

Alles ändert sich!



Zwei Arten von Optimierungen

Realität	Internet
<p>Wie sieht das Geschäft aus?</p> 	<ul style="list-style-type: none"> • Wie soll ich meine Website gestalten? • Flash oder doch „langweiliges“ HTML?
<p>Eintrag in ein Verzeichnis</p>  <p><small>Bürobedarf, Papeterie ÖB&U AG Papeterie, Bürobedarf Papeterie, Bürobedarf Zentrum 47 50101 Witten AG</small></p>	<p>Wo kann ich meine Seite eintragen?</p>  

Optimierung auf der Seite

Realität	Internet
<p>Wie sieht das Geschäft aus?</p> 	<ul style="list-style-type: none"> • Wie soll ich meine Website gestalten? • Flash oder doch „langweiliges“ HTML?

Dokumententyp

Realität	Internet
<p>Wo möchten Sie Ihre Produkte lieber kaufen?</p> <p>In einem Haus oder ...</p>  <p>... in einem Zelt?</p> 	<ul style="list-style-type: none"> • Statische Dokumente sind einfacher zu indexieren und werden deshalb bevorzugt. • Dynamischen Inhalt hinter einer statischen Fassade verstecken (.html anstelle von .jsp). • Ein Format wählen, das von den Suchmaschinen verstanden wird (kein Flash!)

Schlüsselwörter

Realität	Internet
<p>Wo liegt der Schwerpunkt?</p>  <p>Auf welche Kunden bin ich ausgerichtet?</p>	<ul style="list-style-type: none"> • Was tippt ein potenzieller Besucher bei Google ein? (-> In den Besucher versetzen!) • Was ist das Thema meiner Website? (-> Auf ein Thema beschränken!) • Substantive eignen sich für die Beschreibung besonders gut.

Schlüsselwörter

- Kann ich aussagekräftige Schlüsselwörter finden?
- Auf welche Suchmaschine möchte ich optimieren? -> Google normalisiert eine Anfrage nicht (-> Haus, Häuser, Hauses).
- Welche Synonyme existieren?
- Wird ein Begriff oft falsch geschrieben? (-> Begriff falsch schreiben)
- Nur auf eine beschränkte Anzahl Schlüsselwörter optimieren. (-> kein Tante-Emma-Laden)

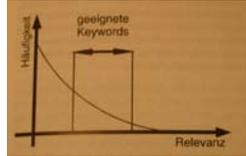
Schlüsselwortdichte

Realität | Internet

- Es gibt eine grosse Auswahl an Karten.
- Karten nehmen einen bedeutenden Teil der Ausstellungsfläche ein.



- Schlüsselwörter sollten gehäuft vorkommen (5%-15% der Anzahl Wörter).
- Schlüsselwörter kommen mit mittlerer Häufigkeit vor.



Titel

Realität | Internet

- Was steht im Schaufenster?
- Wie präsentiere ich mich den Kunden?



- Der Titel (<title>...</title>) ist immer sichtbar und deshalb wichtig.
- Der Titel wird von der Suchmaschine stark bewertet und beim Resultat angezeigt.

[EducETH - Teaching and Learning Unterrichtsmaterialien](#)
 EducETH - eine Plattform für den Austausch von Unterrichtsmaterialien und Informationen im Bereich der Sekundarstufe II ...
[www.educeth.ch/ - 116 - 3 Juni 2004 - Im Cache - Ähnliche Seiten](#)

Titel

- Schlüsselwörter verwenden, die im Dokument vorkommen.
- Für jede Seite einen individuellen Dokumententitel verwenden.
- Schlüsselwörter in der Reihenfolge der Wichtigkeit sortieren.

-> Der Titel ist der wertvollste Platz bei der Optimierung!

Textformatierung

Realität | Internet

Wie sind die Gegenstände im Regal angeordnet?



Text so schreiben, dass die Suchmaschine merkt, welche Stellen wichtig sind.

```
<h2>The Project</h2>
<h3>What is it all about?</h3>
<p>The primary goal of our project is to spread volcanological data and research results, mainly about Stromboli, an island volcano. We would like to give you the feeling of (almost) being part of on-going research on this beautiful and permanently active Italian volcano. In particular we are trying to enrich and stimulate the teaching of Earth Sciences.</p>
```

Wo stehen die Schlüsselwörter?

Realität | Internet

Die Schlüsselwörter (die Karten) sind vom Eingang aus gut sichtbar.



Schlüsselwörter
 •im Titel (<title>)
 •in den h-Tags (<h1>, ..., <h6>)
 •am Anfang der Seite

Zusammengehörende Schlüsselwörter nahe beieinander.

Metainformation

Realität | Internet

- Werbung
- Homepage
- Was erzählt der Besitzer über das Geschäft?



```
<meta content="EducETH - eine Plattform faunlit den Austausch von Unterrichtsmaterialien und Informationen im Bereich der Sekundarstufe II" name="description">
```

[EducETH - Teaching and Learning Unterrichtsmaterialien](#)
 EducETH - eine Plattform für den Austausch von Unterrichtsmaterialien und Informationen im Bereich der Sekundarstufe II ...
[www.educeth.ch/ - 116 - 3 Juni 2004 - Im Cache - Ähnliche Seiten](#)

Wird von Suchmaschinen wegen Missbrauchs nicht mehr bewertet!

Bildbeschreibung

Realität | Internet

Regalbeschriftung bei farbigen Blättern: Kunden sollen Zitronen- und Goldgelb unterscheiden können.



Was ist auf dem Bild zu sehen?
-> Die Suchmaschine kann das Bild „nicht sehen“!

```

```



Interne Linkstrukturen

Realität | Internet

Wo sind die Schulrucksäcke?

Schulrucksäcke im 2. Stock



Das Menu muss für die Suchmaschine verständlich sein:

- keine Javascript-Menus
- keine Imagemaps
- Links wenn möglich als Text
- wichtige Dokumente möglichst hoch im Verzeichnisbaum

Sprachen

Realität | Internet

Nur Dinge zusammen ins Regal stellen, die zueinander passen!



Suchmaschinen erkennen Sprachen und ordnen ein Dokument einer Sprache zu.

-> Nur eine Sprache pro Dokument!

Korrektes HTML

Realität | Internet

Wer kann IKEA-Regale richtig zusammenbauen?



Suchmaschinen müssen HTML interpretieren können.

-> Falsches HTML kann zu Interpretationsfehlern führen!

```
<table rowSpan=3>
  <tr
    <td cellPadding="5*"
      <a href="/informatik/">Informatik
    </a alt="Informatik">
    </td>
  </tr>
</table>
```

Optimierung ausserhalb der Seite

Realität | Internet

Eintrag in ein Verzeichnis



Bürobedarf Papeterie
Oldani AG
Papeterie Bürobedarf
Papeterie Bürobedarf
Zentrstr. 47
5510 Wilen AG

Wo kann ich meine Seite eintragen?

dmolzi open directory project



Namen

Realität | Internet

Namen und Logo einprägsam und einfach!



www.papeterie-oldani.ch
www.papeterie.ch/karten/
www.schreibbedarf.ch

Schlüsselwörter in der URL werden besonders stark gewichtet.

Serververfügbarkeit

Realität | Internet

Wann ist offen?



- Oft nicht verfügbare Seiten sind für die Benutzer nicht attraktiv.
- Wie soll eine Suchmaschine zwischen gelöschten und temporär nicht verfügbaren Seiten unterscheiden?

-> EducETH :-)

Linkpopularität (PageRank)

Realität | Internet

Wo finde ich Hinweise auf das Geschäft?



Die oberste: Papeterie
 Ubbau AG
 Papeterie-Eurobedarf
 Papeterie-Eurobedarf
 Zentralstr. 47
 5610 Wahn AG

Eintrag bei den grossen Verzeichnisdiensten ist wichtig.



Klickpopularität

Realität | Internet

Wie viele Kunden sind im Geschäft?



Wie viele Besucher auf einer Seite?

-> Je mehr Besucher desto besser.

Webspam



Spam oder nicht?

Testfrage:

Ist etwas nur mit dem Ziel erstellt worden, mich zu erreichen und auf ein „Angebot“ aufmerksam zu machen, von dem ich vielleicht gar nichts wissen will?

Spam?

- Fernseh-, Radio und Zeitungswerbung?
- Werbeprospekte?
- Werbefaxe?
- Telefonverkauf?
- Verkauf im Bahnhof?

Unterscheidung zwischen Spam und Optimierung ist fließend und damit eine Gratwanderung!

Spam im Internet

- Nicht angeforderte Emails mit dubiosen Angeboten von Unbekannten.
- Erstellung von Webseiten mit dem alleinigen Ziel, eine Seite in prominente Ränge bei einer Suchmaschine zu bringen.
- Einladung von Fremden beim Instant Messaging.

Wieso gibt es Spam?



Alles ändert sich!



Realität | Internet



@



Google



Zwei Arten von Webspam

Realität | Internet

Dinge versprechen, die nicht gehalten werden können. (-> leere Druckerschachteln)



Texte anbieten, die nichts mit dem Ziel der Webseite zu tun haben (ausser Benutzer anzulocken).

Verteilen von Werbung und Missachten des Stoppklebers.

STOPPI
Keine Werbung einwerfen!
Wir vermeiden Müll.

- Verlinken auf möglichst vielen Seiten.
- Werbung bei Google.

Spamming auf der Seite

Realität | Internet

Dinge versprechen, die nicht gehalten werden können. (-> leere Druckerschachteln)



Texte anbieten, die nichts mit dem Ziel der Webseite zu tun haben (ausser Benutzer anzulocken).

Schlüsselwörter anhäufen

Realität | Internet

Viele Briefumschläge auf einem Haufen.



Schlüsselwörter mehrmals aufführen. -> kein Mehrwert für den Menschen!

```
<h1>Auto Auto Auto BMW Honda
VW Renault Auto Auto PKW PKW
Auto Auto Mercedes Auto
kaufen kaufen kaufen kaufen
kaufen kaufen kaufen kaufen
kaufen</h1>
```

Anhäufen von Schlüsselwörtern verbergen

Realität | Internet

Verteilen Sie die Couverts über den ganzen Laden. Die Anzahl der Couverts ist gleich, für den Kunden ist es aber nicht so einfach erkennbar, dass Sie in jedem Regal 10 Couverts haben.

Schlüsselwörter in einem Text verstecken. -> Klauen den Text eines beliebigen Angebotes und ergänzen ihn mit Schlüsselwörtern.

In hotel the popular car eye, lucerne the hotel most car common lucerne form hotel of spam hotel is that car delivered hotel in e-mail lucerne as a car form of hotel commercial hotel advertising.

Beliebtes Angebot kopieren

Realität | Internet

„Wir sind eine Filiale von MediaMarkt.“



Das Angebot eines beliebten Dienstes kopieren (Yahoo, DMOZ, Wikipedia, ...).

-> Viele Anfragen können auf die eigene Seite umgeleitet werden.

Blindtexte

Realität | Internet

Das Lager wird nur ganz speziellen Kunden gezeigt...



Texte in der Hintergrundfarbe sind nur für Suchmaschinen sichtbar.

Normaler Benutzer:



Suchmaschine:



Tür-Seiten

Realität | Internet

Der geheime Hintereingang



Die Benutzer mit speziell konstruierten Fangseiten auf das eigene Angebot locken. Die Fangseiten werden auf ein Schlüsselwort hin optimiert.



Adressen-Unterscheidung

Realität | Internet

Je nach Investor das Angebot verändern.



Je nach Browserkennung eine andere Seite ausliefern. Suchmaschinen verwenden eigene Browsererkennungen (z.B. Google: googlebot).

Spamming ausserhalb der Seite

Realität | Internet

Verteilen von Werbung und Missachten des Stoppklebers.

STOPPI!
Keine Werbung anwerfen!
Wir vermeiden Müll.

Verlinken auf möglichst vielen Seiten.

Google Bombing

Realität | Internet

Plakate an allen möglichen und unmöglichen Orten aufhängen.



PageRank

Link mit einem Schlüsselwort auf möglichst vielen Seiten.

„miserable failure“



```
<a href="www.whitehouse.gov/president/gubbio.html">miserable failure</a>
```

Links verstecken

Realität | Internet

Unsichtbare Verbindung zwischen verschiedenen Personen.



- Links als Blindtexte (Text hat die selbe Farbe wie der Hintergrund).
- Links mit 1x1-Bildern in der Hintergrundfarbe.

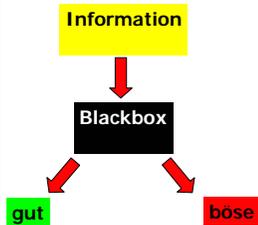
„buy a car lucerne“

- „Automatisches Stehlen“ von anderen Webseiten
- Schlüsselwörter verstecken
- Viele verschiedene „Türseiten“



Gewinnt „Gut“ über „Böse“ oder gibt es eine Lösung des Spamproblems?

Das Problem liegt in der Unterscheidung zwischen guter und böser Information!



Eine solche Blackbox zu entwickeln ist sehr schwierig bis unmöglich!

Was ist gut und böse in welchem Kontext?

-> Es müsste eine Theorie über die Information entwickelt werden!

Kurzfristige Lösung

1. Die Spammer entwickeln eine neue Technik.
2. Die Suchmaschinenbetreiber und andere Institutionen entwickeln eine Gegenmassnahme.

-> Endloser Kampf, der wohl nicht gewonnen werden kann!

Was macht Google gegen Webspam?

Die meisten vorgestellten Spamingverfahren werden von Google erkannt (ausser Googlebombing). Google gibt möglichst wenig Information bekannt, um gegenüber den Spammern einen Schritt voraus zu sein.

Falls Spaming erkannt wird, kann Google eine Seite (oder einen ganzen Server) aus dem Suchindex entfernen.



“Remember the good old days when Spam only came through your computer?”