

GOGGLE, GOSSIP & PR-OSTITUTION

DAS GESCHÄFT EINER SUCHMASCHINE

HENDRIK SPECK UND FRÉDÉRIC PHILIPP THIELE

Jean Baudrillard, der französische Philosoph und Hyperrealist, beschreibt in »Disney World Company« eine Welt, die sich durch die Globalisierung, Vernetzung, und Synchronisierung von Informationen auszeichnet. Der schnelle, uneingeschränkte Zugriff auf neueste Informationen wird darin immer wichtiger und zwanghafter.

Einzig automatisierte Suchmaschinen scheinen in der Lage, dieser ständig anwachsenden Informationsflut Paroli zu bieten und das schnelle Auffinden von Informationen zu gewährleisten. Ein Drittel der Bevölkerung sucht täglich auf diesem Wege nach gewünschten Resultaten. Der Marktführer Google verzeichnet täglich über 200 Millionen Suchanfragen [1]. Der Verlust der editorialen Filter- und Kontrollebene wird jedoch von der Mehrheit noch nicht wahrgenommen – die medienpolitische Einordnung von Suchmaschinen findet kaum statt. Sie werden als unfehlbare, allwissende Götter angesehen.

Der relativ junge Suchmaschinenmarkt hat sich in den letzten Jahren stark gewandelt: In seiner Sturm- und Drangphase noch durch eine Vielzahl kleinerer Suchmaschinen gekennzeichnet, wird der Markt momentan von Google, Microsoft und Yahoo! klar dominiert. Diese Monopolstellung und damit zwangsläufig verbundene Machtpositionen beeinflussen unseren heutigen Zugang zu Wissen. Wenige Suchmaschinenbetreiber entscheiden dabei über die Auswahl, Reihenfolge und Bewertung von Informationen, ohne dass die zu Grunde liegenden Maßstäbe transparent gemacht werden. Verschärfend kommt hinzu, dass das Ranking gerade im kommerziellen Bereich erhebliche finanzielle Konsequenzen haben kann, die kommerzielle Anbieter zu massiven »Lobby- und Manipulationstätigkeiten, so genanntem Search-Engine-Marketing motivieren. Diese Search-Engine-Optimierungen (SEO) sind verstärkt im Vermittler- und Resellerbereich zu verzeichnen, dem so genannten Affiliate Business, wo solche Manipulationen kombiniert mit der Allmacht der Suchmaschinen über die von ihnen indizierten Dokumente, die wesentlichen Probleme und Schwachstellen heutiger Suchmaschinen charakterisieren.

»Our defenses against information glut have broken down; our information immune system is inoperable. We don't know how to filter it out; we don't know how to reduce it; we don't know to use it. We suffer from a kind of cultural AIDS« (Neil Postman 1990).

Die Gestaltung eines freien Informationszugangs erweist sich als wesentliche Herausforderung der Informationsgesellschaft. Zum jetzigen Zeitpunkt sind Suchmaschinen weder strukturell noch ethisch für die vor ihnen liegenden Aufgaben gewappnet. Weder eine entsprechende Selbstverpflichtung, noch ein ihrer Bedeutung entsprechender Kodex liegen vor – die internen Abhängigkeiten und Beziehungen zwischen den einzelnen Suchmaschinen sind gleichfalls nicht dokumentiert bzw. werden durch unterschiedliche Marken kaschiert.

PR-ostitution

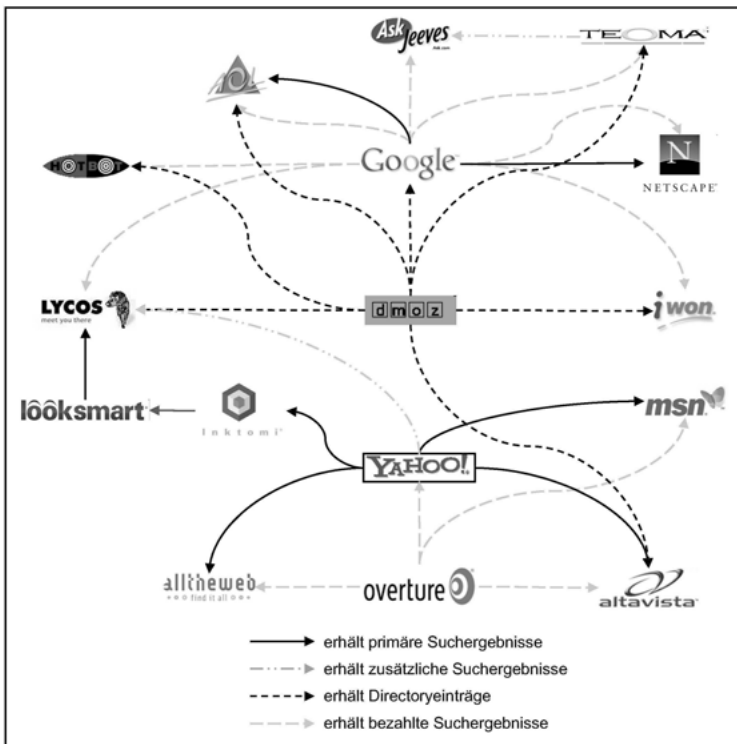
Das Geschäft der Suchmaschinen

Suchmaschinen präsentieren sich heutzutage wie allwissende Götter, sie versprechen innerhalb von Sekundenbruchteilen auf jede nur erdenkliche Frage die perfekte Antwort zu liefern. Doch leider sieht die Realität ganz anders aus. Heutige Suchmaschinen erfüllen ihre theoretisch angedachten Aufgaben nur teilweise und haben noch mit einer Reihe gravierender Probleme zu kämpfen. Qualität, Aktualität, Ordnung und Umfang der Suchergebnisse bestimmen im Wesentlichen die Qualität von Suchmaschinen – kommerzielle Orientierung, Zensur, technologischer Entwicklungsrückstand und fehlende Transparenz der Rankingverfahren definieren hingegen die Entfernung von der ursprünglichen Idee einer idealen Suchmaschine.

Die Qualität der heutigen Suchmaschinen leidet stark darunter, dass die proprietären Evaluationskriterien der einzelnen Suchmaschinen nicht dokumentiert sind und sich somit keiner öffentlichen Auseinandersetzung stellen müssen. Dies führt zu Marktverzerrungen insbesondere in kommerziell ertragreichen Marktsegmenten, da einzelne Marktteilnehmer durch Investition und Manipulation Vorteile gegenüber anderen Wettbewerbern gewinnen können – auf Kosten der Objektivität der Suchergebnisse und zu Lasten des Verbrauchers. Suchmaschinen profitieren sogar noch durch diesen Trend, die von ihnen angebotene beziehungsweise in den Suchergebnissen versteckte Werbung ist für den Großteil ihrer (enormen) Gewinne verantwortlich.

Mittlerweise ist eine ganze Branche mit kommerziellen Angeboten zum gezielten Manipulieren einzelner Marktsegmente entstanden. In der Folge liefern viele Suchanfragen nur noch unbrauchbaren, kommerziel-

len Spam, der für den Benutzer kaum Informationsinhalt besitzt. Durch eine fehlende Transparenz der Rankingalgorithmen der einzelnen Suchmaschinen ist es für den Nutzer nicht nachvollziehbar, warum eine bestimmte Seite ein bestimmtes Ranking hat. Er weiß nicht, ob ein gutes Ranking auf Grund des guten Inhalts zu Stande kommt, oder ob das Ranking durch den Besitzer der Seite erkaufte wurde. Auf eine gezielte Produktsuche bekommt der Nutzer meist nur noch kaum verschlüsselte Kaufangebote diverser Affiliatestrukturen, bei denen verschiedene Mittelsmänner am Vertrieb ein und desselben Produktes durch Provisionen beteiligt sind – eine Informationsvermittlung findet nicht statt.



Grafik 1: Suchmaschinenmarkt

Die direkten Geschäftsbeziehungen, insbesondere »Paid-Results« und Werbung, zwischen Suchmaschinen und Content-Providern, bilden dabei ein wesentliches Hindernis für die Beseitigung des Problems. Nutzern ist im Allgemeinen auch nicht bewusst, welche Suchergebnisse von welchen Datenbanken und Inhaltsanbietern zugeliefert werden (vgl. Grafik 1) – manche Suchmaschinen vermischen außerdem Suchergebnisse, Verzeichniseinträge, bezahlte Suchergebnisse und Werbung.

Problematisch erweisen sich zusätzlich die Ausrichtung und die Schwachstellen der gegenwärtigen Suchmaschinen-Algorithmen, die, soweit publiziert, auf Verfahren beruhen, die auf Linkpopularität basieren. Diese Algorithmen betrachten im Allgemeinen Links als »Stimmen« bzw. als Qualitätsindikatoren für den Inhalt bestimmter Dokumente. Sie sind nicht in der Lage, den Unterschied zwischen dem populärsten Dokument und dem qualitativ hochwertigsten Dokument festzustellen – ein auf Popularität bzw. PageRank (PR) (vgl. Brin/Page 1998) basiertes Verfahren setzt diese im Gegenteil gleich; Gossip wird dabei ein dem Inhalt nicht entsprechender Informationsgehalt zugeschrieben. Dies führt zu weiteren Manipulationsansätzen, der so genannten PR-ostitution, bei der Internet-Content-Provider ihren PageRank künstlich durch technische bzw. kaufmännische Methoden erhöhen, um Suchmaschinen eine höhere Qualität vorzuspiegeln.

Die proprietären Monopolstrukturen und die fehlende öffentliche Auseinandersetzung dokumentieren sich auch in der unzureichenden Evolutionsgeschwindigkeit heutiger Suchmaschinen [2]. Die Anpassung an Nutzerbedürfnisse, neue Technologien oder neue Dokumenttypen demonstriert die fehlende Innovationsbereitschaft der Informationsmonopole. Die mangelhafte Integration und Indizierung von Macromedia-Flash-Dateien oder Open-Office-Dateien in den heutigen Suchmaschinen zeigt, wie langsam auf entsprechende Veränderungen im Netz reagiert wird. Das Multimedia-Format Macromedia Flash wird von einer Mehrheit der Browser interpretiert; viele Websites setzen das Format für interaktive Webapplikationen ein; der Hersteller Macromedia stellt ein seit Jahren dokumentiertes Interface zu Indizierung von Flash-Dateien zur Verfügung [3] – bis zum jetzigen Zeitpunkt sind Flash-Inhalte jedoch in kaum einer Suchmaschine komplett integriert.

Die oben bereits erwähnte Profitorientierung von Suchmaschinen steht nicht nur im krassen Gegensatz zur europäischen Aufklärung und zu offenen Bildungssystemen – sie konterkariert auch die Bestrebungen objektive Suchergebnisse zu liefern. Das Hauptziel einer Suchmaschine ist oft nicht der Nutzen für den Anwender, sondern die kommerzielle Verwertbarkeit – und wird auch offen eingestanden: »Currently, the predominant business model for commercial search engines is advertising. The goals of the advertising business model do not always correspond to providing quality search to users« (Brin/Page 1998) [4].

Börsengang, Shareholder-Orientierung und Kommerzialisierung des Marktführers Google sowie die Gegenbestrebungen anderer Wettbewerber und des Betriebssystemmonopolisten Microsoft führen das Ideal einer objektiven Suchmaschine ad absurdum. Microsoft hat dabei insbesondere die Bedeutung einer integrierten Suchfunktion erkannt und bindet

entsprechende Schnittstellen in den Windows-XP-Nachfolger Longhorn ein. Der Vorteil für den Nutzer ist jedoch nicht klar ersichtlich, da Microsoft seine kommerziellen Interessen wahrscheinlich auch hier über dessen Interessen stellen wird – das von vielen Anti-Trust-Gerichten verurteilte Verhalten Microsofts im Browserkrieg Netscape/ Internet Explorer lässt dabei nichts Gutes erwarten.

Zensur und Abhängigkeit

Die Marktführer der Suchmaschinenindustrie sind global operierende Unternehmen, deren Einkommensquellen im Wesentlichen durch Werbung, bezahlte bzw. manipulierte Ergebnisse, ausgespähte Nutzerprofile sowie demographische Informationen bestimmt werden. Diese bieten ihre Produkte weltweit an und verfügen über Verkaufsbüros, Firmensitze und Gesellschaften in vielen Ländern. Dies zwingt sie im gleichen Atemzug zur Einhaltung der jeweiligen nationalen bzw. regionalen gesetzlichen Bestimmungen. Einzelne Nationalstaaten versuchen dabei ihre nationalen, politisch oder religiös diktierten Interessen über die Interessen eines hypernationalen Mediums zu stellen.

Dies gilt insbesondere für die Vereinigten Staaten von Amerika, die mit ihren entwicklungshemmenden Patentverfahren und einseitigen Auslegungen von Intellectual Property de facto eine ökonomische Kriegsführung (»Lawfare«) auf der Basis von Handelsabkommen und -begünstigungen führen. Da alle Marktführer ihre Firmenzentralen in den USA haben, wird dem internationalen Nutzer praktisch das amerikanische Werte- und Rechtssystem aufgezwungen. Das Internet wandelt sich damit von einem nationalitätsfreien Raum zu einem Raum, in dem Kleinstaaterei zum politischen Konzept gemacht wird. Problematisch ist, dass diese Zensur für den Nutzer nicht transparent ist. Dies wird unter anderem an der Zensur von Suchergebnissen mit erotischem Inhalt, der chinesischen »Firewall« [45], den Kontrollbestrebungen von Scientology sowie den Zensurversuchen regionaler Regierungsbehörden in Deutschland deutlich.

Da das World Wide Web auf seinen dezentralen, paketerorientierten Protokollen und Strukturen aufbaut, sind regionale Zensurbestrebungen meist zum Scheitern verurteilt. Nur wenige Zensurprojekte erweisen sich dabei als so ambitioniert wie die »Firewall« der chinesischen Regierung, zumeist genügt die Ausnutzung der regionalen Gesetzgebungen um Suchmaschinen-Betreiber mit Hinweis auf die jeweiligen Marktchancen zum vorauseilenden Gehorsam und zur Selbstzensur zu bewegen. Verschiedene Filterbestrebungen und Zensur versuchen dennoch politisch oder religiös unliebsame Meinungen zu unterdrücken – dies findet z.B. direkt am Arbeits-PC statt, wie in amerikanischen Bibliotheken, denen

staatliche Fördergelder vorenthalten werden, wenn keine – undokumentierten – Zensurprogramme installiert sind (vgl. Wallace 1998).

Diese Selbstzensur und Bevormundung der Bürger zeigt sich sehr deutlich beim Anfang 2005 gestarteten Dienst »MSN Search« von Microsoft [6], der sich bei sexuellen, anstößigen oder anderweitig zensurbedürftigen Suchbegriffen einerseits sehr bedeckt verhält (geben Sie einmal »Adolf Hitler« oder »Taschengeld« als Suchbegriff ein), andererseits direkt auf »Geschäftspartner« von MSN verlinkt. Microsoft begründet diese Zensur mit dem deutschen Jugendmedienschutz-Staatsvertrag, der am 1. April 2004 in Kraft trat. Das fatale Versagen der angewandten Filter wird durch hinreichend dokumentierte Suchbegriffe deutlich, die katastrophale semantische Beziehungen herstellen, auf Grund obskurer Begründungen eine Zensur durchführen und gegen das Grundgesetz der Bundesrepublik Deutschland verstoßen. Im Fall von MSN Search bleibt zu hoffen, dass Microsoft seine Interpretation des deutschen Jugendschutzes überdenkt und eine Diskussion über die eingesetzten semantischen Filter befördert.

Datenschutz

Suchmaschinen sind mitverantwortlich für den schleichenden Verlust der Privatsphäre. Kommerzielle Anbieter, die versuchen, Werbung und Produktempfehlungen immer mehr auf die Bedürfnisse einzelner Nutzergruppen zuzuschneiden, schaffen immer detailliertere Profile von Nutzern. Beispielhaft setzen sich dabei die Suchmaschine A9 von Amazon, der E-Mail-Service Gmail von Google sowie deren Toolbar (mit erweiterten Funktionen) über Datenschutzbedenken hinweg und schaffen die technischen Voraussetzungen für die Bespitzelung der Nutzer. Global agierende Konzerne wie Amazon und Google verfügen dabei über Daten von Nutzern aus der ganzen Welt (vgl. Grafik 2). Google kann zur Erzielung seiner Profite nicht nur auf die (unpersönlichen) Nutzerdaten von Suchportal und Toolbar zählen, sondern diese mit den Interessen, Beziehungen, Konsumgewohnheiten und Adressdaten der Nutzer von Froggle, Google News, Blogger, AdWords/AdSense, Gmail, und Orkut abgleichen. Die angebotenen Desktop-Suchen von Microsoft und Google können dabei endgültig die noch verbliebenen Unterschiede von Öffentlichkeit und Privatsphäre aufheben.

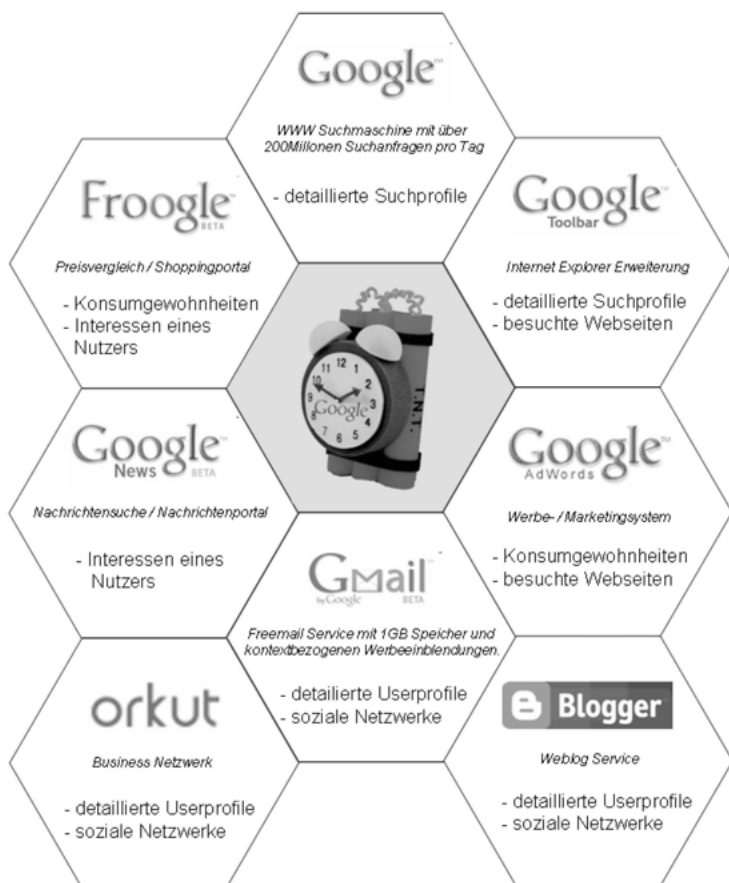
»Some features available on the Google Toolbar can enhance your search and browsing experience, but may require that we have some understanding of the web pages that you are viewing. For example [...] knowing which web page you are viewing [...]. Google can also use this information about the pages you have viewed [...] We believe these features will greatly enhance

your browsing experience. We understand, however, you may not agree that the benefits provided by these features outweigh the information we must collect« (Google Toolbar Privacy Policy) [7].

Mit seiner im April 2004 veröffentlichten Suchmaschine A9.com geht Amazon [8] dabei weiter als die Konkurrenz. Um bei Amazon Zugriff auf erweiterte Suchfunktionen zu haben, muss der Nutzer sich mit seinen (Amazon-)Kundendaten einloggen. Amazon kann dadurch sowohl die Suchanfragen, die besuchten Webseiten, Interessen und Bestellungen als auch die individuellen Nutzerdaten einschließlich Bankverbindung und Zahlverhalten verknüpfen. Dies geschieht laut Amazon natürlich nur, um dem Nutzer die Möglichkeit zu geben, später Seiten leichter wieder zu finden. Amazon weiß dabei nicht nur welche Informationen, Produkte, Bücher, DVDs oder CDs einen Nutzer interessieren, sondern verfolgt gleichzeitig, welche Webseiten er wann besucht, wie lange er sich auf der diesen aufgehalten hat, welche Themengebiete ihn interessieren und noch vieles mehr.

Amazon hat dabei äußerst erfolgreich die eigene Produktpalette diversifiziert, um zum Beispiel dank der angebotenen A9-Toolbar, des integrierten Ranking-Systems von Alexa und der Einbindung der Internet Movie Database weitere Nutzerinformationen zu gewinnen. Amazon bietet außerdem eine Visa-Kreditkarte an, die mit ihrem auf Prozentpunkten basierten Rabattsystem das komplette Käuferverhalten der einzelnen Nutzer auch außerhalb von Amazon transparent macht. Zum Firmenimperium von Amazon gehören neben den nationalen Amazon-Portalen auch die Portale von Borders, CDNow, Waldenbooks, Virginmega, Target und Marshall Field sowie die gleichfalls integrierten Angebote von Circuit City, J&R Music and Computer World, Office Depot, Toys R Us, babiesRus, Expedia, Hotwire, Drugstore, Ofoto, Egghead und JCPenny. Die Auswirkungen sind dem einzelnen Nutzer dabei kaum bewusst.

Mit seinem gleichfalls im Test befindlichen E-Mail-Service Gmail überschreitet auch Google die Grenzen: Um weiteren Webspace für das Anbieten seiner Werbebanner bereitstellen zu können, bietet Google interessierten Nutzern einen mit einem Speicherplatz von einem Gigabyte ausgestatteten kostenlosen E-Mail-Account an.



Grafik 2: Google und zugehörige Unternehmen

Google verheißt eine lebenslange Aufbewahrung aller E-Mails, die eigentlichen Geschäftsabsichten und Profilierungsbestrebungen aber verschwinden demgegenüber im Kleingedruckten. Gmail benutzt ein automatisches System, welches die E-Mails des Nutzers analysiert, entsprechende Nutzerprofile erstellt und dann zum Inhalt jeder E-Mail passende Werbung einblendet.

Die auf der Basis der gewonnenen Nutzerdaten angebotene Werbung erwirtschaftet höhere Gewinne für Google, als die unpersonalisierten Werbeeinblendungen der Konkurrenz. Google hat damit die technischen Möglichkeiten, um E-Mail und Suchanfragen sowie das individuelle Surfverhalten zu verknüpfen und könnte dabei die auf Millionen Webseiten erscheinende Google-AdWord-Werbung an das jeweilige Profil anpassen.

Auf ihr Image bedachte Unternehmen wie Amazon oder Google versichern dabei durchaus, dass die gewonnenen Daten nur zu Zwecken der Verbesserung des Service genutzt werden und nicht an Dritte weitergegeben werden. Dabei wird aber eine kleine Hintertür verschwiegen: Auf Grund der Bestimmungen des Patriot Act (Provide Appropriate Tools Required to Intercept and Obstruct Terrorism), der infolge der Anschläge vom 11. September 2001 zur Bekämpfung des Terrorismus erlassen wurde, sind alle US-Firmen dazu verpflichtet, auf Verlangen entsprechender Regierungsstellen alle gespeicherten Nutzerdaten an diese weiterzugeben. Der Verdacht ist dabei hinreichend, Firmen sind verpflichtet über den Vorgang absolutes Stillschweigen zu bewahren, eine gerichtliche Verordnung ist nicht erforderlich, der (internationale) Nutzer wird nicht informiert.

Bei Amazon sowie bei Gmail [9] gibt es keine Möglichkeit, seinen Account und die damit verbundenen Daten komplett zu löschen – der Nutzer hat das Recht über seine informationelle Selbstbestimmung endgültig verloren – demokratische Grundsätze sind außer Kraft gesetzt. Die rechtliche Zulässigkeit und die Auswirkungen dieser Gesetze sind umstritten. Kritiker prophezeien ein Orwell'sches Szenario, bedenklich sind in jedem Fall die internationalen Folgen derartiger Bestrebungen. Firmen wie Amazon und Google, aber auch Fluglinien, Kreditkartenfirmen oder Electronic-Payment-Provider verfügen nicht nur über die Daten ihrer jeweiligen Staatsbürger, sondern über Käufergewohnheiten, Interessen, Adressen, Bankverbindungen, und religiöse Zugehörigkeit von Nutzern aus der ganzen Welt. Zumindest regionale und nationale Polizeibehörden, Institutionen und Geheimdienste haben im rechtlichen Graubereich Zugriff auf umfangreiche Profile von einem Großteil der Weltbevölkerung und umgehen teilweise rechtsstaatliche Kontrollinstanzen.

Sollten die dort eingesetzten Filtermechanismen auf derselben niedrigen Evolutionsstufe operieren wie die im Suchmaschinen-Sektor vorgeführten Beispiele – und einige bekannt gewordene Vorfälle sprechen dafür – dann könnte zum Beispiel die Reisefreiheit oder freie Wohnortnahme durch die Einreise- oder Visa-Verweigerung auf Grund ähnlich absurder und nicht transparenter Gründe behindert werden. Die »Recording Industry Association of America« andererseits hat mit Sicherheit Interesse an Nutzerprofilen, die die Suchbegriffe DivX, Xvid oder MP3 beinhalten.

Fazit

Zum jetzigen Zeitpunkt findet eine Verknüpfung einzelner Nutzerdaten nur begrenzt bzw. in Einzelfällen statt. Allein die Existenz derartig umfangreicher Datensammlungen ist aber schon problematisch, insbesondere wenn technologischer Fortschritt und kommerzielle Interessen die Privatsphäre mit immer neuen Produkten immer weiter unterhöheln. Der Nutzer verliert mehr und mehr die Möglichkeit zur informationellen Selbstbestimmung und wird damit zum Spielball rechtlich fragwürdiger Interessensgruppen.

Diese Entwicklung verdreht das allgemeine Organisationsmodell des Internet, in dessen dezentralem System eines weltweiten Datenaustausches und entsprechender Kommunikation nur der Zugangspower nach genau festgelegten Rahmenkriterien Nutzerdaten für einen eng umgrenzten Zeitraum speichern darf und kann. Auskünfte seitens der Zugangspower sind strikt reglementiert und sollten nur unter gerichtlicher Aufsicht stattfinden. Durch die globalen Datensammler jedoch findet eine Erosion des dezentralen Prinzips des Internet sowie der Schutzrechte des Einzelnen statt. Kritiker vergleichen bereits die Verletzungen der Privatsphäre durch Suchmaschinen mit Orwell'schen Szenarien, den Anti-Terror-Gesetzen des 21. Jahrhunderts oder dem Dritten Reich. Der Nutzer wird in allen Szenarien zum gläsernen Bürger. Big Brother is watching you!

Literatur

- Postman, Neil (1990): »Informing Ourselves To Death«. German Informatics Society. Stuttgart.
 [WWW-Dokument: www.eff.org/Net_culture/Criticisms/informing_ourselves_to_death.paper - Zugriff: 23.11.2004].
 Wallace, Jonathan (1998): »The Library Lawsuit«. In: *Wired Magazine*.
 [WWW-Dokument: <http://www.wired.com/news/politics/0,1283,11926,00.html> - Zugriff: 23.11.2004].

Digitale Verweise

- [@1] www.google.com/ads/indepth.html
- [@2] www.heise.de/newsticker/meldung/54151
- [@3] www.macromedia.com/software/flash/download/search_engine
- [@4] www-db.stanford.edu/pub/papers/google.pdf
- [@5] <http://cyber.law.harvard.edu/filtering/china>
- [@6] http://dmoz.org/World/Deutsch/Gesellschaft/Religion_und_Spiritualitaet/Religions_und_Sektenkritik/Scientology/Zeitungsartikel/Zensur
- [@7] <http://www.heise.de/tp/r4/artikel/14/14511/1.html>
- [@6] <http://beta.search.msn.de/>
- [@7] <http://toolbar.google.com/privacy.html>
- [@8] <http://a9.com>
- [@9] www.gmail.com