

# I link non sono tutti uguali: 10 illustrazioni sulla valutazione dei link da parte dei motori di ricerca

Di **Rand Fishkin**



Articolo originale:

## All Links are Not Created Equal: 10 Illustrations on Search Engines' Valuation of Links

27 Maggio 2010

<http://www.seomoz.org/blog/10-illustrations-on-search-engines-valuation-of-links>



Foto di [geishaboy500](#)

Nel 1997 i fondatori di Google crearono un [metodo algoritmico](#)<sup>1</sup> per determinare l'importanza e la popolarità di una pagina sulla base di alcuni principi cardinali:

- Sul web i link possono essere interpretati come voti espressi dalla fonte, nei confronti del destinatario
- Inizialmente tutti i voti vengono considerati uguali
- Eseguendo l'algoritmo su un grafo di link, le pagine che ricevono più voti diventano più importanti

- Le pagine più importanti danno voti più importanti
- I voti una pagina che può dare, sono una funzione dell'importanza della pagina, divisa per il numero dei voti (link) dati.

Tale algoritmo, naturalmente, era il PageRank, e cambiò il corso della ricerca sul web, valorizzando in modo straordinario i primi tentativi di Google di miglioramento della qualità e della pertinenza dei risultati. Al diffondersi della conoscenza del PageRank, quelli che avevano un tornaconto ad influenzare l'ordine dei risultati delle ricerche (i SEO) trovarono il modo di sfruttare queste informazioni per i loro siti e pagine web.

Ma Google non stette a dormire sugli allori nel campo dell'analisi dei link; si rinnovarono, sfruttando i segnali come il testo dei link, la fiducia, lo schema "[hubs & authorities](#)"<sup>2</sup>, la classificazione dell'argomento e perfino l'attività umana, per influenzare il valore che un link può veicolare. Eppure, purtroppo, molti che lavorano nel campo SEO, sono ancora all'oscuro di questi cambiamenti e del modo in cui questi impattano il marketing e le best practices di acquisizione link.

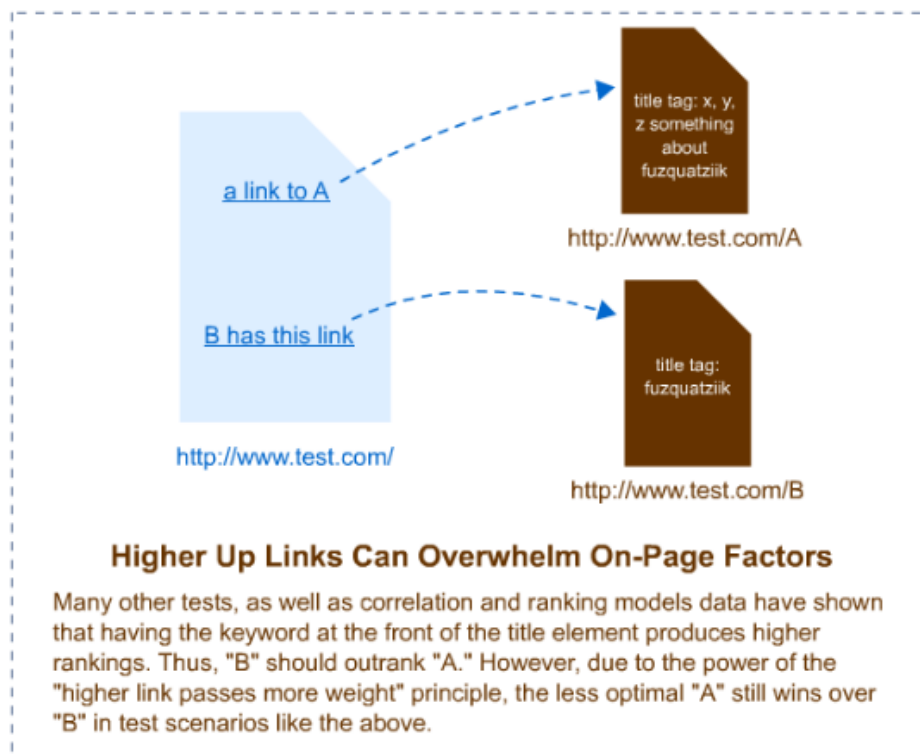
In questo post, ho intenzione di passare in rivista i dieci principi di valutazione dei link che si possono osservare, testare e che in qualche caso sono stati brevettati. Vorrei esprimere un particolare ringraziamento a Bill Slawski di [SEO By the Sea](#)<sup>3</sup>, i cui ultimi

<sup>1</sup> <http://infolab.stanford.edu/~backrub/google.html>

<sup>2</sup> [http://en.wikipedia.org/wiki/Hubs\\_and\\_authorities](http://en.wikipedia.org/wiki/Hubs_and_authorities)

<sup>3</sup> <http://www.seobythesea.com/>



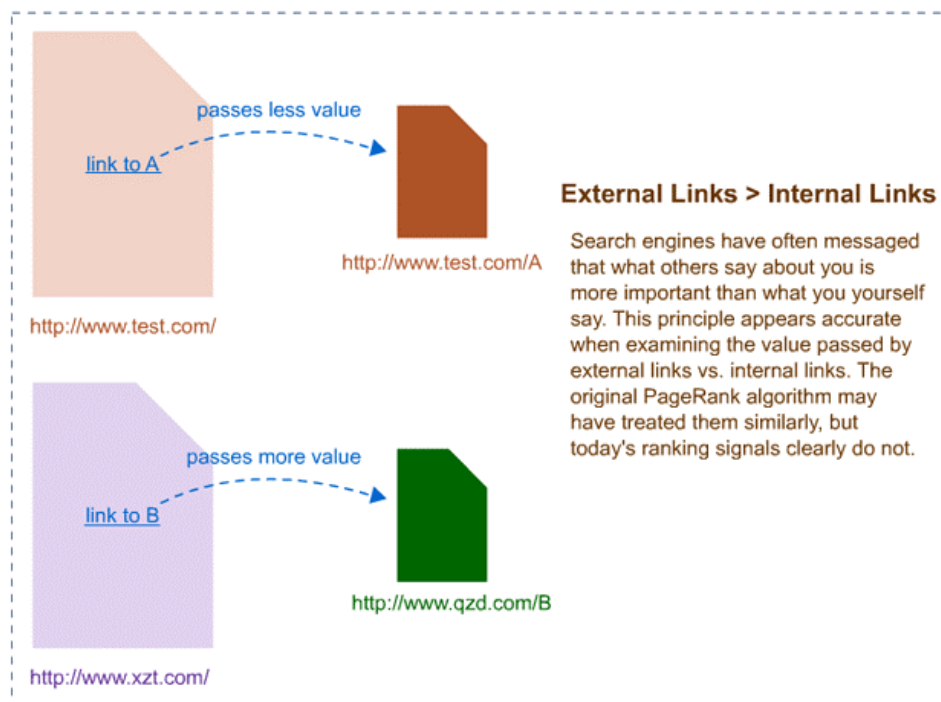


***Il link più in alto possono soverchiare i fattori "on-page"***

*I dati di molti altri test come la correlazione ed i modelli di ranking, hanno mostrato che avere la keyword all'inizio del titolo, produce un ranking più elevato. Quindi "B" dovrebbe superare "A". Comunque, in virtù della regola che "il link più in alto ha più peso", la pagina non ottimale "A", vince ancora su "B" in uno scenario di test come quello sopra.*

Coloro che hanno sfruttato ambienti di test spesso hanno anche combattuto contro il fenomeno "il link più alto vince" e ci può volere un lavoro sorprendente di ottimizzazione sulla pagina, per superare il potere che il link più in alto porta.

## 2 – I link esterni sono più influenti dei link interni



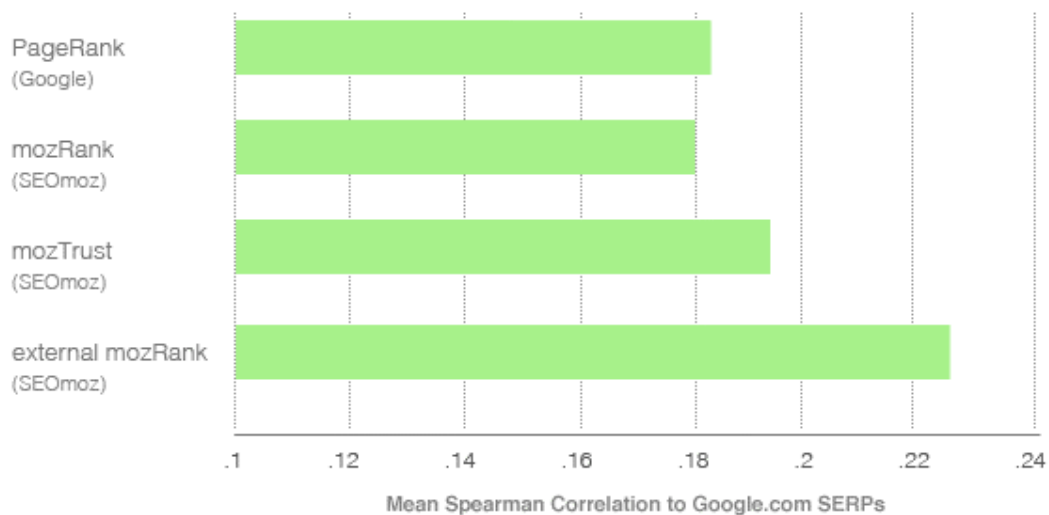
### **Link esterni > link interni**

*I motori di ricerca hanno spesso mandato messaggi sul fatto che quello che gli altri dicono di voi è più importante di quello che dite di voi stessi. Questo principio appare accurato quando si esamina il valore passato dai link esterni rispetto ai link interni. L'algoritmo PageRank originale può averli trattati in modo simile, ma i segnali sui posizionamenti di oggi dicono chiaramente di No.*

Qui Non c'è grande sorpresa, ma se vi ricordate, il concetto originale di PageRank non fa menzione di un diverso conteggio dei link esterni rispetto a quelli interni. E' molto probabile che metriche, più recenti (post-1997) premino i link esterni rispetto ai link interni. Si può vedere questo nel [dati di correlazione dal nostro post](#)<sup>7</sup> di poche settimane fa, in cui si vede che il mozRank esterno (il "PageRank" inviato da pagine di siti esterni) aveva una correlazione molto più alta con il ranking, rispetto al mozRank standard (PageRank):

<sup>7</sup> <http://www.seomoz.org/blog/the-science-of-ranking-correlations>

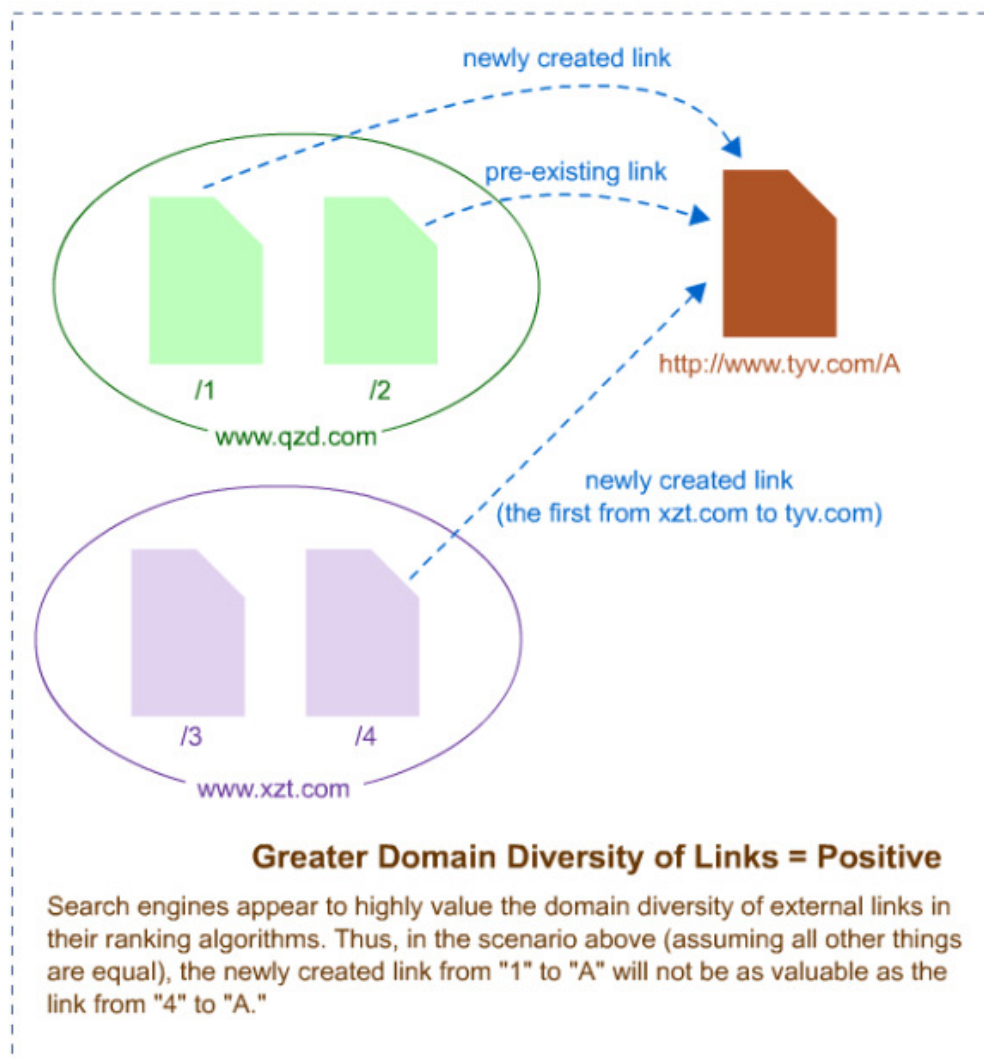
### ● Iterative PR Like Algorithms



#### ***Algoritmi iterativi del tipo PageRank***

Io non credo che sia troppo azzardato, immaginare che Google calcoli separatamente il PageRank esterno e quello interno e che possa usarlo in modi differenti per la valutazione del ranking di una pagina.

### 3 - Link da domini unici contano più dei link da siti che ce ne avevano già forniti in precedenza

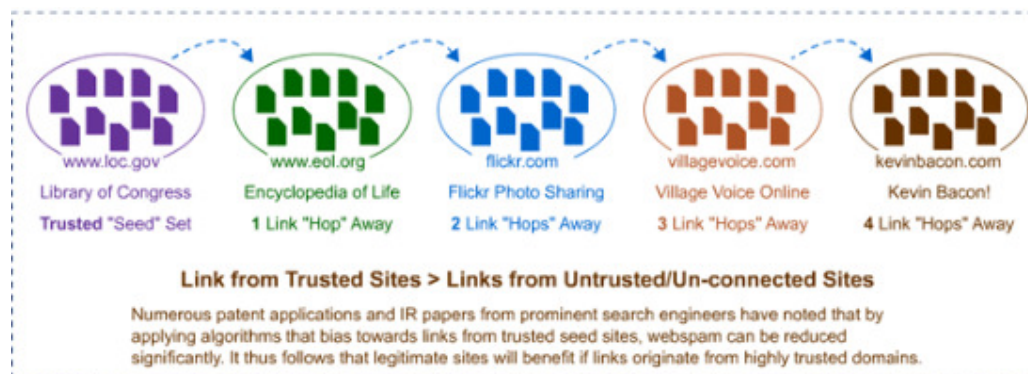


#### **Link con maggiore diversità di domini = Positivo**

*Sembra che i motori di ricerca diano un valore elevato alla diversità dei link esterni, nei loro algoritmi. Pertanto nello scenario sopra illustrato (a parità di tutte le altre condizioni), il nuovo link da "1" verso "A" non avrà lo stesso valore del link da "4" verso "A"*

Parlando dei dati di correlazione, nessuna singola semplice metrica è meglio correlata con il ranking nei risultati di Google, rispetto al numero di domini unici contenenti un link esterno a una determinata pagina. Questo suggerisce fortemente che la diversità sia un elemento in gioco nei sistemi di posizionamento e che è meglio avere 50 link da 50 domini diversi, che avere più 500 link da uno stesso sito. Curiosamente ancora una volta, l'algoritmo PageRank originale questo non lo prevede, che potrebbe essere uno dei motivi per cui i link provenienti a tappeto da un sito con molte pagine ad elevato PageRank funzionavano così bene in quei primi anni dopo il lancio di Google.

#### 4 - Link da siti più vicini al set del Trusted Seed (seme di partenza) passano più valore



#### **Link da siti affidabili > link da siti non affidabili o siti non connessi**

Numerose domande di brevetto e pubblicazioni di Information Retrieval di ingegneri di rilievo nel mondo della ricerca, hanno rilevato che applicando algoritmi che premiano link dai siti del seme di partenza, può ridurre significativamente lo spam. Ne consegue che i siti validi (non spam) avranno un beneficio se i link proverranno domini altamente affidabili.

Abbiamo [parlato in precedenza di TrustRank](#)<sup>8</sup> su SEOmoz ed abbiamo fatto riferimenti in generale agli articoli di Yahoo! sulla ricerca: [Combattere lo span del web con il TrustRank](#)<sup>9</sup>. Tuttavia, certamente anche Google ha fatto molto su questo fronte (come [Bill espone qui](#)<sup>10</sup>) e questa [richiesta di brevetto sulla selezione dei siti per il seme di partenza](#)<sup>11</sup> dice molto sulla necessità corrente ed il valore di questa metodologia. Il [Punteggio mozTrust](#)<sup>12</sup> di Linkscape funziona proprio in questo modo, utilizzando un algoritmo simile a PageRank, che è influenzato solo dai link provenienti da siti affidabili del seme di partenza, invece che determinato in modo uniforme da tutto il web.

<sup>8</sup> <http://www.seomoz.org/blog/whiteboard-friday-domain-trust-authority>

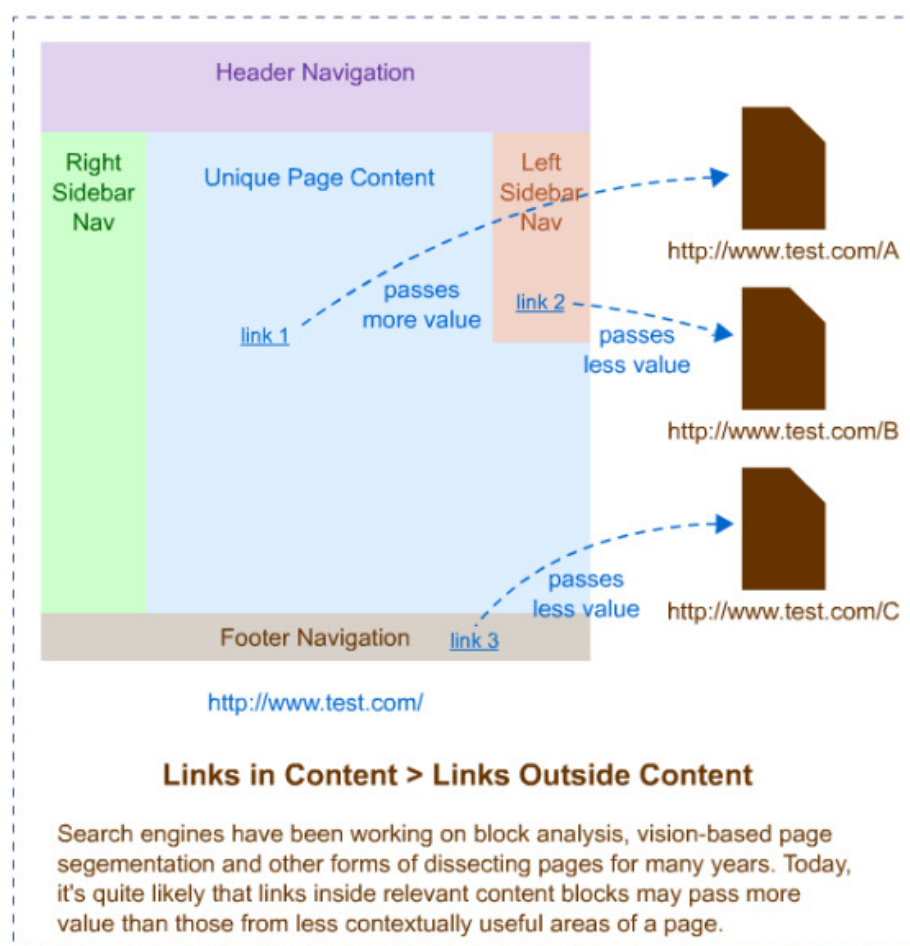
<sup>9</sup> <http://www.vldb.org/conf/2004/RS15P3.PDF>

<sup>10</sup> <http://www.seobythesea.com/?p=3015>

<sup>11</sup> <http://appft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&u=%2Fnetacgi/nph-adv.html&r=1&p=1&f=G&l=50&d=PG01&S1=20100114858.PGNR.&OS=dn/20100114858&RS=DN/20100114858>

<sup>12</sup> <http://www.seomoz.org/linkscape/help/metrics>

## 5 – I link dai contenuti unici localizzati nel corpo della pagina, trasmettono più valore di quelli nel piè di pagina, nella colonna laterale o nella navigazione.



### **Link nel contenuto > link Fuori contenuto**

*I motori di ricerca lavorano da molti anni sull'analisi dei blocchi, sulla segmentazione della pagina basata sulla visione ed altre forme di sezionamento delle pagine. Oggi è estremamente probabile che i link inclusi nei blocchi di contenuto rilevante possano passare più valore di quelli che si trovano in aree della pagina meno contestualmente utili.*

Ricerche come [VIPS di Microsoft](http://research.microsoft.com/apps/pubs/default.aspx?id=70027)<sup>13</sup> (Vision Based Page Segmentation), [Documenti ordinamento dei documenti basato sulla distanza semantica](http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&u=%2Fnetacgi/nph-adv.htm&r=1&p=1&f=G&i=50&d=PTXT&S1=7,716,216.PN.&OS=pn/7,716,216&mp;RS=PN/7,716,216)<sup>14</sup> di Google ed il recente [Google's Reasonable Surfer](http://www.seobythesea.com/?p=3806)<sup>15</sup>, suggeriscono tutte che valorizzare di più i link dai contenuti di quelli laterali o nel piè di pagina, può avere un impatto positivo netto nella lotta allo spamming ed alla manipolazione. Come webmaster e SEO, possiamo certamente confermare il fatto che ci sono un sacco di link a pagamento presenti in queste sezioni e che inserire link di questo genere all'interno del contenuto è molto più difficile.

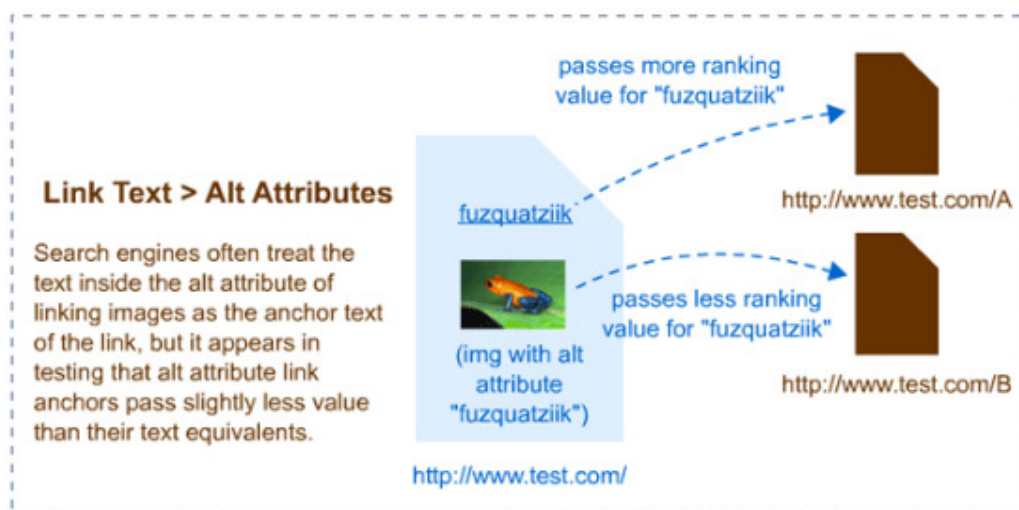
<sup>13</sup> <http://research.microsoft.com/apps/pubs/default.aspx?id=70027>

<sup>14</sup> <http://patft.uspto.gov/netacgi/nph->

[Parser?Sect1=PTO2&Sect2=HITOFF&u=%2Fnetacgi/nph-adv.htm&r=1&p=1&f=G&i=50&d=PTXT&S1=7,716,216.PN.&OS=pn/7,716,216&mp;RS=PN/7,716,216](http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&u=%2Fnetacgi/nph-adv.htm&r=1&p=1&f=G&i=50&d=PTXT&S1=7,716,216.PN.&OS=pn/7,716,216&mp;RS=PN/7,716,216)

<sup>15</sup> <http://www.seobythesea.com/?p=3806>

## 6 – Le parole chiave nel testo HTML passano più valore rispetto a quelle presenti nei tag ALT delle immagini collegate.

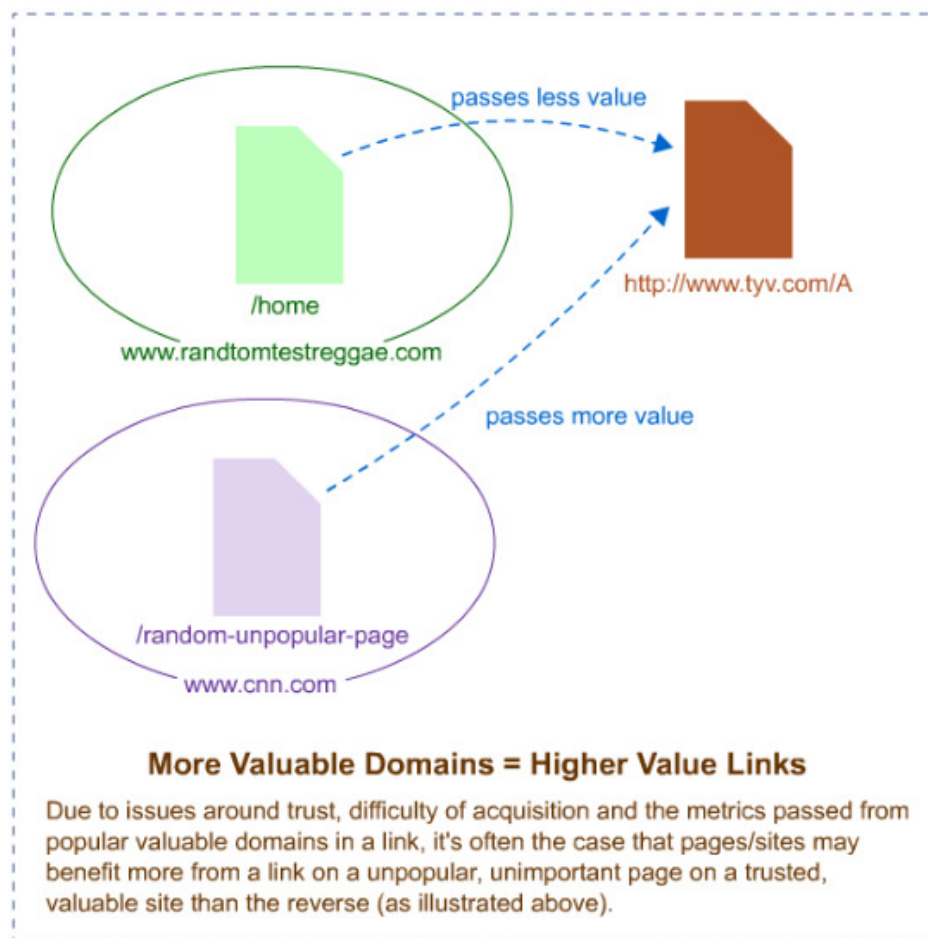


### **Link di testo > attributo ALT**

*I motori di ricerca spesso trattano il testo dentro l'attributo ALT delle immagini inserite in un link, come il suo testo, ma nei test appare che il testo degli attributi ALT passa un po' meno valore dei link puramente testuali.*

Questo non è trattato in alcuna ricerca o brevetto (di mia conoscenza), ma il nostro test ha dimostrato (e viene suffragato dai test altri), che il testo dei link effettuato tramite HTML è in qualche modo più potente o valutato di più, del testo presente nell'attributo alt del link di tipo immagine. Questo non vuol dire che dovremmo correre fuori a buttare in un fosso i link di tipo immagine, le icone o gli attributi alt che portano. E' solo una buona cosa essere consapevoli del fatto che Google sembra avere questa preferenza (forse sarà temporanea).

## 7 – I link da siti più importanti, popolari, attendibili passano più valore (anche se da pagine meno importanti)



### **Domini di maggior valore = link di maggior valore**

*Per delle considerazioni sull'affidabilità, la difficoltà di acquisire link da siti attendibili e le metriche trasmesse dai link di domini popolari e importanti, spesso succede che delle pagine o interi siti possano beneficiare di più da un link proveniente da una pagina sconosciuta e poco importante di un sito affidabile e rilevante, che l'inverso (come illustrato sopra).*

Probabilmente abbiamo tutti sperimentato la sensazione di sprofondare vedendo un concorrente con meno link e da pagine apparentemente meno forti, sorpassarci nelle posizioni. Ciò può essere in parte spiegato, dal valore che un dominio trasmette attraverso un link e che potrebbe essere non pienamente giustificato dalle metriche della pagina sorgente. Può anche aiutare i motori di ricerca a combattere lo spam e fornire risultati più attendibili in generale. Se i link provenienti da siti che raramente mettono link verso siti spazzatura passano significativamente più valore di quelli le cui politiche di link impattano sulla rete globale in modo discutibile, allora i motori di ricerca possono controllare molto meglio la qualità.

NOTA: Ho problemi a scavare nella rete per trovare rapporti di ricerca o brevetti su questo argomento, cercherò meglio domani e vedrò di rivisitare il post.

## 8 – I link contenuti all'interno dei tag NoScript trasmettono basso (e forse nessun) valore.



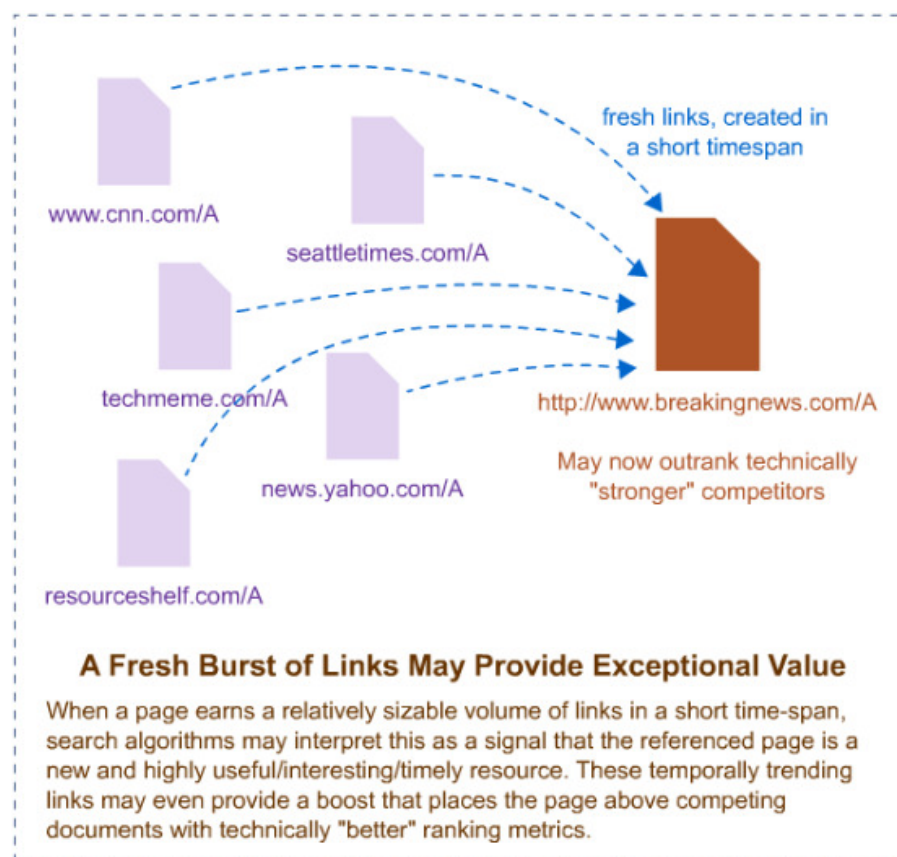
### ***I link nei tag noscript possono passare poco o nessun valore***

*I link nei tag NOSCRIPT non possono passare molto valore (se lo passano).*

*Siccome questi link non sono visibili dai visitatori (e poiché vengono da una storia di manipolazioni), i link dentro i tag noscript sembrano passare frequentemente un incremento minimale del ranking o nessuno del tutto.*

Nel corso degli anni, questo fenomeno è stato segnalato e contraddetto più volte. I nostri test hanno sicuramente suggerito che i link noscript non passano valore, ma che può non essere vero in ogni caso. Ecco perché abbiamo inserito la possibilità di filtrare i link noscript in Linkscape, ma la quantità di link globali sul web all'interno di questo tag è esigua.

## 9 - Una raffica di nuovi collegamenti può consentire ad un documento di superare i concorrenti più forti, temporaneamente (o per sempre)



### **Una raffica di nuovi link può fornire un valore eccezionale**

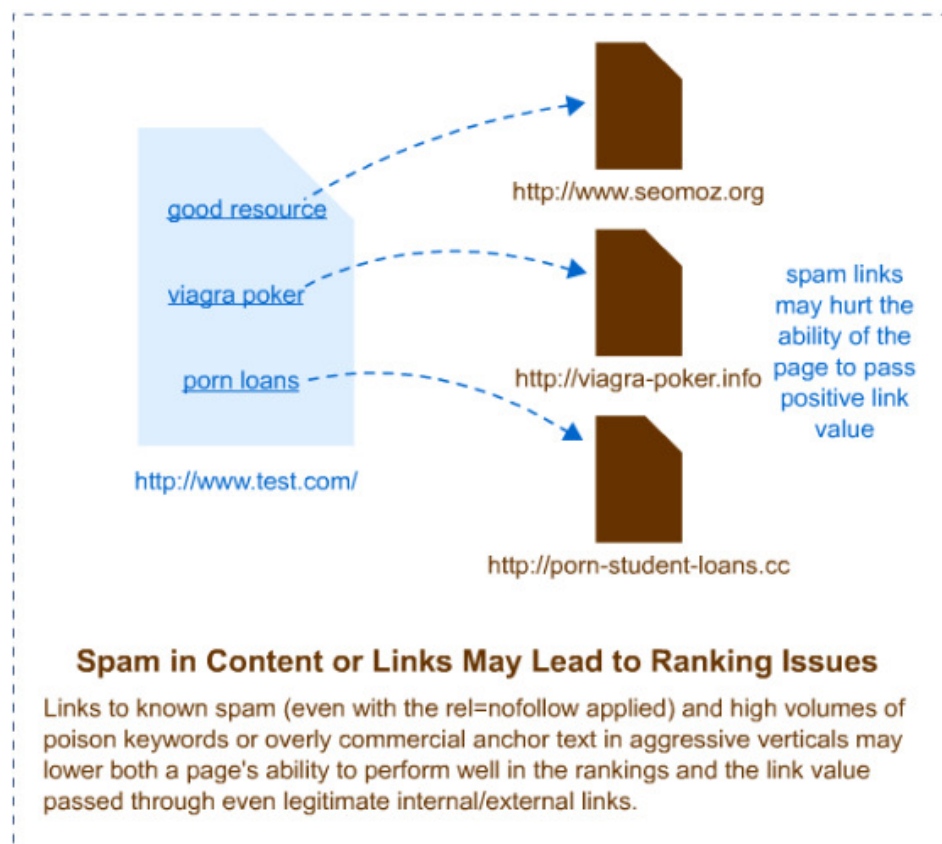
Quando una pagina si guadagna un volume abbastanza grande di link in un breve lasso di tempo, gli algoritmi dei motori di ricerca possono interpretarlo come un segnale che la pagina referenziata è nuova ed è una risorsa altamente utile, interessante e attuale. Questo particolare andamento temporale dei link può anche fornire una spinta che piazza la pagina sopra i documenti concorrenti che hanno metriche tecnicamente migliori.

A parte anche l'algoritmo [Google QDF<sup>16</sup>](#) (Query Deserves Freshness – la ricerca merita freschezza dei risultati), che può dare più valore ai contenuti creati e linkati di recente in alcune ricerche temporaneamente popolari, pare che il motore sfrutti anche i segnali temporali che riguardano i link, sia per valutare lo spam e la manipolazione, sia per ricompensare le pagine che guadagnano un gran numero di riferimenti in un breve periodo di tempo. Il brevetto di Google sui [Information Retrieval sulla base dei dati storici<sup>17</sup>](#) suggerì per primo l'utilizzo dei dati temporali, ma da allora il modello ha probabilmente superato diverse revisioni ed ha raggiunto la raffinatezza necessaria.

<sup>16</sup> <http://www.seomoz.org/blog/whiteboard-friday-query-deserves-freshness>

<sup>17</sup> <http://appft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PG01&p=1&u=%2Fnethtml%2FPTO%2Fsrchnum.html&mp;r=1&f=G&l=50&s1=%2220050071741%22.PGNR.&OS=DN/20050071741&RS=DN/20050071741>

## 10 – La pagine che puntano a spam possono svalutare gli altri link che ospitano



### **Lo spam nel contenuto o nei link può portare a problemi di ranking**

*Link a noti siti spam (anche applicando rel=nofollow), ed un alto livello di keyword "avvelenate" o un uso eccessivo di testo commerciale nei link di banner verticali aggressivi, può abbassare sia le prestazioni della pagina in termini di ranking, sia il valore che i suoi link legittimi passano alle pagine interne ed esterne.*

Sono rimasto affascinato a guardare gli esperimenti di Richard Baxter su questo tema nel suo post [Penalità sul livello della pagina di Google per lo spam nei commenti](#)<sup>18</sup>. Da allora, ho tenuto d'occhio alcuni post popolari e preziosi sui alcuni blog, che hanno ricevuto commenti spam altrettanto travolgenti e... evviva! Lo schema sembra verificabile. Gli webmasters farebbero bene a tenersi aggiornati sulla rimozione dei commenti spam per evitare di suscitare potenziali sanzioni sul ranking di Google (e la possibile perdita di valore dei propri link).

E che si può dire del classico "PageRank", il punteggio che di cui abbiamo un minuscolo accenno nei pixel verdi nella toolbar di Google? Potrei effettivamente supporre che mentre molte (e forse tutte) le caratteristiche sui link di cui sopra si stanno facendo strada nel processo di posizionamento, il PR è rimasto relativamente invariato rispetto al suo concetto originale. Il motivo del mio ragionamento? Il mozRank di SEOmoz, si correla molto bene con il PR della toolbar (vicino ad una media di 0,42 dove lo 0,25 è "perfetto" per le 2 cifre significative extra che mostriamo) ed è calcolato sulla base di un'intuizione molto simile a quella del PageRank dell'articolo originale. Se dovessi indovinare (e sto davvero tirando a indovinare), direi

<sup>18</sup> <http://seogadget.co.uk/google-page-penalty-for-comment-spam-rankings-and-traffic-drop/>

che Google ha mantenuto il classico PR perché trovano che sia una euristica semplice ed utile per alcuni compiti (inclusi probabilmente la priorità del crawling e dell'indicizzazione) ed hanno invece adottato molte più metriche da inserire nel calderone dell'algoritmo.

Come sempre, siamo in attesa di vostri commenti e spero che alcuni di voi raccolgano la sfida di provare tutto questo sui vostri siti o all'interno di ambienti di test e di riferire i risultati.

P.S.

Ho finito questo post quasi alle 03:00 (ed ho una riunione del consiglio di amministrazione domani), per cui mi scuso per eventuali refusi dispari o link interrotti. Speriamo che Jen prenderà la penna rossa e farà le correzioni in mattinata!.



**Rand Fishkin**

traduzione di **Marco Dini**

SEOMoz non è affiliato con questo sito e non ha visionato questa traduzione. SEOMoz fornisce i migliori [strumenti](http://www.seomoz.org/tools) [http://www.seomoz.org/tools] e [risorse](http://www.seomoz.org/blog) [http://www.seomoz.org/blog] SEO del Web.

**Questo articolo si trova all'indirizzo:**

<http://www.ideawebitalia.it/seo/1984/>

**Altri articoli di Rand Fishkin:**

<http://www.ideawebitalia.it/author/rand-fishkin/>

**Immagine in prima pagina:**

Sorgente: [geishaboy500](#)

Immagine originale: <http://www.flickr.com/photos/geishaboy500/2371323740/>

Licenza d'uso: <http://creativecommons.org/licenses/by/2.0/deed.it>