



Software Heritage nous détaille la convention entre Inria et l'UNESCO, une étape importante

Visuel indisponible

Software Heritage se définit comme une bibliothèque de code source. Elle propose aujourd'hui plus de trois milliards de fichiers, représentant près de 60 millions de projets. En recherche de reconnaissance, le projet a été distingué par l'UNESCO. Son directeur Roberto Di Cosmo revient avec nous sur ce partenariat.

L'été dernier, Inria présentait officiellement un vaste projet qui avait été lancé en 2015 par l'Institut national de recherche en informatique et en automatique (Inria) : Software Heritage. Son ambition est de « collecter, organiser, préserver, et rendre accessible à tous le code source de tous les logiciels disponibles ». Protéger notre patrimoine logiciel représente donc un travail titanesque.

Dès le lancement, plusieurs partenaires sont rapidement arrivés, Microsoft et le DANS de la Royal Academy des Pays-Bas. Aujourd'hui, ce projet passe un nouveau cap puisqu'Inria a signé une convention de partenariat avec l'UNESCO. Contacté par nos soins, Roberto Di Cosmo – directeur de ce projet – nous explique en quoi c'est une étape importante.

Sur trois piliers, il en manquait un

Il nous explique que Software Heritage repose sur « trois piliers » fondamentaux : science et technologie qui est principalement fournie par Inria, des ressources (argents, temps ingénieur, stockage, etc.) amenées par des partenaires et enfin « la reconnaissance de l'importance de la mission ».

Si les deux premiers étaient déjà présents dès l'annonce officielle, le troisième manquait cruellement à l'appel. Après un « long long combat pour le faire reconnaître », cette signature de partenariat avec l'UNESCO permet à Software Heritage de prendre de l'importance dans les « dimensions politiques et sociétales ». Roberto Di Cosmo ne cache d'ailleurs pas son enthousiasme et sa surprise : « on ne s'attendait pas à un truc pareil », lâche-t-il.

Concrètement, ce partenariat avec l'UNESCO débouchera prochainement sur des interventions communes, une première de deux jours étant planifiée pour septembre. Le but sera d'expliquer pourquoi il est important de sauvegarder le code source des logiciels.

Visuel indisponible

De nouveaux partenaires, certains prestigieux

Dans tous les cas, Software Heritage n'a pas attendu que l'UNESCO se penche sur son cas pour avancer. Depuis l'annonce officielle du projet l'été dernier, il a été rejoint par Intel Corporation, qui a mis 100 000 euros sur la table, ainsi que par la Société Générale aussi avec 100 000 euros.

D'autres contributions plus modestes sont également de la partie avec le chinois Huawei (25 000 euros), l'Université de Bologne et Nokia Bell Labs. De deux, il est désormais question de sept partenaires. D'autres peuvent évidemment s'ajouter à la liste.

Si Inria et la Société Générale sont deux groupes français, ce n'est pas le cas des autres partenaires qui viennent du reste de l'Europe, d'Asie et des États-Unis. Cette ouverture est d'ailleurs une bonne chose pour

[Visualiser l'article](#)

Roberto Di Cosmo, qui nous explique qu'« *il ne faut pas que ce soit perçu comme un projet franco-français* », afin qu'il puisse se développer à l'international.

Au niveau du développement, Software Heritage a également des contributions extérieures. Certains viennent simplement donner un coup de main, d'autres s'investissent davantage. C'est notamment le cas de l'ADULLACT (Association des développeurs et utilisateurs de logiciels libres pour les administrations et les collectivités territoriales) qui a développé un plug-in complet pour ce projet.

Si l'aventure vous tente, vous trouverez tous les détails pour contacter l'équipe en charge du développement par ici . Bien évidemment, l'ensemble du code de Software Heritage est open source et disponible par là .

Plus de trois milliards de fichiers et près de 60 millions de projets

Aujourd'hui, Software Heritage revendique une base de données comprenant plus de 3,2 milliards de fichiers uniques représentant 58 254 852 projets. Concernant ce dernier chiffre, on peut remarquer qu'il a fait un bond de plus de 50 % entre 24 octobre et le 3 novembre 2016, alors que le nombre de commits et de codes sources n'a que peu changé.

Le directeur du projet nous explique que ce changement est intervenu lorsque « *chaque fork était compté comme un projet de plus* ». Le nombre de projets explose logiquement, mais comme chaque code source n'est sauvegardé qu'une seule fois – tout est dédoublé –, le nombre de fichiers ne suit donc pas la même tendance.

Visuel indisponible

Software Heritage « c'est aussi le futur »

Enfin, Roberto Di Cosmo nous indique que Software Heritage ne veut pas seulement être une sauvegarde du passé, mais aussi avoir les yeux tournés vers le futur. Pour lui, le but est de « *savoir ce qui se passe dans le monde logiciel* », d'où nous venons et où nous allons ; le tout dans un endroit unique.

Si Software Heritage propose notamment le code d'Apollo 11 (disponible sur GitHub depuis juillet 2016), tout le monde n'a pas une fusée Apollo sous la main pour le mettre en pratique lâche-t-il sur le ton de l'humour. Le directeur présente ce projet comme une grande bibliothèque : l'équipe ajoute des livres dans les rayons, au public maintenant de venir les lire.

Le partenariat avec l'UNESCO pourrait ainsi être perçu comme un premier signe d'ouverture vers le grand public. L'avenir nous le dira.