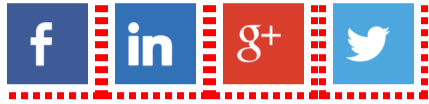


Sei qui: [EVE Milano](#) / [Posizionamento Motori di Ricerca](#) / Screaming Frog Guida Completa in italiano

Screaming Frog Guida Completa in italiano



17/12/2013 SCRITTO DA GIOVANNI SACHELI

29 COMMENTI

Ultimo Aggiornamento: Luglio 2016

Le attività di [analisi SEO](#) sono tantissime e spaziano dalle più tecniche alle più concettuali. Come in ogni cosa non tutto può piacere e anche per gli specialisti SEO ci sono attività che si svolgono più volentieri di altre... Ecco, a me piace molto la SEO tecnica, i tools di analisi, ottimizzare il crawling e gli Audit SEO. Mi rilassa e mi viene relativamente facile cercare **errori nel codice HTML, mancanze tecniche e strutturali che possono rendere la vita difficile a Googlebot**. Questa è una delle attività che considero più divertenti ed appassionanti, una sorta di caccia al tesoro!



Guida Completa a Screaming Frog

Il processo di revisione tecnica di un sito web può essere una attività scoraggiante quando si è alle prime armi poichè **richiede una metodologia molto rigida**, tanto tempo e i giusti strumenti.

La prima attività con cui inizio un Audit SEO è solitamente il crawling del sito web. Ci sono diversi strumenti utili a questo scopo, uno dei più famosi ed utilizzati è senza dubbio [Screaming Frog SEO Spider](#) [↗](#), disponibile per Windows, Mac e Ubuntu. L'interfaccia del tool è molto semplice e user friendly, è veloce da usare e da interpretare. Le numerose opzioni disponibili nel pannello di configurazione tuttavia potrebbero confondere i meno esperti, soprattutto le prime volte che si affrontano crawling *importanti*. Capire il significato di ogni opzione di Screaming Frog è **indispensabile per fare dei crawling corretti**, ogni sito ha le sue caratteristiche e peculiarità e tramite le opzioni dello spider si deve adattare il tool all'ambiente che andrà a visitare per permettergli di fare scansioni corrette.

“ Scegliere una opzione piuttosto che un'altra può **cambiare radicalmente il risultato del crawling**.

Ho scritto questa **Guida Completa di Screaming Frog** usando come base di partenza quella pubblicata da [SEER Interactive](#) [↗](#), ho quindi aggiunto procedimenti testati personalmente ed altre informazioni ottenute con l'uso quotidiano del tool. In futuro mi tornerà utile per ricordarmi tutti i *trucchetti del mestiere* che volta per volta imparo, e spero possa servire anche ad altri webmaster per capire le potenzialità di questo potente crawler per le analisi SEO.

Dato che l'articolo è molto lungo e copre molte funzioni disponibili in Screaming Frog, ho pensato fosse meglio creare un indice per agevolare la navigazione. Per iniziare scegli l'argomento che ti interessa e goditi la guida!

1. [Le Basi del Crawling con Screaming Frog](#)

- 1.1 [I comandi base del crawler](#)
- 1.2 [Avviare il crawling dell'intero sito web](#)
- 1.3 [Come fare crawling di una singola sotto-directory](#)
- 1.4 [Come fare crawling di sotto-domini o sotto-directory](#)
- 1.5 [Come ottenere una lista di tutte le pagine del sito web](#)
- 1.6 [Come ottenere una lista di tutte le pagine contenute in una specifica sotto-directory](#)
- 1.7 [Come ottenere una lista di domini che i competitors usano per redirezionare traffico verso i loro siti principali](#)
- 1.8 [Come trovare tutti i sotto-domini di un sito e verificare la struttura dei link interni](#)
- 1.9 [Come fare crawling di siti con più di 50.000 pagine](#)
- 1.10 [Come fare crawling di un sito ospitato su un server lento o vecchio](#)
- 1.11 [Come fare crawling di un sito che richiede i cookies](#)
- 1.12 [Come fare crawling attraverso un server proxy](#)
- 1.13 [Come fare crawling di pagine che richiedono l'autenticazione](#)

2. [Analizzare i link interni](#)

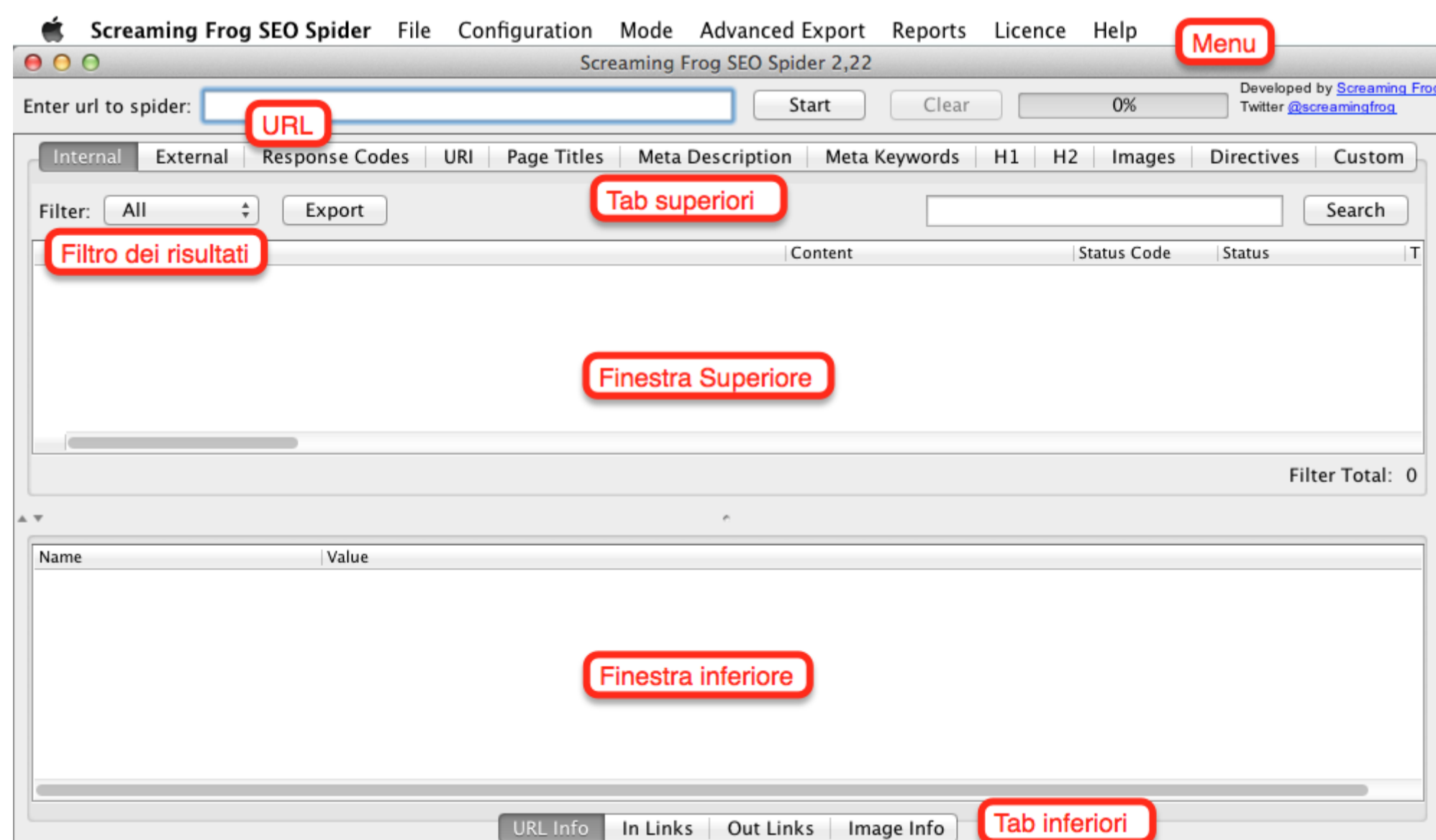
- 2.1 [Come raccogliere informazioni sui link interni ed esterni del mio sito \(anchor text, direttive, links etc.\)](#)
- 2.2 [Come trovare i link interni rotti su una pagina o su un sito](#)
- 2.3 [Come trovare tutti i link in uscita rotti](#)
- 2.4 [Come trovare link che vengono redirezionati](#)
- 2.5 [Come trovare opportunità di linking interno](#)
- 2.6 [Come fare crawling solo degli URL canonici evitando le pagine parametrizzate \(es: "?replyto=..." di WordPress\)](#)

3. [Analizzare i contenuti del Sito](#)

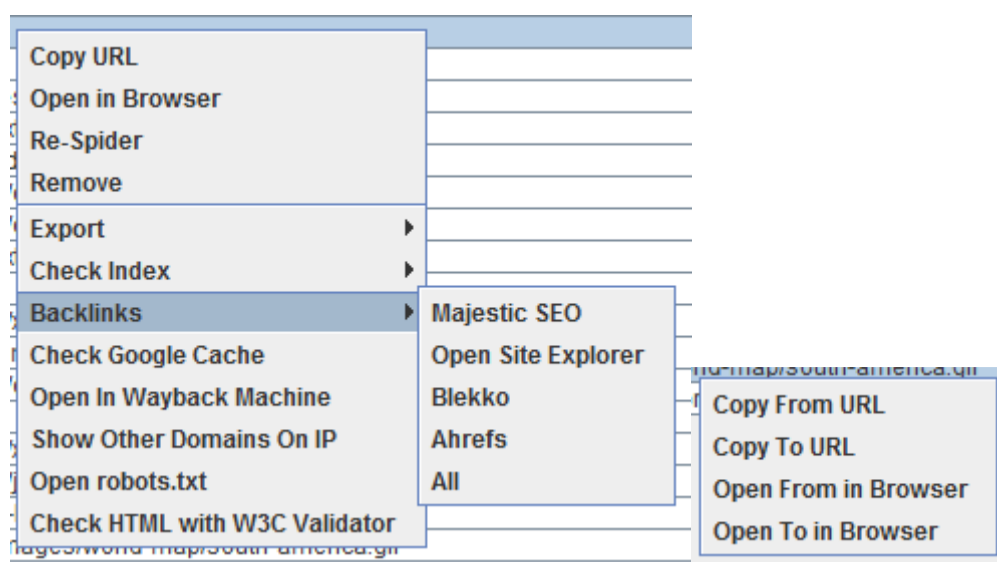
- 3.1 [Come trovare pagine che hanno poco contenuto](#)
- 3.2 [Come ottenere una lista dei link interni provenienti da immagini](#)
- 3.3 [Come trovare immagini a cui manca l'alt tag oppure con alt tag troppo lunghe](#)
- 3.4 [Come trovare tutti i file CSS usati sul sito](#)
- 3.5 [Come trovare tutti i file JavaScript usati sul sito](#)
- 3.6 [Come identificare tutti i plugin jQuery usati sul sito e in quali pagine vengono usati](#)
- 3.7 [Come trovare file in Flash](#)
- 3.8 [Come trovare tutti i file PDF contenuti nel sito](#)

- 3.9 [Come capire la suddivisione dei contenuti di un sito o di un gruppo di pagine](#)
- 3.10 [Come trovare le pagine che usano i pulsanti social](#)
- 3.11 [Come trovare pagine che usano iframes](#)
- 3.12 [Come trovare pagine che contengono video o audio in embed](#)
- 4. **Analizzare i Meta Dati e le Direttive dei crawler**
- 4.1 [Come identificare le pagine con title tag, meta descriptions o URLs troppo lunghi](#)
- 4.2 [Come trovare title tags, meta description oppure URLs duplicati](#)
- 4.3 [Come trovare contenuti duplicati e/o URLs che hanno bisogno di essere riscritti/redirezionati o canonicalizzati](#)
- 4.4 [Come identificare tutte le pagine che contengono direttive dei crawler \(nofollow/noindex/noodp/canonical/etc.\)](#)
- 4.5 [Come verificare che il robots.txt funzioni secondo quanto desiderato](#)
- 4.6 [Come trovare o verificare tutti i markup di Schema.org o altri microdati sul sito](#)
- 5. **Funzioni per Sitemap XML**
- 5.1 [Come creare una sitemap XML](#)
- 5.2 [Come controllare la mia Sitemap.xml esistente](#)
- 6. **Problematiche Generali**
- 6.1 [Individuare i motivi per cui certe pagine non vengono indicizzate](#)
- 6.2 [Come verificare che la migrazione o il redesign del sito sia andata a buon fine](#)
- 6.3 [Come trovare pagine che hanno un tempo di caricamento troppo alto](#)
- 6.4 [Come individuare malware o spam sul sito](#)
- 6.5 [Come aumentare la memoria allocata a Screaming Frog](#)
- 7. **PPC & Google Analytics**
- 7.1 [Come verificare che il codice Analytics sia presente in tutte le pagine o in alcune pagine specifiche](#)
- 7.2 [Come validare una lista di landing page per il PPC](#)
- 8. **Scraping di contenuti**
- 8.1 [Come fare scraping di tutti i meta dati presenti in una lista di pagine](#)
- 8.2 [Come fare scraping di tutte le pagine di un sito che contengono uno specifico footprint](#)
- 9. **Riscrittura degli URL**
- 9.1 [Come trovare e rimuovere parametri e session id dagli URL raccolti da Screaming Frog](#)
- 9.2 [Come fare rewriting degli URL scansionati, ad esempio sostituendo .com con .co.uk, oppure riscrivere tutti gli URL in maiuscolo](#)
- 10. **Analisi delle parole chiave**
- 10.1 [Come sapere quali sono le pagine più importanti dei siti dei competitor](#)
- 10.2 [Come sapere quali anchor text usano i competitors nei link interni](#)
- 10.3 [Come trovare le meta keywords dei competitor \(se usate\)](#)
- 10.4 [Fare scraping della pagina dei risultati di Google](#)
- 11. **Funzioni per la Link Building**
- 11.1 [Come trovare opportunità per inserire link](#)
- 11.2 [Come trovare link non funzionanti](#)
- 11.3 [Come verificare i backlinks e anchor text](#)
- 11.4 [Come verificare l'appartenenza a network di link](#)
- 11.5 [Come verificare la rimozione di backlink](#)
- 11.6 [Come trovare i backlink dofollow tra quelli segnalati da Google Webmaster Tools](#)
- 12. **Estrazioni Customizzate**
- 12.1 [Autori e commenti](#)
- 12.3 [Google Analytics ID](#)
- 12.4 [Heading H1-H6](#)
- 12.5 [Annotazioni Mobile](#)
- 12.6 [Tag Alternate Hreflang](#)
- 12.7 [Social Meta Tags: Open Graph e Twitter Card](#)
- 12.8 [Schema.org](#)
- 12.8 [Indirizzi Email](#)
- 12.9 [Prerender](#)
- 13. **Collegare Google Analytics**
- 13.1 [Come collegare Google Analytics](#)

1. Le Basi del Crawling con Screaming Frog



1.1 I comandi base



Cliccando con il tasto destro su qualsiasi URL nella finestra superiore dei risultati, potrai eseguire le seguenti azioni:

- Copiare oppure aprire l'URL
- Fare di nuovo crawling dell'URL oppure rimuoverlo dalla lista
- Esportare le informazioni di un URL specifico, i link in entrata, in uscita oppure le informazioni dell'immagine.
- Controllare l'indicizzazione della pagina in Google, Bing e Yahoo
- Controllare i backlinks della pagina con Majestic, OSE, Ahrefs e Blekko
- Aprire la pagina nella cache di Google per verificarne la versione e la data
- Vedere le vecchie versioni della pagina salvate da archivi online
- Validare il codice HTML della pagina
- Aprire il file robots.txt relativo al sito che si sta analizzando
- Cercare altri domini ospitati sullo stesso IP

Nello stesso modo, nella finestra inferiore, cliccando con il pulsante destro potrai:

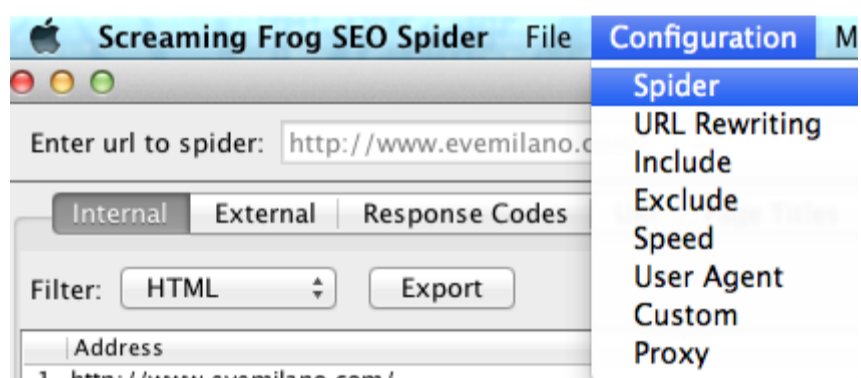
- Copiare oppure aprire l'URL nelle colonne "To" e "From" per la riga selezionata.

1.2 Avviare il crawling dell'intero sito web

Come impostazione di base, Screaming Frog esegue il crawling solo del sotto-dominio che inserisci. Qualsiasi sotto-dominio aggiuntivo che lo spider potrebbe incontrare sarà trattato come un link esterno. Per abilitare lo spider a fare crawling di tutti i sotto-domini, devi entrare nel pannello di amministrazione e cambiare una opzione della sezione relativa alla configurazione dello spider.

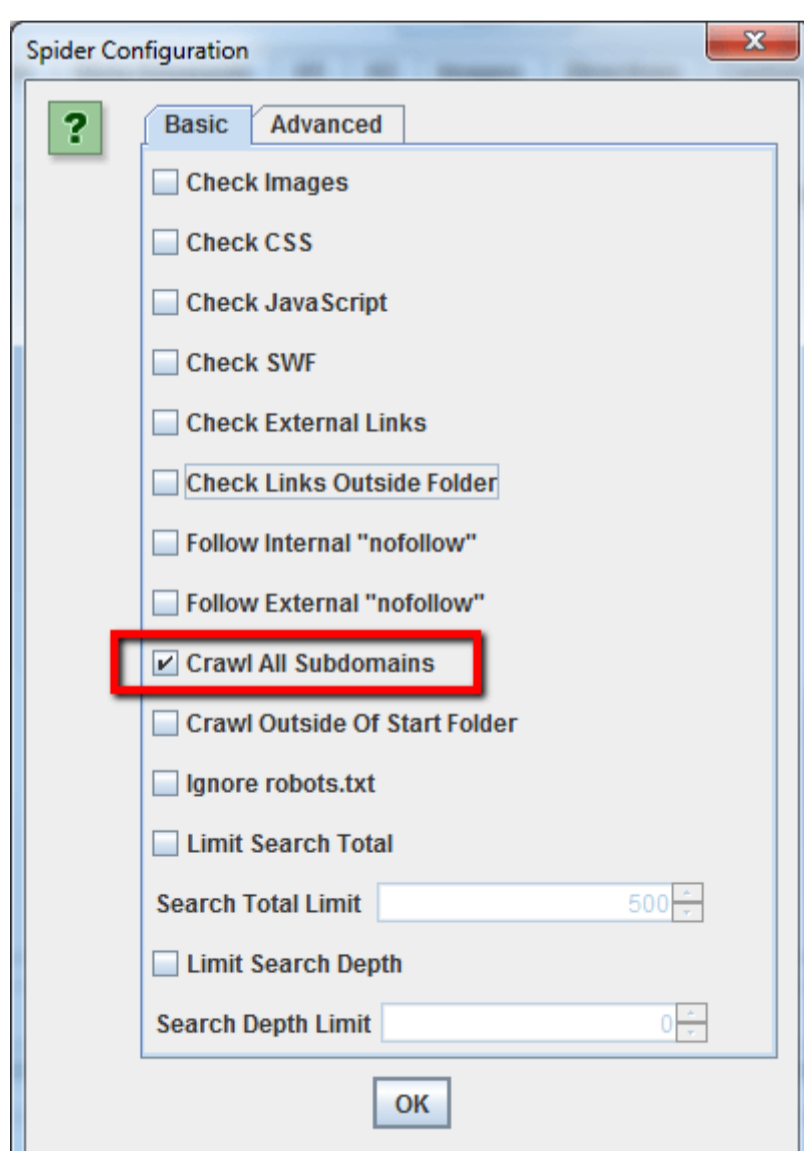
Selezionando **"Crawl All Subdomains"** lo spider farà il crawling anche di tutte le pagine, linkate internamente, appartenenti a sotto-domini.

Primo Step:



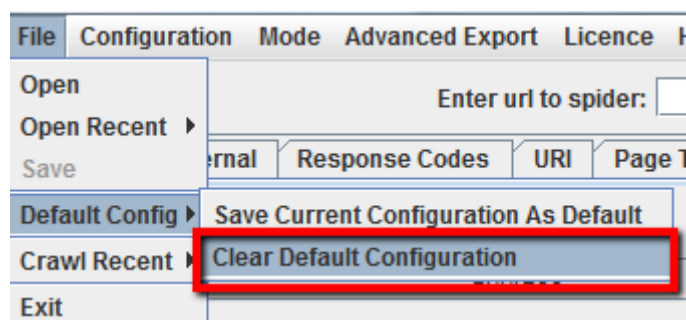
Secondo Step:

Per velocizzare il processo di crawling ti consiglio di **non fare scansionare a Screaming Frog le immagini, CSS, JavaScript, SWF, oppure link esterni**. Puoi modificare le opzioni nel pannello di controllo dello spider.

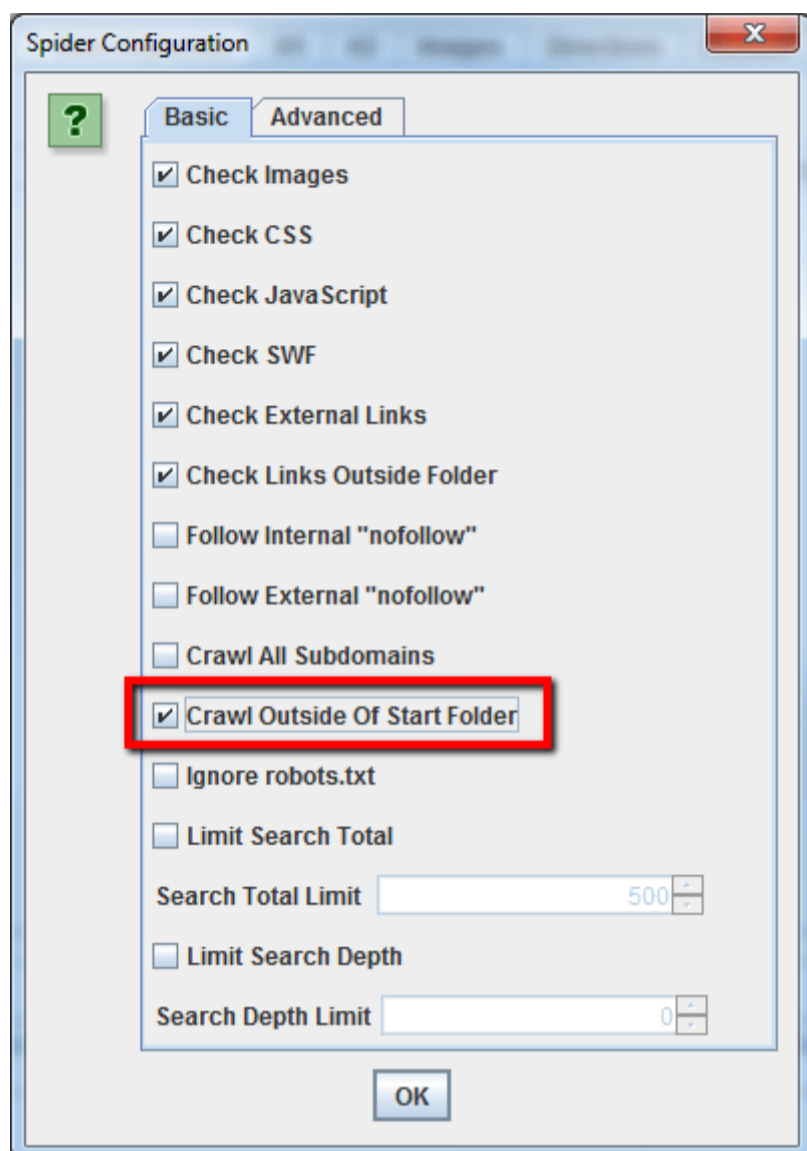


1.3 Come fare crawling di una singola sotto-directory

Se vuoi limitare l'azione dello spider ad una singola cartella (sub-folder), ti basterà inserire l'URL esatto del sub-folder e premere Start senza cambiare nessuna impostazione di default. Se in precedenza hai sovrascritto le impostazioni di base ti basterà resettare la configurazione di default tramite l'opzione dedicata nel menu "File", oppure spunta tutte le opzioni in modo da non avere niente selezionato nella configurazione dello spider.



Se preferisci iniziare a fare crawling dentro una cartella specifica, ma vuoi recuperare anche le altre pagine del sotto-dominio, assicurati di aver selezionato l'opzione **"Crawl Outside Of Start Folder"** nel pannello di configurazione dello spider, prima di inserire l'URL di partenza.



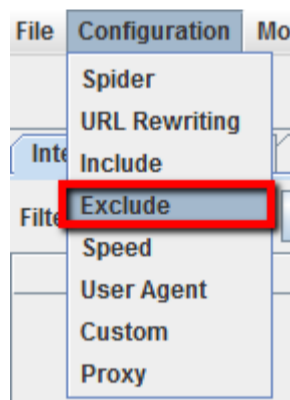
1.4 Come fare crawling di specifici sotto-domini o sotto-directories

Se vuoi limitare il crawling a specifici sotto-domini o sotto-directories puoi usare il comando **RegEx**. Questo comando si usa per impostare le regole nella sezione "Include/Exclude" del pannello di configurazione dello spider.

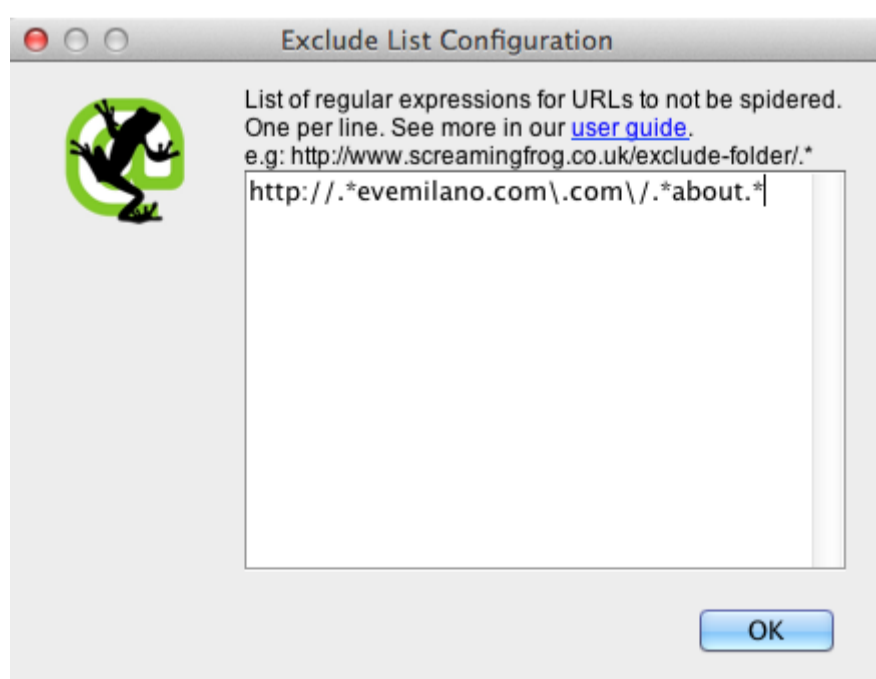
Esclusione:

In questo esempio vogliamo fare crawling di tutte le pagine del sito tranne la pagina "about", in tutti i sotto-domini.

Step 1:

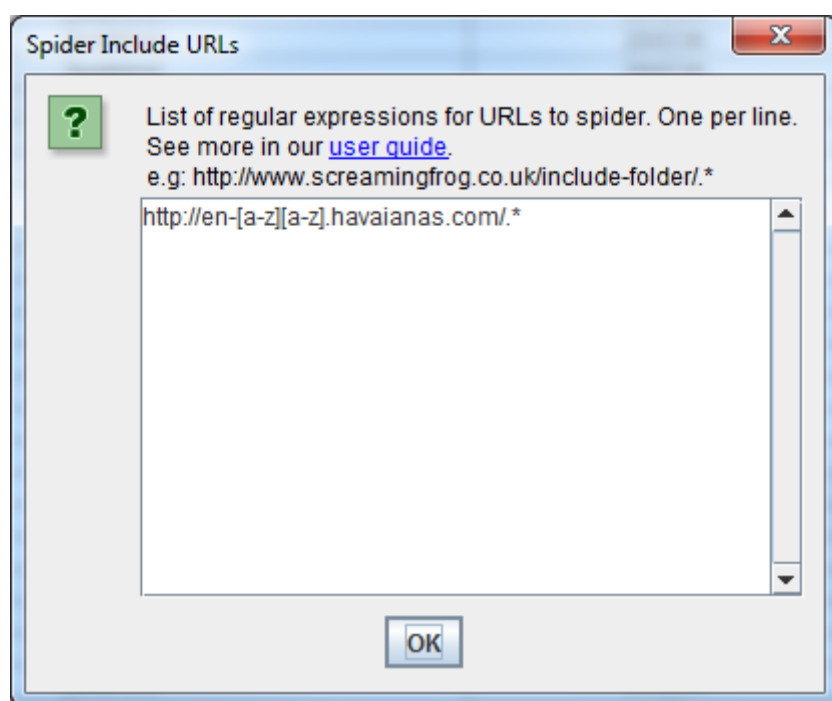


Step 2:



Inclusione:

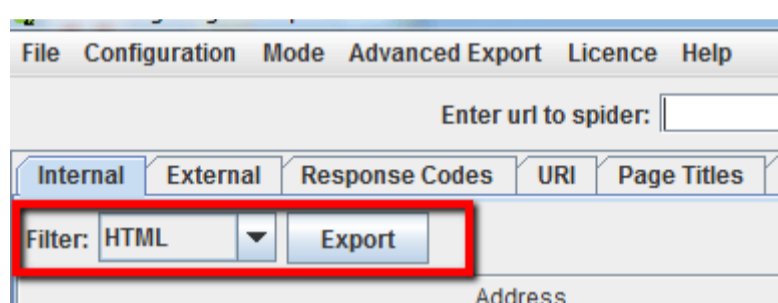
Una situazione opposta a quella presentata precedentemente: l'inclusione, ovvero fare crawling soltanto di certe pagine. In questo esempio vogliamo fare crawling **soltanto** del sotto-dominio in lingua Inglese del sito havaianas.com.



1.5 Voglio una lista di tutte le pagine del mio sito

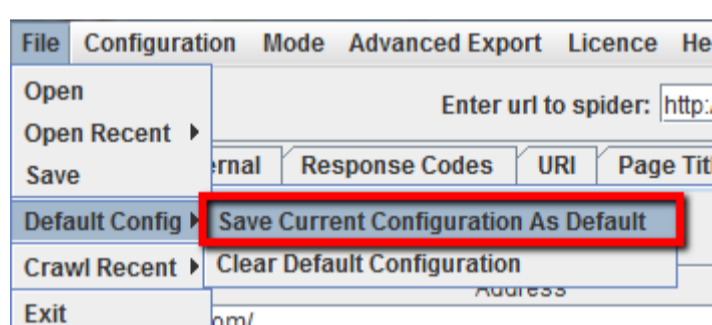
Come impostazione di base, Screaming Frog fa crawling di tutte le immagini, JavaScript, CSS e file flash che incontra. Per raccogliere soltanto le pagine HTML **devi deselezionare** "Check Images", "Check CSS", "Check JavaScript" e "Check SWF" nella configurazione dello spider.

Utilizzare lo spider in questo modo permette di creare una lista di tutte le pagine (URLs) del sito linkate internamente. Una volta terminato il processo di crawling spostati nella tab *Internal* e filtra i risultati per "HTML". Clicca quindi "Export" e otterrai la lista completa delle pagine del sito in formato .CSV.



Consiglio:

Se di abitudine usi il crawler con opzioni customizzate, puoi salvare la tua configurazione dello spider:



1.6 Voglio una lista di tutte le pagine in una specifica sotto-directory

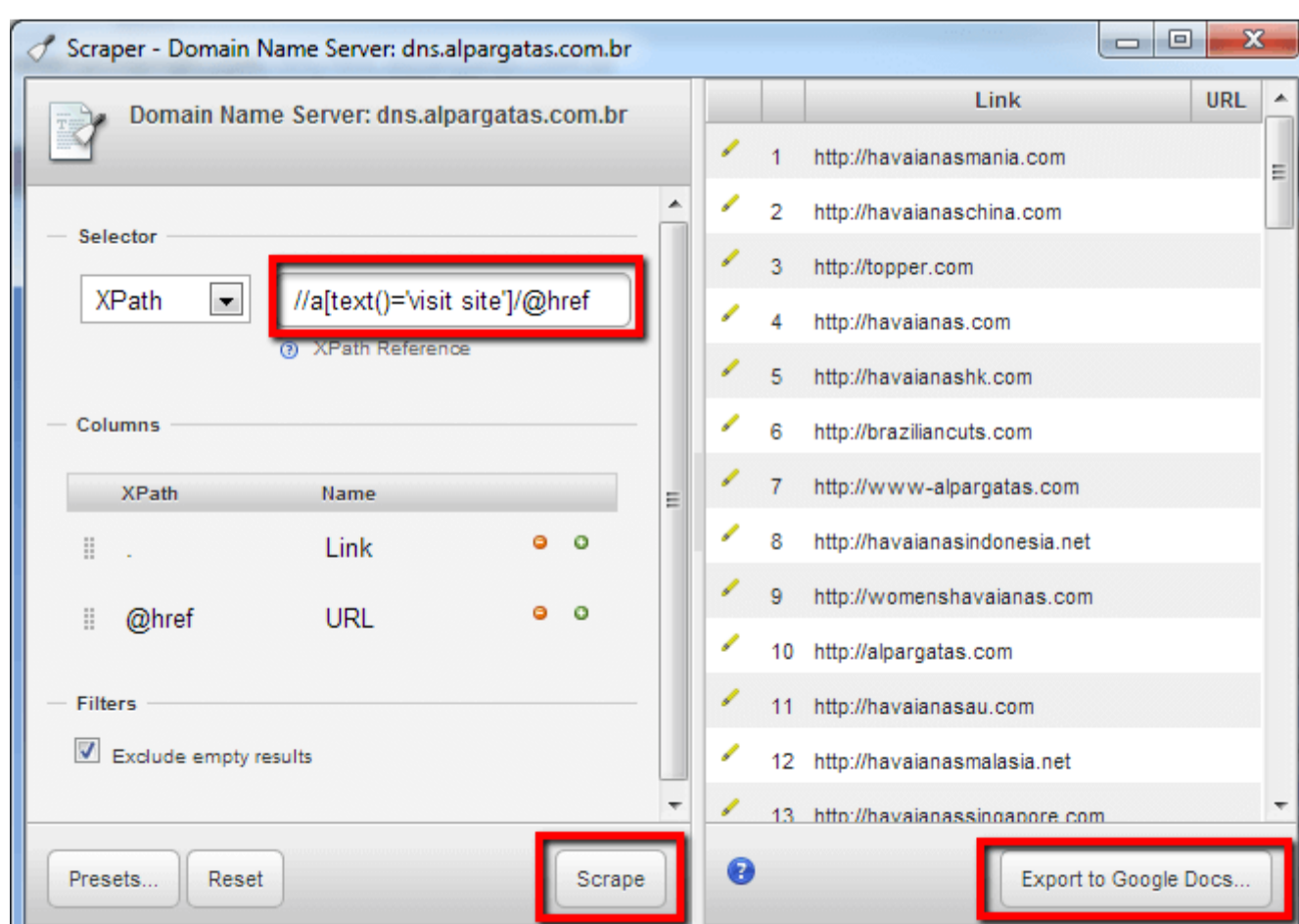
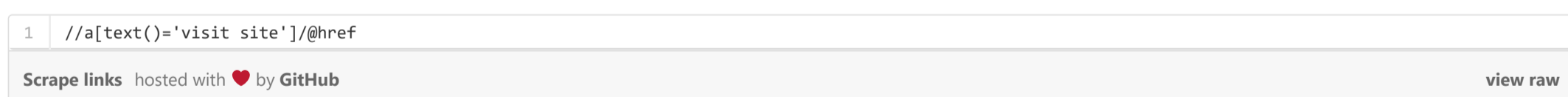
Oltre a deselezionare "Check Images", "Check CSS", "Check JavaScript" e "Check SWF", dovrai anche deselezionare "**Check Links Outside Folder**" nel pannello di configurazione dello spider.

Lanciando Screaming Frog con queste impostazioni otterrai una lista di pagine contenute soltanto nella sotto-directory. Le pagine che non ottengono link interni non saranno incluse nella lista.

1.7 Come trovare una lista di domini che un competitor redireziona sul suo sito principale

Inserisci il sito principale (quello dove avvengono le conversioni) del competitor che vuoi analizzare nel tool [ReverseInternet](#), poi clicca i link nella parte superiore della tabella per trovare altri siti che condividono lo stesso indirizzo IP, il *nameserver* oppure il **codice di tracciamento Analytics**.

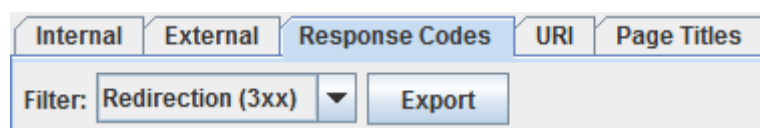
Usando l'estensione per Chrome [Scraper](#) puoi ottenere la lista di URL che usano specifiche anchor text direttamente da ReverseInternet. Se hai già installato Scraper ti basterà fare click con il pulsante destro sopra la sezione di domini da esportare e scegliere "Scrape similar". Nella finestra di pop-up dovrai cambiare la query XPath per estrarre solo i link, come mostrato di seguito:



A questo punto premi "Next", poi "Scrape" ed in fine "Export to Google Docs". Da Google Doc ora puoi scaricare il file .CSV.

Fai l'upload del file .CSV in Screaming Frog, seleziona la modalità a lista "List mode" per verificare la lista di URLs.

Quando lo spider avrà terminato, vedrai lo status code nella tab *Internal*, oppure puoi usare la tab *Response Codes* e filtrare per "Redirection" per vedere tutti i domini che attualmente stanno redirezionando al sito principale del tuo competitor.



NB: Quando carichi il file .CSV in Screaming Frog, ricordati di selezionare "CSV" come tipologia di file, altrimenti il programma si chiuderà con un errore.

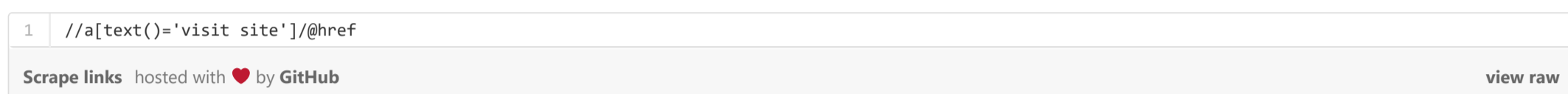
Consiglio:

Puoi usare questo metodo anche per identificare i domini appartenenti ai tuoi competitors e soprattutto per capire l'uso che ne stanno facendo.

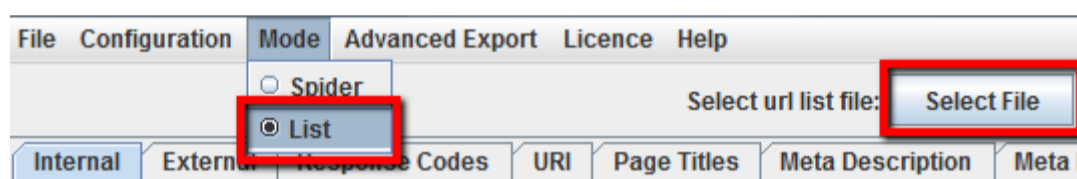
1.8 Come trovare tutti i sotto-domini di un sito e verificare i link interni

Inserisci l'URL della root nel tool [ReverseInternet](#), poi clicca sulla tab *Subdomains* per vedere la lista dei sotto-domini.

Ora usando [Scrape Similar](#) puoi raccogliere la lista di URLs tramite la query XPath:



Esporta i risultati in .CSV e carica il file in Screaming Frog usando la "List mode". Lancia il crawling, quando lo spider avrà finito potrai verificare gli status codes, tutti i link nella homepages del sotto-dominio, anchor texts e la tag title duplicate (e altre info).



1.9 Come fare crawling di un eCommerce oppure di sito di grandi dimensioni, con molte pagine

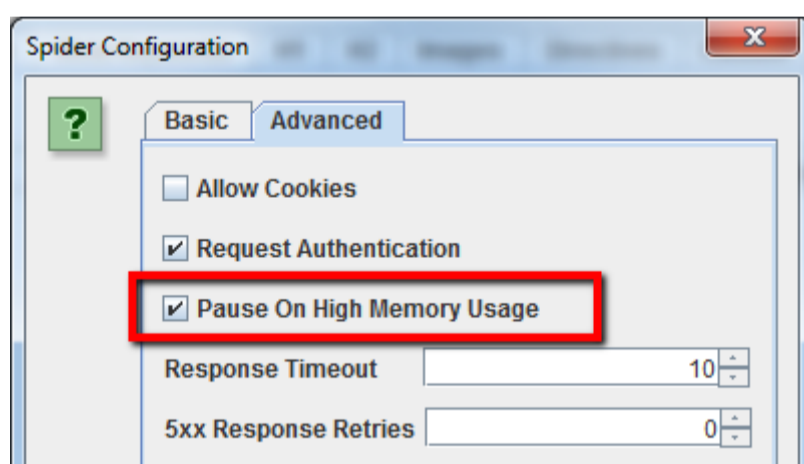
Screaming Frog non è studiato per fare crawling di centinaia di migliaia di pagine a causa di limiti di memoria fisica del computer sul quale viene eseguito, in particolare mi riferisco ai limiti di memoria per applicativi in Java. Tuttavia ci sono degli accorgimenti da prendere per evitare che il programma si blocchi quando fa crawling di grossi siti, ovvero con più di 50.000 pagine.

Come prima cosa puoi [aumentare l'allocazione della memoria](#) per lo spider. Come seconda opzione potresti spezzettare il sito in sotto-directories a fare crawling di una singola sezione per volta. Come ultimo accorgimento potresti dire allo spider di non scansionare immagini, JavaScript, CSS e flash in modo da salvare memoria da dedicare alle sole pagine HTML.

Consiglio:

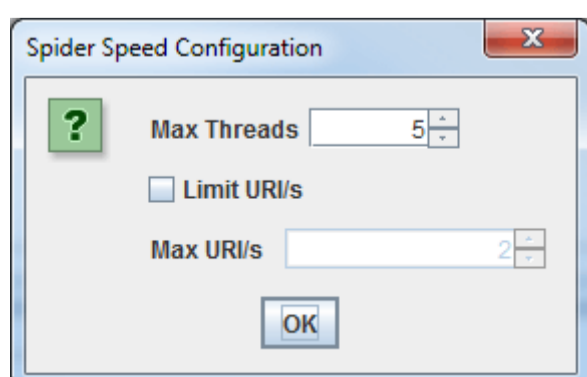
Con versioni precedenti di Screaming Frog poteva capitare che il crawler andava in time-out su siti molto grossi. Dalla versione 2.11 puoi dire al programma di mettersi in pausa quando la memoria si sta per esaurire. Queste impostazioni di sicurezza evitano che Screaming Frog vada in crash prima di aver salvato oppure prima di aver aumentato la memoria allocata.

Queste sono le impostazioni di default, se stai per iniziare un crawling molto esteso ricordati di attivare "**Pause On High Memory Usage**" nelle impostazioni avanzate dello spider.



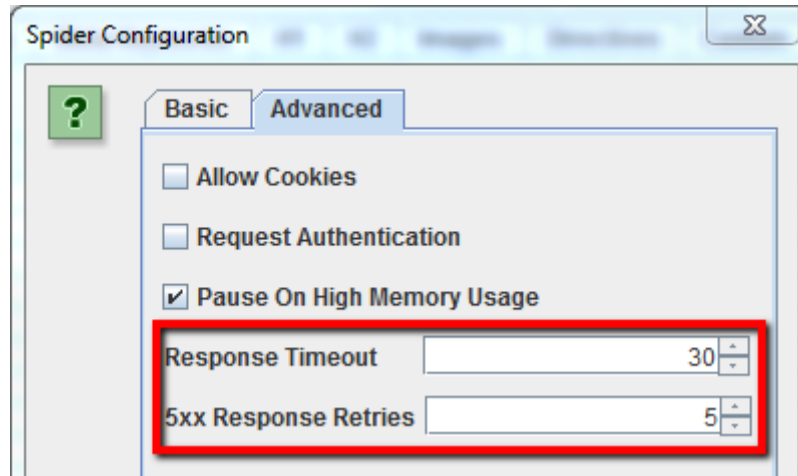
1.10 Come fare crawling di un sito ospitato su un vecchio server

In alcuni casi, i vecchi server potrebbero non essere in grado di gestire le numerose richieste simultanee di URLs di Screaming Frog. Per ridurre la velocità di crawling dello spider basta andare nella sezione "Speed" delle opzioni e scegliere un valore inferiore cercando di trovare il giusto equilibrio con il web server. Ovviamente da questo menu puoi anche aumentare il numero di richieste per **velocizzare il crawling su server particolarmente potenti**.



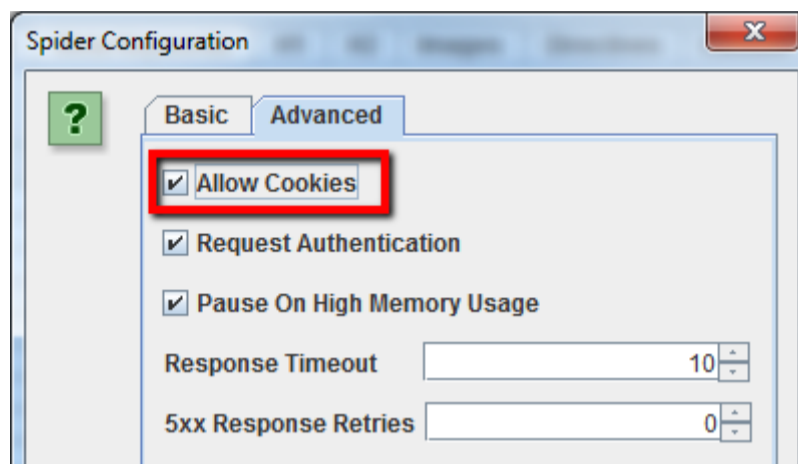
Consiglio:

Se noti che Screaming Frog restituisce troppi falsi errori del server (ovvero pagine in errore che testate una seconda volta restituiscono codice 200), vai nelle opzioni "Avanzate" dello spider e aumenta il "Response Timeout" e le "5xx Response Retries".



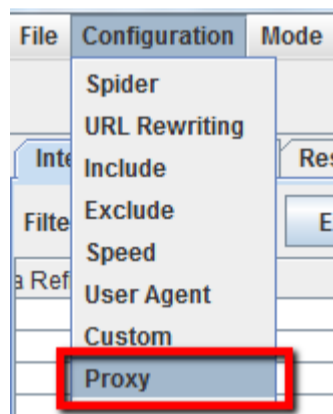
1.11 Come fare crawling di siti che richiedono i cookies

Sebbene gli spiders dei motori di ricerca non accettano i cookies di default, se stai facendo crawling di un sito che richiede necessariamente i cookies ti basterà selezionare "Allow Cookies" nella tab *Advanced* delle opzioni dello spider.

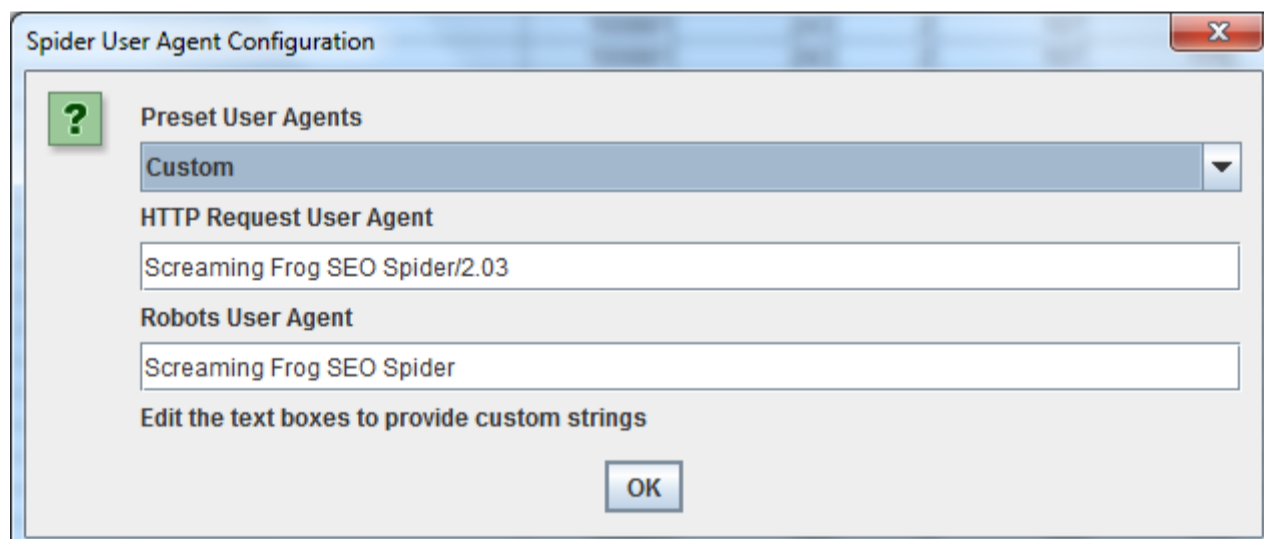


1.12 Come fare crawling usando un proxy oppure un user-agent differente

Per fare crawling usando un proxy, seleziona "Proxy" nel menu di configurazione ed inserisci le informazioni del server Proxy.



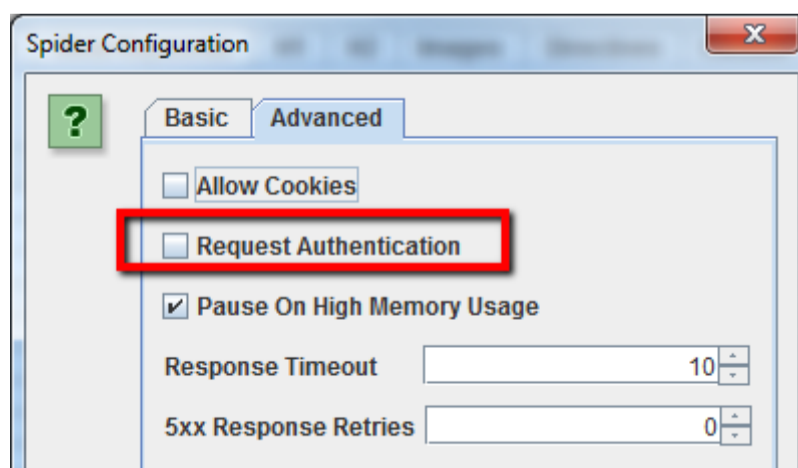
Per fare crawling usando un user-agent differente seleziona "User Agent" nel menu di configurazione, scegli il bot di ricerca dal menu a tendina oppure inserisci l'user-agent desiderato.



1.13 Come fare crawling di pagine protette da password

Quando lo spider di Screaming Frog trova pagine protette da password, verrà mostrato un pop-up nel quale potrai inserire il nome utente e la password per accedere ai contenuti protetti.

Per disattivare il pop-up ti basta togliere la spunta da "Request Authentication" nella tab *Advanced* del pannello di controllo dello spider.



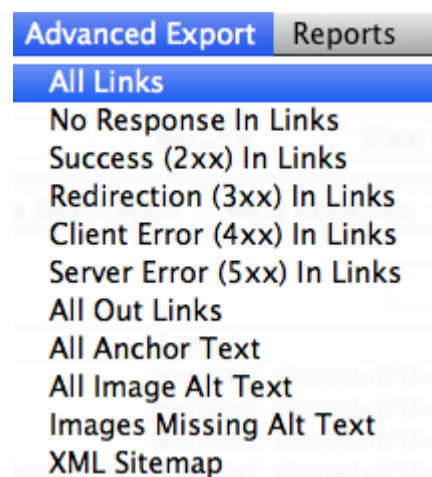
2. Analizzare i Link Interni

2.1 Come raccogliere informazioni sui link interni ed esterni del mio sito (anchor text, direttive, links per pagina etc.)

Se non hai bisogno di verificare le immagini, JavaScript, flash o CSS sul sito, deseleziona queste opzioni nel pannello di controllo dello spider per salvare tempo e memoria.

Una volta che lo spider avrà finito di fare crawling, usa la voce "Advanced Export" del menu per esportare tutti i link "All Links" in formato .CSV.

Il file conterrà tutti gli URL dei links, l'anchor text, le direttive etc.



Per trovare velocemente il numero di link presenti in ogni pagina vai nella tab *Internal* e ordina i dati della colonna "Outlinks".

Consiglio: Una vecchia linea guida di Google diceva di non superare i 100 link per pagina, ora sappiamo che quel limite è stato superato ma è sempre buona norma non esagerare.

Size	Word Count	Level	Inlinks	Outlinks	External Outlinks
65692	499	1	4758	82	17
65658	476	2	4203	82	17
66045	468	2	4203	82	17
65953	509	2	4203	82	17
65590	492	2	4203	82	17
66018	517	2	4203	82	17
0	0	4	2874	0	0
76064	3618	2	2319	80	20
130829	3967	2	1860	153	20
119592	3917	2	1766	127	20
125498	4011	2	1667	147	20
116358	3812	2	1655	136	20
111778	3784	2	1608	121	20
115761	3785	2	1606	131	20
106976	4154	2	1553	82	20
63731	1154	2	1551	80	14
52365	588	2	1550	80	14

Conteggio dei link interni

2.2 Come trovare link interni rotti (broken internal links) su una pagina o un sito

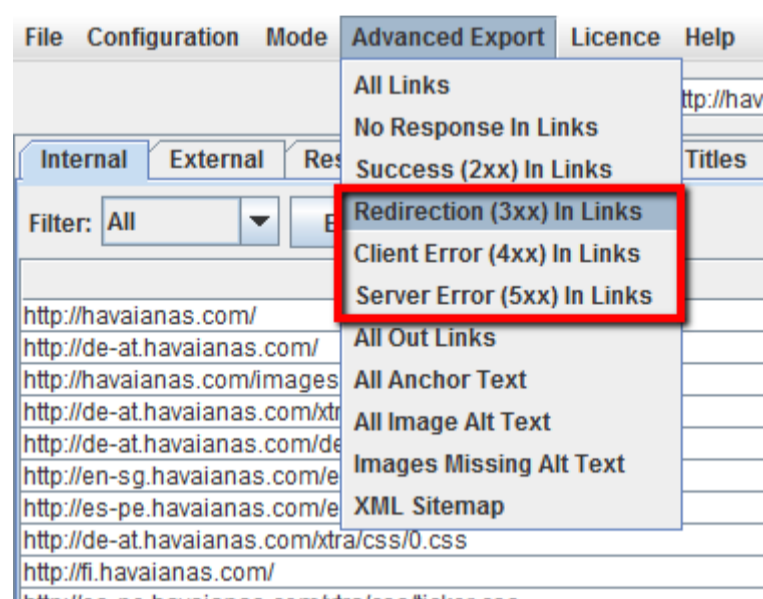
Se non hai bisogno di verificare le immagini, JavaScript, flash o CSS sul sito, deseleziona queste opzioni nel pannello di controllo dello spider per salvare tempo e memoria.

Una volta che lo spider avrà finito di fare crawling, ordina i dati della colonna "Status Code" nella tabella "Internal". Tutti gli errori 404, 301 oppure altri Status Code possono essere facilmente individuati e filtrati.

Puoi visualizzare ulteriori informazioni sui link interni nella parte bassa dello schermo, semplicemente cliccando su qualsiasi URL nella lista dei risultati del crawling nella finestra in alto dello schermo.

Dopo aver selezionato un URL, puoi cliccare sulla tab *In Links* per trovare la lista di pagine che linkano all'url selezionato, le anchor text e le direttive di quei links. Questa funzionalità è molto utile per trovare pagine che hanno link interni che devono essere aggiornati.

Per esportare in .CSV la lista completa delle pagine che includono link rotti o redirezionati, scegli "Redirection (3xx) In Links" oppure "Client Error (4xx) In Links" oppure "Server Error (5xx) In Links" nella voce del menu "Advanced Export".



Alcuni webmaster utilizzano i **link relativi** all'interno del loro sito, questa svista se accompagnata ad un errore errore comune quello di inserire i link senza prefisso HTTP può generare **catene di link infiniti** in grado di consumare il **crawl budget dei motori di ricerca** inutilmente. Per fortuna questo non è un grande problema perchè con Screaming Frog è possibile individuare e correggere questi link in un attimo. Per filtrare i link errati basta cliccare sul menu di configurazione del crawler ed impostare l'inclusione dei soli caratteri:

```
1 <href="www.
```

Relative link fix hosted with ❤ by GitHub [view raw](#)

I risultati del crawling mostreranno tutte le pagine che linkano con la sintassi errata.

2.3 Come trovare link esterni rotti (broken outbound links) su una pagina o un sito ed in genere tutti i link che puntano fuori

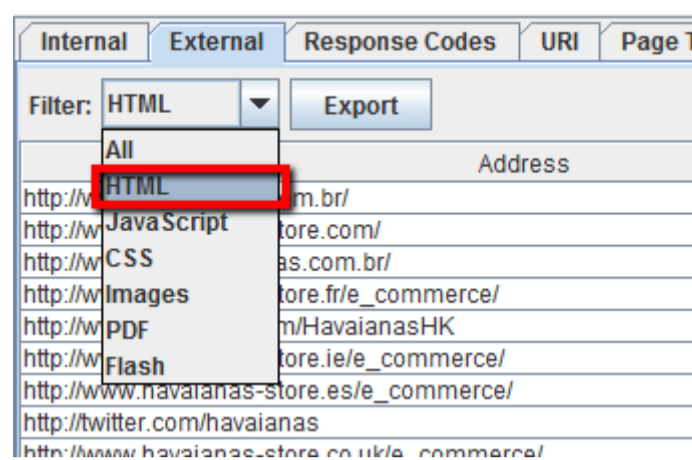
Dopo aver deselezionato "Check Images", "Check CSS", "Check JavaScript" e "Check SWF" nel pannello di configurazione dello spider, assicurati di **selezionare la voce "Check External Links"**.

Una volta che lo spider avrà finito di fare crawling, clicca sulla tab *External* nella finestra superiore, ordina i risultati per "Status Code" e potrai trovare velocemente tutti gli URL con Status Code diverso da 200. Cliccando su qualsiasi URL nei risultati del crawling e successivamente sulla tab *In Links*

nella finestra inferiore, troverai la lista di pagine che puntano a quello specifico URL che hai selezionato. Puoi usare questa funzione per identificare pagine che richiedono un aggiornamento dei link esterni.

Per esportare la lista completa degli outbound links, clicca su "Export" nella tab *Internal*. Puoi anche impostare un filtro per esportare link che puntano ad immagini esterne, file JavaScript esterni, file CSS esterni, file Flash e PDF esterni.

Per limitare il filtro alle sole pagine seleziona "HTML".



Per una lista completa di tutte le pagine che contengono i link outbound e le relative anchor text, seleziona "All Out Links" dalla voce del menu "Advanced Export" e poi filtra la colonna "Destination" del .CSV in modo da escludere il proprio dominio.

2.4 Come trovare link che vengono redirezionati

Dopo che lo spider ha finito di fare crawling, seleziona la tab *Response Codes* nella finestra superiore di Screaming Frog, poi filtra per "Redirection (3xx)". Questa tab ti fornirà una lista di tutti i link interni ed esterni che vengono redirezionati. Ordina per "Status Code" e potrai distinguere ogni singolo errore e suddividere i redirect per categoria.

Clicca un link qualsiasi nella finestra superiore "Response Codes" e poi naviga nella tab *In Links* nella finestra inferiore per vedere quali sono le pagine che contengono il link che viene redirezionato. Se esporti direttamente da questa tab potrai vedere soltanto le informazioni mostrate nella finestra superiore (Address, Status Code e Redirect URL).

Per esportare la lista completa delle pagine che contengono link che vengono redirezionati, devi selezionare "Redirection (3xx) In Links" nel menu "Advanced Export". Il file esportato includerà le locations di tutti i link redirezionati. Per mostrare solo i redirect interni al sito, filtra la colonna "Destination" nel file .CSV in modo da includere soltanto il tuo dominio.

Consiglio:

Usa la funzione VLOOKUP di Excel tra due files esportati per aggiungere alle colonne "Source" e "Destination" anche la location finale dell'URL.

La formula di esempio è:

```
1 =VLOOKUP([@Destination],'response_codes_redirection_(3xx).csv'!$A$3:$F$50,6,FALSE)
```

Vlookup Excel hosted with ❤️ by GitHub [view raw](#)

- "response_codes_redirection_(3xx).csv" è il file .CSV che contiene gli URL redirezionati
- "50" è il numero di righe presenti in quel file

2.5 Opportunità di linking interno

Ti consiglio di leggere questa guida [Scaling Internal Link Building with Screaming Frog & Majestic](#)

2.6 Voglio scansionare soltanto gli URL canonici evitando tutte le pagine parametrizzate

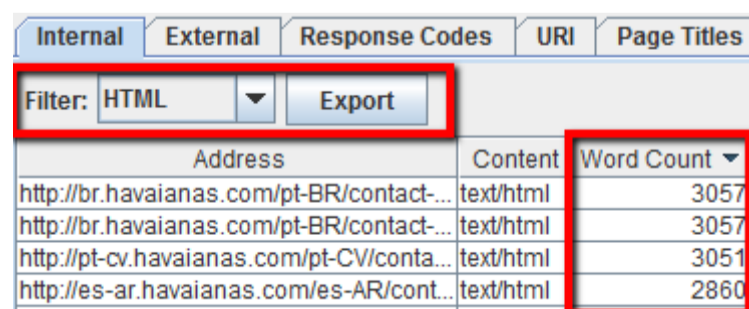
Dal pannello di controllo dello spider accedi alla tab *Advanced* e seleziona "Respect Canonical"

3. Analizzare i contenuti del sito

3.1 Come identificare pagine con poco contenuto

Dopo che lo spider ha finito di fare crawling, vai nella tab *Internal* e filtra solo i file HTML. A questo punto scrolla a destra fino a quando non trovi la colonna "Word Count".

Ordina la colonna dal minore al maggiore per vedere quali sono le pagine con poco testo. Puoi trascinare la colonna a sinistra per affiancarla agli URL. Clicca "Export" nella tab *Internal* se preferisci lavorare su un file .CSV con Excel.



3.2 Voglio la lista di tutte le immagini contenute in alcune pagine

Se hai già fatto crawling di un intero sito o di una sotto-cartella, ti basta selezionare nella finestra superiore la pagina che vuoi analizzare e quindi cliccare sulla tab *Image Info* nella finestra inferiore.

In questa lista puoi vedere tutte le immagini che sono state trovate in quella pagina. Le immagini sono elencate sotto la colonna "To".



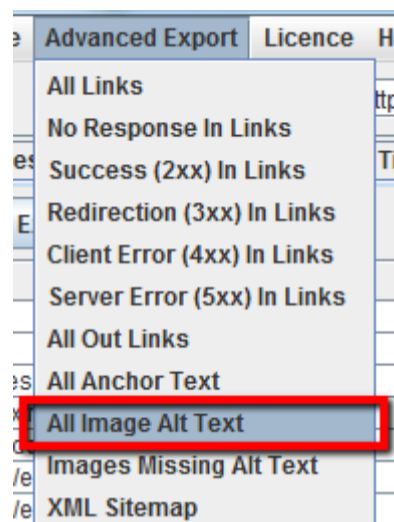
Consiglio:

Clicca con il tasto destro su qualsiasi risultato nella finestra inferiore per copiare o aprire un URL.

In alternativa puoi vedere tutte le immagini contenute in una pagina semplicemente facendo crawling di solo quell'URL. Assicurati di aver impostato il valore di profondità del crawler (Crawl Depth) a **1** nelle impostazioni dello spider.

Una volta che il crawler ha terminato puoi cliccare sulla tab *Images* per vedere l'elenco di tutte le immagini che sono state trovate da Screaming Frog.

Per esportare i dati usa la voce del menu "Advanced Export" e clicca su **"All Image Alt Text"** per avere la lista completa delle immagini e delle rispettive alt tag.

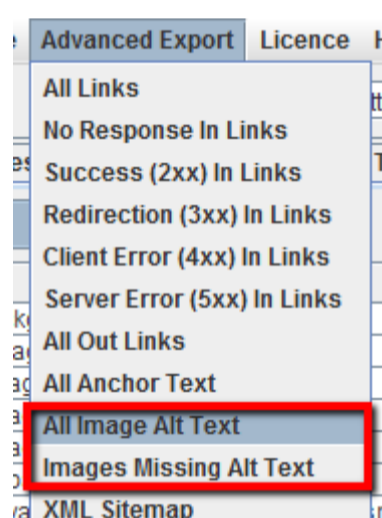


3.3 Come trovare tutte le immagini a cui manca l'alt tag oppure con alt tag con troppo testo

Per prima cosa assicurati di aver selezionato la voce "Check Images" nel pannello di configurazione dello spider. Dopo che lo spider avrà finito di fare crawling vai nella tab *Images* e filtra per "Missing Alt Text" oppure "Alt Text Over 100 Characters". Ora potrai individuare le pagine che contengono le immagini semplicemente cliccando nella tab *Image Info* nella parte inferiore dello schermo. Le pagine saranno elencate nella colonna "From".

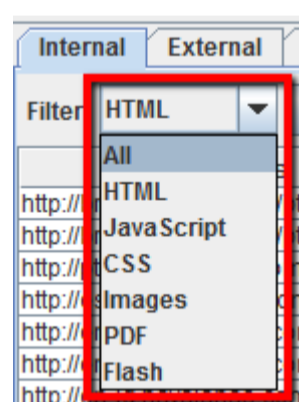
In alternativa, nella voce del menu "Advanced Export", puoi esportare "All Image Alt Text" oppure "Images Missing Alt Text" direttamente in formato .CSV.

Il file scaricato mostrerà la lista di tutte le pagine che contengono le immagini usate nel sito.



3.4 Come trovare tutti i file .CSS nel mio sito

Nel pannello di configurazione dello spider seleziona soltanto "Check CSS" prima di lanciare il crawling. A processo terminato naviga nella tab *Internal* e filtra per "CSS".

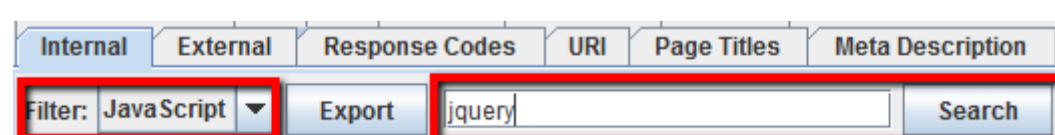


3.5 Come trovare tutti i file JavaScript nel mio sito

Nel pannello di configurazione dello spider seleziona soltanto "Check JavaScript" prima di lanciare il crawling. A processo terminato naviga nella tab *Internal* e filtra per "JavaScript".

3.6 Come identificare tutti i plugin jQuery usati sul sito ed in quali pagine vengono usati

Prima di tutto assicurati di aver selezionato l'opzione "Check JavaScript" nel pannello di configurazione dello spider. Dopo che lo spider ha finito il suo lavoro, filtra solo i files "JavaScript" nella tab *Internal* poi cerca "jQuery". La lista conterrà tutti i files dei plugin, ordina i dati con la colonna "Address" poi apri la tab *InLinks* nella parte inferiore dello schermo oppure esporta in un file .CSV tutte le pagine che includono i plugin (campo della colonna "From").



In alternativa puoi usare la funzione "Advanced Export" ed esportare in .CSV la lista "All Links" e filtrare la colonna "Destination" per mostrare solo gli URL con "jQuery".

Consiglio:

Non tutti i plugin jQuery sono negativi per la SEO. Se vedi un sito che usa jQuery, la cosa migliore da controllare è che il contenuto che ti interessa indicizzare sia incluso nel codice sorgente e mostrato al caricamento della pagina e non successivamente. Se sei insicuro controlla la cache di Google oppure informati sul funzionamento di quello specifico plugin.

3.7 Come trovare i file Flash inseriti nel sito

Nel menu di configurazione dello spider seleziona "Check SWF" prima di iniziare il crawling. A processo terminato vai nella tab *Internal* e filtra per "Flash".

NB: Questo metodo troverà soltanto i file .SWF che sono linkati in una pagina. Se il file viene richiamato con JavaScript, ti servirà un filtro customizzato.

3.8 Come trovare i file PDF inseriti nel sito

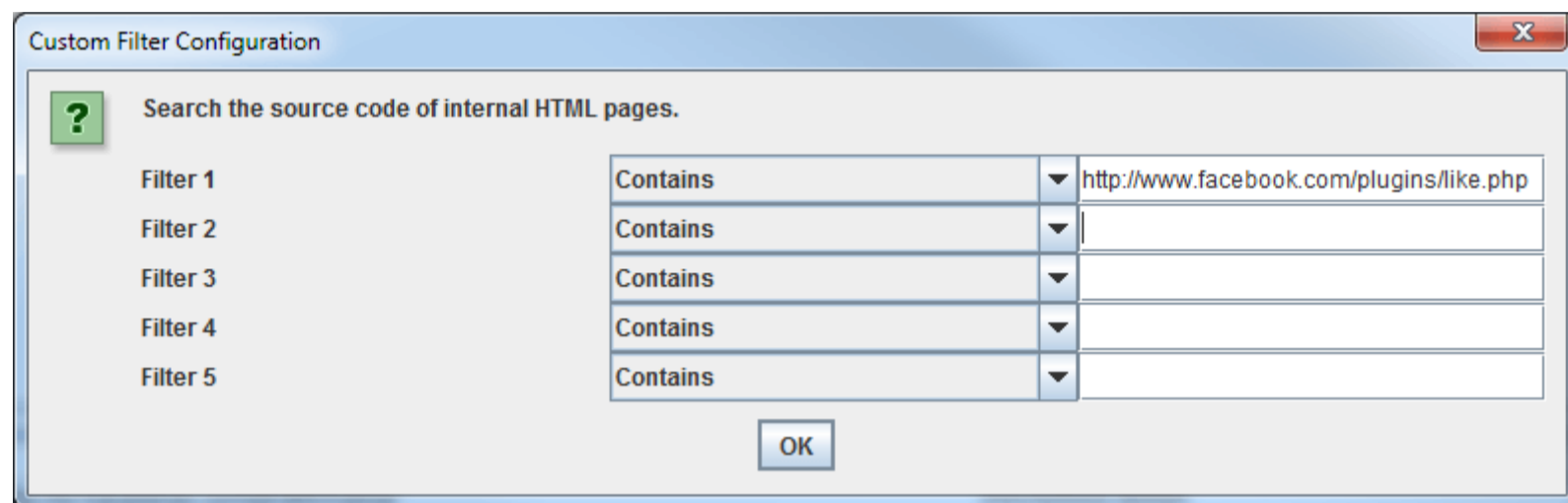
Dopo che lo spider ha finito di fare crawling, filtra i risultati della tab *Internal* selezionando "PDF".

3.9 Come identificare la suddivisione dei contenuti all'interno di un sito o un gruppo di pagine

Se vuoi trovare pagine sul tuo sito che contengono uno specifico contenuto, imposta un filtro customizzato per il footprint HTML che definisce univocamente quelle pagine. L'impostazione deve essere fatta prima di avviare il crawling.

3.10 Come trovare le pagine che includono i pulsanti di condivisione social

Per trovare tutte le pagine che contengono i pulsanti di condivisione social, devi impostare un filtro customizzato prima di avviare il crawler. Per creare un filtro customizzato vai nel menu di configurazione e clicca "Custom". In questa sezione inserisci lo snippet di codice che vuoi ricercare nel codice sorgente delle pagine.



Nell'esempio dello screenshot, si vuole trovare le pagine che contengono il pulsante "Like" di Facebook e per fare ciò filtriamo tutte le pagine che contengono un link verso <http://www.facebook.com/plugins/like.php>.

3.11 Come trovare le pagine che usano iframes

Per trovare tutte le pagine che includono un iframe, imposta un filtro customizzato per il codice `<iframe` prima di avviare lo spider.

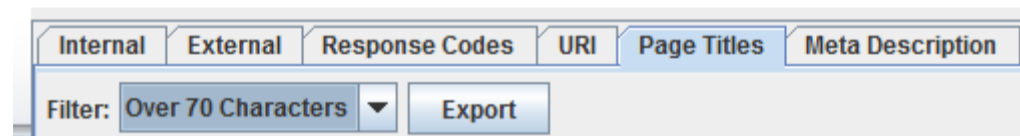
3.12 Come trovare le pagine che contengono video oppure file audio

Per trovare le pagine che contengono video embeddati o file audio, imposta un filtro customizzato con lo snippet di codice usato per embeddare un video YouTube, oppure qualsiasi altro media player che viene usato nel sito.

4. Meta Dati e le Direttive

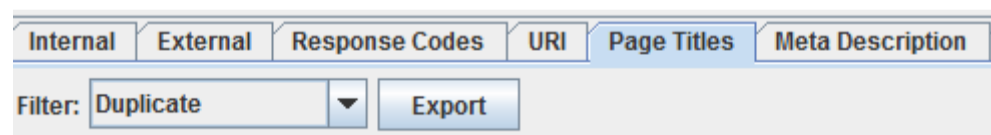
4.1 Come identificare pagine con title tag, meta description oppure URL troppo lunghi e troppo corti

Dopo che lo spider ha finito di fare crawling, vai nella tab *Page Titles* e filtra per "Over 70 Characters" per vedere tutti i title tag che sono troppo lunghi. La stessa cosa può essere fatta nella tab delle "Meta Description" ed in quella degli "URI".



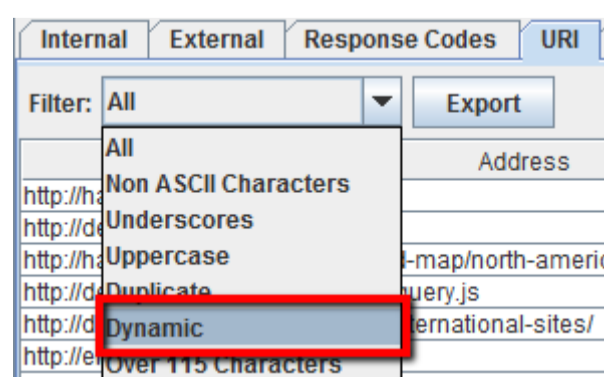
4.2 Come trovare title tag, meta description e URL duplicati

Dopo che lo spider ha finito di fare crawling, vai nella tab *Page Titles* e filtra per "Duplicate". La stessa cosa può essere fatta nelle tab *Meta Description* e "URI".



4.3 Come trovare contenuto duplicato e URL che devono essere riscritti/redirezionati/canonicalizzati

Dopo aver terminato il crawling vai nella tab *URI*, poi filtra per "Underscores", "Uppercase" oppure "Non ASCII Characters" per trovare tutti gli URL che potrebbero venire riscritti in modo più semplice. Filtra per "Duplicate" per trovare tutte le pagine che hanno versioni multiple dello stesso URL. Filtra per "Dynamic" e troverai tutti gli URL che includono parametri.



Inoltre, se vai nella colonna "Hash" nella tab *Internal* e filtri per "HTML" puoi vedere una serie unica di lettere e numeri per ogni pagina. Questo numero seriale serve a contraddistinguere le pagine, Hash identici indicano pagine identiche. Se esporti il .CSV e lo apri in Excel potrai trovare subito gli Hash identici con le formattazioni condizionali.

Level	Hash
	e1a413350b13ba773f97bced6d372523
	2c4a008826d0bab200f1b60db4739d00
	dd966151473fad9a1f3a09eba9eb8bdb
	8501d543c828d7f50ccdd78bd21b13b7
	3231140303ca0ab80a249744fd565b7d
	9a71c0fa4f735267853c0c7ac4bd092
	91812eada58b5912980dd46358a0d13c
	b8250f6c954b3ea2948a47da9a1babbe

4.4 Come identificare le pagine che includono direttive dei crawler, es: nofollow/noindex/noodp/canonical etc.

Dopo aver terminato il crawling clicca sulla tab *Directives*. Puoi usare il filtro per visualizzare soltanto specifiche direttive:

- index
- noindex
- follow
- nofollow
- noarchive
- nosnippet
- noodp
- noydir
- noimageindex
- notranslate
- unavailable_after
- refresh
- canonical

Filter	Export
All	
Index	1 htt
Noindex	1 htt
Follow	1 htt
Nofollow	1 htt
NoArchive	1 htt
NoSnippet	1 htt
NoODP	1 htt

4.5 Come verificare che il file robots.txt funzioni come desiderato

Le opzioni standard di Screaming Frog lo forzano a seguire le direttive del robots.txt che incontra nel sito. Come priorità lo spider seguirà le direttive specifiche assegnate al bot di Screaming Frog. Se non viene definita nessuna direttiva specifica per l'user-agent di Screaming Frog, seguirà le direttive per Googlebot e se nemmeno le direttive per Googlebot sono state assegnate seguirà le direttive globali per tutti gli user-agent. Tieni presente che lo spider seguirà soltanto un set di direttive, **non mischierà direttive assegnate a spider diversi**.

Se vuoi bloccare alcune parti del sito allo spider, usa la sintassi regolare del robots.txt e assegna le direttive all'user-agent "Screaming Frog SEO Spider". Per far sì che lo spider ignori il robots.txt basta selezionare l'opzione dedicata nel pannello di configurazione.

4.6 Come trovare markup di Schema.org e altri microdati

Per trovare tutte le pagine che contengono markup di Schema oppure qualsiasi altro formato di Microdati, bisogna usare i filtri customizzati. Clicca su "Custom" nel menu di configurazione ed inserisci il footprint che stai cercando.

- Per trovare tutte le pagine che contengono markup di Schema.org, inserisci questo snippet nel filtro customizzato: `itemtype=http://schema.org`
- Per trovare markup specifici come ad esempio il markup di Schema per il rating, usa il filtro customizzato ``

Custom Filter Configuration

Search the source code of internal HTML pages.

Filter 1	Contains	<div itemscope itemtype="http://schema.org/
Filter 2	Contains	
Filter 3	Contains	
Filter 4	Contains	
Filter 5	Contains	

OK

Puoi inserire fino a 5 filtri customizzati differenti. Premi OK e procedi con il crawling del sito o di una lista di pagine.

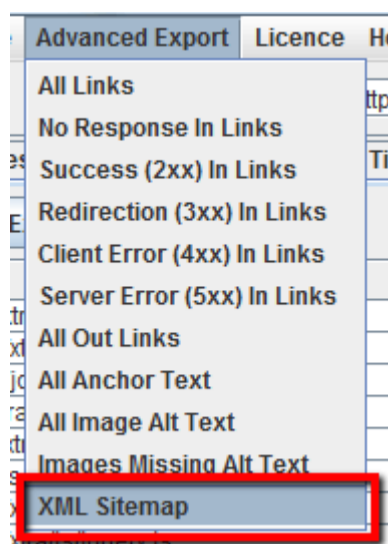
Quando il crawling è terminato seleziona la tab *Custom* nella parte superiore dello schermo per trovare tutte le pagine che contengono il tuo footprint. Se hai usato più di un filtro customizzato puoi vederne uno a uno cambiando il filtro nei risultati.

Filter	Export
Filter 1	
Filter 2	Address:
Filter 3	om/newarrivals/
Filter 4	om/kids-baby-sandals/baby-
Filter 5	om/kids-baby-sandals/

5. Utilità per Sitemap XML

5.1 Come creare una Sitemap XML

Dopo che lo spider ha terminato il crawling, clicca sul menu "Advanced Export" e seleziona "XML Sitemap".



Salva la sitemap e aprila con Excel. Seleziona "Read Only" e apri il file come tabella XML. Potresti ricevere un avvertimento di incompatibilità con il foglio di lavoro, premi "Yes" e procedi. Ora che la sitemap è stata importata puoi modificare facilmente i valori di frequency, priority ed eventualmente altri. Assicurati che la sitemap contenga una sola versione di ogni URL (canonico), senza parametri o fattori di duplicazione. Quando avrai finito di modificare il file, salva il lavoro come XML.

5.2 Come verificare la Sitemap XML

Prima di iniziare dovrai avere una copia della sitemap.xml salvata nel tuo computer. Apri Screaming Frog e abilita la modalità "List" cliccando la voce del menu "Mode". Importa il file della sitemap.xml cliccando su "Select File" e poi avvia lo spider. Una volta terminato il crawling potrai vedere se il sito ha pagine che finiscono in errore 404, se vengono redirezionate, se ci sono URL duplicati e altri errori che potrebbero *sporcare* la sitemap.

6. Problematiche generali

6.1 Individuare i motivi per cui certe pagine non vengono indicizzate

Ti stai chiedendo perchè certe pagine del tuo sito non vengono indicizzate? Per prima cosa assicurati che il robots.txt non blocchi per errore le pagine che vuoi indicizzare poi controlla di non avere tag noindex nell'head. Ora bisogna verificare che queste pagine siano facilmente raggiungibili dagli spider dei motori di ricerca attraverso i link interni al sito. Dopo che Screaming Frog ha terminato il crawling vai nella tab *Internal* e scorri a destra fino alla colonna "Inlinks". I valori riportati in colonna rappresentano il numero di link interni che ogni pagina riceve. Controlla che le pagine più importanti ricevano in proporzione più link interni rispetto a pagine secondarie.

Esporta la lista di URL interni in .CSV usando il filtro "HTML" nella tab *Internal*. Apri il file .CSV con Excel, in un secondo foglio incolla tutti gli URL che non sono indicizzati. Con la funzione VLOOKUP verifica che le pagine non indicizzate siano state raggiunte dal crawler di Screaming Frog.

- In caso negativo devi rivedere i link interni verso quella pagina poichè sembra non averne
- In caso affermativo quella pagina potrebbe ricevere troppo pochi link interni. Cerca di aumentarne il numero in proporzione alle altre pagine di minore importanza.

Consiglio:

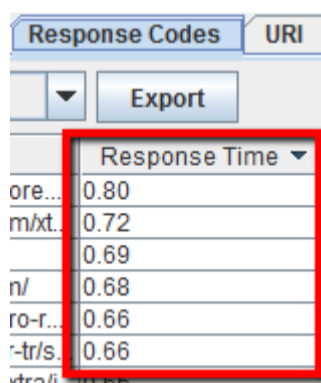
Su Google Doc potete trovare questo file da usare in Excel: [Pages Not Indexed](#). Una volta importati i dati di Screaming Frog, il foglio analizza i dati e fornisce possibili spiegazioni per cui le pagine che avete indicato non vengono indicizzate.

6.2 Come verificare la corretta migrazione di un sito web

Segui questa procedura per verificare la [corretta migrazione del tuo sito web](#). Prima di iniziare il crawling del nuovo devi procurarti la vecchia sitemap.xml del sito pre-migrazione. Premi sul tasto "Mode" e attiva la modalità "List". Carica la vecchia sitemap.xml e avvia lo spider. Per la verifica finale dovrai controllare che tutti i vecchi URL presenti in sitemap.xml vengano correttamente redirezionati verso i nuovi URL. Eventuali errori 404 indicano che tali URL non hanno regole di redirect.

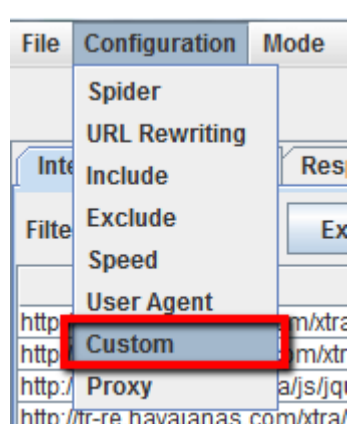
6.3 Come trovare pagine che hanno tempi lunghi di caricamento

Dopo che lo spider ha terminato il crawling, apri la tab *Response Codes* e ordina i dati con la colonna "Response Time". Questa visuale ti aiuterà ad identificare subito le pagine che necessitano di troppo tempo per venire caricate

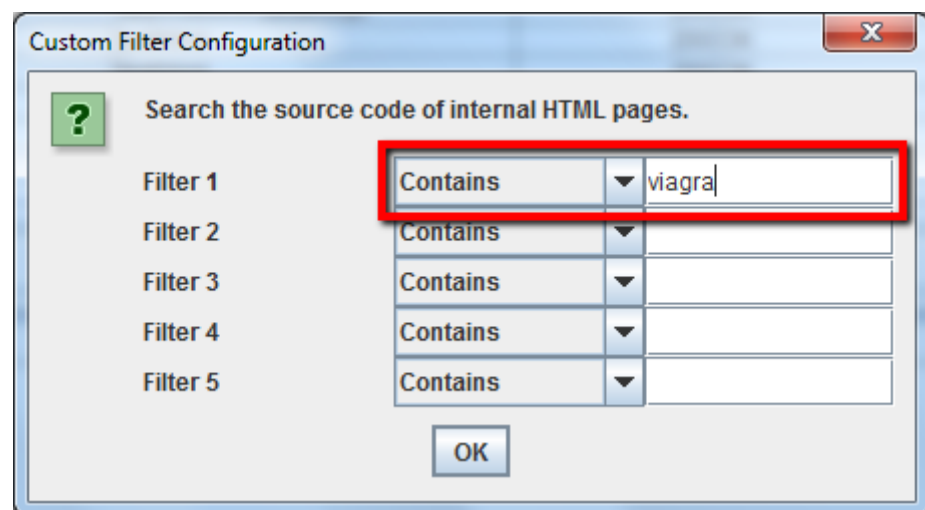


6.4 Come trovare malware e spam sul sito web

Per prima cosa devi identificare lo snippet di codice che rappresenta il malware o lo spam. Successivamente vai nel menu di configurazione e clicca su "Custom" per modificare i filtri customizzati. ed inserisci il footprint che ti interessa individuare.



Puoi inserire fino a 5 filtri customizzati. Avvia il crawler premendo OK per processare il sito o la lista di pagine.



Quando lo spider avrà terminato il processo, seleziona la tab *Custom* nella parte superiore dello schermo per vedere tutte le pagine che contengono il footprint. Se hai inserito più filtri puoi passare da uno all'altro semplicemente cambiando il filtro dei risultati.

6.5 Come aumentare la memoria allocata a Screaming Frog

Windows 32 & 64-bit: Prima di tutto, se hai una macchina a 64-bit, assicurati di aver scaricato ed installato la versione di Java a 64-bit, altrimenti non potrai allocare più memoria di una macchina a 32-bit. Apri la cartella dove hai installato lo spider (solitamente C:\Program Files\Screaming Frog SEO Spider), dovresti trovare 4 file: due applicazioni (install and uninstall), un file .JAR e poi il file .INI di configurazione che ci interessa editare, chiamato "ScreamingFrogSEOSpider.I4j.ini". Apri questo file con un editor di testo e noterai una riga di codice con questo numero "-Xmx512M" che rappresenta il massimo di memoria assegnata allo spider, ovvero "512Mb". Il valore pre-impostato è "512M", quindi per raddoppiare la memoria puoi semplicemente raddoppiare il numero sostituendolo con "1024". Per allocare, ad esempio, 6GB sostituisci il numero con "6144" (attenzione a non modificare il codice antecedente al numero, quindi per 1024 dovrai scrivere -Xmx1024M). Considera che questo valore si riferisce alla memoria RAM e non del disco, quindi se assegni più memoria di quella disponibile nel tuo computer lo spider non si avvierà. Se dovesse capitarti questa situazione apri di nuovo il file di configurazione ed inserisci un valore realistico basato sulla memoria disponibile nella tua macchina.

In genere le macchine Windows a 32-bit sono limitate ad un massimo di 4GB RAM. Questo tradotto vuole dire che potrai allocare un valore compreso tra 1,024mb e 1,500mb.

Per le macchine a 64-bit potrai allocare decisamente più memoria (relativamente a quanta installata). Screaming Frog è programmato per lavorare su macchine a 32 e 64-bit, ricordati di installare Java a 64-bit se la tua macchina è a 64-bit altrimenti non potrai allocare tutta la memoria potenziale dei 64-bit.

Se Windows non ti permette di modificare il file di configurazione (probabilmente ti servono i diritti di amministratore), prova a copiare il file sul desktop, modificalo e rimettilo nella cartella dello spider sostituendo il file vecchio..

Per verificare che la memoria sia effettivamente aumentata riavvia Screaming Frog, clicca su "Help" e successivamente su "Debug" e osserva il valore "Max". Il valore è sempre leggermente inferiore a quanto allocato, è normale e dipende da JVM management.

Mac: Esistono due metodi per incrementare la memoria di Screaming Frog sul Mac e dipende dalla versione del crawler che si sta utilizzando. Se stai usando l'ultima versione (oppure dalla 2.55 in avanti) puoi semplicemente aprire il Terminale (lo trovi dentro "Applications" e "Utilities") e digitare `uk.co.screamingfrog.seo.spider Memory 1g`.

Questo comando alloca 1 GB di memoria a Screaming Frog. Per allocare 9 GB scrivi: `uk.co.screamingfrog.seo.spider Memory 8g`.

Puoi anche specificare la memoria in megabytes, usando la stringa: `uk.co.screamingfrog.seo.spider Memory 2048m`.

Queste impostazioni verranno ricordate anche dopo un aggiornamento e non sarà necessario ripetere l'operazione. Puoi verificare le impostazioni digitando nel terminale: `defaults read uk.co.screamingfrog.seo.spider Memory`. Il risultato sarà tipo: 8g.

Se stai utilizzando un Mac con versione OS inferiore alla 10.7.3 (32-Bit Macs) dovrai usare una versione uguale o inferiore alla 2.40 di Screaming Frog e dovrai editare il file di configurazione come segue.

Apri in un editor di testo il file "Info.plist" dentro la cartella "/Applications/Screaming Frog SEO Spider", trova la stringa -Xmx512Mm. Aumenta il valore in base alla memoria disponibile sul tuo sistema, per raddoppiare l'allocazione di memoria ad esempio inserisci -Xmx1024Mm

Se non riesci a trovare il file "info.plist" vai nella cartella "applications" con il finder. Clicca con il pulsante destro su "ScreamingFrogSEOSpider" e seleziona "Show package contents". Il file "info.plist" è dentro la cartella "contents". Clicca con il pulsante destro sul file e aprilo con un editor di testo.

Ubuntu: per aumentare la memoria e fare crawling di siti più grandi devi modificare il numero nel file .screamingfrogseospider situato nella cartella "Home" dell'utente. Il file viene creato la prima volta che Screaming Frog viene avviato.

Il file è impostato di default su questo valore: -Xmx512M. Puoi incrementare il valore come spiegato sopra.

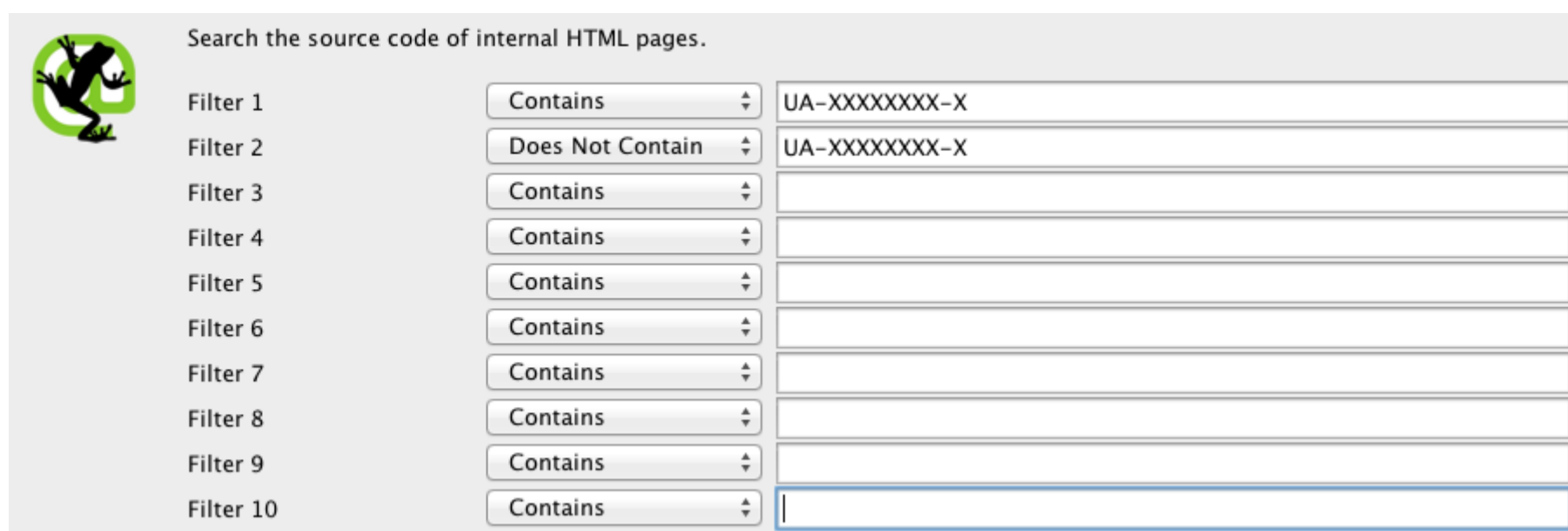
Se non sei in grado di aumentare la memoria allocata e devi fare crawling di siti molto estesi, ti consiglio di dividere il sito in sezioni. Puoi usare il menu di configurazione per fare crawling soltanto delle pagine HTML (evitando quindi immagini, CSS e JS) oppure escludendo alcune sezioni del sito. Come alternativa se il sito ha una struttura di URL ben definita puoi dividere il crawling in cartelle (/holidays/, /blog/ etc) oppure usare la funzione Regex per includere soltanto certi URL.

Screaming Frog non è stato sviluppato per fare crawling di siti con centinaia di migliaia di pagine poichè non alloca tutta la memoria in database. Per studiare siti molto grossi si possono usare servizi di crawling online come [80legs](#) oppure [DeepCrawl](#).

7. PPC & Analytics

7.1 Come verificare che il codice di tracciamento di Google Analytics sia presente in tutte le pagine

Prima di avviare il crawler apri i filtri customizzati e nei primi due campi inserisci il codice UA di Analytics che vuoi verificare:

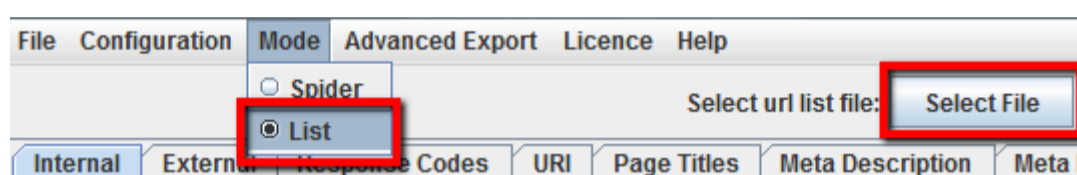


- Contains: UA-XXXXXXXX-X
- Does Not Contain: UA-XXXXXXXX-X

Avvia Screaming Frog e quando ha terminato vai nella tab *Custom*. Il filtro così impostato ti permetterà di passare da una visualizzazione all'altra semplicemente cambiando il filtro dei risultati.

7.2 Come verificare una lista di landing page usate per il PPC

Salva la tua lista in formato .TXT oppure .CSV poi cambia Screaming Frog in modalità "List".



Successivamente seleziona il file da caricare e premi "Start" per avviare lo spider. A processo terminato apri la tab *Internal* e controlla gli status code.

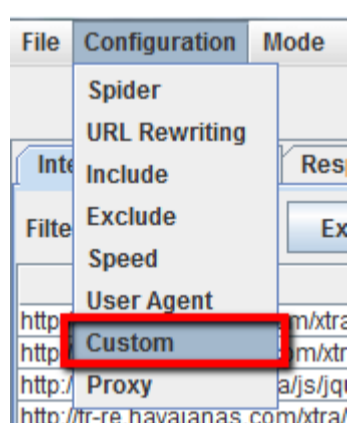
8. Scraping di contenuti

8.1 Come fare scraping dei meta dati di una lista di pagine

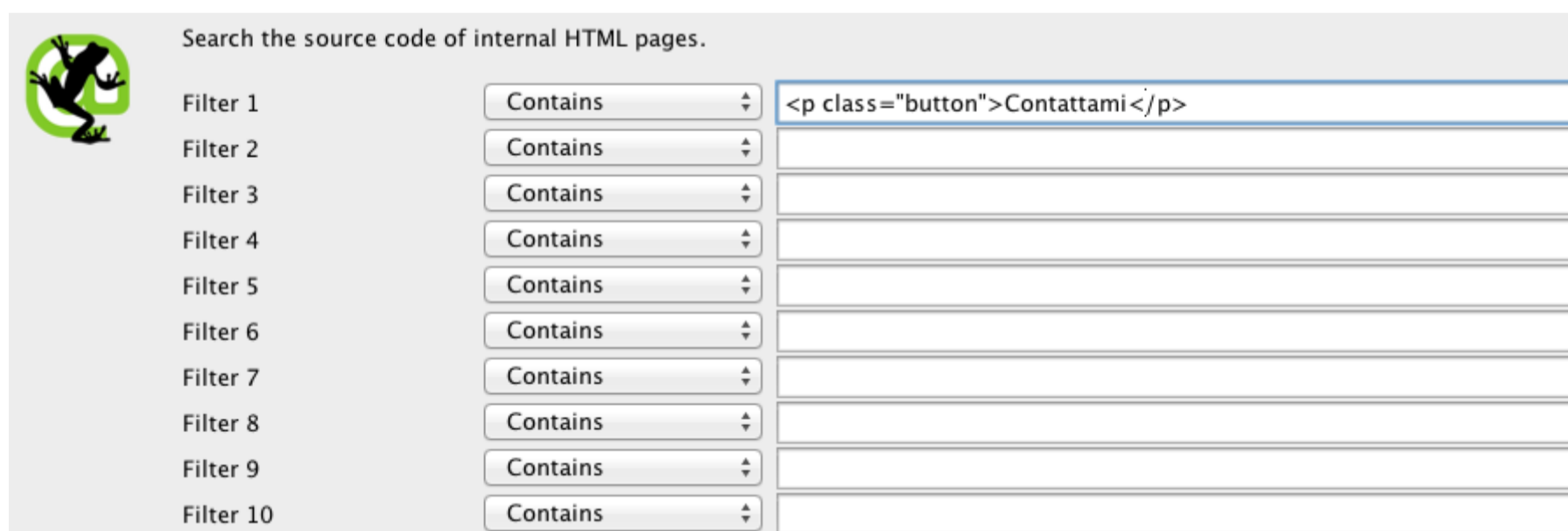
Per raccogliere informazioni e meta dati da una lista di URL devi attivare la modalità "List", caricare la lista di pagine in formato .TXT oppure .CSV e poi avviare lo spider. A processo terminato potrai analizzare gli status code, i link in uscita, il conteggio delle parole contenute nelle singole pagine e ovviamente tutti i meta dati.

8.2 Come fare scraping di tutte le pagine che contengono un determinato footprint

Per prima cosa devi identificare il footprint che ti serve. Salva il codice HTML dell'elemento che vuoi cercare, poi apri il pannello di configurazione, clicca su "Custom" ed inserisci il footprint.



Puoi inserire fino a 5 filtri customizzati nello spider. Premi OK e procedi con il crawling del sito o delle pagine in elenco. In questo specifico esempio voglio identificare tutte le pagine che includono il pulsante "Contattami", ho quindi cercato il codice del pulsante nelle pagine HTML e l'ho inserito nel campo customizzato del filtro.



Quando lo spider ha terminato di fare crawling, seleziona la tab *Custom* nella parte superiore dello schermo per vedere tutte le pagine che contengono il footprint. Se hai inserito più di un filtro customizzato puoi visualizzare i rispettivi dati cambiando il filtro dei risultati.

8.3 Come fare scraping di tutte le pagine che contengono un determinato footprint

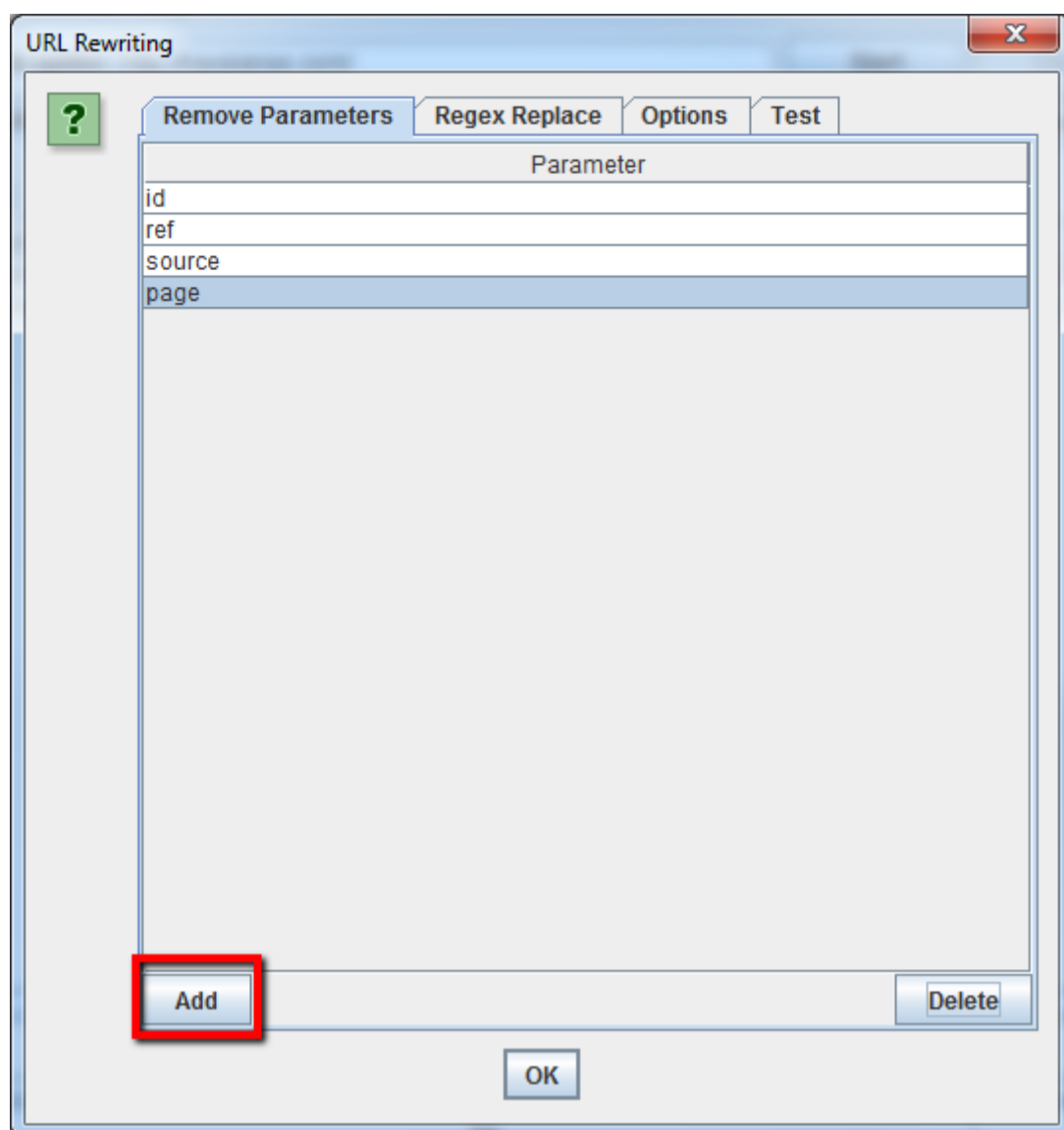
Per prima cosa devi identificare il footprint che ti serve. Salva il codice HTML dell'elemento che vuoi cercare, poi apri il pannello di configurazione, clicca su "Custom" ed inserisci il footprint.

9. Riscrittura degli URL

9.1 Come trovare e rimuovere i parametri e le session id dagli URL scansionati

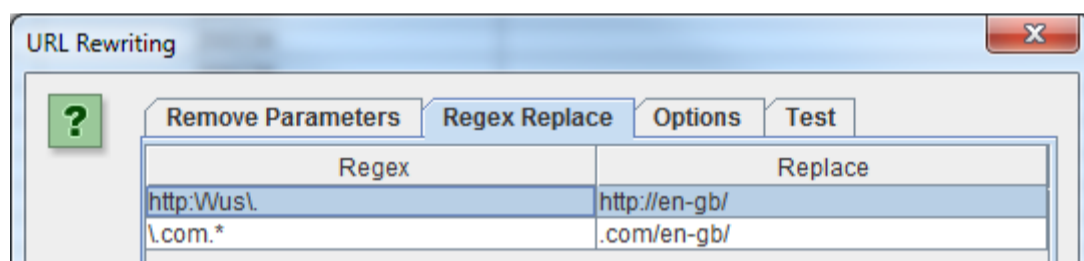
Per identificare tutti gli URL che contengono ID di sessione e parametri, avvia SCreaming Frog con le impostazioni di default. Quando lo spider ha terminato clicca sulla tab *URI* e filtra per "Dynamic" per vedere tutti gli URL che includono parametri.

Per rimuovere i parametri dagli URL raccolti seleziona "URL Rewriting" nel menu di configurazione, apri la tab *Remove Parameters* e clicca su "Add" per aggiungere i parametri che vuoi rimuovere dagli URL e premi OK. Dovrai lanciare di nuovo il crawling con questi settaggi per ottenere la riscrittura degli URL.

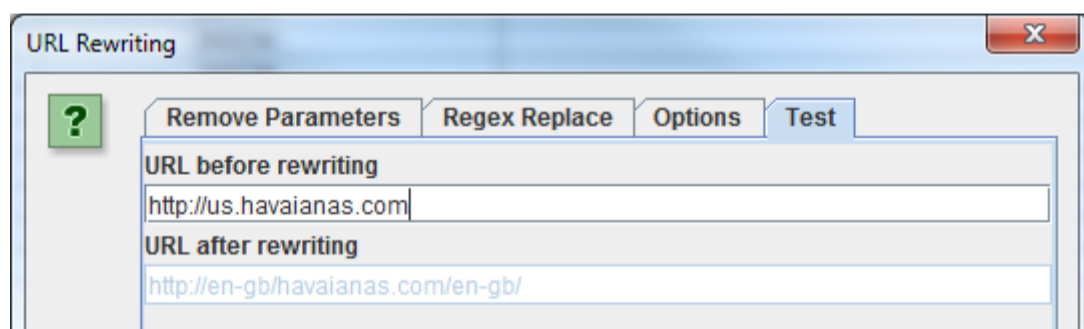


9.2 Come riscrivere gli URL raccolti dal crawler (es: sostituire .com con .co.uk, oppure riscrivere tutti gli URL in minuscolo)

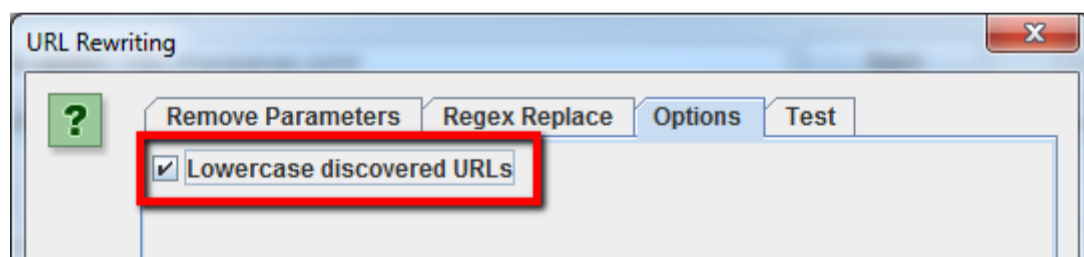
Per riscrivere tutti gli URL archiviati dallo spider, seleziona "URL Rewriting" nel menu di configurazione. Nella tab "Regex Replace" clicca su "Add" per aggiungere una regola RegEx che definisca quello che vuoi sostituire.



Una volta che hai definito tutte le tue regole di riscrittura ti consiglio di testarle. Vai nella tab *Test*, inserisci l'URL che vuoi provare nello spazio "URL before rewriting" ed il campo "URL after rewriting" si aggiornerà automaticamente seguendo le regole che hai impostato.



Per rimuovere dal crawling tutti gli URL duplicati da caratteri maiuscoli devi selezionare "Lowercase discovered URLs" nella finestra delle opzioni.



Ricorda che le impostazioni vanno decise prima di avviare lo spider per abilitare la riscrittura degli URL.

10. Analisi delle parole chiave

10.1 Come trovare le pagine più importanti nei siti competitor

Generalmente i siti dei competitor cercheranno di guidare *link popularity* e traffico verso le loro pagine più importanti attraverso link interni. I competitor orientati alla SEO probabilmente linkeranno le pagine più importanti dal loro blog.

Per trovare la pagine più importanti di un sito avvia il crawling e poi nella tab *Internal* ordina la colonna "Inlinks". Le pagine più linkate internamente saranno *presumibilmente* anche le più importanti per il competitor.

Word Count	Level	Inlinks	Outlinks	External Outlinks
176		282	35	52
596		282	41	70
124		228	40	55
208		223	159	16
626		167	26	27
208		164	160	0
32		154	13	20

Per trovare le pagine linkate del blog del tuo competitor, non selezionare l'opzione "Check links outside folder" nel menu di configurazione dello spider poi avvia il crawling della cartella o sottodominio del blog. A questo punto nella tab *External* filtra i risultati cercando l'URL del dominio principale. Scorri la schermata a destra e ordina la colonna "Inlinks" per vedere quali pagine sono linkate più spesso.

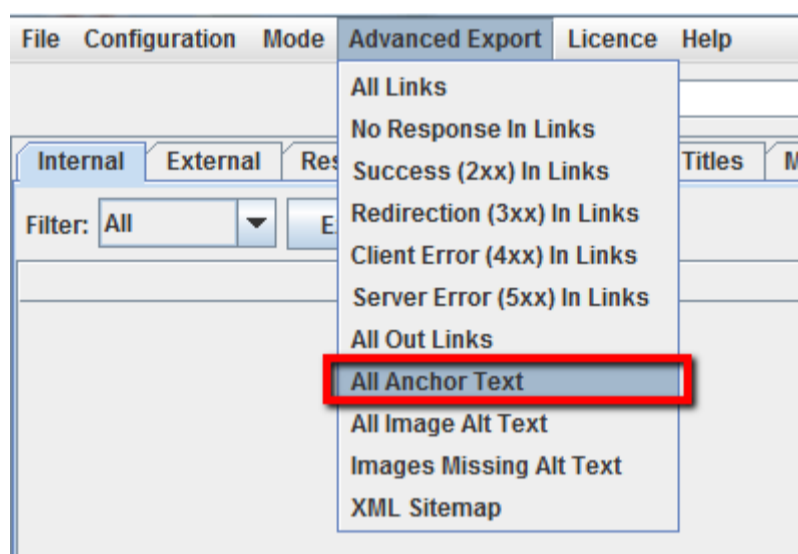
Inlinks	Level	Status
1406	2	OK
140	2	OK
108	3	OK
48	2	OK
25	2	OK
19	2	OK

Consiglio:

Puoi ottimizzare la visualizzazione dei dati semplicemente trascinando le colonne nella posizione desiderata.

10.2 Come sapere quali anchor text usano i competitors nei link interni

Avvia il crawling sul sito del competitor. A processo terminato vai nel menu "Advanced Export" e seleziona "All Anchor Text" per esportare il file .CSV con l'elenco degli anchor text usati nel sito, dove sono usati e a cosa linkano.



10.3 Come trovare le parole chiave (Meta Keywords) usate dai competitor

Dopo che Screaming Frog ha terminato il crawling del sito del competitor, guarda la tab *Meta Keywords* per trovare la lista delle Meta Keyword usate in ogni pagina. Ordina la colonna "Meta Keyword 1" per trovare le tag vuote.

10.4 Fare scraping della pagina dei risultati di Google

Inserisci un URL di ricerca con Google in Screaming Frog, ad esempio <https://www.google.it/search?q=hotel+milano>. Nel menu di configurazione dello spider **seleziona soltanto** "Ignore robots.txt" mentre nella tab *Advanced* non selezionare niente. Con l'opzione "Include" puoi scegliere se fare scraping di tutte le pagine che lo spider troverà, oppure decidere se restare su Google.it per analizzare soltanto la prima o tutte le pagine dei risultati (attento a non esagerare oppure Google bannerà il tuo IP temporaneamente). L'ultima impostazione da cambiare è sotto il menu "User Agent", entra e seleziona **Chrome**. A questo punto puoi avviare lo spider.

Clicca sull'unico URL trovato (se hai forzato l'analisi solo sulla prima pagina) e nella parte inferiore dello schermo, nella tab *Out Links*, troverai molte informazioni estratte dalla prima pagina di Google:

- i termini correlati
- i top 10 competitor
- i titoli delle loro pagine
- i siti correlati
- la cache e molto altro.

Con un minimo di fantasia potrete trovare altri trucchi interessanti per recuperare dati da Google.

I più attenti avranno notato in mezzo alla lista alcuni URL con il parametro `?q=related:www.esempio.com`. Questi URL sono molto importanti per capire come Google relaziona argomenti e siti web. Il parametro *related* permette di ottenere, dato un sito web di partenza, una lista di altri siti web correlati, ovvero molto vicini per argomenti e tipologia. Questa info torna molto utile per pianificare strategie di [link building](#) e content marketing poiché aiuta ad individuare i siti più attinenti sui quali pubblicare testi e backlink.

Un altro gruppo di URL che può essere analizzato è *webcache* poiché permette di verificare se l'URL è presente nella cache di Google oppure no.

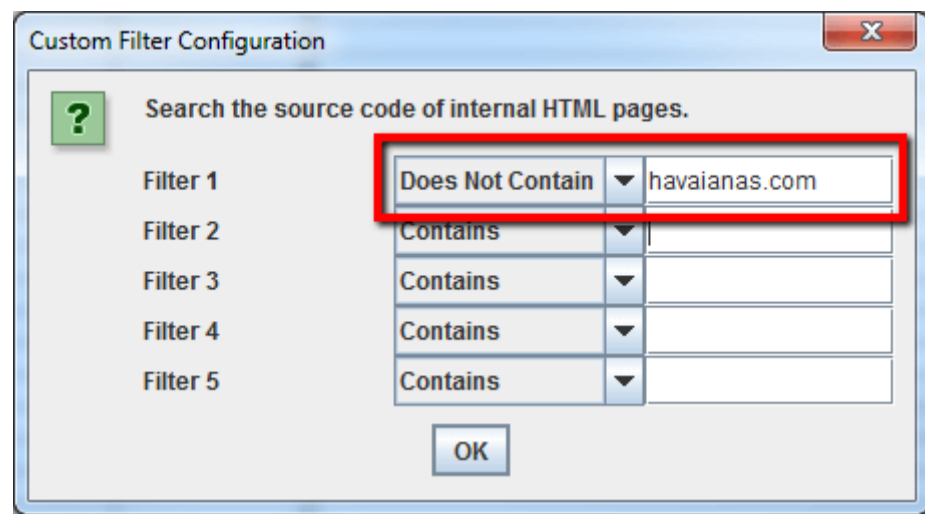
11. Funzioni per la Link Building

11.1 Come analizzare una lista di URL a cui puntano i link

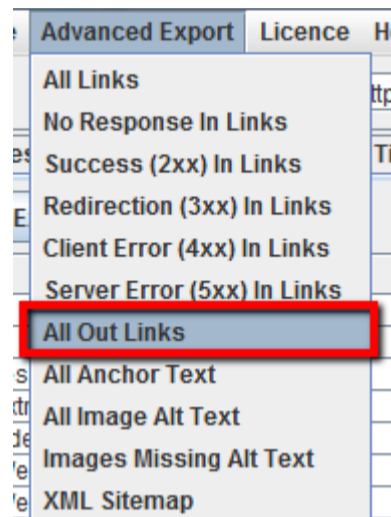
Se hai fatto scraping oppure hai una lista di URL da controllare, puoi caricare la lista in Screaming Frog tramite la modalità "List" raccogliere alcune informazioni su quelle pagine. Dopo che lo spider avrà finito il suo lavoro vai nella tab *Response Codes* e analizza gli Status Code. Potrai ottenere info riguardanti tutti gli outbound links, i tipi di link, le anchor text e le direttive nofollow nella tab *Out Links* nella parte inferiore dello schermo. La lista ti

aiuterà a capire a quali tipologie di siti linkano queste pagine e soprattutto come. Per controllare la tab *Out Links* assicurati di aver selezionato l'URL che ti interessa nella finestra superiore.

Puoi anche usare un filtro customizzato per determinare se queste pagine ti stanno già linkando o meno.



Puoi anche esportare la lista completa dei link in uscita dal menu "Advanced Export Menu" cliccando sulla voce "All Out Links". Il file conterrà tutti i link che puntano all'esterno.



11.2 Come trovare link rotti ed usarli a proprio vantaggio

Hai trovato un sito dal quale ti piacerebbe ricevere un backlink? Usa Screaming Frog per trovare collegamenti rotti su quel sito. Contatta successivamente il webmaster e se hai contenuti che possono sostituire la pagina di destinazione del link rotto proponila, se ti va bene riceverai un backlink altrimenti avrai comunque fatto un gesto che verrà apprezzato e da cosa nasce cosa, chi lo sa :)

11.3 Come posso verificare i miei backlinks e le relative anchor text

Carica la tua lista di backlink in Screaming Frog e lancia lo spider in modalità "List". A processo terminato esporta la lista completa di link cliccando "All Out Links" nella tab *Advanced Export Menu*. La lista contiene tutti gli URL e le anchor text/alt tag per tutti i link in queste pagine. Ora puoi usare un filtro con Excel sulla colonna "Destination" per determinare se il tuo sito è linkato e quale anchor text/alt tag è incluso.

11.4 Come verificare se il tuo sito fa parte di un network di link

Vuoi verificare se un gruppo di siti si sta linkando reciprocamente? Leggi questa guida [visualizing link networks using Screaming Frog and Fusion Tables](#).

11.5 Sto facendo pulizia di backlink e voglio controllare che siano stati tutti definitivamente cancellati

Imposta un filtro customizzato che contiene l'URL della root del tuo dominio, poi carica la lista di backlink in Screaming Frog e lancia lo spider in modalità "List Mode". Quando lo spider ha finito di fare crawling, seleziona la tab *Custom* per vedere tutte le pagine che stanno ancora linkando verso il tuo sito.

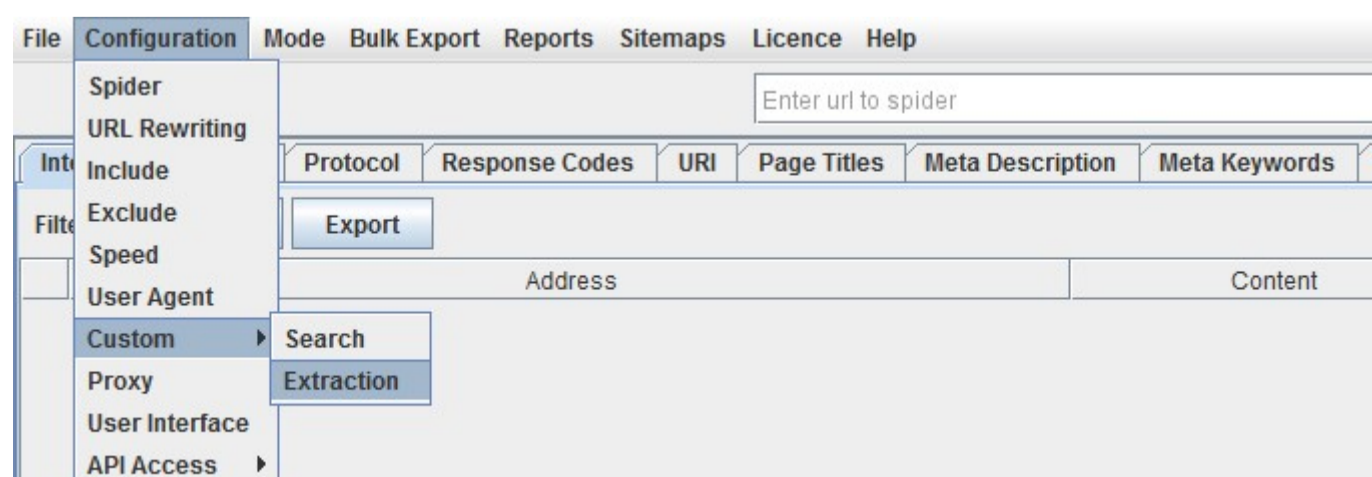
11.6 Come trovare i backlink dofollow tra quelli segnalati da Google Webmaster Tools

Scarica l'elenco dei backlink da Webmaster Tools, avvia Screaming Frog e dal menu "Mode" attiva la visione "List". Clicca il bottone "Select File", importa la lista di backlink e avvia lo spider.

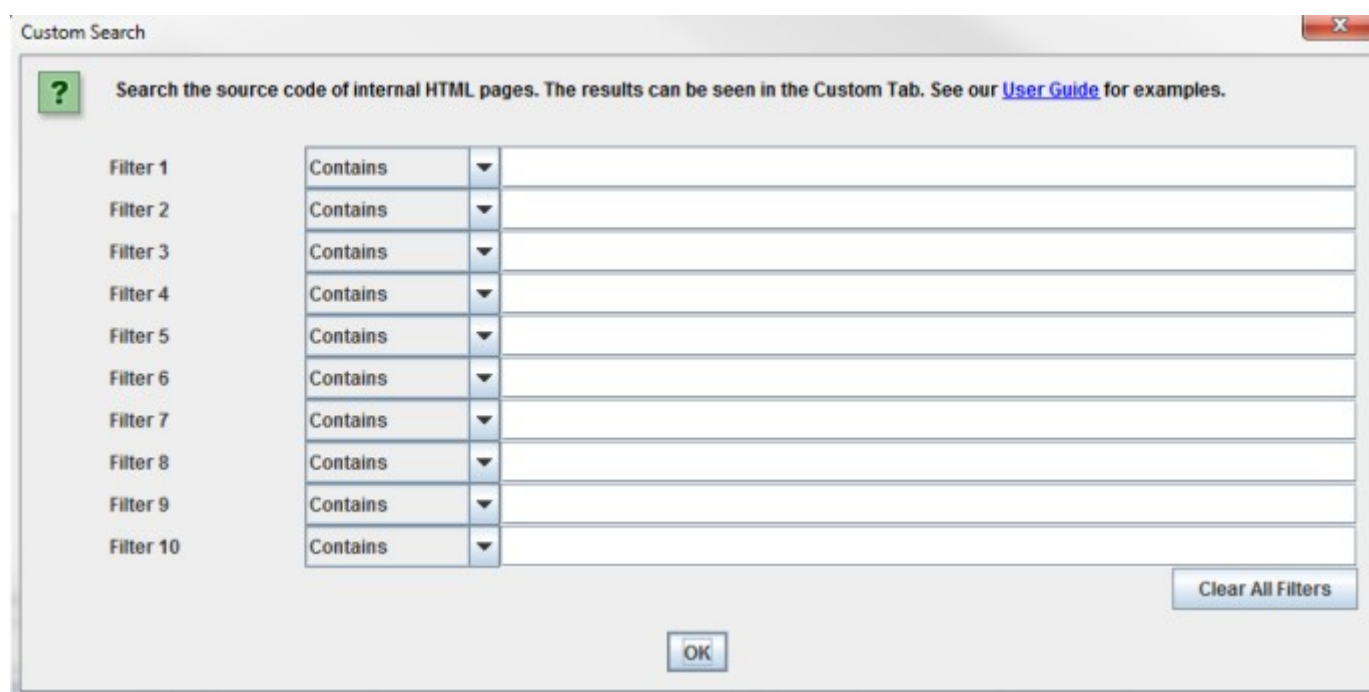
A processo terminato clicca su "Advanced Export" e seleziona "All Out Links". Apri il file .CSV con Excel, con il comando dati/text-to-columns ordina i dati in colonne e poi applica i filtri. Dalla colonna "Destination" filtra soltanto gli URL del tuo dominio e nella colonna "Follow" puoi vedere quali backlink sono dofollow e quali no.

Estrazioni customizzate

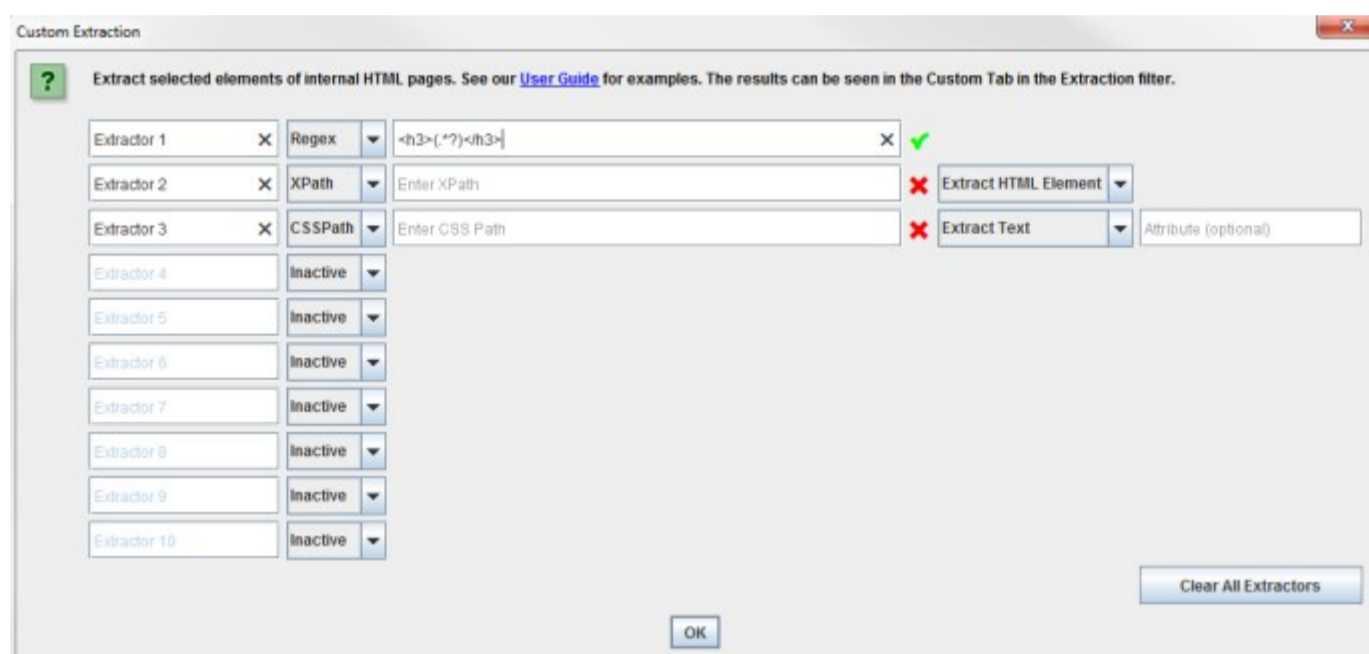
La nuova funzione "custom extraction" permette di collezionare qualsiasi informazione estratta dall'HTML di un URL. Se sei familiare a strumenti di scraping che importano l'XML e Xpath (SeoTools per Excel, Scraper per Chrome, ...) allora sai di cosa sto parlando. Questa funzione la trovi sotto "Configuration > Custom".



"Search" è la classica ricerca nel codice HTML.



“Extraction” è simile alla funzione Search, hai a disposizione 10 campi per estrarre qualsiasi cosa dall’HTML della pagina usando sia XPath, CSS path oppure RegEx.



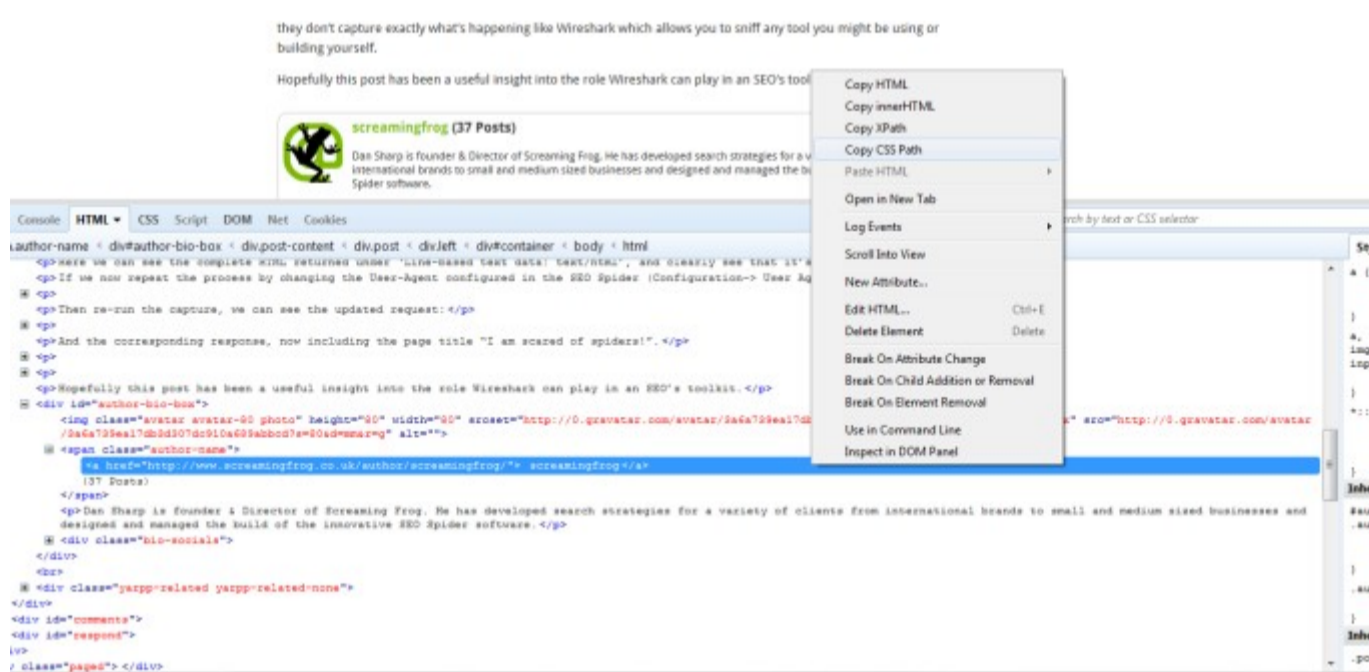
Quando usi XPath o CSS Path per fare scraping dell’HTML, puoi decidere cosa estrarre:

- **Estrarre elementi HTML:** estrae l’elemento completo e l’HTML contenuto in esso.
- **Estrarre Inner HTML:** estrae il contenuto dell’HTML dell’elemento selezionato. Se l’elemento selezionato contiene un altro elemento, questo verrà incluso.
- **Estrarre semplice testo:** estrae il testo contenuto nell’elemento ed il testo di sotto elementi (essenzialmente viene estratto tutto l’HTML).

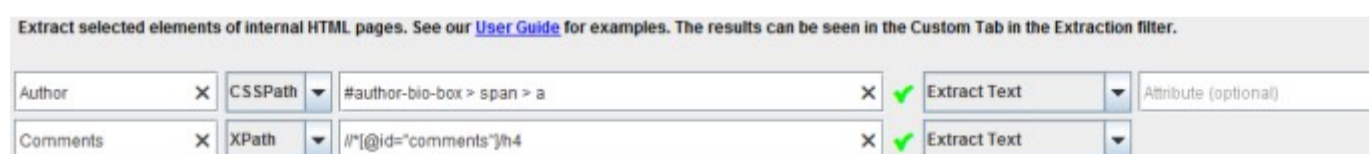
Come verifica della sintassi di XPath, CSS path e RegEx inserita nei filtri viene attivata una virgoletta verde se è corretta, altrimenti una x rossa se la sintassi è errata.

Autori & Commenti

Ora Screaming Frog ci può dire quale sia l’autore di ogni post e quanti commenti ogni post abbia ricevuto, bello no? Tutto quello che ti serve fare è aprire un post su Chrome, cliccare col tasto destro e selezionare “Ispeziona elemento” sul dato che vuoi collezionare e copiare l’ XPath o CSS path che ti interessa.



Puoi anche rinominare gli “estrattori” assegnando un nome alle colonne. In questo caso ho usato “CSS Path” e “XPath”.



Il nome autore e la somma dei commenti vengono mostrati sotto il filtro “extraction” dentro la tab “Custom”.

Internal	External	Protocol	Response Codes	URI	Page Titles	Meta Description	Meta Keywords	H1	H2	Images	Directives	AJAX	Custom
				Address	Status Code	Status							Mobile Annotation
1				http://www.huffingtonpost.co.uk/	200	OK							http://m.huffpost.com/uk/
2				http://www.huffingtonpost.co.uk/news/food/	200	OK							http://m.huffpost.com/uk/news/uk-food
3				http://www.huffingtonpost.co.uk/young-voices/	200	OK							http://m.huffpost.com/vertical/uk-universities-education/
4				http://www.huffingtonpost.co.uk/news/pics-giggles/	200	OK							http://m.huffpost.com/uk/news/pics-giggles/
5				http://www.huffingtonpost.co.uk/news/property/	200	OK							http://m.huffpost.com/uk/news/property
6				http://www.huffingtonpost.co.uk/news/skills-space/	200	OK							http://m.huffpost.com/uk/news/skills-space
7				http://www.huffingtonpost.co.uk/huffingtonpostcouk-faq.html	200	OK							http://m.huffpost.com/uk/entry/
8				http://www.huffingtonpost.co.uk/karim-goudiaby/	200	OK							http://m.huffpost.com/uk/author/karim-goudiaby/
9				http://www.huffingtonpost.co.uk/news/innovation/	200	OK							http://m.huffpost.com/uk/news/uk-innovation
10				http://www.huffingtonpost.co.uk/huffingtonpostcouk-careers.html	200	OK							http://m.huffpost.com/uk/entry/
11				http://www.huffingtonpost.co.uk/news/men/	200	OK							http://m.huffpost.com/uk/news/uk-men
12				http://www.huffingtonpost.co.uk/david-emm/	200	OK							http://m.huffpost.com/uk/author/david-emm/
13				http://www.huffingtonpost.co.uk/sul-breaks/	200	OK							http://m.huffpost.com/uk/author/sul-breaks/

Tag Multilingua Hreflang

Con la nuova versione 4.0 di Screaming Frog possiamo estrarre tutte le tag Hreflang.

```
(//*[@hreflang])[1]
```

```
(//*[@hreflang])[2]
```

```
--
```

Alternativa:

```
//link[@rel="alternate"][1]/@hreflang
```

```
//link[@rel="alternate"][2]/@hreflang
```

```
//link[@rel="alternate"][3]/@hreflang
```

```
//link[@rel="alternate"][4]/@hreflang
```

```
//link[@rel="alternate"][...]/@hreflang
```

Alternativa con definizione del country code:

```
//link[@hreflang="en-US"]/@hreflang
```

```
//link[@hreflang="en-AU"]/@hreflang
```

Extract selected elements of internal HTML pages. See our [User Guide](#) for examples. The results can be seen in the Custom Tab in the Extraction filter.

Label	XPath	Filter	Action
hreflang	XPath	(//*[hreflang])[1]	Extract HTML Element
hreflang 2	XPath	(//*[hreflang])[2]	Extract HTML Element
hreflang 3	XPath	(//*[hreflang])[3]	Extract HTML Element

La stringa che segue raccoglie tutto l'elemento HTML, con il link al valore Hreflang.

Internal	External	Protocol	Response Codes	URI	Page Titles	Meta Description	Meta Keywords	H1	H2	Images	Directives	AJAX	Custom
				Address	Status Code	Status							hreflang
1				http://uk.moo.com/	200	OK							<link href="http://uk.moo.com" hreflang="en-GB" rel="alternate">
2				http://uk.moo.com/pickup/	200	OK							<link href="http://uk.moo.com/pickup" hreflang="en-GB" rel="alternate">
3				http://uk.moo.com/products/minicards.html	200	OK							<link href="http://uk.moo.com/products/minicards.html" hreflang="en-GB" rel="alternate">
4				http://uk.moo.com/products/minicards.html	200	OK							<link href="http://uk.moo.com/products/minicards.html" hreflang="en-GB" rel="alternate">
5				http://uk.moo.com/products/greeting-cards.html	200	OK							<link href="http://uk.moo.com/products/greeting-cards.html" hreflang="en-GB" rel="alternate">
6				http://uk.moo.com/luxe-project/index.html	200	OK							<link href="http://uk.moo.com/luxe-project/index.html" hreflang="en-GB" rel="alternate">
7				http://uk.moo.com/products/marketing-and-stationery.html	200	OK							<link href="http://uk.moo.com/products/marketing-and-stationery.html" hreflang="en-GB" rel="alternate">
8				http://uk.moo.com/design-templates/minicards/pack/fingerprints.html	200	OK							<link href="http://uk.moo.com/design-templates/minicards/pack/fingerprints.html" hreflang="en-GB" rel="alternate">
9				http://uk.moo.com/products/stickerbooks.html	200	OK							<link href="http://uk.moo.com/products/stickerbooks.html" hreflang="en-GB" rel="alternate">
10				http://uk.moo.com/products/postcards.html	200	OK							<link href="http://uk.moo.com/products/postcards.html" hreflang="en-GB" rel="alternate">
11				http://uk.moo.com/design-templates/luxe/business-cards/back/ghostpapers-black.html	200	OK							<link href="http://uk.moo.com/design-templates/luxe/business-cards/back/ghostpapers-black.html" hreflang="en-GB" rel="alternate">
12				http://uk.moo.com/products/flyers.html	200	OK							<link href="http://uk.moo.com/products/flyers.html" hreflang="en-GB" rel="alternate">
13				http://uk.moo.com/ideas/girls-that-travel.html	200	OK							<link href="http://uk.moo.com/ideas/girls-that-travel.html" hreflang="en-GB" rel="alternate">
14				http://uk.moo.com/affiliates/	200	OK							<link href="http://uk.moo.com/affiliates/" hreflang="en-GB" rel="alternate">
15				http://uk.moo.com/ideas/stocking-fillers-for-kids.html	200	OK							<link href="http://uk.moo.com/ideas/stocking-fillers-for-kids.html" hreflang="en-GB" rel="alternate">
16				http://uk.moo.com/accessories/notebooks.html	200	OK							<link href="http://uk.moo.com/accessories/notebooks.html" hreflang="en-GB" rel="alternate">
17				http://uk.moo.com/ideas/christmas-ideas.html	200	OK							<link href="http://uk.moo.com/ideas/christmas-ideas.html" hreflang="en-GB" rel="alternate">

Se invece vuoi solo recuperare il valore della tag Hreflang, puoi specificare l'attributo Xpath @hreflang:

Extract selected elements of internal HTML pages. See our [User Guide](#) for examples. The results can be seen in the Custom Tab in the Extraction filter.

Label	XPath	Filter	Action
hreflang	XPath	(//*[hreflang])[1]/@hreflang	Extract Inner HTML
hreflang 2	XPath	(//*[hreflang])[2]/@hreflang	Extract Inner HTML
hreflang 3	XPath	(//*[hreflang])[3]/@hreflang	Extract Inner HTML

Questo filtro colleziona soltanto i valori dei linguaggi inseriti nella tag:

Internal	External	Protocol	Response Codes	URI	Page Titles	Meta Description	Meta Keywords	H1	H2	Images	Directives	AJAX	Custom
				Address	Status Code	Status	hreflang	hreflang 2	hreflang 3				
1				http://uk.moo.com/	200	OK	en-GB	fr-FR	de-DE				
2				http://uk.moo.com/affiliates/	200	OK	en-GB	fr-FR	de-DE				
3				http://uk.moo.com/affiliates/contact.html	200	OK	en-GB	fr-FR	de-DE				
4				http://uk.moo.com/de/about/terms-conditions/	200	OK	en-GB	fr-FR	de-DE				
5				http://uk.moo.com/de/accessories/strathmore-notebook-large.html	200	OK	en-GB	fr-FR	de-DE				
6				http://uk.moo.com/de/design-templates/business-cards/pack/slater.html	200	OK	en-GB	fr-FR	de-DE				
7				http://uk.moo.com/de/design-templates/luxe/minicards/	200	OK	en-GB	fr-FR	de-DE				
8				http://uk.moo.com/de/design-templates/square-business-cards/	200	OK	en-GB	fr-FR	de-DE				
9				http://uk.moo.com/de/design-templates/square-business-cards/pack/marble-marvel.html	200	OK	en-GB	fr-FR	de-DE				
10				http://uk.moo.com/de/products/business-services/	200	OK	en-GB	fr-FR	de-DE				
11				http://uk.moo.com/de/products/luxe/postcards.html	200	OK	en-GB	fr-FR	de-DE				
12				http://uk.moo.com/de/products/original-business-cards.html	200	OK	en-GB	fr-FR	de-DE				
13				http://uk.moo.com/de/products/stickers-range.html	200	OK	en-GB	fr-FR	de-DE				
14				http://uk.moo.com/design-templates/luxe/business-cards/pack/ghostpapers-black.html?...	200	OK	en-GB	fr-FR	de-DE				
15				http://uk.moo.com/design-templates/luxe/fletterheads/pack/jules-hufig.html	200	OK	en-GB	fr-FR	de-DE				
16				http://uk.moo.com/design-templates/luxe/minicards/	200	OK	en-GB	fr-FR	de-DE				
17				http://uk.moo.com/design-templates/luxe/notecards/pack/bertram-hotel-notecards.html	200	OK	en-GB	fr-FR	de-DE				

Social Meta Tags: Open Graph e Twitter Card

Per estrarre le tag dedicate ai social media, come Open Grap e Twitter card, puoi usare i seguenti filtri:

```
//meta[starts-with(@property, 'og:title')][1]/@content
//meta[starts-with(@property, 'og:description')][1]/@content
//meta[starts-with(@property, 'og:type')][1]/@content
...
```

Extract selected elements of internal HTML pages. See our [User Guide](#) for examples. The results can be seen in the Custom Tab in the Extraction filter.

OG Title	XPath	Filter	Action
X	//meta[starts-with(@property, 'og:title')][1]/@content	X	Extract Inner HTML
X	//meta[starts-with(@property, 'og:description')][1]/@content	X	Extract Inner HTML
X	//meta[starts-with(@property, 'og:type')][1]/@content	X	Extract Inner HTML
X	//meta[starts-with(@property, 'twitter:title')][1]/@content	X	Extract Inner HTML
X	//meta[starts-with(@property, 'twitter:description')][1]/@content	X	Extract Inner HTML
X	//meta[starts-with(@property, 'twitter:image:src')][1]/@content	X	Extract Inner HTML
X	//meta[starts-with(@property, 'twitter:creator')][1]/@content	X	Extract Inner HTML

Ad esempio, sul sito di Moz questi sono i risultati di scansione delle tag social:

File Configuration Mode Bulk Export Reports Sitemaps Licence Help

https://moz.com/blog

Resume Clear

Internal	External	Protocol	Response Codes	URI	Page Titles	Meta Description	Meta Keywords	H1	H2	Images	Directives	AJAX	Custom
Filter: Extraction Export													
				Address	OG Title	OG Description							
1				https://moz.com/blog									
2				https://moz.com/blog/7-dead-simple-ways-to-improve-slide-decks	7 Dead-Simple Ways to Improve Slide Decks	Improve your presentation slide decks with the							
3				https://moz.com/blog/8-rules-for-exceptional-slide-presentations...	8 Rules for Exceptional Slide Presentations - Whiteboard Fri...	In last week's Whiteboard Friday, we talked abo							
4				https://moz.com/blog/accepting-mozcon-ignite-pitches	We Want Your Stories: Accepting MozCon Ignite Pitches	Announcing a new Tuesday night event at Moz!							
5				https://moz.com/blog/a-sneak-preview-for-mozcon-2015	A Sneak Preview of #MozCon 2015	Get a preview of MozCon - July 13-15th in Seatl							
6				https://moz.com/blog/category/advanced-seo									
7				https://moz.com/blog/category/branding									
8				https://moz.com/blog/category/competitive-research									
9				https://moz.com/blog/category/events									
10				https://moz.com/blog/category/search-engines									
11				https://moz.com/blog/digital-marketing-from-the-agile-manifesto	What Digital Marketing Managers Should Learn from the Agil...	The Agile movement has done a world of good							
12				https://moz.com/blog/eye-tracking-in-2014-how-users-view-and-l...	Eye Tracking in 2014: How Users View and Interact with Tod...	With Google's SERPs regularly testing new fea							
13				https://moz.com/blog/eye-tracking-in-2014-how-users-view-and-l...									
14				https://moz.com/blog/google-tracking-in-2014-how-users-view-and-l...	How Google Pulls Structured Snippets from Websites' Tables	There are many pages on the Web that are fille							
15				https://moz.com/blog/hiring-seo-how-to-hire-someone-with-no-e...	Hiring for SEO: How to Find and Hire Someone with Little or ...	SEOs are in high demand, but that doesn't me							
16				https://moz.com/blog/i-cant-drive-155-meta-descriptions-in-2015	I Can't Drive 155: Meta Descriptions in 2015	For years, we've suggested a 155-character lin							
17				https://moz.com/blog/keyword-research-for-the-modern-custome...	Keyword Research for the Modern Customer Journey	As searcher behavior gets more complicated, i							

Schema.org

Per fare scraping del markup di Schema.org dentro l'HTML usa questi filtri:

```
(//*[@itemtype])[1]/@itemtype
(//*[@itemtype])[2]/@itemtype
...
```

Extract selected elements of internal HTML pages. See our [User Guide](#) for examples. The results can be seen in the Custom Tab in the Extraction filter.

Item Type 1	XPath	Filter	Action
X	//*[@itemtype][1]/@itemtype	X	Extract Inner HTML
X	//*[@itemtype][2]/@itemtype	X	Extract Inner HTML

Così è come appaiono i dati raccolti:

File Configuration Mode Bulk Export Reports Sitemaps Licence Help

www.plumbworld.co.uk

Resume

Internal	External	Protocol	Response Codes	URI	Page Titles	Meta Description	Meta Keywords	H1	H2	Images	Directives	AJAX	Custom
Filter: Extraction Export													
				Address	Status Code	Status	Item Type 1	Item Type 2					
1				http://www.plumbworld.co.uk/greenwood-airvac-affresco-1943-18629	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
2				http://www.plumbworld.co.uk/affine-vienne-bathroom-suite-18579-30284	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
3				http://www.plumbworld.co.uk/600mm-x-800mm-single-panel-18646-22942	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
4				http://www.plumbworld.co.uk/mira-beat-shower-slide-bar-17928-1244159	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
5				http://www.plumbworld.co.uk/home-of-ultra-push-button-18793-30075	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
6				http://www.plumbworld.co.uk/mira-beat-shower-slide-bar-17928-1244158	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
7				http://www.plumbworld.co.uk/essentials-acrylic-bath-panel-2089-30644	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
8				http://www.plumbworld.co.uk/essentials-acrylic-bath-panel-2089-30645	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
9				http://www.plumbworld.co.uk/xpelair-lv100-low-voltage-1940-16386	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
10				http://www.plumbworld.co.uk/architect-square-single-spray-1734-17423	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
11				http://www.plumbworld.co.uk/chrone-chrome-rolux-pull-out-326-25448	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
12				http://www.plumbworld.co.uk/600mm-x-1100mm-single-panel-18646-22945	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
13				http://www.plumbworld.co.uk/greenwood-airvac-r90bw-2392-17145	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
14				http://www.plumbworld.co.uk/affine-touton-bathroom-suite-18579-30275	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
15				http://www.plumbworld.co.uk/affine-touton-bathroom-suite-18579-30274	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
16				http://www.plumbworld.co.uk/park-lane-bond-bathroom-suite-18727-29644	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
17				http://www.plumbworld.co.uk/vasari-novara-gloss-white-wall-17875-26946	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
18				http://www.plumbworld.co.uk/vasari-calvados-straight-bath-2089-22086	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
19				http://www.plumbworld.co.uk/500mm-x-800mm-radiator-18647-22913	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					
20				http://www.plumbworld.co.uk/600mm-x-1600mm-radiator-double-18647-22961	200	OK	http://schema.org/OfferItemCondition	http://schema.org/Product					

is -

Indirizzi Email

Ebbene sì, Screaming Frog può anche fare scraping di tutti gli indirizzi email che trova scansionando le pagine del sito. Per impostare il filtro basta usare questo Xpath:

```
//a[starts-with(@href, 'mailto')][1]
//a[starts-with(@href, 'mailto')][2]
...
```

Extract selected elements of internal HTML pages. See our [User Guide](#) for examples. The results can be seen in the Custom Tab in the Extraction filter.

Email 1	XPath	Filter	Action
X	//a[starts-with(@href, 'mailto')][1]	X	Extract Inner HTML
X	//a[starts-with(@href, 'mailto')][2]	X	Extract Inner HTML

I risultati appaiono come segue:

Internal	External	Protocol	Response Codes	URI	Page Titles	Meta Description	Meta Keywords	H1	H2	Images	Directives	AJAX	Custom
				Address	Status Code	Status	Email 1	Email 2					
1				http://www.screamingfrog.co.uk/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
2				http://www.screamingfrog.co.uk/meet-the-team/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
3				http://www.screamingfrog.co.uk/link-building/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
4				http://www.screamingfrog.co.uk/accreditations/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
5				http://www.screamingfrog.co.uk/seo-spider-3-0/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
6				http://www.screamingfrog.co.uk/blog/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
7				http://www.screamingfrog.co.uk/disclaimer/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
8				http://www.screamingfrog.co.uk/category/screaming-frog/screaming-frog-seo-spider/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
9				http://www.screamingfrog.co.uk/how-to-schedule-a-crawl/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
10				http://www.screamingfrog.co.uk/category/blog/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
11				http://www.screamingfrog.co.uk/screaming-frog-seo-spider-update-version-2-30/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
12				http://www.screamingfrog.co.uk/seo-spideruser-guide/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
13				http://www.screamingfrog.co.uk/search-awards-2014/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
14				http://www.screamingfrog.co.uk/screaming-frog-seo-spider-update-version-2-55/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
15				http://www.screamingfrog.co.uk/category/events/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
16				http://www.screamingfrog.co.uk/screaming-frog-seo-spider-update-version-1-60/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					
17				http://www.screamingfrog.co.uk/search-engine-optimisation/	200	OK	info@screamingfrog.co.uk	support@screamingfrog.co.uk					

Prerender

Per estrarre l'URL indicato nel link Prerender imposta la seguente custom extraction:

```
//link[@rel="prerender"][1]/@href
//link[@rel="prerender"][2]/@href
//link[@rel="prerender"][3]/@href
//link[@rel="alternate"][...]/@hreflang
```

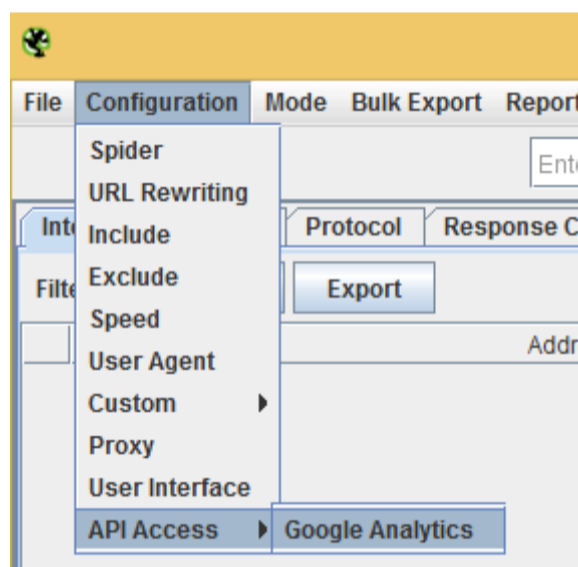
Collegare Google Analytics

Dalla versione 4.0 di Screaming Frog è possibile scaricare via API i dati di navigazione per ogni singolo URL. Finalmente diciamo basta a faticosi merging con vlookup del database Screaming Frog con il database di Google Analytics :D

Come collegare Google Analytics

Per collegare Screaming Frog a Google Analytics devi cliccare nel menu "Configuration" e poi su "API access". Nella nuova schermata clicca su "connect account" per essere redirezionato sulla pagina di conferma di Google per accettare la condivisione dei dati con l'applicazione. Ora è disponibile il pannello delle opzioni di Google Analytics dove puoi impostare tra le altre cose, il range di date, le metriche e le dimensioni da mostrare.

A questo punto basta scegliere l'account, la proprietà e la visuale del profilo Analytics del sito che stiamo scansionando ed il gioco è fatto. I dati sono mostrati nella tab dedicata "Analytics".



Come collegare Screaming Frog e Google Analytics

Spero che questa guida ti abbia aiutato a comprendere meglio le potenzialità di Screaming Frog, io non potrei farne a meno! Hai suggerimenti che vorresti aggiungere? Lascia un commento e verrà inserito nella guida e **se le info ti sono servite condividi questa pagina sui social!**

Richiedi ora un [SEO Audit](#) per il tuo sito, compila il [form di contatto](#) e riceverai un preventivo rapido e specifico per il tuo business!

Articoli correlati che potrebbero interessarti:

- [SEO Checklist: i fattori SEO da Ottimizzare \(29.4\)](#)
- [Come organizzare un SEO Audit \(21\)](#)
- [SEO e Prestashop, tutto ciò che devi sapere \(20.7\)](#)
- [Come verificare la correttezza di una Sitemap.xml \(20.1\)](#)
- [Guida base all'ottimizzazione per i motori di ricerca \(18.3\)](#)
- [Screaming Frog vs Visual SEO, intervista a Federico Sasso \(18.1\)](#)
- [Guida a Google Search Console \(15.5\)](#)
- [Guida alla SEO tecnica per siti mobile \(14.8\)](#)

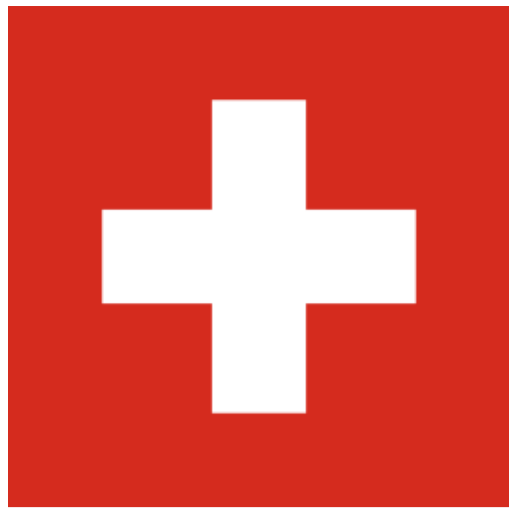
FILED UNDER: POSIZIONAMENTO MOTORI DI RICERCA

TAGGED WITH: ANALISI SEO , CRAWLING , CSS , DATI STRUTTURATI , GOOGLE DOC , JAVASCRIPT , REGEX , SCHEMA.ORG , SCRAPING , SCREAMING FROG , SEO TOOLS , USER-AGENT

AGENZIA PARTNER IN SVIZZERA

Searcus Swiss Sagi agenzia di [Consulenza SEO a Lugano](#) dal 2009.

EVE Milano collabora con Searcus per progetti **SEO multi lingua**.
Cerchi una agenzia SEO in Ticino? Ti consiglio Searcus.ch e avremo modo di lavorare assieme. Abbiamo una profonda



esperienza della SEO e Google AdWords e conosciamo il mercato Svizzero ed Europeo.


About Giovanni Sacheli

[Giovanni Sacheli](#) è **consulente SEO** e SEM per Searcus Swiss Sagl ed EVE Milano. Professionista Certificato Google AdWords e grande appassionato di **analisi SEO** tecniche ed ottimizzazioni di siti eCommerce multilingua. Dal 2009 condivide su questo blog le sue esperienze e le nozioni tecniche più interessanti per posizionare siti web su Google.


Commenta con Google+

Questo articolo è stato pubblicato prima della migrazione su HTTPS avvenuta il 1/1/2016. Questi commenti riguardano l'URL vecchio, senza HTTPS. Non è possibile aggiungere altri commenti con Google+, usa i commenti nativi di WordPress.

49 commenti Google+


 Aggiungi un commento come Giovanni Sacheli


Commenti più popolari 🌐 🔗


 **Giovanni Sacheli** tramite Google+ 2 anni fa - [Condivisione pubblica](#)
Partendo dalla guida di SEER Interactive ho ampliato e approfondito gli aspetti di analisi SEO per fornire la **prima Guida Completa di Screaming Frog in Italiano**

+17 +1 · Rispondi


Visualizza tutte le risposte (12) ▾


 **Giovanni Sacheli** 2 anni fa (modificato)
Quindi **+Giorgio Taverniti** mi confermi il requisito di 9.000 parole per articoli in-depth? Le uniche info ufficiali che trovo non parlano di lunghezza, quanto più di markup <http://googlewebmastercentral.blogspot.it/2013/08/in-depth-articles-in-search-results.html>


 **Giorgio Taverniti** 2 anni fa
Ah no, era una battuta in base a quello che tu avevi detto. Non conosco il limite, mi spiace :(

 **Giovanni Sacheli** 2 anni fa - [SEO Italia \(SEO Audit & planning\)](#)
Guida in Italiano a Screaming Frog

+5 +1 · Rispondi


 **Marco Garulli (mitico185WebMaster)** 2 anni fa +7
+Giovanni Sacheli ciao scusami il se ti disturbo ho visto che hai postato questo programma io l'ho scaricato ma prima di comprarlo volevo sapere se esiste una guida in Italiano e ancora meglio dei tutorial

 **Marco Garulli (mitico185WebMaster)** 2 anni fa +7
+Giovanni Sacheli hahhahahhah scusami non ho cliccato dopo ho visto che hai fatto la guida troppo pigro in questi giorni sarà il **#natale**


 **Giovanni Sacheli** tramite Google+ 2 anni fa - [Condivisione pubblica](#)
6.5 Come aumentare la memoria allocata a Screaming Frog


Windows 32 & 64-bit: Prima di tutto, se hai una macchina a 64-bit, assicurati di aver scaricato ed installato la versione di Java a 64-bit, altrimenti non potrai allocare più memoria di una macchina a 32-bit. Apri la cartella dove hai installato lo spider

+4 +1 · Rispondi


 **Giovanni Sacheli** 2 anni fa - [SEO Italia \(Discussione\)](#)
Guida in Italiano a Screaming Frog

+3 +1 · Rispondi


 **Martino Mosna** tramite Google+ 2 anni fa - [Condivisione pubblica](#)

 **Giovanni Sacheli** ha condiviso inizialmente questo post
Partendo dalla guida di SEER Interactive ho ampliato e approfondito gli aspetti di analisi SEO per fornire la **prima Guida Completa di Screaming Frog in Italiano**


+5 +1 · Rispondi

 **Giovanni Sacheli** 2 anni fa

 **+Gianmaria Allisiardi** grazie correggo subito!

 **Diego D. De Zan** 2 anni fa
Visto...interessante...Grazie ad entrambi.


 **Andrea Scardino** tramite Google+ 2 anni fa - Condivisione pubblica
Proprio l'altro giorno a [#ConvegnoGT](#) mi chiedevo come mai non ci fosse una buona guida in italiano di Screaming Frog. Ed eccola qui. Complimenti **+Giovanni Sacheli**


 **Giovanni Sacheli** ha condiviso inizialmente questo post
Partendo dalla guida di SEER Interactive ho ampliato e approfondito gli aspetti di analisi SEO per fornire la **prima Guida Completa di Screaming Frog in Italiano**

+4  Rispondi

Visualizza tutte le risposte (6) ▾


 **Maurizio ZioPal Palermo** 2 anni fa +2
è una risorsa di quelle che valgono complimenti a **+Giovanni Sacheli** che l'ha fatta e **+Andrea Scardino** che l'ha proposta

 **Andrea Scardino** 2 anni fa +1
+Giovanni Sacheli fammi sapere quando farai la prossima revisione. Non me la perderò. E complimenti ancora.


 **Argoserv** tramite Google+ 2 anni fa - Condivisione pubblica
Screaming Frog: la guida completa

+1  Rispondi


 **Stefano Rigazio** tramite Google+ 2 anni fa - Condivisione pubblica

 **Giovanni Sacheli** ha condiviso inizialmente questo post
Partendo dalla guida di SEER Interactive ho ampliato e approfondito gli aspetti di analisi SEO per fornire la **prima Guida Completa di Screaming Frog in Italiano**


+6  Rispondi

 **Carlo Rubino** tramite Google+ 2 anni fa - Condivisione pubblica
Guida completa di Screaming Frog in italiano!


+1  Rispondi

 **Alessandro Romano** tramite Google+ 2 anni fa - Condivisione pubblica
mo Articolo


+1  Rispondi

 Condiviso da **Patrizia Frattini** tramite Google+ 2 anni fa - Condivisione pubblica

+1  Rispondi

 **Gianluca Punzi** tramite Google+ 2 anni fa - Condivisione pubblica
[#seo](#)

+1  Rispondi

 Condiviso da **Davide Bove** tramite Google+ 2 anni fa - Condivisione pubblica


+1  Rispondi


 Condiviso da **Carlo Casciaro** 2 anni fa - Condiviso in privato

+1  Rispondi


 Condiviso da **fiorica gasparini** tramite Google+ 2 anni fa - Condiviso in privato


+1  Rispondi


 **Maurizio ZioPal Palermo** tramite Google+ 2 anni fa - Condivisione pubblica
La risorsa è di quelle che valgono ;)

 **Giovanni Sacheli** ha condiviso inizialmente questo post
Partendo dalla guida di SEER Interactive ho ampliato e approfondito gli aspetti di analisi SEO per fornire la **prima Guida Completa di Screaming Frog in Italiano**


+7  Rispondi


 **Luca De Berardinis** 2 anni fa
Mooooooooolto interessante!
· Traduci

 **Mario Carabotta** tramite Google+ 2 anni fa - Condivisione pubblica
Uau, ce n'è da imparare. Bel regalo di natale!

 **Giovanni Sacheli** ha condiviso inizialmente questo post
Guida in Italiano a Screaming Frog

+3  Rispondi

 **Massimo Fattoretto** tramite Google+ 2 anni fa - Condivisione pubblica
La guida compelta di Screaming Frog in Italiano by **+Giovanni Sacheli**

 **Giovanni Sacheli** ha condiviso inizialmente questo post
Partendo dalla guida di SEER Interactive ho ampliato e approfondito gli

aspetti di analisi SEO per fornire la **prima Guida Completa di Screaming Frog in Italiano**

+2  Rispondi



Cristian Brunori tramite Google+ · 2 anni fa · Condivisione pubblica
Grazie a **+Giovanni Sacheli** per questa risorsa!



Giovanni Sacheli ha condiviso inizialmente questo post
Partendo dalla guida di SEER Interactive ho ampliato e approfondito gli aspetti di analisi SEO per fornire la **prima Guida Completa di Screaming Frog in Italiano**

+1  Rispondi



Andrea Monti tramite Google+ · 2 anni fa · Condivisione pubblica
Ottima guida di **#screamingfrog**! Complimenti **+Giovanni Sacheli**

+1  Rispondi

[Mostra più cerchie](#)

COMMENTS

Mario

[20/02/2015 at 12:43](#)

[\(Edit\)](#)

Buon giorno, complimenti e grazie per l'articolo. Volevo chiederti però cosa intendi tu per tempi lunghi di caricamento pagina? mi spigo meglio scusa, quale il tempo massimo di caricamento pagina che un sito non dovrebbe mai superare secondo la tua esperienza. Es: 0,80 0,90 e i siti che superano tali limiti subiscono penalizzazioni? Grazie di cuore

[Rispondi](#)

Giovanni Sacheli

[20/02/2015 at 12:52](#)

[\(Edit\)](#)

Ciao Mario,

un page load time buono è sotto i 3-4 secondi, se invece ti riferisci ai tempi del crawler (indicati anche in GWMT – statistiche di scansione) allora è meglio stare sotto il secondo. Però preferisco rispondere linkandoti alcuni video ufficiali di Matt Cutts che parlano dell'argomento:

<http://youtu.be/B3zmP0W26M0>

<http://youtu.be/muSizHum4U>

<http://youtu.be/oG7y87SVG6A>

<http://youtu.be/nndi1mxl4vE>

<http://youtu.be/slyM7NFIDWM>

<http://youtu.be/-l4rWnQxxkM>

<http://youtu.be/a-9pCfyYPdQ>

[Rispondi](#)

Mario

[20/02/2015 at 14:14](#)

[\(Edit\)](#)

Grazie Giovanni

[Rispondi](#)

Virgilio

[24/02/2015 at 14:45](#)

[\(Edit\)](#)

Ciao ragazzi avrei un problema, dal web master tool risulta che un mio sito riceve link da un altro mio sito, e io non voglio (forse ci sono capitati per errore). Come faccio a sapere quali sono le pagine dove stanno i link, se faccio un analisi con Screaming Frog mi mostra dove stanno i link in modo che posso eliminarli?

[Rispondi](#)

Giovanni Sacheli

[24/02/2015 at 14:57](#)

[\(Edit\)](#)

Ciao Virgilio, esegui la scansione con SF e quando termina passa nella tab superiore "external". Individua il link verso il tuo sito (con il filtro è veloce) e cliccalo, nella tab inferiore trovi "Inlinks". Dentro Inlinks trovi tutte le pagine che contengono quel link.

:D buona giornata!

[Rispondi](#)

Virgilio

[24/02/2015 at 16:45](#)

[\(Edit\)](#)

Grazie Giovanni

Virgilio

[18/03/2015 at 08:14](#)

[\(Edit\)](#)

Buon giorno Giovanni, la scansione del mio sito si blocca al 80% che faccio? da premettere ho già preso tutte le precauzioni, ho deselezionato Images, CSS, JavaScript, SWF e ho aumentato la memoria del mio computer un i5 con 10 GB di ram, che altro posso fare? tu nel art sopra consigli di dividere il sito, ma così facendo non rischio di perdere eventuali duplicazioni di title?. Grazie di tutto

Giovanni Sacheli

[18/03/2015 at 09:32](#)

[\(Edit\)](#)

Ciao Virgilio, alcune domande:

- 1) a quanti URL si ferma Screaming Frog? Con 10 GB di RAM allocati dovresti arrivare tranquillamente a scansionare 20-25.000 URL (con immagini, CSS, JS & Co.) oppure circa 200.000 di soli URL HTML
- 2) nella guida spiego come verificare che la memoria sia effettivamente allocata a SF, hai verificato?
- 3) SF è un software in Java che NON archivia i dati in un database ma in memoria, questo lo rende poco efficiente RAM-parlando. Se hai un sito con centinaia di migliaia o peggio ancora milioni di URL è impensabile usare questo software su PC con dotazione standard. Per siti molto estesi ci sono crawler online, ti elenco i migliori:

<http://80legs.com/>

<https://www.deepcrawl.com/>

Dividere il sito è una soluzione, poi comunque dovrai esportare i dati e riunirli in Excel per evitare, come dici tu, di perdere dei dati come titoli duplicati. Se dividi il sito in sezioni l'analisi con Excel diventa obbligatoria perchè non potrai più basarti solo sui dati mostrati da SF essendo dati parziali.

Virgilio

[18/03/2015 at 10:00](#)

[\(Edit\)](#)

Giovanni sei un pozzo di scienza, grazie mille

[Rispondi](#)

Giovanni Sacheli

[18/03/2015 at 10:03](#)

[\(Edit\)](#)

eheheh grazie :D ma non mi hai detto se hai risolto e come!

[Rispondi](#)

Virgilio

[18/03/2015 at 10:59](#)

[\(Edit\)](#)

Voglio provare a dividere il sito, i servizi che mi hai consigliato sono troppo cari. Ti terrò aggiornato. A dimenticavo la scansione è arrivata a 676.121 url

Giovanni Sacheli

[18/03/2015 at 11:23](#)

[\(Edit\)](#)

676.121 URL? Mezzo milione di URL non sono pochi con 10 GB RAM :D

Virgilio

[18/03/2015 at 12:38](#)

[\(Edit\)](#)

Quindi l'unica soluzione senza spendere altri soldi è excel giusto?

Giovanni Sacheli

[18/03/2015 at 12:49](#)

[\(Edit\)](#)

eh si, scansioni settoriali e merging dei database in Excel :)

Virgilio

[18/03/2015 at 12:46](#)

[\(Edit\)](#)

Si è un sito grosso ma ha problemi di indicizzazione, pensa che dal wmt risultano indicizzate solo 1.555 URL in 6 mesi circa il dominio è anziano 5 anni. Devo trovare il problema

[Rispondi](#)

Michela

[31/03/2015 at 09:18](#)

[\(Edit\)](#)

Buon giorno, ottimo art. bravo, bravo. Potrei avere una delucidazione? io ci provo. Scansionando il mio sito mi ritrovo con una miriade di 302 found tutti gravatar.com ma è normale? Giovanni SOS

[Rispondi](#)

Giovanni Sacheli

[31/03/2015 at 09:27](#)

[\(Edit\)](#)

Ciao Michela, gravatar è il servizio che associa le foto alle email nei commenti. È un servizio esterno al tuo blog quindi non mi preoccuperei.

[Rispondi](#)

Michela

[31/03/2015 at 09:40](#)

[\(Edit\)](#)

Grazie giovanni, quindi lascio tutto così, anche se sono tanti 302?

Giovanni Sacheli

[02/04/2015 at 16:30](#)

[\(Edit\)](#)

Ciao Michela, se sai come entrare nel template e correggere i link verso Gravatar puoi anche fixare. Clicca il link attuale sugli avatar, controlla cosa cambia nel path dell'URL dal vecchio al nuovo e dovresti risolvere :D

Jedi82

[25/04/2015 at 22:51](#)

[\(Edit\)](#)

Ciao Giovanni, che guida!!!!!! Sei un grande! Provo a chiederlo anche a te: provando a scansionare il mio sito con il programma, mi restituisce sempre questo errore:

Status code 406 Status Not Acceptable

Ma come cavolo è possibile? Ho anche contattato l'assistenza del mio host ma non hanno trovato una soluzione.,,

Grazie!!!!

[Rispondi](#)

Giovanni Sacheli

[27/04/2015 at 14:32](#)

[\(Edit\)](#)

Grazie Jedi :D

Dato che Screaming Frog non definisce niente nell'intestazione HTTP come "accept headers", l'errore 406 probabilmente è causato da errata configurazione del webserver.

Tecnicamente l'errore dice: la risorsa richiesta è solo in grado di generare contenuti non accettabili secondo la header Accept inviato nella richiesta.

Ma anche via browser ricevi 406? Prova a disabilitare qualche plugin a turno.

[Rispondi](#)

Giuseppe

[01/08/2015 at 20:24](#)

[\(Edit\)](#)

Salve Giovanni,

per prima cosa mi sento in dovere di ringraziarti per quanto condiviso!

Sono alle prime armi tuttavia so riconoscere la qualità di un contenuto e sia questo che le tue altre guide sono fantastiche!!

Passo ora con la domanda!

Quando provo ad eseguire il passaggio 1.2 "Avviare il crawling dell'intero sito web" nel momento in cui vado su Configurations/Spider/Basic non mi è possibile abilitare nessuna voce compreso "Crawl All Subdomains".

Mi compare il messaggio "These options are only available to licenced users".

Cosa dovrei fare per abilitarlo?

E nel caso si facesse crawling di un competitor come è possibile (se possibile) abilitare tali voci?

Ringraziandoti ancora ti saluto

[Rispondi](#)

Giovanni Sacheli

[02/08/2015 at 19:57](#)

[\(Edit\)](#)

Ciao Giuseppe, grazie per il commento!

“ These options are only available to licenced users

vuol dire che la funzione che cerchi è disponibile soltanto nella versione a pagamento di Screaming Frog.

Come sai quel programma, nella versione gratuita, è limitato. Per fruire di tutte le sue funzioni ti serve la licenza.

Spero di averti aiutato, buona serata!

[Rispondi](#)

Giuseppe

[04/08/2015 at 12:45](#)

[\(Edit\)](#)

Ti ringrazio Giovanni, pensavo che la versione free permettesse di fare qualcosina in più.

Un saluto e grazie per il materiale messo a disposizione!!

[Rispondi](#)

luketto

[21/09/2015 at 15:55](#)

[\(Edit\)](#)

Salve Giovanni e complimenti per l'ottima guida, avrei una domanda: ho modificato il file hosts della workstation per visualizzare un sito web da analizzare che sta subendo un restyling sia grafico che di architettura informativa. Se lancio la scansione, Screaming Frog cosa mi analizza? il dominio vecchia versione o interpreta la modifica la mio file hosts e mi scansiona la nuova versione? grazie

[Rispondi](#)

Giovanni Sacheli

22/09/2015 at 12:14

[\(Edit\)](#)

Ciao Luketto, grazie per il commento.

In teoria se modifichi correttamente il file host, Screaming Frog dovrebbe navigare sul sito di test – ovvero sull'IP segnato nel file HOST.

[Rispondi](#)

luketto

22/09/2015 at 12:59

[\(Edit\)](#)

Perfetto, grazie, tutto chiaro!

marco

12/07/2016 at 17:45

[\(Edit\)](#)

Grazie Giovanni per l'ottima guida.

Ho un problema: ho avviato screaming frog per analizzare un sito su protocollo https, appena avvio mi scarica solo l'home page con risposta: connessione rifiutata. Ho provato anche a cambiare vari user agent, ma con lo stesso risultato, cosa può essere secondo te? c'è qualche impostazione da fare per https?

Con xenu riesco a spiderizzarlo

Grazie dell'aiuto

[Rispondi](#)

Giovanni Sacheli

12/07/2016 at 18:19

[\(Edit\)](#)

Ciao Marco grazie per il commento.

Screaming Frog può essere bloccato via robots.txt o da redirezioni sull'user-agent a livello di http-header.

Il robots.txt blocca SF?

Se hai già provato a lanciare la scansione con diversi user-agent senza successo, allora il problema potrebbe essere lato server. Ad esempio il server potrebbe richiedere i cookies per fornire le pagine, in questo caso ti basta spuntare l'opzione "accetta cookies". Un altro esempio: i link del sito web sono in JavaScript e Screaming Frog non vedendoli si ferma alla homepage.

L'errore che ha scritto "connessione rifiutata" è della serie 5xx. La serie di status code 5xx è relativa ad errori sul web server, hai controllato il file .htaccess?

[Rispondi](#)

LASCIA UN COMMENTO.

[Autenticato come Giovanni Sacheli](#) · [Uscire?](#)

Commento

COMMENTO ALL'ARTICOLO

CONTATTI

Consulenza SEO e Web Marketing

Via Pannilani 37/D, Como, 22100 - Italia

Telefono: (0039) 339-3668879

Email: info (at) evemilano.com

La società titolare di EVE Milano è [Searcus Swiss](#)

[Sagl](#) 

SEARCH MARKETING

Il core business di EVE Milano è la **consulenza SEO** con focus sull'ottimizzazione tecnica dei siti web per migliorarne il posizionamento nei risultati dei motori di ricerca. Attraverso l'[analisi delle parole chiave](#) riusciamo ad identificare le

le keyword più efficaci per generare traffico naturale. Con l'[analisi dei competitor](#) studiamo le

PARTNER GOOGLE ADWORDS



Sei interessato ad una consulenza per la [gestione campagne AdWords](#) ? Giovanni Sacheli è membro

[Agenzia SEO](#) e SEM specializzata in **Search Marketing**, ti possiamo aiutare con analisi SEO professionali, [local SEO](#) e campagne PPC con Google AdWords e Facebook. Abbiamo anche sviluppato uno specifico [corso SEO](#) e SEM che si adatta alle tue competenze di partenza e ti guida attraverso le più moderne ed efficaci tecniche di [posizionamento sui motori di ricerca](#) .

CONSULENZA SEO E PPC

[Servizi SEO](#)

[Web Marketing](#)

[Pubblicità su Google](#)

[Pubblicità su Facebook](#)

[Blog Aziendali](#)

mosse vincenti dei TOP player più visibili su Google e abbiamo sviluppato molti altri servizi utili a definire strategie di web marketing vincenti. Per ottenere il massimo da Google è necessario un sito web autorevole con una struttura efficiente e noi possiamo aiutarti con specifiche [analisi SEO tecniche](#) e servizi di [link building](#) professionali.

CATEGORIE DEL BLOG

[Guide WordPress](#)

[La Fidanzata del SEO](#)

[Posizionamento Motori di Ricerca](#)

[Search Engine Marketing](#)

[Social Media Marketing](#)

[Web Analytics](#)

del programma di Certificazione AdWords [Google Partner](#) per garantire ai propri clienti competenze specifiche ed una professionalità comprovata. Molte Web Agency offrono AdWords nei loro servizi ma soltanto una minima percentuale è certificata da Google. Richiedi una [consulenza AdWords](#) professionale e certificata!

ARCHIVIO

Archivio

Seleziona mese