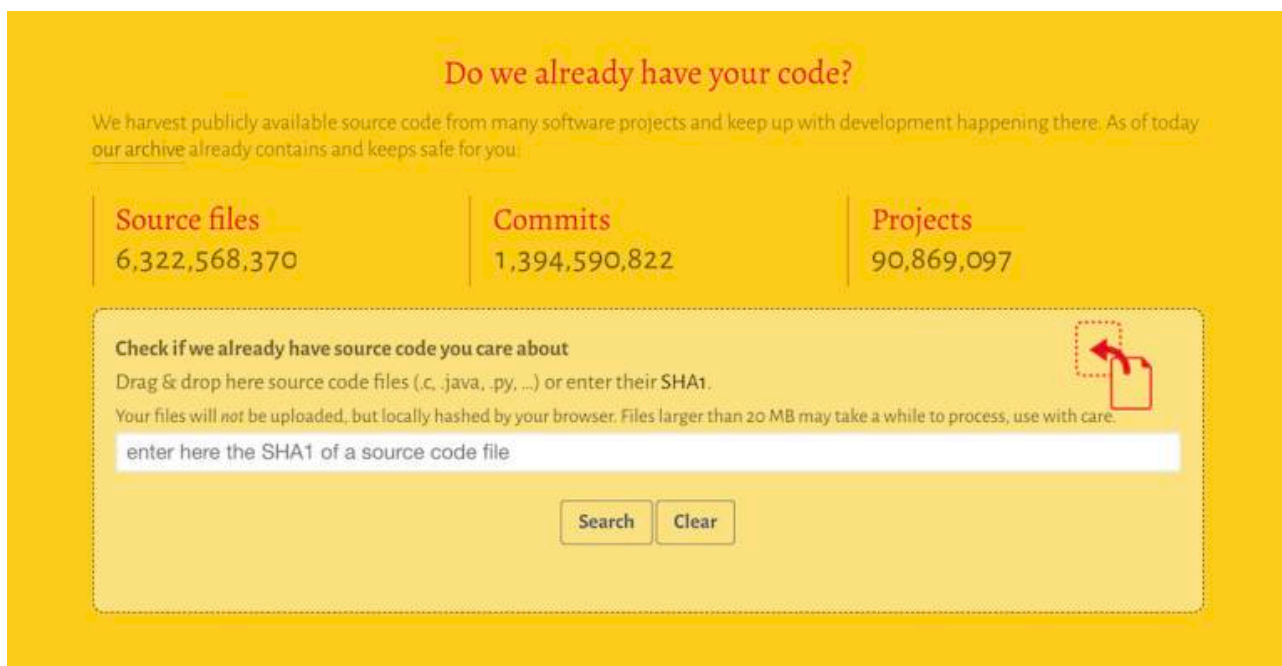


# Software Heritage, la biblioteca mondiale del codice sorgente avrà un centro ricerche a Bologna

di **Dario D'Elia**

“**Software Heritage**“, il progetto che punta a **raccogliere, archiviare e rendere accessibile il codice sorgente di tutti i software** disponibili nel mondo, avrà un centro ricerca a Bologna – esattamente presso la sede locale dell’**ENEA** (Agenzia nazionale per le nuove tecnologie, l’energia e lo sviluppo economico). Non è casuale il coinvolgimento dell’Agenzia, poiché nel tempo si è distinta nell’elaborazione di modelli previsionali, computing ad alte prestazioni grazie all’ENEA Grid Infrastructure (CRESCO) e nello sviluppo software di altissimo profilo.



The screenshot shows the Software Heritage website interface. At the top, it asks "Do we already have your code?". Below this, it states: "We harvest publicly available source code from many software projects and keep up with development happening there. As of today our archive already contains and keeps safe for you:". There are three statistics displayed: Source files (6,322,568,370), Commits (1,394,590,822), and Projects (90,869,097). Below the statistics is a section titled "Check if we already have source code you care about". It includes instructions: "Drag & drop here source code files (.c, .java, .py, ...) or enter their SHA1." and a note: "Your files will not be uploaded, but locally hashed by your browser. Files larger than 20 MB may take a while to process, use with care." There is a text input field with the placeholder "enter here the SHA1 of a source code file" and two buttons: "Search" and "Clear".

“Il Centro ENEA di Bologna non solo contribuirà alla sua sostenibilità a lungo termine: **porterà anche la sua esperienza nei Big Data** per studiare e analizzare il

“Big Code” contenuto in questa libreria di codici sorgente e algoritmi che non ha precedenti”, spiega il team di Software Heritage.

Nato su proposta degli informatici italiani Roberto Di Cosmo e Stefano Zacchiroli e con il sostegno dell’Istituto nazionale francese per la ricerca nel campo dell’informatica e dell’automazione (Inria), nonché l’Unesco, “Software Heritage” punta a diventare una sorta di Biblioteca di Alessandria del codice sorgente. Il riferimento storico è calzante poiché il polo ellenico fu a lungo considerata la più grande biblioteca del mondo antico.

Oggi l’obiettivo è di preservare i codici sfruttando una rete di strutture situate in diverse parti del mondo, in modo che collaborino più istituzioni impiegando domini e tecnologie di archiviazione diverse. Il centro di Bologna sarà uno di questi, **gestendo una copia degli archivi che oggi contengono più di 6 miliardi di file** di codice sorgente unico legati a 90,8 milioni di software sviluppati da 25 milioni di programmatori. Fra questi vi sono codici storici come ad esempio quello del sistema di navigazione dell’Apollo 11 e quello del primo “browser” Mosaic, da cui è poi nato Netscape.

Domani, come anticipa [La Repubblica](#), Roberto Di Cosmo, professore ordinario di Informatica all’università di Parigi e direttore di Software Heritage e Simonetta Pagnutti, della Divisione Ict dell’Enea, presenteranno ufficialmente il progetto in una conferenza stampa a Bologna.