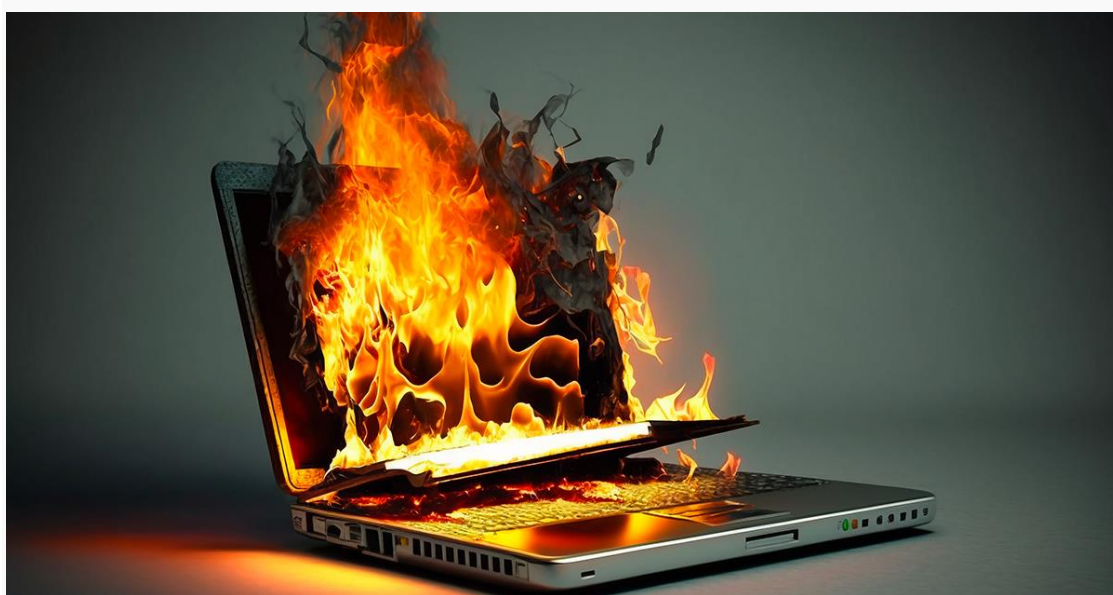


Gli articoli scientifici rischiano di sparire da Internet

di **Antonella De Robbio**

<https://ilbolive.unipd.it/it/news/articoli-scientifici-rischiano-sparire-internet>



La conservazione della documentazione accademica è stata una tematica rilevante fin dai primi stadi della produzione del sapere. Tradizionalmente, con le pubblicazioni cartacee, il compito di conservazione ricadeva principalmente sulle istituzioni bibliotecarie. Tuttavia, con il passaggio all'editoria digitale e, in particolare, con l'introduzione dell'accesso aperto (*Open access*, OA), si sono verificate ambiguità e complessità aggiuntive. Questa transizione ha comportato sfide significative per l'accessibilità a lungo termine delle riviste scientifiche, che potrebbero addirittura sparire completamente dal web.

A quasi un ventennio di distanza dall'azione intrapresa dall'Europa verso un accesso aperto al 100% si riscontrano infatti ancora diversi ostacoli al cambiamento, derivanti soprattutto da carenze organizzative e dalla mancanza di chiarezza riguardo alle responsabilità. Sebbene siano intanto emerse infrastrutture digitali per il deposito di articoli scientifici e dati in tutto il mondo, spesso mancano modelli finanziari e organizzativi capaci di garantire un accesso a lungo termine. Le strutture di supporto alla conservazione si configurano per progetti specifici, con finanziamenti limitati a periodi definiti: risorse frammentarie e non in grado di fornire soluzioni a lungo termine. Dal punto di vista tecnico permangono criticità riguardanti la conservazione di grandi quantità di dati, particolarmente evidenti in discipline come l'astronomia e le scienze della

terra; persiste intanto il problema dell'interoperabilità tra discipline, ma va anche considerato il processo di adattamento delle normative e delle prassi nazionali di ciascun membro dell'UE in materia di deposito obbligatorio dei dati, sia in termini di materiali inclusi che di modalità pratiche adottate.

Se andiamo a ritroso nel tempo, il 17 luglio 2012 la Commissione Europea emanò due importanti documenti rivolti agli Stati membri sull'accesso aperto all'informazione scientifica. Il primo fu la *Comunicazione COM(2012) 401 final*, *Verso un accesso migliore alle informazioni scientifiche: aumentare i benefici dell'investimento pubblico nella ricerca*, definiva gli obiettivi di una policy sull'accesso aperto ai contenuti della ricerca finanziata nel corso del programma quadro *Horizon 2020*. La strategia Europa 2020 per un'economia intelligente, sostenibile e inclusiva, sottolineava il ruolo centrale della conoscenza e dell'innovazione nel generare crescita economica. I risultati della ricerca, inclusi sia le pubblicazioni sia le raccolte di dati, avrebbero dovuto essere diffusi rapidamente e ampiamente utilizzando i *repository* e le piattaforme digitali: ciò avrebbe dovuto accelerare le scoperte scientifiche, consentendo nuove forme di ricerca ad alta intensità di dati e favorendo al contempo l'adozione sistematica dei risultati della ricerca da parte delle imprese e dell'industria europee.

I tre obiettivi principali di *Horizon 2020* erano nell'ordine: (1) dotarsi entro il 2014 da parte di tutti gli Stati membri dell'UE di politiche governative di Open Access; (2) rendere disponibili in OA entro il 2016 il 60% degli articoli di ricerca pubblicati in Europa e (3) il 100% delle pubblicazioni di Orizzonte 2020. Ancora oggi però il tasso e il grado di "apertura" di articoli scientifici pubblicati in OA varia da Paese a Paese e da disciplina a disciplina. Molto c'è ancora da fare e l'orizzonte è ancora lontano.

Successivamente è stato lanciato *Orizzonte Europa 2021-2027*, un programma di finanziamento che durerà fino al 2027 e dispone di un budget di 95,5 miliardi di euro, inoltre nel 2018 l'Unione Europea ha sostenuto un consorzio di organizzazioni nazionali di finanziamento della ricerca nel lancio di *cOAlition S*, che si basa sul PlanS o Piano S che mira a promuovere l'accesso aperto completo e immediato come obiettivo principale. A partire dal 2021, il Piano S impone l'accesso aperto completo e immediato alle pubblicazioni accademiche soggette a revisione paritaria, provenienti da ricerche finanziate sia da sovvenzioni pubbliche che private.

L'idea di Open Access arriva però in Europa già nel 2007 con un'apposita *Comunicazione della Commissione al Parlamento Europeo* si invitavano gli Stati membri ad adottare una serie di misure, sia per l'accesso che per la conservazione di dati e informazioni, delineando anche i tempi previsti. Tuttavia i progressi nel corso di questi quindici anni sono stati disomogenei. Tale disomogeneità emerge dall'ultima edizione 2018 della *Relazione sull'accesso all'informazione scientifica e sulla conservazione in Europa sull'attuazione della raccomandazione C(2018)2375 final della Commissione* che consolida e presenta le informazioni comunicate dagli Stati membri in merito ai loro progressi nell'attuazione della raccomandazione del 2012, poi successivamente riviste con la *Raccomandazione 2018/2375/UE*, (che ha sostituito e rafforzato la *2012/417/UE* con la quale si esortavano gli Stati membri a migliorare le politiche e le pratiche in materia di accesso e conservazione). Dall'indagine condotta nel corso del 2017-2018 emerge come in alcuni Paesi sia fondamentale intensificare gli sforzi per garantire le competenze necessarie per la corretta conservazione delle informazioni scientifiche ai fini di un loro riutilizzo: la formazione è quindi un altro aspetto decisivo da considerare.

Va ricordato che nel 2011 era stato redatto un Documento di lavoro dei servizi della Commissione Europea di accompagnamento alla [Comunicazione della Commissione sulla digitalizzazione e l'accessibilità in rete dei materiali culturali e sulla conservazione digitale, SEC\(2011\) 1274 definitivo](#). In altri termini l'Europa avvertiva come uno degli ostacoli al cambiamento fosse proprio un non adeguato investimento sull'accesso alle informazioni scientifiche e sulla loro conservazione. La Commissione Europea prevedeva che non sarebbe stato possibile realizzare appieno le potenzialità economiche e sociali di un miglior accesso alle informazioni scientifiche se i fondi disponibili per tale accesso e per la loro conservazione a lungo termine fossero stati insufficienti.

È indubbio che la conservazione a lungo termine delle informazioni, delle conoscenze e delle competenze ad uso delle generazioni future presenti vantaggi economici e sociali considerevoli: a tal fine l'Europa sottolinea da tempo come sia necessario dedicare particolare attenzione anche alla conservazione del software e dei modelli scientifici per assicurare la riusabilità e la riproducibilità dei risultati delle ricerche, suggerendo l'adozione di standard e formati aperti e di software a codice sorgente aperto.

In particolare il considerando 7 della Comunicazione del 2018 recita: *“La conservazione dei risultati della ricerca scientifica risponde all'interesse pubblico. Questo compito è affidato in genere alle biblioteche o agli archivi, in particolare alle biblioteche nazionali di deposito legale. Il volume dei risultati della ricerca è in costante crescita. Per consentire la conservazione a lungo termine dei risultati della ricerca in formato digitale occorrono meccanismi, infrastrutture e soluzioni software. Il finanziamento sostenibile della conservazione riveste un'importanza cruciale, perché i costi legati alla cosiddetta curation (raccolta, ordinamento e condivisione) dei contenuti digitalizzati sono ancora relativamente elevati. Vista l'importanza della conservazione per l'uso futuro dei risultati della ricerca, è opportuno raccomandare agli Stati membri l'elaborazione o il rafforzamento di politiche in quest'area”*.

In uno studio esplorativo apparso nel 2020 nella sezione biblioteca digitale del repository internazionale arXiv dall'eloquente titolo *Open is not forever: a study of vanished open access journals* (in seguito pubblicato nel 2021 nella rivista OA [Journal of the Association for Information Science and Technology](#)), Mikael Laakso, Lisa Matthias e Najko Jahn analizzano sistematicamente il fenomeno delle riviste scientifiche scomparse, un ambito di ricerca precedentemente poco indagato. Per condurre tale analisi, sono stati consultati diversi importanti indici bibliografici come [Scopus](#), [Ulrichsweb](#) e la [Directory of Open Access Journals](#) (DOAJ), oltre alle riviste tracciate attraverso la [Wayback Machine di Internet Archive](#). I tre autori hanno individuato 176 riviste OA che, a causa della mancanza di archivi completi e aperti, sono scomparse dal web nel periodo compreso tra il 2000 e il 2019, spaziando attraverso tutte le principali discipline di ricerca e le regioni geografiche del mondo. I risultati emersi dalla loro ricerca sollevano una seria preoccupazione riguardo all'integrità del registro accademico e sottolineano l'urgenza di intraprendere azioni collaborative per garantire un accesso continuo alle conoscenze scientifiche e per prevenire la perdita di dati accademici cruciali.

Il problema della scomparsa delle riviste open access (OA) è un tema importante che richiede una riflessione approfondita, **ma non sono a rischio solo le riviste o gli articoli open, ma anche articoli scientifici non necessariamente open access anche se muniti di DOI**. La conservazione digitale è fondamentale per garantire la persistenza dei collegamenti e delle citazioni accademiche tramite di Digital Object Identifier (DOI), uno standard che consente l'identificazione univoca e duratura di oggetti digitali, associandoli a

metadati strutturati ed estensibili. Questo identificatore facilita la gestione, l'accesso e la transazione di informazioni e documenti all'interno di reti digitali, consentendo anche l'automazione dei processi. Crossref è la più grande agenzia di registrazione DOI, che assegna gli identificatori a circa 20.000 membri, tra cui editori, musei e altre istituzioni. A differenza degli URL, che indicano solo la posizione dell'oggetto, il DOI identifica direttamente l'oggetto stesso come entità primaria. È inoltre distinguibile dagli identificatori bibliografici tradizionali (ISBN, ISSN, ISRC, ecc.) poiché è immediatamente attivabile online e può essere utilizzato per sviluppare servizi specifici come motori di ricerca e certificazioni di autenticità.

L'allarme secondo cui milioni di articoli scientifici rischiano di scomparire da Internet arriva da un recente studio pubblicato il 24 gennaio 2024 nel *Journal of Librarianship and Scholarly Communication* condotto da Martin Paul Eve su 7.438.037 articoli muniti di DOI. Tra le opere esaminate sono state individuate 5,9 milioni di copie distribuite negli archivi considerati nello studio. Complessivamente 4.342.368 delle opere analizzate (58,38%) erano presenti in almeno un archivio. Tuttavia il 27,64% delle opere, pari a 2.056.492, risultavano apparentemente non conservato. Il restante 13,98% dei lavori nel campione è stato escluso per motivi quali la pubblicazione nell'anno corrente, la non appartenenza alla categoria degli articoli di rivista o la mancanza di metadati sufficienti per identificare la fonte.

In generale la mancata conservazione nel tempo delle risorse digitali comporta almeno tre rischi piuttosto seri, non solo per quanto riguarda le pubblicazioni scientifiche: la privatizzazione, la manipolazione delle fonti, la non verificabilità delle citazioni. In particolare nelle pubbliche amministrazioni, per siti web di organi governativi, ministeri e governi locali, le conseguenze possono essere preoccupanti. Una delle principali preoccupazioni riguarda l'obsolescenza del formato, che in futuro potrebbe rendere difficile o addirittura impossibile l'accesso ai contenuti archiviati.

È dunque essenziale sviluppare strategie per la migrazione dei dati verso formati più moderni e compatibili nel tempo. Inoltre, la conservazione digitale non è un'attività statica ma dinamica e continua, e richiede costanti investimenti e aggiornamenti per mantenere i dati accessibili e utilizzabili nel tempo. Impegno costante che dovrebbe essere garantito da parte delle istituzioni e delle organizzazioni che gestiscono queste risorse. È infine essenziale considerare anche gli impatti ambientali, tramite il consumo di energia elettrica e la produzione di rifiuti elettronici, delle strategie di conservazione digitale. Tutte questioni essenziali, sulle quali è urgente aprire un confronto partecipato e soprattutto prendere provvedimenti: **il rischio è veder sparire una parte consistente della conoscenza che produciamo.**