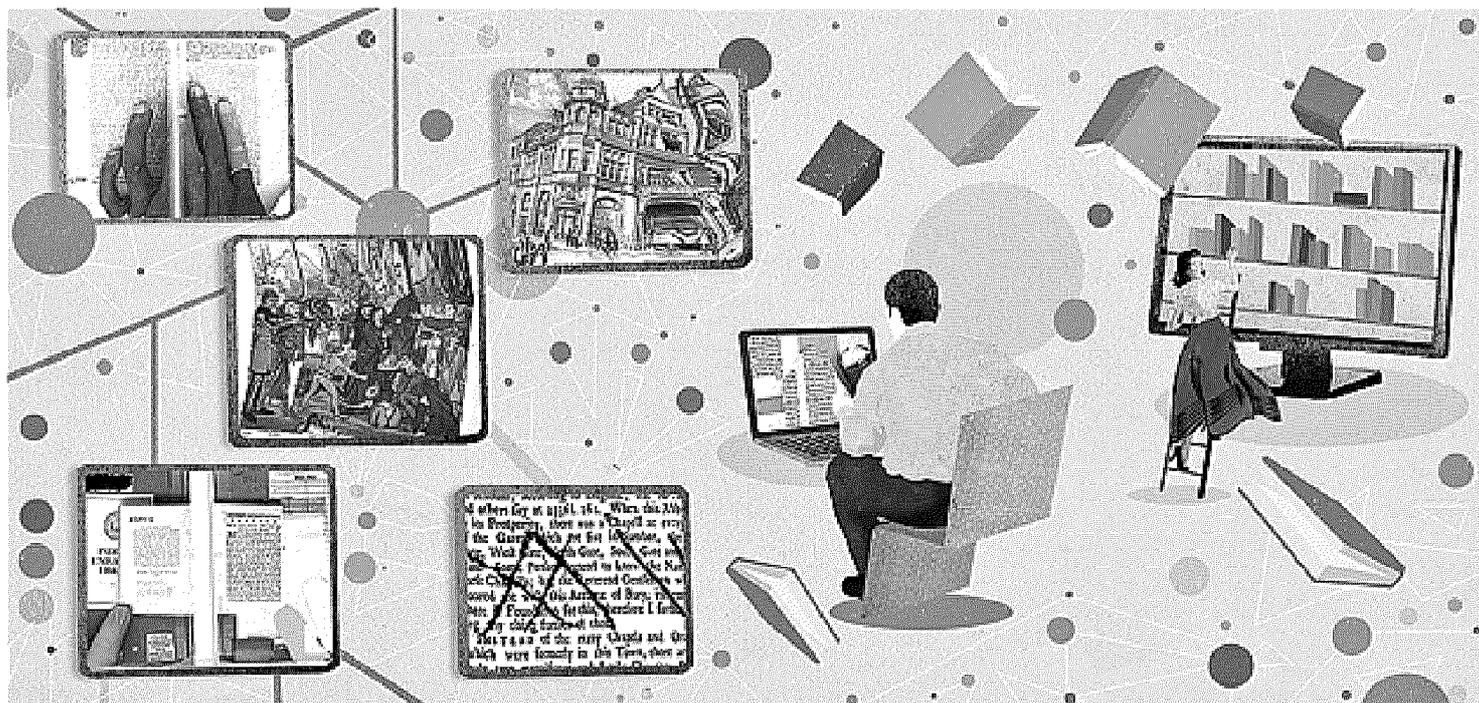


Archivi 2.0

Errori di scansione, obsolescenza dei supporti, monopolio di Google
La digitalizzazione della conoscenza rischia di disperdere il passato



I bit non bastano, il sapere svanisce

di FEDERICO GUERRINI

Parigi, 2020. Una giovane coppia. Lui impegnato a registrare con il telefonino momenti della loro vita, lei che si stufa e se ne va. Qualche tempo dopo, una tempesta elettromagnetica distrugge tutti gli archivi digitali del pianeta e il ragazzo rimane senza più memorie. Di quel che è stato resta solamente una vecchia foto Polaroid che lei gli aveva consegnato prima di sparire. È la trama di *Lost memories*, cortometraggio del regista francese François Ferracci apprezzatissimo l'anno scorso al Trieste Film Festival, ma è anche la metafora di come il nostro mondo — tutto teso a rincorrere il presente e il futuro — rischi di perdere, con la progressiva digitalizzazione della conoscenza, il passato.

Il passaggio da un sistema di archiviazione analogico — nastri, dischi, album fotografici — a uno digitale era visto in origine come una delle strade maestre per la preservazione del sapere (la Carta per la conservazione del patrimonio digitale stilata dall'Unesco è del 2003), tuttavia si è rivelato nell'ultimo decennio un'arma a doppio taglio. Fonte di grandi opportunità in termini di riduzione dello spazio necessario all'archiviazione, facilità di accesso e riproduzione delle informazioni e minori costi, ma foriero anche di emergenze e di contraddizioni.

Nell'illustrazione di **FRANCESCA CAPELLINI** foto di errori di scansione di Google Books tratte da theartofgooglebooks.tumblr.com

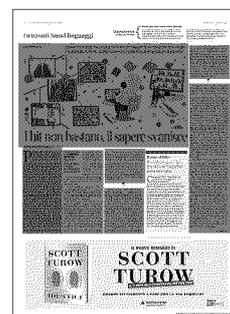
Un libro di carta può durare centinaia d'anni, mentre il ciclo di vita di un supporto magnetico può arrivare al massimo a un secolo.

«Un'altra ragione — prosegue Carr — è che applicazioni, sistemi operativi e formati dei file diventano obsoleti velocemente, rendendo impossibile accedere a materiale multimediale anche solo di pochi anni precedente».

Quanti floppy disk rimangono a prendere polvere, da quando sono stati inventati i masterizzatori? E quante volte aprire un file creato con vecchie versioni di Photoshop o Office si rivela un'impresa improba? Senza contare che nel processo di conversione da analogico a digitale possono intervenire errori difficili da eliminare. Per il progetto «Books», Google è arrivato ormai a scansare quasi 30 milioni di volumi: inevitabile, con tali numeri, incorrere in qualche errore. I nuovi amanuensi non lasciano glosse

come i loro predecessori, ma immagini di dita o persino di intere mani, riprese per errore dalla macchina che effettua la scansione. Quelle che Google chiama *unexpected peculiarities* possono includere anche timbri di biblioteche o di precedenti possessori, scarabocchi. L'effetto è paradossale e non di rado affascinante, tanto che alcuni blog (uno su tutti: *The Art of Google Books*) hanno finito per esaltare questi errori come una forma di «arte».

«La digitalizzazione in realtà può ridurre le possibilità che un'opera sopravviva» spiega a «la Lettura» Nicholas Carr, autore del bestseller *Internet ci rende stupidi?* (edito da Raffaello Cortina).



A mano a mano che la digitalizzazione dei testi avanza — e a essa si accompagnerà la sempre più difficile reperibilità degli originali cartacei — cosa accadrà delle pagine rese illeggibili dagli errori di scansione? Siamo destinati a perdere interi frammenti di conoscenza? Ed è giusto lasciare a una multinazionale privata il monopolio della preservazione della produzione libraria? Risponde a «la Lettura» Jeffrey Schnapp, direttore del MediaLab di Harvard: «Per quanto potente, Google è soltanto una piattaforma e per giunta molto farraginoso. Tuttavia è significativo che solo una grande società privata abbia avuto l'immaginazione e le risorse per lanciare un progetto così ambizioso. Dovrebbero essere le istituzioni pubbliche a prendere l'iniziativa, ma spesso non sono brave a implementare risorse e piattaforme sul web». Ci sono comunque lodevoli eccezioni nate negli ultimi anni su impulso di istituzioni, università e consorzi privati. L'HathiTrust Digital Library, fondata nel 2008, riunisce e rende accessibili (in maggiore o minore misura, a seconda del copyright), più di 10 milioni di volumi provenienti da un'ottantina di atenei e partner privati nordamericani ed europei.

Nell'aprile dello scorso anno, dall'Università di Harvard è partita la Digital Public Library of America, un hub che indicizza e rende facilmente reperibili risorse online provenienti da ogni angolo degli Usa. Nel Vecchio Continente, la biblioteca digitale Europea svolge da un lustro una funzione simile e aggrega ormai più di 30 milioni di oggetti multimediali: non soltanto libri, ma anche quadri, mappe, manoscritti e file sonori. L'Unesco persegue dal 1992 il progetto *Mémoire du monde*, contenente fra l'altro un registro di testimonianze della storia umana che comprende ad oggi 299 documenti e collezioni conservati su tutti i supporti possibili dalla pietra alla celluloidi fino alle registrazioni audiovisive. E oggi si sta interessando particolarmente dei problemi della conservazione digitale. Problemi che non riguardano esclusivamente i reperti del passato e nemmeno soltanto i manufatti culturali. Una recente analisi dell'Università della British Columbia (Canada) su un campione di 516 studi scientifici pubblicati fra il 1991 e il 2011 ha mostrato come, dopo vent'anni, per l'80% di essi i dati originali di ricerca non fossero più reperibili.

Nell'era dei Big Data, ogni giorno vengono prodotti 2,5 quintilioni (1 seguito da 30 zeri) di byte di dati: malgrado le «nuvole informatiche», gli immensi data center e i chip di ultima generazione esiste il rischio concreto di essere incapaci di gestire la mole di informazioni. Anche se è evidente che solo una parte di dati ha davvero valore e, sottolinea Schnapp, «dobbiamo abbandonare l'illusione della "conservazione totale": non l'abbiamo mai praticata, nemmeno nel XVIII e XIX secolo. Preservare è, per definizione, un processo selettivo».

Se gli attuali sistemi di archiviazione non dovessero rivelarsi sufficienti, comunque, potremmo escogitarne di nuovi: un anno fa, due scienziati dell'Istituto europeo di bioinformatica hanno dimostrato come sia tecnicamente possibile, per quanto al momento troppo costoso, immagazzinare informazioni multimediali nel Dna. Questo consentirebbe di accedere a volumi di spazio al momento inimmaginabili. Oppure, Dna a parte, se il digitale non convince proprio si potrebbe pensare di tornare alla carta. Se non altro come ultima ratio.

 @fedeguerrini

© RIPRODUZIONE RISERVATA