

Internet und Webseiten-Gestaltung

Wirtschaftsinformatik 3
Vorlesung 2005-05-27
Dipl.-Inf. (FH) T. Mättig

Suchen im Internet

1. Wie findet man Dinge?
2. Und wie findet man Sie wieder?



2004-05-27

InternetundWebseiten-Gestaltung - T.Mättig

3

Suchen im Internet

- Wer bietet Informationen an?
 - Firmen und kommerzielle Organisationen
 - Nicht kommerzielle Organisationen
 - Hochschulen, Universitäten,
Forschungseinrichtungen
 - Medienanbieter
 - Bibliotheken, Archive
 - Privatpersonen
 - ...

Suchen im Internet

- Zwei Möglichkeiten zur Informationsbeschaffung:
 1. Gezielt nachschlagen
 2. Stöbern („surfen“)
 - Sie brauchen in jedem Fall einen Startpunkt, einen „Anfang“.

2004-05-27

InternetundWebseiten-Gestaltung - T.Mättig

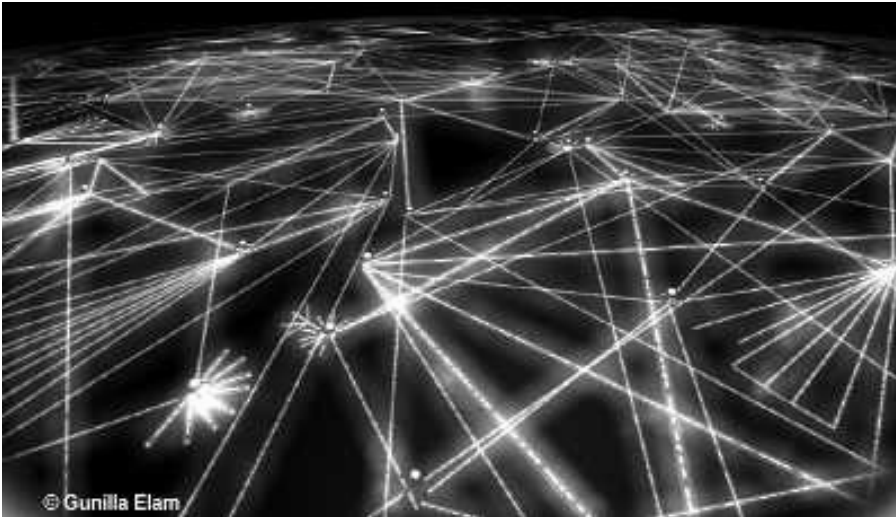
2

2004-05-27

InternetundWebseiten-Gestaltung - T.Mättig

4

Problem: Das Internet hat keinen „Anfang“ ;-)



Lesezeichen/Favoriten

- Lesezeichen konsequent verwenden!
- 1. Schritt, nachdem Sie eine interessante Seite gefunden haben: zu den Lesezeichen hinzufügen!
- Ab und zu thematisch in Ordner einsortieren und ausmisten
 - Besser, als etwas nicht wieder zu finden, von dem Sie genau wissen, dass es da ist, irgendwo zwischen 8 Milliarden anderen Webseiten

2004-05-27

InternetundWebseiten-Gestaltung - T.Mättig

7

Aufbau des Internets

- „Chaotischer“ Zusammenschluss verschiedenster Teil-Netze, verbunden durch gemeinsame Kommunikations-Protokolle
 - TCP/IP: Transport-Protokoll, sorgt dafür, dass abgesendete Daten-Pakete den richtigen Empfänger finden
 - Wird von Ihrem Modem/Netzwerkkarte verstanden
 - HTTP: Spezielles Transport-Protokoll für Web-Inhalte
 - Wird von Ihrem Browser verstanden

Suchen im Internet

- Suchstrategien:
 1. „Domain-Raten“
 2. Stichwortsuche mit Suchmaschinen
 3. Suche in Themenkatalogen

2004-05-27

InternetundWebseiten-Gestaltung - T.Mättig

6

2004-05-27

InternetundWebseiten-Gestaltung - T.Mättig

8

1. „Domain-Raten“

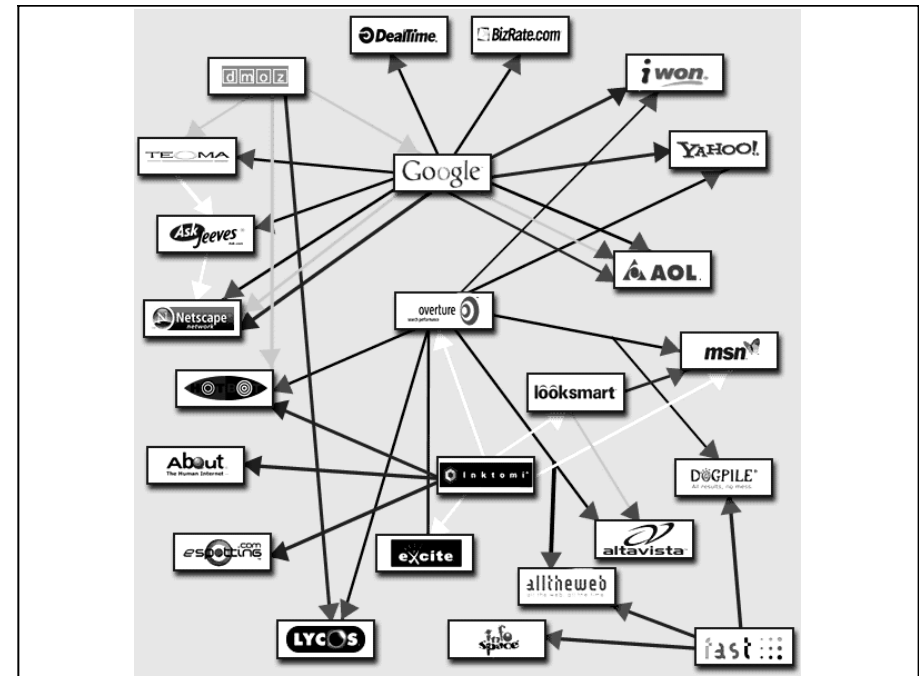
- Finden durch Ausprobieren, z.B.:
 - www.bertelsmann.de
 - www.mercedes.de
 - www.bundesregierung.de
 - Quasi-Konventionen:
 - www.tu-dresden.de
 - www.uni-leipzig.de
 - www.fh-aachen.de
 - www.hs-bremen.de

Alles Google oder was?

- Es gibt nicht nur „das eine Telefonbuch“, sondern:
 - Yahoo (www.yahoo.de) und MSN (www.msn.de)
 - Fireball
 - AltaVista, Lycos, AlltheWeb
 - Oft im Browser voreingestellt:
 - MSN
 - AOL (nutzt Google)
 - T-Online(nutzt Google)
 - Suchmaschinen-Suchmaschine:
 - www.klug-suchen.de

Länder-Domains

- www.bigband.de für „Deutschland“
 - www.bigband.it usw.
- www.bigband.net eigentlich „Netzwerk“
- www.bigband.org eigentlich „Organisation“
- www.bigband.com für „Company“
 - Aber: www.bigband.co.uk
 - Praktisch alles völlig frei verwendet
 - maettig.de war schon belegt ;-)



„Syntax“

- Die Art, wie Suchbegriffe eingegeben und verknüpft werden, nennt man „Syntax (der Anfrage-Sprache)“.
- Als Quasi-Standard hat sich die Syntax von Google durchgesetzt
 - Es gibt jedoch Sonderfälle → Hilfe lesen

Verknüpfung

- Und-Verknüpfung (mit +) ist nicht nötig
 - ‚brot butter‘ findet nur Seiten mit beiden Wörtern
- Wörter ausschließen mit –
 - ‚brot –butter‘ findet nur Seiten ohne ‚butter‘
- Gewichtung durch die Reihenfolge

Anfrage-Sprache

- Groß/Kleinschreibung wird ignoriert
 - kleinschreibung spart tipparbeit ;-)
- Umlaute können umschrieben werden
 - ‚baeren‘ findet auch ‚Bären‘
 - Nebeneffekt: ‚neü fotos‘ findet ‚neue Fotos‘

Phrasensuche

- "genau diese zeichenkette" findet genau das
- Ohne Anführungszeichen wird trotzdem fast genau das gefunden
 - ‚diese‘ wird jedoch ignoriert
 - Gefunden: ‚genau die Zeichenkette‘ etc.
- Wort ignorieren kann durch Stern (Platzhalter) erzwungen werden
 - "brot * butter"
 - Gefunden: ‚brot mit butter‘, ‚brot und butter‘ etc.

Sortierung der Fundstellen

- „Ranking“ (Positionierung)
- Früher: Je häufiger das Wort im Dokument vorkam
 - Leicht manipulierbar
 - Kein wirklich aussagekräftiges Kriterium
- Heute: Je mehr Links auf ein Dokument verweisen
 - Manipulationen kommen vor („Suchmaschinen-Spam“)
- Auch: Gekaufte Suchergebnisse

2004-05-27

InternetundWebseiten-Gestaltung - T.Mättig

18

Such-Taktiken

- Tipp: Zu einem sehr speziellen Wort ein sehr allgemeines Wort hinzunehmen, das die Kategorie beschreibt
 - Beispiel: ‚kohl‘ finde Kauderwelsch, aber ‚kohl kochen‘ findet genau das

2004-05-27

InternetundWebseiten-Gestaltung - T.Mättig

20

Hilfe, ich finde gar nichts

- Suche planen!
 - **Was** genau suche ich eigentlich?
 - Begriffe zurechtlegen, die eindeutig und präzise sind
 - Synonymezurechtlegen
 - **Wer** könnte darüber etwas im Web stehen haben?
 - **Wie** könnte das Gesuchte vorkommen?
 - **Wo** gibt es evtl. Hinweise, die mir weiterhelfen?
- Inhalte suchen, nicht Überschriften
 - Beispiel: nicht ‚Goethe Faust‘ sondern ‚Habe nun ach! Philosophie‘

2004-05-27

InternetundWebseiten-Gestaltung - T.Mättig

19

Hilfe, ich finde nichts

- Es soll wirklich noch Dinge geben, die nicht im Internet stehen
- Kann zu verzerrter Wahrnehmung der Realität führen

2004-05-27

InternetundWebseiten-Gestaltung - T.Mättig

21

Hilfe, ich finde zu viel

- Einschränkung des Suchraumes
 - Nur bestimmte Sprachen
 - Nur in bestimmten Regionen
 - Nur ‚Seiten aus Deutschland‘
 - Nur innerhalb bestimmter Domains
 - Nur im Titel, in Überschriften, in Links
 - Nur bestimmte Dokument-Typen (JPG, PPT, ...)

Thematische Suche

- Neben den allgemeinen Themenkatalogen/Verzeichnissen gibt es „Portale“ zu allen erdenklichen Themen
- Redaktionell gepflegt
- Stark eingeschränkte Thematik
- Dafür qualitativ hochwertige Fundstellen
- Beispiel:
 - www.goerlitz.de

3. Thematische Suche

- Allgemeine Themenkataloge
 - Systematisch geordnet
 - Redaktionell gepflegt
 - Kleine, aber brauchbare Auswahl
 - Nur eher allgemeine Themen
- Beispiele:
 - Yahoo
 - Der älteste Web-Katalog
 - Fireball, Web.de
 - allesklar.de
 - Google-Verzeichnis

Stichwortsuche mit Meta-Suchmaschinen

- Suchen in mehreren normalen Suchmaschinen gleichzeitig
- Es wird mehr gefunden, aber auch mehr Falsches
- Empfehlung: Wenn Sie in der normalen Suchmaschine einfach nichts finden
- Beispiele:
 - MetaGer
 - Metacrawler

Wie macht Google das eigentlich?

- „Such-Roboter“ („Crawler“, „Spider“) surft durchs Internet
 - Das dauert natürlich eine Weile
- Gefundene Seiten werden „indiziert“
 - In einer Wortliste (Index) wird notiert, auf welchen Seiten jedes Wort vorkommt
 - Die paar hunderttausend Wörter sind gut handhabbar ;-)
- Bei Suchanfragen wird nur die erstellte Wortliste verwendet

Ich will gefunden werden!

- Abwarten. Die meisten Suchmaschinen finden Sie von ganz allein.
 - Kann durchaus 2 Monate dauern
- Konzentrieren Sie sich nicht auf die Suchmaschine, sondern auf Ihr Angebot.
 - Wenn viele Leute Ihr Angebot mögen, werden viele Links zu Ihrem Angebot entstehen.
 - Google mag Sie, wenn andere Sie mögen.

Zwischen-Speicher (Cache)

- Suchmaschine speichert Wortlisten und evtl. Kopien der Seiten
- Ihr Browser speichert die Seiten, die Sie vor kurzem besucht haben
- „Proxy (Server)“ leitet Anfragen weiter
 - Wenn die selbe Anfrage gleich wieder kommt, kennt er die Antwort schon und gibt die Seite aus seinem Zwischenspeicher heraus
- → Viele Gründe, warum man manchmal alte Versionen neuer Seiten sieht

In Verzeichnissen gefunden werden

- Verzeichnisse werden von Menschen gepflegt → Sie müssen sich anmelden
- Rubrik auswählen
 - Aber nicht:
 - „Persönliche Homepages > M > Me > Meier“ (was hilft es, wenn Sie da drin stehen?)
 - Sondern:
 - „Freizeit & Lifestyle > Feste & Feiern > Festgestaltung“

Meta-Angaben

- Meta-Angaben im Kopf (Head) jedes einzelnen HTML-Dokumentes
 - Autor
 - Schlüsselwörter
 - Beschreibung
- Hat heute (fast) gar keine Bedeutung mehr
 - Nicht wirklich aussagekräftig
 - Zu leicht manipulierbar, zu oft falsch verwendet
- Aber: Titel, Autor und Beschreibung wird in Katalogen etc. angezeigt
- `<meta name="robots" content="noindex,nofollow">`

Keine Tricks

- Sie können versuchen, Google zu beschummeln („Google ist blind“)
- Aber: Was nützt es, wenn Ihre Seite gefunden wird, in Wirklichkeit aber gar nichts Nützliches da steht?

„Google ist blind“

- Suchmaschinen „sehen“ Ihre Seiten nicht
 - Farben, Bilder, layouttechnische Anordnung sind irrelevant
 - Nur blanker Text wird „gelesen“
 - Auch deswegen: Alternativtexte für Bilder
 - Überschriften werden höher eingestuft
 - Der erste Satz/Absatz wird höher eingestuft

„Suchmaschinen-Optimierung“

- SEOs (Search Engine-Optimierer) bieten ihre Dienste für Geld an
- Oft wird die Web-Präsenz zum Schaden der menschlichen Besucher kaputt-optimiert
- Ihr Seite wird danach gefunden, aber unter Suchbegriffen, die wohl niemand eintippen wird
- Seriöse Optimierung optimiert die Inhalte Ihrer Seiten
 - Was Ihren menschlichen Besuchern nützt, gefällt früher oder später auch den Suchmaschinen