

Comment utiliser Software Heritage dans son activité de recherche ? – Open science : évolutions, enjeux et pratiques

 opendscience.pasteur.fr/2026/03/04/comment-utiliser-software-heritage-dans-son-activite-de-recherche/

CeRIS - Institut Pasteur

4 mars 2026



[Mise à jour d'un article initialement publié en mars 2023]

L'archive **Software Heritage** (SWH, que nous vous présentons [ici](#)) comprend des millions de codes sources de logiciels, dans des domaines variés et dans une grande variété de langages de programmation. Cette archive peut donc constituer une ressource très utile pour votre activité de recherche. Mais comment l'utiliser ? Quelles sont les fonctionnalités qui pourraient vous être utiles ? Voici quelques manières de l'utiliser.

Trouver des codes sources sur l'archive

Vous pouvez rechercher des codes sources existants, ce qui peut vous faire économiser beaucoup de temps et d'efforts de développement. Vous pouvez interroger SWH [par API](#) ou par l'interface web.

Dans la barre de recherche de l'[archive](#), vous pouvez rechercher des termes, soit dans l'URL du logiciel (cette URL correspond à l'emplacement d'origine du code, par exemple dans git), soit dans les métadonnées associées au code (en cochant la case « *search in metadata* »). Par exemple, [en cherchant par le nom de la bactérie Klebsiella](#), vous pourrez trouver si un logiciel a été développé dans le cadre de travaux sur cette bactérie.

Il n'est pas évident de trouver un résultat pertinent dans SWH, car la description des codes n'est pas sa fonction première ([contrairement à HAL](#)). Il est cependant prévu de développer une nouvelle fonctionnalité de recherche améliorée, pour faciliter la

navigation dans l'archive.

Naviguer dans l'interface de consultation du code source

Lorsque vous cliquez sur une URL, vous accédez à l'interface de consultation du code (par exemple [celle-ci](#)). Vous pouvez notamment :

- Accéder directement à l'emplacement où le code source a été capturé, en cliquant sur la flèche à droite de l'URL.
- Consulter les différentes versions du code (*Releases*), l'historique des révisions (*History*) et ainsi parcourir le code source tel qu'il était à chaque révision.
- Télécharger l'ensemble du répertoire que vous êtes en train de consulter au format tar.gz en cliquant sur le bouton *Download*.

Attention, vérifiez bien la licence du logiciel avant de télécharger et de réutiliser le code source. Si vous ne trouvez aucune information sur la licence, considérez que vous n'avez pas le droit de réutiliser le code. Vous devez en demander l'autorisation aux auteurs.

Citer le code ou un morceau de code

Une fonctionnalité très intéressante de SWH est sa capacité à générer des identifiants uniques et pérennes (*Software Hash IDentifiers*, ou SWHIDs) pour permettre de citer facilement différents artefacts logiciels (répertoires, fichiers, fragments de code, etc).

Concrètement, vous pouvez à tout moment cliquer sur le bouton **Permalinks** situé à droite de votre page pour citer l'élément que vous êtes en train de consulter. Par exemple, si vous sélectionnez certaines lignes de code et que vous cliquez sur ce bouton, SWH génère automatiquement un identifiant et un lien pérenne vers ce morceau de code (par exemple, ce [permalien](#) fait le lien vers les lignes 577 à 593 du fichier [finder.py](#)).

Si le répertoire comporte un fichier codemeta.json ou citation.cff, un bouton **Citations** apparaît également sur la droite. Il permet de générer automatiquement une citation du logiciel au format BibTeX, incluant le SWHID de l'artefact sélectionné ([en savoir plus ici](#)).

Vous pouvez ainsi **citer facilement des codes sources dans vos publications** et même incorporer des permaliens vers des parties de code, afin de souligner l'importance de certaines commandes ou l'apport de certains bouts de code dans votre travail.

Pour aller plus loin sur l'utilisation de Software Heritage, n'hésitez pas à [suivre la visite guidée](#) de l'interface.

S'abonner aux actualités

[Lettre d'information bimensuelle](#)

Filtrer par catégories

Rechercher par mots-clés

[Amérique latine](#) [APC](#) [Archivage et pérennité](#) [Archive ouverte](#) [Bibliométrie](#) [Citation](#) [cOAlition S et plan S](#) [Collaboration](#) [Communs](#) [Conférence](#) [Consultation](#) [COVID-19](#) [Data paper](#) [Données à caractère personnel](#) [Editeur prédateur](#) [Entrepôt de données](#) [Europe](#) [Evaluation de la recherche](#) [Formation](#) [France](#) [Gestion des logiciels](#) [HAL](#) [Identifiant numérique](#) [Intégrité](#) [Juridique et éthique](#) [Licence de diffusion](#) [Logiciel libre](#) [Médiation](#) [Métadonnées et standards](#) [Open Access](#) [Open Data](#) [Open peer-reviewing](#) [Plan de Gestion des Données](#) [Plateforme](#) [Politique des financeurs](#) [Politique des éditeurs](#) [Politique institutionnelle](#) [Politique nationale](#) [Pratiques de gestion des données](#) [Pratiques de publication](#) [Principes FAIR](#) [Prépublication](#) [Reproductibilité](#) [Résultats négatifs](#) [Réutilisation des données](#) [Science et société](#) [Transparence](#) [Urgence de santé publique](#) [Valorisation des logiciels](#) [Voie diamant](#)

Archives

Archives

Contact

Pour toute question relative à cette veille ou à la science ouverte à l'Institut Pasteur, contactez la [bibliothèque du CeRIS](#)

Licence

Le blog Science Ouverte réalisé par la [bibliothèque du CeRIS de l'Institut Pasteur](#) est mis à disposition selon les termes de la licence [CC BY 4.0](#) sauf mention contraire