



Gli hard disk di tutto il mondo faticano a stare al passo con la crescita vertiginosa dei dati digitali.

# NON PERDIAMOCI I RICORDI (DIGITALI)

Ormai quasi tutto ciò che riguarda la «registrazione» dell'attività umana (dalle fotografie a ogni genere di documento) avviene su file virtuali. E impegna, su server sparsi nel mondo, una quantità di byte inimmaginabile. Ma è una memoria a rischio di sparizione. Come è già successo.

di Luca Sciortino

In *Blade Runner 2049*, film che ha riproposto in diversa chiave narrativa il celebre *Blade Runner* di Ridley Scott di 42 anni fa, il mondo viene sconvolto da un disastro apocalittico che cancella ogni dato digitale. L'umanità viene così privata della sua memoria collettiva e ricordi artificiali devono essere impiantati nella mente di replicanti, copie artificiali di esseri umani.

La fantascienza ha in questo modo previsto uno scenario da tutti temuto, e forse persino possibile: ora che la memoria di ciascuno di noi, per esempio fotografie, post, messaggi e altri ricordi personali, ma anche dati medici, genetici, legali e molto altro, risiede quasi interamente in uno spazio virtuale, che cosa succederebbe se anche solo una parte venisse cancellata per sempre?

Diversi fatti mostrano che si tratta di una questione sempre più urgente. Un giornale americano che analizza l'evolversi della mole degli studi accademici online (*Journal of Librarianship and Scholarly Communication*) ha appena riportato che oltre un quarto di tutti gli articoli di ricerca identificati da un Doi (sorta di codice univoco di identificazione) sono scomparsi o a rischio di scomparire da Internet. Così, altri articoli che li citano, dando per scontata la plausibilità dei loro risultati, saranno molto meno credibili: si baseranno su conclusioni di cui non esisterà più prova. In un certo senso, l'edificio della nostra conoscenza viene minato alle fondamenta.

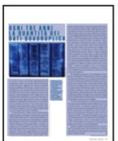
Bisogna poi ricordare i numerosi casi nei quali disastri ambientali hanno cancellato per sempre in-

formazioni sensibili di privati cittadini distruggendo anche dati di compagnie e uffici pubblici, come nell'incendio di dimensioni colossali scoppiato nel nord della California nel 2018 o il terremoto con tsunami del 2011 in Giappone.

Infine, il nostro sforzo di immagazzinamento negli hard disk fatica a stare al passo con la crescita vertiginosa dei dati digitali. È come se dovessimo continuamente aumentare la memoria del telefonino per far fronte alla crescita vertiginosa di foto e video da conservare. Secondo la UC Berkeley School of Information, dall'epoca dei documenti commerciali e amministrativi dei Sumeri impressi nelle tavolette cuneiformi fino all'invenzione del computer, l'umanità aveva accumulato una memoria di circa 12 esabyte, cioè 1 seguito da 18 zeri, più o meno l'informazione di un dvd della durata di 50 mila anni.

Ma dall'avvento del computer fino al 2006 siamo arrivati a 180 esabyte, dal 2006 al 2011 a 1.600 esabyte e dal 2011 al 2015 a 8 zettabyte (uno zettabyte equivale a mille esabyte). Allo stato attuale, ogni tre anni il numero di dati quadruplica. Non a caso si parla di Big Data: immaginate di riempire ogni giorno 10 volte tutte le biblioteche d'Europa.

A fronte di questo aumento, la quantità di dati



memorizzati per centimetro quadrato raddoppia ogni 18 mesi, una tendenza in rallentamento rispetto alle osservazioni dell'ingegnere elettronico americano Mark Kryder che, in un articolo del 2005 su *Scientific American*, aveva riportato una proiezione di un raddoppiamento ogni 13 mesi. Di conseguenza, gli hard disk sono sempre più grandi e costosi. Più una compagnia è povera più tenderà a cancellare dati per tagliare i costi dell'archiviazione.

Come afferma Luciano Floridi, filosofo dell'etica dell'informazione dell'Università di Bologna e dell'Università americana di Yale, «nella nostra epoca il problema non è cosa salvare ma cosa cancellare. Dal momento che la nostra capacità di immagazzinamento è insufficiente, qualcosa deve essere cancellato, riscritto o neppure registrato. Il nuovo che entra è il primo a uscire: pagine web aggiornate cancellano quelle vecchie, le nuove foto rendono obsolete le precedenti, i nuovi messaggi sono registrati sopra i vecchi».

**Questo mutato rapporto con la memoria sta ri-plasmando il nostro sé.** Cinquant'anni fa i giovani conservavano gelosamente una manciata di foto cartacee dei propri familiari sulle quali costruivano una narrazione di sé stessi, che li autodefiniva. Di contro, i giovani di oggi sono inondati da migliaia e migliaia di fotografie alle quali, proprio per il loro enorme numero e la volatilità, non può essere attribuito il valore di un tempo. «La perdita di memoria su Internet avviene a ritmo continuo» dice Federico Cabitza, ingegnere informatico del dipartimento di Informatica, sistemistica e comunicazione all'Università degli studi di Milano-Bicocca. «Basti pensare al "link rot", cioè quel fenomeno per cui un link cessa di puntare a dati esistenti, cosicché cliccandovi sopra si ottiene una pagina inesistente. Dopo sette anni un link su due è rotto, e in dieci anni il numero aumenta a due terzi».

Uno studio sul *Columbia Journalism Review* ha trovato che gli articoli del *New York Times* pubblicati tra il 1996 e il 2019 avevano un tempo di dimezzamento di circa 15 anni. «Per quanto riguarda gli articoli accademici, l'oblio è in un certo senso il benvenuto» continua Cabitza. «Oggi per fare carriera accademica bisogna pubblicare di continuo, e anche in fretta. La conseguenza è una crescita abnorme di articoli, molti dei quali non sono di valore, oppure lo sono ma non possono essere facilmente notati. Oggi esistono ben 50 mila riviste scientifiche e sono stati scritti quasi 70 milioni di articoli negli ultimi 30 anni; ne vengono pubblicati 5 milioni di nuovi all'anno e 10 al minuto, con un aumento di circa il 5 per cento ogni anno. L'Italia in questa produzione è sesta al mondo, prima di Francia e Giappone, che ha il doppio degli abitanti e una vocazione non meno scientifica».

Il racconto *Funes, o della memoria*, dello scrittore argentino Jorge Luis Borges, rimanda alla condizione umana della nostra epoca. A causa di un incidente, un giovane aveva una memoria così prodigiosa da cogliere ogni dettaglio di ciò che lo circondava. Ma questo lo rendeva incapace di formulare idee e costruire un

rapporto con i suoi simili. La sua morte è in fondo la morte dell'uomo schiacciato dal peso di un eccesso di memoria. D'altra parte, oggi abbiamo una necessità continua di preservare certi tipi di dati che riguardano le nostre vite, dai dati genetici a quelli medici, fino a quelli che riguardano le nostre occupazioni principali. «È il cosiddetto problema della "digital preservation" sul quale si fa ricerca in tre direzioni» afferma Franco Maria Nardini, primo ricercatore dell'Istituto di Scienza e tecnologie dell'informazione «Alessandro Faedo» del Cnr. «Un filone indaga come costruire copie delle informazioni a rischio di essere cancellate. L'obiettivo principale è memorizzare i dati in data center robusti, a prova di incendi e altri eventi catastrofici. Un'altra linea studia come preservare l'integrità del byte attraverso l'inserimento di codici all'interno dell'informazione stessa. Infine, è molto esplorata anche la gestione del dato stesso e delle tecnologie per usufruirne. Un esempio è Knowledge Graph, una sorta di strumento di integrazione dell'informazione. Ne vediamo gli effetti quando cerchiamo una parola su Google. Se si digita su Google "quanto è alta la torre Eiffel?" compare una pagina in cui c'è una risposta strutturata con informazione a contorno. La risposta è una integrazione di informazioni esistenti sul web in un tutto organico».

Tecnologie come queste, che manipolano, classificano o trasmettono l'informazione, sono necessarie per orientarci all'interno di una massa enorme di dati. La nostra civiltà è prevalentemente basata su di esse e questa è una novità storica: l'espansione della civiltà sumera non dipese tanto da tecniche di fruizione e conservazione dell'informazione sotto forma di tavolette cuneiformi, quanto dall'uso di tecnologie in agricoltura o dalla matematica nella soluzione dei problemi quotidiani. Ma oggi siamo strettamente legati alle tecnologie dell'informazione dato l'estremo grado di dipendenza dalla memoria digitale.

Le tracce che lasciamo su social come Facebook sono una porzione della memoria collettiva di cui non è chiara la sorte. «Al momento i grandi servizi web hanno enormi data center geograficamente distribuiti che conservano la memoria di un dato social su disco, per esempio post, fotografie, video e così via» continua Nardini. «Possiamo immaginare che meno un determinato post è consultato più è probabile che venga accantonato, anche se non proprio rimosso; e che i social tendano a conservare i nostri post per fidelizzare il cliente a scopi pubblicitari».

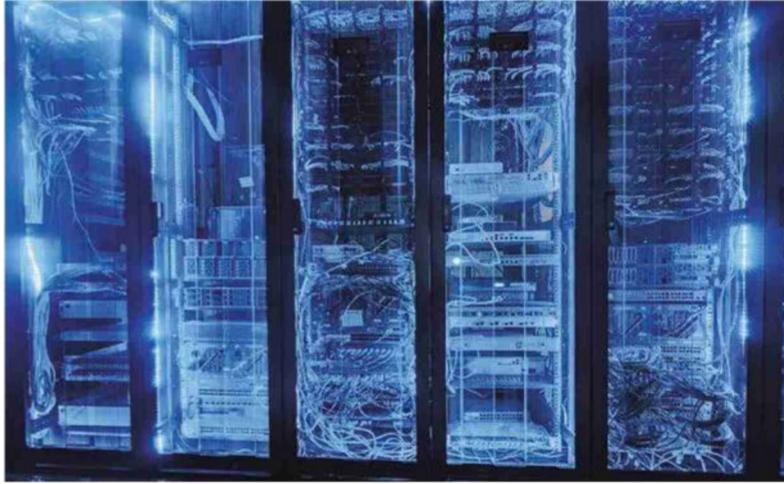
Nessuno sembra curarsene più di tanto. La memoria è l'anima dell'uomo, diceva Umberto Eco. Forse è per questo che siamo disposti a tutto pur di conservarla il più a lungo possibile. ■

© RIPRODUZIONE RISERVATA

**OGNI TRE ANNI  
LA QUANTITÀ DEI  
DATI QUADRUPPLICA**



Peso: 56-100%, 58-75%, 59-83%



Getty Images, iStock

**Abbiamo una  
necessità vitale  
di conservare  
i dati che  
ci riguardano,  
a partire  
da quelli medici.  
È il cosiddetto  
problema  
della «digital  
preservation»,  
oggi allo studio.**



Peso:56-100%,58-75%,59-83%