# Crawl Budget

Stand: 04.07.2022

**Crawling gehört zu den wichtigsten Aufgaben der Suchmaschinen, um neue Inhalte zu erkennen und zu indexieren. Die Bots durchforsten das World Wide Web unbemerkt und sind in der Lage, alles Neue zu finden und zu bewerten. Aus der Arbeit der Crawl-Bots ergibt sich das Ranking in den Suchmaschinen. Doch ganz ohne finanziellen Aufwand durch das Crawl Budget funktioniert das Vorgehen nicht.**

## Definition Crawl Budget

Das **Crawl Budget ist ein Guthaben, welches Google jeder Website zuspricht**. Die Höhe ist unterschiedlich und **hängt vom PageRank ab**. Je höher dieser ausfällt, desto mehr Crawl Budget steht für die Website zur Verfügung – damit können die Bots mehr Seiten einer URL unter die Lupe nehmen. Kleinen Websites mit einem geringen PageRank steht weniger Crawl Budget zur Verfügung.

Für den Vorgang ist es wichtig, die Website sorgfältig zu optimieren. Der Wechsel einer Domain oder unterschiedliche Serverfarmen erfordern ein höheres Crawl Budget als beispielsweise das Veröffentlichen neuer Inhalte.

## Abgrenzung Crawl Budget / Index Budget

Neben dem Crawl Budget schreibt Google auch das Index Budget aus. **Der Unterschied besteht darin, dass für das Erkennen von Webseiten ein anderes Guthaben definiert ist**. Damit wird festgelegt, wie viele Webseiten indexiert werden können.

Beide Budget-Arten sind eng miteinander verzahnt. Mithilfe des Crawl Budgets werden Webseiten auf der ersten Ebene ausfindig gemacht, das Deep Crawling wird bestimmt und anschließend dem Index zugeführt – das Index Budget sortiert das Ranking in den SERPs, wendet die gültigen Algorithmen an und leitet die passenden Suchergebnisse an den Index weiter.

URLs, die regelmäßig von den Crawlern durch das Crawl Budget erfasst werden, bleiben anschließend auch dauerhaft im Index der Suchmaschine.

## Auswirkungen auf Websites

Die Inanspruchnahme des Crawl Budgets ist maßgeblich dafür verantwortlich, dass **Websites beständig in den Suchergebnissen von Google auftauchen**. Das Guthaben ist begrenzt und für jede Seite unterschiedlich. Wird also der PageRank bei einer URL höher eingestuft, ist ebenfalls mehr Crawl Budget vorhanden. Der PageRank zeigt die eingehenden Links auf. Google legt dabei großen Wert auf eine natürliche Struktur und ehrlich erworbene Popularität. Dies geschieht jedoch nur, wenn auch die Inhalte dazu einladen, eine Website weiterzuempfehlen.

Webseitenbetreiber sind dazu angehalten, die **Rankingfaktoren so gut wie möglich auf ihrer Website zu berücksichtigen**, damit sich andere Seiten dafür interessieren. Gekaufte Backlinks hingegen führen zu einer Abwertung des PageRank und somit zu einer Minimierung des Crawl Budgets.

## Einflussfaktoren auf das Crawl Budget

**Erkennt Google schädliche Faktoren, wird das Crawl Budget reduziert**. Die Folgen daraus können fatal sein. Die Bots minimieren die Aktivitäten, auf einer Website nach neuen Inhalten zu suchen. Frischer Content bleibt damit unter Umständen unentdeckt und zieht keine neuen User an. Dieser kann zudem nicht über Backlinks weiterempfohlen werden und der PageRank sinkt.

Google selbst hat in einem Blogartikel die Merkmale genannt, die sich negativ auf das Crawl Budget auswirken:

- Duplicate Content: doppelte Inhalte
- gehackte Seiten: schädliche Websites
- Soft Error: Seiten mit falschem Inhalt oder 404-Fehler
- Content von niedriger Qualität: Spam-Inhalte oder Thin Content
- Infinity Spaces: Content, der eine unendliche URL erzeugt
- Facettennavigation & Session-ID: unterschiedliche und undefinierte URLs
- Weiterleitungen: zu viele notwendige Sprünge, die die Crawl-Bots verwirren
- fehlende Redirects: Erreichbarkeit von Inhalten unter einer festgelegten URL

## Optimierung des Crawl Budgets

Um sich für die eigene Website ein möglichst hohes Crawl Budget zu sichern, können im Rahmen der Suchmaschinenoptimierung einige hilfreiche Maßnahmen ergriffen werden.

- Backlinks: vertrauenswürdige Verlinkungen aus einem themenrelevanten Umfeld
- Seitenstruktur: Realisierung einer flachen Klickhierarchie
- Pagespeed: Ladezeiten der Website verbessern

- Serverfehler: sinnvolle Weiterleitungen einrichten, um Fehlermeldungen zu verhindern
- interne Verlinkung: Häufigkeit der Links an Relevanz anpassen
- Seitenausschluss: unwichtige Inhalte mittels robots.txt vom Crawling ausschließen
- Metatags: follow und nofollow überlegt einsetzen
- XML-Sitemap: Datei erstellen und über die Google Search Console einreichen

## Relevanz für die SEO

Mit dem Crawl Budget zeigt sich recht deutlich, **wie eng alle Rankingfaktoren miteinander zusammenhängen**. Hochwertige Inhalte, schnelle Ladezeiten und ein userfreundlicher Aufbau laden dazu ein, die Website weiterzuempfehlen. Nur dann steigt das Crawl Budget und die Bots durchsuchen die Seite nach neuen Inhalten für den Index.

Google ist eine Maschine, die mit den richtigen Anweisungen gefüttert wird. Um Inhalte zu erkennen, müssen **die Signale eindeutig gesetzt werden**. Webseitenbetreiber haben es selbst in der Hand, das Crawl Budget zu erhöhen. Das passiert nur dann, wenn Google die Webseiten verstehen kann. Hochwertige Inhalte müssen schnell erreichbar sein, um sie zu indexieren. Je besser die SEO-Maßnahmen realisiert werden und greifen, desto höher ist auch das Crawl Budget. Und das wiederum wirkt sich positiv auf das Ranking in den Suchergebnissen aus.