

GET SEEN

Analisi e Soluzioni Seo Sito: Efficacemente.com



Controllo Codice

Dopo un controllo del codice html visibile del sito e della struttura di navigazione alla ricerca di involontarie strozzature del passaggio degli spider, non è stato rilevato nulla di insolito. L'utilizzo di Wordpress come Cms, del resto, mette il webmaster/sviluppatore e poi il gestore del sito al riparo da errori di programmazione o mark up delle pagine web che potrebbero mandare in tilt i motori di ricerca.

Tema grafico

Il tema grafico scelto, inoltre, pur non essendo chiaramente studiato per finalità di ranking sui motori (lo si denota innanzitutto per una disinvolta scelta delle intestazioni h1 e h2), non presenta tuttavia alcun impedimento fisico al caching delle pagine.

In ottica di miglioramento generale della visibilità online molto si potrebbe fare sul layout scelto per Efficacemente.com, ma il problema del sito, come denunciato dal titolare, risiede nell'aver perso le chiavi legate al brand e un crollo nei referrer dai motori di ricerca, Google in primis.

Controllo pagine indicizzate su Google

Google tende ad accogliere nel suo indice un po' di tutto, per scelta aziendale persino la validazione W3C non è requisito essenziale per il posizionamento. Laddove vi sia un contenuto testuale unico, il motore farà di tutto per assicurarlo ai suoi utenti. Eppure Efficacemente.com presenta una notevole carenza di pagine indicizzate (indicizzazione = download e inserimento in archivio di una copia testuale della pagina web), a partire dall'home page.

<http://209.85.129.132/search?hl=en&q=cache%3Ahttp%3A%2F%2Fwww.efficacemente.com%2F&btnG=Search>

L'home page, non presente negli archivi, presenta un Pagerank visibile sulla Google Toolbar pari a 3. Segno, questo, che l'Url dell'home page è noto a Google. I contenuti no.

Da una query site: (l'operatore di Google che elenca le pagine effettivamente indicizzate), spuntano alcuni url di articoli:

http://www.google.it/search?q=site%3Aefficacemente.com&sourceid=navclient-ff&ie=UTF-8&rlz=1B3GGGL_enIT300IT301

Però visualizzando il primo risultato: <http://www.efficacemente.com/2009/04/definizione-obiettivi/> e volendone visualizzare la copia cache, non si ottiene nulla:

<http://209.85.129.132/search?sourceid=navclient-ff&ie=UTF-8&q=cache%3Ahttp%3A%2F%2Fwww.efficacemente.com%2F2009%2F04%2Fdefinizione-obiettivi%2F>

L'ipotesi quindi che il problema sia da ricercarsi ad un livello superiore, ossia nel server/hosting del sito, prende quindi forma.

Controllo degli Header

L'header è un segmento del pacchetto dati parte integrante della comunicazione tra browser e spider (che funzionano grosso modo come dei browser) e il server che ospita i dati richiesti (alias, l'hosting del sito web). Gli header restituiti dal server possono farci capire molto del problema.

A titolo esemplificativo, di seguito il segmento header di un sito perfettamente indicizzato su Google, con caratteristiche equivalenti a efficacemente.com : server unix, Apache, php, sito su Wordpress con i pretty-permalink attivati.

```
#1 Server Response: http://www.abcfinanze.com/  
HTTP Status Code: HTTP/1.1 200 OK  
Date: Wed, 20 May 2009 09:29:33 GMT  
Server: Apache/2.2
```

X-Powered-By: PHP/4.4.9
Expires: Thu, 19 Nov 1981 08:52:00 GMT
Cache-Control: no-store, no-cache, must-revalidate, post-check=0, pre-check=0
Pragma: no-cache
X-Pingback: <http://www.abcfinanze.com/xmlrpc.php>
Set-Cookie: PHPSESSID=be2e3e137a3a7de43c7c45aa982a283b; path=/
Connection: close
Content-Type: text/html; charset=UTF-8

Come si vede, il server ad un richiamo di Url restituisce tutta una serie di informazioni utili agli spider, come l'url di riferimento (www.abcfinanze.com) , lo status http (200 OK , ossia la pagina esiste ed è scaricabile), la data di accesso ecc..

Ora facciamo un controllo degli headers per [Efficacemente.com](http://www.efficacemente.com) , che come detto ha le stesse caratteristiche del sito prima analizzato:

#1 Server Response: <http://www.efficacemente.com/> The operation timed out

L'operazione di richiesta pacchetto header quindi va in time out immediatamente. Nonostante il sito sia perfettamente visualizzabile e navigabile. Abbiamo riprovato la richiesta in momenti diversi della giornata di ieri, e di stamattina, ottenendo lo stesso risultato.

Ipotesi finale

Premettendo che non sono stati fatti controlli su link in entrata e in uscita per la semplice ragione che qualsiasi errore, abuso, linkaggio sospetto avrebbe portato il sito ad una penalizzazione ma non alla cancellazione dell'indice, e avendo verificato che non vi sono tecniche illecite (spam) sulle pagine del sito, **il problema attuale del sito deriva dall'HOSTING**

Anche se negli ultimi tempi non è facile osservare behaviour di questo tipo, non è infrequente che alcuni servizi di hosting "blocchino" il passaggio degli spider sui siti ospitati presso i propri server. Questo per preservare la banda disponibile e la velocità di accesso: per bloccare uno spider infatti non è necessario un blocco specifico, basta negare agli spider la visualizzazione degli header. A quel punto i browser normali (che scaricano questa informazione, ma non la ritengono conditio sine qua non per la visualizzazione della pagina) potranno visualizzare correttamente il sito, ma per gli spider quel sito non esiste.

Questo il motivo per il quale [efficacemente.com](http://www.efficacemente.com) piano piano sta perdendo tutte le pagine da Google, cominciando dall'home page, come detto prima non indicizzata.

Soluzione

- a. Informare il proprio hosting, chiedendo all'assistenza tecnica, del problema reale rilevato: il server non restituisce gli header (chiedere al supporto di fare una prova immettendo l'url <http://www.efficacemente.com/> in un tool di analisi degli header, oppure in una finestra DOS). Chiedere quindi che il problema venga risolto altrimenti il sito non potrà essere visitato dagli spider e posizionato sui motori (di fatto, per i motori , il sito non esiste al momento)
- b. Se il problema persiste, ossia l'hoster non cambia la situazione, senza produrre motivazioni, è vitale cambiare hosting.

Eseguito uno dei punti sopra indicati, e quando gli header del sito saranno correttamente visualizzati (gli header di [efficacemente.com](http://www.efficacemente.com) DEVONO essere pressochè identici a quelli di [Abcfinanze.com](http://www.abcfinanze.com)), il sito tornerà tranquillamente nell'indice di Google e il problema sarà risolto.