

Searchengines:

Wichtige Meta-Tags für HTML-Dokumente

Wolfgang Nigischer

In diesem Artikel werde ich die Wirkung der Meta-Tags vor allem aus der Sicht der Arbeitsweise einiger Searchengines beleuchten.

Gerade für die Such-Roboter im Internet wurden diese Meta-Tags eingeführt (so richtig losgetreten hat das z.B.: Alta Vista), um eine gewisse themenmäßige Strukturierung der Inhalte seiner Datenbank zu erreichen.

Quellen

- http://www.searchcode.de/artikel/metatags_info.shtml
- dem deutschsprachigen HTML-Papst und Autor von: SELFHTML Stefan Münz, <http://www.teamone.de/selfaktuell/>
- Seiten für Suchmaschinen optimieren: <http://www.ideenreich.com/dr-gef2.shtml>
- Suchfibel: Gefunden werden <http://www.suchfibel.de/6gefundn/>
- sowie von Hinweisen diverser Searchengines und Internet-Provider.
- Meta-Tag-Generatoren finden Sie unter:
 - <http://www.fireball.de/metagenerator.html>
 - <http://www.searchcode.de/metatags.htm>
 - <http://www.riese.net/Metagenerator/index.htm> (sehr gut!) (Trotz Meta-Tags ist es notwendig, sich bei Link-Verzeichnissen (z.B.: aller YAHOO!) sich selbst in richtige Verzeichnis einzutragen!)
 - <http://www.suchfibel.de/gefundn/metaedit.htm>

Wo werden Meta-Tags platziert?

Wie Sie im unteren Beispiel ersehen können, Im Header-Teil (head..../head) des Dokuments.

Da mitunter einige Searchengines nur die ersten 100-250 Zeichen eines Web-Dokuments zur Indizierung benutzen, ist es vorteilhaft die Meta-Tags gleich nach dem <title>-tag zu setzen.

Unten ein Beispiel eines programm-spezifischen Meta-Tag-Generators, man kann diese aber mit jedem Text, oder HTML-Werkzeug einfügen.

```

<META NAME="Title" CONTENT="Titel der Seite">
<META NAME="Author" CONTENT="Wolfgang Nigischer">
<META NAME="Publisher" CONTENT="PC - News">
<META NAME="Copyright" CONTENT="Firma Nigischer">
<META NAME="Keywords" CONTENT="Keyword - Liste (bis zu 1000 Z)>
<META NAME="Description" CONTENT="Beschreibung der Seite">
<META NAME="Abstract" CONTENT="Beschreibung der Seite">
  
```

Wichtig: es dürfen keine Zeilenumbrüche von Hand innerhalb eines Meta-Tags eingefügt werden, da dies mitunter zu Missinterpretationen bei einigen Searchengines führen kann!

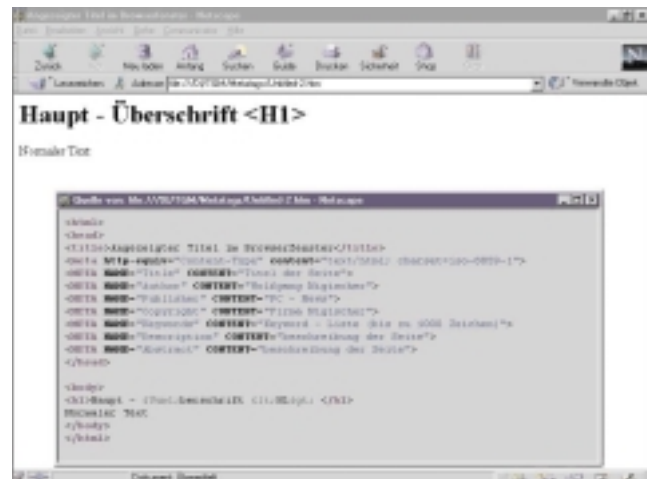
Erst nach diesen Meta-Informationen ist es ratsam, diverse Java, oder andere Scripte einzusetzen.

Der hierbei erzeugte Source-Code, richtig eingetragen, sieht folgendermaßen aus:

```

<html>
<head>
<title>Angenehmer Titel im Browserfenster</title>
<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1">
<META NAME="Title" CONTENT="Titel der Seite">
<META NAME="Author" CONTENT="Wolfgang Nigischer">
<META NAME="Publisher" CONTENT="PC - News">
<META NAME="Copyright" CONTENT="Firma Nigischer">
<META NAME="Keywords" CONTENT="Keyword - Liste (bis zu 1000 Zeichen)">
<META NAME="Description" CONTENT="Beschreibung der Seite">
<META NAME="Abstract" CONTENT="Beschreibung der Seite">
-----
</head>
<body>
<h1>Haupt - &Uuml;rnschrift &lt;H1&gt; </h1>
Normale Text
-----
</body>
</html>
  
```

und im Browser:



Die wichtigsten Meta-Tags

- **META NAME="TITLE" CONTENT="Ihr Text"**
Sinnvollerweise maximal bis zu 8 Worte, möglichst Aussagekräftig!
- **META NAME="AUTHOR" CONTENT="Autor des Dokuments"**
Hier sollte der Inhaltlich verantwortliche Autor stehen. Empfohlen: maximal 64 Zeichen (Beistriche und Leerzeichen zählen mit!).
- **META NAME="PUBLISHER" CONTENT=""**
- **META NAME="COPYRIGHT" CONTENT="Copyrighteintrag"**
Der Name sagt wohl alles... Empfohlen: maximal 64 Zeichen.
- **META NAME="KEYWORDS" CONTENT="eine, Liste, von, Wörtern"**
KEYWORDS sind die eigentlichen Begriffe, wonach man die Seite finden soll. Hier können Sie bis zu 1000 Zeichen (Infoseek) eintragen, wobei die für die WWW-Seite unwichtigeren ab der einhundertsten Stelle gereiht werden sollten, da 1000 Zeichen nicht von jeder Suchmaschine verarbeitet werden (Viele Verarbeiten "nur" 200! Der Rest erhöht nur die Ladezeit!). Die einzelnen Wörter werden mittels Komma getrennt, Phrasen müssen unter "Anführungszeichen" gesetzt werden. Auch ist es sinnvoll, möglichst viele Synonyme hier zu inkludieren. Zum Beispiel, wenn die WWW-Seite etwas über Autos enthält, sollten auch Worte, wie: Fahrzeug, Automobil, Kraftfahrzeug, Personenuagen, PKW, usw. vorkommen, also alles, wonach diese Seite von den Inter-

net-Usern gesucht werden könnte. Sie können auch eventuelle Tippfehler mit berücksichtigen! Online steht hierzu eine Keyword-Datenbank zur Verfügung:

http://www.webmasterplan.com/cgi-local/router.cgi?l=de&p=wmp.de&s=opt_key

Schreiben Sie keine Begriffe, wie *SEX* [eines der häufigsten Suchbegriffe (siehe **PCNEWS-68** Suchwortstatistiken, Seite 65)] u.dgl. in diese Auflistung, da die **KEYWORDS** oft miteinander verglichen werden. Stellt die Suchmaschine fest, dass ein Wort überhaupt nicht in Zusammenhang mit den anderen steht, so wird die Seite zurückgereiht, oder gar nicht in das Verzeichnis aufgenommen!

- **META NAME="DESCRIPTION" CONTENT="Beschreibung der Seite"**

Hier sollte eine aussagekräftige Beschreibung der Seite eingetragen werden, die diese bei richtiger Interpretation bei der Suchmaschine im Ergebnis einer Suche mit angezeigt wird. Auch sollte der Eintrag nicht mehr als 200 Zeichen betragen, und kann auch als Satz formuliert werden.

Grundsätzlich ist es ratsam, Begriffe, die unter **KEYWORDS** verwendet wurden, in **DESCRIPTION** wieder zu verwenden (erhöht oft das Ranking)

- **META NAME="ABSTRACT" CONTENT="..."**

In einigen Suchmaschinen ist **Description** zu **dc.description**, **doc.com** und **abstract synonym**. D.h. alle Informationen darin werden auf ein einziges Feld zusammengelegt.

- **META NAME="ROBOTS" CONTENT="index, follow"**

Hier sind 3 verschiedene Eintragungen möglich:

INDEX: Hier wird nur diese eine WWW-Seite in das Suchregister der Searchengine aufgenommen

FOLLOW: Mit dieser Anweisung veranlassen Sie den Suchroboter, nicht nur die angegebene Seite in sein Suchregister aufzunehmen, sondern auch weiterführende Links, die Sie in Ihrer WWW-Seite aufgelistet haben, weiter zu verfolgen und diese ebenfalls in sein Register aufzunehmen.

NOFOLLOW: Hier werden keine weiteren Links, die sich auf der Seite befinden weiterverfolgt (werden meistens Navigation-Frames sein). Kombinationen, wie im obigen Beispiel sind möglich.

Auch kann: **ALL** für **INDEX** + **FOLLOW**, bzw: **NONE** für **NOINDEX** und **NOFOLLOW** verwendet werden.

- **META http-equiv="expires" content="0"**

Mit diesem Tag veranlasst man eine Seite immer von der Originaladresse zu laden. Sinnvoll für sich häufig ändernde Seiten, oder WWW-Seiten mit Counter (da der Proxy-Server dadurch übergangen werden soll.).

Es versteht sich von selbst, mit diesem Eintrag in der WWW-Seite sparsam vorzugehen!

- **META NAME="REVISIT-AFTER" CONTENT="xx days"**

Revisit veranlasst den Suchroboter alle **xx** Tage auf diese WWW-Seite zurückzukehren.

Auch hier bitte genau überlegen, was man dem Suchroboter antut! Viele Spider folgen daher Ihren eigenen Regeln.

Das gleiche gilt auch für:

- **META http-equiv="pragma" CONTENT="no-cache"**

Veranlasst den Proxy-Agenten, das WWW-Dokument nicht auf den Proxy-Server zu speichern.

- **META http-equiv="Content-Language" content="de"**

Bestimmt die Sprache eines Dokuments (vor allem für Suchroboter mit Suchmöglichkeit nach verschiedenen Landesprachen wichtig)

Wichtig: das z.B. von Frontpage 2000 erstellte: **meta http-equiv="Content-Language" content="de-at"** ist kein reglementierter Bestandteil von HTML 4.0 und wird wahrscheinlich auch von den meisten Searchengines nicht richtig erkannt!

Den richtigen Einsatz der META-Tags können Sie hier online überprüfen lassen:

http://www.webmasterplan.com/cgi-local/router.cgi?l=de&p=wmp.de&s=opt_pre

Abschließend zu den Meta-Tags möchte ich noch sagen, dass natürlich nicht jeder Such-Roboter alle hier vorgestellten Befehle kennt, bzw. auswertet.

Da die HTML-Tags (wie **HTML**, **URL**, **META**,...) jedoch ignoriert werden und die WWW-Seite oft nur die ersten 200 Zeichen gespannt wird, liegen Sie zumindest mit den Begriffen, die Sie in **CONTENT** und **DESCRIPTION** eingetragen haben richtig, denn diese werden sowieso ausgelesen.

Was bezüglich Ranking noch wichtig sein kann

- **TITLE-Tag**

Der Titel einer Seite, eines der wichtigsten Elemente. Kurz, prägnant, aussagekräftig! Der Titel kann unter Umständen beim Ranking hilfreich sein. (Das ist auch jener Teil, der in der Titel-Leiste des Browser-Fensters steht!). Der Titel wird in zahlreichen Suchindizes äußerst hoch gewichtet. Tragen Sie deshalb hier etwas Sinnvolles ein!

- Heben Sie die Wichtigkeit eines Textabschnittes mit Überschriften hervor. So wie im eingangs gezeigten Beispiel. Gestalten Sie Kapitelüberschriften nicht mit **FONT**-Tags, da manche Suchmaschinen auch Überschriften wie **<H1>** für deren Indizes herausfiltern. Suchmaschinen bewerten nicht das Aussehen einer WWW-Seite!

- Beginnen Sie den Text mit einer kurzen prägnanten Einführung und fügen erst später das Bild ein (z.B. Firmenlogo). Oft werden die ersten 200 Zeichen im **BODY**-Tag durchleuchtet!

- Wenn Sie mit den 1000 Suchbegriffs-Wörtern nicht auskommen, schreiben Sie keinen unsichtbaren Text in das WWW-Dokument (=Text in der gleichen Farbe, wie der Hintergrund). Das kann von einigen Suchmaschinenbetreibern als Manipulationsversuch gewertet werden, und die Seite wird "gesperrt".

- Verwenden Sie auch **Keyworte** im **ALT**-Attribut der Bilder.

- **Brückenseiten:** Das sind WWW-Dokumente, welche aus keinen Tabellen, Bildern, Java-Scripten, oder ähnlichen Animationen bestehen, dafür aber für Suchmaschinen leicht zu bearbeiten sind und die alle oben genannten Merkmale enthalten. Diese Seiten verweisen dann mit einem Link auf die eigentliche Seite (nicht **META-REFRESH** verwenden). Der Nachteil dieser Seite ist sicherlich die Optik...

Ich glaube hiermit die wichtigsten Fakten zum Thema **META-Tags** für Suchmaschinen dargebracht zu haben, auch wenn der Artikel gleichzeitig ein Ausflug in die oft debattierte Ranking-Problematik wurde, doch meiner Ansicht nach passiert wird hier noch viel Unsinn produziert.

META-TAGS