

Yahoo

Das Ende von Alta Vista - das Ende von All the Web

Wie erwartet hält die Firma Yahoo nach dem Start ihrer eigenen Suchmaschine die Datenbanken von All the Web und Alta Vista nicht mehr aufrecht. Beide wurden Ende März abgeschaltet. Die bekannten Suchseiten bestehen jedoch weiter und basieren nun auf dem Yahoo-Index. Mit der Abschaltung der beiden Suchmaschinen werden die Wahlmöglichkeiten der Suchenden wieder deutlich kleiner. Zwar gibt es weiterhin viele Suchangebote. Vor allem die großen Portalangebote basieren jedoch auf den Datenbanken von Yahoo oder Google. AltaVista war der Pionier der Internet-Suche. AltaVista startete 1995 als erste Volltextsuchmaschine mit dem damals größten Index und mit Suchmöglichkeiten, die sich am professionellen Bedarf orientierten und teilweise bis heute von keiner anderen Suchmaschine angeboten werden. Lange Zeit galt AltaVista als der Stern am Suchmaschinen-Himmel. Im Zuge mehrerer Verkäufe und Strategiewechsel begann dieser Stern allerdings zu sinken; während AltaVista im Jahr 2000 gerade auf Portalfunktionen setzte (ein Fehler, den zu dieser Zeit einige Suchmaschinen begingen), kam Konkurrenz durch neue Suchmaschinen, allen voran Google. In den letzten Jahren konnte sich AltaVista nie vom Verlust großer Anteile seiner Nutzer erholen, blieb jedoch wegen seines eigenen Indexes und der vorbildlichen Suchmöglichkeiten eine wichtige Suchmaschine. Mit dem Ende von AltaVista fehlt nicht nur der zugrunde liegende Index. Vielmehr fehlen vor allem solche Suchmöglichkeiten wie Trunkierung, Abstandsoperatoren sowie die Beachtung der Groß- und Kleinschreibung bei Suchanfragen. All the Web konnte nie einen Bekanntheitsgrad wie die anderen "großen" Suchmaschinen erreichen, obwohl diese Suchmaschine sicher das Zeug dazu gehabt hätte. Die zugrunde liegende eigene Datenbank war seit dem Start der Suchmaschine 1999 immer eine der international größten Suchmaschinen, zeitweise

war sie sogar führend. All the Web war die erste Suchmaschine, die über zwei Milliarden Dokumente indexiert hatte; zuletzt lag der Index bei etwa 3,2 Milliarden Dokumenten. Ein Pionier war All the Web beim Aufbau von Bilder-, Video- und FTP-Indizes. Die ursprünglich von der norwegischen Firma FAST Search & Transfer betriebene Suchmaschine belieferte Portale wie Lycos und T-Online mit ihren Suchergebnissen. Im Jahr 2003 wurde die Web Search Division von FAST von Overture (und schließlich von Yahoo) übernommen. FAST als eigenständige Firma konzentriert sich seitdem auf den Geschäftskundenbereich. Mit All the Web verliert der Web-Sucher einerseits eine der größten Suchmaschinen mit einigen

Google (1)

Streit um Gmail

Google hat einen eigenen Freemail-Dienst angekündigt. Durch die Einblendung von kontextbezogener Werbung neben den Nachrichtentexten sorgt dieser jedoch schon im Vorfeld für Kontroversen.

Gmail (<http://gmail.google.com>) befindet sich im Moment in einer Testphase, an der nur ausgewählte Nutzer teilnehmen können. Wann der Dienst für die Allgemeinheit verfügbar sein wird, ist nicht bekannt. Gmail bietet dem Nutzer im Gegensatz zu den bekannten Freemail-Anbietern wie Yahoo und GMX mit 1.000 Megabyte Speicherplatz nahezu unbeschränkte Archivierungsmöglichkeiten privater Nachrichten. Dazu kommt, dass der Dienst ausdrücklich suchbasiert ist, d.h. dass das lästige Ablegen der Nachrichten entfallen soll. Statt dessen können Nachrichten einfach gesucht werden, die Ergebnisse werden im Konversations-Kontext dargestellt. Dazu soll ein effizienter Spam-Filter kommen, der unerwünschte Nachrichten aussortiert.

Für den Nutzer ist der Dienst kostenlos, allerdings erscheinen bei der Anzeige von E-Mails kontextbezogene Anzeigen, wie sie schon von der Google-Webseite her bekannt sind. Da dafür der Text der E-Mails ausgewertet werden muss,

interessanten Spezial-Indizes, andererseits einzigartige Suchfunktionen wie die Größeneinschränkung und die Suche nach eingebetteten Inhalten.

Mit dem Weggang von All the Web und Alta Vista existieren nur noch zwei bedeutende Suchmaschinen-Indizes: Google und Yahoo. Als weitere relativ große Datenbank ist Teoma zu empfehlen. Hier ist allerdings die Aktualisierung des Index und die starke Ausrichtung auf englischsprachige Dokumente zu bemängeln. Für die angesprochenen erweiterten Suchfunktionen findet sich leider kein Ersatz - vielleicht darf man ja auf Yahoo hoffen. Dort hat man inzwischen wieder damit angefangen, so altmodische Dinge wie Metadaten zu berücksichtigen.

halten Datenschützer den Dienst für problematisch wenn nicht sogar das Fernmeldegeheimnis verletzend und damit für unzulässig. Die Firma Google diskutiert mittlerweile Änderungen an dem Dienst.

Die hinter dem E-Mail-Dienst stehende Frage lautet, warum das Unternehmen Google, das sich bisher einzig und allein als Suchmaschinen-Betreiber ansah, einen solchen Dienst entwickelt. Schon wurde der Umbau zu einem Portal am Horizont gesehen. In der Diskussion untergegangen sind bisher allerdings die Kombinationsmöglichkeiten zwischen Web-Suche und Suche in persönlichen Dokumenten. So zeigen die von Google veröffentlichten Screenshots nicht nur kontextbezogene Werbung, sondern auch kontextbezogene Suchergebnisse aus Googles Web-Index an.

Allgemein ist für die Zukunft mit einer Integration der Suche in unterschiedlichen Datenbeständen unter einer Oberfläche zu rechnen. So bietet das Projekt "Stuff I've seen" von Microsoft Research eine gleichzeitige Suche in (bereits einmal angesehenen) E-Mails, Dateien und Webseiten. Die HotBot Deskbar (<http://www.hotbot.com/tools/desk-top/>) durchsucht gleichzeitig lokale Dateien und E-Mails und das Web.

Neue Dienste

Google bietet als neue Dienste Web Alerts, Personalized Web Search und einen neuen Operator an.

Mit den Google Web Alerts (<http://www.google.com/webalerts>) ist es möglich, bestimmte Suchanfragen dauerhaft zu verfolgen. Bisher wurde eine solche Monitoring-Funktion für Google nur von einem Fremdanbieter (<http://www.googlealert.com/>) angeboten. Dieser Dienst erfreute sich bereits einiger Beliebtheit, so dass Google nun nachzog. Damit ist es möglich, sich einmal täglich oder wöchentlich neu gefundene Suchergebnisse per E-Mail zuschicken zu lassen. Dabei werden allerdings nur die ersten zwanzig Suchergebnisse berücksichtigt, so dass bei Anfragen, die mehr als zwanzig Treffer ergeben, neue Ergebnisse nur gemeldet werden, wenn sie von Google in den Top 20 gelistet werden. Das gleiche Problem besteht auch bei [Googlealert.com](http://www.googlealert.com/), hier werden allerdings fünfzig Ergebnisse berücksichtigt. Fazit: Die Google Web Alerts bringen nichts

Neues und keine Verbesserung gegenüber dem bereits angebotenen Dienst des Fremdanbieters. Positiv dürfte allerdings die Wirkung auf die Konkurrenz sein, die in nächster Zeit mit eigenen Alert-Diensten nachziehen dürfte.

Bei der Personalized Web Search (<http://labs.google.com/personalized>) setzt Google auf Voreinstellungen des Nutzers, der seine Interessen in einem persönlichen Profil festlegt, welches in einem Cookie gespeichert wird. Die Interessen können aus einem Kategorienschema ausgewählt werden. Dieses ist allerdings sehr allgemein gehalten. Google ordnet gefundene Webseiten den Kategorien zu und bewertet die Übereinstimmungen in der Suchanfrage. In der Trefferlistenanzeige kann eingestellt werden, wie stark die Ergebnisse dem persönlichen Profil angepasst werden sollen. Probleme bestehen allerdings, wenn der Nutzer mehrere Interessen in seinem Profil ausgewählt hat. In diesem Fall kommt es oft zu Ergebnissen, die keine Verbesserung gegenü-

ber den regulären Web-Ergebnissen darstellen. Insbesondere gegenüber Netzwerk-Ansätzen wie [Eurekster.com](http://www.eurekster.com) ist die personalisierte Suche bei Google recht primitiv und stark verbesserungswürdig.

Weiterhin hat Google einen neuen Operator eingeführt, mit dem sich nach numerischen Informationen suchen lässt, und zwar insbesondere mittels "Von-bis-Einschränkung". Werden zwei Zahlen mit zwei Punkten verbunden (ohne Leerzeichen), so werden Dokumente gefunden, die eine Zahl zwischen den beiden Werten enthalten. Besonders interessant ist diese Suchmöglichkeit für Jahresangaben, beispielsweise bei einer Suche nach Geschichtlichem zu Suchmaschinen: das Suchargument könnte hier "suchmaschinen 1993..1996" lauten.

*Dirk Lewandowski,
dirk.lewandowski@uni-duesseldorf.de*