

Corsa contro il tempo per salvare il sapere digitale

Conoscenza. Solo il 44% delle pubblicazioni accademiche è archiviato in modo sicuro La minaccia arriva da hacker e formati obsoleti

Gianluca Dotti



Milioni di articoli scientifici rischiano di scomparire dal web, mettendo a repentaglio il progresso della ricerca. Secondo l'International Association of Scientific, Technical and Medical Publishers, solo il 44% delle pubblicazioni accademiche digitali è archiviato in modo sicuro. Nel 2017, il servizio di archiviazione Lockss ha registrato la perdita irreversibile di oltre 50mila paper. È come se un'intera biblioteca fosse stata data alle fiamme, senza possibilità di ricostruirla.

«L'infrastruttura digitale del mondo della ricerca è estremamente vulnerabile e la politica di chiusura da parte di riviste e piattaforme può comportare la perdita definitiva di decenni di conoscenza», spiega Martin Eve, professore di tecnologia dell'editoria all'università di Londra, interpellato anche da Nature alla fine dell'anno scorso in un accorato appello a intervenire in fretta. I dati mostrano che il 13% degli articoli pubblicati su riviste scientifiche chiuse tra il 2000 e il 2015 non è più reperibile. E il rischio è amplificato dall'obsolescenza tecnologica: formati digitali diventano rapidamente incompatibili con i nuovi sistemi, rendendo impossibile l'accesso ai contenuti. «I documenti stampati resistono per secoli, mentre i file digitali richiedono un costante trasferimento e aggiornamento», spiega Marco Leona, responsabile della ricerca scientifica al Metropolitan Museum of Art (Met) di New York. «Basta pensare ai papiri di Ercolano, che anche bruciati e carbonizzati possono essere letti grazie a tecnologie di frontiera, mentre un contenuto digitale se viene perso lo è in modo irreversibile». Rischiamo, insomma, di arrivare all'assurdo: se da un lato possiamo leggere manoscritti medievali, dall'altro rischiamo di perdere per sempre

informazioni contenute in floppy disk degli anni Novanta.

La digitalizzazione ha comunque offerto enormi vantaggi, velocizzando la diffusione delle pubblicazioni scientifiche: oggi il 90% delle riviste è solo digitale, consentendo ai ricercatori di accedere a migliaia di articoli in pochi secondi. Tuttavia, questa transizione ha reso la conservazione un problema sempre più urgente da affrontare. In assenza di misure adeguate, gli archivi digitali potrebbero diventare una fragile illusione, un castello di sabbia destinato a sgretolarsi nel tempo, o al primo hackeraggio. Molte istituzioni accademiche stanno cercando di affrontare il problema attraverso collaborazioni internazionali per creare archivi digitali ridondanti, in modo da evitare che la perdita di una piattaforma comprometta l'accesso globale alla conoscenza. «Senza strategie strutturate di conservazione digitale, una quantità significativa di conoscenza accumulata sarebbe in bilico», specifica Eve. «Se le pubblicazioni digitali non vengono preservate adeguatamente, rischiamo di perdere collegamenti cruciali nella documentazione accademica, compromettendo la capacità di verificare e costruire sulla ricerca esistente: un disastro per la comunità scientifica globale. Il panorama è reso ancora più critico dall'aumento della frequenza di violazioni informatiche nel mondo accademico».

Nel 2022, più di un quinto delle università ha subito attacchi ransomware, con un aumento del 50% rispetto all'anno precedente. Gli hacker bloccano l'accesso ai server delle istituzioni e chiedono riscatti milionari per restituire i dati. L'Università della California ha pagato 1,14 milioni di dollari per recuperare file cifrati da un attacco nel 2020. «I musei e le istituzioni della scienza e della cultura sono particolarmente vulnerabili: i fondi sono destinati alla tutela delle opere, non alla conservazione delle ricerche», chiarisce Leona. «La ricerca nel nostro settore è spesso poco finanziata e la sicurezza informatica non è una priorità, rendendoci bersagli facili per i criminali informatici».

Un altro aspetto centrale è la diffusione dell'open access: il modello aperto ha reso disponibili milioni di articoli senza barriere economiche, ma la possibilità di scaricare liberamente i contenuti aumenta la loro diffusione e preservazione attraverso copie distribuite. Questo potrebbe aiutare a impedire lo scenario distopico in cui intere biblioteche digitali diventino inaccessibili a causa di collassi informatici o di interruzioni energetiche. Del resto, la memoria e l'archiviazione della conoscenza sono alla base del progresso: perdere milioni di pubblicazioni significa disperdere secoli di studi e dati, lasciando le nuove generazioni di ricercatori senza le indispensabili fondamenta su cui costruire il futuro. A preoccupare gli esperti è anche l'obsolescenza del formato, che in futuro potrebbe rendere difficile o addirittura impossibile l'accesso persino a contenuti correttamente archiviati e protetti. È dunque essenziale sviluppare strategie per la migrazione dei dati verso formati più moderni e

compatibili nel tempo. Progetti innovativi come Archiver, finanziato dalla Ue, stanno sviluppando infrastrutture per garantire la conservazione a lungo termine delle pubblicazioni, adottando tecnologie che permettono la migrazione automatizzata dei dati per contrastare l'obsolescenza. Si tratta di un'attività dinamica e continua, che richiede aggiornamenti per mantenere le pubblicazioni accessibili e anche adeguati e continui investimenti.

© RIPRODUZIONE RISERVATA