

MEMORIA DIGITALE



di Silvia Bencivelli

PARIGI. È come la Biblioteca di Alessandria: ha l'ambizione di contenere tutto (o quasi) lo scibile umano. Ma qui non c'è niente che si possa toccare. Non un libro, non un manoscritto, niente su cui si poserà mai la polvere. Perché in questo immenso archivio dell'intelligenza umana il sapere è immateriale: è il software che silenziosamente muove le nostre vite digitali.

Questo è il senso dell'iniziativa Software Heritage, nata dal basso nella comunità degli informatici e oggi sostenuta dall'Unesco, che ha cominciato a riconoscere come patrimonio culturale il lavoro certosino di non si sa quanti anonimi programmatori, autori di milioni, o forse miliardi, di programmi. O meglio, di quelli che tecnicamente si chiamano "codici sorgente", cioè i "testi" dei programmi che fanno fare alle macchine quello che vogliamo, e che brulicano sotto la superficie delle nostre vite: ci permettono di mandare messaggi, prenotare aerei, scattare fotografie, fare shopping, giocare e persino flirtare. Ma soprattutto fanno funzionare industrie, pubblica amministrazione, trasporti, comunicazione, sanità, ricerca... praticamente ogni comparto della società.

Questi codici rischiano di dissolversi nel niente, di scomparire senza la-

TUTTO IL SAPERE IMMATERIALE STA DENTRO UNA BIBLIOTECA

SI CHIAMA **SOFTWARE HERITAGE** ED È NATA DALL'IDEA DI UN ITALIANO A PARIGI, ROBERTO DI COSMO. PUNTA A PROTEGGERE DALL'OBLIO I PROGRAMMI DI MILIONI DI INFORMATICI. COMPRESI I CODICI DI FERMI

sciare traccia in quel mondo digitale sommerso del quale noi utenti comuni ignoriamo beatamente l'esistenza. Ecco perché li si vuole raccogliere e proteggere in una biblioteca universale.

L'idea iniziale è stata di Roberto Di Cosmo: informatico italiano di nascita, ma parigino per accento, portamento e soprattutto affiliazione (è professore all'Université Paris Cité, oltre che direttore di Software Heritage), dal 2015 sta lavorando alacremente per convincere ministri e istituzioni culturali francesi e internazionali a sostenerlo. Riuscendoci. Così il primo febbraio la sede dell'Unesco a Parigi ha aperto le porte a un folto gruppo di informatici (e non solo) chiamati a discutere degli sviluppi dell'ini-

ziativa, e delle enormi novità che arrivano dalla crescita dell'intelligenza artificiale. Perché qui non si tratta solo di evitare di disperdere il lavoro già fatto: dietro c'è un'idea di proprietà intellettuale e di gestione collettiva della scienza, della sua utilità per lo sviluppo umano, e un nuovo pensiero ecologico sul costo ambientale della tecnologia.

Com'è andata? «Scrivere un codice sorgente richiede una quantità enorme di conoscenze» spiega Di Cosmo. «A un certo punto ci siamo resi conto che, a differenza di altre espressioni della nostra cultura, questa conoscenza non veniva vista, e quindi non veniva protetta». Era più o meno il 2014. L'anno dopo Di Cosmo e i suoi



Roberto Di Cosmo, ideatore e direttore di Software Heritage



hanno cominciato ad archiviare il software rilasciato in rete, giusto in tempo: «Esattamente in quel momento ci siamo accorti che quasi un milione di progetti software stavano per essere cancellati tutti insieme». Succedeva perché, con pochissimo preavviso e senza deroghe per nessuno, stavano chiudendo due delle piattaforme commerciali dove vengono depositati i progetti. «Come se d'un tratto chiudessero le piattaforme in cloud dove salviamo i nostri file, tipo Dropbox o Google Drive» prosegue Di Cosmo.

«Oggi siamo a circa 18 miliardi di codici sorgente (contenuti in 270 milioni di progetti) prodotti da circa settanta milioni di programmatori di tutto il mondo». Dentro a questo archivio c'è anche un sacco di storia, come i programmi usati per guidare la missione Apollo 11 sulla Luna. Così all'università di Pisa, una delle tre partner accademiche italiane del progetto con l'università di Bologna e la Scuola Normale Superiore di Pisa, ci si è messi a cercare tra i codici più antichi e si è riusciti a ritrovare alcuni tra i primi contributi italiani all'informatica. Per esempio le linee di codice scritte per la mitica CEP, Calcolatrice Elettronica Pisana, il primo calcolatore scientifico costruito in Italia su impulso di Enrico Fermi e inaugurato dal presidente della Repubblica Giovanni Gronchi nel 1961. E i primi programmi per comporre musica elettronica negli anni Settanta.

Poi la storia è andata avanti e i numeri sono cresciuti. «Molto del software oggi prodotto viene pubblicato su Internet e reso libero: chiunque può riprenderlo, modificarlo, usarlo come vuole», spiega Stefano Zacchioli, un altro italiano in Francia, oggi professore all'Istituto Politecnico di Parigi nonché cto di Software Heritage. «Tutti questi progetti definiti *open source* sono un enorme bene comune, con cui si costruiscono telefonini, si fa andare la Stazione Spaziale Internazionale e così via».

Una biblioteca universale come quella di Software Heritage consente anche di fare tante altre cose: «Per esempio di studiare come si sviluppano i software, dove e chi li scrive (e si scopre che il 92 per cento dei programmatori sono maschi, ma con una tendenza al miglioramento)», prosegue Zacchioli. «Si può studiare la storia del software mondiale e renderne più facile lo studio. Si possono cercare errori e problemi di sicurezza, segnalarli per facilitarne la correzione. Si possono studiare le tendenze del momento e dedurre le evoluzioni future».

«Insomma Software Heritage è anche come uno di quei grandi telescopi con cui si osservano le galassie»,

chiosa Di Cosmo, «solo che noi osserviamo la galassia del software». Con un pizzico di orgoglio, «perché qui c'è la differenza cruciale tra noi informatici e i matematici» aggiunge Chiara Bodei, professoressa all'università di Pisa. «Ai matematici, di un programma, interessano soprattutto i dati che entrano e quelli che escono. A noi, invece, la procedura interessa quanto il risultato. Siamo un po' artigiani, anche perché "riutilizziamo" gli strumenti (cioè i codici) scambiandoceli e costruendoci nuovi programmi...».

Tutto questo è immateriale, ma le

macchine servono eccome: i server preposti allo stoccaggio, alla gestione e alla elaborazione del traffico di informazioni (sono come enormi "cabine armadio" per file), e servono siti mirror, cioè altre macchine su cui si mette, per sicurezza, copia dei dati. In Italia abbiamo il primo mirror istituzionale di Software Heritage, attivato a dicembre scorso al centro di ricerche Enea di Bologna. Qui è già conservata «una miniera di codici e algoritmi», spiega Giovanni Ponti,

responsabile della Divisione per lo sviluppo dei sistemi per l'informatica e l'Ict di Enea, «e con questa enorme quantità di dati si potrà anche addestrare la prossima intelligenza artificiale». Una prospettiva, ha spiegato all'Unesco Gilles Matthiaeu, rappresentante del ministero francese per l'Università e la Ricerca, che aiuta a pensare la IA generativa come più democratica e trasparente, perché basata su dati "open", a disposizione di tutti.

Questo la biblioteca di Alessandria non poteva farlo. Non poteva difendersi dagli incendi con una "bi-

blioteca mirror". E non poteva fare altro che conservare. Il software, invece, può essere messo da parte ed essere riciclato per il futuro. Provando, se ci si riesce, a renderlo migliore. □



Il logo di Software Heritage: per l'Italia collaborano l'università di Pisa, la Scuola Normale Superiore, sempre di Pisa, e l'università di Bologna

«A DIFFERENZA
DI ALTRE
ESPRESSIONI
DELLA CULTURA,
QUESTA
NON VENIVA
PRESERVATA»