



COMMUNIQUÉ DE PRESSE

Paris, le 30 juin 2016



Collecter, organiser, préserver et partager le patrimoine logiciel de l'humanité

Inria annonce aujourd'hui l'ouverture de Software Heritage, une initiative ambitieuse qui a pour objectif de collecter, organiser, préserver, et rendre accessible le code source de tous les logiciels disponibles publiquement.

Pour échanger des messages avec notre famille et nos amis, payer des factures, faire des achats, accéder au divertissement, faire des démarches administratives, trouver des informations, ou planifier des voyages, pratiquement tous nos actes quotidiens dépendent de logiciels exécutés par des ordinateurs.

Mais ce n'est que la partie émergée de l'iceberg : le logiciel contrôle le système embarqué dans nos moyens de transport ou de communication, les échanges commerciaux et financiers. Il est au cœur des équipements et des dispositifs médicaux ; il assure le bon fonctionnement des réseaux de transport et de communication, des banques et des établissements financiers.

Le logiciel est crucial dans le fonctionnement des organisations, qu'elles soient publiques ou privées, que ce soit sur des terminaux mobiles ou sur le « cloud ».

En un mot, le logiciel joue un rôle central dans les sociétés modernes pour notre vie quotidienne, notre industrie et notre société.

Un défi sociétal, technique et scientifique

Inria dévoile aujourd'hui le projet **Software Heritage**, dont l'objectif est de construire à la fois une moderne « bibliothèque d'Alexandrie » du logiciel, un référentiel unique du code source et un grand instrument de recherche pour l'Informatique : un défi sociétal, technique et scientifique.

Software Heritage est une composante essentielle pour préserver, développer et diffuser la connaissance qui se trouve aujourd'hui encodée dans le logiciel, ainsi que notre capacité d'accéder à l'ensemble de l'information numérique.

Software Heritage s'appuiera sur une infrastructure distribuée, de manière à garantir la robustesse et la disponibilité.

Software Heritage constitue le socle pour une base de données référençant tous les codes sources des logiciels utilisés dans l'industrie, garantie d'une meilleure gestion du cycle de vie et d'une accessibilité à long terme de la production logicielle industrielle. Lorsqu'il bénéficiera de

mises à jour en temps réel, **Software Heritage** deviendra le catalogue de code de référence pour tous les utilisateurs industriels, permettant ainsi aux développeurs de nouveaux logiciels de trouver, ré-utiliser, et archiver de nouveaux codes source.

Software Heritage est la fondation sur laquelle nous pouvons construire un grand instrument de recherche en informatique qui permettra des avancées significatives pour la qualité, la sûreté de fonctionnement et la sécurité des logiciels que nous utilisons tous les jours.

Software Heritage aujourd'hui

Inria, organisme de recherche dédié aux sciences du numérique, a décidé de relever le défi, en mettant en place une équipe dédiée avec les moyens nécessaires pour démarrer le projet.

A ce jour, **Software Heritage** a déjà collecté plus de **20 millions de projets logiciels**, avec plus de **deux milliards et demi de fichiers sources uniques archivés**, ainsi que tout l'historique de leur développement, ce qui en fait d'ores et déjà **l'archive de code source la plus riche de la planète**.

Pour **Antoine Petit, président directeur général d'Inria**, « *nous avons décidé de lancer Software Heritage il y a plus d'un an, et nous avons montré sa faisabilité. Afin de le déployer à l'échelle mondiale, il est temps maintenant d'ouvrir le projet à la contribution la plus large, nationale et internationale* ».

Les premiers partenaires s'engagent

Deux premiers partenaires internationaux se sont déjà engagés pour soutenir le projet et l'aider à grandir : Microsoft, l'un des plus grands industriels de l'informatique dans le monde, et DANS, une institution de l'Académie Royale des Arts et des Sciences et l'organisation néerlandaise dédiée à la préservation et à la promotion de l'accès aux données des recherches sur le numérique.

"Microsoft a été impliqué dans des initiatives open source en facilitant, en intégrant, et en contribuant à de nombreux projets et communautés open source depuis plus d'une décennie", explique **Jean Paoli, General Manager chez Microsoft Corp**. «*Nous applaudissons le projet Software Heritage comme un projet ouvert qui contribuera à organiser et conserver la connaissance sous la forme de code source pour les générations futures, et aidera les générations actuelles de développeurs à trouver et à réutiliser le code dans le monde entier. Nous sommes fiers d'être l'un des premiers partenaires industriels à soutenir cette initiative et de mettre à disposition l'infrastructure Cloud Azure pour contribuer à assurer la robustesse et la disponibilité des données*».

« *La pérennité des logiciels est un des éléments essentiels pour accélérer la recherche. La reproductibilité des résultats scientifiques, mais aussi et la possibilité de retrouver et réutiliser les logiciels utilisés sont des principes fondamentaux de la science que DANS soutient fermement* », selon **Peter Doorn, Directeur de DANS**, « *nous sommes engagés dans la préservation à long-terme des informations scientifiques, et nous sommes ravis d'être le premier partenaire académique international à rejoindre l'appel de Software Heritage* ».

Tous concernés, tous contributeurs

Software Heritage a déjà reçu le soutien de scientifiques, d'industriels, de sociétés savantes, de fondations, d'organisations indépendantes et institutionnelles.

Après avoir lancé le projet, avoir démontré sa faisabilité et avoir établi les premiers partenariats, Inria lance dès aujourd'hui un appel mondial à tous les contributeurs à rejoindre cette initiative et ouvre le site :

<https://www.softwareheritage.org/>

Software Heritage est un projet ambitieux, qui nécessite la contribution d'un grand nombre d'acteurs :

- Pour collecter tout le code source disponible : nous aider à identifier les milliers de sites disparates dans lesquels se retrouve aujourd'hui éparpillé le patrimoine logiciel mondial.
- Pour contribuer au développement : l'équipe de **Software Heritage** a une longue tradition de collaboration et est bien connue dans le monde du logiciel libre ; dans les prochains jours, nous allons ouvrir notre propre code source au monde entier et nous serons ravis d'accueillir les développeurs qui voudront partager leur vision et nous aider dans cette mission.
- La résolution des défis scientifiques posés par la construction d'une archive universelle du code source à partir d'informations dispersées va nécessiter de nouvelles contributions ; la participation des chercheurs de toutes les disciplines sera déterminante pour réussir.
- Préserver sur le long terme et partager à travers le monde le contenu de cette archive nécessite des ressources conséquentes, sur un plan humain, matériel et financier, ainsi que des partenaires internationaux.

S'engager avec **Software Heritage**, c'est l'occasion unique de relever un défi sociétal majeur, de soutenir une initiative d'envergure mondiale et de participer à la préservation de notre patrimoine logiciel mondial.

A propos d'Inria

Inria, institut national de recherche dédié au numérique, promeut « l'excellence scientifique au service du transfert technologique et de la société ». Inria emploie 2700 collaborateurs issus des meilleures universités mondiales, qui relèvent les défis des sciences informatiques et mathématiques. Son modèle ouvert et agile lui permet d'explorer des voies originales avec ses partenaires industriels et académiques. Inria répond ainsi efficacement aux enjeux pluridisciplinaires et applicatifs de la transition numérique. Inria est à l'origine de nombreuses innovations créatrices de valeur et d'emplois.

Contacts presse :

Laurence Goussu – laurence.goussu@inria.fr
01 39 63 57 29 – 06 81 44 17 33

Hélène Latour – helene.latour@meanings.fr
07 60 46 66 24



DOSSIER DE PRESSE

Paris, le 30/06/2016



Collecter, organiser, préserver et partager le patrimoine logiciel de l'humanité

<https://www.softwareheritage.org/>

Software Heritage a pour objectif de **collecter, organiser, préserver, et rendre accessible à tous** le code source de tous **les logiciels** disponibles, un enjeu de portée mondiale. En construisant une archive universelle et pérenne du logiciel, **Software Heritage** vise à mettre en place une **infrastructure essentielle** au service de la société, de la science et de l'industrie.

Le logiciel au cœur de notre société

Pour échanger des messages avec notre famille et nos amis, payer des factures, faire des achats, accéder au divertissement, faire des démarches administratives, trouver des informations, ou planifier des voyages, pratiquement tous nos actes quotidiens dépendent de logiciels exécutés par des ordinateurs.

Mais ce n'est que la partie émergée de l'iceberg : le logiciel contrôle le système embarqué dans nos moyens de transport ou de communication, les échanges commerciaux et financiers. Il est au cœur des équipements et des dispositifs médicaux ; il assure le bon fonctionnement des réseaux de transport et de communication, des banques et des établissements financiers.

Le logiciel est crucial dans le fonctionnement des organisations, qu'elles soient publiques ou privées, que ce soit sur des terminaux mobiles ou sur le « cloud ».

Le logiciel est au cœur de tout développement technologique et est devenu indispensable pour la recherche scientifique, dans tous les domaines. Il joue donc un rôle central et même critique dans notre vie quotidienne, notre industrie et notre société.

La transformation numérique pousse à convertir en représentation numérique tout aspect de nos connaissances et de notre patrimoine culturel, mais comme le dit **Vinton Cerf, co-créateur de l'Internet**, « *sans le logiciel capable d'interpréter ces représentations, toutes ces connaissances ne seraient que des suites inintelligibles de bits sans valeur* ».

Afin de préserver ce patrimoine et de répondre aux défis technologiques et scientifiques de demain, il est essentiel de construire dès aujourd'hui une **archive universelle et pérenne** du

logiciel.

Software Heritage au service de la société, de la science et de l'industrie

Le rôle central du logiciel dans les sociétés modernes fait de Software Heritage une infrastructure de portée sociétale universelle.

Préserver le logiciel, rendre accessible la connaissance numérique



Les logiciels que nous utilisons forment une partie importante de notre culture moderne, et les logiciels libres y jouent un rôle majeur. Jusqu'ici un grand nombre de projets logiciels ont été développés et distribués en utilisant des infrastructures variées et disparates qui changent souvent et ne sont pas pérennes. Il s'agit d'un pari risqué, et cela est devenu évident il y a quelques mois seulement, quand deux dépôts importants, qui contenaient des centaines de milliers de projets

logiciels, ont fermé leur porte avec un bref préavis: Gitorious et Google Code.

Aujourd'hui, le besoin de construire une infrastructure capable de référencer tout le code source publiquement disponible et qui s'inscrit dans la durée est clair. L'enjeu n'est pas seulement de préserver les anciens codes sources, mais d'aider les développeurs de nouveaux projets logiciels à trouver, réutiliser, référencer et archiver les nouveaux codes sources. **Software Heritage** s'est donné cet objectif.

Mettre à disposition une archive unique et universelle rendant disponibles les codes source de ces logiciels facilitera l'accès à la connaissance qu'ils contiennent, soutiendra la formation à la programmation, et de créera un catalogue de référence avec l'ensemble des connaissances sur ces logiciels, une sorte de Wikipedia du logiciel. **Software Heritage** s'appuiera sur une infrastructure distribuée, avec des partenaires internationaux, de manière à garantir la robustesse et la disponibilité de l'archive.

Préserver et partager le logiciel pour la Science

Une très grande partie du savoir scientifique est produite par du logiciel, ou contenue dans du logiciel. Pourtant, les logiciels scientifiques sont peu et mal archivés, mal ou non référencés, au point que la reproductibilité des expériences scientifiques qui les utilisent n'est pas assurée. Pour préserver la connaissance technologique et scientifique, il est maintenant impératif de pouvoir garder non seulement les articles publiés, mais aussi le code source des logiciels décrits dans les articles scientifiques et utilisés pour traiter les données.



La recherche en Informatique a fait de gros progrès. Il est maintenant possible d'identifier automatiquement des erreurs de programmation, des défauts de conception ou des attaques informatiques. Pourtant, à la différence des disciplines comme la physique, elle ne dispose pas encore de grands instruments de recherche permettant d'explorer systématiquement l'ensemble des codes source de tous les logiciels.

Software Heritage constitue une infrastructure de recherche fondamentale dont l'objectif est de devenir la référence pour la mise à disposition des logiciels utilisés dans toute la recherche scientifique, complétant la démarche de l'Open Access, qui vise à rendre les articles scientifiques librement accessibles pour tous.

Software Heritage constitue la fondation sur laquelle nous pourrons construire un **grand instrument de recherche** sur l'ensemble des codes source des logiciels, apportant des progrès significatifs dans tous les domaines des sciences du numérique, dans un objectif

d'amélioration de la qualité, de la sûreté de fonctionnement et de la sécurité des logiciels que nous utilisons dans notre vie quotidienne.



Améliorer la qualité des logiciels pour l'industrie

La sûreté et la qualité logicielles sont cruciales dans de nombreux secteurs industriels comme l'aérospatiale, la médecine, l'énergie, le transport, la sécurité, qui sont au cœur de notre société. Actuellement, les logiciels industriels sont développés en assemblant de nombreux composants, dont il est impératif d'assurer la **traçabilité** et la **qualité**.

Software Heritage constitue une base de données de référence pour tous les logiciels open-source, garantie d'une meilleure gestion du cycle de vie des logiciels et d'une accessibilité à long terme pour le développement des logiciels industriels présents et futurs. L'ensemble des artefacts déposés dans **Software Heritage** obtiendront une référence unique et persistante servant de base pour les services liés à la sécurité, la certification, la traçabilité et la qualification du code. Lorsqu'il bénéficiera de mises à jour en temps réel, Software Heritage deviendra le catalogue de code source de référence pour tous les utilisateurs industriels.

Software Heritage aujourd'hui

Ce projet a été initié par **Inria**, institut de recherche dédié au numérique et acteur incontournable du logiciel depuis de nombreuses années, convaincu du rôle essentiel du logiciel dans le développement de la société numérique. Inria a souhaité relever le défi, en mettant en place une équipe dédiée et les moyens nécessaires pour démarrer le projet.



Comme le rappelle **Antoine Petit, Président-directeur général d'Inria**, « *les logiciels sont aujourd'hui au cœur de la plupart des activités humaines, de la médecine aux loisirs, des avions à l'agriculture... Ce sont eux qui rendent "intelligents" les transports, les villes, les habitats ou encore les usines. Ils jouent ainsi un rôle central et même critique dans notre vie quotidienne. Les chercheurs d'Inria, institut de recherche dédié aux sciences du numérique, ont produit depuis près de 50 ans plusieurs milliers de logiciels, dont certains ont permis des avancées scientifiques majeures et d'autres sont à la base de produits commerciaux, en particulier dans les 130 spin-offs issues d'Inria. Il est donc légitime pour Inria de se soucier de la préservation de toute la connaissance technologique et scientifique liée au logiciel, d'aider à construire et conserver le patrimoine mondial des logiciels, et de le mettre au service de la société, de l'industrie, de la science et de l'éducation* ».

Pour **Roberto Di Cosmo, Directeur de Software Heritage**, « *La connaissance technologique de l'humanité est indissociable des logiciels : toute notre société et toute notre culture en dépendent. Le code source des logiciels est précieux : il s'agit d'une forme de connaissance qui est, de par sa conception, à la fois exécutable par des machines, et compréhensible par des êtres humains. Nous avons le devoir de préserver cette connaissance. Soutenir Software Heritage, c'est prendre part à un projet essentiel qui préserve le passé pour préparer le futur.* »

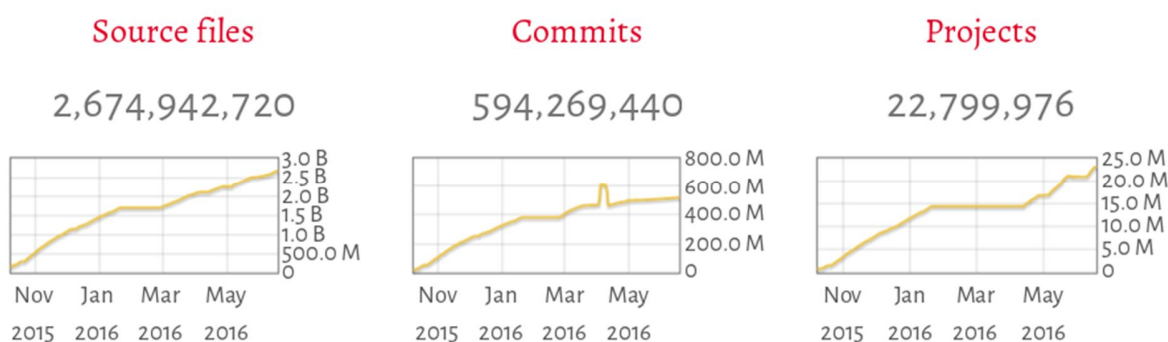
Projet d'envergure mondiale, **Software Heritage** est confronté à de nombreux défis.

Organiser et indexer de façon efficace l'ensemble de tout le code source disponible, pour le rendre plus facilement accessible, et pour en extraire des informations utiles afin d'en améliorer la qualité pose des nombreuses questions scientifiques.

D'un point de vue technique, il s'agit par exemple d'assurer la pérennité des données par une réplication massive de l'infrastructure, de fournir des références stables à chaque composant logiciel, et de prendre en compte les spécificités du code source du logiciel, qui est différent des autres objets numériques.

D'un point de vue stratégique, il faut bien sûr une légitimité institutionnelle, mais aussi une véritable capacité d'ouverture pour fédérer un consensus large. Cela a déjà été le cas avec l'expérience du World Wide Web consortium, dont Inria fut le premier hôte européen.

Pour **Jean-François Abramatic, directeur de recherche Inria, et ancien Président du W3C**, «*Il y a plus de 20 ans, le W3C était accueilli en Europe, avec la vocation de garantir un Web indépendant, libre, ouvert et accessible à tous. Aujourd'hui, le projet Software Heritage s'ouvre avec cette même volonté de contribution internationale dans un objectif de préserver et nourrir le bien commun des logiciels de manière pérenne et sur un plan international*».



A ce jour, **Software Heritage** a déjà collecté plus de **20 millions de projets logiciels**, avec plus de **deux milliards et demi de fichiers sources uniques**, ainsi que tout l'historique de leur développement, ce qui en fait déjà **l'archive de code source la plus importante de la planète** (voir <https://www.softwareheritage.org/archive/>).

Pour **Stefano Zacchiroli, directeur de la technologie de Software Heritage**, «*notre ambition est de collecter, préserver, et partager la totalité du bien commun logiciel, dont le logiciel libre forme la vaste majorité. Une approche ouverte et transparente est indispensable pour réussir. C'est la raison pour laquelle notre propre code sera donc entièrement libre et nous invitons d'ores et déjà tous les acteurs du monde du logiciel libre à nous joindre dans la poursuite de cet objectif ambitieux* ».

Ce n'est que le début : en visitant www.softwareheritage.org, vous pourrez suivre en direct l'évolution de la base de code collectée et découvrir comment y contribuer

Les premiers partenaires

Nous avons le plaisir d'annoncer aujourd'hui que deux premiers partenaires internationaux viennent de rejoindre le projet Software Heritage pour le soutenir et l'aider à grandir : **Microsoft**, l'un des plus grands industriels de l'informatique dans le monde, et **DANS**, une institution de l'Académie Royale des Arts et des Sciences et l'organisation néerlandaise dédiée à la préservation et à la promotion de l'accès aux données des recherches sur le numérique.

"Microsoft a été impliqué dans des initiatives open source en facilitant, en intégrant, et en contribuant à de nombreux projets et communautés open source depuis plus d'une décennie", explique **Jean Paoli, General Manager chez Microsoft Corp**. «*Nous applaudissons le projet Software Heritage comme un projet ouvert qui contribuera à organiser et conserver la connaissance sous la forme de code source pour les générations futures, et aidera les générations actuelles de développeurs à trouver et à réutiliser le code dans le monde entier.*

Nous sommes fiers d'être l'un des premiers partenaires industriels à soutenir cette initiative et de mettre à disposition l'infrastructure Cloud Azure pour contribuer à assurer la robustesse et la disponibilité des données".

*« La pérennité des logiciels est un des éléments essentiels pour accélérer la recherche. La reproductibilité des résultats scientifiques, mais aussi et la possibilité de retrouver et réutiliser les logiciels utilisés sont des principes fondamentaux de la science que DANS soutient fermement. », selon **Peter Doorn, Directeur de DANS**, « Nous sommes engagés dans la préservation à long-terme des informations scientifiques, et nous sommes ravis d'être le premier partenaire académique international à rejoindre l'appel de Software Heritage. »*

Software Heritage : nous sommes tous concernés, et tout le monde peut contribuer !

Software Heritage a déjà reçu le soutien de scientifiques, d'industriels, de sociétés savantes, de fondations, et d'une grande variété d'organisations publiques et privées. Tous les témoignages sont disponibles sur le site :

www.softwareheritage.org/support/testimonials/

Après avoir donné l'impulsion nécessaire au démarrage de ce projet, démontré sa faisabilité, et établi les premiers partenariats, Inria lance dès aujourd'hui un appel à rejoindre cette initiative et ouvre le site :

<https://www.softwareheritage.org/>

L'objectif est de donner naissance rapidement à une communauté internationale, qui a vocation à être gérée par un consortium autonome, comme cela a été par exemple le cas du W3C. Les premiers concernés sont les acteurs clés du numérique, les fondations d'intérêt général, les organismes scientifiques et les acteurs publics, français et internationaux.

Mais les contributions de toute nature sont attendues :

- Pour collecter tout le code source disponible : nous aider à identifier les milliers de sites disparates dans lesquels se retrouve aujourd'hui éparpillé le patrimoine logiciel mondial.
- Pour contribuer au développement : l'équipe de **Software Heritage** a une longue tradition de collaboration et est bien connue dans le monde du logiciel libre ; dans les prochains jours, nous allons ouvrir notre propre code source au monde entier et nous serons ravis d'accueillir les développeurs qui voudront partager leur vision et nous aider dans cette mission.
- La résolution des défis scientifiques posés par la construction d'une archive universelle du code source à partir d'informations dispersées va nécessiter de nouvelles contributions ; la participation des chercheurs de toutes les disciplines sera déterminante pour réussir.
- Préserver sur le long terme et partager à travers le monde le contenu de cette archive nécessite des ressources conséquentes, sur un plan humain, matériel et financier, ainsi que des partenaires internationaux.

Dans les prochains jours, **Software Heritage** va ouvrir son propre code source pour que des contributeurs externes, en particulier issus du secteur du Logiciel Libre, puissent participer à l'effort de développement de cette infrastructure commune, et aider à développer l'archive.

S'engager avec **Software Heritage**, c'est l'occasion unique de relever un défi sociétal majeur, de soutenir une initiative d'envergure mondiale et de participer à la préservation de notre patrimoine logiciel mondial.