

Così il software Ngram Viewer calcola le "presenze" nei libri digitalizzati

IL CANONE di GOOGLE CONTA PIÙ PLATONE O KANT? CE LO DICE UN ALGORITMO

MAURIZIO FERRARIS

Cosa cerchiamo quando cerchiamo su google? E cosa cerchiamo di più nella biblioteca d'Alessandria 2.0, cioè su Google Books? Insomma qual è il canone di Google? Grazie al Google Ngram Viewer possiamo capirlo. Si tratta di uno strumento disponibile dal dicembre 2010 che permette di realizzare grafici in cui si visualizza la frequenza anno per anno di combinazioni di lettere (n-grams), parole ("spread") o frasi ("Impero Romano") presenti negli oltre 5,2 milioni di libri digitalizzati da Google sino al 2008, che corrispondono al 4% di tutti i libri pubblicati. L'estensione cronologica dei libri copre il mezzo millennio della diffusione massiccia della stampa, dal 1500 a oggi, e si riferisce a cinquecento miliardi di parole in inglese (nelle varianti britannica e americana), francese, tedesco, spagnolo, russo e cinese. Le parole italiane sono contate in base alla loro occorrenza in altre lingue. La cosa interessante, suggerisce la voce di wikipedia da cui traggio queste informazioni, è che si possono effettuare comparazioni attraverso dei grafici: di persone (Galileo, Darwin, Freud, Einstein) di composizioni musicali (ad esempio la Quinta e la Nona di Beethoven), o di entità fittizie (Pegaso è più popolare di Sherlock Holmes?).

Il browser permette di cercare in differenti raccolte di libri, i "corpora", cioè il plurale di "corpus", un insieme omogeneo di libri. C'è dunque una differenza tra il corpus dei libri britannici e quello dei libri americani, come si capisce facilmente se si traccia un grafico delle occorrenze di "centre" e di "center" (rispettivamente la

estensori della voce, è anche possibile trovare prove di attività censorie, per esempio del nome "Marc Chagall" nella Germania nazista.

La base cronologica più attendibile, avvertono i curatori, è il periodo tra il 1800 e il 2000 nel corpus inglese (sommatoria di quello britannico e americano). Prima di quella data la quantità di libri scansionati in Google Books è statisticamente insufficiente, dopo il 2000 i corpora sono ancora in formazione, e dunque la presenza di libri è legata a motivi contingenti. Una volta che si siano prese tutte le precauzioni del caso può nascere una nuova scienza, la "culturonomia", ossia l'economia dei flussi culturali.

Google Ngram Viewer è basato sul browser "Bookworm" creato dall'osservatorio culturale di Harvard, ed è dettagliatamente descritto sull'edizione online di *Science* del 16 dicembre 2010, «Quantitative Analysis of Culture Using Millions of Digitized Books». A portare la mia attenzione su Google Ngram Viewer è stato però, qualche settimana fa, un collega inglese, l'ontologo analitico Barry Smith, che mi ha mandato una mail che recitava: Aristotele diventa più importante di Derrida. Aggiungendo che «solo l'ontologia cresce in termini assoluti». Sotto c'era il link a un diagramma da cui si evince che nel periodo 1984-2008 (dunque in un periodo problematico, data l'inaffidabilità dei dati dopo il 2000) e nel corpus inglese (cioè sul corpus più affidabile) Aristotele è più citato di Marx, Foucault e Derrida, mentre l'ontologia, sia pure in basso nella classifica, traccia la linea di una crescita continua. Ho chiesto a Smith da dove avesse tratto il grafico, lui mi ha detto di Google Ngram Viewer, e ha aggiunto un

gli articoli non ne parlano affatto. Ma, sinceramente, mi sembrano obiezioni deboli, perché suppongono una inverosimile disarmonia nei processi culturali. Ed è così che, per gioco ma sul serio, ho passato un paio d'ore a fare "ricerche" di cui vorrei comunicarvi i risultati.

Ho prima di tutto cercato di fare confronti in ambiti omogenei. Per esempio fra i tre principali autori post-strutturalisti, Foucault, Derrida e Deleuze nel periodo 1990-2008. Dove si conferma la maggiore presenza di Foucault rispetto a Derrida che si vedeva nel grafico su Aristotele, mentre Deleuze è all'ultimo posto sia in inglese, sia in francese, ma qui Foucault è più in alto e Deleuze molto più vicino a Derrida.

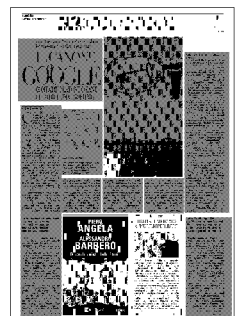
Il che, in effetti, confermerebbe una ipotesi intuitiva, e cioè che Foucault è penetrato più capillarmente, negli studi degli storici, dei sociologi, degli psichiatri, e nel dibattito politico. Questi dati sono confermati dai risultati in francese dal 1960 (epoca della fioritura) al 2000 (termine del periodo di affidabilità). Vale la pena di osservare che la morte di Foucault alla metà degli anni Ottanta e quella di Deleuze a metà degli anni Novanta non intacca l'ascesa relativa. Dunque fu vera gloria.

Se però dalla triade dei post-strutturalisti passiamo, sullo stesso periodo, a quelli che in un certo senso sono i loro eredi (beninteso, in materia di impatto mediatico), e cioè la triade pop Michel Onfray, il La Mettrie redivivo, Peter Sloterdijk, neoapocalittico spengleriano, e Slavoj Žižek, psicoanalista dialettico, emergono dati forse meno prevedibili. In inglese

Lo strumento fa nascere una nuova scienza la "culturonomia" ossia l'economia dei flussi culturali

dizione britannica e americana di "center") nei due corpora. E, osservano gli

link riferito al periodo (più affidabile) 1700-2008 commentando che «dimostra come l'ontologia, dopo 300 anni, stia finalmente trionfando sulla metafisica». Ovviamente si potrebbe obiettare che il trionfo dell'ontologia potrebbe anche essere un successo malfamato, in cui in montagna di libri si dissuadono i lettori dall'occuparsi di ontologia. O che magari mentre i libri parlano tanto di ontologia



se Zizek straccia Sloterdijk e Onfray sia in termini assoluti di crescita, sia in termini relativi. In italiano (per quel che vale) la classifica è la stessa, ma tutti e tre risultano in salita netta dalla fine degli anni Novanta, che in effetti coincidono con il momento della loro esposizione mediatica.

Ma veniamo ai grandi nomi, agli evergreen della filosofia. È forte la tentazione di confrontare Platone, Kant e Wittgenstein (un antico, un moderno, un contemporaneo). In inglese Platone e Kant, l'antico e il moderno, sono molto vicini, Wittgenstein sotto, molto staccato. I classici continuano ad avere la meglio, anche nel periodo 1990-2008, dunque con un confronto ad armi pari (non avrebbe senso cercare Wittgenstein nell'Ottocento, o Kant nel Cinquecento).

Veniamo alla triade degli idealisti, Fichte, Schelling, Hegel. In inglese non c'è gara, stravince Hegel, gli altri due sono vicinissimi in valori molto bassi. In tedesco, Hegel è in vantaggio, ma meno, e comunque è in declino. E per la coppia dei neodealisti italiani, in inglese, Croce ha un vantaggio, ma cala, Gentile recupera (situazione molto simile in italiano).

E come vanno le cose se confrontiamo gli antichi tra loro? Per esempio Platone, Aristotele, Epicuro. In inglese Aristotele ha la meglio, in italiano, invece, ha la meglio Platone. La tendenza si ribadisce in tedesco, dove Platone ha un vantaggio enorme. E si conferma una dissonanza culturale significativa e profonda: gli analitici sono aristotelici, i continentali platonici.

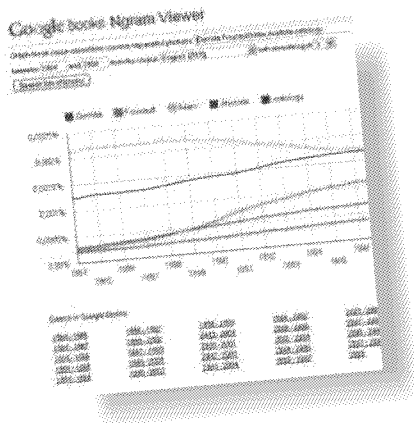
E, invece, come tengono Nietzsche e Marx, tra il 1990 e il 2008? In tedesco Nietzsche è stabilmente piatto e Marx è in de-

Tra i duelli, quello tra Marx e Nietzsche nel '900: sia in inglese che in tedesco il primo è in vantaggio sul secondo

clino. Situazione pressoché identica in inglese. Invece in francese dal 1997 inizia per entrambi una rimonta (ma comunque anche qui Marx è sopra Nietzsche di parecchio). In italiano con la svolta del secolo Marx ha una lieve rimonta, Nietzsche un lieve declino, e soprattutto, sia pure con Marx in vantaggio, il distacco tra i due non è così grande come in altre tradizioni. Il dato davvero sorprendente lo abbiamo in russo, dove Nietzsche è uniformemente basso, ma Marx è in discesa sino al 1997, ma poi inizia una risalita vertiginosa. A questo punto è forte la tentazione di confrontare Nietzsche e Marx dal 1900 al 2000, esperimento che ha dell'istruttivo, perché ne vien fuori il diagramma ideologico del secolo. In inglese, Marx è sempre in vantaggio, con un picco nel 1975, poi inizia la discesa. Non è così diverso in tedesco, dove si mantiene il vantaggio di Marx su Nietzsche. L'italiano si distingue per una ascesa (relativa) di Nietzsche dagli anni Settanta, e non stupisce pensando a tutti i libri che sono stati scritti nel periodo.

Fin qui i duelli seri. Ma utilizzando questo strumento si può fare il confronto tra Topolino e Paperino o tra Frankenstein e Frodo. Un modo per capire un po' le tendenze alte e basse del sapere (digitale).

© RIPRODUZIONE RISERVATA



Le ricerche

Una videata di Google Ngram Viewer dove sono messe a confronto le "frequenze" di Derrida (blu), Foucault (rosso), Marx (arancione), Aristotele (verde), e "ontologia" (viola).

