

L'evoluzione della ricerca

Danny Sullivan



Articolo originale:

The Evolution of Search

<http://moz.com/blog/the-evolution-of-search>

28 Giugno 2013

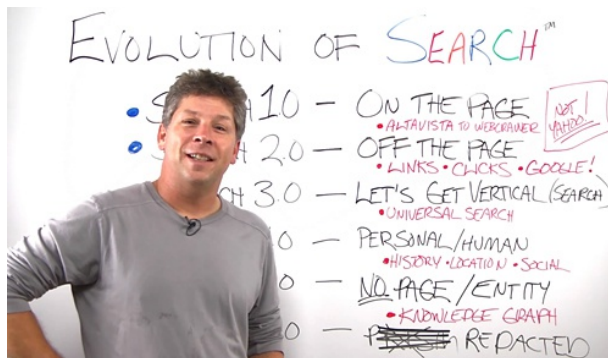


Salve Moz fan, benvenuti a Whiteboard Friday. Io non sono Rand, sono Danny Sullivan, il fondatore di SearchEngineLand.com e MarketingLand.com. Siccome qui a Seattle ci sono 8.000 gradi, Rand ha deciso di non stare nei paraggi ed io sono qui sudato come un maiale, per aver camminato fin qui.

Sono molto emozionato all'idea di fare un Whiteboard Friday. Questo è il mio primo pezzo da solista assoluto e mi hanno detto che devo farlo in 11 minuti, e in 1.5 riprese. No, in una sola ripresa. L'argomento di oggi sarà l'evoluzione della ricerca, il marchio registrato da Google. Non è vero, non sono i proprietari della ricerca.

C'è stato un tempo in cui non padroneggiavano la ricerca e questo ci porta alla "Ricerca 1.0".

Ricerca 1.0



Lo sapevate, ragazzi, che i motori di ricerca erano molteplici, che non esistevano termini come Googling? Usavamo cose come Alta Vista, Lycos, e WebCrawler. Vi ricordate quei nomi? Ci sono state cose come OpenText e quell'altro, Magellan. Bene, questi erano i motori di ricerca che esistevano prima di Google; andavano in giro sul web, facevano il crawling di tutte le

pagine (circa qualche dozzina di pagine che esistevano al tempo) e dopo potevamo farci le nostre ricerche e trovare il modo di farle salire di posizione.

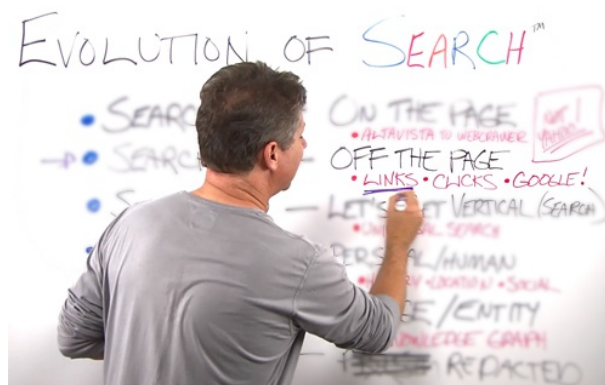
Tutto questo era determinato solo dalle parole che stavano nella pagina. Quindi, se si voleva posizionarsi bene per, non so, qualcosa come "film", si sarebbero dovuti mettere i termini "film" sulla pagina, 100 volte di fila. Poi, se qualcun altro avesse voluto posizionarsi prima di noi, doveva mettercelo 150 volte di seguito, perché un motore di ricerca, ragionava così: "noi pensiamo che la pertinenza stia tutta nel numero di ripetizioni di una parola sulla pagina ed un pochino anche nella posizione di quelle parole". Le parole nella parte superiore della pagina potrebbero contare un po'

di più, che se stessero più in basso.

Il risultato era che questo criterio rendeva facile lo spamming. I motori di ricerca non volevano veramente che vi posizionaste meglio per la keyword "film", perché avevate ripetuto la parola "film" 150 volte, rispetto a qualcun altro che l'aveva ripetuta 100 volte. Avevano bisogno di tirar fuori un segnale migliore che indicasse la rilevanza. Si presero tutto il tempo per cercare quel segnale.

Per farla breve, non stavano ricavando un granché di soldi dalla ricerca, per cui non prestavano troppa attenzione al problema. Ma i fondatori di Google si misero a pensare: "guarda un po', se creiamo un motore di ricerca, un giorno qualcuno potrebbe farci un film con Owen Wilson e Vince Vaughn, quindi diamoci da fare e troviamo un sistema migliore" e questo ci porta alla "Ricerca 2.0".

Ricerca 2.0



La ricerca 2.0 ha iniziato a guardare alle cose che indichiamo come fattori di ranking "off-the-page", perché tutti gli elementi "on-the-page" erano sotto il completo controllo di chi pubblicava la pagina. Il proprietario della pagina poteva cambiare tutto ciò che voleva. C'è stato anche un tempo, quando si usava Infoseek, in cui si poteva sottomettergli una pagina web, che veniva immediatamente aggiunta

all'indice e si poteva vedere subito come si posizionava. Se non vi piaceva, potevate immediatamente effettuare una modifica e rimetterla su di nuovo. Si poteva salire di posizione in questo modo. Quindi, il criterio "off-the-page" è una cosa del genere: "andiamo in giro e raccogliamo delle raccomandazioni da qualcuno che non sia il proprietario e stabiliamo cosa gli altri pensano di queste pagine web, perché forse questo criterio è meno facile da aggirare e potrebbe darci risultati di ricerca di migliore qualità".

A proposito, qui non ho menzionato Yahoo, perché sto parlando di motori di ricerca crawler-based, quelli che utilizzano l'automazione per trovare e raccogliere in giro le pagine web. Yahoo per la maggior parte del tempo (o almeno a me pare così) era una directory, vale a dire, un motore di ricerca basato sul lavoro di esseri umani, dove le cose sono elencate, solo perché un essere umano è andato realmente su un sito web, ha scritto una recensione e lo ha aggiunto.

Ma torniamo alla ricerca 2.0. Google andò avanti e cominciò a fare molto uso di qualcosa che si chiama analisi dei link. Anche gli altri motori di ricerca ci provarono, ma non riuscivano a trovare la formula giusta e non facevano dipendere molto i loro risultati da questo.

Ma Google, il nuovo arrivato nella compagnia, disse, "lo useremo moltissimo. Considereremo i link come voti e le persone con un sacco di link che li puntano, forse otterranno un sacco di voti, e daremo loro un punteggio più alto per posizionarli meglio". Non si trattava solo di un puro criterio quantitativo, però. Google poi ha anche voluto sapere chi ha i voti migliori, chi è la reale autorità. Così hanno anche osservato la qualità di quei link.

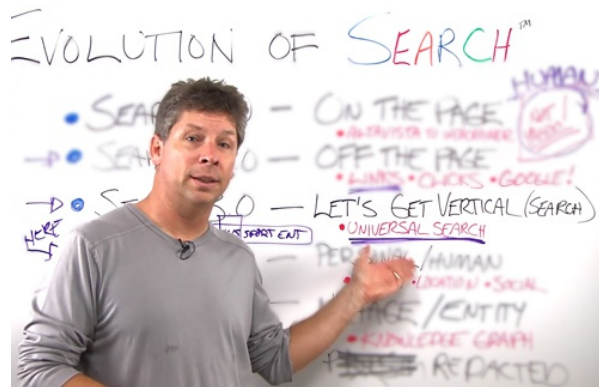
C'erano altri che stavano facendo un po' di cose off-the-page. Uno di loro, se vi

ricordate, era Direct Hit. Loro in realtà guardavano cose come i clic. Li osservavano e dicevano: "abbiamo esaminato 10 risultati della ricerca e possiamo vedere che le persone fanno clic sul terzo risultato della ricerca, in modo del tutto sproporzionato rispetto a come ci aspetteremmo. Invece di raccogliere, diciamo, il 20% dei clic, sta attirando l'80% dei clic". Questo suggeriva loro di spostarlo in alto fino al numero uno e quindi, di spostare gli altri risultati un po' più in basso.

Queste sono alcune delle cose che si cominciavano a fare, ma sono stati i link che ci hanno portato avanti per circa un decennio. Ora i link, i segnali off-the-page, sono stati alla base dei risultati della ricerca e ancora oggi determinano i risultati della ricerca ed il modo in cui hanno cominciato a migliorare; ma ora abbiamo un momento di una pausa, che possiamo chiamare, o meglio, che io chiamo: Ricerca 3.0.

A proposito, ho buttato giù tutto questo schema, soprattutto perché possiate anche essere in disaccordo con questa visione, o magari perché vi accorgiate di essere d'accordo su qualcosa; comunque qualche anno fa, stavo cercando di spiegare come vedevo l'evoluzione della ricerca e alcuni di questi cambiamenti che stavano venendo fuori.

Ricerca 3.0



Quello che è successo in questa era della Ricerca 3.0 è che, anche se stavamo usando i link e stavamo ottenendo risultati di qualità migliore, c'era anche una moltitudine di informazioni che stava entrando, e quei soli segnali non erano sufficienti. Avevamo bisogno di un altro modo per ricavare più pertinenza e il modo in cui i motori di ricerca hanno cominciato a realizzarlo era questo: "invece di cercare su 100 miliardi di pagine, cerchiamo

su una piccola collezione di pagine di solo contenuto ben focalizzato". Questo si chiama ricerca verticale.

Nella ricerca orizzontale, avremmo fatto una ricerca su cose come notizie, sport, intrattenimento, shopping e poi avremmo buttato tutto in una grande casella di ricerca. Vai in giro e cerca di tornare con tutte le pagine, da tutto il Web, che pensi siano rilevanti per ciò che cerco. La ricerca verticale, è come una fetta verticale del web; quella fetta verticale del web è solo e soltanto, per esempio, il contenuto delle notizie. Così quando si fa una ricerca per qualcosa come NSA, si va solo a guardare il contenuto delle notizie, in modo da trovare le risposte sulla notizia che sono relative alla NSA in questo momento. Non si va a vedere se magari c'è qualche informazione sportiva o informazioni di shopping che possono corrispondere ad NSA.

Tra l'altro, questo è particolarmente importante in questi giorni. Si sente parlare che sta accendendo una cosa come PRISM. Si tratta di un programma di spionaggio o di un programma di intercettazioni o di un programma di data mining, a seconda di chi è il vostro interlocutore, che il governo degli Stati Uniti sta usando. Prism è anche un oggetto che si usa per filtrare la luce (prisma) e quindi se stiamo facendo una ricerca per ottenere informazioni sul filtraggio della luce, probabilmente non ci si rivolgerebbe ad un motore di ricerca per le notizie, perché in questo momento le notizie sono piene di contenuti riguardo al sistema PRISM. Invece, se vogliamo le ultime informazioni su ciò che sta accadendo proprio in tutta questa storia del PRISM, allora rivolgersi al

motore di ricerca per le notizie è importante, perché non verranno fuori le altre cose a cui non è necessariamente PRISM è correlato.

Quindi abbiamo questa cosa che si chiama Ricerca 3.0, o ricerca verticale, o come in particolare la chiama Google, universal search. Cerca di risolvere il problema che sorge quando qualcuno digita in un campo di ricerca "immagini di fiori"; dovrebbe mostrare immagini di fiori, piuttosto che 10 link che forse portano a immagini di fiori. Oggi questa funzionalità è abbastanza robusta. Anche Bing fa una cosa di questo genere, hanno il loro modo di miscelare i risultati.

Ricerca 4.0

Poi c'è la Ricerca 4.0. E' una specie di ritorno a ciò che faceva Yahoo, ossia all'uomo. A proposito, io non scrivo più molto bene con la penna, perché digito sempre su una tastiera.

Riguardo all'impiego di esseri umani, uno delle cose più straordinarie che sono successe con l'avvento dei motori di ricerca, è che, in un periodo molto breve, hanno completamente cambiato il modo in cui cerchiamo le informazioni. Per migliaia di anni, se si aveva bisogno di sapere qualcosa, si doveva parlare con un essere umano. Anche quando abbiamo avuto le biblioteche e la gente disponeva di tutte quelle informazioni, tipicamente andando in una libreria ci si doveva rivolgere ad un bibliotecario e chiedere: "sto cercando alcune informazioni su questo e quest'altro". Oppure, se si ha bisogno di un idraulico, si può chiedere a qualcuno: "conosci un buon idraulico"? Una Babysitter, un medico? O magari, si tratta di un buon prodotto? Qualcuno conosce questo televisore? Fa bene questo lavoro? Devo comprare quello? Si tende a rivolgersi agli esseri umani o a cose scritte da esseri umani.

Poi tutto ad un tratto abbiamo visto arrivare questi motori di ricerca che si limitavano a raccogliere tutte quelle pagine in giro, ma in realtà non utilizzavano una quantità enorme di informazioni umane. D'accordo, i link sono stati messi lì da esseri umani e certamente, qualche essere umano ha dovuto scrivere quel contenuto, ma in qualche modo abbiamo perso un altro aspetto dell'elemento umano che era là fuori: il grande volume di raccomandazioni che stavano lì.

Questa cosa sta accedendo con la Ricerca 4.0. La prima cosa che la Ricerca 4.0 sta facendo è che i motori di ricerca hanno iniziato a prendere in considerazione le ricerche che abbiamo fatto nel corso del tempo. Se riescono a stabilire che siamo tornati spesso su un sito informatico, come Diverge o CNET, allora possono dire: "la prossima volta che cerchi qualcosa, daremo maggior peso a quei siti per dar loro una spinta in alto nei risultati, perché sembra davvero che ti piaccia il contenuto che c'è lì. Per cui diamo loro una sorta di ricompensa per questo". Oppure "Vediamo che cerchi viaggi in questo momento e che hai appena cercato New York. Invece di far finta che queste cose siano scollegate, le metto insieme alle tue ricerche successive perché probabilmente sei alla ricerca di informazioni su un viaggio a New York, anche se non hai usato tutte quelle parole. Così faccio uso della vostra storia che sta già lì".

L'altra cosa che hanno fatto (e parte di questa era mischiata insieme al resto, in un primo tempo), è la ricerca localizzata intorno alla vostra posizione geografica. Se fate una ricerca sul "football" nel Regno Unito, probabilmente, la maggior parte delle volte non volete ottenere informazioni sulla NFL (Lega nazionale del football americano). Volete informazioni su quello che gli americani chiamerebbero "soccer" (calcio). Quindi, sapere che siete nel Regno Unito, quando eseguite una ricerca per il "football", aiuta il motore di ricerca dire: "dovremmo andare fino in fondo e venire fuori con informazioni che sono rilevanti per il Regno Unito, o rilevanti per la Stati Uniti, in base a dove ti trovi". Questo però ha subito un forte cambiamento e oggi si va a verificare anche l'area metropolitana. Se fate una ricerca per lo zoo e siete a Seattle, avrete informazioni sui giardini zoologici che stanno a Seattle, piuttosto che sullo Zoo

di Washington, o sui giardini zoologici di Detroit o così via.

L'ultima cosa, quella davvero molto interessante è l'uso dei social network, di cui i motori di ricerca stanno ancora cercando venire a capo. Ho parlato in precedenza dell'idea di considerare i link come dei voti e mi piace sempre usare questa analogia: i link sono come i voti e sono in qualche modo la democrazia del web (che è come Google ancora li descrive su alcune delle sue pagine), quindi la democrazia del web è com'era all'inizio, la democrazia negli Stati Uniti, quando per votare, si dovevano avere almeno 25 anni, essere bianchi e possedere una proprietà. Non era realmente rappresentativo di tutti gli individui.

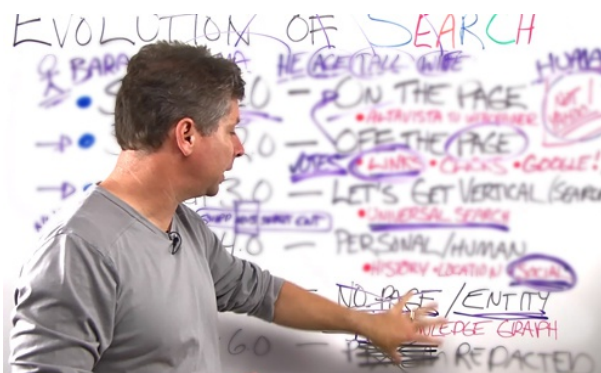
In un sistema di questo tipo, votare era così: "accidenti, era un ottimo ristorante quello in cui sono andato! Voglio scrivere un post sul blog su questo ristorante e ho intenzione di mettere un link al sito del ristorante. Inoltre, voglio assicurarmi di usare una piattaforma che non imponga automaticamente cose come 'no follow' sul link, perché impedirebbe al mio sito di passargli il credito. Oh, e siccome è un ottimo ristorante, mi ricorderò di fare in modo che il testo di ancoraggio, o le parole vicino al testo di ancoraggio, siano qualcosa come 'ottimo ristorante', perché voglio fare in modo che il link sia rilevante e risulti appartenere al contesto semantico giusto. Ora, quando avrò fatto tutto questo, avrò dato il mio voto".

Probabilmente le altre 99 persone che sono andate al ristorante non faranno tutto questo. Ma ciò che invece queste persone tendono a fare è un "Mi piace" su Facebook, un +1 su Google+, una raccomandazione su Yelp, o utilizzare uno qualsiasi del grande numero di sistemi sociali che permettono efficacemente la gente di votare molto più facilmente. Penso che gran parte del futuro in cui ci accingiamo ad entrare, sia in questa direzione "sociale". Questi segnali sociali sono molto, molto importanti per il futuro di come i motori di ricerca determineranno quali sono le migliori pagine che ci sono in giro.

Purtroppo, ci hanno messo dentro così tanto in tutto questo sistema complesso di link, per stabilire se questo è un buon link, o se questo è un cattivo link, se questo è un link che ci accingiamo a destituire di credito (disavow), o se questo è un link che voi stessi avete rinnegato e così via, che hanno ancora bisogno di lavorarci su, per rendere migliori tutti questi elementi sociali. Questo aspetto diventerà importante. Non dico che i link scompariranno come segnali, ma credo che i segnali sociali stiano per entrare in scena molto più pesantemente, nel prossimo futuro.

Ora, lungo la strada fino qui, pensavo una cosa; siccome mi è stato chiesto: "vuoi parlare dell'evoluzione della ricerca?" Ho risposto, "sì, non è un problema perché ho già parlato della ricerca da 1.0 a 4.0". C'è un intero [post sul blog se si cerca "Search 4.0"](http://searchengineland.com/search-40-putting-humans-back-in-search-14086) [http://searchengineland.com/search-40-putting-humans-back-in-search-14086].

Ricerca 5.0



Stavo pensando: "Cosa viene dopo?" Mentre sudavo nel salire le scale, non la scala dell'edificio di MOZ. C'è una scala, perché stavo al livello del mare e ho dovuto salire, a quanto pare, di 100 metri per arrivare a piedi fin qui, dove si trova l'edificio Moz. Se ci fosse stata una swear jar (barattolo dove chi dice parolacce paga pegno mettendo dei soldi), ci avrei messo un dollaro.

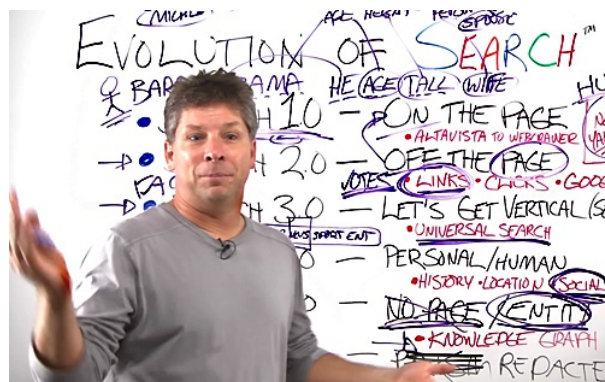
Ricerca 5.0 e questa è davvero la ricerca dove le pagine non ci sono proprio. Ricordate i fattori on-the-page ed i fattori, off-the-page, che sono fuori da una pagina, ma comunque, su qualche altra pagina? In questo caso non mi interessa nemmeno che si tratti di una pagina. Ho scritto un post sul blog, di cui non riesco a ricordarne il titolo, ma se cercate "[Google conversational search](http://searchengineland.com/googles-impressive-conversational-search-goes-live-on-chrome-160445)" [http://searchengineland.com/googles-impressive-conversational-search-goes-live-on-chrome-160445] lo troverete. Se non lo trovate, vuol dire che chiaramente Google è un pessimo motore di ricerca.

Nella cosa della ricerca conversazionale che dimostravo, se avete Chrome e fate clic sul microfono, ora sui desktop, potete parlare a Google, un po' come si può fare al telefono. Si può dire: "Barack Obama" e Google allora vi mostrerà i risultati per Barack Obama, e poi potete parlare di nuovo voi e dire: "Barack Obama è il presidente degli Stati Uniti," bla bla bla bla. Vi mostrerà un piccolo box tematico su di lui, con la sua immagine e un po' di testo estratto da Wikipedia.

Quindi si può dirgli: "quanti anni ha?", o qualcosa di molto simile. Allora il motore di ricerca risponderà: "Barack Obama ha ... anni" (non ricordo quanti anni ha). Dovete provare a fare questa ricerca vocale, vi risponderà dicendo Barack Obama ha 51 anni. Si può continuare e chiedere: "bene, quanto è alto?" Risponderà: "Barack Obama è alto ..." Penso che sia 1 metro e 85. Poi dite: "con chi è sposato?" e Google risponde: "Barack Obama è sposato con Michelle Obama". Voi chiedete: "quanti anni ha?" Allora Google risponde e dice: "è una cosa molto scortese chiedere l'età ad una donna, ma lei ha ... anni". Credo 39. Sì, di solito con questo numero si va sul sicuro.

Per riuscire a fare tutto questo, Google deve capire che Barack Obama, quando hai cercato lui, non era solo un insieme di lettere su una pagina web. Doveva capire che lui è una persona, che è una entità, se volete, una persona, un luogo, o una cosa. Non un sostantivo, ma un soggetto. Deve capire che c'è una cosa in giro chiamata Barack Obama, che si può collegare alla ricerca e di cui si può sapere qualcosa. Quando chiedete la sua età, ed avete detto: "How old is he", ha dovuto capire che "he" non era solo una parola, ma che in realtà "he" si riferiva a una entità che si era specificata prima, l'entità Barack Obama. Quando avete detto "his age", "age" non erano solo un mucchio di lettere che corrispondono a una pagina web, ma l'età è uguale ad un valore che Google conosce perché l'entità Barack Obama possiede un valore di età e la collega alla domanda.

Quando avete detto: "How tall is he" ha fatto la stessa cosa. Quel "tall" non erano solo lettere, ma "tall" è in realtà un elemento di altezza che conosce. Questo la dice altezza, fidatevi. Quando avete detto: "Who's his wife" la moglie, ha capito che sua moglie è una persona che significa coniuge, che è una cosa che Google comprende, un'entità. Ancora una volta, non si tratta solo di parole. E' come una cosa che comprende davvero: è Michelle e sa tutte queste cose su di lei.



Tutto ciò è molto diverso dalla Ricerca 1.0 dove eravamo davvero solo in cerca di lettere su una pagina. Quando avete cercate "film", rispondeva "quante pagine ho che hanno queste cinque lettere in questa sequenza? Iniziamo a contarle ed a metterle insieme".

Qui cerchiamo le entità ed il grafo della conoscenza di Google (knowledge graph) è una dimostrazione di dove stiamo

andando. E' anche tutto molto entusiasmante e per chi lavora nel marketing è sempre stimolante quando il mondo dove si muove sta cambiando, perché se si sta sulla

cresta dell'onda e si vede dove siamo diretti, si prospettano sempre nuove opportunità. E' anche eccitante perché alcuni dei vecchi modi non funzionano più, e la Ricerca 5.0 potrà meglio remunerare le cose che stanno venendo avanti.

E' anche un po' preoccupante che Google venga a conoscenza delle entità ed impari a conoscere cose come i fatti, che se siete alla ricerca di film in un certo posto, possa magari anche decidere che: "ho un database di tutti quei film. Non ho più bisogno di puntare a una pagina web che ha questo genere di informazioni". La lezione che se ne ricava è che se il vostro lavoro consiste solo nel creare pagine web che riportano tutti fatti già noti che si trovano in giro, allora diventerà sempre più difficile, perché le persone non verranno più indirizzate da voi sui fatti che stanno fuori da Google. La gente verrà indirizzata verso fatti a cui Google può rispondere direttamente. Il vostro compito è quello di assicurarvi di avere sempre le informazioni che Google non ha, i fatti che non si trovano facilmente in giro.

Ricerca 6.0

Per quanto riguarda la ricerca 6.0, ha coinvolto il sistema PRISM, ma non possiamo parlarne più, così in qualche modo è scomparsa ed è rimasta fuori della discussione. In pochi anni da adesso, non avrà alcun senso. In questo momento, si spera, sia ancora molto attuale.

Penso che questo sia tutto. Quindi vi ringrazio per la vostra indulgenza con la mia prima Whiteboard Friday solista. Spero di non essere andato troppo veloce. Spero che tutto per voi abbia un senso, e vi ringrazio molto.



Danny Sullivan

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 qui:
www.idealwebitalia.it/seo/5208/