

Änderungsanalyse in den Suchergebnissen einer fachspezifischen Recherche mit den Suchmaschinen Google und Scirus

Vergleich der Jahre 2004 und 2005 für eine Fragestellung aus dem Bereich der Biotechnologie

M. Luisa Doldi, Wien

Das Ziel dieser Arbeit war die Identifizierung und Quantifizierung der Veränderungen in den Suchergebnissen aus ausgewählten Suchmaschinen im Jahresabstand.

Eine Recherche mit der gleichen Suchabfrage ist mit der allgemeinen Suchmaschine Google und mit der wissenschaftlichen Suchmaschine Scirus im August 2004 und April 2005 durchgeführt worden. Die ersten 40 Treffer sind nach vordefinierten qualitativen und quantitativen Parametern analysiert worden. Weiters ist die Überlappung der Suchergebnisse für die oben genannten Abfragezeitpunkte je nach Suchmaschine getrennt bewertet worden.

Die aus dieser Studie resultierenden Ergebnisse zeigen einerseits die Existenz von positiven Tendenzen in den Ergebnislisten der untersuchten Suchmaschinen (Verbesserungen in der Relevanz der Treffer, eine Verringerung der Anzahl an nicht abrufbaren Links, sowie ein Zuwachs an erzielter Information) und andererseits die Rolle der Websuchinstrumente selbst hinsichtlich den Veränderungen von Suchergebnissen bei einer Recherche im Web.

Aim of this work was the analysis of the changes occurring in hit lists of selected search engines at two different points of time. The same search query was applied to the search engine Google and to the science search engine Scirus in August 2004 and in April 2005. The first 40 hits of each search have been analysed on the basis of predefined qualitative and quantitative parameters. Further an analysis of overlapping between search hits at the above mentioned time points has been carried out for each search engine separately.

The results of this study has shown on one hand the existence of positive trends in the result lists of both search engines (less dead links, a higher precision in results) and on the other hand the role played by search engines themselves in the dynamic of result lists. The method applied here has shown to be useful for describing and measuring trends and changes in the search results of search engines over a time span.

Einleitung und Problemstellung: die dynamische Natur der Information im Web

Eine der markantesten Eigenschaften des Webs ist seine dynamische Natur. Das Web wächst ständig und ändert sich schnell (Lawrence, 1999). Dokumente tauchen auf und verschwinden wieder mit einer Geschwindigkeit, die kein anderes Medium ermöglicht.

Man muss aber eines nicht vergessen: man schaut auf dem Web immer durch das vom jeweiligen Suchinstrument angebotene „Fenster“. Und was wäre dann, wenn das Instrument selber zumindest für einen Teil dieser Dynamik verantwortlich wäre, die wir dem Web zuschreiben? Anders gesagt: findet man ein Dokument nicht mehr, weil es wirklich vom Web entfernt wurde oder weil die verwendete Suchmaschine, die inzwischen zum Beispiel ihre Suchalgorithmen geändert hat, das Dokument nicht mehr finden kann?

In mehreren Studien sind die Suchergebnisse von Suchmaschinen analysiert worden. Gordon (1999) und Dresel

(2001) haben Vergleiche zwischen Ergebnissen aus mehreren Suchmaschinen für die gleiche Suchanfrage durchgeführt. Lo Grasso (2005) berichtet über Ergebnisse aus derselben Suchmaschine für unterschiedliche Suchanfragen. Es sind allerdings kaum Studien durchgeführt worden, um über eine längere Zeitspanne die Suchergebnisse von denselben Suchanfragen aus der gleichen Suchmaschine zu beobachten. Eine solche Analyse würde es ermöglichen, Trends und Tendenzen in der Entwicklung einer Suchmaschine aus Nutzer Sicht zu identifizieren und Veränderungen in der Wiedergabe von Suchergebnissen zu beobachten und quantitativ zu dokumentieren.

In diesem Kontext stellt sich diese Arbeit das Ziel, eine Analyse der qualitativen und quantitativen Änderungen in den Suchergebnissen aus ausgewählten Suchmaschinen für die gleiche Suchabfrage, zu zwei unterschiedlichen Zeitpunkten durchzuführen.

Diese Studie baut auf einer Untersuchung des Jahres 2004 auf (Doldi et al. 2005), wo die allgemeine Suchmaschine Google und die auf Wissenschaft spezialisierte Suchmaschine Scirus für eine ganz spezifische Abfrage in dem wissenschaftlichen Bereich der Biotechnologie in der Pflanzenproduktion verwendet worden ist. Die sorgfältige Analyse und Dokumentation der damaligen Ergebnisse hat es ermöglicht im Jahr 2005 die gleiche Recherche wieder durchzuführen und die Ergebnisse zu vergleichen. Obwohl der hier beispielhaft dargestellte Informationsbedarf zu einem sehr spezifischen wissenschaftlichen Bereich gehört, haben die Erkenntnisse aus dieser Studie allgemeine Gültigkeit indem:

- die Aussagen dieser Studien über qualitative und quantitative Veränderungen der Suchergebnisse vom untersuchten Bereich unabhängig sind;
- die in dieser Studie beschriebene Methode als Benchmark für die Erhebung und Analyse von Trends in der Entwicklung einer Suchmaschine aus Nutzer Sicht gesehen werden kann.

Es soll hier betont werden, dass diese Arbeit keinen Vergleich zwischen Suchmaschinen anstrebt, sondern einen Vergleich zwischen Suchergebnissen aus der gleichen Suchmaschine zu zwei unterschiedlichen Zeitpunkten.

Für diese Studie sind die allgemeine Suchmaschine Google (www.google.com) und die auf Wissenschaft spezialisierte Suchmaschine Scirus (www.scirus.com) ausgewählt worden. Die erste gilt als repräsentatives Instrument der Benutzergewohnheiten (Machill and Welp, 2003), die zweite -eine der wenigen auf Wissenschaft spezialisierten Suchmaschinen- eignet sich gut für den Bereich des Informationsbedarfes. Für detaillierte Informationen über die ausgewählten Suchmaschinen wird zu den jeweiligen Webseiten verwiesen.

Methode

Zwei Arbeitsphasen charakterisieren diese Studie: die Phase der Recherche und die Phase des Vergleichens der Ergebnisse.

Recherche

Die Recherche Phase bestand aus drei Arbeitsschritten:

- * Vorbereitung der Recherche
- * Durchführung der Recherche
- * Analyse der Suchergebnisse

Bei der Vorbereitung der Recherche ist der Informationsbedarf identifiziert und analysiert worden. Für diese Studie ist ein konkreter Informationsbedarf aus dem Bereich „Biotechnologie in der Pflanzenzüchtung“ genommen worden, und zwar:

„Welche quantitativen Merkmale der Pflanze Sojabohne (*Glycine max L.*) sind im Zeitraum Januar 2000- Dezember 2003 durch molekulare Marker analysiert worden. Hinweise aus der Literatur“.

Die Fragestellung ist sowohl einer inhaltlichen als auch einer syntaktischen und semantischen Analyse unterzogen worden. Als Ergebnis dieser Analysen ist der Informationsbedarf klar definiert und abgegrenzt und eine Suchabfragekette definiert worden. Die daraus resultierende Suchabfragekette

(soybean OR soyabean OR „Glycine max“) AND
 (marker OR markers)
 AND
 (qtl OR qtls OR „quantitative trait“ OR „quantitative traits“ OR „quantitative locus“ OR „quantitative loci“).
 wurde nach Anpassung an die Suchfunktionen und Suchmöglichkeiten des jeweiligen Suchinstrumentes als Basis für die Recherche verwendet.

In beiden Suchinstrumenten wurde der erweiterte Suchmodus verwendet und der Zeitbereich zwischen 2000 und 2003 als Einschränkungsfaktor gesetzt. Wir sind uns dessen bewusst, dass diese Einschränkung für die Suchmaschine nicht unbedingt das Erscheinungsdatum der erzielten Dokumenten ist, sondern entweder das Datum der Aufnahme des Dokumentes bzw. der Webseite in den Suchmaschinenindex oder das Datum der letzten Aktualisierung der jeweiligen Webseite. Diese Tatsache bringt notwendigerweise eine gewisse Unschärfe der Suchergebnisse mit sich, mit der man aber leben muss, da unserem Wissen nach (noch) keine Suchmaschine die Fähigkeit hat, das Erschei-

nungsdatum der in den indexierten Webseiten enthaltenen Dokumenten zu erschließen.

Für die Durchführung der Suche in Google ist die Oberfläche von Faganfinder (Bates, 2004) verwendet worden, wo alle Optionen der erweiterten Suche auf einer einzigen, benutzerfreundlichen Bedienoberfläche dargestellt sind, und wo auch die Möglichkeit besteht, eine genaue Zeiteinschränkung zu setzen. Faganfinder ist im Web freiverfügbar unter <http://www.faganfinder.com/google.html>.

Die Recherche ist im August 2004 und im April 2005 durchgeführt worden. Die gleiche Suchabfragekette und die gleichen Einstellungen sind zu beiden Zeitpunkten verwendet worden. Die ersten 40 Treffer sind bei der jeweiligen Suche analysiert worden (Corrao, 2004).

Vergleich der Suchergebnisse

Eine getrennte Analyse der Ergebnisse ist für jede Suchmaschine durchgeführt worden. Die Ergebnisse der jeweiligen Suchmaschine aus dem Jahr 2004 und aus dem Jahr 2005 sind verglichen worden, um jegliche Unterschiede bzw. Überlappungen zu identifizieren. Für den Vergleich sind folgende Parameter betrachtet worden:

- * Quantitative Parameter:
 - o Anzahl der gesamten Treffer
 - o Anzahl der inaktiven Links
 - o Anzahl der für den Informationsbedarf relevanten Treffer bzw. Dokumente (Literaturhinweise)
 - o Anzahl der redundanten Seiten bzw. Dokumente
- * Qualitative Parameter:

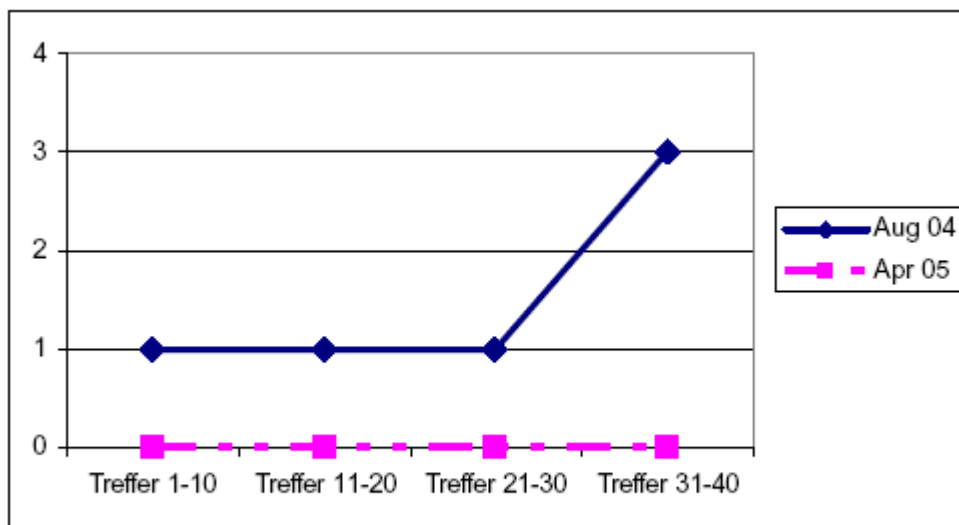


Abbildung 1: Anzahl der inaktiven Links und deren Verteilung im Jahr 2004 und 2005 aus der Suchmaschine Google

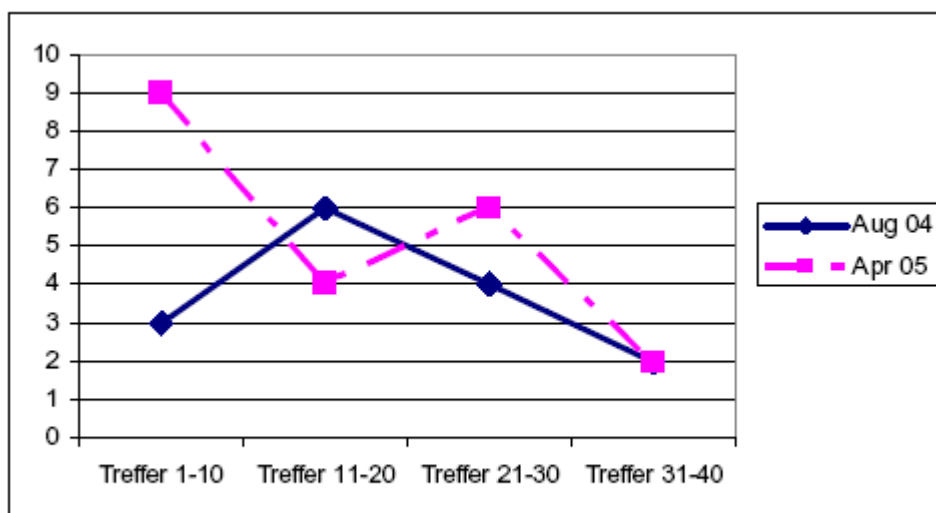


Abbildung 2: Verteilung der für den Informationsbedarf relevanten Treffer in Google aus den Jahren 2004 und 2005

- o Relevanz der erzielten Seiten bzw. Dokumente für den Informationsbedarf

- o Seriosität und Autorität der erzielten relevanten Dokumente
- o Analyse der relevanten Dokumente hinsichtlich geographischen Ursprung, Domain-Kategorie, Sprache, Art

* Analyse der Überlappungen zwischen Ergebnissen aus dem Jahr 2004 und aus dem Jahr 2005 aus den jeweiligen Suchmaschinen.

Aus dem Vergleich zwischen Ergebnissen aus dem Jahr 2004 und aus dem Jahr 2005 sind Indikationen von Tendenzen für die jeweilige Suchmaschine abgeleitet worden.

Ergebnisse

Google: Vergleich der Suchergebnisse aus dem Jahr 2004 und aus dem Jahr 2005

Quantitative Parameter

Die gleiche Suchabfrage führte im Jahr 2004 zu ~17900 Treffern während im Jahr 2005, nachdem Google seinen Index verdoppelt hatte, zu ~895 Treffern. Eine Analyse der ersten 40 Treffer führte im Jahr 2004 zu 6 inaktiven Links, während im Jahr 2005 die Anzahl der inaktiven Links zwischen den ersten 40 Treffern null war. (Abbildung 1).

Die für den Informationsbedarf relevanten Treffer waren im Jahr 2004 insgesamt 15, was eine Relevanzrate von 37,5% ergibt. Im Jahr 2005 lag die Relevanzrate bei 52,5% mit 21 relevanten Treffern aus 40. Abbildung 2 zeigt die Verteilung der relevanten Treffern zwischen

den Ergebnissen sowohl aus dem Jahr 2004 als auch aus dem Jahr 2005

Die größte Veränderung betrifft die ersten 10 Treffer, die auch die von den Benutzern meist analysierten Treffer sind: während im Jahr 2004 hier nur drei relevanten Treffer zu finden waren, stiegen diese bis auf 9 im Jahr 2005.

Nur eine URL Adresse wurde im Jahr 2004 zweimal abgerufen, während im Jahr 2005 5 URL zweimal bzw. dreimal abgerufen worden sind. Allerdings handelte es sich hier immer um unterschiedliche Seiten der gleichen Domäne, die auch unterschiedliche Dokumente geliefert haben. Somit ist die resultierende Redundanz der erzielten Dokumente für das Jahr 2005 bei null.

Was die erzielten Dokumente bzw. Literaturhinweise betrifft, welche das Objekt des Informationsbedarfes waren, sind folgende Klassen von Dokumenten erzielt worden:

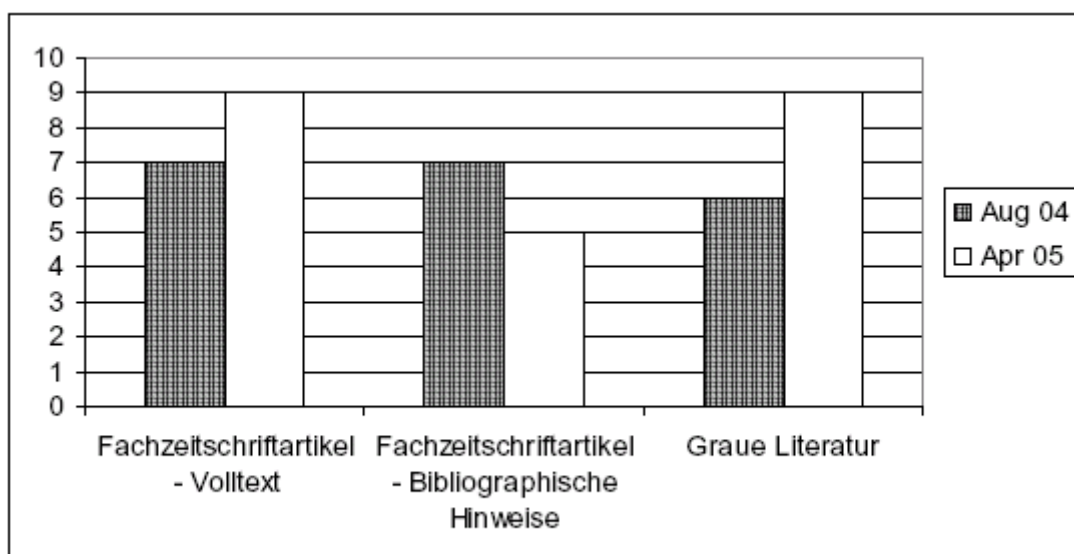


Abbildung 3: Typologie und Anzahl der Literaturdokumente erzielt mit der Suchmaschine Google

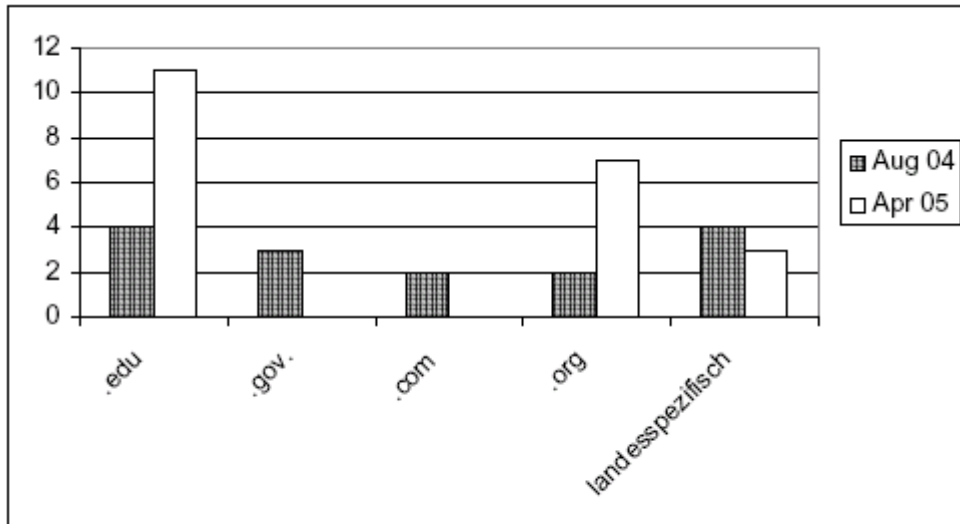


Abbildung 4: Kategorien der Domain Extensionen der erzielten Treffer und deren Anzahl in beiden untersuchten Jahren

- * Volltext Fachzeitschriftartikel
- * Hinweise für Fachzeitschriftartikel (Autor, Quelle, Erscheinungsjahr)
- * Graue Literatur: technische Berichte, Projektberichte, Konferenzliteratur.

Im Jahr 2004 wurden insgesamt 20 Dokumente und im Jahr 2005 insgesamt 23 Dokumente erzielt, deren Typologie und Anzahl in Abbildung 3 dargestellt sind.

Qualitative Parameter

Eine Analyse der relevanten Seiten ist durchgeführt worden hinsichtlich Domäne-Extension, geographischer Ursprung und Sprache. Insgesamt sind in beiden Jahren 5 Kategorien von Domain-Extensionen erzielt worden: .org, .edu, .gov, .com und landesspezifische Extensionen (Abbildung 4). Nachfol-

gend ist eine Auflistung jener Institutionen angeführt, die in dieser Studie hinter den Domain-Extensionen zu finden sind:

- * „.edu“: Universitätswebseiten
- * „.gov“: staatliche Institutionen wie das amerikanische Ministerium für Landwirtschaft (USDA) oder Bibliotheken (National Library of Medicine)
- * „.org“: Fachvereine und nicht gewinnorientierte Organisationen
- * „.com“: gewinnorientierte Verlagshäuser
- „.land“: in den hier erzielten Treffern handelte es sich um Universitäten oder Vereine.

Alle relevanten Webseiten stammen sowohl im Jahr 2005 als auch im Jahr 2004 aus den Vereinigten Staaten und sind dementsprechend auf Englisch.

Auch die Autorität der erzielten Webseiten ist untersucht worden, da diese ein wichtiger Parameter ist, um über die Qualität eines Treffers zu entscheiden. Es sind jene Seiten als „autoritativ“ und „seriös“ bezeichnet worden, wo:

1. ein Verantwortlicher für die Inhalte (Autor, Editor) eindeutig, klar und einfach zu erkennen war;
2. die Ziele der Webseite klar zu identifizieren waren;
3. die Dokumente ein Erscheinungsdatum, Autornamen und Quellenhinweis trugen;
4. Kontaktmöglichkeiten mit dem Autor bzw. Editor oder Webmaster möglich waren.

Alle erzielten relevanten Webseiten haben sowohl im Jahr 2004 als auch im Jahr 2005 diese Parameter erfüllt und sind somit als „autoritativ und seriös“

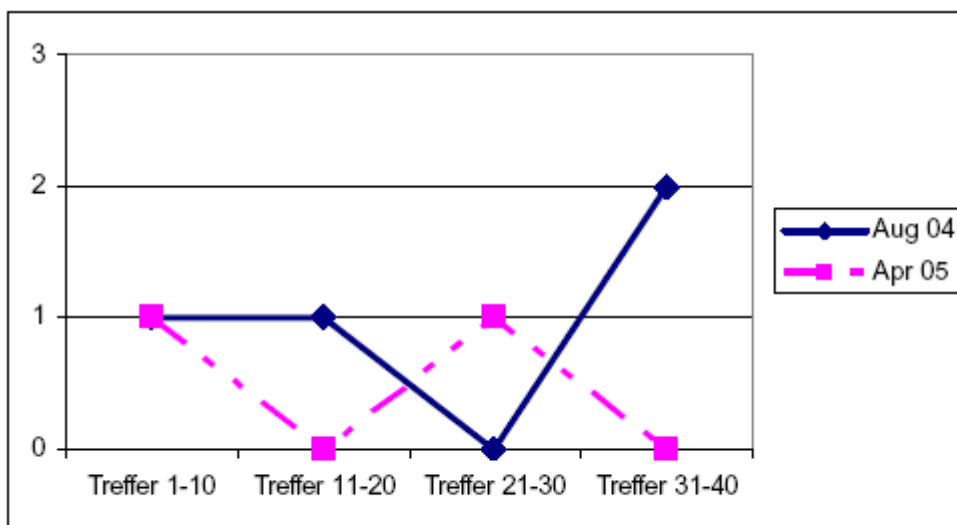


Abbildung 5: Anzahl der inaktiven Links in Scirus und deren Verteilung in den zwei untersuchten Jahren

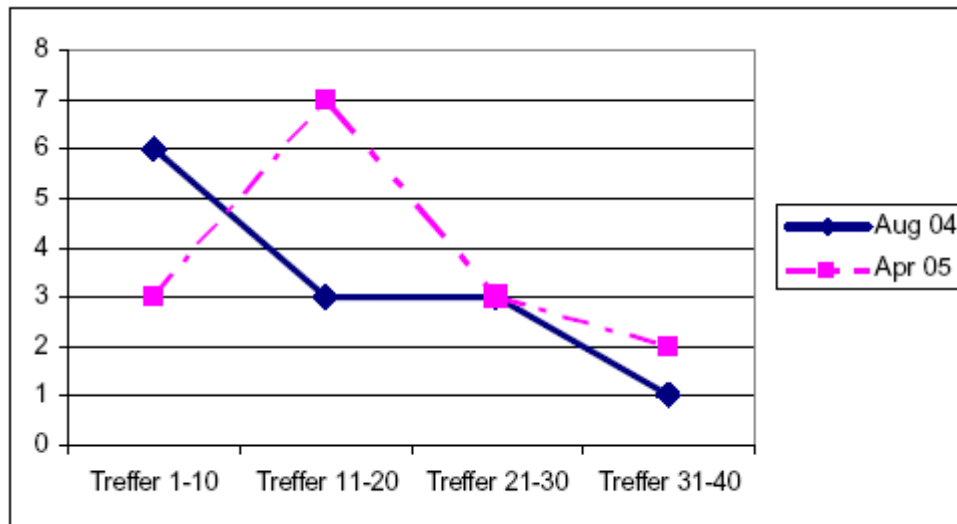


Abbildung 6: Verteilung der für den Informationsbedarf relevanten Treffer in Scirus aus den Jahren 2004 und 2005

eingestuft worden. Der Parameter „Aktualität“ einer Webseite, der auch sehr wichtig ist, um deren Zuverlässigkeit zu beurteilen, spielt hier auf Grund der zeitlichen Einschränkungen des Informationsbedarfes keine Rolle. Nicht relevante Seiten wurden für die Evaluierung der Ergebnisse nicht in Betracht gezogen. Allerdings ist auch eine Analyse dieser Seiten durchgeführt worden, um den Grund deren Irrelevanz zu identifizieren. Diese Analyse hat ergeben, dass es sich hier meistens um Seiten handelte, die mit den zeitlichen Bedingungen des Informationsbedarfes nicht gepasst haben.

Überlappung zwischen den Treffern aus dem Jahr 2004 und aus dem Jahr 2005

Eine Analyse der Überlappung zwischen den Suchergebnissen aus den zwei Jahrgängen ist durchgeführt worden. Folgende Erkenntnisse sind draus resultiert:

* insgesamt sind sechs relevante Webseiten sowohl im Jahr 2004 als auch im Jahr 2005 abgerufen worden. Diese Seiten haben auch die gleichen Dokumente in beiden Recherchen geliefert. Allerdings sind diese Seite in den untersuchten Jahre anders gewichtet worden: im Jahr 2005 zwischen den ersten 10 Treffern und im Jahr 2004 ab den 17. Treffer. Die beobachtete Überlappungsrate beträgt 14,5%.

* Durch die Recherche des Jahres 2005 sind 15 relevante neue Seiten im Ver-

gleich mit dem Jahr 2004 erzielt worden. Allerdings sind 9 relevante Seiten aus dem Jahr 2004 im Jahr 2005 nicht mehr zwischen den ersten 40 Treffern zu finden und somit für die Ergebnisse verloren. Eine Analyse dieser 9 relevanten Seiten aber zeigt, dass sie auch im Jahr 2005 noch im Web verfügbar sind.

Scirus: Vergleich der Suchergebnisse aus dem Jahr 2004 und aus dem Jahr 2005

Quantitative Parameter

Die gleiche Suchabfrage führte im Jahr 2004 zu 75 Treffern während im Jahr 2005 86 Treffer erzielt wurden. Eine Analyse der ersten 40 Treffer führte im Jahr 2004 zu 4 und im Jahr 2005 zu 2 inaktiven Links (Abbildung 5).

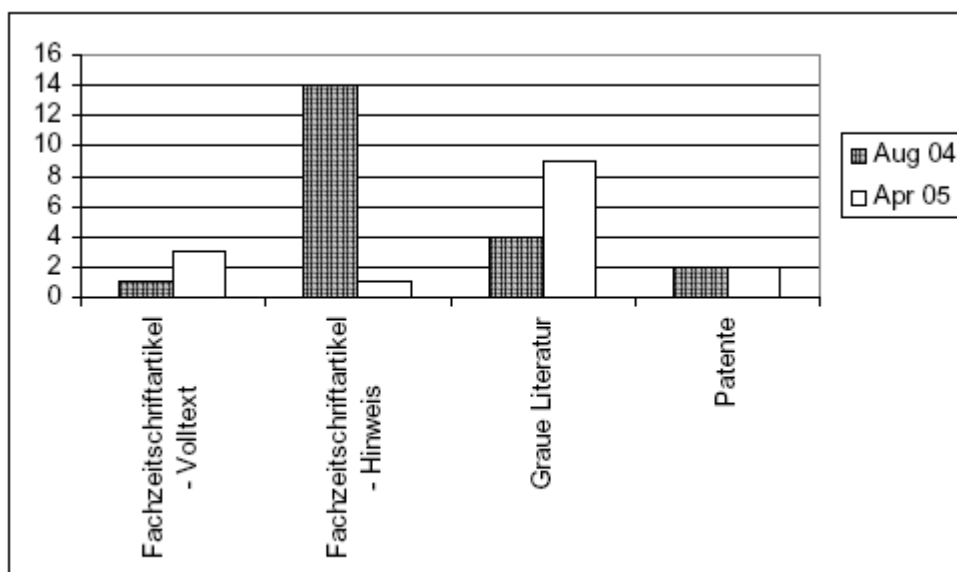


Abbildung 7: Typologie und Anzahl der Literaturdokumente erzielt mit Scirus

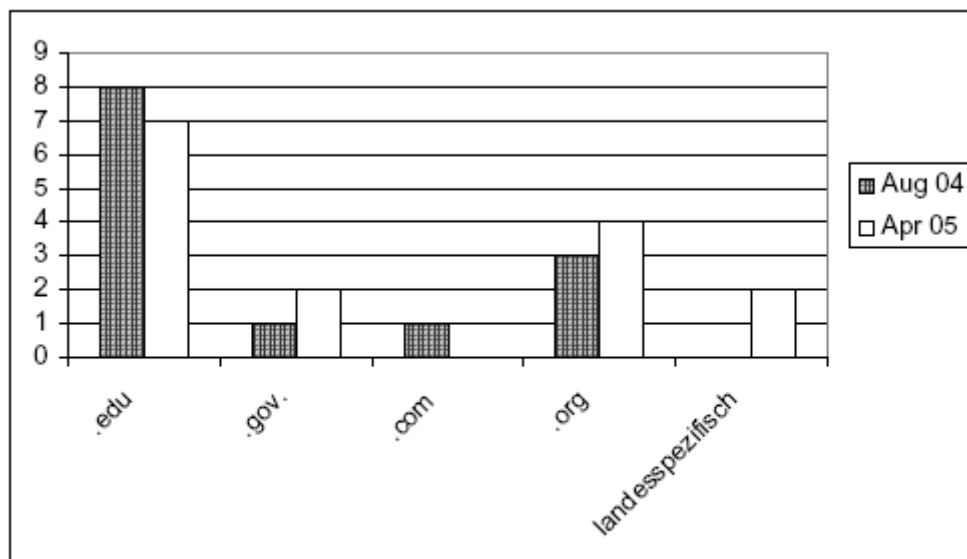


Abbildung 8: Anzahl und Verteilung der Domain Extensionen der relevanten Seiten in den untersuchten Jahren

Die Relevanzrate der Treffer betrug im Jahr 2004 32,5% und im Jahr 2005 37,5%.

Wie in Abbildung 6 dargestellt, hat sich die Verteilung der relevanten Treffer in der Trefferliste geändert. Der Großteil von ihnen wurde von der ersten Gruppe (Treffer 1-10) im Jahr 2004 in die zweite Gruppe (Treffer 11-20) im Jahr 2005 gereiht.

Sowohl im Jahr 2004 als auch im Jahr 2005 sind keine redundanten Seiten zwischen den ersten 40 Treffern gefunden worden.

Die erzielten Dokumente unterteilen sich in folgenden Kategorien:

- * Volltext Fachzeitschriftartikel
- * Hinweise für Fachzeitschriftartikel
- * Graue Literatur: technische Berichte, Projektberichte, Konferenzliteratur
- * Patente in Volltext

Abbildung 7 zeigt die Verteilung dieser Kategorien zu den untersuchten Zeitpunkten

Qualitative Parameter

Die gleichen Domain-Extensionen wie in Google sind auch in Scirus erzielt worden mit einer Verteilung, die in Abbildung 8 dargestellt ist.

Was den geographischen Ursprung der erzielten Seiten betrifft, stammen im Jahr 2004 12 Seiten aus den Vereinigten Staaten und eine aus England; im Jahr 2005 stammen 13 Seiten aus den Vereinigten Staaten, eine aus Brasilien und eine aus Singapur. Nur eine Seite (aus Brasilien) ist auf Portugiesisch, alle anderen sind auf Englisch.

Auch in Scirus sind die erzielten Webseiten bzw. Dokumente einer Seriositätsprüfung unterzogen worden. Es sind die gleichen Parameter wie in Google analysiert worden. Alle erzielten relevanten Seiten haben diese Prüfung bestanden. Eine kurze Analyse der nicht relevanten Seiten ist unternommen worden, um einen Indikator über die Gründe der Unschärfe der Recherche zu bekommen. Zwei Kategorien von irrelevanten Treffern sind gefunden worden:

- * Treffer, die zeitlich mit dem Informationsbedarf nicht passen, da sie Dokumente liefern, die entweder vor dem Jahr 2000 oder nach dem Jahr 2003 erschienen sind. Diese Treffer befinden sich großteils in der Gruppe 1-10 und 11-20;
- * Treffer, die das Thema verfehlen. Diese Treffer befinden sich großteils im Bereich 21-40.

Überlappung zwischen den Treffern aus dem Jahr 2004 und aus dem Jahr 2005

Eine Überlappungsrate von 40% ist zwischen den Ergebnissen aus dem Jahr 2004 und den Ergebnisse aus dem Jahr 2005 beobachtet worden. Die wiederholt erzielten Treffer besetzten in den untersuchten Jahren ähnliche Stellen. Was die anderen Treffer betrifft, sind 5 relevante Treffer aus dem Jahr 2004 im Jahr 2005 nicht mehr zu finden, obwohl sie sich noch im Web befinden.

Diskussion und Schlussfolgerungen

Der Vergleich der Suchergebnisse aus den untersuchten Jahren zeigt eine un-

terschiedliche Situation zwischen den zwei ausgewählten Suchmaschinen.

Die größten Veränderungen zeigt die Suchmaschine Google, die eine eindeutige Verbesserung im Jahr 2005 aufweisen kann:

- * die Anzahl der nicht abrufbaren Links ist im Jahr 2005 deutlich niedriger als im Jahr davor
- * die gesamte Relevanzrate ist eindeutig gestiegen
- * die Anzahl der relevanten Treffer zwischen den ersten 10 ist eindeutig gestiegen.

Viele relevante Treffer, die im Jahr 2004 nicht zu finden waren, sind im Jahr 2005 dazu gekommen. Im Jahr 2005 sind aber auch einige relevante Treffer aus dem Jahr 2004 verloren gegangen. Im Ganzen lässt sich ein Gewinn an Information registrieren.

Eine Überraschung liefert die gesamte Anzahl der erzielten Ergebnisse: während im August 2004, vor der Index-Expansion, ~17900 Treffer erzielt wurden, so reduzierten sich diese im Jahr 2005, nach der Index-Expansion, auf ~895. Die Größe des Indexes ist also vermutlich nicht die einzige Veränderung, die von Google unternommen wurde. Es lassen sich auch Änderungen anderer Natur hinter den Kulissen erkennen, die auf jeden Fall in Richtung einer Verbesserung der Trefferliste gehen und nicht, wie man oft hört, in Richtung dessen Verschlechterung.

Die wissenschaftliche Suchmaschine Scirus zeigt prinzipiell eine gewisse Stabilität in der Wiedergabe der Ergebnisse. Allerdings sind auch einige Verände-

rungen zwischen dem Jahr 2004 und dem Jahr 2005 zu beobachten:

- * die Anzahl der nicht abrufbare Links hat sich von 4 auf 2 reduziert
- * die Relevanzrate ist von 32,5% auf 37% gestiegen
- * die relevanten Treffer befinden sich größtenteils zwischen den ersten 20 Treffern und nicht mehr zwischen den ersten 10.

Die beobachtete Überlappungsrate zwischen den Suchergebnissen aus der gleichen Suchmaschine ist überraschend niedrig, insbesondere für die Suchmaschine Google, die nur 14,5% aufweist. Die Tatsache, dass sich die Dokumente, die beide Suchmaschinen im Jahr 2005 nicht mehr finden, noch im Web befinden, bringt Bewusstsein über eine zu wenig angesprochene Tatsache: die Dynamik der Information ist nicht gänzlich der Kurzlebigkeit der Dokumente im Web zuzuschreiben sondern hängt auch maßgeblich von den Suchmaschinen ab, die durch Änderungen ihrer Indexes und ihrer Suchalgorithmen einen wesentlichen Anteil an der zeitlich begrenzten Informationsverfügbarkeit tragen.

Dagegen kann der Benutzer einige Strategien einsetzen, wie zum Beispiel die Recherche in mehreren Suchmaschinen und - wenn möglich - wiederholt durchzuführen. Als Web- und Suchmaschinenbenutzer ist es auf jeden Fall sehr wichtig, sich dieser Tatsache bewusst zu sein, insbesondere dort wo die Rechercheergebnisse unmittelbaren Einfluss auf die Praxis haben oder wo Themen behandelt werden, die ein historisches oder politisches Gewicht haben können.

Die qualitative Analyse der Suchergebnisse hat die Herrschaft der englischen Sprache im Web bestätigt und hat eine eindeutige Dominanz von Seiten aus verschiedenen Organisationen und In-

stitutionen der Vereinigten Staaten gezeigt. Diese bedienen sich dem Web eindeutig mehr (und besser) als Institutionen anderer Länder (einschließlich EU) um ihre Aktivitäten und sich selber zu präsentieren und das von ihnen erworbene Wissen zu verbreiten.

Die in dieser Studie angewandte Methodik hat sich als geeignet gezeigt, um Tendenzen und Änderungen in den Suchergebnissen einer Suchmaschine über eine längere Zeit zu identifizieren und zu quantifizieren. Die ausgewählten quantitativen und qualitativen Parameter, nach denen die Analyse durchgeführt worden ist, ermöglichen eine eindeutige Darstellung dieser Tendenzen.

Die aus dieser Studie resultierenden Ergebnisse haben keinen Anspruch auf allgemeine Gültigkeit. Viel mehr sollte man, aufbauend auf dieser Studie, eine längere Zeitspanne regelmäßig im Jahresabstand beobachten. Weiters bleiben einige Fragen offen, wie zum Beispiel aus welchen Gründen Treffer aus der Ergebnisliste einer Suchmaschine verschwinden, obwohl sie sich noch im Web befinden. Eine tiefgehendere Analyse als in dieser Studie durchgeführt ist für die Beantwortung dieser Fragen notwendig. Allerdings haben es diese Ergebnisse ermöglicht, eine Methodik zu erproben, um die Existenz von Tendenzen in den Ergebnislisten von Suchmaschinen zu beleuchten und diese zu quantifizieren und damit das Bewusstsein der Nutzer von Websuchinstrumenten hinsichtlich des begrenzten Informationsgehaltes einer Trefferliste zu sensibilisieren.

Literaturhinweise

Bates M.E. (2004): „Hidden Google Tools“ <http://searchenginewatch.com/searchday/article.php/3304771> (Stand: Mai 2005)

Corrao S, Leone F and S Arnone (2004): „Informazione biomedica su internet e motori di ricerca. Risultati di un trial di un anno“. *Recenti progressi di medicina*, Vol 95, 1 Gennaio 2004.

Doldi, M.L. and E. Bratengeyer (2005): „The Web as a source for scientific information: a comparison with fee-based databases“ *Online Information Review* - in press.

Dresel, R., Hörnig, D. Kaluza, H. Petr, A., Roßmann, N. and Sieber W., (2001): „Evaluation deutscher Web-Suchwerkzeuge. Ein vergleichender Retrievaltest“. *Information, Wissenschaft und Praxis* Vol 52, pp. 381-392.

Gordon, M., Pathak, P., (1999): „Finding information in the World Wide Web: the retrieval effectiveness of search engines“. *Information, Processing and Management* Vol 35 pp. 141-180

Lawrence, S. and Giles, C.L. (1999), „Accessibility of information on the Web“, *Nature* Vol. 400 No. 8, pp. 107-109

Lo Grasso, L. and Wahlig, H. (2005), „Google und seine Suchparameter. Eine Top 20 Precision Analyse anhand repräsentativ ausgewählter Anfrage“, *Information Wissenschaft und Praxis* Vol 56, No 2, pp. 77-86

Machill, M. and Welp, C. (Eds) (2003), „Wegweiser im Netz“, Verlag Bertelsmann Stiftung, Gütersloh.

Dr. M. Luisa Doldi
 Gloriettegasse 45/1/9
 A-1130 Wien
 Tel/Fax: + 43 -(0) 1 877 5762
 E-Mail: mldoldi@mldoldi.com