

## Die Datei robots.txt

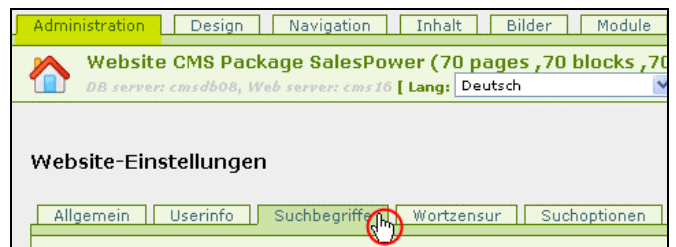
Ein Robot ist ein Computerprogramm, das von Suchmaschinen eingesetzt wird, um das World Wide Web zu durchsuchen und Webseiten zu analysieren. Dabei werden die gefundenen Seiten indiziert und in den Datenbestand der Suchmaschine aufgenommen.

Der Robot (auch Webcrawler oder Spider genannt) liest beim Auffinden einer Website zuerst die Datei robots.txt. In dieser Datei kann festgelegt werden, ob und wie die Website von einem Robot besucht werden darf. Der Website-Betreiber hat so die Möglichkeit, ausgewählte Inhalte generell oder auch nur für bestimmte Suchmaschinen zu sperren.

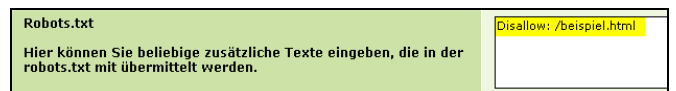
Die robots.txt ist eine einfache Textdatei im Hauptverzeichnis einer Domain, die vom Worldsoft-CMS automatisch generiert wird. Zusätzlich können individuelle Anpassungen vorgenommen werden.

Weitere Informationen zum Thema erhalten Sie unter [http://de.wikipedia.org/wiki/Robots\\_Exclusion\\_Standard](http://de.wikipedia.org/wiki/Robots_Exclusion_Standard)

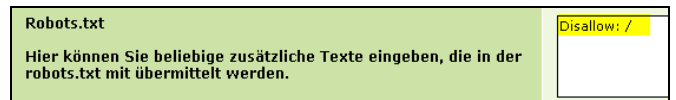
Sie finden die Datei robots.txt in Ihrer Domainadministration unter der Lasche **Administration / Suchbegriffe**.



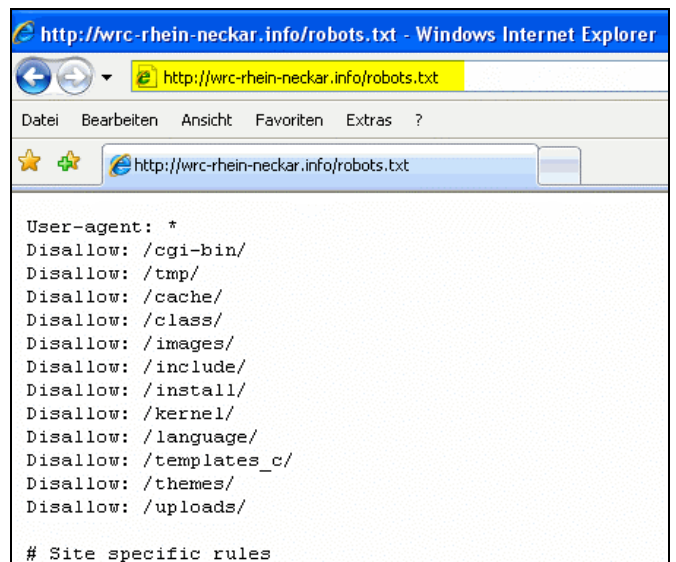
Mit dem Befehl **Disallow: /beispiel.html** würden Sie verhindern, dass die Inhaltsseite beispiel.html von den Suchmaschinen aufgenommen wird.



Der Befehl **Disallow: /** ohne weitere Angabe sperrt die gesamte Website.



Das Worldsoft-CMS generiert automatisch die abgebildete robots.txt. Einige grundlegende Verzeichnisse sind darin gesperrt.



Der Befehl **User-agent: \*** signalisiert, dass die Vorgaben für alle Robots gelten.

Individuelle Sperrungen oder Freigaben lassen sich mit dem Befehl **User-agent: Robot** einstellen. **Der Name Robot muss dabei durch den richtigen Namen ersetzt werden.** Weitere Informationen zum Thema finden Sie hier: [http://de.wikipedia.org/wiki/Robots\\_Exclusion\\_Standard](http://de.wikipedia.org/wiki/Robots_Exclusion_Standard)

Sie können die Datei robots.txt auf Ihrer eigenen Website so aufrufen und kontrollieren: **www.meinedomain.de/robots.txt**

**Viel Spaß und viel Erfolg mit der Datei robots.txt des Worldsoft-CMS!**