

## **Double Content in Webkatalogen und Webverzeichnissen wirkt sich negativ aus – auch für die eigenen Webseiten der eingetragenen Unternehmen**

*Autor: Dirk Günther, Administrator BeChem Webverzeichnis, 14.09.2010*

Unter Double Content (DC) versteht man in diesem Zusammenhang gleichlautende Inhalte, d.h. Texte bzw. Textpassagen, Grafiken, Links oder sonstige Objekte auf verschiedenen Webseiten. Damit ist der Teil der Webseite gemeint, der vom Browser oder Spider interpretiert wird.

Suchmaschinen wie Google benutzen sogenannte Spider, die mehr oder weniger regelmäßig die direkt angemeldeten Webseiten oder diejenigen überprüfen, auf die von irgend einer anderen dem Spider bekannten Seite (auch über Umwege) verlinkt wurde. Diese Überprüfung schließt auch die Prüfung auf DC ein.

Nach der Prüfung erfolgt eine Bewertung, die dazu führt, dass die Seite entweder im Index aufgenommen wird (Voraussetzung zum Vergleich mit Suchanfragen, z.B. Google.-Suche) oder nicht. Nicht aufgenommene Seiten werden natürlich nie gefunden – selbst bei noch so guter Optimierung von URL, Titel, Beschreibung oder Content (Texte). Bei aufgenommenen Webseiten erfolgt trotzdem aber auch eine Art Wertschätzung, die für die Berechnung des Pageranks (PR) verwendet wird. Vermutlich hat eine negative Bewertung auch Einfluss auf die Berücksichtigung bei Suchanfragen. Die dafür angewendeten Methoden, Verfahren, Prüf-routinen oder Algorithmen werden nicht im Detail offengelegt und wohl auch oft geändert.

So kann der Webseitenbetreiber nur durch Erfahrungswerte – die mehr oder weniger auch im Internet verbreitet werden – lernen, was Google & Co. mögen oder nicht. Die u.g. Erfahrungen und Empfehlungen lassen sich aus dem Betrieb des **Themenkataloges BeChem – Bezugsquellen Chemie Pharma Biotechnologie...** ableiten. Nachdem die Katalogeinträge auf DC überprüft wurden bzw. die entsprechenden Kategorienseiten auf eine erfolgreiche Google-Indexierung zeigte sich, dass bei den nicht indexierten Seiten sehr viele Texte genauso in anderen Katalogen oder den eigenen Firmenseiten der eingetragenen Unternehmen vorkommen.

Zunächst einige Angaben zum Double Content im Katalog:

- php-basierte Webseiten mit teilweise gleichen Inhalten (Containern) auf verschiedenen Seiten, etwa Überschrift, Anzeigenwerbung, Textinhalte, die auf mehreren / allen Seiten gleich sein sollen wie Menüleisten etc.
- index.php und ergänzende Seiten direkt unter der Top Level Domain (TLD)
- Seiten für Rubriken und Kategorien mit relativ langen URL in flacher Struktur direkt unter der TLD, alles suchmaschinenfreundliche URL ohne Sonderzeichen und Parameter („sprechende URL“)
- Möglichkeit, Deep Links (URL könnten sehr lang sein) einzutragen
- Beschreibungen wurden bisher nur automatisch auf verbotene Zeichenketten geprüft, etwa URL-Eingaben in dafür nicht vorgesehenen Felder oder auf „Bad Words“ und es erfolgte eine redaktionelle Prüfung / Korrektur lediglich auf Einhaltung der wichtigsten Aufnahmebedingungen hinsichtlich der Thematik sowie auf Sinn, Ausdruck und Rechtschreibung.
- Seiten für leere Kategorien enthalten mehr als 70% DC, da bis auf die Rubrik- und Kategoriebezeichnung alles Andere gleich war.

Nach der redaktionellen Änderung der Texte für die Katalogeinträge mit DC (Überprüfung von Double Content mit Google), hat sich die Google-Indexierung deutlich verbessert, d.h.

Google indexierte mehr Kategorienseiten. Um diesbezüglich ein Ergebnis einer Webseitenänderung festzustellen, muss man allerdings bis 4 Wochen warten.

Am Beispiel Webverzeichnis BeChem wurden folgende Erfahrungen gewonnen:

1. Auch nur teilweise Double Content auf einer Webseite kann sich nachteilig auf die Indexierung bei Google auswirken.
2. Weniger als ca. 30% Double Content stört die Google-Indexierung nicht, wenn der Rest (siehe Pkt. 3 und 4) gut optimiert ist.
3. Die genaue Prozentzahl aus Pkt. 2 hängt von weiteren Faktoren ab, hauptsächlich von Seitentitel, Seitenbeschreibung, URL (auch URL-Länge) und der Qualität der Textinhalte (nicht nur Keyword-Aufzählungen).
4. Lange URL (wie bei Artikelverzeichnissen üblich) mit mehr als 100 Zeichen führen bei der Indexierung oft zum Rauschmiss, wenn Seitentitel, Seitenbeschreibung und Content der Seite nicht wirklich gut sind.

Daraus lassen sich für die Eintragung in Webkataloge / Webverzeichnisse folgende Empfehlungen formulieren:

Vermeiden Sie unbedingt gleichlautende Formulierungen bei der Darstellung Ihres Unternehmens, die schon anderswo im Internet stehen. Double Content schadet nicht nur dem entsprechenden Webkatalog, sondern natürlich auch Ihren eigenen Webseiten bei der nächsten Bewertung durch den Google-Spider. Damit verschlechtern Sie das Auffinden Ihrer eigenen Seiten bzw. der PR und das Google-Ranking (auch für Ihre Seiten) werden schlechter.

Vermeiden Sie die Aufzählung von Stichwörtern / Keywords im Beschreibungstext, dann dadurch erhöht sich die Gefahr der Interpretation als DC.

.

- Ende -