

Qualität

Internet-Suchwerkzeuge im Vergleich

Teil 1: Retrievaltest mit Known Item Searches

Web Search Engines und Web Directories sind die technologische Speerspitze im High-Tech-Bereich Internet. Als Suchassistenten gestatten sie Laien und Profi-Searchern gleichsam einen optimalen Einstieg in die Vielfalt der Informationsquellen des World Wide Web. Sie ermöglichen durch eine Kombination aus hierarchisch geordnetem kontrollierten Vokabular, automatischer Indexierung sowie diverser Synonym-, Homonym- und Fachwörterbücher einen zielgenauen Zugriff auf Web-Sites.

- Stop! Glauben Sie das? - Wir waren skeptisch und haben Suchmaschinen und Suchverzeichnisse genau angeschaut. Durchgeführt wurde ein Retrievaltest, der unterschiedliche Suchsysteme mit jeweils der gleichen Anfrage konfrontierte, auf die wir bereits die Antwort wussten ("Known Item Search"). Skizziert werden zudem bisher ausgearbeitete empirische Erhebungen zur Evaluation von Web-Retrievalsystemen, angefangen von seriösen informationswissenschaftlichen Forschungen bis hin zu den "Tests" in Publikumszeitschriften.

Die Entwicklung von Suchwerkzeugen im Internet boomt. In nur rund sechs Jahren haben sich derartige Retrievalsysteme bei Millionen von Nutzern durchgesetzt - die kommerziellen Systeme haben in den 30 Jahren ihrer Geschichte höchstens einige Zehntausend einer Informationseleite ansprechen können. Die erste erfolgreiche Suchmaschine ist WebCrawler, sie geht Anfang 1994 ins Netz. Bis Ende 1995 folgen Lycos, Infoseek (heute: Go), Open Text (stellt seinen Dienst allerdings bereits 1998 wieder ein), AltaVista und Excite. HotBot arbeitet seit 1996 und Northern Light sowie Alltheweb (heute: FAST) seit 1997. RealNames (ab 1997) spezialisiert sich auf die Suche nach Homepages. Google, das auf linguistische Verfahren beim Ranking verzichtet und statt dessen Links zählt, macht den etablierten Systemen seit 1998 ernste Konkurrenz. Ende 1999 treten mit Oingo und Simpli Systeme auf, die nicht wort-, sondern bedeutungsorientiert arbeiten. Mit Yahoo! (seit 1994) und dem Open Directory-Projekt (seit 1998) ergänzen klassifikatorische, intellektuell vorgehende Systeme die automatisch indexierenden Suchmaschinen.

Auch in Deutschland werden Suchwerkzeuge kreiert. Verwendet Fireball noch die eingekaufte Technik von AltaVista, setzen Systeme wie Acoon, Speedfind oder Spider auf eigene Software. Die meisten amerikanischen Systeme bieten inzwischen eine deutsche Version an.

Viele Anbieter von Search Engines haben ihre Entwicklungen durch Patente schützen lassen. Die IPK-Klasse G06F-017/30 (Information Retrieval; unterste Hierarchieebene der Internationalen Patentklassifikation) hat bei Inpadoc rund 20.000 Einträge (Einzelpatente), die Phra-

se "Search Engine" bringt bei der selben Datenbank ca. 210 Treffer. Die Patentstatistik signalisiert demnach, dass diverse technische Neuerungen bei den Retrievalsystemen zu begrüßen sind. Sind diese Neuerungen wirklich völlig neu? Oder wurden die grundlegenden Aspekte bereits vor Jahren in Dokumentation und Informationswissenschaft thematisiert? Und, vielleicht noch wichtiger, sind diese Neuerungen so gestaltet, dass sie die Probleme, die sie zu lösen verheißen, in der Tat lösen? Können Laien und Information Professionals im WWW gezielt suchen und finden?

Vielfach wird behauptet, dass die Suchmaschinen sich im günstigen Fall langsam Standards nähern, die in der kommerziellen Informationswirtschaft längst erreicht worden sind. "In fact, Internet search engines", schreibt Maryellen Mott Allen, "seem to act like rusted-out, temperamental locomotives than high-tech, streamlined information retrieval engines", um sich direkt anschließend über "the small world of search tools" zu beschweren (Allen 2000, 49).

Aus Unkenntnis oder auch schlichter Ignoranz der dokumentarisch-bibliothekarischen Geschichte würden grundlegende Errungenschaften wie z.B. Homonym- und Synonymbehandlung in Dokumentations-sprachen, das gewichtete Retrieval bis hin zum Aufbau ganzer Klassifikationssysteme neu entwickelt. Hier würden tatsächlich "die Bratkartoffeln neu erfunden". In der Praxis seien die Neuentwicklungen - insbesondere im Vergleich zu den theoretischen Vorgaben der Informationswissenschaft (man denke an SMART von Salton!) - weitgehend untauglich. Dies sei kein systematisches Retrieval, sondern vielmehr ein Stochern im Nebel. Um im (böartigen) Bild zu blei-

ben, nicht nur die Bratkartoffeln werden neu erfunden, nein, man hat zudem Salz und Zwiebeln vergessen. Zeichnen wir hier ein Zerrbild der gegenwärtigen Situation? Schauen wir uns die Web Search Engines und Directories doch genau an!

Aufgaben von Retrievalsystemen im World Wide Web

Retrievalsysteme im WWW stehen vor vier Problemen (vgl. Stock 2000, Kap. 6 und 7):

- (1) Auffinden der Dokumente
- (2) Indexieren der gefundenen Dokumente
- (3) Aktualisierung der Datenbasis
- (4) Angebot eines Suchsystems.

Im World Wide Web gibt es keine zentrale Instanz; jeder kann Dokumente ins Netz stellen. Es gibt damit auch keinen zentral geführten "Katalog" aller Dokumente. Entsprechend ist die erste Aufgabe der Systeme, die Dokumente überhaupt zu suchen und einzusammeln. Der zweite Schritt, die Indexierung, verläuft entweder automatisch via maschinelle Indexierung oder intellektuell durch die Einordnung von Dokumenten in Klassen. Dokumente im WWW verfügen z.T. über eine begrenzte Lebensdauer. Wenn der Autor so will, können sie jederzeit revidiert oder gelöscht werden. Entsprechend müssen die Retrievalsysteme im dritten Schritt "ihre" Dokumente periodisch überprüfen, ob sie überhaupt noch existieren. Letztlich bieten die Systeme eine Oberfläche zum Suchen an, wobei meist mehrere Einstiege etwa in eine "einfache Suche" oder in eine "fortgeschrittene Suche" vorgehalten werden.

Die Qualität eines Suchsystems hängt gleichermaßen von allen vier Aspekten ab. Was im ersten Schritt nicht gesammelt worden ist, ist für einen Nutzer natürlich nie abfragbar. Wenn die Datenbasis nicht häufig aktualisiert wird, bekommen wir Treffer, die dann zu der frustrierenden Meldung "File not found" führen. (Wird die Datenbasis im Gegenzug häufig verändert - was eigentlich positiv ist -, muss der Nutzer damit rechnen, dass ein Retrievalergebnis, das heute unter den ersten 20 eingereicht ist, morgen bei identischer Suchfrage unter den Top 20 nicht mehr aufscheint.) Arbeitet die Indexierung, vor allem der Ranking-

algorithmus, suboptimal, so sieht der Nutzer trotz Vorkommens in der Datenbasis die Volltreffer nicht, weil sie viel zu weit nach hinten sortiert worden sind. Weiß ein Nutzer nicht, was im Suchsystem an welcher Stelle eingegeben werden sollte, so nützt die beste Datenbasis und das beste Indexierungssystem nichts: Er wird kaum etwas Relevantes finden.

Das WWW und der Informationsbedarf von Wissenschaftlern

Als Einstieg in unsere komparative Analyse von Suchwerkzeugen im World Wide Web wollen wir einige der bisher durchgeführten informationswissenschaftlichen Evaluationen der Internet-Retrievalsysteme kurz referieren. (Für einen Überblick zu älteren Studien vgl. Dong/Su 1997.) Zwei der meistzitierten Artikel über die Qualität von Suchmaschinen stammen aus der Feder von Steve Lawrence und C. Lee Giles. Im Sinne unserer vier Aspekte von Suchwerkzeugen betrachten Lawrence und Giles in ihrem "Science"-Beitrag von 1998 zunächst ausschließlich den ersten, die relative Vollständigkeit des Einsammelns. Für 575 Anfragen von Wissenschaftlern aus dem Hause NEC im Dezember 1997 werden die Treffer notiert, allerdings nicht mehr als 600 pro Suchfrage. Die Gesamt-treffermenge wird auf 100% geeicht. Für sechs Suchmaschinen werden deren Anteile an der Gesamttreffermenge berechnet. Da es sich um eine Stichprobe handelt, rechnen die Autoren das 95%-Konfidenzintervall aus; das ist das Intervall, in dem sich der "wahre" Mittelwert - den wir ja nicht kennen - mit einer Wahrscheinlichkeit von 95% befindet. (Für Excite liegt demnach der "wahre" Abdeckungsgrad mit einer Irrtumswahr-

scheinlichkeit von 5% zwischen 22,2% und 24,0%.) Die entstehende Hitparade nach Abdeckungsgrad (siehe Tabelle 1) sieht HotBot mit 57,5% an der Spitze, gefolgt von AltaVista, Northern Light, Excite und Infoseek. Lycos wird mit nur 4,4% aller Antworten kaum von den NEC-Forschern benutzt.

Über die beobachtenden Überlappungen von Suchmaschinen versuchen unsere Autoren eine - allerdings ausgesprochen vage - Schätzung der Abdeckung des gesamten Web. Die Rangordnung aus Tabelle 1 bleibt dabei erhalten, lediglich die Prozentwerte differieren: HotBot 34%, AltaVista 28%, Northern Light 20%, Excite 14%, Infoseek 10%, Lycos 3% (vgl. Lawrence/Giles 1998, 100). Als Nebenergebnis berichten Lawrence und Giles über tote Links und geben damit unserem dritten Aspekt, der Update - Qualität, einen Ausdruck. Die Rangordnung der sechs betrachteten Systeme sieht Lycos mit 1,6% vorne, es folgen Excite (2,0%), AltaVista (2,5%), Infoseek (2,6%), Northern Light (5,0%) und HotBot (5,3%). In einer zweiten Untersuchung, publiziert in "Nature" im nächsten Jahr, zeigt sich für die Autoren ein Absinken des Abdeckungsgrads des Webs; keine Suchmaschine erreicht mehr als 16%. Die Aktualität sowohl des erstmaligen Einsammelns von Seiten als auch des Updating läßt extrem zu wünschen übrig und dauert im Schnitt mehrere Monate. Das Fazit klingt vernichtend: "The current state of search engines can be compared to a phone book which is updated irregularly, is biased toward listing more popular information, and has most of the pages ripped out" (Lawrence/Giles 1999).

Recall und Precision bei Suchwerkzeugen

Tabelle 1: Abdeckungsgrad von Suchmaschinen für wissenschaftliche Anfragen

Suchmaschine	Abdeckungsgrad	95% Konfidenzintervall
HotBot	57,5%	± 1,3 %
AltaVista	46,5%	± 1,3 %
Northern Light	32,9%	± 1,1 %
Excite	23,1%	± 0,9 %
Infoseek	16,5%	± 1,0 %
Lycos	4,4%	± 0,4 %

Quelle: Lawrence/Giles 1998

In einer sehr breit angelegten Studie aus dem Jahr 1999 evaluieren Michael Gordon und Praveen Pathak acht Suchwerkzeuge mittels der klassischen Maße Recall und Precision. Wir konzentrieren uns hier auf einige wenige ausgesuchte Experimente und Resultate. 33 Anfragen von Wissenschaftlern der Michigan Business School werden von Profi-Searchern mit einem Arbeitsaufwand von ca. ein bis zwei Tagen pro Anfrage mittels der acht Suchwerkzeuge abgearbeitet. Pro Frage und Suchsystem speichern die Searcher die jeweils ersten 200 Treffer. Sie verfolgen die Links der ersten 20 Treffer (also 160 für jede Suche), drucken die Web-Sites aus und sortieren sie in einer zufälligen Reihenfolge in einen Sammelband, den sie den Auftraggebern überreichen. Diese entscheiden für jede Web-Site, ob diese für die Suchfrage Relevanz hat oder nicht. (Die Treffer Nr. 21 bis 200, die ja den Wissenschaftlern nicht vorliegen, kommen dann ins Spiel, wenn eine andere Suchmaschine dieselbe Seite unter Nr. 1 bis 20 sortiert. Man kann nunmehr nachsehen, ob die anderen Suchwerkzeuge sie gefunden haben.) Die Präzision ist der Quotient aus der Anzahl der als relevant eingestuftten Dokumente und der Gesamtanzahl der Dokumente. Die Präzision wird für jede Ausgabeposition (Cut-off-Wert) einzeln berechnet (also für den jeweils ersten Treffer, den zweiten, dritten usw. bis zum zweihundertsten), pro Suchwerkzeug wird der Mittelwert aus den 33 Fragen errechnet. Zusätzlich eruieren unsere Autoren eine durchschnittliche Precision für Positionsintervalle (u.a. für 1-5 und 1-10), die sich wiederum als arithmetisches Mittel der gegebenen Mittelwerte darstellt. Die Position-Eins-Präzision schwankt zwischen 66,7% bei Open Text und 12,1% bei Yahoo!. Zwischen Position 1 und 9 differieren die Rangplätze der Suchmaschinen, ab Position 10 zeichnet sich ein eindeutiger Trend ab: AltaVista, Open Text und Lycos sind - in dieser Reihenfolge die "top precision performers";

Tabelle 2: Rangordnungen von Suchmaschinen nach Recall und Precision

Rang	Precision	Rang	relativer Recall
1	AltaVista	1	AltaVista
2	Open Text	2	Open Text
3	Lycos	3	Lycos
4	Infoseek	4	Infoseek
4	Excite	4	Excite
4	Magellan	4	Magellan
4	HotBot	7	Yahoo!
8	Yahoo!	8	HotBot

Quelle: Gordon/Pathak 1999

Yahoo! ist das Schlusslicht; Infoseek, Excite, Magellan und HotBot nehmen im Mittelfeld je nach Cut-off-Wert unterschiedliche Ränge ein (vgl. Gordon/Pathak 1999, 154 ff.). Die Präzision liegt bei AltaVista zwischen den Cut-Off-Werten 10 und 20 bei durchschnittlich etwa 40%, Lycos liegt bereits 10% darunter; Yahoo! erreicht gerade noch 15%. Die Precision-Werte fallen ab der Rangposition 25 zunächst stark, dann (ab etwa 50) nur noch schwach ab. AltaVista unterschreitet die 20%-Grenze bei Rang 50, die 15%-Grenze bei 70 und die 10%-Grenze in etwa bei einem Cut-off von 100. Das heißt: Jede zehnte Web-Site, die auf die hundertste Stelle bei der Trefferanzeige von AltaVista sortiert ist, ist für den Nutzer relevant.

Recall errechnet sich als Quotient der gefundenen relevanten Dokumente und aller relevanter Dokumente. Da sich ein "absoluter Recall" (geeicht auf die Gesamtmenge der relevanten Dokumente im Web) nicht erfassen lässt, gehen Gordon/Pathak von einem "relativen Recall" aus, der sich aus der Summe der relevanten Dokumente aus den acht Suchwerkzeugen ergibt. Die Autoren berechnen wiederum für jeden Cut-off-Wert den Recall, wobei sie kumulativ vorgehen. Die Rangordnung der Suchwerkzeuge nach Recall gleicht der der Precision, lediglich Yahoo! gewinnt (zuungunsten

von HotBot) einen Platz (vgl. ebd., 160 ff.). AltaVista erreicht bis Sortierposition 20 einen Recall von knapp 16% und verbessert sich auf rund 22% bei Cut-off 200.

Die Untersuchung von Gordon und Pathak beschränkt sich auf den Informationsbedarf von universitären Wissenschaftlern. Betrachtet man den mageren Recall und die ebenfalls nicht überzeugende Precision, die in diesem Experiment zu Tage treten, so ist das Fazit der Autoren naheliegend: "Absolute retrieval effectiveness is fairly low" (ebd., 173).

Precision in Nutzersicht

Eine Evaluation von Suchmaschinen aus Anwendersicht legen Louise T. Su, Hsin-liang Chen und Xiaoying Dong bei der Jahrestagung der American Society for Information Science 1998 vor. Elf Endnutzer der Universität Pittsburgh, darunter acht Studenten, führen jeweils eine Recherche nach wissenschaftlichen Aspekten in vier Suchsystemen durch und bewerten ihre Ergebnisse. Die Precision wird für jeden der ersten 20 Treffer der vier Suchmaschinen auf einer Dreierskala (relevant, teilweise relevant, nicht relevant) beurteilt. Zusätzlich werden die Endnutzer durch die Autoren interviewt und dabei gebeten, ein Urteil über die Zufriedenheit mit der Präzision der Suchwerkzeuge zu fällen.

Su et al. berechnen sowohl die Präzision bzgl. der relevanten und teilweise relevanten Dokumente (was wir "schwache Precision" nennen wollen) als auch bzgl. nur der relevanten Web-Sites ("starke Precision"). In Tabelle 3 sind die Mittelwerte aus den Suchläufen notiert. Die höchste (schwache wie starke) Genauigkeit weisen Lycos (Rang 1) und AltaVista

Tabelle 3: Precision und Zufriedenheit von Endnutzern

Indikator	AltaVista	Infoseek	Lycos	OpenText
schwache Precision	68,12% (2)	63,18% (3)	70,00% (1)	60,45% (4)
starke Precision	30,46% (2)	25,00% (4)	35,00% (1)	28,64% (3)
Zufriedenheit	4,91 (1)	4,14 (4)	4,77 (2)	4,36 (3)

Quelle: Su/Chen/Dong 1998; in Klammern der jeweilige Rangplatz

(Rang 2) auf, Infoseek und Open Text tauschen Plätze drei und vier. Spannend ist das Ergebnis, dass die subjektive Zufriedenheitseinschätzung eine andere Rangfolge ergibt. Nunmehr liegt AltaVista vor Lycos. Die Autoren stellen fest: "It is interesting to note that although Lycos has the highest mean precision ratios based on both calculations, users are more satisfied with the precision ratios achieved by AltaVista. Other factors must be at play" (Su/Chen/Dong 1998, 354). Hier wird Forschungsbedarf deutlich: "These contradictory findings need to be further examined" (ebd., 359).

First 20 Precision"

H. Vernon Leighton und Jaideep Srivastava verlassen in ihrem Retrievaltest den ausschließlich wissenschaftlichen Bereich und analysieren 15 Suchfragen, die bei der Auskunft wissenschaftlicher Bibliotheken gestellt worden sind. Dies sind zwar mehrheitlich durchaus wissenschaftliche Probleme, aber auch allgemeinere Sachfragen (wie z.B. über die Nationalgalerie in Prag oder die Gruppe "Queer Nation"). Im Test stehen fünf Suchmaschinen, deren Präzision innerhalb der ersten 20 Dokumente erhoben wird. Jede der ausgegebenen Dokumente wird genau einer Kategorie zugeordnet: Duplikat (gleiche Basis-URL; nicht: Spiegel), inaktiver Link, Kategorie 0 (irrelevant), Kategorie 1 (formal in Ordnung, da die Suchbegriffe auftauchen, inhaltlich aber irrelevant), Kategorie 2 (potentiell relevant) und Kategorie 3 (relevant). Je nach Rangplatz in der Ausgabe wird gewichtet, Positionen 1 bis 3 mit Faktor 20, Positionen 4 bis 10 mit Faktor 17 und

Positionen 11 bis 20 mit Faktor 10. Für die fünf Suchmaschinen berechnen die Autoren den Median (50%-Punkt) aus den 15 Suchfragen. Leighton und Srivastava führen mehrere Experimente durch; von drei Ansätzen wollen wir berichten. In Experiment 1 gelten alle Treffer der Kategorien 1 bis 3 als relevant ("formale Precision", Experiment 2 schränkt auf Kategorien 2 und 3 ein (in unserer Terminologie "schwache Precision"), Experiment 3 auf Kategorie 3 ("starke Precision").

Nehmen wir den Median aller als Schwellenwert, so liegen Excite, AltaVista und Infoseek stets über diesem Wert, HotBot und Lycos darunter. "It is clear looking at the data that, in general, AltaVista, Excite, and Infoseek did a superior job delivering quality relevance" (Leighton/Srivastava 1999, 880).

"Leistungstests" deutscher Suchwerkzeuge

In deutschen Publikumszeitschriften finden sich "Tests", die die breite Masse der Internetnutzer mit Bewertungen von deutschen Suchwerkzeugen unterrichten. Wie seriös sind solche Evaluationen?

PC-Online untersucht sieben Suchmaschinen und fünf Kataloge. Bei international agierenden Werkzeugen werden - so vorhanden - die deutschen Versionen abgefragt. Fünf Kriterien mit jeweils unterschiedlicher Gewichtung (Suche: 10%; Operatoren: 10%; Präsentation: 10%; Extras: 5%; Suchtest: 65%) ergeben für jedes Suchwerkzeug einen numerischen Wert. Der "Suchtest" umfasste fünf Fragen (u.a.: "CD von Pete Namlook", "die günstigsten Internet-by-

Call-Tarife"). Als Kriterium für das Bestehen des Suchtests wird notiert: "Die jeweils ersten zehn Ergebnisse prüften wir auf Qualität. Bei den Katalogen werteten wir, wie nützlich die Informationen sind, und ließen die Struktur in die Ergebnisse einfließen" (PC-Online 2000, 48). Hier wird offenbar ein Relevanztest durchgeführt, wo ein Tester subjektiv zu entscheiden hat, ob innerhalb der ersten zehn Ergebnisse - für ihn - ein Treffer vorliegt. Testsieger bei den Suchmaschinen sind Google, Fireball und AltaVista; etwas dahinter landen Infoseek, Lycos und Excite im Mittelfeld, abgeschlagener Letzter ist Acoon. Bei den deutschen Webkatalogen sehen wir die Reihenfolge Yahoo!, Web.de, Dino, Allesklar und Bellnet.

Neben Laminaten und Dampfbügel-eisen wendet sich die **Stiftung Waren-test** im August-Heft 2000 den Internet-Suchmaschinen zu. Drei Aspekte werden durch Schulnoten bewertet: die Hilfsfunktionen, die Darstellung der Ergebnisse sowie die Resultate eines Suchtests. Es handelt sich dabei um einen Retrievaltest, gestellt werden vier allgemeine Fragen (wie "Infos für Briefmarkensammler" oder "einen Vergleich der Stromtarife") sowie vier konkrete Informationswünsche (u.a. die Antrittsrede von Johannes Rau und Nachrichten über den Besuch von Verona Feldbusch bei Big Brother). Mit den zehn Fragen werden zehn deutsche Suchmaschinen, vier Metasuchmaschinen (hier übergangen) und vier Kataloge konfrontiert. Wir haben die drei von "Test" vergebenen Schulnoten addiert und einen Mittelwert ausgerechnet. Eine glatte "2" erhalten AltaVista, Fireball und Google, eine "2 minus" Infoseek, MSN und Yahoo!, die schlechtesten Noten erhalten Allesklar (4 plus), Dino (3 minus), HotBot und Lycos (jeweils eine "3").

Kritisch zu sehen ist bei den "Tests" in den besprochenen Publikumszeitschriften die Relevanzbestimmung der "Treffer". Über exakte Relevanzkriterien wird nichts berichtet, so dass den Ergebnissen durchaus eine subjektiv gefärbte Zufälligkeit anhaftet. Alle Tests sind keine reinen Retrievaltests, sondern eingebettet in breitere Fragestellungen. Dies ist jeweils mitnichten verwerflich (obwohl man über gewisse Kriterien - etwa Hilfsfunktionen oder "Extras" - streiten könnte), verhindert aber die Vergleichbarkeit der Testergebnisse. Es wird nur ein "deutscher Gesichtspunkt"

Tabelle 4: First 20 Precision bei unterschiedlichen Relevanzkriterien

	Experiment 1 formale Precision	Experiment 2 schwache Precision	Experiment 3 starke Precision
AltaVista	0,90	0,45	0,06
Excite	0,94	0,47	0,07
HotBot	0,72	0,29	0,04
Infoseek	0,87	0,51	0,10
Lycos	0,61	0,27	0,03
Median aller	0,81	0,40	0,06

Quelle: Leighton/Srivastava 1999

gewählt, der der Internationalität des World Wide Web auf keinem Fall gerecht wird. Bei den informationswissenschaftlichen Studien dominieren wissenschaftliche Suchfragen; hier ist es genau anders herum: "alltägliche" Informationsbedürfnisse werden thematisiert.

Kritik

Die aus informationswissenschaftlicher Perspektive geschriebenen Forschungsartikel haben den Nachteil, dass sie bevorzugt vom Informationsbedarf eines Wissenschaftlers ausgehen. Zudem beschränken sie sich auf die englische Sprache bei der Formulierung der Suchanfragen. Ein zentraler Schwachpunkt aller Precision-Untersuchungen ist die Relevanzbeurteilung, die ein nicht zu unterschätzendes subjektives Moment enthält. Was für die eine Testperson relevant ist, kann für die nächste durchaus irrelevant sein. Deutlich wird dieses methodologische Problem u. a. bei der Inkonsistenz der Urteile, von denen Su et al. berichten. Unsere deutschen Retrievaltests berücksichtigen zwar die deutsche Sprache und nicht-wissenschaftliche Suchfragen, sind aber keine reinen Suchtests, sondern eine bunte Mischung aus mehreren Faktoren, die jede Vergleichbarkeit verbietet.

Ein immenses Problem ist, dass Tests unterschiedlicher Forschergruppen durchaus voneinander abweichende Resultate bringen. So belegt z. B. Lycos mit derselben Fragestellung (nach der Precision) bei Su/Chen/Dong Platz 1 (vor AltaVista), bei Gordon/Pathak Rang 3 hinter AltaVista (hier auf Rang 1) und bei Leighton/Srivastawa Rang 5 sehr weit hinter AltaVista. Hier scheinen große methodologische Probleme vorhanden zu sein.

Unser Retrievaltest: Known Item Search

Unser Retrievaltest geht von Dokumenten aus, deren Existenz wir wissen. (Damit umschiffen wir das Problem der Relevanzbewertung.) Die Search Engines und Directories werden jeweils mit der selben Suchfrage konfrontiert. Recherchiert wird nach Web-Sites, die uns bekannt sind und deren URL am Tag der Suche auch aktiv ist. Ein solcher Test wird in der Bibliotheks- und Informationswissenschaft als "Known Item Search" bezeichnet (vgl. Kantor 1976), die Trefferquote zeigt die Verfügbarkeit (Availability) der Bestände der jeweiligen Datenbank.

Known Item Searches sind bislang vor allem zur Evaluation von Bibliotheken eingesetzt wor-

Tabelle 5: Rangordnungen von Suchmaschinen in deutschen Publikumszeitschriften

PC-Online	Stiftung Warentest
Rang / Suchwerkzeug / Punkte	Rang / Suchwerkzeug / Note
1 / Google / 82	1 / AltaVista / 2,0
2 / Yahoo! / 79	1 / Fireball / 2,0
3 / Web.de / 77	1 / Google / 2,0
4 / Fireball / 76	4 / Infoseek / 2,3
5 / AltaVista / 74	4 / MSN / 2,3
6 / Dino / 64	4 / Yahoo! / 2,3
7 / Infoseek / 63	7 / Acoon / 2,7
8 / Lycos / 60	7 / Excite / 2,7
9 / Excite / 59	7 / Web.de / 2,7
9 / Allesklar / 59	10 / Aladin / 3,0
11 / Bellnet / 57	10 / HotBot / 3,0
12 / Acoon / 47	10 / Lycos / 3,0
	13 / Dino / 3,3
	14 / Allesklar / 3,7

(Quellen: PC Online; Test 8/2000)

den. Bibliotheksbenutzer, die einen konkreten Ausleihwunsch - eben das "Known Item" - haben, werden befragt, ob ihr Wunsch beim Besuch in der Bibliothek sofort befriedigt werden kann. Der Quotient aus der Anzahl aller Ausleihwünsche und der Anzahl der erledigten Ausleihen ist ein quantitativer Indikator zur Bestandsverfügbarkeit konkreter Bibliotheken. In der Literatur wird über Verfügbarkeitsraten von rund 60% in us-amerikanischen und von gut 40% in deutschen Bibliotheken berichtet. Nicht-Verfügbarkeit kann unterschiedliche Gründe haben: Ein Buch ist überhaupt nicht angeschafft worden, es ist derzeit ausgeliehen, es steht am falschen Standort, der Nutzer findet es nicht. Bei Detailuntersuchungen zur Availability wird so analysiert, wo eine Bibliothek Probleme hat. Analog dazu wäre bei unserer Suchwerkzeug-Availability zu fragen, an welchem Aspekt (siehe oben die vier Problemfelder) die Nicht-Verfügbarkeit gescheitert ist. Unser Test wird ausschließlich die Gesamt-Verfügbarkeitsrate messen, nicht die einzelnen Komponenten. (Hier ist noch reichlich Forschungsbedarf vorhanden.)

Das Experiment

18 Suchsysteme stehen im Test. Ausgewählt werden diese wegen ihrer "Prominenz", operationalisiert durch das Auftreten in einschlägigen Lehrbüchern (wie etwa Glossbrenner/Glossbrenner 1999, Hock 1999 oder Spallek/Kreinacke 2000) bzw. bei den "Beobachtungsposten" des Web (vgl. Notess, Sullivan),

oder wegen besonders interessanter Vorgaben wie das Beherrschen "semantischen Retrievals" bei Oingo und Simpli (vgl. Feldman 2000). Zusätzlich kommen die deutschen Suchmaschinen Fireball (eine Tochter von Gruner + Jahr in Hamburg), Acoon (ein Unternehmen der Drillisch AG in Marl), Spider (mit der Suchmaschine Abacho; Teil der Endemann!! Internet AG in Neuß) sowie Speedfind (ein Produkt der freenet AG in Hamburg) ins Visier. Keine Berücksichtigung finden in diesem Experiment Meta-Suchmaschinen, können diese doch nur so gut sein wie die von ihnen angesprochenen Suchwerkzeuge. Es geht uns ausschließlich um die **Suchfunktionalität** der getesteten Systeme, nicht um Eigenschaften wie redaktionelle Informationen (Nachrichten usw.), Portfolios (zur Überwachung von Börsenkursen) oder Dienste zum E-Commerce.

Insgesamt zwanzig Know Items werden gesucht; die Suchargumente sind so ausgewählt, dass sie erstens auf der Web Site an prominenter Stelle (etwa im Titel oder im ersten Textabschnitt) vorkommen und zweitens diese Site recht genau beschreiben. Die 20 Sites verteilen sich je zur Hälfte auf deutsch- und englischsprachige Quellen (1-10 deutsch; 11-20 englisch) sowie auf Literatur- und Produktquellen (1-5, 11-15 Literatur; 6-10, 16-20 Produkte). Hat eine Suchmaschine eine deutsche und eine englischsprachige Oberfläche, so werden die deutschen Quellen mittels der deutschen Variante, die restlichen bei der englischen Version gesucht. Mittels der Zweisprachigkeit wollen wir der sprachlichen "Einseitigkeit" der bisher durchgeführten Tests entgegenwirken. Das

WWW ist - soviel ist nach den Vorarbeiten überdeutlich - nur am Rande ein Hort wissenschaftlicher Informationen. Deshalb unterteilen wir die Suchfragen nach (auch wissenschaftlichen) Literaturquellen sowie dem Marketing zuzurechnenden Produktinformationen. Tabelle 6 listet die URL der Known Items sowie die Suchargumente auf. Der Retrievaltest wird in den ersten beiden Oktoberwochen 2000 durchgeführt.

Zwei Testläufe pro Suchwerkzeug werden (mit einer Ausnahme) angegangen. Testlauf 1: In der "einfachen Suche" der Suchmaschinen werden die Suchargumente, durch Leerzeichen getrennt, ohne jeden Operator eingegeben. Testlauf 2: Bietet ein Retrievalsystem zusätzlich eine "fortgeschrittene (oder Profi-) Suche" an, so wird unter optimaler Ausnutzung des Befehlsumfangs (also einschließlich Truncation, Booleschen Operatoren, Abstandsoperatoren, Phrasenmarkierung, Homonymkontrolle usw.) recherchiert. Gibt es keine ausgewiesene Profi-Suche, aber bestimmte Operatoren (etwa "+" oder "-"), so werden diese im zweiten Testlauf eingesetzt. Insgesamt werden 700 Suchläufe gestartet.

Als "Treffer" wird gezählt, wenn die bekannte URL innerhalb der ersten zwanzig Anzeigen vorkommt, ansonsten wird

"Fehlanzeige" notiert. Spiegelungen inhaltgleicher Sites gelten als Treffer. Tabelle 7 zeigt die Ergebnisse. Eine "1" oder "0" signalisiert Treffer bzw. Fehlanzeige; zwei Zeichen bei einer Spalte bedeuten "einfache Suche" (links) und "Profi-Suche" (rechts). RealNames bietet keinerlei Profi-Suche an; deshalb ist dort nur eine Zahl angegeben.

Ergebnisse nach Suchfragen

Suchfrage 1 nach der "Mitarbeitermotivation in Bibliotheken" scheitert bei allen Suchwerkzeugen. Im August 2000 wurde eine ältere HTML-Version zugunsten einer neuen PDF-Version (mit HTML-Titelblatt) unter einer neuen URL aufgelegt. Die Update-Zyklen der Retrievalsysteme sind offenbar so lang, dass innerhalb von zwei Monaten noch keine Aktualisierung erfolgte. Auch Frage 13 nach "Robin Hood und Sherwood Forest" bringt keinen einzigen Treffer. Möglicherweise liegt hier der Fehler bei uns; es gibt sehr viele Sites zum Thema. Unser Known Item kommt - vielleicht weil zuviele ähnliche Seiten vorliegen - nie unter die ersten 20. Ein ähnliches Problem könnte zu dem schlechten Ergebnis bei Nummer 19 ("Nike's Fußballschuhe") geführt haben.

Peinlich erscheint uns das Ergebnis zu Frage 16 zur diesjährigen Online Information in London. Der Algorithmus zum Einsammeln von Dokumenten arbeitet offenbar genauso schlecht wie der zum Update. **Aktuelle Ereignisse können mit Suchwerkzeugen kaum erfolgreich recherchiert werden.** Der Informationsbedarf nach "Star Wars II" (Nr. 18) wird mehrfach durch das Angebot anderer Star Wars Episoden "gedeckt", so dass die Vermutung naheliegt, dass die Suchmaschinen mit der kurzen Zeichenfolge "II" nicht klarkommen. Frage 15 stellt eine Herausforderung an Suchmaschinen dar: Shakespeares "to be or not to be" enthält sowohl Operatoren (or und not) als auch Stoppwörter (to und be). Die Phrase wird (bei der einfachen Suche) von keinem System erkannt. Erst bei der Profi-Suche, wo wir die Phrase markieren, kommen Ergebnisse.

Wesentliche Unterschiede zwischen einfacher und fortgeschrittener Suche realisieren die Testläufe 11, 12 und 18. Bei den anderen Testfragen erzielt die Profi-Suche entweder die gleiche Menge an Treffern oder ein bis zwei weitere (über alle Suchwerkzeuge hinweg). Unverständlich ist, dass die fortgeschrittene Suche bei Frage 20 beim Klassifikationssystem des "Open Directory" keinen Treffer bringt, wo doch

Tabelle 6: Known Items (URLs) und entsprechende Suchfragen

Nr.	URL (http://)	Suchfrage
1	www.fbi.fh-koeln.de/fachbereich/papers/kabi/band.cfm?ID=26	Mitarbeitermotivation, Bibliothek
2	www.medicine-worldwide.de/krankheiten/orthop_erkrankungen/klumpfuss.html	Pes equino-varus congenitus
3	www.erft.de/schulen/rendsburger/recht/srjoef.htm	Jugendlicher, Gaststätte, Nachtclub, Gesetz
4	www.abi99-dbg.purespace.de/referate/zucker/zucker.htm	Zucker, Zuckerrübe, Raffinadezucker
5	www.mherger.com/hackbrett/	chromatisches Hackbrett
6	www.volkswagen.de/sharan/index_.htm	VW, Sharan
7	www.semket.com/winzer/winzer.htm	Flaschetikettierer, Winzer
8	www.wein-juwel.de/produkte/346642414981.html	Schilcher, Blauer Wildbacher, Steiermark
9	www.lambertz.de/html/body_saisonartikel1.html	Lambertz Nuss-Honig Saftprinten
10	www.bierclub.de/BDM/11-99/bdm11-99.htm	Hohenfelder Pilsener
11	www.wchstv.com/gmarecipes/plumpudding.html	plum pudding, christmas, recipe
12	www.onlineinc.com/onlinemag/OL1998/hock5.html	field searching, search engine
13	www.mansfieldpages.co.uk/Robin.htm	Robin Hood, Sherwood Forest
14	www.sna.net.au/greatmoments/bios/hogan/greatmom_bio_hogan.html	Paul Hogan, biography
15	andromeda.rutgers.edu/~jlynch/Texts/hamlet.html	To be or not to be, Samuel Johnson
16	www.online-information.co.uk/ol00/index.asp	Online Information 2000 London
17	bvkmall.com/bvkmall/maupcounmanc.html	mango chutney, Hawaii
18	www.starwars.com/episode-ii/	Star Wars II
19	nike.com/football/	Nike, football, footwear
20	www.angebears.com/	handmade Teddy bear, Angie Rogers

Tabelle 7: Testergebnisse der Known-Item-Search

Nr.	Accom	FAST	AltaVista	Excite	Fireball	Google	HotBot	Infoseek/ Lycos	NorthernL.	Dingo	Simpli	Speedfind	Spider	WebCrawler	RealNames	Yahoo!	ODP	Trefferanzahl	Availability
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%
2	0	0	1	0	1	1	0	0	1	0	0	0	0	0	0	0	0	4	22%
3	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	11%
4	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	17%
5	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	3	17%
6	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	4	22%
7	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	11%
8	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	6%
9	0	1	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	7	39%
10	0	1	0	0	1	1	0	0	1	0	0	0	1	0	0	0	0	6	33%
11	0	1	1	0	1	1	0	0	0	0	0	0	0	1	0	0	0	7	39%
12	0	1	0	0	1	1	0	0	1	0	0	0	0	1	0	0	0	5	28%
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%
14	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	7	39%
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	39%
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%
18	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	6%
19	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	22%
20	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	33%
	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%
	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	72%

Trefferanzahl	0	4	8	10	8	12	4	4	7	10	12	13	6	6	4	7	5	6	9	11	3	8	2	2	1	3	1	1	2	3	1	0	75	102
Availability	0%	20%	40%	50%	40%	60%	20%	20%	30%	30%	65%	30%	65%	60%	65%	30%	35%	20%	35%	25%	30%	45%	55%	15%	40%	10%	10%	0%	5%	10%	5%	15%	5%	0%
Rang*	10	4	2	10	4	1	0	0	7	0	14	16	14	12	16	12	10	16	10	16	6	12	14	14	14	12	12	12	12	12	12	16	16	16

* = 700 Suchbänfe; 16 Suchwerkzeuge; "entlastete Suche"; 17 Suchwerkzeuge; "fortgeschrittene Suche"; jeweils 20 Suchfragen; * = Rang nach bestem Ergebnis; * bei Yahoo! = Treffer via Google

Tabelle 8: Availability beim Einsatz von zwei und drei Suchwerkzeugen

Google (65%) + AltaVista (60%)	80%
Google (65%) + Northern Light (55%)	80%
Google (65%) + FAST (50%)	80%
Google (65%) + Fireball (50%)	75%
AltaVista (60%) + Northern Light (55%)	65%
AltaVista (60%) + FAST (50%)	75%
AltaVista (60%) + Fireball (50%)	60%
Northern Light (55%) + FAST (50%)	75%
Northern Light (55%) + Fireball (50%)	65%
FAST (50%) + Fireball (50%)	70%
Google + AltaVista + Northern Light	80%
Google + AltaVista + FAST	90%
Google + NorthernLight + FAST	90%

jeweils fortgeschrittene Suche;
in Klammern Trefferhäufigkeit des Werkzeugs allein; Known Item Search

die einfache Suche erfolgreich ist. Im Prinzip kann man feststellen, dass es durchaus lohnt, unter voller Ausnutzung der angebotenen Suchfunktionalität des jeweiligen Werkzeugs "fortgeschritten" zu recherchieren: Oingo bringt so fünf Treffer mehr, Acoon und AltaVista bringen vier weitere sowie Fireball und Infoseek/Go je drei weitere zutreffende Ergebnisse.

Ergebnisse nach Suchwerkzeugen

Die größte Availability hat Google; 65% unserer Known Items werden bei der fortgeschrittenen Suche gefunden. Auf den Plätzen landen AltaVista (60%) und Northern Light (55%). Fireball (50%), FAST (ebenfalls 50%) sowie Oingo (40%) bringen noch recht brauchbare Resultate. Infoseek/Go (35%), Lycos (30%) und HotBot (30%) liegen schon bei nur etwa der Hälfte der Availability des Klassenbesten. Indiskutabel schlecht schneiden der Rest der analysierten Suchwerkzeuge ab. 20% Availability bringen Acoon und Excite, 15% WebCrawler und Yahoo!, 10% Simpli und Spider sowie blamable 5% Speedfind, RealNames und Open Directory. (Etwas entschuldigt sind die Kataloge Yahoo! und Open Directory sowie die Homepage-Suche von RealNames, die durch die intellektuelle Erschließung nicht in der Lage sind, mit der "Massenproduktion" der Suchmaschinen mitzuhalten. Zudem hat Yahoo! durch den Zugriff auf eine Teildatenbank von Google eine Gesamt-Availability von 45%.)

Durch die jeweils suboptimalen Availa-

bility-Raten einzelner Suchwerkzeuge liegt es nahe, mehrere in Kombination abzufragen. Tabelle 8 zeigt Kombinationen von zwei bzw. drei unserer Top-Player. Mit Google und FAST sowie entweder AltaVista oder Northern Light erreichen wir eine Gesamt-Availability von 90%; und dies ist ein Wert, der sich durchaus sehen lassen kann.

Durch unsere unterschiedlichen Typen von Fragestellungen können wir die Ergebnisse nach Informationsart und Sprache auswerten. Um nicht mit zu kleinen Availability-Raten konfrontiert zu werden, beschränken wir die Analyse auf die ergiebigsten neun Suchwerkzeuge. Produktinformationen werden im Durchschnitt der neun Suchsysteme zu 43% exakt recherchiert, Literaturinformationen zu 49%. Für die einzelnen Systeme gibt es z. T. große Unterschiede in der Produktinformations- und Literatur-Availability. Die ersten drei Produkt-Finder sind Google (70%), FAST (60%) und Northern Light (50%), die besten Literatur-Finder sind AltaVista (80%), Fireball (70%), Google (60%) sowie Northern Light (auch 60%). Insgesamt sind deutschsprachige Informationen (durchschnittliche Availability 37%) schlechter zu finden als englischsprachige (56%). Während für die englische Sprache das Ergebnis für alle Top-Search Engines mit 60% gleich ist, sind die Differenzen für deutschsprachige Informationen groß: Google findet 70% (und damit mehr als in englisch), AltaVista 60% und Northern Light 50%.

Fragen wir noch genauer nach Kombinationen aus Sprache und Informationsart, so können wir gewisse Spezialisten erkennen. (Das Resultat sollte wegen

der kleinen Fallzahl nicht überbewertet werden.) Englische Literaturinformationen werden gut bei AltaVista, Fireball, Northern Light und Lycos gefunden (jeweils 80%), deutsche Literaturinformationen am besten bei AltaVista (80%). Bei deutschsprachigen Produktinformationen ist Google (80%) führend, bei englischsprachigen Produktinformationen wiederum Google (60%), aber auch gleichauf FAST und Oingo.

Ausblick

Sowohl Suchmaschinen als auch Web-Kataloge haben Schwierigkeiten, eine große Anzahl von bekannten URLs (Known Items) zu finden. Allerdings wartet unser Testsieger, Google, mit einer Availability-Rate von 65% mit einem Wert auf, der - bei analogen bibliothekswissenschaftlichen Erhebungen - nur bei guten Bibliotheken erreicht wird. Zumindest die **Top-Suchwerkzeuge** (und hierzu zählen wir nach den referierten Studien und unseren eigenen Resultaten **Google, AltaVista, Northern Light, FAST** und **Oingo** sowie - dank der zusätzlichen Google-Treffer - **Yahoo!**) haben einen durchaus brauchbaren Standard erreicht. Es gibt spezifische Vorteile; AltaVista scheint bei der Literatursuche kaum schlagbar zu sein, Google und FAST liefern bei Produktinformationen starke Ergebnismengen. Es liegt demnach nahe, Suchwerkzeuge in Kombination abzufragen. Diesen Aspekt haben auch die Betreiber der Suchsysteme bemerkt und setzen auf Kooperation. Wir erinnern hier nur an die Zusammenarbeit von Yahoo! und Google sowie an Oingos Kooperationen mit AltaVista und dem Open Directory Projekt. Bei den deutschen Suchwerkzeugen dominiert Fireball, das auch im internationalen Vergleich mithalten kann. Die anderen deutschen Ansätze erweisen sich (noch) als Nullnummern. ■

Mechtild Stock &
Wolfgang G. Stock

In Password 12/2000: Die Top- Suchmaschinen im Detail

Tabelle 9: Availability nach Informationsart und Sprache

Produkt	Literatur	deutsch	englisch
1 / Google / 70%	1 / AltaVista / 80%	1 / Google / 70%	1 / AltaVista / 60%
2 / FAST / 60%	2 / Fireball / 70%	2 / AltaVista / 60%	1 / FAST / 60%
3 / North.L. / 50%	3 / Google / 60%	3 / North. L. / 50%	1 / Fireball / 60%
4 / AltaVista / 40%	3 / North.L. / 60%	4 / FAST / 40%	1 / Google / 60%
4 / Oingo / 40%	5 / FAST / 40%	4 / Fireball / 40%	1 / North.L. / 60%
4 / Infoseek / 40%	5 / Oingo / 40%	6 / Oingo / 30%	1 / Lycos / 60%
4 / HotBot / 40%	5 / Lycos / 40%	7 / Infoseek / 20%	7 / Oingo / 50%
8 / Fireball / 30%	8 / Infoseek / 30%	7 / HotBot / 20%	7 / Infoseek / 50%
9 / Lycos / 20%	9 / Hotbot / 20%	9 / Lycos / 0%	9 / HotBot / 40%
Durchschnitt aller 9: 43%	Durchschnitt aller 9: 49%	Durchschnitt aller 9: 37%	Durchschnitt aller 9: 56%

(Rang / Suchmaschine / Trefferhäufigkeit)

GIN

ALLECO geht auf Plattform der Informationsbranche

Der Online-Anbieter ALLECO (= All Economy Online) in Dortmund hat sich auf GIN, der Plattform der Informationsbranche, begeben. ALLECO ist ein Dienst der Deutschen Telekom AG in Zusammenarbeit mit der ECOFIS Wirtschaftsinformationen GmbH. Der Online-Dienst zeichnet sich durch eine schnelle Verfügbarkeit aktueller Unternehmensinformationen aus. Er wurde in einer aktuellen vergleichenden Studie von IIE und Infratest über Firmeninformationsanbieter im Internet hoch bewertet.

Literatur

anonym: Test. Internet-Suchmaschinen. - In: test / Stiftung Warentest Nr. 8 (2000), S. 20-23.

anonym: Finden statt suchen. Schnüffler im Test. - In: PC-Online Nr. 10 (2000), S. 48-50.

Maryellen Mott Allen: Six degrees of affiliation. The small world of search tools. - In: Online 24 (2000), Nr. 5, S. 49-52.

Xiaoying Dong; Louise T. Su: Search engines on the World Wide Web and information retrieval from the Internet: A review and evaluation. - In: Online & CDROM Review 21 (1997), S. 67-82.

Susan Feldman: Find what I mean, not what I say. Meaning based search tools. - In: Online 24 (2000), Nr. 3, S. 49-56.

Michael Gordon; Praveen Pathak: Finding information on the World Wide Web: the retrieval effectiveness of search engines. - In: Information Processing & Management 35 (1999), S. 141-180.

Alfred Glossbrenner; Emily Glossbrenner: Search Engines for the World Wide Web. - Peachpit Press, 1999.

Randolph Hock: The Extreme Searcher's Guide to Web Search Engines. - Medford, NJ: CyberAge Books, 1999.

Paul B. Kantor: Availability analysis. - In: Journal of the American Society for Information Science 27 (1976), S. 311-319.

Steve Lawrence; C. Lee Giles: Searching the World Wide Web. - In: Science 280 (1998), Nr. 5360, S. 98-100.

Steve Lawrence; C. Lee Giles: Accessibility of information on the Web. - In: Nature 400 (1999), Nr. 6740, S. 107-109.

H. Vernon Leighton; Jaideep Srivastava: First 20 precision among World Wide Web search services (search engines). - In: Journal of the American Society for Information Science 50 (1999), S. 870-881.

Greg R. Notess: Search Engine Showdown. - URL: <http://www.searchengineshowdown.com/>

Andre Michael Spallek; Marcos Kreinacke: Suchmaschinen. Gezielt recherchieren im Internet. - München: Beck - DTV, 2000.

Wolfgang G. Stock: Informationswirtschaft. Management externen Wissen. - München; Wien: Oldenbourg, 2000.

Luise T. Su; Hsin-liang Chen; Xiaoying Dong: Evaluation of Web-based search engines from the end-user's perspective: A pilot study. - In: Proceedings of the 61st ASIS Annual Meeting; Vol. 35: Information Access in the Global Information Economy. - Medford, NJ: Information Today, 1998, S. 348-361.

Danny Sullivan: Search Engine Watch. - URL: <http://searchengine-watch.com/>

Start der Diplomarbeitsdatenbank und des News-Archivs - Zwölf Hochschulen in GIN eingebunden.

Unter der Rubrik "Suche" wurde die GIN-Diplomarbeitsdatenbank mit kostenlosen Nachweisen von Diplomarbeiten aus den Bereichen Information und Dokumentation verfügbar gemacht. Der Input erfolgt direkt über die Hochschulen. Ebenfalls neu ist die Datenbank zum GIN-Newsarchiv. Hier sind kostenfreie Recherchen im Bestand der GIN-News seit dem 1. Juni möglich. Die Datenbank wird wöchentlich aktualisiert.

Im Rahmen der GIN-Hochschulinitiative haben sich bereits zwölf Hochschulen verlinkt. Den Studierenden werden kostenfreie Recherchen in Infodok und in Kooperation mit der FH Potsdam kostenfreie Recherchen in Teilen von Infodata eingeräumt.

Erstmals wurde GIN als Medium zur Befragung von Studentinnen im Rahmen einer Diplomarbeit über GIN genutzt. Der Fragebogen in GIN wurde gut genutzt und führte zu interessanten Ergebnissen. Diese werden nach Abschluss der Diplomarbeit auch in GIN publiziert.