

# MOTORE IN

**il punto**  
di Matteo Fago

G\$\$\$gle

Circa 20 anni fa nasceva il *world wide web* al Cern di Ginevra. Inventato come strumento di condivisione delle informazioni scientifiche, divennero ben presto evidenti le sue potenzialità di comunicazione commerciale. Oggi è un grande mercato in cui ogni giorno si effettuano milioni di transazioni. Per orientarsi nella grande Rete furono costruiti degli indici, il più celebre e innovativo dei quali, Google, compie 10 anni in questi giorni. Google ha avuto uno straordinario successo perché ha inventato un nuovo modo di catalogare internet, rendendo la ricerca di informazioni più precisa rispetto ai predecessori. Ma non solo. L'altra grande invenzione è stata un nuovo modo di fare pubblicità. Invece di vendere spazi, si "vendono" visitatori. Ogni volta che lo usiamo, Google ci propone pubblicità in base a quello che stiamo cercando. Chi compra la pubblicità paga soltanto per il potenziale cliente che arriva sul proprio sito web. Chiaro quindi che la pubblicità da mostrare deve essere quella giusta per attirare la nostra attenzione di consumatori. Ed è tanto più azzeccata tante più sono le informazioni che Google ha su di noi. L'accusa dalle associazioni dei consumatori ma anche dall'Unione europea è quella di fornire servizi gratuiti (posta, documenti, mappe, fotografie, video, blogs, statistiche ora anche un browser) in cambio di informazioni personali. Ma anche di avere il potere di censurare a piacimento parti di internet. Paranoie all'americana? Forse. Di certo ci sono i soldi, tantissimi: 16 miliardi di dollari in ricavi pubblicitari nel 2007, 10 nel primo semestre 2008 e prospettive di crescita a 2 cifre per i prossimi anni.



Una storia di grande successo. Che come tutte le ascese che si rispettino, contiene luci e ombre. Un algoritmo sceglie cosa farci leggere, in base solo alla popolarità. Il perché ce lo spiegano gli autori di *Dark side of Google* di Gruppo di ricerca Ippolita - [ippolita.net](http://ippolita.net)

**S**ono in molti a pensare che Google e la Rete siano più o meno la stessa cosa, o meglio che il motore di Mountain view "contenga" tutto il web. Ne è testimone il bottone "Mi sento fortunato", che ci catapulta direttamente alla pagina che stiamo cercando. Incredibile. Google sostiene di poter intercettare ed esaudire perfettamente i nostri

desideri di ricerca.

"Mi sento fortunato" è l'esemplificazione della filosofia oracolare googoliana. Questa religione nascente ha un bug concettuale, di fondo.

Semplificando: il sistema "democratico" e "oggettivo" di Google per ordinare il mondo si fonda sul Pagerank. Questo algoritmo ordina i risultati in base al numero di link in entrata per

# PANNE



IL GRUPPO DI STUDIO

## Chi sono

Il gruppo di ricerca Ippolita è nato nel 2005 a Milano attorno alla stesura del testo copyleft *Open non è free, comunità digitali tra etica, hacker e mercato globale*, Eleuthera edizioni. La scelta del nome collettivo, la rinuncia all'autorialità individuale sta proprio a sottolineare come l'opera, se di buona qualità, non abbia alcun bisogno di padri. Autrice collettiva dunque perché come dicono loro: «Bisogna mettere in comune discipline diverse, tradurre e tradire l'accademia e metterci a fare divulgazione». Il loro secondo volume *Luci e ombre di Google. Futuro e passato dell'industria dei metadati*, Feltrinelli 2007 è l'unico libro in circolazione in Italia che sottolinei gli aspetti critici e le problematiche relative al fenomeno Google. Il testo è stato tradotto in francese nel 2008 per la casa editrice Payot&Rivage, ed è in attesa di traduzione in spagnolo.

I libri di Ippolita sono anche scaricabili gratuitamente dal sito del gruppo: <http://Ippolita.net>, oggi sono state scaricate più di 30mila copie tra download diretti sul sito e altri esterni.

ogni sito e in base all'importanza che i siti linkanti a loro volta posseggono. Più un nodo sarà connesso più acquisterà posizioni di testa nella lista dei risultati. Google parla esplicitamente di "voto", poiché ogni link equivale a un voto. Un sito fortemente votato (linkato) si troverà in una posizione di vantaggio nel ranking rispetto a un altro meno conosciuto. Scambiando la popolarità e il filtro tecnologico per democrazia, si può arrivare a definire Google una "democrazia elettronica". Infatti, sostengono a Mountain view, «lo spider non si interessa dei contenuti di una pagina ma solo dei link»: sembra proprio che ci troviamo di fronte a una tecnologia in grado di garantire l'oggettività di un risultato. Il ragionamento sottinteso è che il Page-rank, un algoritmo matematico, in

quanto oggetto tecnologico, non giudica, non può mentire, è buono per natura. Peccato che in questo meccanismo di assegnazione matematica di un valore si scorga la convinzione di fondo che "più saranno conosciuti più saranno famosi", un diktat del nostro tempo. Anziché dar voce alla creatività e alla ricerca scientifica, Google legittima il primato della Paris Hilton di turno, la quale non ha ovviamente nessun merito se non quello di essere giustappunto famosa.

**Ma entriamo nel merito.** Quest'idea del link come espressione di un voto positivo ha una sua precisa origine culturale. Si tratta infatti di una traduzione in

## Anziché dar voce alla creatività, legittima la "Paris Hilton" di turno

senso matematico della citazione scientifica (il citation index ndr): più un articolo scientifico viene accolto in maniera positiva da altri scienziati (che si suppone lavorino in maniera indipendente), più l'articolo stesso viene con-

siderato importante e meritevole. Brin e Page, i fondatori di Google, traslano nel proprio campo di ricerca questo approccio e sviluppano la teoria secondo cui il numero di link che conduce a una pagina è un modo per valutare il valore della pagina stessa e in un certo senso la sua qualità. Il criterio dell'autorevolezza del recensore scientifico diventa più banalmente la misurazione della popolarità di un sito. Ma l'e-

spressione della democrazia digitale è ben più complessa. Il metodo che teoricamente offre le garanzie maggiori di oggettività è il cosiddetto "doppio cieco" (*double-blind*): prima di essere accolto in una rivista scientifica, un articolo viene sottoposto ad almeno due ricercatori di chiara fama, che non devono conoscere il nome dell'autore. Il secondo passaggio riguarda l'estensore dell'articolo che non deve conoscere il nome dei recensori. Insomma nella "Repubblica scientifica" il sistema è un po' più complesso, e non può essere applicato a tutte le risorse della Rete, perché queste non hanno un valore univoco. La risorsa "Paris Hilton", in un sistema di ricerca che si picca di gestire oggettivamente l'interesse delle conoscenze umane in Rete, non ha lo stesso valore della risorsa "Rita Levi Montalcini". Il Pagerank di Google si presenta

perciò come una sorta di *citation index* dei poveri e veicola un'idea di democrazia quantomeno demagogica.

**L'algoritmo del Pagerank** dunque è utile per rintracciare il senso comune riguardo a un certo argomento. Insomma, funziona per ricerche generaliste su risorse note. Ma quando la nostra intenzione è di tipo esplorativo, ossia quando non sappiamo esattamente cosa cerchiamo, ma siamo in una fase di ricognizione sulle informazioni disponibili, ecco che il sistema di valutazione googoliano perde di senso. Sia perché scomodo da utilizzare, sia perché superficiale. Se dobbiamo, per esempio, scrivere una tesi è quasi certo che i primi risultati ci interessino poco. Cerchiamo

## Il Pagerank funziona per ricerche generaliste su risorse note

probabilmente una risorsa poco linkata, un articolo specifico, magari scritto in gergo tecnico e classificato da Google tra le ultime pagine. Anzi, una risorsa può essere poco conosciuta ma interessante, se non anche fondamentale, su un piano di ricerca tutto individuale.

Improvvisamente l'interfaccia di Google diventa scomoda: impossibile vedere la relazione fra i risultati forniti in una classifica a liste, distinta solo per alto e basso. L'unica cosa che rimane da fare è sfogliare l'universo dei risultati a mano, pagina dopo pagina.

**Aggiungete a tutto questo** che, attraverso le tecniche di profilazione (la raccolta dei dati sull'utente), Google è talmente sicuro di sapere cosa stiamo cercando



Una manifestazione contro la censura che il motore

## il profiling Come difendersi

**Per profilazione si intende la raccolta di dati che identifica un utente volta a fornire servizi e pubblicità personalizzati online.** Questi dati non ci possono essere sottratti, ma li forniamo noi per necessità o per leggerezza, non hanno alcun valore se

raccolti in un singolo momento, lo ottengono se collezionati per un lungo periodo. Il profiling non è illegale ma gli utenti non hanno alcun controllo e voce in capitolo, motivo che ha spinto gruppi e singoli programmatori a sviluppare alcune contromisure per contrastare il tracciamento. Ecco ne alcuni:

**Track me not** è un'estensione per Firefox, funziona per tutti i motori, non solo per Google. TMN invia periodicamente ricer-

che casuali. Il motore riceverà quelle reali mescolate ad altre. In questo modo Google non sarà in grado di recepire gli argomenti che ci interessano, poiché questi compariranno in modo offuscato.  
<http://www.mrl.nyu.edu/~dhowe/trackmenot/>

**Scroogle** è un sito internet che si colloca tra l'utente e Google rendendo anonime le ricerche. Passando attraverso Scroogle si diventa praticamente invisibili a Google

che ci ripropone sempre gli stessi risultati. Chi sostiene il contrario non ha mai condotto una ricerca per motivi di studio o lavoro o gioco al di là delle rotte più navigate. Profiling più Pagerank uguale piccolo stagno di Google: altro che tutto il web. Le ricerche tramite Pagerank sono tendenzialmente funzionali, evitano al massimo la possibilità di broken link (link rotti) o di informazioni diverse da quelle precedentemente archiviate. Il problema è che in questo modo gli utenti sono indotti a credere erroneamente che internet sia un mondo chiuso, connesso, completo, privo di strade poco illuminate o di percorsi preferenziali, poiché sembrerebbe che, data un'interrogazione, si giunga sempre al risultato "giusto". Tuttavia questo non significa affatto che su internet siano assenti isole di dati. L'integrità referenziale proposta dalla base dati di Google sottintende l'i-

dea di un mondo unico per tutti, chiuso e finito. Al contrario, tracciare un percorso in una Rete complessa significa compiere un'esplorazione che determina sempre dei risultati relativi e parziali.

Il sogno di un Google contenitore di tutta internet è un'idea particolarmente comoda, utile per sostenere tutte le caratteristiche che lo rendono un "servizio unico", un dispensatore di verità. Questo meccanismo ci illude che la libertà consista nell'ottenere una qualità totale a costo zero, immediata, perché ci sentiamo fortunati. Sappiamo invece che in un sistema reticolare complesso come la Rete non esistono verità assolute, ma solo autorità distribuite a seconda del percorso che si desidera af-

frontare, o anche solamente in funzione del tempo che si è disposti a investire nella ricerca. La qualità dipende interamente dalla nostra soggettiva percezione dell'accettabilità del risultato.

**Google usa le sue potenzialità** soprattutto per creare comportamenti standard, profilare e ridurre il bacino a un'ideale risposta "esatta". D'altra parte, spesso gli utenti si accontentano di ricerche superficiali o peggio usano le proprie capacità tecnologiche per procacciarsi il migliore posizionamento all'interno del modo localizzato da Google. Perciò se chiediamo cambiamento, dobbiamo essere i primi a cambiare mentalità, cominciando per esempio a interessarci alle tecniche per garantire anonimato e privacy (vedi box). Ci sono, si installano sul browser e non sono

di difficile utilizzo. Ne trovate alcune sul nostro sito. Ma non esistono bacchette magiche: quello che deve cambiare davvero è il nostro approccio alla ricerca e alle tecnologie digitali in generale. Per farlo occorre cominciare a informarsi e informare: inseguire l'ultima applicazione favolosa sul mercato, con spirito techno-entusiasta, peggiora solo la situazione di passività digitale e alimenta la tecnocrazia. Intendiamoci, noi amiamo l'informatica e per questo vi invitiamo a sviluppare un'attitudine "hacker": metteteci le mani sopra, andate a vedere come funziona, cosa c'è dentro, cosa fa esattamente quella roba lì. Fermatevi a valutare.

## Vederlo come un contenitore di tutta internet è un'utopia



di ricerca applica in alcuni Paesi. A sinistra, tre momenti di lavoro e relax a GooglePleex

perché sarà il sito stesso a raggiungere il motore di ricerca per noi. Navigando su Scroogle si avranno gli stessi risultati di Mountain view, senza pubblicità e profilazione. Scroogle è stato creato da Daniel Brandt, ideatore di un altro sito anti-google: Google watch, osservatorio che si occupa di raccogliere malefatte più o meno note. <http://scroogle.org/>

**Scookies** è un'estensione per Firefox e funziona per tutti i motori di ricerca e siti inter-

net che cercano di tracciare gli utenti tramite i cookie. Scookies cambia i cookie degli utenti mescolandoli gli uni agli altri. In questo modo vengono alterati i profili di ciascuno intorbidendo il tracciato originale. È una creazione di Andrea Marchesini che lo ha immaginato come uno strumento comunitario, infatti ogni utente può segnalare siti che fanno profilazione contribuendo a migliorare l'anonimato di tutti gli altri. <http://www.autistici.org/bakunin/scookies/>





GoogleEarth, il primo di una serie di applicativi per osservare il nostro pianeta, la Luna e Marte attraverso satelliti e missioni spaziali

## IDATI

## Numeri da capogiro

**Google Inc.** Società quotata in borsa

**Nascita:** 7 settembre 1998

**Quartier generale:** Mountain view, California

**Fondatori:** Sergey Brin e Larry Page

**Impiegati:** 19.604 (giugno 2008)

**Fatturato 2004:** 2,5 miliardi di dollari (ingresso in borsa)

**Fatturato 2006:** 10,6 miliardi di dollari

**Utile netto 2006:** 3 miliardi di dollari

La moneta della conoscenza siamo noi, sono i nostri percorsi, le nostre esplorazioni, la nostra unica e inimitabile impronta digitale. Google è la pellicola più sensibile mai costruita e tutti i grandi player dell'informatica che investono in social network si sono adeguati a questo standard. Google non è buono, a dieci anni di distanza lo slogan "don't be evil" è più ridicolo che mai. Ma Brin e Page non sembrano turbati. Mentre pranzano seduti al ristorante vegetariano del GooglePleex, dietro la locandina "Free Tibet", la loro ultima creatura, il browser Chrome, ci imbastisce con l'ennesima balla su bontà e sicurezza. Un altro capitolo della guerra contro Microsoft per il controllo del web. Ci fideremo, certo. Senza dubbio.

## Quando usiamo i loro servizi, forniamo informazioni su noi stessi

I servizi che usiamo li paghiamo con qualcosa di più prezioso del denaro. Ogni volta che utilizziamo i servizi di ricerca, di posta elettronica, YouTube, Blogspot, Office, news... e tutte le centinaia di applicazioni collegate, ecco, in ognuno di quei momenti noi stiamo fornendo a Google informazioni su noi stessi. Si chiama profilazione e ufficialmente Google la usa per farci pubblicità mirata. Noi non abbiamo alcun controllo su quei dati, non sappiamo nemmeno come vengano gestiti. Questo non è affatto buono o democratico. Google vigila costantemente su ognuna delle miriadi di richieste fatte in tutto il globo terrestre e ci guadagna parecchio denaro, più di quanto si possa immaginare. Ogni volta che immettiamo una parola nella barra di ricerca, ogni volta che facciamo click o riceviamo un'e-mail da un account Gmail, ogni volta che si muove una foglia sul web, Google guadagna dei soldi.

L'industria dei meta-dati è tutto ciò che non riguarda il dato in sé, ma il complesso delle informazioni che vi ruotano attorno (chi, dove, in relazione a cosa, per quale comunità etc...). Questo è il nuovo filone aurifero dell'economia informatica di oggi. Questo tipo di mercato conta sull'inconsapevolezza dell'utente, sulla leggerezza con la quale espone le sue informazioni personali e quelle di coloro che lo circondano. È ur-

gente elaborare una visione complessiva, trasversale, critica. Internet non si sta espandendo, Google non ci sta rendendo più intelligenti, ci sono solo un pugno di protagonisti che ricombinano lo spazio in sottoreti comunitarie sempre più omogenee. Foto, video, blog, geolocalizzazione sono solo contenitori entro cui accumulare informazioni sui gusti, le tendenze, le paure degli utenti, i preziosi meta-dati che fluttuando foraggiano una fetta notevole dei mercati finanziari. Non è difficile intuire come, una volta costruite le infrastrutture e gli standard di portabilità, sia possibile eterodirigere i soggetti, creando desiderio indotto, stimolando a fornire sempre più informazioni, fino a una vera e propria formazione a distanza, all'utilizzo di sempre nuove piattaforme. Nulla di nuovo, ma i mezzi con cui la creazione spontanea di senso viene convogliata e messa in produzione sono più sofisticati. Soprattutto, il mondo di internet non si limita a questo. La Rete non è uno spazio-tempo fatto solo per il dio denaro, ci sono un sacco di altre cose. Come nella realtà non virtuale, molte non sono affatto buone, mentre altre potrebbero servirvi per difendervi dai sedicenti benefattori. Sta a voi cercarle. Come si diceva: cambiare mentalità, magari a cominciare dall'uso del buon vecchio software libero. Addirittura, "trovare" potrebbe essere meno interessante dell'atto stesso di "cercare"; anzi, forse è piacevole non trovare affatto, perché se non si riesce a trovare significa che si è presi nel gioco del percorso, appassionatamente impegnati a cercare. ■