

15

64

MÄRZ 2005

# Suchmaschinen: Das Tor zum Netz

Hintergrund und Information

## Impressum

Herausgeberin	Bündnis 90/Die Grünen Bundestagsfraktion Platz der Republik 1 11011 Berlin <a href="http://www.gruene-fraktion.de">www.gruene-fraktion.de</a>
Verantwortlich	Grietje Bettin, MdB bildungs- und medienpolitische Sprecherin Silke Stokar, MdB innenpolitische Sprecherin BÜNDNIS 90/DIE GRÜNEN Bundestagsfraktion Platz der Republik 1 11011 Berlin E-Mail: <a href="mailto:grietje.bettin@bundestag.de">grietje.bettin@bundestag.de</a>
Autor	Dr. Harald Neymanns
Redaktion	Büro MdB Bettin: Julia v. Dobeneck
Bezug	Bündnis 90/Die Grünen Bundestagsfraktion Info-Dienst Platz der Republik 1 11011 Berlin Fax: 030 / 227 56566 E-Mail: <a href="mailto:public@gruene-fraktion.de">public@gruene-fraktion.de</a>
Schutzgebühr	€ 1,--
Redaktionsschluss	März 2005

# Inhalt

<b>Vorwort</b> .....	<b>3</b>
<b>1. Suchmaschinen und Suchverhalten</b> .....	<b>5</b>
<b>2. Die Dominanz von Google</b> .....	<b>6</b>
<b>3. Suchmaschinen – allwissend und neutral?</b> .....	<b>7</b>
Die Weiten des Raumes .....	7
Intransparente Algorithmen und der Schein der Neutralität .....	8
Redaktion oder Werbung? .....	9
<b>4. Datenschutz bei Suchmaschinen</b> .....	<b>10</b>
<b>5. Vielfalt fördern, Standards halten und Kompetenz stärken</b> .....	<b>12</b>
Vielfalt fördern und Alternativen schaffen .....	12
Transparenz fördern und Aufsicht stärken .....	14
Datenschutzstandards etablieren .....	14
Medienkompetenz stärken.....	15
<b>6. Kurz vorgestellt: Google-Tipps und Alternativen</b> .....	<b>17</b>
<b>7. Glossar</b> .....	<b>19</b>



## Vorwort

Recherche im Internet findet heutzutage zum Großteil über Suchmaschinen statt. Google ist für viele Internetnutzerinnen und -nutzer das Synonym für Suchmaschinen schlechthin: 75% suchen über Google, wenn sie eine Information benötigen. Das Verb „googeln“ wurde sogar in den Duden aufgenommen.

Suchmaschinen erleichtern den Zugriff auf die Vielfalt an Informationen im Internet. Allerdings tun sie nicht nur das: Suchmaschinen entscheiden auch darüber, welche Information in welcher Reihenfolge angezeigt wird - und welche nicht.

BÜNDNIS 90/DIE GRÜNEN haben sich immer für einen uneingeschränkten Zugang zu Informationen ausgesprochen. Bürgerinnen und Bürger müssen sich umfassend und aus verschiedenen Quellen informieren können. Dies ist für uns Grundvoraussetzung für die Teilhabe an unserer demokratischen Gesellschaftsform.

Gerade das Internet ist beispiellos in seiner Vielfalt an Informationen und Meinungen. Hier hat sich ein Medium etabliert, bei dem der direkte Zugang zu Informationen möglich ist. Kommunikatoren sind nicht mehr auf Gatekeeper angewiesen, die ihre Aussagen selektieren und prüfen, sondern treten selbst als Informationsanbieter auf.

Der ungefilterte Zugang zu diesem vielfältigen Informationsangebot muss auch vonseiten der Suchmaschinenbetreiber gewährleistet sein.

Suchmaschinenanbieter müssen sich ihrer großen Verantwortung bewusst sein. Sie entscheiden über die Hierarchisierung der Suchergebnisse und steuern die Aufmerksamkeitslenkung im Internet. Suchmaschinenanbieter müssen daher das Zustandekommen ihrer Ergebnisse so transparent wie möglich halten und Ergebnisse, die auf kommerziellen Vereinbarungen beruhen, kennzeichnen. Darüber hinaus müssen sie sparsam mit Nutzerdaten umgehen.

Um die Abhängigkeit von einer Meinung zu vermeiden, halten BÜNDNIS 90/DIE GRÜNEN es für wichtig, die Alternativen zu Suchmaschinen-Monopolisten zu stärken und zu unterstützen. Insbesondere Initiativen, die es sich zum Ziel gesetzt haben, bereits bestehende Datenbanken und Kataloge von Universitäten und Bibliotheken miteinander zu vernetzen und in ein gemeinsames Portal zu integrieren, halten wir für überaus förderungswert.

Das richtige Beherrschen von Suchtechniken und entsprechenden Tools gehören inzwischen zu den Grundlagen in Sachen Medienkompetenz. Daher ist es uns besonders wichtig, dass Nutzerinnen und Nutzer über die Funktionsweise der jeweiligen Suchmaschine und die verschiedenen Alternativen am Markt Bescheid wissen.

Die wenigsten Nutzerinnen und Nutzer wissen darum, wie ihre Such-Ergebnisse zustande kommen und wie viele sinnvolle Alternativen es zu den derzeitigen Marktführern bei den Suchmaschinen inzwischen gibt.

Die wenigsten wissen auch um die Problematik der Speicherung von Benutzereinstellungen bei Suchmaschinen. Die Datenschutzproblematik spielt im Zusammenhang mit Suchmaschinen eine gewichtige Rolle. Z. B. wird Google immer wieder vorgeworfen, durch so genannte "langlebige Cookies" umfangreiche Nutzerprofile zu erstellen. Außerdem werden mit jeder Suchanfrage Informationen wie Tageszeit, Browser-Typ, Browser-Sprache und IP-Adresse mitgeliefert.

Im Zuge des Börsengangs von Google im August 2004 ist die Debatte über die Marktdominanz von Suchmaschinen, undurchsichtige Hierarchisierung von Suchergebnissen und Nutzerdatenverwertung erneut entflammt.

Es haben sich verschiedene Initiativen gegründet, die sich mit der Funktionsweise vorherrschender Suchmaschinenanbieter beschäftigen und sich etwa um dezentrale Erfassung bereits vorhandener Netzinhalte bemühen. Ebenso werden Suchmaschinenbetreiber zu Selbstverpflichtungserklärungen angehalten, um Standards wie Transparenz und Datensparsamkeit zu gewährleisten. Alle diese Initiativen unterstützen wir.

Wir möchten mit dieser Info-Broschüre unsererseits Nutzerinnen und Nutzer über die Suchmaschinenproblematik informieren und Alternativen aufzeigen und bekannt machen.

Denn nur wer weiß, nach welchen Kriterien Suchergebnisse zustande kommen und wie das vorhandene Angebot am sinnvollsten genutzt wird, kann von der Vielfalt des Informationsangebots im Internet frei profitieren.

Grietje Bettin

Silke Stokar

# 1. Suchmaschinen und Suchverhalten

Mehr als 8 Milliarden Webseiten, über 280 Millionen weltweit registrierte Internet-Adressen (Domains)<sup>1</sup> und über 8 Millionen deutsche Adressen, die auf .de enden, müssen sortiert werden. Suchmaschinen sind für die Auffindbarkeit von Information im Internet notwendig, erst sie machen das World Wide Web (WWW)<sup>2</sup> zu der Informationsquelle, wie sie heute bekannt ist. Über die zurzeit bekannteste Suchmaschine, Google, kommen noch über 880 Millionen Bilder und mehr als 845 Millionen Nachrichten aus Diskussionsforen dazu.<sup>3</sup> Gerade wegen dieser kaum fassbaren und niemals zu bewältigenden Menge an Daten und Informationen ist es wichtig, plausible Ordnungssysteme zu entwickeln.

Hauptsächlich private Unternehmen versuchen, Struktur in das dezentrale Internet, vor allem das WWW, zu bringen. In den letzten Jahren haben viele Unternehmen zumindest für eine Zeit den Markt der Ordner im Netz bestimmt. Yahoo, Excite, Hotbot, AltaVista, Infoseek, Inktomi sind nur einige der bekannten Namen. Heute ist Google mit einem Marktanteil von über 75 Prozent in Deutschland die dominierende Suchmaschine. Anders ausgedrückt: 75 Prozent der deutschen Nutzerinnen und Nutzer lassen sich das Web von Google strukturieren, Google ist ihr Tor zum Netz.

Ein kurzes Beispiel zum Nachahmen soll verdeutlichen, dass es um mehr geht als um die bloße Auffindbarkeit von Information. Die Suche nach „George W. Bush“ bei den drei Suchmaschinen Excite.com, Yahoo.com und Google.com bringt sehr unterschiedliche Ergebnisse hervor: unter den ersten zehn Treffern gibt es zwischen Google und Yahoo nur zwei gleiche Treffer, zwischen Excite und Google sind es immerhin vier, wobei der erste Treffer von Google bei Yahoo auf Platz vier und bei Excite auf Platz sechs geführt wird. Acht der gefundenen Ergebnisse von Excite werden bei den anderen beiden nicht unter den Top 10 geführt.<sup>4</sup>

In den Augen vieler Nutzerinnen und Nutzer stellen Suchmaschinen neutrale Instanzen dar, die mathematisch verlässlich und inhaltsneutral die besten Ergebnisse liefern. Doch die oben genannten Suchbeispiele nach George W. Bush zeigen, dass manche Suchmaschinen manche Ergebnisse wichtiger bewerten als andere.

Auch das folgende Beispiel zeigt, dass es nicht nur um gute und inhaltlich neutrale Suchmethoden geht: vor einigen Monaten hat Google seine Algorithmen angepasst, um ein verbessertes Ranking zu realisieren und Manipulationsversuche zu erschweren. Dadurch wurden die Ergebnislisten verändert, frühere erste Plätze fanden sich auf den hinteren Rängen wieder. Nach Protest der „Herabgestuften“ wurden die Veränderungen weitgehend zurückgenommen.

Ein Großteil der Nutzerinnen und Nutzer nutzt nur eine Suchmaschine. Lediglich 40 Prozent von ihnen ziehen eine Zweite zu Rate. Außerdem werden im Durchschnitt nur die ersten zwanzig Ergebnisse beachtet. Durch Suchmaschinen wird also nicht nur eine Strukturierung, sondern auch eine Wertung von Information vorgenommen: wenn man das Nutzungsverhalten in Betracht zieht, entscheiden „die Suchmaschinen“, welche Information im Netz gefunden und genutzt wird und welche nicht. Ergebnisse, die es nicht unter die ersten zwanzig schaffen, existieren für den Großteil der Nutzerinnen und Nutzer nicht.

---

<sup>1</sup> <http://www.isc.org/index.pl?/ops/ds/>. Stand: 12.12.2004.

<sup>2</sup> Technische Fachbegriffe sind kursiv gedruckt und im Glossar ab Seite 19 erklärt.

<sup>3</sup> <http://www.google.com/corporate/facts.html>. Stand: 12.12.2004.

<sup>4</sup> Die Suche wurde durchgeführt am 10.12.2004.

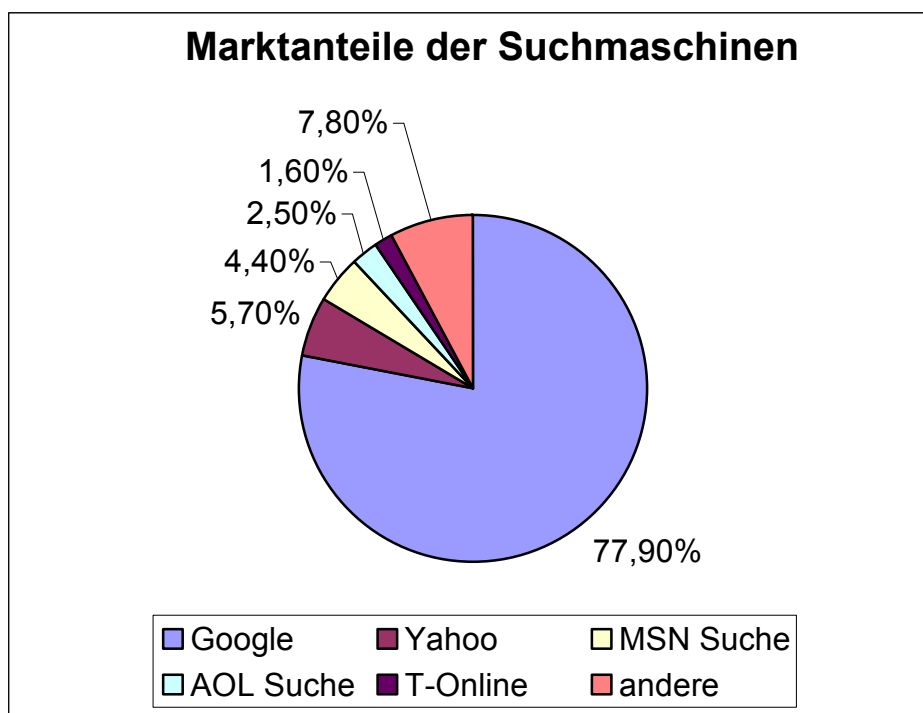
Man kann diesen Zusammenhang leicht mit dem Fernsehen vergleichen: die Programme auf den Tasten eins bis neun werden deutlich mehr genutzt als die auf zweistelligen Programmplätzen. Oder man stelle sich ein Regal im Supermarkt vor: die (teureren) Markenprodukte stehen in der Regel auf Augenhöhe, für „no name“ Produkte muss man sich bücken oder recken. Im Internet ist es durch die Dominanz von Google aber eher so, als wären die „no name“ Produkte in dem einzigen zugänglichen Supermarkt hinter einem Vorhang versteckt.

Das Netz gilt weiterhin als Alternative zu großen Medienkonzernen, in denen wenige Gatekeeper die Agenda kontrollieren. Zwar finden sich auch heute schon Abbildungen der herkömmlichen Medienstruktur im Internet, wie sich an der Popularität von Spiegel.de oder ähnlichen Online-Formaten ablesen lässt. Es besteht die Gefahr, dass sich diese Tendenz durch die Dominanz einer Suchmaschine fortsetzt. Die Gatekeeper sind dann allerdings nicht mehr Redakteure, sondern Programmierer und deren Algorithmen. Die „anderen“ Informationsquellen würden zwar weiter existieren, nur immer mehr in den Hintergrund rücken. Das Netz ist zwar ein dezentrales Medium, doch wenn es de facto nur einen Einstieg gibt, verliert es für den Großteil der Nutzerinnen und Nutzer seine Dezentralität.

## 2. Die Dominanz von Google

Trotz zahlreicher Alternativen beantwortet Google zurzeit  $\frac{3}{4}$  aller deutschen Suchanfragen. Das restliche  $\frac{1}{4}$  teilen sich die zahlreichen Konkurrenten, von denen Yahoo mit 5,7 Prozent abgeschlagen der stärkste Wettbewerber ist.

**Chart 1: Marktanteile Suchanfragen Deutschland**



Quelle: Webhits.de, Stand 3.12.2004,<sup>5</sup>

Dass diese Marktdominanz nicht unangefochten ist, zeigen die Ambitionen des Software-Giganten Microsoft, der im November 2004 einen Neustart seiner Suche auf search.msn.com

<sup>5</sup> <http://www.webhits.de/deutsch/index.shtml?/deutsch/webstats.html>. Stand: 20.12.2004.

ankündigte und im Februar 2005 startete. Das Ziel ist es, die Vorherrschaft von Google zu beenden. Auch Google ruht sich nicht auf dem Erreichten aus: in der zweiten Jahreshälfte 2004 startete Google zahlreiche neue Suchtools, etwa die Suche nach Produkten (Froogle), oder die Suche auf dem heimischen Rechner (Google Desktop Search). In den USA laufen bereits erste Versuche, Googles Suchtechnologien per SMS auch auf Mobiltelefonen anzubieten. Im Dezember 2004 kündigte Google unter großer Beachtung der Öffentlichkeit an, Vereinbarungen mit verschiedenen Universitätsbibliotheken getroffen zu haben, um große Teile der Buchbestände zu digitalisieren und per Suchfunktion zugänglich zu machen.

Auch wenn der Markt noch nicht gefestigt ist und sich die Machtverhältnisse schnell verändern können, wird deutlich, dass es nur theoretisch unendlich viele Einstiegsmöglichkeiten in den Wissenspool WWW gibt. Für die „Durchschnittsnutzer“ existieren lediglich ein oder zwei Tore zum Netz, während die anderen verschlossen bleiben. Ähnlich wie bei Betriebssystemen oder Webbrowsern sollte auch hier gelten, dass Vielfalt und Verschiedenheit von Vorteil ist.

### **3. Suchmaschinen – allwissend und neutral?**

Wenngleich Suchmaschinen in den letzten Jahren permanent weiter entwickelt wurden, haben sie inhärente Schwächen. Diese führen dazu, dass die Nutzerinnen und Nutzer je nach Fehleranfälligkeit beziehungsweise Reichweite der Suchmaschinen verzerrte Ergebnisse bekommen.

#### **Die Weiten des Raumes**

Suchmaschinen erfassen nur einen Teil der im WWW vorhandenen Informationen. Schätzungen gehen davon aus, dass Google nur etwa 60 Prozent der deutschsprachigen Webangebote katalogisiert hat. Dieser durch Suchmaschinen sichtbar gemachte Teil des WWW wird auch Visible Net genannt. Dabei handelt es sich vor allem um die statischen Inhalte im WWW, also normale Webseiten.

Zunehmend werden Seiten jedoch dynamisch erstellt, das heißt sie werden ständig aktualisiert und aus Datenbankbeständen erst bei einer konkreten Anfrage einer Nutzerin erzeugt. Diese Seiten werden von den meisten Suchmaschinen leider kaum oder nur unvollständig erfasst. Das gilt auch für bestimmte Dateiformate wie Flash oder Bildinhalte. Dieser für viele dadurch unsichtbare Bereich wird auch als Invisible Net oder Deep Web bezeichnet. Schätzungen aus dem Jahr 2000 gehen sogar davon aus, dass er 400 bis 550 Mal größer ist als der sichtbare Bereich und die stärksten Wachstumszahlen aufweist.<sup>6</sup>

Viele dieser über Suchmaschinen nicht unmittelbar zugänglichen Seiten enthalten professionell aufbereitete und qualitativ hochwertige Informationen. So werden beispielsweise die Webseiten und informationsreichen Archive von Tageszeitungen von den meisten gängigen Suchmaschinen nicht erfasst. Auch Berichte aus vielen Forschungsdatenbanken werden nicht aufgenommen.

Weitere Teile des Invisible Web werden von den Betreibern selbst von der Indexierung in Suchmaschinen ausgeschlossen. Diese Bereiche des WWW sind:

- Opaque Web („undurchsichtig“): Einschränkung der Aufnahme in die Suchmaschine durch den Seitenbetreiber
- Private Web: Passwortgeschützte private Information

---

<sup>6</sup> siehe Wikipedia, [http://de.wikipedia.org/wiki/Host\\_\(Datenbankanbieter\)#Deep\\_Web](http://de.wikipedia.org/wiki/Host_(Datenbankanbieter)#Deep_Web). Stand: 29.12.2004.

- Proprietary Web („urheberrechtlich geschützt“): Bewusste Einschränkung, beispielsweise wegen Registrierungspflicht oder kostenpflichtigen Inhalten. Insgesamt wird also nur ein sehr begrenzter Teil einer potenziell verfügbaren Vielfalt an Information erfasst. Allwissend sind die Suchmaschinen also keinesfalls.

### **Intransparente Algorithmen und der Schein der Neutralität**

Im Groben sind die Bewertungsmechanismen bei Google und den anderen Suchmaschinen bekannt. Details jedoch werden nicht veröffentlicht. Dafür gibt es zwei Gründe:

- Einerseits gelten die *Algorithmen* als Betriebsgeheimnis und könnten bei Offenlegen problemlos kopiert werden. Aus Angst vor der Konkurrenz hält man sie lieber geheim.
- Andererseits würden sich bei transparenten Bewertungsverfahren möglicherweise zahlreiche neue Manipulationsmöglichkeiten ergeben.

Die Suchmaschinenbetreiber stehen vor der schwierigen Aufgabe, täglich Millionen von Anfragen zu bearbeiten. Viele Anfragen sind mehrdeutig und müssen interpretiert werden. Ob die Suche nach dem Begriff „Bank“ etwa eine Sitzgelegenheit oder ein Geldinstitut meint, ist nicht klar. Ob die Suche nach bestimmten Produkten gleich eine Kaufintention suggeriert oder nur kritische Auseinandersetzungen mit den Funktionalitäten eines Geräts meint, lässt sich manchmal schwer abgrenzen. Je spezifischer die Anfrage, desto besser die Ergebnisse. Für die Nutzerinnen und Nutzer bleibt die Auswahl letztendlich leider intransparent, sie müssen das Vertrauen haben, dass die Suchmaschine „nach bestem Wissen und Gewissen“ sortiert.

Dieser Sachverhalt widerspricht der Metapher der Maschine, die suggeriert, dass die Maschine neutral ihre Arbeit verrichtet. Dass sie das keinesfalls tut, wurde bei der beispielhaften Suche nach gleichen Begriffen mit verschiedenen Suchmaschinen deutlich.

Trotz ausgefeilter Mechanismen ist es weiterhin möglich, die Trefferlisten zu beeinflussen und ein Angebot besser zu platzieren, als es die eigentlichen Inhalte und Vernetzung erlauben würden.<sup>7</sup> Die folgenden Manipulationsmöglichkeiten sind die bekanntesten:

- Cloaking: beim Cloaking wird dem *Spider*programm eine andere Webseite geliefert, als dem User nachher angezeigt wird. Der Vorteil: man kann eine für die Suchmaschine optimierte Seite erstellen, den Nutzerinnen und Nutzern aber grafisch anspruchsvolle Seiten liefern.
- Doorway-Pages: ähnlich wie beim Cloaking werden hier Seiten angelegt, die nur den Zweck haben, auf eine weitere Seite zu verweisen. Bei Doorway Pages werden tausende leicht unterschiedliche Seiten angelegt, die, wenn von den Nutzerinnen und Nutzern angeklickt, auf eine Seite verweisen und deren Positionierung verbessern.
- Linkfarmen oder Linkfabriken: das *Page-Ranking* Verfahren, das von Google als erstes genutzt wird, stellt die Bedeutung einer bestimmten Seite auch anhand der Anzahl darauf verweisender Links her. Link-Farmen haben den einzigen Zweck, auf eine Seite zu verweisen, und somit zu einem guten Ergebnis beim *Page-Ranking* Verfahren zu sorgen.

Viele Suchmaschinen bemühen sich, die Manipulationsmethoden zu begrenzen: Wird ein Manipulationsversuch erkannt, führt das in der Regel zu einem schlechten *Ranking* und möglicherweise zur Verbannung aus den Trefferlisten. Doch nicht alle Manipulationsversuche

---

<sup>7</sup> Ausführlichere Informationen zu den Manipulationsmöglichkeiten finden sich unter: <http://de.wikipedia.org/wiki/Suchmaschinen-Spamming>. Stand: 9.1.2005.

können zuverlässig und umfassend entlarvt werden, da die Betreiber von Webseiten immer neue Manipulationsvarianten erfinden und oft auf technisch höchstem Niveau vorgehen.

### Redaktion oder Werbung?

Die meisten Suchmaschinen finanzieren ihr Angebot hauptsächlich über Werbung – ebenso wie sich auch Fernsehen, Zeitungen und Zeitschriften über den Verkauf von Anzeigen (co-)finanzieren. Bei Rundfunk und Fernsehen ist die Trennung von redaktionellen Inhalten und Werbung zur Sicherung der Unabhängigkeit der Programmgestaltung und der Einhaltung der Neutralität im Rundfunkstaatsvertrag und Pressekodex klar geregelt.<sup>8</sup> Auch für Suchmaschinen gibt es hier Regeln, die besagen, dass kommerzielle Kommunikation als solche erkennbar sein muss.<sup>9</sup>

Trotzdem kommen immer wieder Vorwürfe auf, dass sich eine gute Platzierung in den Trefferlisten erkaufen lässt. Das lässt sich allerdings selten beweisen. Grund für diese Vorwürfe mögen auch die verschiedenen, für die Nutzerinnen und Nutzer undurchsichtigen Werbemaßnahmen sein.<sup>10</sup> Diese sind:

- Paid placements sind vom redaktionellen, also den „regulären“ Suchergebnissen klar abgetrennte, in der Regel thematisch auf die Suchbegriffe abgestimmte Werbelinks. Für die Nutzerinnen und Nutzer ist der Unterschied erkennbar, die platzierte Werbung ist als solche gekennzeichnet.
- Paid inclusions und paid submissions: Hierunter versteht man die bezahlte Aufnahme in die Trefferlisten bzw. Webkataloge. „Neutrale“ Suchalgorithmen werden also die zumindest modifiziert, indem bezahlte Angebote aufgenommen werden. Nicht zwangsläufig haben die Webseitenbetreiber Manipulation im Sinn: durch paid inclusions können sie beispielsweise rascher als im 2-4 Wochenzyklus der *Spider*programme aufgenommen werden, auch kann die Bezahlung sicherstellen, dass alle Unterseiten indiziert werden. Zwar wird bei paid inclusions normalerweise keine Platzierungsgarantie gegeben, gleichzeitig kann aber davon ausgegangen werden, dass die Kunden auch eine vernünftige Platzierung sehen wollen, wenn sie schon dafür bezahlen. Leider ist dieses Verfahren für die Nutzerinnen und Nutzer höchst intransparent.

Die folgende Tabelle zeigt auf, welche Suchmaschinen bzw. *Kataloge* welche Form von Werbung erlauben.

Suchmaschine	Paid Placements	Paid Inclusions
AltaVista	„Sponsored Matches“	Kann bei den Webseitentreffern vorkommen

<sup>8</sup> § 7 Absatz 2 Rundfunkstaatsvertrag: „Werbung oder Werbeinhalte dürfen das übrige Programm inhaltlich und redaktionell nicht beeinflussen.“

Pressekodex, Ziffer 7: „Die Verantwortung der Presse gegenüber der Öffentlichkeit gebietet, dass redaktionelle Veröffentlichungen nicht durch private oder geschäftliche Interessen Dritter oder durch persönliche wirtschaftliche Interessen der Journalistinnen und Journalisten beeinflusst werden. Verleger und Redakteure wehren derartige Versuche ab und achten auf eine klare Trennung zwischen redaktionellem Text und Veröffentlichungen zu werblichen Zwecken.“

<sup>9</sup> § 7 Teledienste-Gesetz: „[Besondere Informationspflichten bei kommerziellen Kommunikationen]

Diensteanbieter haben bei kommerziellen Kommunikationen, die Bestandteil eines Teledienstes sind oder die einen solchen Dienst darstellen, mindestens die nachfolgenden Voraussetzungen zu beachten.

1. Kommerzielle Kommunikationen müssen klar als solche zu erkennen sein.“

§ 4 Gesetz gegen den unlauteren Wettbewerb: „Beispiele unlauteren Wettbewerbs

Unlauter im Sinne von § 3 handelt insbesondere, wer [...]

3. den Werbecharakter von Wettbewerbshandlungen verschleiert;“

<sup>10</sup> Vgl. Danny Sullivan: *Buying Your Way In: Search Engine Advertising Chart*, November 22, 2004.

<http://searchenginewatch.com/webmasters/article.php/2167941>. Stand: 9.1.2005.

	zeigt bezahlte Einträge	
AOL Search	„Sponsored Links“ zeigt bezahlte Einträge	nein
AskJeeves	„Featured Sponsor & Sponsored Web Results“ Zeigen bezahlte Einträge	Kann bei den Webseitentreffern vorkommen
Google	„Sponsored Links“ Neben den „Anzeigen“	nein
Infospace	„Sponsored Links“ Einzelnen gekennzeichnet	Möglich (konnte nicht zweifelsfrei überprüft werden)
MSN Search	„Gesponsorte Webseiten“ zeigt bezahlte Einträge	„Gesponsorte Webseiten“ Kann bei den Webseitentreffern vorkommen
Teoma	„Sponsored Links“ zeigt bezahlte Einträge	Möglich (konnte nicht zweifelsfrei überprüft werden)
Yahoo	„Sponsored Results“ zeigt bezahlte Einträge oder in abgetrennten Boxen als Werbung gekennzeichnet	Über „Infos zu dieser Seite“ finden sich Hinweise auf Paid Inclusions

Quelle: SearchengineWatch<sup>11</sup>

Für die Nutzerinnen und Nutzer ist das Problem offensichtlich: Redaktionelle Inhalte und Werbung sind nicht immer voneinander zu unterscheiden, die Grenzen verschwimmen häufig. Bei den traditionellen Medien Zeitung und Rundfunk wäre eine solche Situation nicht hinnehmbar. Hier hängt das Netz noch deutlich hinterher.

#### 4. Datenschutz bei Suchmaschinen

In Bezug auf Datenschutz lassen sich erhebliche Bedenken gegen den gängigen Umgang der Suchmaschinenbetreiber mit Nutzer- und Nutzungsdaten formulieren. Google beispielsweise hat es wegen der Speicherung von Nutzerdaten in den USA zu einer Nominierung zum Big Brother Award gebracht – einem „Preis“ für Firmen und Individuen, die sich grober Datenschutzverletzungen verdächtig gemacht haben. Bemängelt wurde neben der Haltbarkeit des *Cookies* bis ins Jahr 2038 vor allem, dass die *IP-Adresse*, die zur Identifizierung des Rechners herangezogen werden kann, sowie das Datum, die Zeit und gar der Inhalt der Suchanfrage gespeichert werden. Auch Informationen über die *Browserkonfiguration* werden gesammelt. Nutzerinnen und Nutzer können sich dagegen kaum wehren.

Auch andere Suchmaschinen kommen schnell in den Verdacht, die Nutzungsdaten sehr großzügig auszuwerten. Eine neue Suchmaschine namens Blinkx, ein Startup aus Kalifornien,

<sup>11</sup> <http://searchenginewatch.com/webmasters/article.php/2167941>. Stand: 15.12.2004.

rühmt sich zu erkennen, was die Nutzerin oder der Nutzer am Rechner gerade treibt und kann dazu die entsprechenden Ergänzungen liefern. Das Programm analysiert die Inhalte der gerade geöffneten Dokumente oder die gerade per *Browser* angesteuerten Seiten. Dann liefert es automatisch Tipps und Hinweise zu den jeweiligen Themen. Dass mit dem Programm erhebliche datenschutzrechtliche Fragen aufgeworfen werden, ist klar: das Verhalten der Nutzerinnen und Nutzer wird intensiv analysiert, die Inhalte von Dokumenten werden automatisch ausgelesen. Zusätzlich dazu findet die Suche nach passenden Tipps noch auf den Rechnern von Blinks statt. Es ist also möglich, ausführliche Benutzerprofile zu erstellen. Auch die Ausweitung der angebotenen Dienstleistungen von Google und anderen Suchmaschinenbetreibern wirft datenschutzrechtliche Fragen auf. Drei Beispiele sollen hier aufgeführt werden: Die Suche in den *Newsgroups*, die Erweiterung der Suchfunktionalitäten auf den heimischen *Desktop* und der geplante Aufbau von Googles E-Mail Service GMail. Anfang 2001 kaufte Google die Suchmaschine für *Usenet Newsgroups*, DejaNews, und integrierte deren Funktionen ins eigene Angebot. Über den Link „Groups“ über dem Eingabefenster lassen sich die *Usenet* Diskussionsforen durchsuchen. Die Diskussionsforen im *Usenet*, das schon 1979 entstand und somit deutlich älter ist als das *WWW*, sind E-Mail basiert und sind öffentlich zugänglich. Ursprünglich benötigte man spezielle Programme, sog. *Newsreader*, um Zugriff auf die Gruppen zu haben. Heute lassen sich viele Foren per *Browser* nutzen. Wie lange die Beiträge in den jeweiligen Foren aufbewahrt wurden lag in der Entscheidung der jeweiligen Serverbetreiber. Über DejaNews und Google, zusammen unter dem Label Google Groups vereint, wurden die Archive wieder zugänglich. Mit Google Groups lassen sich Diskussionen durchsuchen, die vor Jahren geführt wurden – das konnte den Nutzerinnen und Nutzern damals nicht klar sein. Die Kommunikation im *Usenet* lebt von einer gewissen Geschwindigkeit und ist eher ein Gespräch als eine Presseerklärung. Im Unterschied zum *WWW*, wo man eigene Information bereitstellen und diese dann auch wieder löschen kann, ist das bei Diskussionsforen nur sehr schwer möglich. Informationen über Personen, die unter einer bekannten E-Mail Adresse oder dem Realnamen geschrieben haben, sind „Wort für Wort“ auffindbar.

Im Herbst 2004 erweiterte Google sein Angebot: ein neues Programm bringt die Suchtools von Google auf den PC. Die Nutzerinnen und Nutzer können sich den Google *Desktop* auf den Rechner laden und damit die integrierte Windows-Suchfunktion verbessern. Das wird vom Virtuellen Datenschutzbüro, einer zentralen Informationsstelle, die von zahlreichen offiziellen Datenschutzinstitutionen getragen wird, folgendermaßen kommentiert:

„Auch wenn Google derzeit nicht auf die vom Programm erstellte Index- und Archivierungsdatei zugreift, ist es durch die standardmäßige Verknüpfung von Web- und *Desktop*suche in Verbindung mit der einzigartigen, jedem Google-Suchprogramm eigene Identifikationsnummer möglich, Rückschlüsse auf Computerinhalte der Nutzerin oder des Nutzers zu ziehen.“<sup>12</sup>

Auch der kostenlose Mail-Service von Google, Gmail, hat datenschutzrechtlich zweifelhafte Funktionen, denn die E-Mails werden automatisch ausgelesen und auf werberelevante Begriffe durchsucht. Mit diesen Informationen lassen sich umfassende Profile der Nutzerinnen und Nutzer erstellen und personalisieren, da die E-Mail Adresse ja bekannt ist. Von US-amerikanischen Datenschützerinnen wurde kritisiert, dass Google keine genauen Angaben dazu veröffentlicht hat, in welcher Form die Daten an welche Unternehmen weitergeleitet werden. Gmail verstößt gegen das deutsche Fernmeldegeheimnis, da es nach dem Gesetz keinen Unterschied macht, ob eine Maschine oder ein Mensch die Mails mitliest.

---

<sup>12</sup> Virtuelles Datenschutzbüro unter: <http://www.datenschutz.de/news/alle/detail/?nid=1377>. Stand: 15.1.2005

## 5. Vielfalt fördern, Standards halten und Kompetenz stärken

Die hier angesprochenen Probleme sind wichtige Herausforderungen, denen sich Politik und Wirtschaft ebenso wie Nutzerinnen und Nutzer stellen müssen. Durch die Ausweitung der angebotenen Dienste auf Produktsuche, die Suche auf dem heimischen Rechner oder die Suche mit Mobiltelefonen, steigt die Bedeutung der Suchmaschinen. Als ein schwieriges Problem stellt sich dar, dass sich allgemeinverbindliche Regulierungen im Internet schwerlich nationalstaatlich vornehmen lassen. Dies liegt an der internationalen Reichweite des Internet und damit auch von Suchmaschinen. Auch die Dezentralität, die den besonderen Charakter des Internet ausmacht und die BÜNDNIS 90/DIE GRÜNEN für sehr entscheidend und wichtig halten, stellt ein Problem dar, da bei Weitem nicht alle Inhalte von einer Suchmaschine erfasst werden. Dadurch geht ein großer Teil hervorragender Informationen verloren.

Die Abhängigkeit von einem den Markt dominierenden Akteur und die daraus resultierenden Einschränkungen müssen, so gut es geht, abgeschwächt werden. Vielfalt und Alternativen im Bereich der Suchmaschinen müssen gefördert werden, indem bereits bestehende Alternativen bekannter gemacht und neue Suchtechnologien unterstützt werden. Darüber hinaus müssen Datenschutzstandards sowie die „redaktionelle Unabhängigkeit“ auch bei Suchmaschinen eingehalten werden. Am wichtigsten jedoch ist die Förderung von Medienkompetenz: Nutzerinnen und Nutzer müssen sich über die Probleme bei den beliebtesten Suchmöglichkeiten bewusst sein, sie sollten die bestehenden Möglichkeiten, die Suche spezifisch einzugrenzen, stärker nutzen und nicht nur auf Google setzen.

### Vielfalt fördern und Alternativen schaffen

Für das Fernsehen kontrolliert die Kommission zur Ermittlung der Konzentration im Medienbereich (KEK) die „Einhaltung der Bestimmungen zur Sicherung der Meinungsvielfalt im Fernsehen“.<sup>13</sup> Der Grenzwert, bei dem eine Konzentration als Gefahr für die Meinungsvielfalt angesehen wird, liegt bei 30 Prozent Zuschaueranteil. Im Pressebereich greift das Gesetz gegen Wettbewerbsbeschränkungen (GWB). Nach §19 Satz 3 GWB wird eine marktbeherrschende Stellung angenommen, wenn das Unternehmen einen Marktanteil von 33 Prozent hat.<sup>14</sup> Vor diesen Vergleichszahlen stellt sich der Anteil von Google mit 75 Prozent auf dem deutschen Markt noch drastischer dar. Hier muss darüber nachgedacht werden, welchen Regulierungsmechanismen Suchmaschinen unterworfen werden können.

Um die Abhängigkeit von einer Meinung zu vermeiden, ist es wichtig, die Alternativen zum momentanen Marktführer Google zu stärken:

1. In einem Test der Computerzeitschrift c't vom Dezember 2004 wurden drei große und 15 kleinere (Meta-) Suchmaschinen untersucht und verglichen. Im Ergebnis lieferten alle getesteten Suchmaschinen „passable bis gute Ergebnisse“.<sup>15</sup> Einige Aufgaben wurden von manchen Suchmaschinen sogar besser gelöst als von Google. Dieser Test zeigt: Es gibt genügend Auswahlmöglichkeiten. Daher werden einige Alternativen, die gute Ergebnisse liefern und ebenso nutzbar sind wie Google, im Anhang aufgeführt.
2. Diese Alternativen sind zu wenig bekannt. Der Bekanntheitsgrad der Alternativen zu den gängigen Suchmaschinen sollte gesteigert werden, indem die ver-

---

<sup>13</sup> <http://www.kek-online.de/cqi-bin/resi/k-auf/index.html>, Stand: 25.1.2005.

<sup>14</sup> § 19 Gesetz gegen Wettbewerbsbeschränkungen: „Missbrauch einer marktbeherrschenden Stellung [...] (3) Es wird vermutet, dass ein Unternehmen marktbeherrschend ist, wenn es einen Marktanteil von mindestens einem Drittel hat.“ Diese 33 Prozent werden bei einer geplanten Fusion relevant. Dann hat das Kartellamt zu entscheiden.

<sup>15</sup> c't, 26, 2004, S. 156f.

schiedenen Suchmaschinen in die *Browser* integriert werden. Beim neuen *Browser* Firefox, der von einer weltweiten Entwicklergemeinde entworfen wurde, sind in einem Suchfenster am oberen Rand einige Suchmaschinen voreingestellt. Je mehr Suchmaschinen dort eingestellt und je leichter neue hinzuzufügen sind, desto mehr Nutzerinnen und Nutzer werden auf diese zugreifen.

3. Suchangebote müssen nach Meinung von BÜNDNIS 90/DIE GRÜNEN nicht so zentralisiert sein wie bei Google. Vielmehr bietet die dezentrale Struktur des Internet einen großen Vorteil, der sich auch in den Suchmaschinen widerspiegeln sollte. Es existieren bereits *Kataloge* und Suchmaschinen, die dezentral aufgebaut sind und auf bereits vorhandene Informationen zurückgreifen. Sie bauen stärker auf Kooperation als auf Hierarchie und entsprechen damit der dezentralen und vielfach kooperativen Struktur des Internet. Im Juli 2004 startete beispielsweise eine akademische Suchmaschine mit dem Namen „Bielefeld Academic Search Engine“ (BASE). Im Forschungsprojekt wurde die Eignung von Suchmaschinentechnologien für die akademische Suche in digitalen Bibliotheken und Forschungsportalen untersucht. Ein erster Prototyp der Suchmaschine ist mittlerweile online. Mit BASE sollen auch Teile des *Invisible Web* erschlossen werden.<sup>16</sup> Von großer Bedeutung ist auch die Idee, dass eine solche, dezentrale Suchmaschine nicht monopolisierbar ist und mit *Open Source* Produkten betrieben werden soll. Darüber hinaus sollen im Projekt BASE neue Verfahren zur Bewertung wissenschaftlicher Arbeiten entwickelt werden, die denen herkömmlicher Suchmaschinen deutlich überlegen sind. Diese vorhandenen Strukturen und Projekte wie BASE müssen stärker ausgebaut und miteinander verknüpft werden.
4. Zunehmend werden spezialisierte Suchmaschinen wichtiger, die bestimmte Anfragen besser lösen können als die Generalisten. Diese Suchmaschinen dringen in Bereiche vor, die bei den gängigen Suchmaschinen bislang oft vernachlässigt werden. Außerdem greifen sie auf bestehende, qualitativ hochwertige Informationen zu. Universitäten und Bibliotheken bieten schon heute hervorragende und spezialisierte Suchfunktionen. Bislang sind sie in der Regel nur lokal dem akademischen Publikum zugänglich. Dieses Potenzial könnte durch Vernetzung deutlich besser ausgeschöpft werden. Die Deutsche Internetbibliothek<sup>17</sup> ist ein bereits funktionierender, dezentral aufgebauter und modular erweiterbarer Webkatalog. Über diesen Aufbau wird Vielfalt garantiert. Momentan betreuen 70 öffentliche Bibliotheken den Katalog und die Hilfsfunktionen.
5. Glücklicherweise hat auch der *Open Source* Gedanke bei Suchmaschinen Einzug gehalten. Bei *Open Source* geht es darum, dass der Programmcode offen gelegt wird und damit überprüft werden kann. Damit ist auch die Möglichkeit gegeben, den Programmcode weiterzuentwickeln und kooperativ zu verbessern. Bei Suchmaschinen hat das den Vorteil, dass Verzerrungen im *Algorithmus*, mit dem das *Ranking* verändert würde, erkennbar sind. Vorwürfe beispielsweise, dass die Platzierungen erkaufte werden können, sind sofort nachprüfbar. Das Open Directory Project<sup>18</sup>, ein *Webkatalog*, baut auf den *Open Source* Gedanken auf. Im Projekt Nutch<sup>19</sup> wird zum Beispiel versucht, eine

---

<sup>16</sup> <http://base.ub.uni-bielefeld.de/index.html>.

<sup>17</sup> <http://www.internetbibliothek.de>. Die Deutsche Internetbibliothek entstand in Zusammenarbeit zwischen dem Deutschen Bibliotheksverband und der Bertelsmann Stiftung.

<sup>18</sup> <http://dmoz.org>.

<sup>19</sup> <http://www.nutch.org>.

Suchmaschine mit *Open Source* auf non-profit Basis aufzubauen. BÜNDNIS 90/DIE GRÜNEN halten *Open Source* Suchmaschinen für unterstützenswert und notwendig, um Transparenz zu schaffen und Monopolbildungen zu vermeiden.

6. Google und die anderen Suchmaschinen lassen viele und vor allem qualitativ hochwertige Bereiche unentdeckt. Diese Inseln im Netz bleiben für den Großteil der Nutzerinnen und Nutzer unerreichbar. Eine interessante Möglichkeit, diesem Problem zu begegnen, bietet folgende Idee: alle deutschen *Domains*, das heißt alle Internetadressen, die auf .de enden, sind bei Denic<sup>20</sup> registriert. Würde man diese registrierten *Domains* als Ausgangspunkt für die Indizierung der Suchmaschinen nehmen, könnte man einen deutlich größeren Teil des deutschsprachigen Internets erschließen, als es Google und allen anderen Suchmaschinen bislang möglich ist.

### **Transparenz fördern und Aufsicht stärken**

Generell gilt, dass die Suchmaschinenbetreiber so transparent wie möglich agieren müssen und sich ihrer gesellschaftlichen Verantwortung stellen. In Bezug auf die Vermischung von redaktionellen Inhalten und Werbung ist die Verantwortung sehr deutlich geworden. Hier besteht immenser Nachholbedarf. Suchmaschinenbetreiber müssen die Nutzerinnen und Nutzer über die Funktionsweise der Suchmaschine aufklären und insbesondere die Kriterien der *Rankings* offen legen. Sie müssen ihre Ergebnisseiten möglichst transparent gestalten. Inhalte, deren gute Positionierung in der Ergebnisliste erkaufte wurde, müssen als Werbung erkennbar sein.

Es kann nicht sein, dass in Rundfunk und Presse die Kennzeichnung von Werbung gängige Praxis ist, die Grenzen zwischen Werbung und Inhalten im Bereich der Suchmaschinen jedoch verschwimmen.

Der Verhaltenskodex in Sachen Transparenz, den die Freiwillige Selbstkontrolle Multimedia<sup>21</sup> erarbeitet hat und der auch bereits von den wesentlichen Anbietern unterzeichnet wurde, ist ein richtiger und wichtiger Schritt, um diese Standards zu gewährleisten. Nun kommt es darauf an, dass die Unterzeichnenden ihre Selbstverpflichtung auch umsetzen.

### **Datenschutzstandards etablieren**

Bei den Suchmaschinen sammeln sich riesige Mengen an Nutzerdaten. Gerade wenn Suchmaschinenbetreiber ihre Produkte auf andere Anwendungen und Plattformen ausdehnen wird es wichtiger, einen angemessenen Standard beim Datenschutz zu erreichen. Hier müssen klare Regeln ausgearbeitet und konsequent vertreten werden. Neue Suchtechnologien, wie die oben erwähnte neue Suchmaschine Blinkx, die die Aktivitäten der Nutzerinnen und Nutzer deutet und daraufhin Suchergebnisse liefert, verheißen wenig Gutes, wenn die Nutzung der Daten nicht beschränkt wird.

Schon heute lassen sich von Internetnutzern in wenigen Minuten umfassende Profile aus den frei im Web verfügbaren Informationen erstellen. Zwei Prinzipien dürfen dabei auf keinen Fall vergessen werden:

- „auch personenbezogene Daten, die der Nutzer freiwillig veröffentlicht hat, unterliegen den für sie geltenden Schutzbestimmungen; [...]
- der Einzelne sollte in jedem Fall und zu jedem Zeitpunkt das Recht haben, der Veröffentlichung seiner personenbezogenen Daten in einem Internet-Angebot

---

<sup>20</sup> <http://www.denic.de>.

<sup>21</sup> <http://www.fsm.de>.

zu widersprechen. Er oder sie sollte das Recht haben zu verlangen, dass der Zweck respektiert wird, für den die Daten veröffentlicht worden sind.“<sup>22</sup>

Wir wollen, dass die Gesetze zum Datenschutz konsequent eingehalten werden. Dazu sollten die Betreiber von Suchmaschinen durch klare Offenlegung ihrer Datennutzung verpflichtet werden. Auch hier sollte eine Selbstverpflichtungserklärung für Offenheit sorgen, in der der Grundsatz der Datensparsamkeit für Suchmaschinenbetreiber festgelegt wird und die Verwendung der Daten für weitere eigene Produkte, die über die Nutzung einer Suchmaschine hinausgehen, nur mit ausdrücklicher Zustimmung der Nutzerinnen und Nutzer erlaubt sind. Es kann auch darüber nachgedacht werden, in Anlehnung an §6 Teledienste-Gesetz<sup>23</sup> klare Regulierungen zu Pflichtangaben bei Suchmaschinen zu entwickeln. Eine besondere Behandlung von Suchmaschinen ist für uns aufgrund ihrer besonderen Stellung gerechtfertigt. Sinnvoll ist es auch, eine Label „datenschutzgeprüft“ einzuführen, mit der sich mit deutschen Standards konforme Seiten schmücken dürfen. Ein solches Label bietet eine Orientierung für die Nutzerinnen und Nutzer und stärkt den Wettbewerb der Betreiber untereinander. Das neu geregelte Bundesdatenschutzgesetz, das noch in dieser Legislaturperiode verabschiedet werden soll, sieht in §9a die Umsetzung von entsprechenden Auditkriterien vor, ähnlich denen aus dem Landesdatenschutzgesetz von Schleswig-Holstein, die die Datenschutzkonzepte von Webseiten überprüfen und bewerten sollen.<sup>24</sup> Es wäre sinnvoll, dieses Verfahren für das Label „datenschutzgeprüft“ zu nutzen.

### **Medienkompetenz stärken**

Am Wichtigsten ist uns jedoch, die Nutzerinnen und Nutzer mit den Gefahren einer Monokultur bei den Suchmaschinen vertraut zu machen und alle aufzufordern, Alternativen zu nutzen sowie sparsam mit der Datenvergabe zu sein. Noch sind die Möglichkeiten der Nutzung alternativer Suchmöglichkeiten über Webkataloge und Metasuchmaschinen oder die Eingrenzung des Suchraums oder der –begriffe leider nicht hinreichend bekannt. Abgesehen von Google gibt es unzählige weitere Suchmaschinen, die die Navigation im Internet ermöglichen. Dabei gibt es Generalisten, wie beispielsweise Google, Yahoo, Web.de oder MSN. Über diese bekannten Suchmaschinen hinaus erledigen allein in Deutschland geschätzte 2000 bis 3000 Suchmaschinen spezialisierte Aufgaben, in dem sie sich etwa auf ein Themengebiet oder ein Branche konzentrieren. In Ergänzung dazu gibt es Suchmaschinen, die dezentral operieren oder den Gedanken der *Open Source* Bewegung aufgreifen, in dem der Programmcode und die *Algorithmen* veröffentlicht werden. Im Kapitel „0

---

<sup>22</sup> [http://www.datenschutz-berlin.de/doc/int/iwgdpt/find\\_de.htm](http://www.datenschutz-berlin.de/doc/int/iwgdpt/find_de.htm). Stand: 7.2.2005.

<sup>23</sup> §6 Teledienste-Gesetz: § 6 Anbieterkennzeichnung

„Diensteanbieter haben für ihre geschäftsmäßigen Angebote anzugeben:

1. Namen und Anschrift sowie
2. bei Personenvereinigungen und –gruppen auch Namen und Anschrift des Vertretungsberechtigten.

Weitergehende Informationspflichten nach dem Fernabsatzgesetz, dem Fernunterrichtsschutzgesetz oder dem Teilzeit-Wohnrechtgesetz bleiben unberührt.“

<sup>24</sup> <http://www.datenschutzzentrum.de/audit/>. Stand: 7.2.2005.

6. Kurz vorgestellt: Google-Tipps und Alternativen“ werden einige der Suchmaschinen erläutert und Tipps zur Nutzung gegeben.

Viele schöpfen die Möglichkeiten von Suchmaschinen nicht aus: Lediglich die ersten Treffer zwischen 1 und 20 finden größte Beachtung und werden verwertet. Es ist wichtig, die Suche so genau wie möglich durchzuführen. Hierbei ist zu beachten:

- Alle Suchmaschinen bieten jenseits vom einfachen Eingabefeld zahlreiche Möglichkeiten, die Resultate zielgerecht einzugrenzen.
- Sofort sichtbar ist auf einigen Startseiten der Suchmaschinen zum Beispiel die Option, die Suche auf Webseiten aus Deutschland oder deutschsprachige Angebote zu beschränken.
- Dass darüber hinaus auch nur spezifische Seiten durchsucht oder auch nicht durchsucht werden können, versteckt sich hinter den „Suchoptionen“.
- Oft lässt sich auch einstellen, welche Art von Dokumenten gesucht wird und wie alt diese sein dürfen.

Diese „versteckten“ Funktionen der Suchmaschinen müssen stärker beworben werden. Hier können die Suchmaschinen selbst eine wichtige Rolle spielen. Sie können die Nutzerinnen und Nutzer „an die Hand“ nehmen und ihnen die gängigen Spezifizierungstricks nahe bringen. Denkbar wäre hier eine „Guided Tour“ in der Art einer „Bibliotheksführung“, in der die Techniken und Tricks grafisch aufbereitet oder animiert erklärt werden. Ganz nebenbei würden sie damit auch ihr eigenes Produkt stärken.

Nicht zuletzt liegt die Verantwortung im Umgang mit den Nutzungsdaten auch bei den Nutzerinnen und Nutzern selbst. Daher müssen auch die Abwehrmöglichkeiten der Nutzerinnen und Nutzer gestärkt werden. Um diesen Schutz zu gewährleisten, müssen Möglichkeiten des anonymen und pseudonymen Surfens stärker bekannt gemacht werden. Es gibt bereits jetzt viele Möglichkeiten, mit denen die Nutzerinnen und Nutzer ihre Identität verschleiern können.<sup>25</sup> So lassen sich beispielsweise für die Nutzung von Diskussionsforen zweite E-Mail Adressen einrichten, in der die Nutzerin oder der Nutzer die Schreibweise des Namens variiert – damit ist die Identität zumindest verschleiert. Auch die Beiträge, die man vor Jahren im *Usenet* geschrieben hat, lassen sich zumindest aus der Suche mit Google News entfernen, indem man Google direkt dazu auffordert oder ein kleines Hilfsprogramm nutzt.<sup>26</sup>

Der Zugang zum Internet muss so offen und transparent wie möglich sein. Wir halten viele verschiedene Angebote hier für wichtig. In einer Dominanz, wie Google sie derzeit hat, liegt eine große Gefahr für den freien Informationszugang. Es muss sorgfältig überprüft werden, wie mit einer solchen Dominanz eines Marktteilnehmers in Zukunft umgegangen wird. Alle relevanten Akteure sind hier gefragt, das nötige Werkzeug zu entwickeln, um den negativen Effekten entgegenwirken zu können.

---

<sup>25</sup> Dem Staat blieben auch bei aktiver Förderung des anonymen oder pseudonymen Surfens bei strafrechtlich relevanten Inhalten ausreichende Möglichkeiten der Verfolgung.

<sup>26</sup> <http://groups.google.com/googlegroups/help.html#9>. Stand: 5.2.2005.

## 6. Kurz vorgestellt: Google-Tipps und Alternativen

Gute Überblicke über Suchmaschinen bei <http://www.suchlexikon.de/> und <http://www.toool.de/>.

### Allgemeine Suchmaschinen

Google

[www.google.de](http://www.google.de)

- ⇒ Wie sich die Suche bei Google genauer eingrenzen lässt und beispielsweise bestimmte Wörter oder Seiten ausschließt, wird erklärt unter:  
<http://www.google.de/intl/de/help/refinesearch.html>.  
[www.google.de](http://www.google.de) → „Alles über Google“ → „Suchhilfe“ → „Verfeinern Ihrer Suche“
- ⇒ Über eine Eingabemaske lassen sich bestimmte Kriterien genauer definieren, z.B. die Sprache, welche Dateiformate gesucht werden oder wie aktuell die Information sein sollen.  
[http://www.google.de/advanced\\_search?hl=de](http://www.google.de/advanced_search?hl=de)  
[www.google.de](http://www.google.de) → „Erweiterte Suche“
- ⇒ <http://www.google.de/intl/de/features.html>: hier werden alle Google-Funktionen vorgestellt. [www.google.de](http://www.google.de) → „Alles über Google“ → „Google-Funktionen“

### Weitere allgemeine Suchmaschinen

Yahoo: [www.yahoo.de](http://www.yahoo.de)

MSN: [search.msn.de](http://search.msn.de)

A9: [www.a9.com](http://www.a9.com)

Fireball: [www.fireball.de](http://www.fireball.de)

Web.de: [www.web.de](http://www.web.de)

### Metasuchmaschinen

Metager

[www.metager.de](http://www.metager.de)

Metager ist wohl die bekannteste deutsche *Metasuchmaschine*. Gleich auf der Startseite werden zahlreiche Optionen gegeben, mit denen sich die Suche eingrenzen lässt. So können beispielsweise die Suchmaschinen ausgewählt werden, die Metager in die Suche einbezieht. Es lassen sich die Anzahl der Treffer ebenso festlegen wie die Anzahl möglicher doppelter Treffer.

- ⇒ Die jeweiligen Präferenzen lassen sich dauerhaft einstellen und sind bei jeder Suche verfügbar.  
<http://www.metager.de/index0-config.html>  
[www.metager.de](http://www.metager.de) → „Persönliche Einstellungen speichern“
- ⇒ Der Q-Assoziator gibt Hilfe, wenn man mit den Suchwörtern nicht die gewünschten Ergebnisse findet, und liefert verwandte Begriffe.  
<http://nhf4.rrzn.uni-hannover.de/~mr/asso.html>  
[www.metager.de](http://www.metager.de) → „Q-Assoziator“

### **Weitere Metasuchmaschinen**

Metacrawler: [www.metacrawler.de](http://www.metacrawler.de)

Metaspinner: [www.metaspinner.de](http://www.metaspinner.de)

### **Spezialisierte Suchmaschinen**

Forschung

[www.forschungsportal.net](http://www.forschungsportal.net)

Forschungsportal.net ist ein Projekt des Bundesministeriums für Bildung und Forschung. Die Suche lässt sich auf Forschungsinstitute oder Texte beschränken.

Praktischerweise sind die Operatoren, wie man Suchbegriffe kombinieren kann, unter dem Eingabefenster aufgeführt.

## 7. Glossar

Algorithmus	<p>„Ein Algorithmus ist eine genau definierte Berechnungsvorschrift zur Lösung eines Problems oder einer bestimmten Art von Problemen.“ Aus: Wikipedia, <a href="http://de.wikipedia.org/wiki/Algorithmus">http://de.wikipedia.org/wiki/Algorithmus</a>. Stand: 27.12.2004.</p>
Betriebssystem	<p>„Ein Betriebssystem (BS) ist die Software, die die Verwendung (den Betrieb) eines Computers ermöglicht. Es verwaltet Betriebsmittel wie Speicher, Ein- und Ausgabegeräte und steuert die Ausführung von Programmen.“ Aus Wikipedia, <a href="http://de.wikipedia.org/wiki/Betriebssystem">http://de.wikipedia.org/wiki/Betriebssystem</a>. Stand: 17.1.2005.</p>
Browser	<p>„Webbrowser oder Browser (engl. Für „Stöberer“) sind Computerprogramme, die zum Betrachten verschiedener Arten von Dokumenten dienen. Vorwiegend werden sie verwendet, um (HTML-)Seiten aus dem Internet anzuzeigen. [...]</p> <p>Ursprünglich bezeichnet der aus dem Englischen entlehnte Begriff browsen (engl. To browse = stöbern, sich umsehen, schmökern) am Computer lediglich das Benutzen von Navigations-Elementen (Vor, Zurück, Index, ...) zum Lesen von Texten.“ Aus: Wikipedia, <a href="http://de.wikipedia.org/wiki/Browser">http://de.wikipedia.org/wiki/Browser</a>. Stand: 17.1.2005.</p> <p>Gängige Browserprogramme sind: Netscape Navigator, Microsoft Internet Explorer, Firefox, Mozilla, Opera, Safari, u.a.</p>
Cookie	<p>„Ein Cookie [...] ist ein kurzer Eintrag in einer meist kleinen Datenbank auf einem Computer und dient dem Austausch von Informationen zwischen Computerprogrammen oder der zeitlich beschränkten Archivierung von Informationen. [...]Webbrowser stellen eine Cookie-Datenbank zur Verfügung, die Cookie Cache genannt wird; dort kann der Webserver einer besuchten Webseite Informationen in Form von HTTP-Cookies hinterlegen und bei einem Wiederbesuch der Seite auslesen.“ Aus: Wikipedia, <a href="http://de.wikipedia.org/wiki/Cookie">http://de.wikipedia.org/wiki/Cookie</a>. Stand: 9.1.2005.</p>
Deep Web	Siehe Visible Net
Desktop	<p>Der englische Begriff Desktop wird für Computer mit grafischer Oberfläche benutzt. Darüber hinaus kann noch ein bestimmter Teil der grafischen Oberfläche gemeint sein, nämlich alle Dateien oder Programme, die auf dem ersten sichtbaren Bildschirm aufgeführt sind („alles was auf dem Schreibtisch liegt“). Auch der gesamte stationäre Rechner wird, im Gegensatz zum tragbaren Laptop, als Desktop bezeichnet („auf dem Tisch“ anstelle von „auf dem Schoß“).</p>
Domain	<p>„Eine Domäne (englisch: domain) ist ein zusammenhängender Teilbereich des [...] Namensraumes. Eine Domäne umfasst ausgehend vom ihrem Domänennamen immer die gesamte untergeordnete Baumstruktur. [...] Domänen der höchsten Ebene (z.B. .de) werden auch Top-Level-Domains genannt. Dementsprechend heißen die Domänen zweiter</p>

und dritter Ebene Second- beziehungsweise Third-Level-Domains.“ Aus: Wikipedia, <http://de.wikipedia.org/wiki/Domain>. Stand: 15.1.2005.

Alle Domains, die auf .de enden, werden bei Denic ([www.denic.de](http://www.denic.de)), der zentralen Registrierungsstelle in Deutschland, registriert.

Flash	Flash-Inhalte sind in der Regel animierte Grafiken oder Filme. Flash-Animationen finden sich heute auf vielen Webseiten. Sie können auch dafür genutzt werden, komplexe Zusammenhänge in animierten Grafiken darzustellen.
Gatekeeper	„Als Gatekeeper [...] bezeichnet man in der Publizistik- und Kommunikationswissenschaft metaphorisch einen (meist personellen) Einflussfaktor, der darüber entscheiden kann, welche Nachricht in den Medien erscheint.“ Aus: Wikipedia, <a href="http://de.wikipedia.org/wiki/Gatekeeper">http://de.wikipedia.org/wiki/Gatekeeper</a> . Stand: 27.12.2004.
Invisible Net	Siehe Visible Net
IP-Adresse	Jeder mit dem Internet verbundene Computer hat für die Dauer der Sitzung eine eindeutige Identifikationsnummer, die IP-Nummer. IP steht für Internet Protocol, einer grundlegenden Vereinbarung zum Austausch von Daten zwischen zwei Rechnern. Die Nummer besteht momentan aus vier Blöcken je drei Ziffern, also zum Beispiel: 160.219.70.35. Anhand der IP-Adresse ist ein einzelner Rechner oder ein Rechnernetz eindeutig identifizierbar.
(Web-)Kataloge	Bei Webkatalogen sammeln und bewerten Redakteure neue oder veränderte Webseiten und kategorisieren diese. Bei den meisten Katalogen finden sich kurze Zusammenfassungen der einzelnen Seiten und Beschreibungen der Kategorien. Darüber hinaus werden oft Links zu verwandten Themen angegeben, oder, wie bei Yahoo, Links zu den Kategorien im englischsprachigen Angebot. Leider erfordern diese Arten von Webkatalogen einen sehr hohen Personalaufwand und sind daher meist auf bestimmte Sachgebiete spezialisiert.
Metasuchmaschine	Eine Metasuchmaschine leitet die Suchanfrage an mehrere Suchmaschinen weiter und stellt die besten Treffer der jeweiligen Suchmaschine auf einer Ergebnisseite dar. Das Problem der Abhängigkeit von einer Suchmaschine wird mit Metasuchmaschinen somit abgeschwächt.
Newsgroup / Newsreader	„Newsgroups [...] dienen dem Austausch von Nachrichten [...]“ und sind so eine Art öffentlicher Plattform, analog vielleicht zum Schwarzen Brett. „Weltweit existieren seit Jahrzehnten verschiedene News-Systeme. Das bekannteste ist dabei das Usenet, das unzählige Newsserver weltweit verbindet. Oft sind Newsgroups aber auch rein regional verteilt, oder werden in Intranets zur Verfügung gestellt. Newsgroups sind wie E-Mail ein Dienst im Internet. Das Usenet ist deutlich älter als das World Wide Web [...]“. Aus: Wikipedia, <a href="http://de.wikipedia.org/wiki/Newsgroup">http://de.wikipedia.org/wiki/Newsgroup</a> . Stand: 22.1.2005.  Die Anzahl der weltweiten Newsgroups ist nicht exakt zu beziffern und geht weit über 50.000 hinaus und es werden sowohl für die Öffentlichkeit interessante als auch eher private Themen diskutiert.
Open Source	„Der Ausdruck Open Source [...] steht für quelloffen, einerseits in dem

Sinne, dass der Quelltext eines Programms frei erhältlich ist, andererseits für ‚offene Quelle‘, also dass ein Werk frei zur Verfügung steht. [...]

Dabei beschränkt sich der Begriff aber nicht ausschließlich auf Software, sondern kann auch auf Wissen und Information allgemein ausgedehnt werden. [...] Der Ausdruck Open-Source-Software [...] wird zu Recht als Synonym für „freie Software“ verwendet, wenn auch die Verfechter der jeweiligen Begriffe verschiedene Schwerpunkte setzen.“ Aus: Wikipedia, [http://de.wikipedia.org/wiki/Open\\_source](http://de.wikipedia.org/wiki/Open_source). Stand: 20.1.2005.

Ranking	„(B)ei der Abfrage von Suchmaschinen werden die Ergebnisse nach einem „ranking“ dargestellt, nämlich in einer Reihenfolge, wo die ersten Treffer die größte Relevanz für den Suchbegriff bilden. [...] Welche Kriterien für das ranking der Suchmaschinen herangezogen werden, ist meist ein Firmengeheimnis der jeweiligen Suchmaschine.“ Aus: Wikipedia, <a href="http://de.wikipedia.org/wiki/Ranking">http://de.wikipedia.org/wiki/Ranking</a> . Stand: 17.1.2005.
Spider (-programm)	Spider, auch Robots, Bots oder Crawler genannt, sind kleine Programme, die das WWW systematisch nach Webseiten durchsuchen. Dabei „hangeln“ sich die Programme von Link zu Link und geben die gefundenen Daten an die zentralen Rechner der Suchmaschinenbetreiber weiter. Diese Daten werden dann nach Schlagwörtern durchsucht und indiziert
Usenet	Siehe Newsgroup
Visible Net	Die über Suchmaschinen zugänglichen Webseiten werden Visible Net (Sichtbares Netz) genannt. Es wird geschätzt, dass Google etwa 60 Prozent der deutschen Webseiten erfasst. Die verbleibenden 40 Prozent gehören zum so genannten Deep Web oder Invisible Net. Vor allem Datenbanken zählen zum Invisible Net, weil diese die Seiten erst nach der konkreten Anfrager der Nutzerinnen und Nutzer erstellt werden und so nicht von den Suchmaschinen indiziert werden können. Deswegen sind viele hochwertigen Informationen nicht über heutige Suchmaschinen verfügbar.
World Wide Web	„Das World Wide Web (kurz das Web oder WWW, wörtlich: Weltweites Gewebe/Netz) ist ein Hypertext-System, das über das Internet abgerufen werden kann. [...] Der Benutzer kann den Hyperlinks im Dokument folgen, die auf andere Dokumente verweisen, egal ob sie auf dem gleichen Webserver oder einem anderen gespeichert sind. Hierdurch ergibt sich ein weltweites Netz oder Gewebe aus Webseiten. Das Verfolgen der Hyperlinks wird oft als "Surfen im Web" bezeichnet.  Das WWW wird im allgemeinen Sprachgebrauch oft mit dem Internet gleichgesetzt, obwohl es jünger und nur eine mögliche Nutzung des Internets ist [...]. Aus: Wikipedia, <a href="http://de.wikipedia.org/wiki/Www">http://de.wikipedia.org/wiki/Www</a> . Stand: 17.1.2005.

