



SCIENCES

L'Inria crée la bibliothèque d'Alexandrie des logiciels libres

Cette initiative vise à regrouper plusieurs dizaines de millions de programmes informatiques.

MARC CHERKI mcherki@lefigaro.fr

NUMERIQUE Faible coût, mais grandes ambitions. Pour 1,5 million d'euros sur trois ans, y compris le salaire des quatre chercheurs et de deux étudiants qui travaillent sur le projet depuis dix-huit mois, l'Institut national de la recherche en informatique et en automatique (Inria) a décidé de créer une sorte de bibliothèque d'Alexandrie de tous les logiciels libres, dont le code source est disponible publiquement (dit « open source »), depuis la naissance de l'informatique.

Le logiciel est aujourd'hui au cœur de toutes les activités humaines, explique en substance Roberto Di Cosmo, directeur de ce projet baptisé « Software Heritage » (accessible via le site software-heritage.org). Mais ces connaissances sont fragiles et éphémères : elles évoluent au gré des langages de programmation, de l'arrivée de nouveaux outils, tels les téléphones mobiles, dont la puissance de calcul dépasse celle des ordinateurs d'il y a dix ans, ou de nouveaux objets connectés. De plus, les logiciels sont utilisés dans toutes les disciplines scientifiques, y compris les sciences humaines, pour bâtir des scénarios, conduire des recherches et superviser ou faire

fonctionner des appareils.

L'idée est donc de créer « une sorte de grand instrument à vocation mondiale. Mais, à terme, l'Inria va s'effacer » derrière sa création, indique Antoine Petit, le PDG de l'Inria, qui espère tout de même que le logo de l'organisme reste associé à cette initiative. Le logiciel est un « bien commun », renchérit Jean-François Abramatic, directeur de recherche à l'Inria et ancien patron du World Wide Web Consortium, la principale organisation d'Internet qui a contribué à créer des langages majeurs (html, xml, etc.). Jean-François Abramatic est l'un des trois conseillers scientifiques de Software Heritage, aux côtés de ses collègues Serge Abiteboul, professeur à l'École normale supérieure, et Gérard Berry, professeur au Collège de France.

Surtout, cette initiative se justifie car de nombreux sites, souvent américains, avaient décidé de jouer ce rôle pour héberger les logiciels écrits collectivement, utilisables et modifiables par tous... jusqu'à ce que ces entreprises décident de clore cette activité par simple décision stratégique. « Il y a un an, Google a fermé sa plateforme GoogleCode, où il y avait



1,5 million de projets, en prévenant les contributeurs qu'ils avaient trois semaines pour récupérer leurs codes», rappelle Roberto Di Cosmo.

Le but de Software Heritage est donc de «collecter, organiser, préserver et partager les codes sources de tous les logiciels». Cette bibliothèque pourra aider à gagner un temps considérable au développement d'applications de recherche, car il n'y aura pas la nécessité de recommencer à écrire ce qui a déjà été bien fait par d'autres... Cette bibliothèque devrait également améliorer la qualité des travaux, car les chercheurs pourront comparer plus facilement leurs approches.

Software Heritage fédère déjà plus de 22 millions de projets logiciels, qui ont donné lieu à près de 600 millions de modifications et à 2,7 milliards de fichiers. Ces données occupent quelque 200 téraoctets (l'équivalent de 200 disques durs d'ordinateurs personnels).

Les logiciels sont hébergés par l'Inria et seront également dupliqués, en Europe, par Microsoft, converti au logiciel li-

«Collecter, organiser, préserver et partager les codes sources de tous les logiciels»

LA MISSION DE SOFTWARE HERITAGE

bre et partenaire du projet. Toutes les autres entreprises informatiques au monde pourront contribuer à cette initiative majeure. L'institut Dans (Data Archiving and Networked Services) de l'Académie royale des arts et des sciences des Pays-Bas a décidé d'y participer. Bien sûr, Software Heritage, qui doit se muer en fondation, est ouvert à toutes les institutions et aux centres de recherches.

Maintenant que l'Inria a démontré que le projet est réalisable, il reste encore des détails techniques à finaliser pour permettre à tous d'accéder, en lecture seule, aux nombreux programmes qui y seront archivés. Cette prochaine phase devrait aboutir dans six mois. Plus tard, les programmes pourront aussi être directement téléchargés. Déjà, il est possible d'interroger cette base pour savoir si tel logiciel y est présent ou non. Un autre prolongement de cette bibliothèque

pourra être celui du «dépôt légal» : un auteur (ou une entreprise) ne pourra plus revendiquer la paternité d'un logiciel écrit par d'autres auparavant... ■



L'Institut national de la recherche en informatique et en automatique (Inria), à Paris.

VINCENT BOISOT
POUR LE FIGARO