

Verkaufte Suchwörter, verkaufte Links. Retrieval nach Homepages optimiert?

Selten wurde ein Dienst so stark diskutiert wie "RealNames", da sowohl seine Suchwörter als auch die darauf aufsetzenden Links auf Homepages verkauft werden. Machen wir uns doch nichts vor: Das World Wide Web ist ein kommerzielles Unternehmen, die Dienstleistungen müssen finanziert werden. Und warum soll nicht ein Privatunternehmen mit Inhaltserschließung Gewinne erzielen? RealNames mit seinem "Internet Keyword System" ist eine interessante Idee, das Retrieval nach Homepages zielgenau durchzuführen. Was bringt RealNames den Nutzern? Wann sollte jemand, der Sites im World Wide Web unterhält, Suchwörter und Links bei RealNames abonnieren (und was kostet das)? Im Oktober 1999 haben wir RealNames genauer angeschaut. Analysiert wurde das gezielte Retrieval nach Homepages bei AltaVista, Fireball und bei RealNames.com.

Suche nach Homepages mittels Suchmaschinen

Unterstellen wir, ein Nutzer sucht nach genau einer Homepage, etwa von einem Unternehmen, einem Produkt, einer Stadt oder einer Person. Die Adresse im World Wide Web (URL - Uniform Resource Locator) sei unbekannt. Nehmen wir an, jemand sucht nach Bill Clintons Homepage oder nach der Internet-Site der Camcorder von Sony. Gibt man in einer Suchmaschine, sagen wir AltaVista, als Suchargumente "Bill Clinton"

oder "Sony Camcorder" ein, so erhalten wir im ersten Fall mehrere Hunderttausend Treffer, wo Clinton vorkommt, und im zweiten Fall immer noch so viele Links, dass das Durcharbeiten Mühe bereitet. Die gesuchte Homepage ist in der Treffermenge bei AltaVista enthalten, aber versteckt wie die Nadel im Heuhaufen. Dieses Problem der Inhaltserschließung hat RealNames auf seine Art gelöst.

RealNames Internet Keyword System

Im April 1997 wurde von Keith Teare die Firma "Centraal" gegründet. Die Geschäftsidee war einfach: Da das Merken und Eingeben von (durchaus recht langen) URL Probleme bereitet, sollte der neue Dienst statt mit URL mit natürlich-sprachigen Suchwörtern arbeiten. Diverse Suchmaschinen, darunter AltaVista und Fireball, setzten das Centraal-Produkt mit dem Namen "RealNames" als ergänzenden Dienst ein. Seit August 1999 firmiert Centraal unter dem Namen seines Produktes als "RealNames Corporation"; Sitz ist nach wie vor San Carlos in Kalifornien. RealNames schreibt über sein Unternehmen, "RealNames Corporation goal is to make the Internet a more enjoyable place to visit and navigate for everyone", und recht selbstbewußt, "Internet Keywords represents the first major step forward in internet navigation since the introduction of the URL".

RealNames speist seine Datenbank aus unterschiedlichen Quellen: ● die verkauften Suchwörter mit den verkauften

Links (einschließlich gratis vergebener Suchwörter an Privatpersonen und öffentliche Bildungsinstitutionen); ● von RealNames intellektuell erschlossene Homepages; ● (automatische) Analyse des Domainnamens; ● Kooperation mit anderen Informationsproduzenten; derzeit: Internal Revenue Service (IRS); Steuervordrucke; US Securities and Exchange Commission (SEC); Lageberichte und Bilanzen börsennotierter Unternehmen in den USA und US Patent & Trademark Office; Patente.

Zielgenaue Suche nach Internet-Sites

Eine zielgenaue Suche geht - ohne RealNames - ausschließlich über die URL. Wenn man

http://whitehouse.gov/WH/EOP/OP/html/OP_Home.html eingibt, landet man im Weißen Haus bei der Homepage des Präsidenten; unter <http://www.sel.sony.com/SEL/consumer/camcorder>

liegen die Produktbeschreibungen von Sonys Camcorder. Nun kann man keineswegs voraussetzen, dass solche Buchstabenkolonnen allgemein bekannt oder etwa gut merkbar wären. RealNames verknüpft die URL mit den jeweils bestpassenden Suchargumenten, also etwa "Bill Clinton" mit der URL des amerikanischen Präsidenten; "Sony Camcorder" mit der entsprechenden SEL-Sony-Site. **Abbildung 1** zeigt den Ergebnisschirm der Suche nach "Bill Clinton" (als Phrase) bei der Simple Search von

AltaVista. (Derzeit läuft RealNames bei AltaVista nur bei der Simple Search, nicht bei der Advanced Search.) Markiert ist der RealNames-Link stets durch das hochgestellte "RN" hinter dem Link.

Das Suchergebnis bei RealNames wird stets (ohne Nummerierung) vor die Trefferliste sortiert. Findet RealNames nur genau eine Homepage (wie im Falle vom **Abbildung 1**), so wird der direkte Link angegeben. Bei "Bill Clinton" war dieses Verfahren ausgesprochen günstig, fanden wir doch seine Homepage unter den ersten 50 Treffern von AltaVista nicht. Das Relevance Ranking führte vielmehr nach Treffern folgenden Inhalts:

1. Anba Tonè Bill Clinton; The Haiti Global Village Discussion Forum ...
2. Vietnam Veterans' War Stories: Bill Clinton's ROTC Letter ...
3. Bill Clinton vs. Thomas Jefferson
4. End Bill Clinton's War
5. Bill Clinton -- Libertarian?
6. Bill Clinton, Aggressor
7. Bill Clinton's Underpants usw.

Letztlich sind dies Treffer, die einen Nutzer, der sich ausschließlich für Clintons Homepage interessiert, mitnichten zufriedenstellen werden. Ohne das Ver-

fahren von RealNames ist mit einem - zugebenermaßen primitivem - Suchargument wie "Bill Clinton" die erwünschte Homepage praktisch kaum zu finden.

RealNames arbeitet sowohl mit exakter Suche ("exact matching") als auch mittels "best-matching"-Verfahren. Solange das exact matching zielführend ist, wird ausschließlich damit gearbeitet. D.h., stimmen Suchargument und das entsprechende RealNames-Schlagwort zeichengenau überein, gibt es stets genau einen Treffer. Gibt es keine exakte Übereinstimmung, wird nach dem "best-matching" Verfahren weitergearbeitet und nach ähnlichen Zeichenfolgen gesucht. Die zusätzlichen "best-matching"-Suchen sind derzeit nur bei AltaVista, Fireball und bei RealNames.com realisiert; andere Suchmaschinen (darunter Go / Infoseek, MSN Search und LookSmart) bieten nur "exact-matching" an.

In den meisten Fällen arbeiten die RealNames-Links zuverlässig. Es gibt jedoch ein Problem. Ist das Suchargument eine Phrase, so sucht RealNames auch nach den Bestandteilen der Phrase (auch dann, wenn die Phrase eigens durch Anführungszeichen markiert ist). So führte unsere Suche nach "fbi fh köln" nicht zum Fachbereich Bibliotheks- und Informationswesen der FH Köln (der kein

Suchwort und somit keinen Link abonniert hat), sondern zum FBI.

Ein analoges Problem tritt auf, wenn mehrere (vergebene) Suchwörter die selben Zeichenfolgen beinhalten. Wir probierten "information technology" und erhielten eine recht große Liste von RealNames-Links (**Abbildung 2**). Sollte die Zahl der Abonnenten stark ansteigen, werden solche Überschneidungen häufiger. Irgendwann wird dann auch RealNames entweder (1.) eine Ausgabe nach Relevanz anbieten müssen oder (2.) nur nach dem "exact-match"-Retrieval arbeiten.

Unsere Variante (2.) - das "exact-match"-Verfahren - hat bei großen Treffermengen Vorteile, bei kleinen jedoch nicht. Nehmen wir an, es existieren die zwei zu vermietenden Suchwörter "Lausitzer Braunkohle" sowie "Bundesverband Braunkohle". Suchen wir exakt, bekommen wir jeweils genau den richtigen Link und nur diesen. Wissen wir das Suchwort jedoch nur ungefähr und versuchen die Recherche mit "Braunkohle", so erhalten wir derzeit eine Liste mit zwei Treffern, beim "exact-match"-Verfahren würden wir keinen Treffer erhalten. Hier wäre das heutige Verfahren für den Nutzer besser. Man darf gespannt darauf sein, wie RealNames seine Software angesichts der riesigen Mengen an Suchwörtern optimieren wird.

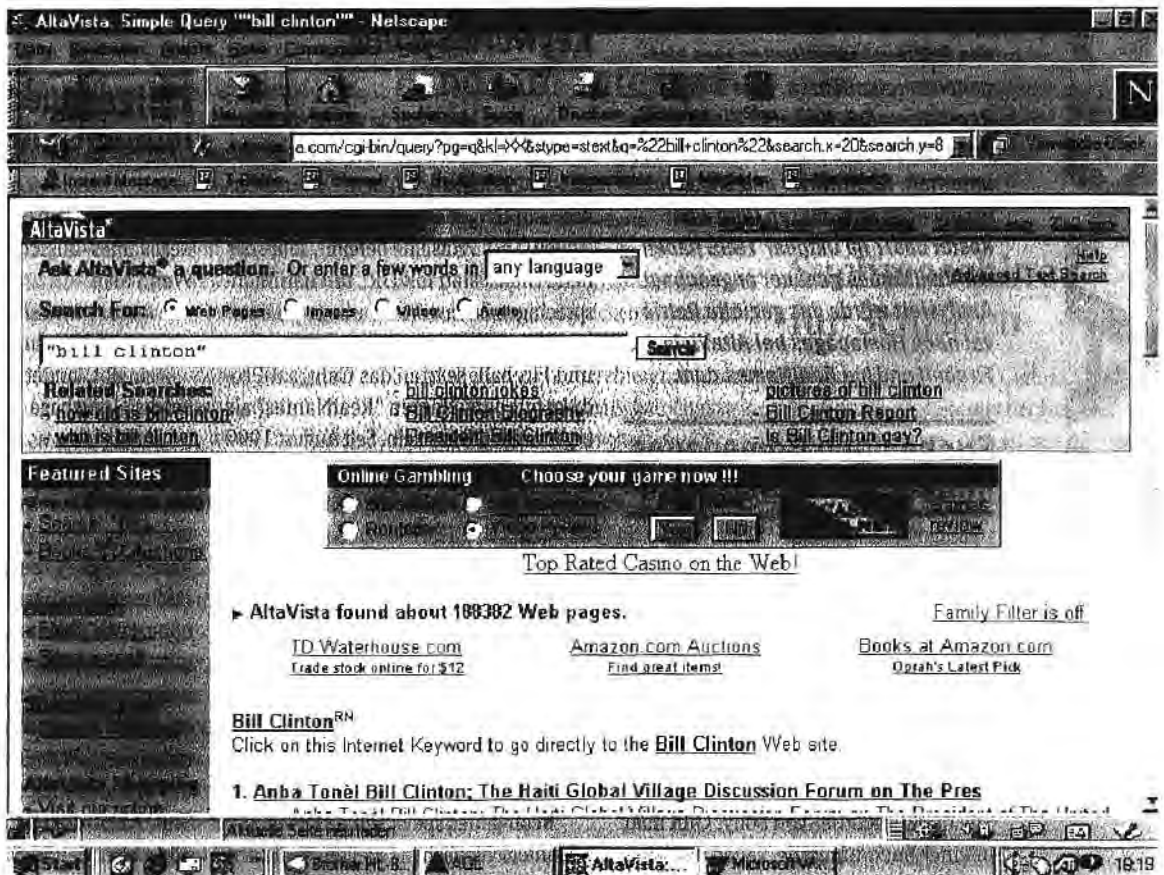


Abb. 1: Der RealNames-Link bei der Ergebnisanzeige von AltaVista

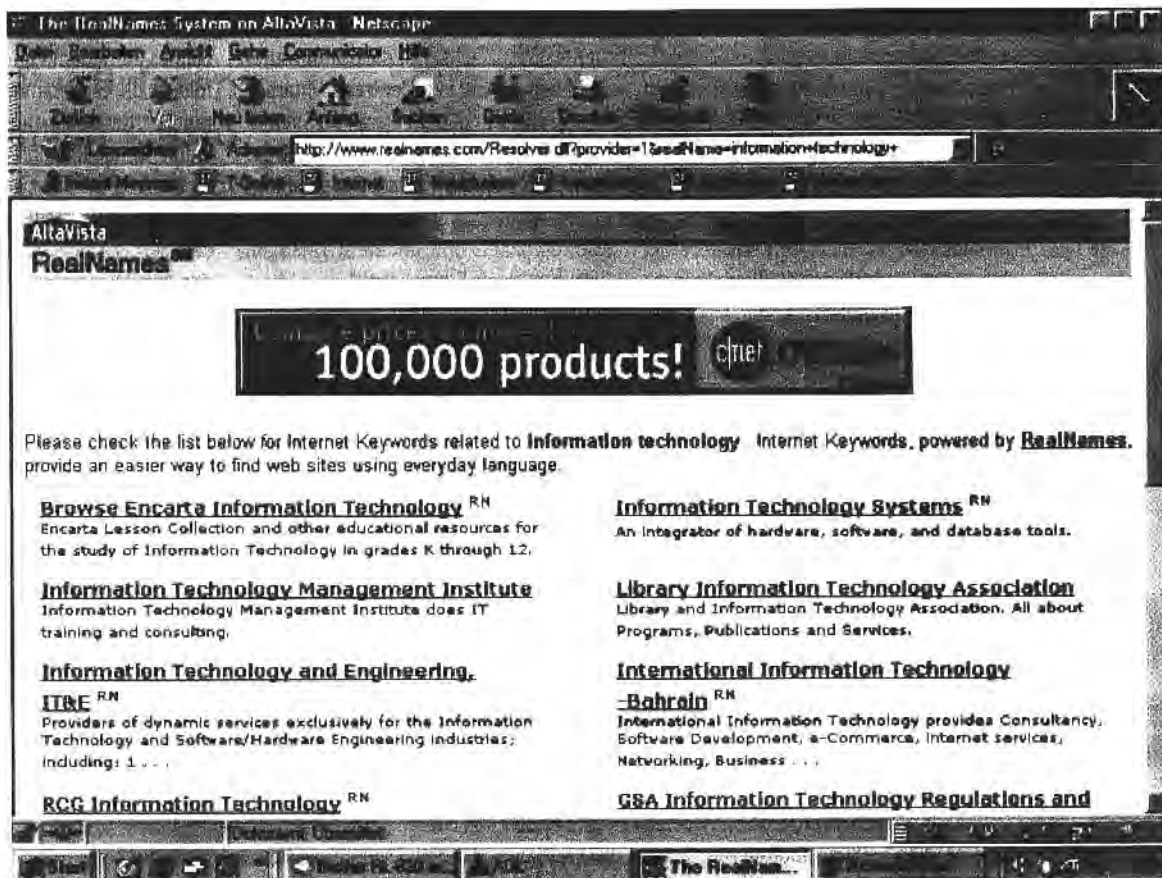


Abb. 2: RealNames-Links bei mehreren Treffern

Durchschalten zu anderen Informationsproduzenten

Durch Kooperationen sichert sich RealNames den direkten Zugang zu weiteren Internet-Sites. So ist es etwa möglich, sich (amerikanische) Steuervordrucke ausgeben zu lassen. Wichtiger ist der Link zu SECs EDGAR Datenbank, in der die vorgeschriebenen Lageberichte aller Unternehmen abgelegt sind, die an amerikanischen Börsen notiert sind. Kennt man die Nummer eines gewünschten US-Patentes, so wird man direkt zur Datenbank des amerikanischen Patentamtes weitergeschaltet und dort zum Patent geführt. Das Durchschalten läuft nicht bei Fireball, wohl aber bei AltaVista.

Hier ist ein gewisses Vorwissen bei der Recherche nötig. Recherchiert wird grundsätzlich als Phrase, also das gesamte Suchargument in Anführungszeichen. Je nach Informationsproduzent ist der erste Teil der Phrase "IRS", "SEC" bzw. "US Patent". Bei den Steuererklärungen folgt die Nummer des Formblattes (etwa: "IRS 1040"), bei SEC-Dokumenten der Unternehmensname (z.B. "SEC Amazon") und bei Patenten dessen Nummer (Beispiel: "US Patent 5,745,899").

Landesspezifische Varianten

Die Menge der Suchwörter, der "Name Space", und die daraus abgeleiteten Links sind bei gewissen Namen landesspezifisch ausgelegt. Die Eingabe ein und desselben Sucharguments führt bei diesen Termen bei AltaVista zu einem anderen Link als bei Fireball. Solche Varianten fanden wir bei den Homepages multinationaler Unternehmen, wo der Nutzer auf die jeweils landessprachliche Homepage geleitet wird. Unsere Eingabe von "Siemens" führte uns bei Fireball zu <http://www.siemens.de/de/home>, bei AltaVista gelangten wir zu <http://www.siemens.de/en/home/index.shtml>.

Abonnement von Suchwörtern

RealNames bietet drei Dienstleistungen an: ● Router-Software, die einen Nutzer, der ein registriertes Suchargument eingibt, direkt zur "passenden" Homepage navigiert; ● Datenanalyse für Abonnenten (im wesentlichen eine Benutzerstatistik) und ● Abbonnentenservice (Zuordnung von Suchargumenten zu einer URL).

Der Abbonnentenservice ist für Wirtschaftsunternehmen kostenpflichtig. Die Datenanalyse ist in diesen Kosten enthalten. Für den Nutzer entstehen keine Kosten. Die Preise sind gestaffelt: ● für Großunternehmen: abhängig von der Anzahl der vermittelten Links; ● für kleine und mittlere Unternehmen: Subskriptionspreis von \$ 100 pro Suchargument und Jahr; ● für Ausbildungseinrichtungen: kostenlos; ● für Individuen: kostenlos.

Ausbildungseinrichtungen werden bei der Anmeldung automatisch registriert. Institutionen außerhalb der USA haben - wegen ihrer URL - "Legitimationsprobleme". Individuen können nur dann kostenlos eingetragen werden, wenn sie gleichzeitig eine Homepage via "Homestead.com" (auch diese kostenlos) anlegen.

Der Preis von \$ 100 beinhaltet 2.500 Zugriffe auf das abonnierte Suchwort pro Monat. Wird die Grenze überschritten, gilt der Kunde als "Großunternehmen" und muß anhand der Anzahl der "Klicks" über den Preis neu verhandeln. Das Abbonnement gilt nur für ein Jahr. Welcher Preis dann im nächsten Jahr verlangt wird, ist offen. Sollte sich das Verfahren bewähren, sind Preissteigerungen nicht auszuschließen. Zu bedenken ist, dass die RealNames Corp. Monopolist für diese Art Dienstleistung ist.

Für wen lohnt das Abbonnement von Suchargumenten? Zunächst natürlich für alle, die kosten-

los die Dienste nutzen können. Für alle anderen kommt RealNames notwendig dann infrage, wenn die eigene URL nicht "sprechend" ist, also etwa nicht den (allseits bekannten) Unternehmensnamen enthält. Zudem sollten Firmen, die eigene Sites für ihre Produkte im Web pflegen, diese Seiten mit dem Produktnamen sowie mit der exakt passenden Beschreibung der Produktart via RealNames zugänglich machen. Eigentlich sollten alle Eigennamen (wie z.B. Unternehmens-, Produkt- oder Personennamen), als Suchargument eingegeben, zur Homepage des Namensträgers führen.

Darüber hinaus kann es zielführend sein, auch außerhalb der Eigennamen aussagekräftige Allgemeinbegriffe zu abonnieren. So wäre es doch durchaus sinnvoll, wenn vom Suchargument "Braunkohle" ein direkter Link zur Homepage der Rheinbraun AG führen würde. Hier müßte von RealNames sichergestellt sein, dass der Link auch thematisch einschlägig ist, da sonst der Dienst ad absurdum geführt würde.

Es ist möglich, mehrere Suchwörter für eine Web-Seite zu abonnieren.

RealNames garantiert die Einzigartigkeit des abonnierten Suchworts. Die Terme werden demnach nicht automatisch vergeben, sondern nach den Vorgaben der Abonnenten intellektuell geprüft. Unternehmens- bzw. Produktnamen können nur von denjenigen abonniert werden, die die Rechte an diesen Namen besitzen. Hier könnte ein Problem stecken: Dieselben Produktnamen, vergeben in unterschiedlichen Branchen, sind markenrechtlich möglich. Denken wir nur an "GBI": In der Informationswirtschaft ist es die "Gesellschaft für betriebswirtschaftliche Information", bei den Totengräbern das "Große Beerdigungsinstitut".

Im Gegensatz zum Zeichensatz bei den Suchmaschinen berücksichtigt RealNames Umlaute und Sonderzeichen. Die maximal zulässige Anzahl von Zeichen ist 80.

RealNames betont: "Der RealNames Service ermöglicht schnelle und anwenderfreundliche Navigation im Internet per Schlagwort. Mit umgangssprachlichen Namen kann ein User intuitiv zu privaten

und kommerziellen Web-Sites gelangen. Der Namespace (das ist die Summe aller RealNames) wird von Centraals Mitarbeitern gepflegt und nicht etwa von Suchrobotern zusammengestellt (wie es z.B. bei Suchmaschinen der Fall ist). Das garantiert qualitativ gute Navigation, den Schutz von Markennamen und anderem geistigen Eigentum sowie Einmaligkeit".

RealNames in Deutschland

RealNames Corp. ist in Silicon Valley zuhause. Nachdem das Produkt der RealNames-Links im amerikanischen Markt erfolgreich die Phase der Produkteinführung hinter sich gelassen hat, wendet sich das Unternehmen den Auslandsmärkten zu. Angesichts der globalen Aspekte des World Wide Web ist dieser Schritt naheliegend. In Deutschland hat man mit Fireball bereits einen Partner gefunden. Seit Mai 1999 besteht eine strategische Allianz zwischen RealNames Corp. und Gruner & Jahr Electronic Media Service, dem Betreiber von Fireball. Ansprechpartner bei RealNames ist der aus Deutschland stammende Michael Wibbeke.

Fazit

Aus Nutzersicht ist der Service von RealNames ausschließlich zu begrüßen. Es handelt sich hierbei um einen Dienst, der bei gewissen Suchargumenten den Nutzer direkt zur einschlägigen Homepage leitet. Führt man RealNames bei einer Suchmaschine wie AltaVista aus, ist keine gesonderte Eingabe nötig. Liegt ein RealNames-Link vor, wird dieser Link - abgesetzt von der Liste der Treffer und als RealNames-Link gekennzeichnet - an den Anfang gesetzt. Der Aufwand für den Nutzer ist demnach zu vernachlässigen; die Leistung kann hoch sein (wenn man genau diese Homepage suchte): fünf Punkte von fünf möglichen. Der Bedienungskomfort ist unschlagbar: Man muß gar nichts machen. Etwas problematisch ist allenfalls, dass man - beim "best-matching"-Verfahren - eine Liste von Homepages durchsehen muß. Wenn man nicht möchte, kann man die RealNames-Links auch ignorieren: ebenso fünf Punkte. Die Qualität der Inhalte läßt noch zu wünschen übrig. Dies hängt vor allem damit zusammen, dass noch zu wenige Unternehmen, Behörden und Individuen "ihre" Suchargumente abonniert haben. Der eine Punkt für die Inhalte ergibt sich also nur aus dem gegenwärtigen Stand aus deutscher Sicht.

Aus Abonnentensicht eignet sich RealNames auf jedem Fall für alle Suchargumente, die Eigennamen des Abonnenten repräsentieren (Abonnement von "Rheinbraun" für die Rheinbraun AG, von "Persil" für Henkel, von "Kerpen" für die Stadt Kerpen, von "Willi Bredemeier" für Willi Bredemeier). Vorsicht dürfte bei der Subskription von Allgemeinwörtern geboten sein: Hier sollten nur voll zutreffende und aussagekräftige Terme benutzt werden, die sicherstellen, dass ein Nutzer, der sie eingibt, von der entsprechenden Homepage nicht enttäuscht wird. ■

Wolfgang G. Stock

Weiterführende Informationen

RealNames Corporation
Keith Teare oder Michael Wibbeke
2 Circle Star Way
San Carlos, CA 94070-1350 U.S.A.
Tel.: 001/650-298-8080
Fax: 001/650-298-8035
homepage:
<http://www.realnames.com>
E-Mail: mwibbeke@realnames.com

Literatur

James A. Martin: RealNames Offers Home Page Finder. A free service makes it easy for the world to find your Web site. - In: *PC World Online* vom 20. Apr. 1999. - URL: http://www.pcworld.com/shared/printable_articles/0,1440,10565,00.html.

Danny Sullivan: Real Name Tops at AltaVista. - In: *The Search Engine Report* (June 3, 1998). - URL: <http://searchenginewatch.internet.com/sereport/9806-realname.html>.

Danny Sullivan: RealNames Expands Listings. - In: *The Search Engine Report* (Sept. 2, 1998). - URL: <http://searchenginewatch.internet.com/sereport/9809-realnames.html>.

Danny Sullivan: RealNames Expands Namespaces. - In: *The Search Engine Report* (Apr. 5, 1999). - URL: <http://searchenginewatch.internet.com/sereport/99/04-realnames.html>.

Danny Sullivan: Using RealName Links. - In: *The Search Engine Report* (Sept. 27, 1999). - URL: <http://searchenginewatch.internet.com/facts/realnames.html>.

Was ist RealNames? Centraal Corp., 1999. - URL: <http://heavymetal.fireball.de:4080/>.

Auf dem ersten Blick

RealNames

Aufwand-Leistungsverhältnis ○○○○○○

Bedienungskomfort ○○○○○○

Qualität der Inhalte ○