

Die ungebremste Kreativität der Web-Suche

Zeit und Ort: 25.04.2006, EEAR

Referent: Rigo Wenning, Justitiar World Wide Web Consortium (W3C)

Protokoll: Ass. jur. Iris Speiser

Rigo Wenning ging zunächst kurz auf den Vortrag von Prof. Weikum ein und brachte zum Ausdruck, dass für ihn die Frage des Datenschutzes ein Hauptpunkt bei den neuen Suchtechnologien sei.

Google habe in den letzten Jahren eine wahre Hysterie ausgelöst. Die Suchmaschine erfreue sich unter den Internetnutzern einer sehr großen Bekanntheit und sei Marktführer in diesem Marktsegment. Dieser Umstand löse gewisse Ängste aus. Google habe es in 10 Jahren geschafft, vom kleinen Start-up-Unternehmen, bestehend aus zwei Personen zu einem marktführenden Konzern aufzusteigen. Derzeit expandiere Google in alle Bereiche des Information-Managements. Die neuesten Produkte seien z.B. GoogleEarth, GoogleBase, Blogsearch und GoogleDesktop.

Es stelle sich daher die Frage nach dem Erfolgsrezept. Google habe Bedürfnisanalysen durchgeführt und ein Produkt entwickelt, das eine innovative Kombination existierender Technologien darstelle. Google habe aber dennoch das gleiche Problem, mit dem alle Datenbankanbieter zu kämpfen haben: Das Verhältnis von Precision und Recall. Google erfasse längst nicht alle Dokumente im Netz – Schätzungen gingen davon aus, dass lediglich ca. 20% aller Webseiten im Google-Index enthalten seien, obwohl der Index täglich wachse. Zudem sei der Anteil der irrelevanten Dokumente bei Google sehr groß – sogar wesentlich größer als der bei Fachdatenbanken.

Warum sollten Anbieter kommerzieller Datenbanken in Anbetracht dieser Unzulänglichkeiten Angst haben? Die Nutzung von Google ist für den Nutzer kostenlos, während die Nutzung kommerzieller Datenbanken mit Nutzungsentgelten verbunden ist. Weiterer Vorteil ist das sehr einfache Interface, das keine komplexen logischen Abfragen erfordere und dennoch regelmäßig einschlägige und wertvolle

Treffer liefere. Google sei ein Spiegel der Community und zeichne sich durch Exzellenz aus. Trotz der kostenlosen Nutzungsmöglichkeit sei Google aber nur auf den ersten Blick kostenlos. Google erwirtschaftete seine Einnahmen aus Werbung, der Erstellung von Profilen und durch den Verkauf von Service-Dienstleistungen an große Unternehmen. Diese Entgeltstruktur ermögliche keinen direkten Vergleich mit den kommerziellen Datenbankanbietern. Es bestehe aber durchaus ein Grund für Datenbankanbieter, Google als Konkurrenz zu betrachten.

Während viele kommerzielle Datenbanken komplexe Suchformulare anböten und vom Nutzer den Aufbau logischer Suchabfrage verlangen, lieferten bei Google auch einfache Suchergebnisse akzeptable Ergebnisse. Zwar erhalte man auch jede Menge nicht-einschlägiger Ergebnisse; die Sortierung nach Relevanz führe jedoch in der Regel dazu, dass die ersten Ergebnisseiten überwiegend einschlägige Treffer enthielten. Das Interface von Google setze auf moderne Web-Technologien wie AJAX. Zudem spiegele sich der Community-Gedanke, wonach man sich gegenseitig zitiert, darin wieder, dass die Wertschätzung als Kriterium für das Ranking verwendet werde.

Inzwischen integriere Google auch Semantic Web Informationen. Als Beispiel verwies Wenning auf die Creative Commons Suche, die eine Suche nach Dokumenten erlaube, die einer bestimmten Creative Commons Lizenz unterworfen seien.

Der Service von Google werde von allen benutzt und sei überall und jederzeit verfügbar – in Versionen für Mobilgeräte, entsprechend dem WAI-Standard. Google sei Pilot der Web-Entwicklung und offen für Partner – mit dem Erfolg, dass inzwischen in jeder Browser-Anwendung standardmäßig ein Google-Suchfeld integriert sei. Über die Google-API könnten Entwickler Google-Abfragen automatisiert übermitteln und die Ergebnisse in Ihren Entwicklungen verwenden; Google profitiere wiederum von derartigen Entwicklungen.

Im Gegensatz zu Datenbankanbietern sei Google jedoch nicht Anbieter; sondern erschließe lediglich unsere Dokumente. Durch Google werden Dokumente besser und durch bessere Dokumente werde Google besser.

Im Hinblick auf juristische Daten stellte Wenning fest, dass immer mehr juristische Daten frei online gestellt würden – sowohl von Gerichten und Behörden als auch von Anwälten und Unternehmen. Auch bilde sich zunehmend eine juristische Community aus, die durch den genannten Mechanismus zur Verbesserung von Google beitrage.

Für die Datenbankanbieter habe diese Entwicklung zur Folge, dass sich die Konkurrenz verstärke. Google werde zunehmend vollständiger und über die Community erfolge eine Qualitätskontrolle. So sei es möglich, Google-Suchabfragen auf bestimmte, vertrauenswürdige Websites wie z.B. die der Bundesgerichte zu beschränken. Das Problem der Anbieter seien die direkten Kosten – während Google sich indirekt finanziere.

Um der zunehmenden Konkurrenz durch Google zu begegnen, müssten die Datenbankanbieter geeignete Abwehrstrategien entwickeln. Zunächst müsse man sich auf die eigenen Stärken besinnen. Juristische Fachdatenbanken seien innerhalb ihres definierten Erschließungsgebiets vollständig und hätten eine hohe Qualität. Was teilweise fehle, sei die Web-Integration. Die Anbieter müssten innovative Geschäftsmodelle entwickeln, die auch für Gelegenheitsnutzer attraktiv seien. Juristische Fachdatenbanken sprächen gezielt die juristische Community an, würden den Kontakt zu dieser aber noch viel zu wenig pflegen. Während Google oder Microsoft inzwischen ihre neuen Produkte der Community zum Beta-Test zur Verfügung stellen und über neue Entwicklungen in Blogs informieren, werde dieses Potential bei den juristischen Fachdatenbanken derzeit nicht genutzt.

Im Hinblick auf die Qualität sollten die Anbieter insbesondere auf definierte Formate setzen, die einen einfachen Import der gefundenen Dokumente ermögliche.

Die Informationsaufbereitung könne z.B. dadurch verbessert werden, dass Zitierungen bewertet werden. So müsse eine Zitierung durch ein Bundesgericht als höher bewertet werden als eine durch ein Instanzgericht. Aus derartigen Bewertungen ließe sich ein Ranking bilden, das das Auffinden wichtiger Entscheidungen erleichtere. Ein weiterer Ansatz wäre, dem Nutzer die Klassifizierung als "wichtig" zu überlassen.

In technischer Hinsicht könnten die Anbieter die Interfaces verbessern, um eine intuitivere Bedienbarkeit zu ermöglichen und mobile Technologien einsetzen; es bestehe auch ein Bedarf nach mobil zugänglichen juristischen Informationen. Bei Web-Angeboten sei auf persistente URLs zu achten.

Schließlich fordert Wenning die Einführung kreativerer Abrechnungsmodelle. Früher wurde nach Zeiteinheiten oder Zeichenanzahl abgerechnet – auch wenn die Recherche im Ergebnis erfolglos war. Was zähle, sei die Information. Dem würden die bisherigen Abrechnungsmodelle nicht gerecht. Auch hier könnte die Einbeziehung der Community einen Innovationsschub liefern.

Zum Abschluss rief Wenning die Datenbankanbieter dazu auf, von Google zu lernen und ihre Mitarbeiter "spielen" zu lassen. Mitarbeiter sollte eine aktive Rolle in der Community übernehmen. Den Mitarbeitern sollten keine Grenzen für Innovationen gesetzt werden. Google räume seinen Mitarbeitern einen Tag pro Woche für eigene Entwicklungen ein. An den Ergebnissen solcher Testentwicklungen lasse man die Community frühzeitig teilhaben. So seien etliche Aufsehen erregenden Services von Google wie GoogleEarth, GoogleMail, etc. Ergebnisse des "Spielens" von Mitarbeitern.