

Das Web hat sich inzwischen in einen Datenmoloch aus 1 Milliarde Dokumenten verwandelt, in dem niemand mehr eine Orientierung hat



Verloren im Web

SUCHMASCHINEN führen im Internet immer seltener zum Ziel – virtuelle Trampelpfade und Software mit gesundem Menschenverstand sollen das ändern

VON NIELS BOEING

Als Ozean aus Informationen wird das World Wide Web immer wieder beschrieben. Welch treffender Vergleich: So wie ein Schiffbrüchiger auf dem Pazifik verdursten kann, weil er nur von Salzwasser umgeben ist, verzweifelt der Websurfer auf der Suche nach der richtigen Information, weil er mit Unmengen von irrelevanten Daten überschüttet wird. Oft genug von Suchmaschinen wie Altavista, Lycos oder Excite, die genau das eigentlich verhindern wollen.

Schlimmer noch: Die herkömmlichen Suchmaschinen kommen mit dem Wachstum des Webs nicht mehr mit, das – ähnlich dem Universum nach dem Urknall – mit atemberaubendem Tempo expandiert. Wie zwei Physiker des japanischen Hightech-Konzerns NEC herausfanden, waren der besten Suchmaschine im März 1999 nur noch ein Sechstel der zu dieser Zeit 800 Millionen Web-Dokumente überhaupt bekannt. Anderthalb Jahre zuvor waren es noch ein Drittel.

Für Hawoong Jeong, Physiker an der University of Notre Dame im US-Bundesstaat Indiana, liegt das an der bislang vorherrschenden Technik der Suchmaschinen: Sie filzen alle ihnen bekannten Web-Dokumente nach der Kette aus Buchstaben und Ziffern, die der User als Suchbegriff eingegeben hat. Das Ergebnis wird nur danach bewertet, ob und wie oft der Treffer im Haupttext, im Titel oder sogar in den verborgenen Zusatzdaten eines Web-Dokuments gelandet wurde. „Diese Methode lässt die hochgradig vernetzte Struktur des Webs völlig ungenutzt“, sagt Jeong. „Sie kann nur Erfolg haben, wenn so viel wie möglich vom Web erfasst wird.“

Doch selbst dann ist diese Suchtechnik ziemlich ineffizient. Jeong und Kollegen fanden bei einer Analyse der Verlinkung von Websites heraus, dass die Welt im World Wide Web wahrlich klein ist: Es handelt sich um ein „Small World Network“, wie Mathematiker sagen, bei dem man mit durchschnittlich nur 18 Klicks von jeder Website zu einer x-beliebigen anderen kommt.

Wenn das Web aber wirklich so „klein“ ist, müssten sich auch Trampelpfade durch den Cyberspace aufspüren lassen, entlang derer die Nutzer schneller zu

wichtigen Informationen gelangen. „Wie man die Struktur eines solchen Netzwerks beim Suchen ausnutzen kann, zeigt Alexa“, sagt Jeong. Das ist ein Orientierungssystem der Firma Alexa Internet (www.alexa.com) aus San Francisco, das etwa 700 000 registrierte Nutzer hat.

Deren Bewegung durchs Netz wird von den Alexa-Servern festgehalten und ständig ausgewertet. Als Ergebnis bekommen die User zu jeder Homepage, die sie aufsuchen, eine Liste mit thematisch ähnlichen Websites auf den Bildschirm geschickt – eine Art virtuellen Trampelpfad. „Wenn genügend Leute denselben Weg genommen haben, wird der wohl brauchbar sein“, erklärt Brewster Kahle, Präsident von Alexa Internet, die Idee. „Und warum sollten Sie nicht von der Suche anderer Nutzer profitieren?“

Voraussetzung ist ein kleines, 1,7 Megabyte umfassendes Programm, das von der Alexa-Website heruntergeladen und auf dem PC installiert wird. Die Software ergänzt dann den Browser um einen kleinen Balken, mit dem zusätzliche Informationen zur jeweiligen Website aufgerufen werden können: allen voran der „Trampelpfad“ in Form der Linkliste, aber auch die Kontaktadresse des Seiten-Betreibers, der Umfang der Website und ihre Bewertung durch die User.

Besonderer Clou ist der Zugriff auf das Alexa-Internet-Archiv, das auch nicht mehr existente Seiten speichert hat. Klickt der Alexa-Nutzer auf einen toten Link, kann die entsprechende Seite aus dem Archiv rekonstruiert werden. 18 Terabyte (Billionen Byte) Information sind hier inzwischen festgehalten, so viel wie in der renommierten Bibliothek des US-Kongresses auf Buchseiten gebannt sind. „Wir wollen dem Internet auch ein Gedächtnis einbauen“, so Kahle. Alexa ist nicht der einzige Ansatz, die

hirnlose Statistik der heutigen Suchmaschinen zu überwinden. Auch Douglas Lenat hat sich dies zum Ziel gesetzt – doch der Informatiker und ehemalige Stanford-Professor geht noch weiter: Er will den Computern beibringen, Begriffe wirklich zu verstehen. „Menschen können aus Informationen eine Bedeutung erschließen. Computer sind dagegen nur Idioten, die zwar viel wissen, aber keine Ahnung haben, was all die Bits bedeuten“, formuliert Lenat das Grundproblem.

Was den Rechnern fehlt, ist der gesunde Menschenverstand, der einen Kontext erfasst, ohne dass er ausgesprochen werden muss: Dass der Begriff „Baum“ eine ziemlich große Pflanze meint, die in Erde verwurzelt ist, auf deren grüne Blätter Sonnenlicht fällt. Dass andererseits mit „Baum“ auch Ex-Bundesinnenminister Gerhart Baum gemeint sein kann, der einer politischen Partei, der FDP, angehört und seinen Amtssitz in der Stadt Bonn hatte.

Um das Kontextproblem zu lösen, hat Lenats Firma Cycorp (www.cyc.com) in akribischer Forschungsarbeit seit 1984 den „Cyc-Knowledge-Server“ entwickelt: Der besteht aus einem 100 Megabyte großen Programm namens Cyc und einer Datenbank aus Zigtausenden solcher Allerweltskontexte.

Als erster Suchdienst wird Hotbot (www.hotbot.com) die Cycorp-Technologie verwenden: In den nächsten Wochen soll ein Cyc-Knowledge-Server mit der Suchmaschine gekoppelt werden. „Der Suchbegriff, den der User eingibt, geht zuerst in den Cyc-Server. Der entscheidet dann nach Rückfrage beim User, welche Anfrage an die Suchmaschine weitergeleitet wird“, beschreibt Lenat die Zusammenarbeit der beiden Systeme. Beispiel: Der englische Suchbegriff *vets*. Das kann für *veterinarian* (Tierarzt), aber auch für *veteran* (Kriegsveteran) stehen.

Cyc fragt den User, in welchem der beiden Kontexte er sucht. Im Falle des Tierarztes fügt Cyc dann weitere Suchbegriffe wie Tiernamen oder medizinisches Gerät hinzu, die mit dem Begriff Tierarzt verbunden sind. Der Nutzer muss sich nicht mehr den Kopf zerbrechen, welche Begriffe er logisch wie verknüpft, um ein brauchbares Ergebnis zu bekommen. „Damit bringen wir das Wissen, das jeder User hat, endlich in den Computer“, sagt Lenat.

BITLAND

DER ONLINE-SOUVERÄN

PETER GLASER über die unheimliche Vielfalt im Internet, die den Nutzer zum Regisseur des eigenen Lebens macht

Der kreative Anspruch der neuen Medien ist erdrückend. Mehr und mehr Information in Form fragmentierter und segmentierter Teile, Artikel, Texte, Bilder, *soundfiles*, die man sich selbst nach Belieben zusammensetzen kann – und oft muss. Alles, was bislang Kino, TV und Zeitung für einen erledigt haben: Auswahl, Zusammenstellung, Gestaltung, muss der Online-Mensch nun leider selbst machen. Die neue Freiheit verwirrt.

Wo bisher ein vollständiges Bild, ein Konsens geliefert wurde – ein Film mit nach den Regeln der Filmgrammatik geschnittenen Bildern und festgelegter Handlung, eine Zeitung mit redaktioneller Auswahl und Layout –, muss nun selbst aus den Teilen, die im Netz zu finden sind, etwas zusammengesetzt werden. „Finden“ ist bedeutsam: Das Allerwenigste wird online gesucht, das meiste bloß gefunden.

Der Online-Mensch wird ein positives Lebensgefühl nur dann entwickeln, wenn er eine Souveränität gegenüber den Informationsströmen empfindet. Wenn sich dieses Gefühl nicht durchsetzt, wird sich das Interesse an den

neuen Medien und ihren kommunikativen Angeboten schließen wie eine Muschel. Wesentlich ist, die Souveränität über seine Zeit zu behalten – das Internet als Sieg des individuellen Rhythmus über den institutionellen.

Das Netz ist in der Lage, sich den unterschiedlichen Formen persönlicher Zeitökonomie anzupassen. Neue Arten von „Zeitgenossen“ bilden sich drumherum aus. Es gibt „Zwischenzeitler“, die ihren Tag lieber in viele kleine, pausenregelhaftige Genusspotenziale aufteilen als sich an klassischen Großrhythmen à la Frühstück, Mittagessen, Abendbrot zu orientieren.

Es gibt den „Mainac“, den Hauptsächler. Er äußert sich ausschließlich in zentralen Sätzen und schickt Telegramme aus dem Herzen der Wirklichkeit. Ansonsten schweigt er. Oder es gibt die „Betweenager“, deren Leben vor allem aus den Stellen besteht, die aus den großen Hollywoodfilmen rausgeschnitten bzw. die gar nicht erst gedreht wurden.

Vom Online-Universum fühlen diese Menschen sich, als von einem gleichartigen Gegenüber, alle verstanden.



NEWS

INTERNET

Auf dem Web reiten

Der Berliner Netzanbieter DATANGO zeigt moderierte Internet-Shows

Internet so einfach wie Fernsehen“ versprechen sechs Berliner Jungunternehmer dem Laien: Ab 21. Oktober kann der sich unter www.datango.de moderierte Webshows ansehen. Auf diesen „Webrides“, wie Datango das nennt, stellen „Webjockeys“ ausgewählte Sites zu speziellen Themen vor, während im Hintergrund Musik läuft. Um die Webrides abspielen zu können, müssen User zunächst ein 1,1 Megabyte großes Zusatzprogramm, den *datango-player* für den Browser runterladen und installieren. Die Shows lassen sich dann per Button-Druck anhalten oder vor- und zurückspulen, werden aber immer als Live-Datenstrom auf den eigenen Rechner geschickt. Gepusht wird Datango unter anderem von dem TV-Sender Pro Sieben und dem Internet-Auktionshaus E-Bay. Wer keine Lust auf quasselnde Webjockeys und Software-Installationen hat, kann sich unter www.netzpiloten.de coole Internet-Seiten als Dia-Show vorführen lassen – die Piloten-Software läuft automatisch online.



HARDWARE

Tragbarer Speicher an der Leine

Der neue poppige USB ZIP DRIVE von Iomega fasst 250 Megabyte



Design-verwöhnte iMac-Fans dürfen sich freuen: Mit dem neuen Zip Drive von Iomega für USB-Schnittstellen gibt es endlich ein optisch würdiges Laufwerk mit viel Speicherplatz. 250 Megabyte passen auf die Zip-Disks. Das sind gut vier Stunden Musik im MP3-Format oder, je nach Auflösung, zwischen 300 und 1200 digitale Bilder. Über einen zusätzlichen Adapter kann das Laufwerk auch mit der PCMCIA-Schnittstelle von Laptops oder PCs verbunden werden. Auch ältere 100-Megabyte-Zip-Disks lassen sich mit dem neuen Zip Drive nutzen. Erforderliche Betriebssysteme: Windows 95, 98 oder NT sowie Mac OS 8.1, Preis: 429 Mark (219 €) für das Laufwerk, 99 Mark (50,62 €) für die PC/Laptop-Adapterkarte, 43 Mark (22 €) für die Diskette.

WAS IST EIGENTLICH ...

... eine SOUNDKARTE?

Vorbei sind die Tage, an denen PCs blechern quäkten. Heute bringen Soundkarten Klangvielfalt beim Spielen und Komponieren. Mit eigenen Chips und zusätzlichem Speicher erzeugen sie Geräusche durch Mischung einfacher Töne („FM-Synthese“) oder greifen auf abgespeicherte Klänge zurück (die so genannte „Wavetable“-Technik). Eine gute Violine belegt etwa mehrere Megabyte Speicher. Über Midi-Schnittstellen lassen sich außerdem Musikinstrumente anschließen.

