
Suchmaschinen Aussperren Robots.txt und Robots Meta Tags

Prof. Dr. Clemens H. Cap

<http://www.internet-prof.de>

Search Engines einladen

Anmeldeseiten

- Suchmaschinen haben Anmeldeseiten
- Eintragung ist möglich, aber selten sinnvoll
- Crawler sind heute sehr mächtig
- Domänen-Anmeldung reicht bereits zur Einladung

**Davon sind aber
nur 2 oder 3 wirklich
von Relevanz !**

Meta-Anmeldeseiten und Anmeldesoftware

- Können nicht wirklich schaden
- Nutzen aber selten dem Aufwand und den Kosten entsprechend

Mit unserer selbstentwickelten und langjährig erprobten Technik positionieren wir Ihre Webseite in Suchmaschinen optimal.

- Regelmäßige Eintragung in über 900 Suchdienste

Ongate GmbH übernimmt die Eintragung in über 900 Suchmaschinen. Das ist vergleichbar mit einer Registrierung im Internet und unerlässlich für den (Such-)Erfolg Ihrer Webseiten. Ihre Webseite wird dadurch automatisch in den Suchergebnissen nach oben steigen. Diese Eintragungen wiederholen wir in regelmäßigen Abständen, sodass die Anzahl Ihrer Verweisseiten steigt und damit Ihre Platzierung in den wichtigsten Suchmaschinen.

Search Engines ausladen

Mögliche Gründe

- Last auf dem Server
- Privacy Gründe

Beachte

- Aussperren einer Suchmaschine erlaubt weiterhin den Zugriff auf die Seiten !
- Zum Sichern des Zugriffs auf eine Seite ist Sperre auf Ebene des Servers nötig
Dazu andere Techniken als die hier beschriebenen einsetzen !

Search Engines aussperren

Der Robot Exclusion File

Datei **robots.txt** ins Rootverzeichnis der Domäne stellen

Beispiel 1:

```
User-agent: BeispielRobot
Disallow: /temp/          # Die Dateien sind sehr kurzlebig
Disallow: /logfiles/     # Die ändern sich jeden Tag
Disallow: /bilder/      # Bilder nicht downloaden
```

Beispiel 2: Vermeide Aufnahme eines bestimmten Bildes in Google Bildersuche

```
User-agent: Googlebot-Image
Disallow: /images/private.jpg
```

Beispiel 3: Vermeide Aufnahme aller Bilder in Google Bildersuche

```
User-agent: Googlebot-Image
Disallow: /
```

Search Engines aussperren

Robot Meta Tags

Spezielle Tags, die im <head> der HTML Seite zu plazieren sind

- **Index:** Seite soll indiziert werden
- **Follow:** Den Links auf dieser Seite soll gefolgt werden
- `<meta name="robots" content="index, follow">`
- `<meta name="robots" content="noindex, follow">`
- `<meta name="robots" content="index, nofollow">`
- `<meta name="robots" content="noindex, nofollow">`

Search Engines aussperren

Googlebot Meta Tags

Einschränkung der Meta Tag Wirkung auf Google

- name = "googlebot" statt name="robots"

Google-spezifische Erweiterungen des Robot Meta Tag

- Bestimmte Tag-Inhalte nur von Google implementiert
- Nur zugänglich, wenn name = "googlebot"

- **Unterdrückung von Snippets**

Snippet: Auszug der Ergebnisseite, der alle Suchbegriffe enthält

Erhöht typischerweise User Click-Rate

```
<meta name="googlebot" content="nosnippet">
```

Weitere Steuerungsmöglichkeiten

Caching / Archivierung

- Anweisung an Suchmaschine, eine Seite nicht zu archivieren / cachen
- `<meta name="robots" content="noarchive">`

Erneuter Besuch

- Empfehlung, wann eine Suchmaschine die Seite erneut betrachten soll
- `<meta name="revisit-after" content="30 days">`