

5 June 2005

By:



Cum functioneaza Google Sitemaps ?

Softpedia va explica

Scopul declarat al proiectului Google Sitemap este de a oferi webmasterilor posibilitatea de a informa si directiona motorul de cautare Google prin site-ul propriu. Prin folosirea unui sitemap crawler-ii Google primesc informatii despre structura site-ului, precum si date despre paginile acestuia (o explicatie detaliata poate fi gasita mai jos), ceea ce duce, spune Google, la o indexare mai eficienta. Un alt beneficiu al sitemap-urilor este ca webmasterii pot „forta” indexarea mai rapida a unor pagini prin publicarea lor in sitemap, fara sa mai astepte „vizita” obisnuita a crawler-ilor, tehnica denumita „content pushing”. Atentie insa, sitemap-urile NU vor avea nici un efect direct in ceea ce priveste plasarea paginilor in rezultatele returnate de Google. Faptul ca un site are sitemap nu-i ofera vreun "bonus" la pozitionarea in rezultate. Sitemap-urile se prezinta sub forma unor fisiere tip XML care contin cateva informatii generale despre website si o lista de adrese web (URL-uri) plus informatii adiacente acestora. Numarul de adrese per sitemap nu poate depasi 50.000, iar dimensiunea maxima a sitemap-ului trebuie sa fie maxim 10MB in forma necomprimata (pentru a eficientiza accesul la sitemap-uri, webmasterii le pot comprima folosind utilitarul gzip, obtinand un fisier mult mai mic datorita compresiei bune a continutului de tip text). In caz ca nu va puteti incadra in aceste limite, puteti crea mai multe sitemap-uri care vor fi listate intr-un „sitemap index”. Fiecare pagina web inclusa in sitemap este definita de urmatoarele caracteristici: - Changefreq: frecventa de modificare a continutului paginii - Lastmod: data ultimei modificari - Loc: adresa paginii web - Priority: importanta/prioritatea paginii relativ la celelalte pagini din acelasi website Pentru a usura crearea sitemap-urilor, Google pune la dispozitie utilitarul Sitemap Generator (sitemap_gen.py) scris in limbajul Python. Acest utilitar poate genera sitemap-uri pe baza unor parametri: - prin citirea unui fisier text care contine lista tuturor URL-urilor de inclus in sitemap - prin parcurgerea unei liste de directoare de pe webserver - prin examinarea log-urilor webserverului Utilizatorii avansati pot profita de formatul XML al sitemap-urilor pentru a le genera automat (prin integrarea, de exemplu, a suportului pentru autogenerarea de sitemap-uri in backend-ul de administrare al unui content management system (CMS)). Sitemap-urile pot folosi si „Open Archives Initiative (OAI) protocol for metadata harvesting”, un protocol standard deja popular, asa ca cei care detin sitemap-uri tip OAI-PMH 2.0 le pot trimite direct la Google in aceasta forma. Pentru cei care considera sitemap-urile XML/OAI prea complexe, exista si o solutie simpla, si anume de a oferi Google doar o lista de URL-uri. Pentru mai multe informatii despre protocolul Google Sitemap vizitati adresele:

<http://www.google.com/webmasters/sitemaps/docs/en/protocol.html>

<http://www.google.com/webmasters/sitemaps/docs/en/faq.html>

In final, un lucru demn de luat in seama: proiectul Google Sitemap este in stadiu Beta (testare) iar Google avertizeaza ca nu garanteaza indexarea tuturor paginilor din sitemap-uri. Totusi beneficiile sunt suficient de importante pentru a incuraja webmasterii sa foloseasca sitemap-urile.