

Quellcode-Archiv sammelt Open Source

 mittelstandswiki.de/2016/07/software-heritage-archiv-fuer-offene-software-sammelt-open-source/

Software Heritage

© Inria

Das Projekt [Software Heritage](#) hat sich vorgenommen, den Sourcecode von Software aus aller Welt in einem zentralen Archiv zu sammeln und für die Forschung freizugeben. Die Idee entstand am [Institut national de recherche en informatique et en automatique](#) (Inria), einer staatlichen französischen Forschungseinrichtung, die mit den deutschen Max-Planck-Instituten vergleichbar ist. Mittlerweile wurde das Projekt ausgegliedert, die Initiatoren arbeiten nun exklusiv für Software Heritage.



Wie sich schon im Namen andeutet, sieht Software Heritage Programme als Teil des menschlichen Erbes, der erhalten werden muss. Da Software jedoch veraltet, nicht mehr eingesetzt wird und in Vergessenheit gerät, ist es erforderlich, sie zu konservieren. Zu diesem Zweck hat das Projekt ein Archiv mit einer redundanten und ausfallsicheren Infrastruktur aufgebaut. Dort lagert mittlerweile der Quellcode von 22,7 Mio. Projekten.

Da den Benutzern des Archivs der Code frei zur Verfügung gestellt werden soll, kann ausschließlich freie Software gesammelt werden. Die aktuellen Bestände umfassen die Repositorien von [GitHub](#) (ohne die Forks), die [Debian-Pakete](#) auf dem Stand vom August 2015 und den Sourcecode des Betriebssystems [GNU](#), ebenfalls auf dem Stand vom August 2015. Weitere Programme sollen in Kürze dazukommen. Sämtliche Software wird [unter freien Lizenzen](#) der Allgemeinheit zur Verfügung gestellt.

Die Archivierung und Aktualisierung des Bestands geschieht automatisiert, die Benutzer können aber auch eigene Programme hochladen. Damit man beschädigte Dateien identifizieren kann, wurde sämtliche Software mit Hashes indexiert. Der SHA1-Hash wird zudem für die Suche verwendet. In den nächsten Tagen soll eine Möglichkeit zum Download freigegeben werden. Weitere geplante Funktionen umfassen unter anderem die Anzeige von Detailinformationen zu einzelnen Dateien und eine Volltextsuche. (Quelle: [Software Heritage/rf](#))