# Datenspuren, Informationssuche und Monopolstrukturen im Internet

Hintergründe und Hinweise zur sicheren Informationsrecherche

Jeder, der irgendetwas im Internet tut, sollte sich darüber im Klaren sein, dass ALLES, was er dort tut, irgendwo gespeichert wird – zumindest kurzzeitig. Das Wort „ALLES" bedeutet wirklich, dass jeder Tastenanschlag und jeder Mausklick grundsätzlich erst einmal gespeichert wird. Diese Tatsache ist immer noch weitgehend unbekannt.

Trotzdem haben wir damit allein noch keinen „Big Brother", der jede unserer Datenspuren kennen würde, denn diese Spuren sind eigentlich auf vielen verschiedenen Rechnern vieler verschiedener Betreiber weit verteilt. Sie werden auch nicht dauerhaft festgehalten, sondern vieles wird nur für sehr kurze Zeiträume gespeichert. Diese Datenspuren können, wenn überhaupt, nur sehr schwer zusammengeführt werden. Das gilt, solange sie wirklich auf vielen Rechnern verteilt sind. Je weiter der Konzentrationsprozess im Internet voran schreitet, desto mehr wird diese inhärente Sicherheit, welche die Struktur des Netzes eigentlich bietet, von den Nutzern jedoch z. T. selbst ausgehebelt.

## Datenspeicherung beim Provider

Die erste Stelle im Internet-Datenstrom des Nutzers ist der Provider: diejenige Firma, welche den Internetzugang bereitstellt. Dort fließt jedes Zeichen der Tastatur und jeder Mausklick über die Verbindungsstelle, den Router, ins globale Internet und wird zumindest kurzzeitig gespeichert. Seit das Gesetz zur Vorratsdatenspeicherung in Kraft ist (ab 1.1.2009), ist jeder Provider sogar verpflichtet, einen Teil dieser Daten für 6 Monate „auf Vorrat" zu speichern. Dazu zählt allerdings nicht jeder Tastenanschlag und jeder Mausklick, sondern „nur":

- die dem Teilnehmer zugewiesene Internet-Adresse, die sog. IP-Adresse – eine eindeutige Kennung des Rechners im Internet aus 4 Ziffernblöcken,
- eine eindeutige Kennung des Anschlusses, über den die Internetnutzung erfolgt (i. A. die Telefonnummer),
- den Beginn und das Ende der Internetnutzung.

Der Rückschluss von den IP-Adressen auf Personen, die diesen Internetzugang benutzen, ist nur auf dem Umweg über die Anschlusskennung (Telefonnummer), von der aus die Einwahl ins Internet erfolgte, möglich. Trotzdem kann der Nutzer meist mit hoher Sicherheit identifiziert werden.

Deutschlands bekanntester Provider ist T-Online, daneben gibt es etliche weitere, die durchaus in echtem Wettbewerb zueinander stehen. Hier hat sich bisher noch kein Monopol etabliert.

## Datenspuren auf Servern

Die nächste Stelle, auf der Internet-Nutzer ihre eindeutigen Spuren hinterlassen, sind die Server: Jeder, der im Netz surft, „bewegt" sich im Allgemeinen mit Mausklicks und Tastatureingaben auf Rechnern, die bestimmte Dienstleistungen bereitstellen – seien es Texte für Kochrezepte, seien es Datingdienste oder was auch immer. Jeder dieser Server speichert nun zunächst einmal jeden Mausklick und jede Tastatureingabe, und dazu immer die Internet-Adresse, von der aus diese Eingaben erfolgen. Auch der Server kennt nicht die Personen hinter diesen Adressen. Er kennt auch die Telefonnummer nicht, von der aus die Einwahl ins Internet erfolgte; er hat es also deutlich schwerer, die Identität der Person zu ermitteln, und in vielen Fällen würde es ihm nur gemeinsam mit dem Provider gelingen, der auch die Telefonnummern kennt. Erst die Datenzusammenführung macht aus den Mosaiksteinchen ein Bild.

Der Serverbetreiber ist nicht zur Speicherung jeglicher Datenspuren verpflichtet. Es ist sogar fraglich, ob er nach deutschem Recht die Internet-Adressen seiner Nutzer überhaupt speichern darf! Hier herrscht derzeit drastische Rechtsunsicherheit in Deutschland: Es gibt Gerichtsurteile, die sich in dieser Frage diametral widersprechen. Trotzdem sieht der Gesetzgeber hier bisher keinen Handlungsbedarf. Man kann nur vermuten, dass die Komplexität der Technik die Politik hiermit überfordert.

Der Serverbetreiber ist nach dem Gesetz zur Vorratsdatenspeicherung jedoch dann zu einer Speicherung verpflichtet, wenn er den Dienst „E-Mail" auf seinem Server anbietet; dann muss er Folgendes für sechs Monate speichern:

- beim Versand einer E-Mail die Absender-IP-Adresse, die E-Mail-Adressen aller Beteiligten und den Zeitpunkt des Versands,
- beim Empfang einer E-Mail auf dem Mailserver wiederum alle involvierten E-Mail-Adressen, die IP-Adresse des Absender-Mailservers und den Zeitpunkt des Empfangs,
- beim Zugriff auf das Postfach den Benutzernamen und die IP-Adresse des Abrufers.

Inwieweit das Gesetz zur Vorratsdatenspeicherung mit dem im Grundgesetz verankerten Post- und Fernmeldegeheimnis vereinbar ist, mag sich mancher

fragen. Daher liegen auch zahlreiche Verfassungsbeschwerden dazu zur Entscheidung beim Bundesverfassungsgericht.

## Datenspuren bei der Informationssuche

Für die Informationssuche bei einer Suchmaschine gilt zunächst genau das Gleiche für Server: Die Software „sieht" die IP-Adresse desjenigen, der seine Abfragen eintippt, und fast alle Suchmaschinen speichern diese Adresse auch (unterschiedlich lange). Der Suchmaschinenbetreiber weiß also, von welcher Adresse aus was gesucht wird. Je nach Software liest er auch mit, auf welche Ergebnisse der Nutzer klickt (wie z. B. Google).

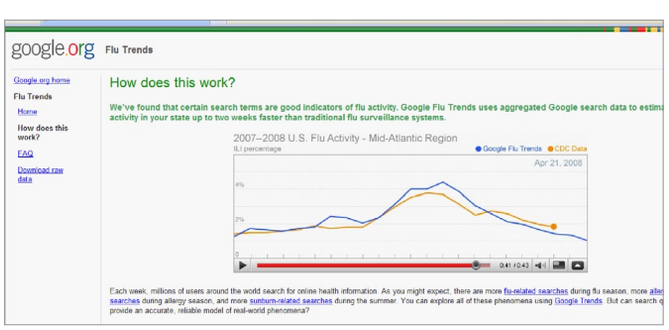
**„Herdenverhalten"**

Hinzu kommt das „Herdenverhalten" der Nutzer: In Deutschland und den meisten westeuropäischen Ländern nutzen mehr als 90 % für die Informationssuche die Server eines einzigen Betreibers: Google. Dies ist aus zwei Gründen wenig sinnvoll:

- Zum einen bekommt der Nutzer damit gesuchte Webseiten immer im Bewertungsmaßstab dieser Suchmaschine präsentiert. Verschiedene Untersuchungen haben gezeigt, dass sich Nutzer im Allgemeinen nur die ersten 10 bis maximal 30 Ergebnisse einer Internet-Suche ansehen. Welches diese ersten 10 bis 30 Ergebnisse nun sind, richtet sich nach den Bewertungskriterien der Suchmaschine. Diese Kriterien sind außerordentlich unterschiedlich. Benutzt man nur eine Suchmaschine, dann sieht man die digitale Welt stets durch deren „spezielle Brille" und das dahinter stehende Weltbild. Allein aus diesem Grund lautet die oberste Grundregel: „Benutzen Sie mehrere Suchmaschinen!"
- Zum zweiten wird durch die Konzentration auf einen Suchmaschinenbetreiber die Anhäufung von Datenspuren an einer zentralen Stelle drastisch erhöht. Dies wird noch dadurch verstärkt, dass die Suche nur noch eines von vielen Angeboten unter Google ist: E-Mail, Textverarbeitung, Tabellenkalkulation, Webseitenstatistik u. v. a. m. Nutzt man neben der Suche auch diese Dienste, dann ist meist eine Anmeldung dazu erforderlich. Damit ist dann auch der Bezug zwischen Person und IP-Adresse, sowie weiteren Datenspuren-Kennungen (User_Agent, Betriebssystem, Versionen usw.) hergestellt. Da darüber hinaus auf dem Rechner des Nutzers auch noch eine Kennung gesetzt wird (das „Cookie"), kann sehr detailliert nachvollzogen werden, was der einzelne Nutzer im Detail tut. All diese Daten sind über die Server einer einzigen Firma miteinander verbunden – damit ist die Zusammenführung der sonst auf den Rechnern vieler verschiedener Betreiber verteilten Daten relativ einfach. Was bei Google nun konkret von jedem Nutzer gespeichert wird, und was wie ausgewertet wird, ist Firmengeheimnis. Während die in Deutschland zz. gesetzlich verordnete Vorratsdatenspeicherung mit ihren sehr weitgehenden Eingriffen in die Privatsphäre genau festlegt, welche Daten wie lange gespeichert werden müssen, ist es hier in das Belieben des Betreibers gestellt. Kritiker bezeichnen die Vorratsdatenspeicherung in Deutschland denn auch als harmlos, verglichen mit den Möglichkeiten eines globalen Quasimonopols.

**Daten-Begehrlichkeiten**

Auf derartige „Datengebirge" von globalen Ausmaßen gibt es natürlich Begehrlichkeiten von verschiedensten Seiten. Zum einen stellen diese Datensammlungen ein enormes wirtschaftliches Kapital dar. Hunderte von Millionen Menschen offenbaren freiwillig – wenn auch oft aus Unkenntnis – dem Betreiber ihre intimsten Wünsche. Das ökonomische Potenzial ist offensichtlich. Google Flu-Trends (Flu = engl. Grippe) zeigt eine öffentlich zugängliche Anwendung (s. Abb. 1), die aus der Auswertung




Quelle: *http://www.google.org/about/flutrends/how.html*

*Abb. 1: Vorhersage von Grippeepedemien aus dem Google-Nutzerverhalten*

des Nutzerverhaltens zur Vorhersage von Grippeepedemien entsteht: *http://www.google.org/about/flutrends/how.html*. Ebenso wie Grippewellen sind natürlich auch Kaufverhaltens-Wellen vorhersagbar.
Aber auch von politischer Seite ist dieses Potenzial erkannt: Durch die Art der eingegebenen Frage und die angeklickten Ergebnisse kann man ein sehr genaues Bild über den einzelnen Menschen, seine Weltanschauung und seine politische Einstellung gewinnen. Bekanntestes resultierendes Beispiel ist die Verurteilung von Dissidenten in China aufgrund von Datenspuren in Suchmaschinen. Auch in den USA gibt der nach dem 11.9.2001 erlassene „Patriot Act" dem Staat weitreichende Möglichkeiten.

## Was kann der Nutzer gegen „Datensammelwut" und Monopolstrukturen tun?

Gegen die vorgeschriebene Protokollierung der Verbindungsdaten beim Provider gibt es bei der Einwahl ins Internet vom heimischen PC aus kein wirksames technisches Mittel. Wer juristische oder politische Mittel in Erwägung zieht, kann beim Arbeitskreis Vorratsdatenspeicherung (*www.vorratsdatenspeicherung.de*) mitarbeiten.
Als einfaches Mittel gegen die Datensammelwut kann der Nutzer den Herdentrieb meiden und solche Suchmaschinen aufsuchen, die seine Daten nicht oder nur kurzzeitig speichern. Gar nicht speichern: *www.metager.de* und *www.metager2.de*, kurzzeitig speichert *www.ixquick.com*.
Um generell nach dem Provider keine personenbeziehbaren Daten auf Servern zu hinterlassen, gibt es Anonymisierungsdienste. Bei diesen schaltet der Nutzer in die Internetverbindung einen weiteren Server dazwischen, von dessen Adresse aus dann alle weiteren Verbindungen zu den eigentlichen Ziel-Servern hergestellt werden. Nachteil: Die Verbindung kann dadurch langsamer werden, und die Anonymisierungsdienste sind im Allgemeinen kostenpflichtig. Oft bieten sie eine Preisstruktur um 10 Euro pro Monat für den einfachen Zugang und sind dann teurer für den schnellen, ungebremsten Zugang (mit Rabatten bei längerer Buchung); bekannte Anbieter sind z. B. *www.findnot.com* und *www.anonymizer.com*. Kostenfrei, aber mit funktionalen Einschränkungen, ist der anonymisierende Zugang über *www.obome.de*.

**Lokale Datenspuren: Browser und Konfiguration**
Auch bei der Wahl des Browsers und vor allem bei dessen Konfiguration kann der Nutzer bereits einiges zum Schutz seiner Privatsphäre tun. Dies betrifft zunächst einmal das, was an Datenspuren auf dem eigenen PC gespeichert wird. Denn es ist eine Illusion zu glauben, die Daten seien auf dem eigenen PC sicher gegen fremden Zugriff: Sobald eine Internetverbindung besteht, kann sie grundsätzlich auch in umgekehrter Richtung genutzt werden.
Ein guter Schutz dagegen ist zunächst einmal das Abblocken eingehender Verbindungen in der Router-Konfiguration. Dies erfordert jedoch einiges an technischem Knowhow; sinnvollerweise sind gängige Router für den Home-Gebrauch jedoch bereits meist dahingehend sicher vorkonfiguriert. Nichtsdestotrotz sollte man selbst testen, ob von außen ein Zugriff möglich ist. Zumindest ein „Portscan" sollte getestet werden: Eine einfache Anleitung findet man z. B. bei *www.heise.de/security/dienste/portscan/*.
Gegen Virusinfektionen des PC durch E-Mails, die unbemerkt Spionagesoftware installieren, hilft all das jedoch nicht. Mit Spionagesoftware können sowohl alle auf dem Rechner gespeicherten Daten, als auch jeder Tastendruck und jeder Mausklick aus der Ferne mitgelesen werden. Hiergegen hilft nur Antivirus-Software. Die heimliche Installation von Spionagesoftware ist mit dem Inkrafttreten der heftig umstrittenen Novelle des Gesetzes für das Bundeskriminalamt zum 1.1.2009 für die Geheimdienste in Deutschland, wie auch in anderen Ländern, zulässig. Inwieweit Antivirus-Software auch hiergegen wirkt, ist eine der vielen offenen Fragen bei der Umsetzung dieses Gesetzes.

**Firefox**
Selbst wenn nun ein Zugriff von außen definitiv unmöglich ist, dann hat natürlich ein anderer lokaler Nutzer – und sei es ein Einbrecher in der heimischen Wohnung – am eigenen PC Zugriffsmöglichkeiten, die wahrscheinlich nicht gewollt sind. Von den derzeit allgemein verfügbaren Browsern bietet auch dagegen der Firefox die besten Möglichkeiten einer sicheren Konfiguration. Aktuell und empfehlenswert ist die Version 3.0 (die Versionsnummer wird angezeigt unter „Hilfe", Menüpunkt „Über Mozilla Firefox").
In den Windows-Installationen findet man die wichtigsten Konfigurationseinstellungen unter „Extras". Dort kann man im Menüpunkt „Private Daten löschen" einstellen, wann diese jeweils gelöscht werden. Zu den privaten Daten gehören:

- die Adressen der besuchten Webseiten,
- die automatisch gespeicherten Seiten (Cache),
- die Cookies und
- Formulardaten.

Unter den Menüpunkten „Sicherheit“ und „Datenschutz“ kann (und sollte) man weitere Voreinstellungen konfigurieren.
Unter dem Menüpunkt „Sicherheit“ gibt es allerdings zwei Voreinstellungen, welche der Sicherheit dienen sollen, jedoch zu hinterfragen sind: „Hinweis anzeigen, falls die besuchte Webseite als attackierende …“ und „… als Betrugsversuch eingeschätzt wird“. Für diese Einschätzungen kommuniziert der Firefox-Browser mit einer Datenbank von Google; daher kann es empfehlenswerter sein, diese Warnungen abzuschalten (Häkchen entfernen). Weiterhin sollte man wissen, dass die Entwicklung des „freien“ Browsers Firefox im Wesentlichen von Google finanziert wird.
Mittlerweile bietet Google auch einen eigenen Browser an – Chrome –, von dessen Verwendung aus datenschutztechnischer Sicht abgeraten wird. Die Standardbrowser der Microsoft-Windows-Systeme – Internet Explorer (aktuell Version 7) – machen immer wieder durch Sicherheitslücken von sich reden, so dass auch von deren Benutzung eher abgeraten wird.

**Monopolstrukturen**

Die digitalen Monopolstrukturen haben also in doppelter Hinsicht negative Rückwirkungen auf den Nutzer:

- einseitige Bewertungsmaßstäbe für die Qualitität von Such-Ergebnisseiten und
- Konzentration der Datenspuren an einer Stelle.

Der erste Schritt aus der „Internet-Kinderstube“ heraus in die reale Welt der Netze besteht darin, nicht nur eine Suchmaschine zu benutzen, sondern mehrere. Der zweite Schritt ist die Installation und Konfiguration eines möglichst sicheren Browsers (zz. Firefox 3).
Ein dritter Schritt für Menschen, die sich diesbezüglich weiter engagieren wollen, kann darin bestehen, den digitalen Pluralismus direkt zu unterstützen: Der „Gemeinnützigen Verein zur Förderung der Suchmaschinen-Technologie und des freien Wissenszugangs“ (SuMa-eV) hat sich dieses zum Ziel gesetzt. Mehr dazu findet man unter *www.suma-ev.de*.

*Dr. Wolfgang Sander-Beuermann, Leiter des Suchmaschinenlabors am Regionalen Rechenzentrum für Niedersachen (RRZN), Leibniz Universität Hannover, sowie Geschäftsführer und Initiator des SuMa-eV.*

*wsb@suma-ev.de*

## Links

- http://www.vorratsdatenspeicherung.de
  Arbeitskreis Vorratsdatenspeicherung
- http://www.metager.de
- http://www.metager2.de
  Suchmaschinen ohne Speicherung der Nutzerdaten
- http://www.ixquick.com
  Suchmaschine mit nur kurzzeitiger Speicherung der Nutzerdaten
- http://www.findnot.com
- http://www.anonymizer.com
  kostenpflichtige Anonymisierungsdienste
- http://www.obome.de
  kostenfreier Anonymisierungsdienst
- http://www.heise.de/security/dienste/portscan/
  Anleitung für einen „Portscan“
- http://www.suma-ev.de
  Website des SuMa-eV

## Klicksafe

**klicksafe (Hrsg.): Knowhow für junge User – Mehr Sicherheit im Umgang mit dem World Wide Web.**
Das klicksafe-Handbuch ist eine praxisnahe Einführung in die weiten Felder der Online- und Netzkommunikationen. Aufbauend auf dem Konzept und den Erfahrungen der klicksafe-Lehrerfortbildungen, bietet es für Lehrer und Multiplikatoren sinnvolle Hilfestellungen und praxisbezogene Tipps für den Unterricht. Zu jedem der Jugendmedienschutz-Themen gibt es drei Unterrichtseinheiten (d. h. Arbeitsblätter zum Kopieren) in verschiedenen Schwierigkeitsgraden – also ca. 90 Arbeitsblätter. Jedes Thema ist gleich aufgebaut mit Sachinformationen, Links, methodisch-didaktischen Hinweisen und den Arbeitsblätter für den Unterricht.

**Bezug**
Bestellung per E-Mail unter *info@klicksafe.de*, die einzelnen Bausteine des Lehrhandbuchs können als PDF-Dateien heruntergeladen werden.