

Il controllo ortografico di Google per le ricerche

Il dizionario per le ricerche

Autore: Rael Dornfest, Paul Bausch e Tara Calishain

Tratto dal libro: [Google Trucchi e segreti](#)

A volte Google si prende la libertà di “correggere” ciò che percepisce come errore ortografico all’interno della query di ricerca. Si tratta del dizionario di Google che quando si sbaglia ascrivere suggerisce quella che ritiene la parola corretta. In questo articolo vedremo come funziona questo sistema e come integrarlo all’interno di uno script in perl.

Controllare l’ortografia

Molti di noi non riuscirebbero a comunicare con il mondo esterno senza un correttore ortografico. Quando inviate un’e-mail o eseguite i ritocchi finali a un documento, un correttore ortografico affidabile si accerterà che non abbiate commesso errori banali.

Anche Google possiede un correttore ortografico incorporato e, quando ritiene di poter scrivere una singola parola o completare una frase nella query di ricerca meglio di voi, suggerisce una ricerca “migliore”, collegandola direttamente alla query.

Conoscete benissimo l’ortografia italiana? Ottimo, ma forse non vale lo stesso per l’inglese. Google vi viene in aiuto particolarmente con questa lingua. Ad esempio, se cercate *hydrecefallus*, Google vi chiederà se intendevate *hydrocephalus*, come mostrato nella *figura 13*.

Suggerimenti a parte, Google presume che sappiate ciò di cui state parlando e restituisce i risultati richiesti, a condizione che la query li abbia generati.

Se non vengono trovati risultati per la query a causa dell’ortografia che avete fornito, e Google ritiene di saperne di più, eseguirà automaticamente una nuova ricerca in base ai suoi stessi suggerimenti. Così, se la ricerca di *hydrecefallus* non trova (e lo credo bene) alcun risultato, viene eseguita la ricerca avviata da Google di *hydrocephalus*.

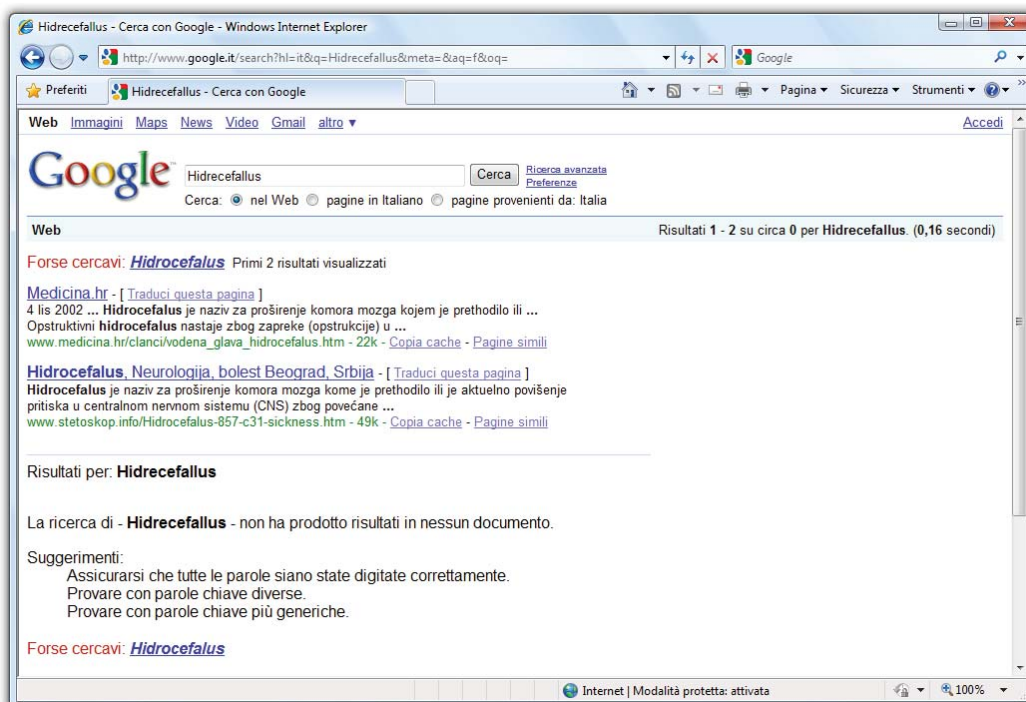


Figura 1: Google pensa di saperne più di voi sull’ortografia inglese

Tenete presente che Google non propone arbitrariamente i suggerimenti, ma li crea in base al proprio database di parole e frasi che trova durante l'indicizzazione del Web. Se cercate una parola senza senso, come *kweghjdlsgaa*, non otterrete alcun risultato né alcun suggerimento.

Questa funzionalità è un piacevole effetto collaterale e rappresenta un facile modo per controllare la frequenza relativa dei modi in cui vengono scritte le parole. Richiedete la modalità di scrittura di un termine particolare e annotate il numero di risultati. Quindi fate clic sulla modalità di scrittura del termine suggerita da Google (**Forse cercavi**) e annotate il numero di risultati. È sorprendente come questi numeri siano a volte vicini, indicando un termine o una frase spesso scritta in modo scorretto.

Accettare gli errori ortografici

Non fate l'errore di scartare automaticamente i risultati del controllo ortografico di una parola scritta in modo errato, in particolare un nome proprio. A titolo di esempio, il nome del noto autore di fumetti Bill Mauldin viene spesso scritto in modo erroneo come "Bill Maudlin". Con una rapida ricerca in Google, sembra che siano molte le persone che commettono questo errore. Non esiste alcuna regola che stabilisce che ogni pagina venga sottoposta al controllo ortografico prima di essere pubblicata online, perciò vale spesso la pena di osservare i risultati nonostante questi errori.

A titolo di esperimento, provate a cercare due parole scritte erroneamente su un argomento correlato, come ad esempio *normotensive hydrocephalis*. Che tipo di informazioni ottenete? Le informazioni ottenute, se presenti, possono essere raggruppate in un particolare genere online?

Al momento della scrittura di questo articolo, la ricerca di *normotensive hydrocephalis* determina solo tre risultati. Il contenuto riguarda generalmente persone che hanno a che fare con vari problemi neurochirurgici. Di nuovo, non esiste alcuna regola che stabilisce che tutto il materiale Web debba essere sottoposto a controllo ortografico.

Utilizzate ciò a vostro vantaggio come ricercatore. Quando cercate resoconti di profani attinenti a malattie o lesioni, il contenuto desiderato potrebbe in effetti essere scritto erroneamente. D'altro canto, se cercate informazioni o riferimenti altamente tecnici da fonti attendibili, filtrare le richieste scritte erroneamente vi porterà più vicini alle informazioni desiderate.

Ortografia sulla riga di comando

Il fatto che Google raccolga l'ortografia dal Web invece che da un dizionario significa che può essere migliore di molti correttori ortografici per e-mail ed elaboratori di testi. Un correttore ortografico per e-mail non rileva il nome errato dell'attore *Dave Shapel* (o è *Dave Chapelle?*), mentre quello di Google lo rileva.

Sebbene questo hack non sostituisca i controllori ortografici standard con Google, il codice riportato qui vi mostrerà come portare il correttore ortografico un po' più vicino al desktop.

Il codice

Questo codice contatta l'API Google e chiede suggerimenti ortografici per la parola o la frase fornita. Se non siete già abituati a utilizzare la riga di comando per eseguire le vostre ricerche, questo hack non renderà il fatto di contattare Google più facile che aprire un browser Web. Ma per coloro che sono ormai esperti nell'utilizzo della riga di comando, è un modo veloce per sfruttare la potenza ortografica di Google.

Salvate il codice seguente come spell.pl e accertatevi di inserire qui la chiave con la vostra chiave API Google:

```
#!/usr/local/bin/perl
# spell.pl
# Contatta Google per suggerimenti ortografici!
# Uso: perl spell.pl <query>

# La chiave dello sviluppatore per l'API Google. my $google_key='inserire qui
la chiave';
# Posizione del file WSDL GoogleSearch. my $google_wsdl = "./GoogleSearch.wsdl";

use strict;

# Usa il modulo SOAP::Lite Perl.

use SOAP::Lite;

# Prende la query dalla riga di comando.
my $query = join(' ',@ARGV) or die "Usage: perl spell.pl <query>\n";

# Crea una nuova istanza SOAP::Lite, portandola in GoogleSearch.wsdl.
my $google_search = SOAP::Lite->service("file:$google_wsdl");

# Interroga Google.
my $results = $google_search -> doSpellingSuggestion($google_key, $query);

# Nessun risultato?
if ($results) { print $results; }
```

Questo script è simile a qualsiasi script Perl di base per contattare l'API Google, ma utilizza il metodo doSpellingSuggestion invece di quello di ricerca standard.

Esecuzione

Eseguite lo script dalla riga di comando, passando qualsiasi parola o frase che desiderate controllare, nel modo seguente:

```
% perl spell.pl inserire parola o frase
```

Passando *Dave Shapel*, potete vedere in che modo Google suggerisce come scrivere questo nome:

```
% perl spell.pl Dave Shapel
Dave Chapelle
```

Se passate una parola corretta, lo script non restituisce alcun suggerimento. Per usare questo script resta sempre da capire quali parole sono dubbie, ma quando dovete verificare velocemente un nome o una frase, potete pensare a Google come al vostro lessiconografo personale (o è lessicografo?).

Conclusioni

In questo articolo abbiamo scoperto come funziona il controllo ortografico che Google effettua sulle query di ricerca degli utenti e come sia possibile usare queste informazioni. In fine abbiamo visto come collegare uno script perl a questa funzione.



Resta sempre aggiornato sulle novità del sito

Per mantenerti sempre aggiornato su nuovi contenuti interessanti, Come fare a... vi offre la possibilità di abbonarvi gratuitamente alla **Newsletter Come fare a...** all'indirizzo: <http://www.comefarea.it/newsletter/> o, se usate Windows Live Messenger, di abbonarvi ai nostri **Windows Live Alerts** all'indirizzo <http://www.comefarea.it/abbonamenti.php>. Per gli utenti di Mac OS X Leopard è anche disponibile gratuitamente un Widget che vi terrà sempre informati sulle ultime novità

Il libro



Google Trucchi e segreti

Google è uno strumento eccezionale per cercare nel Web, ma sapevate che offre anche tantissimi modi per organizzare e lavorare con le informazioni? Dalle grandi immagini di Google Earth alle piccole schermate di Google Mobile, questo libro vi rivela i segreti di Google. Contiene numerosi trucchi per usare i nuovi servizi di Google e utili indicazioni per usare meglio i servizi più tradizionali.

Se desideri acquistare questo libro puoi farlo direttamente online sul sito di Tecniche nuove a questo indirizzo:

http://www.tilibri.com/libri/google_trucchi_e_segreti_703.html

La recensione

Per conoscere e approfondire gli argomenti trattati del libro è disponibile la recensione di questo manuale. Visita la pagina Google Trucchi e segreti disponibile all'indirizzo:

<http://www.comefarea.it/recensioni/googletrucchiesegreti/>

Gli autori

Rael Dornfest: è il responsabile dei sistemi informative in O'Reilly Media. Tra le altre cose, il suo compito prevede anche il monitoraggio di O'Reilly Network. Ha già scritto numerosi testi, tutti presenti nella collana Trucchi e segreti).

Paul Bausch: è uno sviluppatore web indipendente. Quando non si occupa di applicazioni web, ne scrive.

Tara Calishain: è un'esperta di motori di ricerca Internet e del loro utilizzo in situazioni d'affari