Crawl budget: è un problema per te?

come ridurre le scansioni sprecate

Prima capire dove si ferma la scheda, poi correggere la regola

1 Quando il crawl budget è un problema reale

La guida avanzata di Google è rivolta soprattutto a siti con:

oltre 1.000.000
di pagine uniche aggiornate circa ogni settimana

oltre 10.000
pagine uniche aggiornate ogni giorno

ordini di grandezza orientativi, non soglie

Qualche migliaio di schede? Di solito non è la prima causa:
controlla prima accesso, collegamenti, direttive e contenuti

Però una categoria può generare molti più indirizzi dei prodotti:



Scansioni sprecate: varianti senza contenuti utili
contano se intanto le schede importanti aspettano

2 Dove si è fermata la scheda che manca

che cosa trovi	che cosa controllare
Google non la conosce	link seguibili dalle categorie e presenza nella sitemap
Conosciuta, non scansionata	da quanto aspetta, se accade a molte schede, dove vanno le visite
Errore o blocco	risposta del server, protezioni e regole di accesso, prima di tutto
Letta, ma esclusa	direttive, versione principale scelta da Google e contenuto

Solo l’attesa diffusa di schede utili chiama in causa il budget
altre visite non risolvono un’esclusione o una duplicazione

Controllo URL: un test riuscito oggi
non dimostra che la pagina sia già indicizzata

3 Tre strumenti, tre effetti diversi

	Evita le richieste?	Toglie dai risultati?
robots.txt	✓ sì, blocca la scansione	✗ no: l’URL può comparire
noindex	✗ no: Google deve leggerla	✓ sì, la esclude
canonical	✗ non automaticamente	– indica la versione preferita fra copie

robots.txt blocca la lettura del noindex
per escludere una pagina già indicizzata, Google deve potervi accedere

Nuova scansione: da pochi giorni a poche settimane
senza garanzia di inclusione; ripetere la richiesta non accelera

Prima decidi che cosa vuoi ottenere
limitare le visite, togliere una pagina o consolidare copie

Fonti: Google, guide su crawl budget, Controllo URL, robots.txt e canonical

KLC www.klc.it

Fonte: KLC · Crawl budget: è un problema per il tuo sito? Come ridurre le scansioni sprecate
https://www.klc.it/blog/seo/checklist-crawl-budget-e-crawl-waste/