit ce qu'il y a dans Software Heritage et que les développeurs puissent s'en servir. Cette deuxième partie devrait coûter entre 20 et 30 millions d'euros.

Une bibliothèque d'Alexandrie moderne

La signature du partenariat avec l'Unesco a aussi été l'occasion de prendre date. En juin 2017 dans un premier temps, pour une session plénière autour du logiciel, en présence de l'Inria, de l'Unesco mais aussi des partenaires de Software Heritage. Puis une conférence internationale afin de travailler sur la question de la conservation et de la mise à disposition des logiciels, qui aura lieu à la rentrée prochaine.



Pouvoir travailler directement avec ceux qui ont créé certains des logiciels les plus importants

« Le logiciel est une connaissance qui a pris naissance il y a à peine 50 ans, explique Roberto Di Cosmo. C'est la chance pour nous de pouvoir travailler directement avec ceux qui ont créé certains des logiciels les plus importants. » Et ce passionné de code de prendre l'exemple du PHP. À l'origine, quelques lignes de codes écrites par Rasmus Lerdorf pour conserver une trace des visiteurs sur la page de son CV, avant qu'il ne le mette à disposition de tous en 1995. Deux autres programmeurs l'améliorent encore un peu plus et PHP devient l'outil le plus important permettant de faire fonctionner des sites comme Facebook , Wikipédia, et la plupart des sites d'information...

« Personne n'aurait pu prévoir, en 1995, que ces quelques lignes de code allaient devenir si importantes », s'enthousiasme Roberto di Cosmo, justifiant ainsi la nécessité d'archiver et de conserver les logiciels. Un petit logiciel qui aurait pu disparaître si le serveur sur lequel Lerdorf l'avait publié avait dû fermer ou avait été détruit, « comme une bibliothèque qui brûle ». Avec Software Heritage, ce risque ne devrait plus exister. Les codes sources de tous les logiciels seront désormais sauvegardés dans cette archive géante, forme moderne de la bibliothèque d'Alexandrie.