Consulenza SEO Google Ads Training Guide Blog About me Contatti



Sei qui: EVE Milano / Posizionamento Motori di Ricerca / Come creare la Sitemap.xml perfetta

Come creare la Sitemap.xml perfetta

Autore: Giovanni Sacheli Data di pubblicazione: 28/06/2013 Ultimo aggiornamento: 28/09/2023 Tempo di lettura: 12 minuti Commenti: 5 🖼



La sitemap.xml è un elemento molto importante per un sito web. questo file infatti

Richiedi un preventivo

che Google conosca tutte le pagine che hanno del contenuto di qualità e la sitemap.xml è lo strumento perfetto.

Per prima cosa dobbiamo fare in modo che i MdR trovino facilmente la sitemap.xml, quindi bisogna creare un riferimento in un file che tutti i bot leggono (o dovrebbero): il file robots.txt. La sitemap.xml dovrebbe essere linkata dal file robots.txt che è il primo file che i crawler leggono quando atterrano su un sito web:

```
User-agent: *
Disallow: /admin/
Sitemap: http://www.esempio.com/sitemap.xml
```

Google nelle sue linee guida fornisce consigli e risorse dove è possibile studiare la struttura che una sitemap.xml dovrebbe avere, tuttavia molti webmaster ignorano queste indicazioni e generano sitemap.xml spesso rudimentali ed incomplete. Con questo articolo vorrei riassumere le informazioni necessarie per creare una sitemap.xml perfetta.

Cos'è la Sitemap.xml

La sitemap.xml è un file, appunto in formato XML, che contiene l'elenco formattato degli URL del sito web nel quale è contenuta. In questo elenco è possibile specificare, per ogni URL, alcune informazioni utili ai motori di ricerca:

- la data dell'ultimo aggiornamento lastmod
- la frequenza di aggiornamento changefreq
- l'importanza che l'URL ha in relazione a tutte le altre pagine del sito priority
- eventuali immagini presenti nell'URL
- eventuali tag hreflang

Come creare la Sitemap.xml

A volte mi chiedono: "come posso generare una sitemap per il mio sito?"

Per una massima efficacia la Sitemap.xml dovrebbe essere **scritta e aggiornata automaticamente** dal CMS ogni qual volta un contenuto viene aggiornato o aggiunto exnovo. La cosa migliore è quindi avvalersi di plugin o moduli dedicati a questo compito specifici per il proprio CMS.

Non è pratico usare un software per creare una sitemap statica, a meno che non si abbia la certezza che il sito web non verrà mai più modificato o aggiornato.

Dalla versione 5.5 – WordPress crea la sua sitemap.xml senza l'utilizzo di plugin. Puoi raggiungere la sitemap all'indirizzo tuosito.com/wp-sitemap.xml. Se invece preferisci affidarti a plugin dedicati ti consiglio Google XML Sitemaps ♂.

L'opzione che mi piace di meno è generare manualmente la sitemap.xml, dato che per aggiornarla bisogneà ripetere la procedura ad ogni modifica fatta al sito. Per fare una sitemap.xml statica ti consiglio di usare Screaming Frog, qui trovi la guida per generare la sitemap.

Limiti da rispettare

Direnta la compilazione della citamon uni à importante richettare elevai limiti ed recelz

Richiedi un preventivo

AIVIL. TULLIT VAIOTT del dati in una ortemap devono essere *entity-escaped*. Il me stesso deve essere codificato in UTF-8.

Ad esempio, possono venire elencati fino ad un massimo di 50.000 URLs e la dimensione del file Sitemap.xml *non compresso* non deve superare i 50 MB.

Inoltre, la Sitemap deve:

- Iniziare con un tag di apertura <urlset> e terminare con un tag di chiusura </urlset>.
- Specificare il namespace (standard di protocollo) all'interno del tag <urlset>.
- Includere un tag <url> per ogni URL, come tag XML principale.
- Includere un tag <loc> figlio per ogni tag <url> principale.

Tutti gli altri attributi che vedremo a breve sono facoltativi. Il supporto per questi tag opzionali può variare tra i motori di ricerca. Fare riferimento alla documentazione di ciascun motore di ricerca per i dettagli. Inoltre, tutti gli URL in una Sitemap devono provenire da un unico host, come www.example.com o store.example.com.

Vediamo un esempio di Sitemap.xml con due URL:

Come puoi notare la Sitemap.xml inizia con il tag di apertura "urlset" con il quale si dichiara lo **standard del protocollo utilizzato**. Di seguito trovi elencati tutti gli attributi, obbligatori e facoltativi, che una Sitemap.xml completa dovrebbe utilizzare.

Gli attributi della Sitemap.xml

<urlset>

Valore obbligatorio. Comprende il file e fa riferimento agli standard del protocollo corrente.

<url>

Valore obbligatorio. Tag principale per ogni voce URL. I tag restanti sono secondari di questo tag.

<10c>

Richiedi un preventivo

http) e terminare con una barra finale, se richiesto dal server web. Questo valore deve contenere meno di 2048 caratteri.

<lastmod>

Valore facoltativo. Questo campo deve includere la data dell'ultima modifica del file e secondo me è l'attributo facoltativo più importante di tutti perché comunica al motore di ricerca cosa deve essere scansionato con priorità. Questa data dovrebbe avere il formato W3C Datetime, che ti consente di omettere l'ora, se lo desideri, e di utilizzare il formato AAAA-MM-GG. Questo tag è separato dall'intestazione If-Modified-Since (304) che il server può restituire ed è possibile che i motori di ricerca utilizzino queste informazioni da entrambe le sorgenti in modo differente.

<changefreq>

Valore facoltativo. Frequenza con la quale la pagina potrebbe venire modificata. Questo valore fornisce informazioni generiche ai motori di ricerca ed è possibile che non sia esattamente correlato alla frequenza di scansione della pagina. I valori validi sono:

- always
- hourly
- daily
- weekly
- monthly
- yearly
- never

Il valore "always" deve essere utilizzato per descrivere i documenti che subiscono modifiche ogni volta che vi si accede. Il valore "never" deve essere utilizzato per descrivere gli URL archiviati.

Il valore di questo tag viene considerato un suggerimento e non un comando. Anche se i crawler dei motori di ricerca possono prendere in considerazione questa informazione, è possibile che la scansione di pagine contrassegnate come "hourly" avvenga meno di frequente rispetto alla scansione di pagine contrassegnate come "yearly". È possibile che i crawler eseguano la scansione periodica di pagine contrassegnate come "never", in modo da poter gestire modifiche impreviste apportate a tali pagine.

<priority>

Valore facoltativo. La priorità di questo URL rispetto ad altri URL del sito. L'intervallo di valori validi è compreso tra 0,0 e 1,0. Questo valore non influisce sul modo in cui le tue pagine vengono confrontate con le pagine di altri siti, ma ti consente di indicare ai motori di ricerca le pagine che reputi più importanti per i crawler.

La priorità predefinita di una pagina è 0,5.

È probabile che la priorità assegnata a una pagina non influisca sulla posizione degli URL nelle pagine dei risultati di un motore di ricerca. I motori di ricerca possono utilizzare

Richiedi un preventivo

oi

importanti nell'indice di ricerca.

Inoltre, è probabile che l'assegnazione di un'alta priorità a tutti gli URL del tuo sito non sia di aiuto. Poiché la priorità è relativa, essa viene utilizzata solo per selezionare gli URL del tuo sito.

Ci possono essere situazioni dove una sola sitemap non basta, ad esempio per siti con più di 50.000 pagine. In questi casi si devono generare più sitemap ed una sitemapindex che le comprenda tutte. Vediamo un esempio di sitemapindex.xml

La Sitemap_index.xml

Puoi fornire più file Sitemap, infatti se desideri elencare più di 50.000 URL, devi creare più file Sitemap.

Se fornisci più Sitemap, dovresti quindi elencare ciascun file Sitemap in un file Sitemap_index. Anche questa tipologia di sitemap non può contenere più di 50.000 Sitemap e non deve superare i 50 MB (52.428.800 byte) e può essere compressa. Puoi

avere più di un file di indice Sitemap. Il formato XML di un file Sitemap_index è molto simile al formato XML di un file Sitemap.

Il file Sitemap_index deve:

- Iniziare con un tag di apertura <sitemapindex> e terminare con un tag di chiusura
 </sitemapindex>.
- Includere una voce <sitemap > per ciascuna Sitemap come tag XML principale.
- Includere un tag <loc> figlio per ogni tag <sitemap> principale.
- Il tag < loc > facoltativo è disponibile anche per i file di indice Sitemap.

Dove posizionare la Sitemap.xml

Si consiglia di inserire il file Sitemap.xml nella root del web server poichè gli URLs contenuti devono risiedere nella stessa struttura di folder della sitemap. In altre parole, il percorso di un file Sitemap determina l'insieme di URL che può essere inserito nella Sitemap, ad esempio:

" Un file Sitemap posizionato in http://esempio.it/catalog/sitemap.xml può includere tutti gli URL che iniziano con http://esempio.it/catalog/ ma non quelli che iniziano con http://esempio.it/images/.

Richiedi un preventivo

La Sitemap.xml per i siti multi-lingua organizzati in subfolder

Immaginiamo un sito in due lingue, Italiano ed Inglese, organizzato in sub-folder:

- http://www.esempio.com/it/
- http://www.esempio.com/en/

In questo caso si consiglia di generare tre sitemap distinte:

- Un indice di sitemap: http://www.esempio.com/sitemapindex.xml
- La sitemap per gli URL in lingua Italiana: http://www.esempio.com/it/sitemap.xml
- La sitemap per gli URL in lingua Inglese: http://www.esempio.com/en/sitemap.xml

Questa suddivisione aiuta Google ad interpretare e suddividere le lingue nel sito, soprattutto in casi dove il rel alternate non sia ben implementato.

In un sito multi lingua ti consiglio anche di implementare i tag hreflang.

Come riprodurre i caratteri speciali nella Sitemap.xml

Di seguito trovi la versione escape dei caratteri speciali più comuni.

Carattere e Codice di escape:

```
E commerciale &: &
Virgolette semplici ': '
Virgolette doppie ": "
Maggiore di >: >
Minore di <: &lt;
```

Come verificare la correttezza della Sitemap.xml

Ho scritto una guida dedicata alla verifica della sitemap con Screaming Frog che ti consiglio di leggere.

Online esistono diversi tool per testare la validità della Sitemap.xml, ne elenco alcuni che ho provato:

- Google Search Console
- http://www.xml-sitemaps.com/validate-xml-sitemap.html
- http://www.automapit.com/sitemapvalidator.html &
- http://www.xmlcheck.com/
- http://seogadget.com/check-your-xml-sitemap-errors/
- http://www.inspyder.com/free/SitemapValidator.aspx

Linee guida e suggerimenti di Google

Attenzione: i suggerimenti dati da Google potrebbero essere diversi dal <u>protocollo</u> standard .

Utilizza URL uniformi e completi. Google eseguirà la scansione degli URL
 esattamente come indicati. Ad esempio, se il tuo sito si trova all'indirizzo.

Richiedi un preventivo

****** oppore , pagnianiani (<u>otte i ciativo</u>).

- La sitemap.xml può essere pubblicata ovunque sul tuo sito, ma potrà includere solo
 URL figli della directory dove è posizionata. Pertanto, una Sitemap pubblicata nella
 directory principale del sito può includere tutti i file presenti sul sito.
- Non includere ID sessione negli URL contenuti nella tua Sitemap in modo da ridurre la scansione duplicata di tali URL.
- Informa Google della presenza di versioni in altre lingue di un URL usando le annotazioni hreflang.
- Per i file Sitemap devi utilizzare la codifica UTF-8 e gli URL devono contenere i caratteri di escape appropriati.
- Suddividi le Sitemap di grandi dimensioni in Sitemap più piccole: le dimensioni
 massime consentite per ogni Sitemap sono di 50.000 URL/50 MB senza
 compressione. Usa un file Indice Sitemap per elencare tutte le singole Sitemap e
 inviare a Google solo questo file anziché le singole Sitemap.
- Elenca solo gli URL canonici nelle Sitemap. Se hai due versioni di una pagina, elenca solo quella canonica (selezionata da Google) nella Sitemap. Se hai due versioni del sito (ad esempio, www e non www), stabilisci il tuo sito preferito e pubblica lì la Sitemap, quindi aggiungi elementi rel=canonical o reindirizzamenti sull'altro sito.

- Se hai URL diversi per le versioni desktop e per dispositivi mobili di una pagina, ti
 Google consiglia di indirizzare a una sola versione in una Sitemap. Tuttavia, se ritieni
 necessario indirizzare a entrambi gli URL, annota i tuoi URL in modo da indicare la
 versione desktop e quella per i dispositivi mobili.
- Utilizza estensioni Sitemap per indirizzare ad altri tipi di contenuti multimediali quali video, immagini e notizie.
- Se hai pagine alternative per lingue o aree geografiche diverse, puoi usare hreflang in una Sitemap o in tag HTML per indicare gli URL alternativi.
- Caratteri non alfanumerici e non latini: per il file Sitemap deve essere utilizzata la
 codifica UTF-8 (generalmente è possibile applicare la codifica al momento del
 salvataggio del file). Analogamente a quanto accade con i file XML, tutti i valori dei
 dati (inclusi gli URL) devono utilizzare codici di escape per i caratteri elencati nella
 seguente tabella. Una Sitemap può contenere soltanto caratteri ASCII e non
 caratteri ASCII maiuscoli o alcuni codici di controllo e caratteri speciali quali * e {}. Se
 l'URL della Sitemap contiene tali caratteri, verrà generato un errore quando tenterai
 di aggiungere la Sitemap.
- Tutti gli URL (incluso l'URL della tua Sitemap) devono essere codificati in modo da
 consentirne la lettura ai server web su cui si trovano e devono contenere caratteri di
 escape. Se, invece, utilizzi script, strumenti o file di log per generare gli URL (qualsiasi
 metodo, ad eccezione dell'inserimento manuale), l'operazione viene in genere
 eseguita automaticamente. Se invii la Sitemap e ricevi un messaggio di errore
 relativo all'impossibilità per Google di trovare alcuni URL, verifica che gli URL siano
 conformi allo standard RFC-3986 per gli URI, allo standard RFC-3987 per gli IRI e
 allo standard XML.
 - Di seguito viene riportato un esempio di URL in cui viene utilizzato un carattere non ASCII (ü) e un carattere che richiede codici di escape (&):
 - http://www.example.com/ümlat.html&q=name
 - Di seguito viene riportato lo stesso URL con codifica ISO-8859-1 (per l'hosting su un server che utilizza la stessa codifica) e con le forme di escape opportune:
 - http://www.example.com/%FCmlat.html&q=name
 - Di seguito viene riportato lo stesso URL con codifica UTF-8 (per l'hosting su un server che utilizza la stessa codifica) e con le forme di escape opportune:
 - http://www.example.com/%C3%BCmlat.html&q=name
 - Di seguito viene riportato lo stesso URL con codice di escape:

Richiedi un preventivo

pagine che ritieni importanti; Google non si impegna a eseguire la scansione di ogni URL presente in una Sitemap.

- Google ignora i valori priority e changefreq, quindi non preoccuparti di aggiungerli.
- Google legge il valore, ma se lo rappresenti in modo errato, questo valore non verrà più letto.
- La posizione di un URL in una Sitemap non è importante; Google non esegue la scansione degli URL nell'ordine in cui sono visualizzati nella Sitemap.



Guide correlate a Sitemap

- o 5 modi per trovare la Sitemap.xml di un sito web
- o Come verificare la correttezza di una Sitemap.xml

o Cosa significa ![CDATA[e a cosa serve in un un file XML

Autore

Mi chiamo Giovanni Sacheli, sono un **consulente in Search Marketing** (SEO e PPC) e lavoro nella mia agenzia Searcus Swiss Sagl. Co-autore del libro SEO Audit Avanzato, appassionato di **analisi SEO tecniche** e campagne con Google Ads. Dal 2010 condivido in questo blog guide SEO e nozioni tecniche per consulenti, webmaster e addetti ai lavori.

COMMENTI

Emanuele

Ha contribuito con **1** commento 28/06/2014

ciao, Vorrei chiederle delle informazioni in merito a seo sitemap domini e sottodomini, ha un contatto da lasciarmi, grazie mille.

Rispondi

Lorenzo

Ha contribuito con **4** commenti 06/01/2020

Salve Giovanni,

ho un sito dove ci sono molte landing pages, molti articoli e molti profili che sono indicizzati. Premetto che in tutto non si superano i famosi 50.000 URLs, però quello che vorrei fare è: creare una sitemap per la landing, una per i profili ed un'ultima per gli articoli (e poi inserirle tutte sotto una sitemap index). Vorrei sapere se è possibile fare questa operazione e come farla. Ci sono tool online per

Richiedi un preventivo

Rispondi

Giovanni Sacheli

• 749 risposte 07/01/2020

Ciao Lorenzo grazie per la domanda. Per fare le sitemap a mano puoi usare screaming frog oppure uno dei tanti tool online gratuiti, ma non te lo consiglio. La cosa migliore sarebbe avere un sistema sitemap auto aggiornante (tipo https://wordpress.org/plugins/google-sitemap-generator/), non è pratico aggiornarla a mano ogni volta.

Il setup che hai in mente è corretto e otterrebbe il massimo implementando anche il tag lastupdate, corretto e aggiornato tempestivamente.

A presto!

Rispondi

Sergio

Ha contribuito con **1** commento 10/04/2020

Ciao Giovanni, articolo molto interessante, ma ormai hai abituato bene i tuoi lettori!

Ho un dubbio a proposito di sitemap.xml e url di prodotti "out of stock" (se così posso definirli).

Mi occupo di aste on line, quindi per me le pagine prodotto sono quelle delle aste, e mi trovo a dover gestire queste casistiche:

- 1) aste attive (ci si può registrare e partecipare)
- 2) aste chiuse (lotto venduto al miglior offerente, asta terminata)
- 3) aste chiuse temporaneamente (le offerte non hanno raggiunto il prezzo base, l'asta verrà ri-attivata in seguito)

Su ognuna di queste pagine sono presenti link ad "aste correlate".

Attualmente, in tutti e 3 i casi le URL rispondono 200 e sono presenti in sitemap.

Per i casi 2 e 3, ti chiedo se secondo te ha senso mantenere le URL in sitemap oppure è meglio toglierle?

Grazie, a presto!

Rispondi

Giovanni Sacheli



10/04/2020

Ciao Sergio grazie per le domande.

In <u>questa guida</u> consiglio come gestire i vecchi contenuti, che nel tuo caso sono aste. Forse ti può aiutare leggerla.

Per il caso 2: toglierei la pagina dal sito, redirezionandola 301 alla categoria di appartenenza, quindi rimuoverei anche l'URL dalla sitemap.xml.
Per il caso 3: non toccherei nulla dato che la pagina tornerà attiva.

Rispondi

Richiedi un preventivo

LASCIA UN COMMENTO

Il tuo indirizzo email non sarà pubblicato. I campi obbligatori sono contrassegnati *

Commento *

Nome *

Email *

Invia commento

RICHIEDI UN PREVENTIVO

Nome e Cognome	
Telefono	
Total Control	
La tua email aziendale	
Sito Web	
Sono interessato a	
□ Consulenza SEO	
Consulenza Google Ads	
□ Corsi SEO	
☐ Corsi Google Ads	
D. I. J. SEO. DDG	
Budget annuale per SEO e PPC	
Piccole e Medie Aziende: da 10.000€ a 2	
Descrivi la tua richiesta	
☐ Vorrei ricevere la newsletter mensile	
Acconsento al trattamento dei dati personali*	
Quale numero è minore: 3 o 8?	
Invia la richiesta	
* Campo obbligatorio	

Richiedi un preventivo

CLASSIFICA POSIZIONAMENTO MOTORI DI RICERCA

♥ 83 Guida completa alla migrazione SEO di un sito web

♥ 46 web	Ottimizzare i link interni e la struttura del sito
♥ 45 di Goog	<u>PageRank, scopriamo l'algoritmo più famoso</u> gl <u>e</u>
¥ 44	Screaming Frog Guida Completa in Italiano
♥ 42 profess	SEO Audit tecnico: cos'è? Scarica il template sionale
¥ 36	Guida a Microsoft Excel per la SEO
♥ 35	SEO Checklist: i fattori SEO da Ottimizzare
¥ 29	Migliori estensioni SEO per Chrome 2023
¥ 28	Guida completa al Robots meta tag NOINDEX
♥ 25 analisi	<u>Indicizzazione su Google – significato ed</u>

AGGIORNAMENTI IN QUESTA CATEGORIA

Cos'è un Crawler e come funzionano gli spider?

<u>Come impedire a GPT l'accesso ai contenuti del tuo sito</u>

<u>Link pruning: analizzare e correggere i link in ingresso</u> (<u>backlinks)</u>

COMMENTI

Giovanni Sacheli su <u>Analisi dei contenuti - Come</u> <u>eseguire un Content Audit SEO</u>

Carlo Di Somme su <u>Analisi dei contenuti - Come</u> <u>eseguire un Content Audit SEO</u>

Giovanni Sacheli su <u>Guida all'ottimizzazione delle</u> <u>readcrumb per la SEO</u>

nrico su <u>Guida all'ottimizzazione delle Breadcrumb</u> per la SEO

Giovanni Sacheli su <u>Quali sono i migliori strumenti</u> <u>SEO usati dai professionisti?</u>

ARCHIVIO

Seleziona il mese 🔻

Richiedi un preventivo

EVE MILANO BY SEARCUS

PARTNER GOOGLE ADS

Sacheli Giovanni Consulente SEO e PPC

Email: info(at)evemilano.com

La società titolare di questo sito web è <u>Searcus Swiss</u>

<u>Sagl</u>

Agenzia SEO e Google Ads

Searcus Swiss SAGL UID: CHE-180.818.315

Num reg comm: CH-501.4.019.466-4 Amministrazione: Via dei Faggi 4V, Pazzallo - 6912

(CH)

Ufficio: Via Penate 16, Mendrisio - 6850 (CH)

Email: giovanni(at)searcus.ch Mobile: <u>+39 3393668879</u> Tel: <u>+41 (0)79 8643035</u> Soogle Partner

Sei interessato ad una consulenza per la gestione di campagne Google Ads? Mi chiamo Giovanni Sacheli e sono membro del programma di Certificazione Google Ads Partner per garantire ai miei clienti competenze e professionalità.

SEMINARI SISTRIX.COM



Una volta al mese presento Sistrix, lo strumento SEO dei professionisti. Il seminario dura dalle 10:00 alle 13:30. La partecipazione è gratuita. Per informazioni clicca qui.

CONNECT

linkedin 🗗 - discord 🗗

 $Copyright © 2010 - 2023 \ EVE\ Milano\ Consulenza\ SEO\ Consulenza\ SEO\ by Searcus\ Swiss\ Sagl.\ Tutti\ i\ diritti\ riservati\ |\ Privacy\ |\ Sitemap$

Richiedi un preventivo