

Studi Come cambia nei secoli l'uso di parole comuni e nomi propri. Dimenticato Dio, sale Freud, arriva il sushi

Il «genoma culturale» dell'umanità

Harvard ha digitalizzato e selezionato cinque milioni di libri

di ADRIANA BAZZI

Più di cinque milioni di libri, il 4 per cento di tutti i volumi stampati dal 1500 a oggi (classici compresi), ora digitalizzati e raccolti in un immenso archivio elettronico: ecco il genoma culturale dell'umanità. Gli studiosi hanno cominciato ad analizzarlo, concentrandosi sugli ultimi duecento anni e trovando non poche sorprese: la parola Dio, per esempio, ha dominato il XIX secolo, ma è in costante declino, mentre il personaggio più radicato nell'immaginario collettivo è Freud, che supera autorità del calibro di Darwin, Galileo e Einstein.

Si inaugura così l'era della «culturomica», una nuova «omica», in omaggio alla prima, la genomica, che si preoccupa di studiare il funzionamento del Dna umano e di interpretare i tre miliardi di «lettere» che lo compongono. Anche la «culturomica» sfrutta un gigantesco database, quello appunto dei libri digitalizzati, con l'idea di introdurre metodi quantitativi nello studio dei fenomeni culturali, servendosi della potenza dei computer.

Il progetto, i cui primi risultati sono stati appena pubblicati sulla rivista scientifica americana «Science», è partito quattro anni fa alla Harvard University del Massachusetts, in collaborazione con Google e l'Encyclopaedia Britannica.

Finora sono stati convertiti

in formato elettronico 15 milioni di libri (il 12 per cento di tutti quelli pubblicati) e, di questi, ne sono stati selezionati e caratterizzati per data e luogo di pubblicazione, cinque milioni, la maggior parte scritti in inglese, ma con una piccola percentuale anche in francese, spagnolo, tedesco, cinese, russo e ebraico.

I ricercatori di Harvard si sono concentrati su quelli di lingua inglese, apparsi a partire dal 1800 fino ai giorni nostri: romanzi famosi e popolari, come quelli di Mark Twain, ma anche *Lo straniero* di Albert Camus; saggi come quello sulla relatività di Albert Einstein, ma anche (e sono la maggioranza) opere di autori poco conosciuti come *A short dichotomous key to the hitherto unknown species of Eucalyptus*, un manuale sull'eucalipto compilato dal botanico tedesco Johann Georg Luehmann.

Nessun umano sarebbe in grado di leggere tutti questi libri. Chi volesse cimentarsi in un'impresa del genere, limitandosi alle sole novità apparse nel 2000, impiegherebbe (calcolano i ricercatori di Harvard) otto anni, senza interruzioni per mangiare o per dormire, al ritmo di lettura di 200 parole al minuto. I computer sono molto più veloci non soltanto nel leggere, ma anche nel calcolare la frequenza di apparizione di una determinata parola o nel cercare legami fra termini e contesti culturali. A partire dal 1861, la pa-

rola «schiavitù», per esempio, compare più di 21 mila volte, nel corpus dei libri considerati, su un totale di oltre 386 milioni di parole (la frequenza si misura sul rapporto fra singola parola e il totale delle parole), con due picchi: uno, nel periodo della guerra civile americana (all'inizio del 1860) e, l'altro, nell'arco di tempo fra il 1955 e il 1968, quando aveva preso piede il movimento per i diritti civili.

Altre due voci: uomo e donna. Se il divario fra uso della prima e della seconda era enorme all'inizio dell'800, con un'ovvia, netta prevalenza della prima, adesso la situazione si è invertita: il sorpasso della «donna» è avvenuto attorno agli anni Novanta, a conferma che l'onda lunga del femminismo e la moda degli studi di genere hanno avuto un'influenza non da poco.

Nomi propri adesso, per misurare la fama che li accompagna o le eventuali censure nei loro confronti: la scrittrice Virginia Woolf e il padre della teoria dell'evoluzione Charles Darwin, il pittore ebreo Marc Chagall e l'ex presidente americano Bill Clinton, il regista Steven Spielberg e l'attrice Marilyn Monroe, il rivoluzionario Che Guevara e l'astronauta Neil Armstrong, per citare i più noti, analizzati insieme ad altri, un po' meno famosi, di artisti e scienziati, politici e attori. Ecco le conclusioni di Jean Baptiste Michael, ricercatore al Department of Psychology a Harvard e capo del progetto:

la censura lascia impronte ben quantificabili e Marc Chagall, ampiamente citato nei libri a partire dagli anni 10 del Novecento, continua a esserlo in quelli successivamente pubblicati nei Paesi anglosassoni, mentre viene quasi ignorato in Germania dove il suo nome appare solo una volta, fra il 1936 e il 1944.

A proposito di notorietà, i ricercatori osservano che, oggi, si raggiunge più rapidamente che in passato, ma dura meno. I primi a diventare famosi sono gli attori, seguiti dagli scrittori, che però rimangono più a lungo sulla cresta dell'onda. I politici arrivano tardi, ma la pazienza paga: sono i più longevi. Poca fama, invece, per medici e biologi, quasi nessuna per i matematici.

La «culturomica» può, però, diventare un mezzo per indirizzare la ricerca scientifica, sia quella di epidemiologia storica (i «picchi» nell'uso del termine influenza, in letteratura, va di pari passo con le grandi epidemie del passato), sia quella futura sull'alimentazione e sulle diete, con un «menù» che tenga conto di pizza, pasta e sushi (di cui si parla con sempre maggiore frequenza negli ultimi anni), ormai preferiti alla classica bistecca.

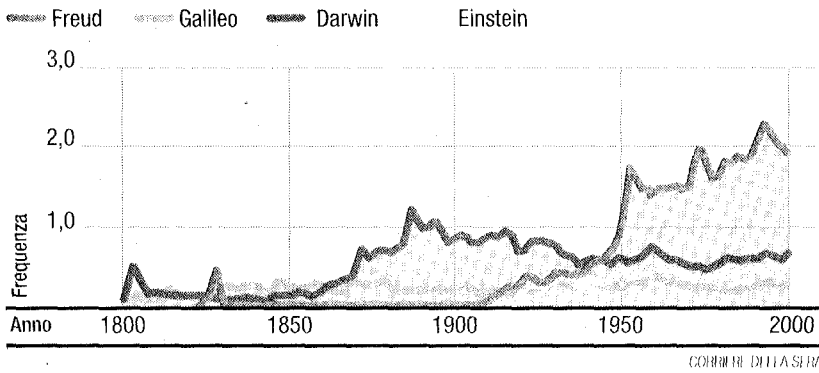
E Dio che fine ha fatto? Non è morto, scrivono alla fine del loro lavoro i ricercatori americani: ha però bisogno di una nuova pubblicistica.

© UNIVERSITÀ HARVARD

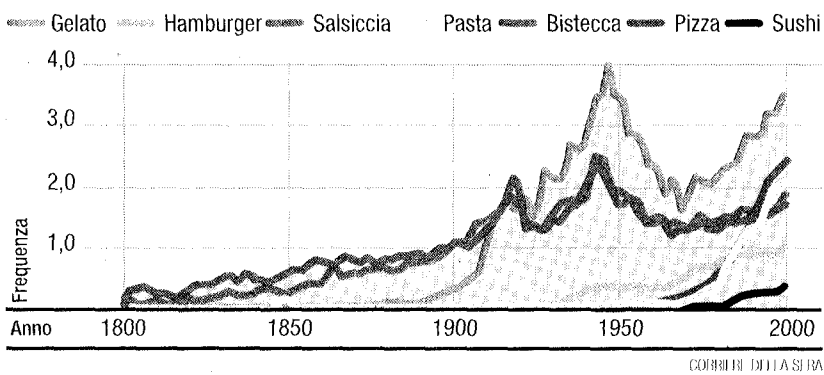
Discipline

La ricerca di «Science» inaugura la «culturomica», disciplina che rende omaggio alla genomica

1900, il secolo di Freud. Il confronto tra le scienze

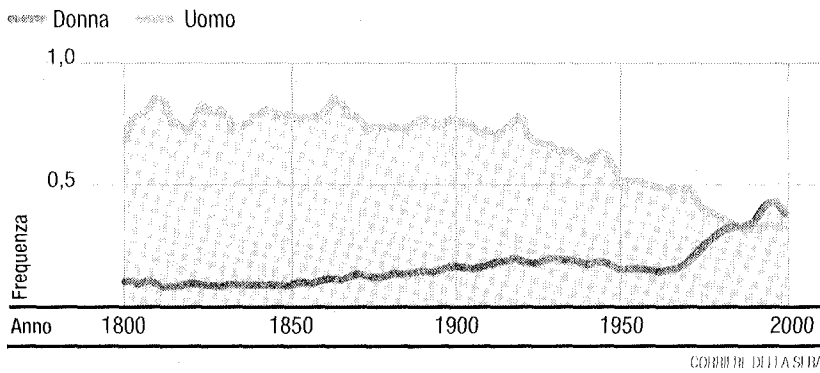


Alimentazione, il cibo dell'anno



Dall'alto:
Mark Twain,
Albert
Camus
e J.G.
Luehmann

Il confronto tra i sessi. La parola «donna» supera «uomo»



1800, il secolo della religione. La parola «Dio»

