

Google, il più famoso dei motori di ricerca, per ordinare i risultati (le pagine web) ottenuti dalla ricerca di un utente, unisce due metodi: la rilevanza e l'importanza. Assegna a ogni pagina un punteggio di rilevanza che misura quanto la pagina è attinente rispetto ai termini inseriti dall'utente. Ma il vero punto di forza di Google, che gli ha consentito di surclassare gli altri motori di ricerca, è il metodo che assegna il giudizio di importanza alle pagine. L'algoritmo prende il nome di PageRank e assegna un giudizio di importanza a ogni pagina indipendentemente dalle interrogazioni degli utenti, ma unicamente in funzione della rete dei collegamenti ipertestuali tra pagine web, rete che forma il world wide web. La tesi adottata dal PageRank è la seguente: "una pagina è importante se viene citata da altre pagine importanti". Il metodo del PageRank è stato proposto da Sergey Brin e Larry Page nel 1998 durante il dottorato all'università di Stanford ed è oggi uno degli ingredienti principali del motore di ricerca Google. Il vero contributo di Sergey Brin e Larry Page è stato quello di partire da una idea non completamente originale e trasformarla, per costruire qualcosa di incommensurabile utilità nell'era dell'informazione. Nel contesto del web, una citazione è un collegamento ipertestuale (link) da una pagina sorgente (la pagina che contiene il link) a una pagina destinatario (la pagina verso cui il link è diretto). La definizione, apparentemente circolare, è in realtà risolvibile usando strumenti matematici profondi, quali il teorema di Perron-Frobenius che garantisce l'esistenza e l'unicità di una soluzione, il metodo delle potenze che permette di calcolarla in modo iterativo. Più precisamente, il metodo del PageRank usa i seguenti tre fattori per determinare l'importanza di una pagina web: 1. il numero di link ricevuti dalla pagina; 2. l'importanza delle pagine sorgenti che contengono i link alla pagina; 3. il numero di link contenuti nelle pagine sorgenti.

Per capire l'importanza del metodo, appliciamolo a una ipotetica rete in cui i giornalisti sono tra loro collegati da giudizi di riconoscimento. Possiamo reinterpretare i tre fattori del PageRank esposti sopra nel seguente modo:

1. un giornalista è tanto più prestigioso quanto più elevato è il numero di riconoscimenti che ottiene da parte di colleghi;
2. riconoscimenti ricevuti da giornalisti importanti valgono più di giudizi ottenuti da giornalisti poco stimati;
3. infine, se un giornalista elargisce abbondanti riconoscimenti, i suoi giudizi sono inflazionati e valgono meno.

*(M. Franceschetti, lavoce.info 30-03-2010)*