

L'algoritmo di Google rivelato

(www.impresapratica.com)

L'algoritmo di ranking di Google rivelato – pag. 2

A scanso di equivoci, vorrei fare una premessa. Prima di tutto nessun esperto SEO, nessun hacker, nessun “guru” conosce con certezza l'algoritmo di ranking di Google (a parte Google ovviamente). Così come nessuno conosce la ricetta della Coca Cola. Certo, molti si sono avvicinati (vedi Pepsi Cola), ma nessuno conosce la formula esatta. Secondo, l'algoritmo di Google non è qualcosa di statico e di immutabile, anzi, al contrario, cambia continuamente. Ad ogni modo, gran parte dei fattori che condizionano il posizionamento nelle SERP di Google, sono sempre validi. Chi non ha mai sentito parlare di SEO, troverà qui alcuni termini “per addetti ai lavori”, come cloaking, stuffing, PageRank, SERP, etc. Vista la lunghezza di questo post, ho omesso la spiegazione di tali termini. Alcune rapide ricerche con Google però, colmeranno la lacuna. Si parte.

Nel lontano 16 ottobre 2006, Rand Fishkin (aka Randfish), CEO di SEOMoz, definì quello che secondo lui era l'algoritmo di ranking (classificazione, posizionamento) di Google:

Punteggio di Google = (Punteggio relativo all'utilizzo delle parole chiave * 0,3) + (Importanza del dominio * 0,25) + (PageRank * 0,25) + (Link in ingresso * 0,25) + (Dati degli utenti * 0,1) + (Qualità del contenuto * 0,1) + (Spinte manuali) – (Penalizzazioni automatiche e manuali)

In questo algoritmo, i fattori nell'utilizzo delle parole chiave (Keywords), sono:

- Parole chiave nel tag del titolo
- Parole chiave nei tag degli header
- Parole chiave nel testo del documento
- Parole chiave nei link interni che puntano alla pagina
- Parole chiave nel nome di dominio e/o nell'indirizzo della pagina (URL)

Quelli relativi al dominio sono:

- Storia della registrazione
- Età del dominio
- Importanza dei link che puntano al dominio
- Rilevanza del dominio sull'argomento basata sui link in entrata ed in uscita
- Utilizzo storico & pattern dei link verso il dominio

Quelli relativi ai link in entrata sono:

- Età dei link
- Qualità dei domini che mandano il link
- Qualità delle pagine che mandano il link
- Testo dei link
- Valutazione della quantità/peso dei link (Pagerank o sue variazioni)
- Rilevanza sull'argomento delle pagine/siti che mandano i link

Quelli relativi agli utenti sono:

- Storia della percentuale dei click (CTR) sulla pagina nei risultati dei motori di ricerca (SERP)

L'algoritmo di ranking di Google rivelato – pag. 3

- Tempo speso dagli utenti sulla pagina
- Numero di ricerche per URL/nome del dominio
- Storia delle visite/utilizzo dell'URL/nome del dominio degli utenti che Google può monitorare (toolbar, wifi, analytics, ecc.)

Infine, quelli relativi al punteggio della qualità del contenuto, sono:

- Potenzialmente dato a mano per le ricerche e per le pagine più popolari
- Fornite da valutatori interni di Google
- Algoritmi automatizzati per valutare il testo (qualità, leggibilità, ecc.)

Recentemente però, Matt Cutts (il portavoce di Google), ha dichiarato che ci sono più di 200 fattori che condizionano il ranking su Google. Nel nuovo algoritmo di Google (Caffeine), uno di questi fattori sarà molto probabilmente la velocità di caricamento della pagina. Non solo: le SERP saranno personalizzate. Ciò significa che a parità di query, ciascuno di noi vedrà dei risultati differenti su Google. I risultati saranno personalizzati in base a:

- Storia delle ricerche effettuate
- Storia dei siti, pagine visitati
- Modifiche alle query (e analisi)
- Scopo delle ricerche (informazioni, navigazione)
- Dati espliciti (preferiti, reader, wiki)
- Interazione con la pubblicità
- Frequenza di navigazione nell'arco della giornata
- PageRank personalizzato (yahoo e google)
- Interazioni tra documenti e SERP

Ma quali sono questi fantomatici 200 fattori che condizionano il ranking di Google? Partendo dal [SEOMoz ranking factors](#) (un sondaggio che viene fatto ogni due anni da SEOMoz, interpellando i maggiori esperti SEO del mondo, 72 per l'esattezza) e unendo [altri elenchi di fattori di ranking](#), ho ottenuto un corposo elenco. Ho ordinato i fattori per importanza, dal più influente al meno influente (dove è possibile, viene riportata la percentuale di importanza di quel fattore). Per alcuni fattori non è disponibile alcun valore di importanza, perciò li ho elencati alla fine. Ciò non vuol dire che siano meno influenti (anzi, alcuni di loro sono a mio avviso molto importanti). Alcuni fattori sono determinanti in senso inverso (cioè non aumentano il posizionamento ma piuttosto la penalizzazione da parte di Google) e li ho evidenziati in grassetto. Sull'importanza di alcuni fattori non sono d'accordo, altri li ritengo davvero... come dire, bizzarri. Il linea di massima però, ci siamo. Alla fine ho tradotto il tutto e ho ottenuto un elenco di 206 fattori che sembrano influenzare il posizionamento di Google nelle SERP:

1. Anchor text con keyword nei link da siti esterni. Importanza: 73%
2. Popolarità e autorevolezza dei link esterni (quantità/qualità dei link esterni).
Importanza: 71%
3. **Dominio bannato dall'indice di Google per spam web. Importanza: 70%**
4. **Codice paese dell'estensione del dominio (es.: .co.uk, .de, .fr, .com.au, etc.).**

L'algoritmo di ranking di Google rivelato – pag. 4

- Importanza: 69%**
5. **Cloaking con intento malevolo/manipolatorio. Importanza: 68%**
 6. Differenziazione delle fonti di link (link da molti domini diversi). Importanza: 67%
 7. Presenza di parole chiave nel tag Title. Importanza: 66%
 8. Fiducia del dominio basata sulla distanza di link da domini fidati (es.: TrustRank, Domain mozTrust, etc.). Importanza: 66%
 9. **Posizionamento del dominio penalizzato su Google per spam web. Importanza: 65%**
 10. Originalità e unicità del contenuto. Importanza: 65%
 11. TrustRank specifico della pagina (se quella singola pagina ha guadagnato link da fonti fidate). Importanza: 65%
 12. Popolarità globale dei link del dominio, basata su algoritmi (es.: PageRank, MozRank, StaticRank, etc.). Importanza: 64%
 13. Diversità dei link (intesa come numero/varietà di domini unici che linkano a pagine del nostro dominio). Importanza: 64%
 14. Link da siti che sono considerati autorevoli in argomenti specifici e simili al nostro (come per l'[algoritmo Hilltop](#)). Importanza: 64%
 15. **I link sono pagati piuttosto che ottenuti spontaneamente. Importanza: 63%**
 16. Posizione della parola chiave nel titolo: keyword posta come prima parola del titolo. Importanza: 63%
 17. **Lingua del contenuto usata sul sito. Importanza: 63%**
 18. PageRank. Importanza: 63%
 19. Parole chiave nel nome del dominio. Importanza: 60%
 20. Link da altri domini dello stesso paese. Importanza: 60%
 21. Rilevanza dell'argomento trattato dalle fonti di link esterni (i link esterni vengono da siti/pagine che hanno rilevanza con l'argomento della mia pagina). Importanza: 58%
 22. Localizzazione geografica dell'indirizzo IP del server che ospita il dominio. Importanza: 57%
 23. **Acquistare link. Link acquisiti da venditori di link conosciuti. Importanza: 56% ([brevetto Google](#))**
 24. Anchor text con keyword nei link interni. Importanza: 55%
 25. **Revisione manuale da parte di personale di Google e/o revisori di qualità. Importanza: 53%**
 26. **Il dominio contiene link ad un quantitativo significativo di siti web che fanno spam. Importanza: 52%**
 27. Crescita nel tempo dei link al dominio (la quantità/qualità dei link guadagnati nel tempo e la loro distribuzione temporale). Importanza: 52% – (brevetto Google) se i link si ottengono troppo in fretta è segno di una crescita non naturale ma fatta ad arte per migliorare il posizionamento. Penalità e sandbox se è un nuovo sito?
 28. Architettura del sito (struttura gerarchica e chiara). Importanza: 52%
 29. Destinazione geografica impostata all'interno degli strumenti per webmaster di Google (paese dei visitatori del sito). Importanza: 52%
 30. **Link DALLA PAGINA verso siti/pagine che fanno spam. Importanza: 51%**
 31. **Cloaking per User Agent. Importanza: 51%**
 32. Posizione della pagina in relazione alla struttura del sito. Regola dei 3 click massimi

L'algoritmo di ranking di Google rivelato – pag. 5

- per raggiungere la pagina. Importanza: 51%
33. Popolarità dei link interni (contando solo i link provenienti da pagine interne verso la pagina principale o homepage). Importanza: 51%
34. **Frequenti interruzioni a livello server e inaccessibilità del sito. Importanza: 51%**
35. Freschezza del contenuto (data di creazione della pagina). Importanza: 50%
36. Keyword nel tag H1. Importanza: 49%
37. **Nascondere il testo con sfondo dello stesso colore. Importanza: 49%**
38. Link dal nostro dominio a siti/pagine che fanno spam. Importanza: 48%
39. Link da domini con estensioni particolari (es.: .edu, .gov, .mil, .ac.uk, etc.). Importanza: 47%
40. Uso nella pagina di keyword negli anchor text dei link interni. Importanza: 47%
41. Uso nella pagina di keyword negli anchor text dei link esterni. Importanza: 46%
42. **Cloaking indirizzo IP. Importanza: 46%**
43. **Eccessive ripetizioni dello stesso anchor text in una percentuale/quantità troppo alta di link esterni verso il nostro sito/pagina. Importanza: 46%**
44. Keyword nella PRIMA parola del tag H1. Importanza: 45%
45. Keyword usata nelle prime 50-100 parole della pagina. Importanza: 45%
46. **Registrazione del sito con Google Local. Importanza: 45%**
47. **Nascondere del testo via CSS, impostando l'offset in un'area della pagina non visibile. Importanza: 44%**
48. **Numero eccessivo di parametri negli URL dinamici. Importanza: 43%**
49. Keyword nel sotto dominio (es.: keyword.seomoz.org). Importanza: 42%
50. Storia del tasso di Click-Through nelle SERP all'URL esatto della pagina. Importanza: 42% – (brevetto Google). Quanti sono i click alla pagina?
51. **Il dominio non ha ancora guadagnato link fidati. Importanza: 41%**
52. **Acquisizione di link grazie a campagne di scambio link e "bait link" (trucchi per guadagnare link). Importanza: 41%**
53. Cross linking interno. Importanza: (41%)
54. **Keyword stuffing nel testo della pagina (troppe keyword = spamdexing). Importanza: 41%**
55. **Troppi link da siti ospitati all'interno dello stesso blocco di indirizzi IP. Importanza: 41%**
56. Address in On-Page Text Content. Importanza: 41% (purtroppo mi sfugge il senso di questo fattore)
57. **Nascondere del testo via CSS, comando display:none; styling. Importanza: 40%**
58. Storia dei cambiamenti di contenuto della pagina (quanto spesso viene modificato il contenuto della pagina). Importanza: 39% – (brevetto Google) – Troppo non va bene
59. **Keyword stuffing nel tag <title>. Importanza: 39%**
60. Storia del tasso di Click-Through nelle SERP all'URL esatto del dominio. Importanza: 39%
61. Keyword usata nell'URL come nome della pagina (es.: seomoz.org/folder/keyword.html). Importanza: 38% Google non si "affida" più a

L'algoritmo di ranking di Google rivelato – pag. 6

- questo tag ma lo usa spesso
62. Keyword usata nell'URL come cartella che contiene la pagina (es.: seomoz.org/keyword/page.html). **Importanza: 37%**
 63. **Keyword stuffing nell'URL. Importanza: 37%**
 64. **I link sono ottenuti grazie a campagne manipolate con widget/badge. Importanza: 37%**
 65. **Cloaking via JavaScript/Rich Media Support Detection. Importanza: 37%**
 66. Link a siti/pagine fidati e autorevoli. **Importanza: 37%**
 67. Durata della registrazione del dominio. 1 anno? 5 anni?. **Importanza: 37%** – (brevetto Google) 5 anni è meglio di 1
 68. Query di ricerca relative al dominio o alla marca associata. **Importanza: 36%**
 69. Storia della registrazione del dominio (da quanto tempo con stesso proprietario, numero di rinnovi, etc.). **Importanza: 36%**
 70. **Cloaking via cookie detection. Importanza: 36%**
 71. **Link ALLA pagina da siti/pagine che fanno spam. Importanza: 36%**
 72. **I link sono ottenuti da directory che vendono link di bassa qualità. Importanza: 36%**
 73. **Numero eccessivo di link da siti posseduti dallo stesso registrant. Importanza: 36%**
 74. Keyword usata in altri tag intestazione diversi da H1 (da <h2> a <h6>). **Importanza: 35%**
 75. Indirizzo associato nel dominio (inteso come indirizzo fisico del proprietario del dominio). **Importanza: 35%**
 76. **Link al dominio da siti/pagine che fanno spam. Importanza: 34%**
 77. **Cloaking con l'intento di migliorare l'esperienza di navigazione degli utenti. Importanza: 33%**
 78. Parole chiave nel tag Alt delle immagini -. **Importanza: 33%**
 79. **Link acquisiti da campagne virali manipolate (cfr. alberi di Natale di Blogbabel). Importanza: 33%**
 80. Keyword usata nel nome dell'immagine contenuta nella pagina (es.: keyword.jpg). **Importanza: 33%**
 81. Keyword usate nel testo della pagina. **Importanza: 33%**
 82. URL dinamici e URL statici. **Importanza: 33%**
 83. Uso di "raffinamenti" nelle query di ricerca. **Importanza: 32%**
 84. Server/Hosting Uptime (se Google non trova il tuo sito, alla lunga viene cancellato dall'indice). **Importanza: 32%**
 85. **Sovra-ottimizzazione degli anchor text nei link interni. Importanza: 32%**
 86. **Uso di "Poison" Keywords negli anchor tex dei link esterni (quali sono quelle italiane? Forse: link, reciproci, clicca, etc.). Importanza: 32%**
 87. **Link acquisiti comprando vecchi domini ed effettuando il redirect. Importanza: 32%**
 88. Hosting. Quali altri domini sono ospitati sullo stesso host. **Importanza: 31%**
 89. Modifiche sulla proprietà del dominio. **Importanza: 31%**
 90. Inclusioni di feed su Google news. **Importanza: 31%**
 91. **Lunghezza dell'URL da mantenere al minimo. Importanza: 30%**

L'algoritmo di ranking di Google rivelato – pag. 7

92. Localizzazione geografica dei visitatori del sito. Importanza: 30%
93. Uso del sitemap XML. Importanza: 29%
94. **Uso di anchor text nei link interni nel footer. Importanza: 27%**
95. Keyword nei tag o . Importanza: 26%
96. **Keyword stuffing nel tag Meta Description. Importanza: 26%**
97. Tempo medio delle visite sulla pagina. Importanza: 26%
98. Uso di HTML razionale (influenza indiretta nel senso di aumento del click-through). Importanza: 25%
99. Storia del proprietario del dominio. Importanza: 25%
100. Quantità e qualità dei link No follow alla pagina. Importanza: 25%
101. Keyword density della pagina tra il 5-20%. (formula: numero di parole chiave usate ÷ numero totale delle parole sulla pagina). Importanza: 25% (alcuni riportano l'influenza dell'argomento: la percentuale di keyword stuffing varia a seconda dell'argomento trattato)
102. **PageRank sculpting: abuso del No follow nei link interni. Importanza: 24%**
103. Registrazione del dominio con Google Local. Importanza: 24%
104. **Link acquisiti comprando vecchi domini ed aggiungendo link. Importanza: 24%**
105. Citazioni del dominio nel testo, anche in assenza di link diretti. Importanza: 24%
106. Inclusione dei feed del dominio in Google Blog Search. Importanza: 24%
107. Citazioni/riferimenti al dominio nella directory di Yahoo!. Importanza: 24%
108. Geo-Tagging of Pages via Meta Data (e.g. Dublin Core Meta Data Initiative). Importanza: 24% (non capisco il significato di questo fattore)
109. Citazioni/riferimenti al dominio su DMOZ.org. Importanza: 23%
110. Keyword usata nei tag degli elenchi puntati () sulla pagina. Importanza: 23%
111. Esistenza del tag Meta Description. Importanza: 22%
112. Keyword usata nei parametri dell'URL (es.: seomoz.org/page.html?keyword). Importanza: 22%
113. **Link ottenuti grazie ai forum (link nella firma, lasciare link nei messaggi, etc.). Importanza: 22%**
114. Citazioni/riferimenti al dominio in Wikipedia. Importanza: 22%
115. Uso dei feed sul dominio. Importanza: 21%
116. Citazioni/riferimenti al dominio nel Librarian's Internet Index – Lii.org. Importanza: 21%
117. Keyword usata nei tag <i> o . Importanza: 21%
118. Percentuale di link Followed in rapporto a quelli Nofollowed che puntano al dominio. Importanza: 21%
119. Presenza del dominio/pagina su Delicious. Importanza: 21%
120. **Tag Title eccessivamente lungo. Importanza: 21% (il Title tag deve contenere tra i 10 e i 60 caratteri, senza includere caratteri speciali)**
121. Presenza del dominio/pagina su SearchWiki di Google (voti, quotazioni, commenti). Importanza: 19%
122. Presenza del dominio/pagina su StumbleUpon. Importanza: 19%
123. Parole chiave nel tag Meta Description (Non Meta Keywords ed entro i 200 caratteri). Importanza: 19%

L'algoritmo di ranking di Google rivelato – pag. 8

- 124.Registrazione del dominio con gli strumenti per webmaster di Google. Importanza: 18%
- 125.Attivazione del Google Image Labeler. Importanza: 17%
- 126.Percentuale di link Follow in relazione a quelli No follow che puntano alla pagina. Importanza: 17%
- 127.Presenza del dominio/pagina su Twitter. Importanza: 17%
- 128.Conformità dell'HTML agli standard W3C. Importanza: 16%
- 129.**Keyword stuffing nel tag Meta Keywords. Importanza: 15%**
- 130.LinkedIn Data About the Domain or Page. Importanza: 15%
- 131.Uso di certificati sul dominio nelle transazioni HTTPS. Importanza: 14%
- 132.Validità degli indirizzi e-mail, numeri di telefono, etc. sulle informazioni del dominio (whois). Importanza: 13%
- 133.Citazioni/riferimenti al dominio su Google Knol. Importanza: 13%
- 134.Uso di elementi Flash (o altri plugin). Importanza: 13%
- 135.Keyword usata nell'estensione del file della pagina (es.: seomoz.org/page.keyword). Importanza: 12%
- 136.Presenza del dominio/pagina su Facebook. Importanza: 12%
- 137.Presenza del dominio/pagina su MySpace. Importanza: 11%
- 138.Uso di pubblicità nella pagina. Importanza: 11%
- 139.Riferimenti/link al dominio su Gmail. Importanza: 9%
- 140.Uso di Google AdSense nella pagina. Importanza: 8%
- 141.Keyword usata nei tag di commento. Importanza: 6%
- 142.Uso di Google AdSense nel dominio. Importanza: 5%
- 143.Uso di Google AdWords per annunci che puntano al dominio. Importanza: 5%
- 144.Alexa Rank del dominio. Importanza: 5%
- 145.Compete.com Rank del dominio. Importanza: 5%
- 146.Keyword usata nel tag Meta Keywords. Importanza: 5%
- 147.Uso delle applicazioni Google sul dominio. Importanza: 3%
- 148.Uso di tag intestazione nell'URL
- 149.Sinonimi relativi al tema del sito/pagina
- 150.Posizione del link (footer, barra di navigazione, corpo del testo)
- 151.Il sito è in un dominio o in un sotto dominio
- 152.File JS
- 153.Velocità di JS
- 154.Uso di CSS esterni
- 155.CSS codificato all'interno della pagina HTML o incluso o su un file esterno
- 156.Navigabilità e architettura o avere 800 link nel template della pagina
- 157.% della pagina modificata
- 158.Sinonimi relativi al tema della pagina/sito
- 159.Prossimità della grafica a supporto del testo
- 160.Risoluzione/qualità delle immagini
- 161.Tag e commenti nei video
- 162.Riferimenti nei social network
- 163.Numero di link interni alla pagina
- 164.Anchor text del PRIMO link

L'algoritmo di ranking di Google rivelato – pag. 9

165. Anzianità del sito. Più vecchio è meglio è. L'opposto della sandbox – (brevetto Google).
166. Rilevanza del contenuto della pagina
167. Attributo Title del link
168. Tag Alt di immagini con link
169. Autorevolezza del dominio (.edu, .gov)
170. Effetto [Sandbox](#) (probabilmente la stessa importanza che ha l'anzianità del dominio)
171. Personalizzazione della ricerca disabilitata (Caffeine)
172. Numeri di impression che riceve il dominio nelle SERP di Google. Especially when that number spikes, it's been mentioned in patents as a possible spam signal.
173. Link JavaScript
174. Interazione con l'utente/soddisfazione con i risultati ottenuti
175. Qualità delle pubblicità contenute nel sito
176. Velocità di caricamento della pagina (Caffeine)
177. Vicinanza delle keyword (per 2 o più keyword)
178. Link non funzionanti (controllali con [Dead-Links](#))
179. Stabilità dei link nel tempo
180. Numero di link nella pagina inferiore a 100
181. Estensione del dominio (.gov, .org, .com)
182. Dimensioni della pagina sotto i 100 Kb (meglio sotto i 40 Kb)
183. Trattini nell'URL (1 o 2 ok, più di 4 inizia a sembrare spam, 10 è considerato sicuramente spam)
184. Anzianità della pagina in relazione all'anzianità del sito e alle altre pagine
185. Pagina senza testo e contenente solo elementi grafici
186. Link a siti di scarsa qualità (es.: siti scambio link)
187. Redirect via metatag refresh
188. Utilizzare immagini/testo non propri (violazioni del copyright)
189. Link di un pixel
190. Pagine gateway e doorway
191. Uso di frasi associate a tecniche di spam conosciute (brevetto Google)
192. Anzianità del link. Vecchio = buono – (brevetto di Google)
193. Frequenza delle modifiche su anchor text. Eccessive modifiche = non buono – (brevetto di Google)
194. Link da siti "esperti" – (brevetto di Google) Algoritmo Hilltop
195. Tematizzazione del sito
196. Tempo trascorso sulla pagina
197. Bookmark – la pagina viene inserita nei preferiti? Buono
198. Traffico del sito – (brevetto di Google) – traffico in aumento = buono
199. Ordine delle keyword (due o più keyword)
200. Dimensione del sito (Google ama i siti grandi)
201. Pagine con frame
202. Posizione del link nella pagina referrer
203. Link in mappe cliccabili
204. La categoria di Dmoz coerente con quella del sito

L'algoritmo di ranking di Google rivelato – pag. 10

205.Sito elencato in altre directory come About, Looksmart, etc.

206.Modalità di abbandono della pagina (bottone back, click su link, etc.)

Significato delle percentuali:

- 65-100% = importanza molto alta
- 55-64% = importanza alta
- 45-54% = importanza moderata
- 35-44% = importanza bassa
- 25-34% = importanza minima
- 0-24% = importanza davvero minima

Fine della guida

Sir Vester

([Impresa Pratica](#))

(Ehi, psst... vuoi sapere come puoi guadagnare lavorando su Internet? Allora prova a dare un'occhiata qui: [Impresa Pratica](#))