Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Fakultät Design Medien und Information

Bachelorthesis

Die Qualität von Retrievalergebnissen im Vergleich - Google Book Search und der OPAC einer wissenschaftlichen Bibliothek

vorgelegt von

Balder Thomsen

Erstbetreuer: Prof. Dr. Lewandowski

Zweitbetreuerin: Prof. Dr. Spree

Abstract

Die Arbeit vergleicht die Qualität von Retrievalergebnissen des OPACs der SUB Hamburg mit der Qualität von Ergebnissen der Google Buchsuche. Der Schwerpunkt liegt dabei auf Relevanzbewertungen von Ergebnissen 50 thematischer Suchanfragen. Stellvertretend für die Ergebnisse werden deren (anonymisierte) Beschreibungen bewertet. Die Bewertung wird von zehn Studenten vorgenommen, die auf zwei Skalen urteilen, einer binären und einer Skala mit 7 Ausprägungen. Die Suchanfragen, die an die Systeme gestellt wurden, wurden aus dem Logfile eines OPAC entnommen. Sie bilden die in der Realität auftretenden Worthäufigkeiten bei OPAC-Suchanfragen ab. Die Ergebnisse zeigen, dass die Ergebnisse des OPACs als relevanter bewertet werden. Der OPAC produziert bei einer größeren Anzahl an Suchanfragen die besseren Ergebnisse. Ergänzend wurden 70 Known-Item-Suchanfragen durchgeführt. Hier erzielt Google Books die besseren Ergebnisse: Die Availability ist höher und das Ranking besser als im OPAC. In der Arbeit wird detailliert auf die Struktur der Retrievalergebnisse eingegangen: Alter, Zugänglichkeit der Ergebnisse und Verfügbarkeit von Kataloganreicherungen werden dargestellt. Bei Ergebnissen der Buchsuche beträgt das Durchschnittsalter 24 Jahre und ist doppelt so hoch wie bei Ergebnissen des OPACs. Das Alter der Ergebnisse bei Google unterscheidet sich stark, abhängig davon in welchem Ansichtsmodus sie verfügbar sind. Kataloganreicherungen sind bei Google deutlich häufiger.

Schlagwörter

Google Books, Google Book Search, Google Buchsuche, OPAC, Information Retrieval, Known-Item-Search, Subject-Search, Volltext, Catalogue Enrichment

In der Arbeit wurde darauf verzichtet, englische Begriffe, die mittlerweile in den deutschen Sprachgebrauch übergegangen sind, als Fremdwörter hervorzuheben. Eigennamen von Firmen und Produkten wurden ebenfalls nicht hervorgehoben.

In der Arbeit wird stets die männliche Form verwendet, dies soll nicht diskriminieren, sondern zu besserer Lesbarkeit führen.

Inhalt

1. Einleitung 1
2. Die untersuchten Angebote 6
2.1. Google Book Search 6
2.1.1. Interface und Benutzung 9
2.1.1.1. Die Suche in Google Books 9
2.1.1.2. Ergebnisspräsentation 11
2.1.1.3. Darstellung eines Buches 13
2.1.1.4. Funktionen zur Arbeit mit Buchtext 17
2.1.1.5. Persönliche Bibliothek 18
2.1.1.6. Offene Schnittstellen 20
2.1.2. Rechtliche Situation 21
2.1.2.1. Bis zur Beilegung des Rechtsstreits mit der Authors Guild 21
2.1.2.2. Rechtliche Situation nach der Einigung mit der Authors Guild 24
2.2. OPAC der Firma OCLC 26
2.2.1. Einfache Suche 27
2.2.2. Erweiterte Suche 28
2.2.3. Suchergebnisse 29
2.3. Catalogue Enrichment 30
3. Ziele der Untersuchung 33
4. Methodik 36
4.1. Auswahl der Bibliothek 37
4.2. Auswahl des OPACs 38
4.3. Entscheidung für eine Bibliothek 39
4.4. Auswahl der Suchanfragen 39
4.4.1. Literatur zum Nutzerverhalten im OPAC 40
4.4.2. Auswahl der Suchanfragen aus dem Logfile 42
4.4.2.1 Worthäufigkeiten 42
4.4.2.2. Thematische Suchanfragen 43
4.4.2.3. Known-Item-Suchanfragen 44
4.4.2.4 Verteilung der Suchbegriffe auf Sprachen und Themen 44
4.5. Datenerfassung und Anonymisierung bei der thematischen Suche 45
4.6. Datenerfassung bei der Known-Item-Search 47
4.7. Relevanzbewertung 47

4.7.1 Die verwendeten Skalen 48
5. Ergebnisse 50
5.1. Subject Search 50
5.1.1. Micro Precision 50
5.1.2. Makro-Precision 51
5.1.3. Sonstige Darstellungen 53
5.2. Vergleich der Ergebnisse der thematischen Suche mit Ergebnissen einer anderen Studie 54
5.3. Known-Item-Search 56
5.4. Arten der Zugänglichkeit 57
5.5. Kataloganreicherungen 59
5.5.1. Rezensionen 61
5.5.2. Inhaltsverzeichnisse 63
5.6. Alter der Treffer 66
6. Fazit 69
Literaturverzeichnis 71
Anlagen ..... A

Abbildungen

Abb. 1: Startseite von Google Books 9

Abb. 2: Trefferliste (in Listenansicht) 12

Abb. 3: Detaillierte Beschreibung eines Buches 15

Abb. 4: Ansicht eines Buches 17

Abb. 5: Persönliche Bibliothek 19

Abb. 6: Trefferliste im OPAC 30

Abb. 7: Inhaltsverzeichnis mit Standard-Überschriften 63

Diagramme

Diagramm 1: Recall-Precision-Graph 51

Diagramm 2: Häufigkeit der Relevanzurteile auf der nicht-binären Skala in absoluten Zahlen 53

Diagramm 3: Anzahl an Suchanfragen mit n relevanten Treffern 54

Diagramm 4: Precision-Ergebnisse dieser Studie im Vergleich mit Ergebnissen der Google Websuche aus LEWANDOWSKI 2008a 55

Diagramm 5: Zugänglichkeit der Bücher im Anzeigemodus „Eingeschränkte Vorschau und Vollansicht“ 57

Diagramm 6: Zugänglichkeit der Bücher im Anzeigemodus „Alle Bücher“ 58

Diagramm 7: Verfügbarkeit der Ergebnisse im OPAC. 59

Diagramm 8: Verteilung der Treffer auf Zeitabschnitte nach Erscheinungsjahr, nach einzelnen Angeboten bzw. Anzeigemodi getrennt 67

Diagramm 9: Verteilung der Treffer auf Zeitabschnitte nach Erscheinungsjahr, nach einzelnen Arten der Zugänglichkeit getrennt 67

Tabellen

Tabelle 1: Bestandsgrößen von Hochschulbibliotheken 38

Tabelle 2: Häufigkeit mit der die einzelnen Datenfelder zur Suche im OPAC genutzt werden 40

Tabelle 3: Worthäufigkeiten im OPAC 41

Tabelle 4: Worthäufigkeiten der ausgewählten Suchanfragen 43

Tabelle 5: Durchschnittliche Precision der Ergebnisse aller Suchanfragen in beiden Angeboten, getrennt nach Anzeigemodus 52

Tabelle 6: Durchschnittliche Relevanz der Ergebnisse aller Suchanfragen getrennt nach Anzeigemodus bzw. Art der Zugänglichkeit. 52

Tabelle 7: Anzahl der Suchanfragen, bei denen die Angebote die besseren Ergebnisse hatten. 52

Tabelle 8: Ergebnisse der *Known-Item*-Suchanfragen 56

Tabelle 9:Verfügbarkeit der einzelnen Kataloganreicherungen in Prozent bei den einstellbaren Anzeigemodi von Google Books sowie im OPAC 60

Tabelle 10: Verfügbarkeit der einzelnen Kataloganreicherungen in Prozent für die einzelnen Arten der Zugänglichkeit bei Google Books 61

Tabelle 11: Anzahl an Rezensionen bei den einstellbaren Anzeigemodi von Google Books sowie im OPAC 62

Tabelle 12: Übersicht über die Häufigkeit des Vorkommens von Inhaltsverzeichnissen sowie deren Qualität bei den einstellbaren Anzeigemodi von Google Books sowie im OPAC in Prozent der Datensätze 64

Tabelle 13.: Übersicht über die Häufigkeit des Vorkommens von Inhaltsverzeichnissen sowie deren Qualität für die einzelnen Arten von Zugänglichkeit bei Google Books in Prozent der Datensätze 65

Tabelle 14: Durchschnittliches Alter der Ergebnisse bei den einstellbaren Anzeigemodi von Google Books sowie im OPAC. 68

Tabelle 15: Durchschnittliches Alter der Ergebnisse für die einzelnen Arten von Zugänglichkeit bei Google Books 68

1. Einleitung

Die Entwicklung und Verbreitung der modernen Informationstechnologien veränderte in den letzten Jahren die Welt. Auch das Informationswesen selbst veränderte sich in allen Bereichen. Klassische Dokumentationseinrichtungen wie Archive, die nicht nur Informationen hüten, sondern Objekte, verwalten ihre Bestände mit Hilfe von Datenbanken. Einrichtungen, deren Geschäft die reine Information ist, stellen ihre Dienstleistung – über einen Datenbankzugang – mittels des Internets der ganzen Welt zur Verfügung. Im Internet wandeln sich die Nutzer von reinen Konsumenten von Informationen zu gleichermaßen Informationen produzierenden Teilnehmern. Aber auch die Institutionen, die das klassische analoge Medium schlechthin - das Buch – verwalten, nutzen neue Technologien. Diese kommen sowohl in Form der neuen Medien wie CDs, DVDs, Ebooks und Datenbanken zum Einsatz, aber auch als Mittel die Bestände zu erschließen. Viele Zettelkataloge wurden durch OPACs (Online Public Access Catalogue) ersetzt, die in der Regel über das Internet verfügbar sind.

Diese neuen Medien und Methoden haben sich in verblüffender Geschwindigkeit entwickelt. Vor etwa 15 Jahren, 1995, wurden Diplomarbeiten über die damals neuartige Technik OPAC, deren Benutzung und Einführung geschrieben (vgl. BEHN 1995). Heutzutage stellt das Internet bei Studierenden die am zweithäufigsten genutzte Methode dar, um nach Literatur zu recherchieren (vgl. KLATT 2001, S. 10). Die Suche in allgemeinen Suchmaschinen wird dabei wesentlich häufiger genutzt als die Recherche in Quellensammlungen oder Katalogen der lokalen Bibliothek (ebd. S. 11).

Informationssuchende haben die Möglichkeit mit weniger Aufwand auf mehr Quellen zuzugreifen als dies früher der Fall war. Um sich orientieren zu können, stellt sich die Frage nach der Qualität der einzelnen Quellen. Gibt es für manche Zwecke bessere Informationslieferanten als für andere? In vielen Fällen ist dies offensichtlich: Geht es um schnelle Informationen, ist die Homepage der meisten Tageszeitungen aktueller als die Zeitung selbst. Geht es um verlässliche Informationen, ist der Brockhaus angebrachter als die Wikipedia.

In anderen Fällen ist die Qualität der Informationsangebote weniger offensichtlich. Dies ist bei der Google Book Search der Fall. Auf den ersten Blick scheint das Angebot die Bibliothek der Zukunft zu sein – es werden, von den Medien gelöst, die Informationen

aus Büchern zur Verfügung gestellt und im Volltext durchsuchbar gemacht. Darüber hinaus werden innovative Leistungen im Katalog angeboten, beispielsweise werden dem Nutzer auf einer Landkarte die Orte angezeigt, die in einem Buch genannt werden, ähnliche Bücher zum aktuellen Treffer angezeigt oder aber besonders beliebte Textstellen als Zitate angeboten.

Weitere Möglichkeiten wie mit digitalen Büchern und Texten gearbeitet werden kann, sind denkbar und sicher auch in Entwicklung. Als Beispiel genannt sei die Funktion des kooperativen Kommentierens. Mit dieser haben die Nutzer die Möglichkeit, Kommentare zu Textstellen von auf Servern gespeicherten Dokumenten zu schreiben. Andere Nutzer können diese Kommentare sehen und ergänzen, eine vernetzte Struktur entsteht (vgl. HEIDENREICH 2009).

Für Google scheinen die Grenzen des Machbaren derzeit nicht technischer Natur zu sein, sondern rechtlicher. In der Verhandlung einer Sammelklage wird in den USA die Frage untersucht, ob das Vorgehen Googles beim Aufbau des Book Search Programms gegen Copyright-Gesetze verstößt.

Während es sicher ist, zu sagen, dass der Bestand der Google Book Search moderner präsentiert wird als der Bestand der meisten Bibliotheken, kann nicht ohne weiteres eine Aussage über die Qualität des Angebots der Google Buchsuche als ganzes getroffen werden. Um die Qualität eines Informationsangebots bewerten zu können, müssen über die Präsentation der Treffer hinaus andere Faktoren bewertet werden: Die Größe und Aktualität des Index, das Ranking-Verfahren und auch die Qualität des Retrievals (vgl. LEWANDOWSKI 2007).

Diese Arbeit untersucht als Teilaspekt der Qualität des gesamten Angebots lediglich die Qualität von Retrievalergebnissen. Die Google Book Search wird dem OPAC einer wissenschaftlichen Bibliothek gegenüber gestellt. Dabei wird vom Nutzungsverhalten in OPACs als Grundlage ausgegangen. Dieses wurde in früheren Untersuchungen durch protokollierte Suchanfragen, sog. *Logfiles*, ermittelt. Auch ein Vergleich zwischen der Buchsuche und der Websuche von Google wäre eine interessante Untersuchung. Hier soll jedoch die klassische Bibliothek mit demjenigen Angebot verglichen werden, dass

wohl am ehesten mit dem vieldiskutierten Begriff „digitale Bibliothek“ bezeichnet werden kann.

Des Weiteren soll auch die Häufigkeit untersucht werden, mit der einige Katalog-Funktionen zur Dokument-Beschreibung (Inhaltsverzeichnisse, Buchcover und Abstracts) vorhanden sind, wobei es sich in dieser Studie bei den beschriebenen Dokumenten immer um Bücher handelt. Dies ist für die Untersuchung der Retrievalergebnisse insofern interessant, als dass „durch eine verbesserte Repräsentation auch die Trefferlisten in ihrer Aussagekraft verbessert werden“ können (LEWANDOWSKI 2005, S. 217). Die Aussagekraft der Trefferlisten beeinflusst, in wie weit sich Nutzer ein korrektes Bild von der Relevanz eines Treffers machen können.

Auf die angesprochenen Katalog-Funktionen soll nur soweit eingegangen werden, wie dies im Rahmen einer quantitativen Untersuchung möglich ist. Die erwähnten Funktionen, die sowohl von der Google Book Search als auch von dem untersuchten OPAC angeboten werden, werden im Hinblick darauf verglichen, bei wie vielen Treffern welche dieser Funktionen vorhanden sind. Eine qualitative Untersuchung darüber, in wie weit diese Katalog-Funktionen dem Nutzer im Einzelnen helfen, die Relevanz eines Treffers zu beurteilen, kann hier nicht erfolgen. Ebenso wenig soll hier das Browsing-Verhalten der Nutzer innerhalb eines OPACs untersucht werden.

Aus der Tatsache, dass das Internet häufiger zur Informationsrecherche verwendet wird als der Bibliothekskatalog, kann gefolgert werden, dass viele Nutzer solchen Medien, die an jedem Ort unmittelbar verfügbar sind, den Vorzug geben vor Medien, die nur in der Bibliothek verfügbar sind. Die Relevanz dieser Quellen wird gesondert untersucht.

Da aber der Schwerpunkt des Dienstleistungsangebots einer Bibliothek nicht auf der Bereitstellung von Online-Medien liegt, wäre ein Vergleich der Buchsuche ausschließlich mit den im OPAC nachgewiesenen Online-Quellen wenig aussagekräftig. Für diese Untersuchung wird deshalb davon ausgegangen, dass Bibliotheksnutzer bereit sind, die Räumlichkeiten der Bibliothek aufzusuchen um Medien zu entleihen.

Bei Suchanfragen innerhalb eines OPACs sind Anfragen mit dem Autorennamen oder dem Buchtitel als Suchargument als *Known-Item-Search* einzuordnen, da der Nutzer

offensichtlich ein bestimmtes Buch bzw. Bücher eines bestimmten Autors sucht. Suchanfragen nach Themen sind hingegen keine *Known-Item-Search*, da davon auszugehen ist, dass der Nutzer kein bestimmtes Buch sucht, sondern alle Bücher zu einem bestimmten Thema finden möchte.

Bei einer *Known-Item-Search* ist nur wichtig, ob das gesuchte Item unter den Treffern ist und an welcher Position es sich befindet (vgl. HAWKING 2001, S. 23). In dieser Untersuchung werden Ranking und Verfügbarkeit ermittelt. Anders als bei einer thematischen Suche ist es bei der Suche nach bekannten Objekten nicht sinnvoll, die Retrievalqualität zu bewerten.

Für die Bewertung von Ergebnissen im Rahmen eines Retrievaltests eignen sich themenorientierte Suchanfragen. In der Untersuchung werden in beiden Katalogen jeweils 50 Suchanfragen durchgeführt und jeweils die ersten zehn Treffer bewertet. Darüber hinaus werden 70 *Known-Item*-Suchanfragen durchgeführt. Bei der Hälfte der Suchanfragen wird nach einem bestimmten Buch gesucht, bei der anderen Hälfte wird nach einer Kombination von Buchtitel und Autorenname gesucht.

Die Frage, ob und warum die Google Buchsuche besser oder schlechter ist, als eine herkömmliche Bibliothek, wird in dieser Arbeit nicht umfassend beantwortet werden können. Zumindest der Teilaspekt Retrievalqualität kann hier aber untersucht werden, was zusammen mit zukünftigen Arbeiten vielleicht einen umfassenderen Vergleich der Buchsuche mit Bibliotheken ermöglichen wird. Daran, dass eine Untersuchung und Bewertung der Buchsuche aus bibliothekswissenschaftlicher Sicht von Interesse ist, kann angesichts der oben angesprochenen und unter 2.1.1 ausführlicher behandelten Entwicklungen kein Zweifel bestehen.

Der Vergleich der Google Book Search mit wissenschaftlichen Bibliotheken liegt nahe, weil unterstellt werden kann, dass das Google-Angebot auf wissenschaftliche Literatur fokussiert ist. Dies zumindest zeigt eine Betrachtung der am Programm teilnehmenden Bibliotheken. Zwar digitalisiert Google unterschiedslos große Teile der Bestände der teilnehmenden Bibliotheken, aber nur bei zweien der 27 [1](vgl. CEYNOWA 2008)

[1] Auf der Homepage des Google Books Angebots (vgl. GOOGLE BOOKS 2009) sind nur 20 Bibliotheken genannt, allerdings erhebt die dort veröffentlichte Liste keinen Anspruch auf Vollständigkeit.

teilnehmenden Bibliotheken handelt es sich um öffentliche Bibliotheken. Es sind die New York Public Library und die Bibliothèque Municipale de Lyon. Wie viele Medien dort jeweils digitalisiert wurden ist nicht bekannt.

Es ist also davon auszugehen, dass mehrere Millionen Bücher für den nicht-akademischen Gebrauch in den Bestand der Google Book Search aufgenommen wurden, der Fokus beim Bestandsaufbau aber auf wissenschaftlichen Bibliotheken lag. Das Google Print Partner Programm ermöglicht es Verlagen auf die Darstellung eigener Bücher in der Book Search Einfluss zu nehmen und darüber hinaus auch eigene Bücher in die Book Search zu integrieren.

Eine Betrachtung der Retrievalqualität des Angebots Google Scholar, das sich ebenfalls an Akademiker richtet, wäre eine interessante Untersuchung. Allerdings werden über Scholar nur im Ausnahmefall Bücher angeboten, stattdessen bilden Artikel wissenschaftlicher Zeitschriften sowie sonstige wissenschaftliche Arbeiten den Bestand (vgl. LEWANDOWSKI 2006, S. 74). Wie beim Angebot Book Search auch, sind die indexierten Dokumente des Google Scholar Angebots im Volltext durchsuchbar. Es werden wie auch bei der Book Search innovative Werkzeuge angeboten, wie etwa die Anzeige, welche Artikel von wem zitiert wurden.

Mit diesem Dienstleistungsangebot bietet Google jedoch keine im engeren Sinne bibliothekarische Dienstleistung an. Während in Bibliothekskatalogen häufig nicht einmal die Titel der einzelnen Artikel in einer Zeitschrift nachgewiesen werden, wird eine Texterschließung im Volltext von Anbietern kommerzieller Datenbanken, wie z.B. SpringerLink, angeboten. Somit müsste Google Scholar mit anderen Volltext-Datenbanken verglichen werden und nicht mit dem OPAC einer Bibliothek.

2. Die untersuchten Angebote

Bevor unter Gliederungspunkt 3. (Ziele der Untersuchung) darauf eingegangen wird, inwiefern sich welche Angebote und Dienstleistungen von Google mit denen eines OPACs vergleichen lassen, werden diese Angebote hier detailliert beschrieben.

2.1. Google Book Search

Die Firma Google Inc. wurde 1998 von zwei Doktoranden der Informatik an der Stanford University gegründet, Sergey Brin und Larry Page (vgl. GOOGLE 2009). Im selben Jahr nahm die gleichnamige Suchmaschine den Betrieb auf. Das Unternehmensziel der Firma aus Mountain View im Silicon Valley besteht nach eigenen Angaben darin „die Informationen der Welt zu organisieren und allgemein nutzbar und zugänglich zu machen“ (GOOGLE 2009a). Die Gründer entwickelten eine Suchmaschine für Webseiten, Usenet-Mitteilungen und ab 2001 auch für Bilder. Heute hat Google Inc. weltweit mehr als 20.000 Mitarbeiter und für den Index der Suchmaschine mehr als 8 Mrd. Webseiten und jeweils 1 Mrd. Bilder und Usenet-Mitteilungen durchsucht (vgl. FAZ FINANCE 2009, GOOGLE 2009a). Im Juli 2008 wurden allein aus Deutschland 100 Mio. Suchanfragen täglich an Google-Server gerichtet (COMSCORE 2008).

2006 betrug der Jahresumsatz ca. 10 Mrd. US-Dollar (vgl. COMPUTERWOCHE 2007, S.3), 2008 waren es 21,8 Mrd. US-Dollar. Der Gewinn konnte innerhalb von 3 Jahren von 3,5 Mrd. USD (2006) auf 6,6 Mrd. USD (2008) gesteigert werden. Der Börsenwert schwankt naturgemäß, derzeit – stand August 2009 - bewegt er sich etwa um 109,3 Mrd. € (vgl. FAZ FINANCE 2009). Zum Vergleich: Der Börsenwert der Thyssen Krupp AG betrug zum gleichen Zeitpunkt 11,3 Mrd. Euro, derjenige der Axel Springer AG 2,1 Mrd. Euro (alle Stand 13.8.2009; vgl. FAZ FINANCE 2009a und FAZ FINANCE 2009b).

Das Angebot, dass heute unter Google Book Search firmiert, wurde unter dem Namen Google Print for Publishers gegründet und später um das Google Print Library Program ergänzt. Google Print wurde später in Google Book Search umbenannt, das Verleger-Programm in Google Book Search Partner Program. Mittlerweile firmiert das Angebot unter dem kürzeren Namen Google Books. Im Folgenden werden die Begriffe Google Books, Book Search und Buchsuche synonym verwendet.

Im Oktober 2004 stellten die Google-Gründer auf der Frankfurter Buchmesse den Plan für den Aufsehen erregenden neuen Dienst Google Print for Publishers vor. Im Rahmen dieses Dienstes wurden Bücher komplett eingescannt oder von Verlegern digital übermittelt. Nutzer sollten die Volltexte der Bücher durchsuchen, diese jedoch nicht ansehen können. Zugriff sollte nur auf kurze Auszüge der Bücher möglich sein (vgl. GOOGLE BOOKS 2009a, GOOGLEBLOG 2004).

Google warb bei den Verlegern für die Teilnahme am Programm und damit für das zur Verfügung stellen von Büchern für Google Books mit zwei Argumenten. Zum einen entstünde durch die Integration der Bücher in den Google-Index kostenlose Werbung für die eigenen Bücher, da Google-Nutzer bei Recherchen darauf stoßen. Zum anderen erhielten die Verleger eine genaue Übersicht darüber, wie oft welches Buch in der Buchsuche aufgerufen wurde und wären so genauestens über das Interesse an den einzelnen Titeln informiert. Dies würde sich dazu nutzen lassen, wesentlich präziser als früher planen zu können, welche vergriffenen Bücher aus dem Verlagsprogramm neu aufgelegt werden sollen (vgl. GOOGLE 2004). Es ist davon auszugehen, dass die erste Behauptung der Wahrheit entspricht und die Teilnahme an Google Books tatsächlich zu mehr Aufmerksamkeit für die Bücher führt. Dass die zweite Behauptung stimmt wurde mittlerweile von Verlagen bestätigt, u.a. von Cambridge University Press (vgl. FUNK 2008, DATEMA 2007). Zurzeit nehmen nach Angaben von Google 20.000 Verleger am Google Books Partner Programm teil (vgl. GOOGLEBLOG 2008).

Im Dezember 2004 folgte die bemerkenswert knappe Ankündigung für den zweiten Teil des Google Print Programms: Google Print for Libraries (vgl. GOOGLE 2004). Mit Hinweis auf das selbst gewählte Ziel der Firma, Informationen zugänglich zu machen, gab Google bekannt, dass Bücher aus den Beständen der mit Google kooperierenden Bibliotheken digitalisiert werden sollten, um dann wie bei beim Google Print Partner Programm auch im Volltext durchsucht werden zu können. Urheberechtsfreie Werke sollten auch im Volltext angesehen oder heruntergeladen werden können. Bei den ersten fünf teilnehmenden Bibliotheken handelte es sich um die Universitätsbibliotheken der Oxford University, die Universitätsbibliotheken Michigan, Harvard, Stanford sowie die New York Public Library.

Die Bibliotheken erhalten für das zur Verfügung stellen ihrer Bestände im Gegenzug eigene Kopien der Digitalisate, sogenannte *library digital copies*. Neben dem erleichterten Zugang zu den Büchern (per Volltextsuche) nennt Google als bedeutenden Vorteil des Programms, dass Nutzer nun auf Bücher zugreifen könnten die seit geraumer Zeit nicht mehr gedruckt werden und deshalb gegenwärtig wenig Leser hätten (vgl. GOOGLEBLOG 2004a, GOOGLE 2004). Mittlerweile ist die Zahl der teilnehmenden Bibliotheken gewachsen, es nehmen nun u.a. auch die Universitätsbibliotheken von Kalifornien, Virginia, Wisconsin, Princeton, Barcelona, der Complutense Madrid sowie der Bayerischen Staatsbibliothek teil (vgl. GOOGLE BOOKS 2009). Insgesamt beteiligen sich gegenwärtig 27 Bibliotheken am Programm, 30 Mio. Medien sollen eingescannt werden (vgl. CEYNOWA 2008, S. 7). Der Gesamtbestand von Google Books beträgt derzeit 7 Mio. Bücher (vgl. GOOGLEBLOG 2008).

Die Mehrzahl der Kooperationsverträge zwischen Google und den Bibliotheken unterliegt einer Verschwiegenheitsvereinbarung (vgl. BAND 2006, S.2 ; CEYNOWA 2008, S. 5). So ist zu erklären, dass es nur wenige Einzelheiten zur genauen Zusammenarbeit gibt. Die meisten Informationen können den Public-Private-Partnership-Verträgen entnommen werden, die Google mit den Universitätsbibliotheken von Michigan und Kalifornien schloss (vgl. UCLA 2006). Warum diese Verträge veröffentlicht werden durften ist unbekannt. Bekannt ist hingegen, dass die University of Michigan technische Vorkehrungen gegen automatisiertes Herunterladen der Digitalisate zu treffen hat (vgl. BAND 2006, S.2). An der Oxford University sollen nur Bücher eingescannt werden die in *public domain*[2] sind, ebenso an der Bibliothek der Universität Harvard (vgl. MILNE 2005, S. 27 ; BAND 2006, S.2).

[2] Der US-amerikanische Rechtsbegriff *in public domain* bedeutet wörtlich übersetzt „in öffentlichem Eigentum“ und entspricht in etwa dem deutschen Rechtsbegriff „gemeinfrei“, bezeichnet also Werke die nie dem Copyright bzw. Urheberrecht unterlegen haben, oder deren Rechtsschutz erloschen ist. Werke werden gemeinfrei, wenn das Urheberrecht erlischt, was in Deutschland 70 Jahre nach dem Tod des Autors der Fall ist (vgl. § 64 Urheberrechtsgesetz). Im US-amerikanischen Rechtsgebiet gibt es, bedingt durch einige Rechtsreformen, je nach Art des Werkes unterschiedliche Fristen, wann ein Werk in *public domain* fällt. Meist entweder, wie in Deutschland, 70 Jahre nach dem Tod des Autors, oder aber 120 Jahre nach Schöpfung bzw. 95 Jahre nach Veröffentlichung (vgl. HIRTLE 2009).Gemeinfreie Werke können von jedermann frei genutzt, bearbeitet und veröffentlicht werden. Es sei darauf hingewiesen, dass sich amerikanisches Copyright und deutsches Urheberrecht voneinander unterscheiden, hier jedoch nicht weiter auf die Einzelheiten eingegangen werden kann.

In Oxford unterliegt die Nutzung der *library digital copy* einigen nicht näher bekannten „ […] constraints designed to protect the investment made by Google […] “ (MILNE 2005, S. 26). Google Books ist zurzeit in 70 Ländern verfügbar (vgl. DATEMA 2007a).

2.1.1. Interface und Benutzung

2.1.1.1. Die Suche in Google Books

Das Angebot Google Books ist über den Browser direkt durch Eingabe der Adresse http://books.google.de zu erreichen. Für die alten Adressen http://print.google.de und http://booksearch.google.de ist eine Weiterleitung auf die aktuelle Adresse eingerichtet, sie funktionieren also noch. Auf der Startseite der Google-Websuche http://www.google.de ist das Google Books Angebot über den in der Kopfzeile platzierten Menüpunkt „Mehr“ zu erreichen (vgl. Abb. 1).

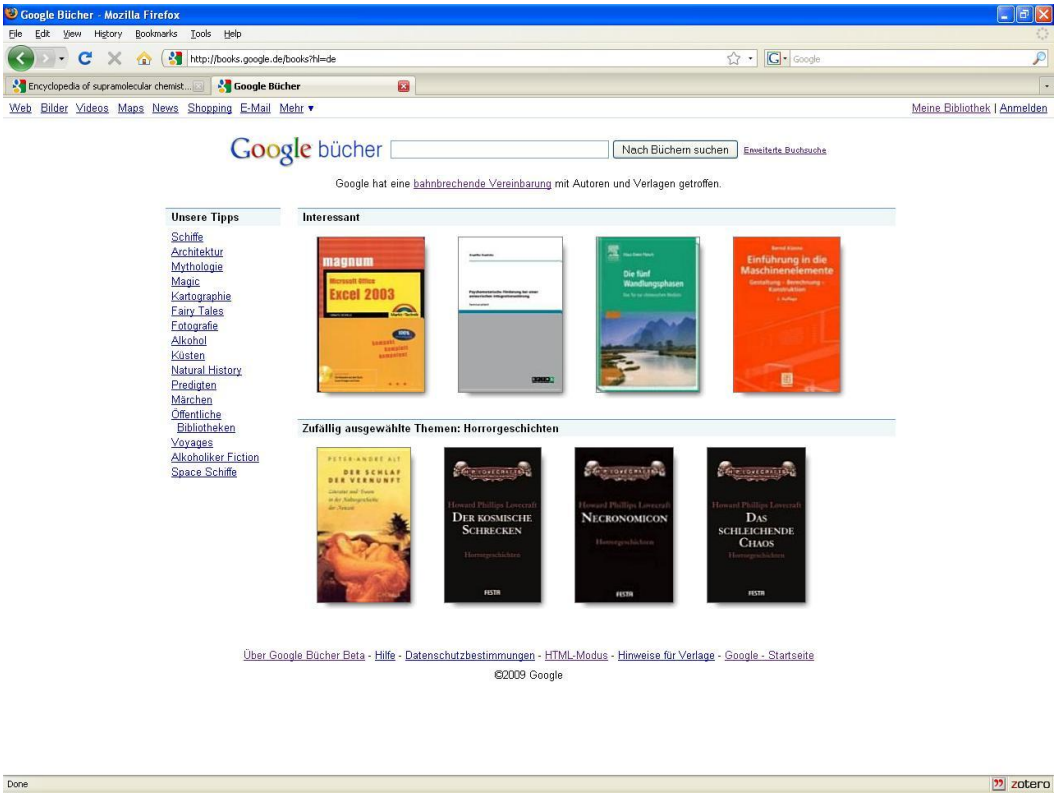


Abb. 1: Startseite von Google Books

Auf der Startseite der Buchsuche finden sich auf der linken Bildschirmseite 16 in Textform dargestellte Links zu Suchanfragen, die mit „Unsere Tipps“ überschrieben sind. Wird ein solcher Link angeklickt, wird die entsprechende Suchanfrage ausgeführt und die Ergebnisseite angezeigt. Bei den Links scheint es sich um Anfragen zu handeln, die

zufällig aus einer umfangreicheren Liste ausgewählt werden und nicht aus einem Protokoll aktueller Suchbegriffe. Dafür, dass diese Suchbegriffe aus einem solchen Protokoll stammen, spricht die Tatsache, dass bei jedem Aufruf der Startseite andere Links angezeigt werden, selbst wenn nur wenige Sekunden zwischen den Aufrufen liegen. Nach mehrmaligem Aufrufen der Startseite zeigt sich aber, dass es Links bzw. Suchbegriffe gibt, die wiederholt in der Liste auftauchen. Dies ist prinzipiell auch bei protokollierten Suchanfragen möglich, die sich wiederholenden Begriffe sind aber derart speziell, dass zufällige Wiederholungen ausgeschlossen werden können. Über mehrere Tage hinweg werden dieselben, sehr speziellen Suchbegriffe angezeigt. Auch das macht es unwahrscheinlich, dass sie aus einem Protokoll aktueller Suchanfragen stammen.

Am oberen Bildschirmrand findet sich die Kopfzeile, die bei allen Google-Angeboten vorhanden ist. Linksbündig enthält sie Links zu den einzelnen Google-Angeboten, dieser Teil der Kopfzeile ist bei allen Google-Angeboten gleich. Rechtsbündig enthält die Kopfzeile ein oder zwei Links zu wichtigen Funktionen des jeweiligen Google-Angebots. Diese unterscheiden sich von Angebot zu Angebot. Bei Google Books sind über diese Links die Funktionen „meine Bibliothek“ und „anmelden“ aufrufbar, die unter Gliederungspunkt 2.1.1.5. erläutert werden.

Unter der Kopfzeile schließt sich das Eingabefeld für Suchanfragen an. Zwischen dem Logo der Buchsuche und dem Button zum Ausführen der Suche (mit „nach Büchern suchen“ beschriftet) findet sich das Suchfeld. Rechts neben dem Button ist ein unauffälliger Link zur „Erweiterten Buchsuche“ platziert.

Den größten Teil des Bildschirms nehmen zwei Reihen von je vier Buchcovern ein. Die erste Reihe von Covern ist überschrieben mit „Interessant“, die zweite Reihe mit „Zufällig ausgewählte Themen:“. Nach dem Doppelpunkt folgt der Name eines Themas, bspw. „Zimmerreinigung“. Die zufällig ausgewählten Themen werden vermutlich zufällig aus derselben Liste ausgewählt, aus der auch die oben erwähnten Links der Rubrik „Unsere Tipps“ ausgewählt werden. Diese Vermutung liegt nahe, da sich an beiden Orten dieselben, teils sehr speziellen Begriffe finden.

Die erweiterte Buchsuche, die unter der Adresse http://books.google.de/advanced_book_search auch direkt per Adresseingabe aufgerufen werden kann, bietet zahlreiche Funktionen, um nach Büchern zu suchen, wie sie auch aus OPACs bekannt sind. Angeboten wird z.B. die Suche nach Sprache, Veröffentlichungszeitraum und ISBN bzw. ISSN. Auch die Einschränkung der Suchanfrage auf bestimmte Datenfelder (Autor, Titel, Verlag) ist möglich, ebenso wie Verknüpfungsoperatoren (AND, OR, NOT) verfügbar sind. Die Suchergebnisse können eingeschränkt werden, so dass ausschließlich Zeitschriftenartikel, ausschließlich Bücher oder aber Treffer aus beiden Medienarten angezeigt werden. Als Kriterium für die Einschränkung der Suchergebnisse steht auch der Anzeigemodus zur Auswahl. Hier kann als erste Option angegeben werden, dass nur Bücher angezeigt werden sollen, die in vollständiger Ansicht vorliegen, als zweite Option, dass Bücher angezeigt werden, die in Form von eingeschränkter Vorschau oder vollständiger Ansicht vorliegen, und als dritte Option, dass alle Bücher angezeigt werden sollen (was auch diejenigen enthält, deren Text nur als Auszug gelesen werden kann). Unter 2.1.2.1. wird detailliert erörtert, warum es unterschiedliche Anzeigemodi gibt.

2.1.1.2. Ergebnisspräsentation

Führt man eine Suche durch, werden die Ergebnisse, also die Liste der Bücher die den Suchbegriff enthalten, standardmäßig in Listenform dargestellt. Es kann auch eine Cover-Ansicht ausgewählt werden. Bei beiden Ansichten findet sich am oberen Bildschirmrand die schon erwähnte Kopfzeile und darunter das Suchfeld mit entsprechendem Button. Daran schließt sich eine Zeile an, die ein Drop-Down-Menü enthält, mit dem die dargestellten Ergebnisse auf Bücher beschränkt werden können, die in einem bestimmten Anzeigemodus vorliegen (s.o.). Neben diesem Menü enthält die Zeile noch die Anzahl der Treffer.

Bei der Coveransicht werden 30 Cover auf einer Seite dargestellt. Für viele Bücher enthält die Google-Datenbank noch kein eingescanntes Cover, hier wird bei der Cover-Ansicht als Ersatz auf simple Grafiken zurückgegriffen, die Autor und Titel des Buches enthalten. Die Darstellung der Ergebnisse in Listenform enthält wesentlich mehr Informationen und ähnelt stark der Darstellung von Ergebnissen der Websuche (vgl. Abb. 2). Die Sortierung der Liste – also die Festlegung der Reihenfolge der einzelnen

Treffer - erfolgt mittels eines Ranking-Algorithmus. Andere Kriterien zur Sortierung, z.B. das Veröffentlichungsdatum, können nicht gewählt werden.

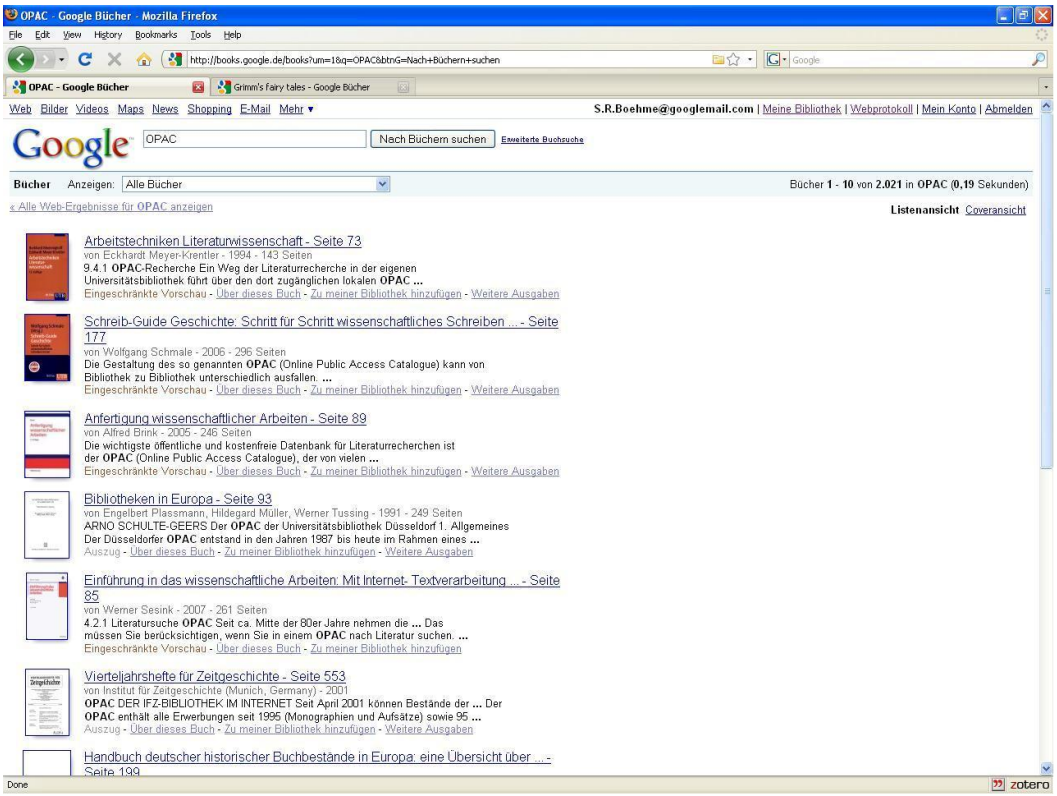


Abb. 2: Trefferliste (in Listenansicht)

Bei der Listendarstellung wird links neben den textuellen Informationen zu jedem Ergebnis ein *Thumbnail* angezeigt, d.h. eine verkleinerte Darstellung der Seite, auf der der Suchbegriff gefunden wurde. Die Überschrift des Treffers ist der Titel des Buches. Daran ist, durch einen Spiegelstrich abgetrennt, die Seitenzahl der Seite auf der das Suchwort gefunden wurde angehängt. Darunter werden in der ersten Zeile der Trefferbeschreibung in grauer Schrift die wichtigsten bibliographischen Informationen dargestellt: Der Name des Autors, das Jahr der Veröffentlichung und die Seitenzahl des Buches. In der zweiten und dritten Zeile sind ein oder zwei Treffer als KWIC (Keyword in Context) angegeben. Dieser Textzusammenhang scheint maximal 80 Zeichen lang zu sein. In der vierten und letzten Zeile steht die Information, in welchem Anzeigemodus das Buch abgerufen werden kann. Desweiteren ermöglichen hier platzierte Links das weitere Arbeiten mit dem Treffer: er kann in die persönliche Bibliothek übernommen werden, es können Informationen über das Buch angezeigt werden und ggf. nach weiteren Ausgaben gesucht werden.

Im Gegensatz zur „klassischen“ Google Websuche gibt es bei der Buchsuche keine als solche kenntlich gemachten *Sponsored Links.* Dabei handelt es sich um Werbung, die dem Nutzer auf der Seite mit den Suchergebnissen in einer Form präsentiert wird, die den nicht gesponserten Links (*Organic Links* genannt) ähnlich sieht. Auch darüber hinaus ist die Ergebnispräsentation der Buchsuche eine „klassische“ Listendarstellung und keine zusammenfassende Darstellung von Treffern aus mehreren Datenquellen im Sinne einer *Universal Search*. Bei der Websuche ist es mittlerweile üblich, Ergebnisse aus mehreren Indizes in einer „universalen“ Liste darzustellen, wenn dies möglich ist. Bspw. werden Ergebnisse der Bildersuche, der Suche im Nachrichtenindex, Videos oder Aktienkurse in die Ergebnisse der Google-Websuche integriert (vgl. QUIRMBACH 2009, S.221). Der Verzicht auf eine solche Darstellung bei der Buchsuche liegt nahe, denn es ist davon auszugehen, dass der Index ausschließlich Bücher und andere Printmedien enthält.

2.1.1.3. Darstellung eines Buches

In der detailierten Beschreibung eines einzelnen Buches sind bei der Google Buchsuche zahlreiche Informationen verfügbar. Unter der obligatorischen Kopfzeile und dem darunterliegenden Suchfeld mit dem Suchbutton ist der Bildschirmbereich vertikal in zwei Frames unterteilt (allerdings ohne dass das HTML-Tag Frame verwendet wird). Der schmalere linke Frame enthält eine verkleinerte Darstellung von Cover oder Titelseite des Buches. Die Titelseite statt des Einbands wird häufig bei älteren Büchern angezeigt, da aufwändige Einbandillustrationen nach dem Mittelalter erst im 20. Jahrhundert wieder modern wurden. Neben Titelblatt oder Cover findet sich das Navigationsmenü, mit dem zwischen Informationen über das gewählte Buch („Übersicht“), der Ansicht des Buchtextes selbst, Rezensionen des Buches und der Kauf-Funktion („Kaufen“) gewechselt werden kann.

Darunter befindet sich eine kleine Leiste von fünf Stern-Symbolen, welche farblich die durchschnittliche Bewertung des Buches durch die Nutzer darstellt. In diesem linken Frame finden sich außerdem ein Suchfeld mit dem ausschließlich das angezeigte Buch durchsucht werden kann und Links zu Händlern, die idealerweise direkt auf die Seiten innerhalb der Onlineshops verweisen sollen, auf denen das gerade ausgewählte Buch zum Verkauf angeboten wird. Unter diesen Links finden sich weitere Elemente, die nicht bei allen Büchern angezeigt werden. Wovon es im Einzelnen abhängt, welche Elemente angezeigt werden und welche nicht, ist nicht bekannt. So werden bei einem Teil der

Bücher die Cover- bzw. Titelblätter von drei „ähnlichen Büchern“ angezeigt. Ebenfalls optional sind (mehr oder weniger) kontextbezogene Werbung sowie ein Logo des veröffentlichenden Verlags.

Der größere rechte Frame dient zur Darstellung des Buchinhalts. Hat der Nutzer eine Suche ausgeführt und ruft ein Suchergebnis auf, wird er direkt zur Ansicht einer das Suchwort enthaltenden Seite geleitet. Welche Seite – im Falle von mehreren Treffern in einem Buch – aufgerufen wird, ist nicht klar (vgl. LEWANDOWSKI 2006a). Im rechten Frame wird auch die detailliertere Buchbeschreibung dargestellt, die durch das Menü im linken Frame oder von der Ergebnisliste aus, durch den Link „Über dieses Buch“, erreicht werden kann (vgl. Abb. 3). Hier findet sich – sofern vorhanden – eine „Buchübersicht“, vermutlich werden unter dieser Überschrift meist Klappentexte dargestellt. [3]

Darunter befindet sich ein Button zur Ansicht des Buchtextes. Die Benennung dieses Buttons richtet sich nach dem Anzeigemodus des Buches. Bei Büchern, die im Volltext angezeigt werden, heißt er „Dieses Buch lesen“. Bei Büchern, die in Teilen gezeigt werden dürfen, heißt er „Voransicht des Buches“. Bei Büchern, die nur in Auszügen dargestellt werden dürfen, ist demensprechend keine Ansicht und somit auch kein Button vorhanden. Ist das Buch im Volltext verfügbar, findet sich zudem ein weiterer Button zum Download des Buches als PDF-Datei.

Für alle weiteren beschreibenden Elemente der Trefferrepräsentation gilt, wie auch für die Elemente im linken Frame, dass sie nur bei einem Teil der Bücher angezeigt werden. Die Anzahl der bei einem Buch angezeigten Elemente schwankt stark. Bei einigen Büchern werden nur ein oder zwei angezeigt, bei anderen fast alle.

Unter dem Button zur Anzeige werden – sofern vorhanden - Rezensionen zum Buch platziert. Es wird bei jeder Rezension kenntlich gemacht, ob es sich um eine von einem Nutzer geschriebene Bewertung oder um eine professionelle Rezension handelt

[3] Dies zumindest lassen Beobachtungen vermuten. Bei fünf von sechs Büchern, die unter der Rubrik „Buchübersicht“ näher beschrieben waren, handelte es sich bei der Beschreibung um Texte, die den Klappentexten der Printausgaben sehr ähnlich waren. Die Tatsache, dass aber in keinem Fall der exakte Klappentext digitalisiert worden war, ist interessant. Über Ursachen dafür könnte nur spekuliert werden. In einem Fall handelte es sich um einen als solchen gekennzeichneten Auszug aus dem Buch.

Abb. 3: Detaillierte Beschreibung eines Buches (Die Abbildung wurde aus den Beschreibungen mehrerer Bücher zusammengestellt um möglichst viele Informationen zu enthalten.)

(„redaktionelle Bewertung“ genannt). Es kann spekuliert werden, dass die professionellen Rezensionen durch Googles Crawler im Internet gefunden werden und automatisiert mit den Titeln in der Buchsuche verknüpft werden.

Unter der Überschrift „Stichwörter und wichtige Sätze“ werden Stichwörter aus dem Text angeboten. Es ist anzunehmen, dass nach einer Ermittlung der Worthäufigkeiten diejenigen Worte als Stichworte ausgewählt werden, die besonders häufig sind und einen besonders hohen Diskriminanzeffekt haben[4]. Zudem werden häufige Begriffe und Phrasen aus dem Buch unter der Überschrift „Common Terms and Phrases“ in Form einer *Tag-Cloud* dargestellt, d.h. proportional zur Häufigkeit des Vorkommens eines Wortes im Buch wird das Wort in größerer Schriftart dargestellt. Es kann spekuliert werden, dass beide Funktionen nur für Bücher zur Verfügung stehen, die von Google im Volltext digitalisiert wurden.

Bei etwa 79 Prozent der Bücher steht ein Inhaltsverzeichnis zur Verfügung. Damit sind unabhängig von den im Buchtext enthaltenen Inhaltsverzeichnissen solche Inhaltsverzeichnisse gemeint, die automatisch aus dem Buchtext ausgelesen werden und dann direkt in die beschreibende „Übersichts“-Seite integriert werden. Unter 5.5.2. wird darauf detaillierter eingegangen.

Weitere Informationen die angeboten werden sind:
- Links zu ähnlichen Büchern. Es werden maximal drei ähnliche Bücher mit Cover präsentiert, es kann aber über einen Link die Ansicht aller ähnlichen Bücher aufgerufen werden. Hier werden die gleichen Bücher angezeigt wie auch im linken Frame, sofern dort das entsprechende Element angezeigt wird.
- Verweise auf das angezeigte Buch aus anderen Büchern oder Zeitschriften innerhalb der Google Buchsuche, aus Medien innerhalb von Google Scholar oder von Webseiten.
- besonders „beliebte Passagen“, die zusammen mit der Angabe, in wie viel Büchern und in welchem Zeitraum die Passage zitiert wurde, angezeigt werden (diese Funktion steht bei Büchern, die nur im Auszug verfügbar sind, nicht zur Verfügung).

[4] Das Wort „Diskriminanzeffekt“ bezeichnet die Tatsache, dass sich seltene Wörter besser zur Beschreibung von Dokumentinhalten eignen als häufige. Besonders häufige Wörter werden deshalb oft auf Stopplisten erfasst und bei der Indexierung nicht berücksichtigt (vgl. STOCK 2007, S. 341).

- Links zu anderen Ausgaben desselben Textes, die ebenfalls bei Google Books digitalisiert wurden.
- „ausgewählte Seiten“, für die *Thumbnails* mit Link-Funktion angeboten werden, die direkt auf diese Seiten verweisen (hier werden besonders oft Seiten mit Grafiken, Bildern oder Illustrationen ausgewählt).
- grundlegende bibliographische Informationen (Titel, Autor, Verlag, Erscheinungsjahr, ISBN und Seitenzahl).

2.1.1.4. Funktionen zur Arbeit mit Buchtext

Betrachtet der Nutzer den Inhalt eines Buches und nicht die Informationen über das Buch, die auf der Übersichts-Seite zu finden sind, wird der Inhalt in dem größeren, rechten Frame dargestellt. Dem Nutzer stehen zahlreiche Funktionen zur Verfügung. Bücher, die nur in Form von Auszügen zugänglich sind, können nicht betrachtet werden.

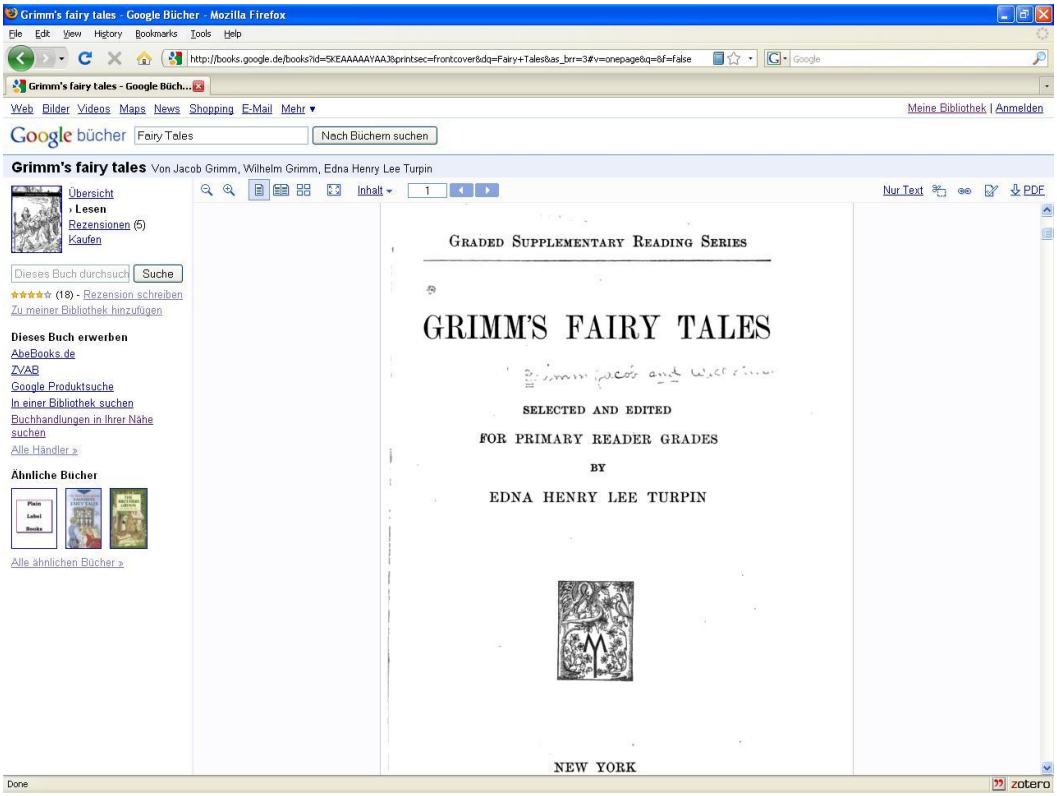


Abb. 4: Ansicht eines Buches

Bei Büchern, die eine eingeschränkte Vorschau erlauben, kann der Vergrößerungsgrad eingestellt werden und wahlweise können eine oder zwei Seiten gleichzeitig angezeigt werden. Es kann ein Vollbild-Modus aktiviert werden, der den linken Frame, die Kopfzeile und das Suchfeld ausblendet. Darüber hinaus kann mittels des Inhaltsver-

zeichnisses navigiert werden oder von Seite zu Seite geblättert werden. Alternativ kann man den Betrachtungsbereich mittels eines Scrollbalkens auch fließend verschieben, so dass die Unterhälfte der ersten und die Oberhälfte der folgenden Seite angezeigt wird.

Für Bücher, die im Volltext zugänglich sind, stehen weitere Funktionen, über die bisher benannten hinaus, zur Verfügung. Dazu gehört eine „nur Text“-Option, die das angezeigte Buch anstatt als Bilddatei als normalen Fließtext darstellt, der z.B. in die Windows-Zwischenablage kopiert werden kann. Es wird eine Option angeboten, mit der ein dauerhaft abrufbarer Link erzeugt werden kann, der auf die jeweils ausgewählte Seite im jeweils ausgewählten Buch verweist. Schließlich wird noch eine Funktion angeboten, die es ermöglicht, frei definierbare Ausschnitte einer Buchseite als Bilddatei aufrufbar zu machen. Dazu wird ein Link erzeugt, der von Nutzern in Webseiten eingebettet werden oder auch per Email verschickt werden kann.

Führt der Nutzer eine Suche innerhalb eines Buches aus, so wird im größeren Frame eine Trefferliste angezeigt. Die Trefferliste kann nach Relevanz oder aufsteigend nach Seitenzahl sortiert werden. Sie enthält Bilddateien, welche die einzelnen Fundstellen des Suchwortes, das durch gelbe Hinterlegung hervorgehoben wird, in ihrem größeren Textzusammenhang zeigen. Auf diesen Auszügen werden zwischen vier und acht Zeilen dargestellt. Bei Büchern, die im Volltext oder als erweiterte Vorschau zur Verfügung stehen, kann mittels eines Links die Seite, von welcher der Auszug stammt, aufgerufen werden. Auch hier ist das Suchwort gelb hervorgehoben.

Bei dieser Ansicht der Fundstellen nicht als Auszug, sondern innerhalb eines Buches, kann von Treffer zu Treffer gesprungen werden. Außerdem stehen alle anderen Funktionen der Buchbetrachtung zur Verfügung.

2.1.1.5. Persönliche Bibliothek

Von allen Webseiten innerhalb von Google Books lassen sich Bücher zur persönlichen Bibliothek hinzufügen. Diese Funktion steht dem Nutzer nur zur Verfügung, wenn er über einen Google-Account verfügt.

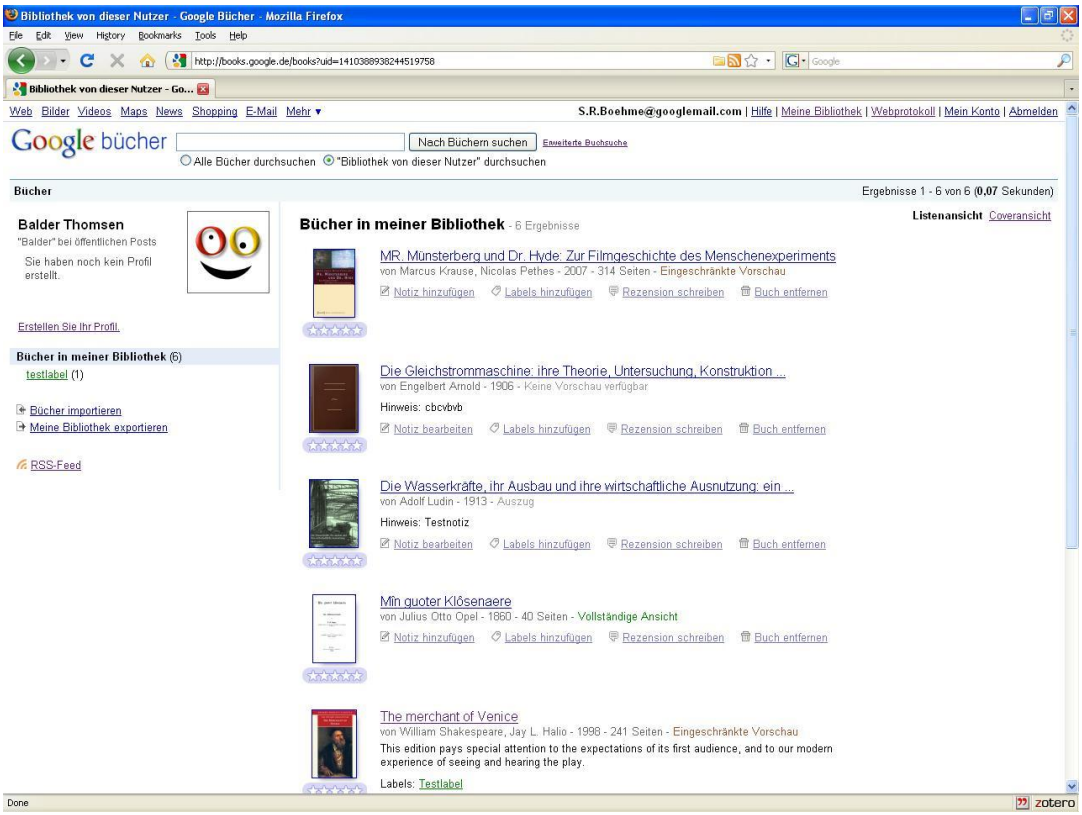


Abb. 5: Persönliche Bibliothek

Die Übersichtsseite der persönlichen Bibliothek ist vertikal in zwei Frames unterteilt. Im linken befindet sich das Menü, der größere rechte dient zur Darstellung des eigentlichen Inhalts der persönlichen Bibliothek. Hier werden die Bücher in der virtuellen Bibliothek in Form einer Liste oder in Form einer Cover-Ansicht aufgeführt, wie bei der Ergebnispräsentation nach einer Suche.

Die Listenform enthält mehr Informationen als die Cover-Ansicht und erlaubt auch das Arbeiten mit den Büchern. Es steht eine Funktion zur Bewertung der Bücher auf einer Skala von eins bis fünf, dargestellt durch Sterne, zur Verfügung. Die Nutzer können Kommentare oder Rezensionen zu den Büchern verfassen. Als Kommentar kann z.B. die Signatur des Buches in einer lokalen Bibliothek gespeichert werden. Die Rezensionen sind öffentlich, d.h. andere Nutzer, die das Buch betrachten, können sie lesen. Die Bücher in der persönlichen Bibliothek können mit Hilfe von Labeln organisiert werden. Jedes Buch kann mit mehreren virtuellen Aufklebern versehen werden. Der Nutzer kann Bücher so zu Gruppen zusammenfassen.

Beim Arbeiten mit der Buchsuche fällt auf, dass vielen Nutzern nicht bewusst zu sein scheint, dass Rezensionen öffentlich lesbare Bewertungen sind. Die entsprechende

Funktion wird genutzt, um Bücher mit privaten Kommentaren wie z.B. „S. 29 Da ist das Zitat von Weber enthalten“ (GOOGLE BOOKS REZENSION) zu versehen. Google bietet neben jeder Rezension einen Link an, um diese als „anstößig“ zu melden. Ein System, mit dem Rezensionen daraufhin bewertet werden können, ob sie hilfreich sind oder nicht, wäre eine sinnvolle Ergänzung. Ein solches System wird z.B. auf der Seite Amazon.com eingesetzt. Weiterhin fällt auf, dass nur 1,8 Prozent der Bücher von Nutzern rezensiert wurden (siehe dazu unter 5.5.2.). Hier ist sicher zu berücksichtigen, dass das Buchsuche-Angebot noch vergleichsweise jung ist.

Neben jeder Rezension wird der selbstgewählte Benutzername des Verfassers angezeigt. Durch Anklicken des Namens können andere Nutzer der Buchsuche die Bibliothek des Verfassers anschauen. Dort wird ihnen angezeigt, welche Bücher die Bibliothek enthält und welche anderen Bücher der Nutzer rezensiert hat (vgl. GOOGLE BOOKS NUTZER-BIBLIOTHEK). Dies erscheint aus datenschutzrechtlicher Sicht bedenklich, da die Nutzer der Buchsuche nicht darauf hingewiesen werden, dass „ihre Bibliothek“ öffentlich einsehbar ist. Beim Online-Buchhändler Amazon.com führte eine ähnliche Situation mit öffentlich einsehbaren Wunschzetteln zu Kritik (vgl. WALKER 2008).

Die gesamten Informationen der Bibliothek – bibliographische Daten der Bücher, Notizen, Label – können als XML-Datei exportiert bzw. lokal gespeichert werden und dann mit anderen Programmen verwendet werden. Jede persönliche Bibliothek steht zudem als RSS-Feed (Real Simple Syndication) zur Verfügung, und kann von Nutzern mit entsprechender Software (einem sog. Feedreader) abonniert werden. Diesen Nutzern wird dann unmittelbar mitgeteilt, wenn ein Buch zur Bibliothek hinzugefügt wird.

2.1.1.6. Offene Schnittstellen

Google bietet seinen Nutzern sog. APIs (Application Programming Interface) an, die es ermöglichen, in Google Books digitalisierte Bücher direkt in die eigene Webseite zu integrieren. Betrachter der Webseite des Kunden müssen sich so in keinster Weise mit Google Books beschäftigen, um es zu nutzen. Dieses Verfahren kommt auch bei Youtube, dem Google Angebot für Videos, zum Einsatz (vgl. GOOGLE BOOKS BLOG 2009). Weiterhin gibt es eine Programmierschnittstelle, die es Bibliotheken erlaubt, im OPAC direkt auf ein Buch in der Buchsuche zu verlinken. Dies macht Sinn, wenn in der Buchsuche zumindest eine eingeschränkte Vorschau möglich ist, der Nutzer des OPAC

kann sich so einen wesentlich präziseren Eindruck des Buches verschaffen (vgl. GOOGLEBLOG 2008a). Bibliotheken können sowohl auf das in der Buchsuche gespeicherte Cover zugreifen, als auch auf die Übersichts-Seite eines Buches (vgl. GBV-WIKI). Die entsprechende API von Google Books ist darauf angewiesen, dass entweder die ISBN oder die PPN übermittelt werden, um das Buch, das angezeigt werden soll, zu identifizieren. Bei der PPN (Pica Production Number) handelt es sich um eine von Anwendern der Bibliothekssoftware PICA vergebene Nummer, die Mediem eindeutig identifiziert.

Enthält eine Webseite einen Link zu einem Buch, das nur in den USA angesehen werden kann (dazu siehe unter 2.1.2.2.), wird außerhalb der USA statt des Buches eine Bilddatei mit dem Schriftzug *Image not available* angezeigt (z.B. hier der Fall: GOOGLE BOOKS BLOG 2009a).

2.1.2. Rechtliche Situation

Google ging bei der Zusammenarbeit mit Bibliotheken im Rahmen des Library Programms teilweise so vor, dass alle Bücher im Bestand (bzw. in einzelnen Sammlungen) eingescannt wurden, unabhängig davon, ob diese noch dem Urheberrecht unterlagen oder nicht. Dies war z.B. in Michigan der Fall (vgl. VISE 2007, S. 229). Bei einigen Bibliotheken, z.B. in Oxford oder München wurden aber auch nur die gemeinfreien Werke bzw. die Werke in *public domain* digitalisiert (vgl. MILNE 2005, S.26).

2.1.2.1. Bis zur Beilegung des Rechtsstreits mit der Authors Guild

Google plante, die digitalisierten Bücher, abhängig von ihrem rechtlichen Status, unterschiedlich darzustellen. Am unkompliziertesten ist die rechtliche Situation bei der Darstellung von gemeinfreien Büchern, da niemand Ansprüche darauf geltend machen kann. Google stellt diese Bücher im Volltext zur Verfügung, der Nutzer kann sie auf seinem PC speichern oder per Internet auf die auf Google-Servern gespeicherte Kopie zugreifen. Bei dieser Art des Zugriffs stehen dem Nutzer alle Funktionen um mit dem Buch zu arbeiten zur Verfügung (wie unter 2.1.1. beschrieben).

Bei Büchern, die noch dem Urheberrecht unterliegen, ist die rechtliche Situation komplizierter. Die Darstellungsart innerhalb der Buchsuche unterscheidet sich danach, auf welchem Weg diese Bücher in den Google-Index gelangten. Solche Bücher, die

digitalisiert wurden, weil sie sich im Bestand der am Bibliotheks-Programm teilnehmenden Bibliotheken befanden, werden nur in – maximal drei - Auszügen dargestellt. Diese Auszüge zeigen den Suchbegriff im Textzusammenhang von vier bis acht Zeilen und werden als Bilddateien dargestellt, so dass der Text nicht ohne weiteres kopiert werden kann. Hat der Nutzer keine Suche ausgeführt und betrachtet das Buch, ohne dass es als Ergebnis einer Suche angezeigt wird, wird ihm kein Inhalt des Buches angezeigt, es steht ihm nur die Übersichts-Seite mit Informationen zur Verfügung (vgl. GOOGLEBLOG 2008, GOOGLEBLOG 2005, GOOGLE BOOKS 2009a).

Bücher, die Verlage oder sonstige Rechteinhaber an Google lizensiert haben, werden in Form einer „eingeschränkten Vorschau“ dargestellt. Die Anzahl der Seiten, die bei solchen Büchern angezeigt werden kann, ist unterschiedlich und schwankt stark (vgl. DRUMMOND 2005, S. 2 ; GOOGLEBLOG 2008). Bei Suchanfragen werden die Fundstellen in unbegrenzt vielen Auszügen dargestellt, die Fundstellen können aber auch im Textzusammenhang der ganzen Seite angezeigt werden.

Allen drei bisher dargestellten Anzeigemodi gemein ist, dass die Bücher im Volltext digitalisiert und indexiert sind, auch wenn sie nicht im Volltext angezeigt werden.

Das dargestellte Vorgehen Googles bei der Digitalisierung in Bibliotheken, auch urheberrechtlich geschütztes Material zu digitalisieren, stellte nach Auffassung der amerikanischen Authors Guild, einer Autorenvereinigung die ca. 8000 Autoren vertritt, eine doppelte Urheberrechtsverletzung dar (vgl. BAND 2006, S. 12). Zum einen sei das Digitalisieren der Werke ungesetzlich, zum anderen das Anzeigen. Die Authors Guild reichte im September 2005 Klage gegen Google ein und forderte darin Schadensersatz sowie eine Unterlassungserklärung (vgl. KLAGESCHRIFT). [5]

Google berief sich als Grundlage des Handelns auf den *Fair-Use*-Paragraphen des amerikanischen Copyrights (vgl. GOOGLEBLOG 2005). Dieser gestattet bestimmte, ansonsten durch das Copyright untersagte Handlungen, wenn diese der öffentlichen Bildung dienen, keinen finanziellen Vorteil generieren und den Marktwert des intellektuellen Gutes nicht beschädigen (vgl. COPYRIGHT, BAND 2005). Ein ähnlich

[5] Eine Übersicht aller offiziellen Dokumente des Rechtsstreits findet sich unter http://news.justia.com/cases/featured/new-york/nysdce/1:2005cv08136/273913.

gelagerter Präzedenzfall, bei dem das verkleinerte Darstellen von urheberrechtlich geschützten Bildern im Internet ohne Genehmigung der Rechteinhaberin vom Gericht auf Grundlage des *Fair-Use*-Paragraphens für rechtens befunden wurde, wurde 2003 geschaffen (vgl. BAND 2006, S. 5 ; GERICHTSURTEIL).

Google kam den Verlagen im August, also noch bevor diese Klage einreichten, entgegen und bot zum einen ein *Opt-Out* genanntes Verfahren an und stellte zum anderen die Digitalisierung von urheberrechtlich geschützten Büchern vorläufig ein (vgl. BAND 2005, GOOGLEBLOG 2005). *Opt-Out* meint, dass die Verlage Google melden müssten, welche Bücher nicht in der Suche erscheinen sollen. Das Gegenteil wäre *Opt-In*, d.h. in der Suche würden nur Bücher angezeigt, deren Erscheinen der Verlag genehmigt hat. Bücher, welche durch den Rechteinhaber mittels *Opt-Out*-Meldung von der Digitalisierung ausgeschlossen wurden, werden von Google nicht im Volltext indexiert, bzw. das Indexat gelöscht. Dementsprechend sind sie nicht im Volltext durchsuchbar. Die in der Google Buchsuche verfügbaren Informationen beschränken sich für diese Bücher auf bibliographische Informationen. Die Authors Guild reichte trotz des Entgegenkommens Klage ein, Google setzte das Digitalisieren von dem Urheberrecht unterliegenden Büchern im November 2005 fort (vgl. GOOGLEBLOG 2005).

Weitere Rechtsstreite gab es in Belgien und Deutschland. In Belgien klagten Vereinigungen von Fotografen und Autoren gegen die Firma, was zu einer außergerichtlichen Einigung führte. In Deutschland war es die Wissenschaftliche Buchgesellschaft WBG die wegen Urheberrechtsverletzungen klagte. Die Klage wurde jedoch zurückgezogen, nachdem von Seiten der Google Inc. eidesstattlich versichert wurde, dass die Daten der Bücher aus dem WBG-Verlagsprogramm gelöscht worden waren und die retrospektive Digitalisierung von Büchern ausschließlich in den USA stattfinde. Deutsche Gerichte hätten demnach keine Zuständigkeit. Des Weiteren ließ das Hamburger Landgericht, vor dem die Verhandlung stattfand, im Vorfeld der Verhandlung verlauten, die von Google angezeigten Auszüge würden nicht die Schöpfungshöhe erreichen, die nötig ist, damit das Urheberrechtsgesetz Anwendung findet (vgl. COMPUTER UND RECHT 2006, S.1).

2.1.2.2. Rechtliche Situation nach der Einigung mit der Authors Guild

Im Oktober 2008 gab Google bekannt, dass eine Einigung im Rechtsstreit um die Book Search erzielt worden sei (vgl. GOOGLE 2008). Dieser Vergleich beendet sowohl die Klage der Authors Guild, als auch eine weitere Klage amerikanischer Verleger. Bei den Gerichtsverfahren handelte es sich um sogenannte *Class Action Lawsuits*, was in etwa einer deutschen Sammelklage entspricht. Bei dieser Art von Verfahren erklären sich die Kläger zum Stellvertreter einer größeren Gruppe (hier: Inhaber von Urheberrechten an Büchern, Zeitschriftenartikeln oder ähnlichem) (vgl. GOOGLE 2009b). Wird in einem *Class Action Lawsuit* eine Einigung erzielt, gilt diese für alle Mitglieder der Gruppe, auch diejenigen, die nicht am Gerichtsverfahren beteiligt waren. Deshalb tritt eine Einigung nicht sofort in Kraft. Bevor sie rechtswirksam wird, gibt es eine Frist, binnen derer Mitglieder der Gruppe erklären können, dass sie der Einigung nicht zustimmen. Die Einigung gilt dann nicht für sie, es steht ihnen offen weitere Gerichtsverfahren anzustrengen. Wer zur Gruppe gehört und nicht innerhalb der Frist widerspricht unterliegt der gerichtlichen Entscheidung. Ein Vergleich in einem *Class Action Lawsuit* muss von dem Gericht, vor dem das Gerichtsverfahren stattfand, genehmigt werden (vgl. EICHHOLTZ 2002, S. 29 ff). In diesem Fall erteilte der U.S. District Court for the Southern District of New York die Genehmigung (vgl. GOOGLE 2008, GOOGLE 2009c).

Die „bahnbrechende Vereinbarung“ (GOOGLE BOOKS STARTSEITE) führt zur Gründung einer neuen US-Behörde, der Book Rights Registry (von Google übersetzt mit „Registrierstelle“). Finanziert wird diese Behörde durch Zahlung von 34,5 Mio. US-Dollar durch Google. Diese Einrichtung soll die korrekte Verteilung eines Teils der Einnahmen der Google Book Search durchführen, d.h. sie soll Tantiemen-Zahlungen an Rechteinhaber weiterleiten. Dazu soll sie eine Datenbank führen, die Inhaber von Urheberrechten mit Adress- bzw. Kontaktdaten speichert. Die Registrierstelle soll ebenfalls dafür zuständig sein, dass Rechteinhaber bestimmte Arten der Nutzung zulassen oder ausschließen können (bspw. Anzeige als eingeschränkte Vorschau oder Auszug) (vgl. GOOGLE 2009d, S. 10 f.).

Die Vereinbarung sieht u.a. vor, dass Rechteinhabern, deren Bücher bis zum 5.5.2009 durch das Bibliotheks-Programm eingescannt wurden, für jedes Buch 60 US-Dollar gezahlt werden. Die Gesamtsumme dieser Zahlung darf lt. Vergleich 45 Mio. US-Dollar nicht unterschreiten. In der Vereinbarung ist auch ein genauer Verteilungsschlüssel für

die Gewinne definiert, die durch Google Books erzielt werden. Von den Einnahmen, die Google durch das Google Books Angebot erzielt, werden 63% an die Rechteinhaber ausgezahlt. Die Google Inc. erhält die verbleibenden 37%. Aus den 63% die an die Rechteinhaber fließen, finanziert sich auch der laufende Betrieb der Registrierstelle, die diese Zahlungen auch durchführt (vgl. GOOGLE 2009d, S.10).

Mit der Vereinbarung wird eine im rechtlichen Kontext neue Unterscheidung der urheberrechtlich geschützten Bücher eingeführt: im Handel lieferbar oder nicht lieferbar. Google bekommt in der Vereinbarung das Recht zugesprochen, Bücher, die nicht mehr lieferbar sind, innerhalb der Buchsuche als Volltext zu nutzen, also auch darzustellen. Der Rechteinhaber kann dieser Nutzung allerdings wiedersprechen. Lieferbare Bücher darf Google nicht nutzen, es sei denn der Rechteinhaber erteilt Google eine entsprechende Genehmigung. Google ist allerdings nicht dazu berechtigt, nicht-zahlenden Benutzern die Volltexte zugänglich zu machen. Diese bleiben zahlenden Nutzern vorbehalten. Nicht-zahlenden Benutzern stehen eingeschränkte Zugriffs-möglichkeiten wie z.B. die eingeschränkte Vorschau zur Verfügung (vgl. GOOGLE 2009d, S. 13).

Es wird geschätzt, dass ca. 4,75 Mio. Bücher von 7 Mio. Büchern im Bestand der Google Buchsuche „verwaist“ sind (vgl. BRANTLEY 2009). Damit ist gemeint, dass diese Bücher urheberrechtlich geschützt sind, aber niemand das Urheberrecht wahrnimmt. Zum Teil, so wird vermutet, ist das Interesse der Rechteinhaber erloschen, zum Teil sind diese verstorben und den Erben ist nicht bewusst, dass die Rechte an sie übergingen. Es ist anzunehmen, dass die überwiegende Mehrzahl der verwaisten Bücher im Handel nicht mehr lieferbar ist. Die Nutzung nicht mehr lieferbarer Bücher durch Google muss wie oben geschildert aktiv ausgeschlossen werden. Da es bei verwaisten Büchern niemand gibt, der dieses aktive Ausschließen vornehmen könnte, räumt die Vereinbarung Google das Recht ein, die Mehrheit der verwaisten Bücher zu nutzen. Hier soll darauf hingewie-sen werden, dass dieses Recht nur für die Google Inc. gilt und nicht für Mitbewerber (ebd.). Google bietet dementsprechend einen Monopol-Zugang zu verwaisten Werken an (bzw. zu der Mehrheit der verwaisten Werke, für die niemand der Nutzung wiederspricht). Weiter soll darauf hingewiesen werden, dass Google sich schon 2005 an einer Anhörung des Copyright Office beteiligte, um auf die Gesetzgebung zum Thema *Orphan Works* Einfluss zu nehmen (vgl. FEDERAL REGISTER 2005, DRUMMOND 2005).

Google hat laut der Vereinbarung das Recht, das Geschäftsmodell von Google Books entscheidend zu erweitern. Nutzungsarten, die ausdrücklich zugelassen sind, umfassen Abonnements für Institutionen, Verkäufe an Privatpersonen, und das Anzeigen von Werbung auf den Seiten der Buchsuche. Die Abonnements sind denen anderer kommerzieller Datenbankanbieter wie LexisNexis vergleichbar, eine Institution erwirbt eine Lizenz für ihre Mitglieder, z.B. eine Universität für die Studenten. Bei den Verkäufen an Privatpersonen wird zunächst nur die Lizenz zur Nutzung des Buches auf dem Google-Server verkauft. Allerdings sind in der Vereinbarung schon weitere Nutzungsmöglich-keiten angesprochen, die zwischen Google und der Registrierstelle ausgehandelt werden können, so z.B. der Verkauf von PDFs zum Herunterladen oder Abonnements für Privatpersonen. Letzteres Angebot kann als „Online-Bibliothek“ bezeichnet werden.

Die Sammelklage wurde vor einem amerikanischen Gericht eingereicht, dementsprech-end entfaltet auch der Vergleich seine Wirkung nur auf amerikanischem Boden. In allen anderen Ländern bleibt es bei den bisherigen Regelungen, die Buchsuche verändert sich nicht. Allerdings nehmen auch Nicht-Amerikaner, die das Urheberrecht an in den USA veröffentlichten Werken halten, am Vergleich teil. Das heißt, sie können sich für bereits digitalisierte Bücher entschädigen lassen und werden in Zukunft an den Einnahmen der Buchsuche beteiligt.

Mit dem Verkauf von Inhalten durch die Buchsuche übertritt Google zum ersten Mal die Grenze zwischen Suchmaschine und Anbieter von Inhalten im Internet (vgl. LEWANDOWSKI 2008). Es ist anzunehmen, dass die Nutzer eine gewisse Objektivität von Suchmaschinen erwarten. Ob diese angesichts der Vermischung von Rollen gegeben ist, ist ungewiss. Ob beim Ranking innerhalb der Buchsuche gewinnbringende Bücher bevorzugt werden, dürfte für den einzelnen Nutzer jedenfalls nicht feststellbar sein.

2.2. OPAC der Firma OCLC

Wird der OPAC im Browser aufgerufen, erscheint die Startseite. Diese ist, wie auch alle anderen Seiten des OPACs, in vier Bereiche geteilt. Am linken Bildschirmrand befindet sich die schmale Navigationsleiste, am oberen Rand die Kopfzeile. In dem Bereich, in dem sich Kopfzeile und Navigationsbereich überschneiden ist meist ein Logo platziert, das zugleich ein anklickbarer Link ist (vgl. Abb. 6). Dieser führt i.d.R. entweder auf die Homepage der Universität, die Homepage der Bibliothek oder zur Startseite des OPACs.

Die Kopfzeile enthält ebenfalls ein Navigationsmenü. Mit diesem kann auf die einzelnen Suchfunktionen zugegriffen werden. Diese sind: Einfache und erweiterte Suche, Anzeige der Suchergebnisse und der Zwischenablage sowie die „Hilfe“-Texte. Abhängig davon, ob die entsprechenden Software-Module in der jeweiligen Bibliothek installiert und konfiguriert sind, finden sich hier auch weitere Links auf die Verwaltung von Benutzerdaten (die Vormerkungen etc. ermöglicht), zu einem Anfrageformular mit dem die Bibliothek kontaktiert werden kann oder zu einer Liste weiterer Kataloge. Stehen die Menüs des Kataloges in anderen Sprachen als Deutsch zur Verfügung, sind die Grafiken der Fahnen, die üblicherweise das Vorhandensein des Textes in der entsprechenden Landessprache anzeigen, ebenfalls in der Kopfzeile platziert.

Welche Links in der Navigationsspalte angezeigt werden, ist von Bibliothek zu Bibliothek verschieden, in der Regel werden hier Links zum Speichern einer bereits durchgeführten Suche angeboten sowie zu andern Katalogen als dem gerade ausgewählten, bspw. dem Verbundkatalog, dem Katalog für elektronische Ressourcen oder einem systematischen Katalog. Einige Bibliotheken bieten hier auch die Möglichkeit zur Trefferanalyse an, die ermittelt, wie sich die Treffer auf Sprachen, Länder oder Materialarten verteilen.

2.2.1. Einfache Suche

Die Suchfunktionen des OPACs werden in der Kopfzeile angezeigt. Bei der einfachen Suche (auch als Standard-Suche bezeichnet) stehen unterschiedliche Suchfunktionen zur Verfügung. Wieder ist ausschlaggebend, welche Funktionen von den Bibliotheken konfiguriert wurden. Das Minimum stellt die Anzeige von drei Feldern dar (z.B. beim OPAC der Uni Kiel[6]): Ein Textfeld zur Eingabe des Suchbegriffes, ein Drop-Down-Menü zur Auswahl des Datenfeldes, das durchsucht werden soll (z.B. Autor oder Titel; Standardeinstellung ist meist „alle Wörter“ [was „alle Datenfelder“ meint]) und ein Drop-Down-Menü, um auszuwählen, mit welchem Operator die Suchwörter verknüpft werden sollen, falls mehrere eingegeben wurden. Hier stehen die klassischen Booleschen Operatoren (Und, Oder, Nicht [„Ausgenommen“ benannt]) zur Verfügung, aber auch die Möglichkeit, eine bereits durchgeführte Suche durch eingeben eines weiteren Wortes entweder einzuschränken oder zu erweitern. Schränkt der Nutzer eine Suche ein, entspricht das einer Und-Verknüpfung, erweitert er eine Suche, entspricht

[6] http://kiopc4.ub.uni-kiel.de:8080/DB=1/LNG=DU/

das einer Oder-Verknüpfung. Die Anzahl der Datenfelder, die durchsucht werden können, ist ebenfalls von der Konfiguration der einzelnen Bibliothek abhängig und schwankt zwischen sieben beim OPAC der SUB Bremen [7] und siebzehn beim OPAC der Uni Gießen[8].

Optionale Einstellungsmöglichkeiten für „einfache Suchanfragen“ sind z.B. die Sortierung der Ergebnisse, die Auswahl der Bibliothek in deren Bestand gesucht werden soll und das Aktivieren einer „unscharfen Suche“. Als Kriterien nach denen die Ergebnisse sortiert werden können, stehen die Relevanz, das Erscheinungsjahr sowie alphabetische Reihenfolgen von Autoren oder Titeln zur Verfügung. Als Standard-einstellung ist meistens das Erscheinungsjahr als Sortierkriterium gewählt. Die Auswahl der Bibliothek macht u.a. in zweischichtigen Bibliothekssystemen Sinn. Mit der Bezeichnung „unscharfe Suche“, wird ein Suchverfahren beschrieben, dass auch Zeichenketten als Ergebnis zurückliefert, die ein oder mehrere Zeichen von der eingegebenen Zeichenkette abweichen. Vorteil ist, dass auch andere Wortformen als die Eingegebene gefunden werden können, z.B. Plural statt Singular oder Akkusativ statt Nominativ. Der englische Begriff ist *Fuzzy Search* (vgl. JELE 2001, S. 46 f.).

2.2.2. Erweiterte Suche

Die erweiterte Suche bietet immer einen Suchfilter mit mehr Optionen als bei der einfachen Suche. Es können bis Eingaben aus bis zu vier Textfeldern mit den schon benannten Booleschen Operatoren verknüpft werden. Für jede Zeichenkette, die gesucht werden soll (also für jede Eingabe in einem Textfeld) kann definiert werden, in welchem Datenfeld der Titeldatensätze nach ihr gesucht werden soll. In den meisten OPACs kann die Ergebnismenge in Drop-Down Menüs nach Sprache oder Land der Veröffentlichung eingeschränkt werden, zudem steht ein Feld zur Einschränkung nach Veröffentlichungsdatum bereit. Durch das Setzen von Haken in Checkboxen kann die Suche auf Materialen bestimmter Art, wie etwa Bücher, Zeitschriften oder Musikalien eingegrenzt werden. Unscharfe Suche und Sortierkriterien stehen in derselben Form zur Verfügung wie bei der einfachen Suche.

[7] http://opac.suub.uni-bremen.de

[8] https://opac.uni-giessen.de

Bei einigen OPACs enthält die erweitere Suche über den umfangreicheren Suchfilter hinaus noch ein weiteres Angebot: das Browsen in der Systematik, bzw. in den Büchern die einer Stelle der Systematik zugeordnet wurden. Anscheinend können zwei Ebenen einer Systematik im System abgebildet werden.

2.2.3. Suchergebnisse

Bei der Darstellung der Suchergebnisse sind die Unterschiede zwischen den OAPCs der einzelnen Bibliotheken weniger groß als bei der einfachen Suche. In der Kurztitelanzeige, die als Standardeinstellung aktiviert ist, werden pro Seite zehn Treffer dargestellt. Es wird jeweils der Titel des Buches als Link angezeigt und darunter - wenn vorhanden - Autor, Auflage, Verlag, Verlagssitz und Jahr der Veröffentlichung (vgl. Abb. 6). Die Treffer sind nummeriert und neben jeder Nummer ist ein Symbol abgebildet, dass die Art der Quelle anzeigt. Möglich sind hier z.B. Bücher, Zeitschriftenartikel, Datenbankinhalte oder Filme. Am Ende jeder Seite der Liste von Kurztiteln sind Möglichkeiten untergebracht, mit denen auf die nächste und vorige Seite der Liste geblättert werden kann und zu einer bestimmten Seite in der Liste gesprungen werden kann. Außerdem wird dort anzeigt, wie viele Treffer welcher Suchbegriff erzielt hat (bei mehreren gesuchten Wörtern). So kann der Nutzer einschätzen, welcher der Suchbegriffe für eine evtl. zu geringe Treffermenge verantwortlich ist und den entsprechenden Suchbegriff aus der Suchanfrage streichen.

Statt der Kurztitelanzeige ist auch die Anzeige der Suchgeschichte möglich, die es erlaubt, die Treffermenge durch Kombination mehrerer Suchanfragen einzuschränken. Es ist auch möglich, sich keine Liste der Ergebnisse anzeigen zu lassen, sondern stattdessen nacheinander jeweils die vollständigen Titelinformationen eines einzelnen Treffers.

Bei der Darstellung aller Titelinformationen eines Treffers unterscheiden sich die Informationen von Bibliothek zu Bibliothek und von Treffer zu Treffer. Zu den typischerweise angezeigten Informationen gehören die vollständigen bibliographischen Daten mit Titel, Verfasserangabe, Verlag, Verlagssitz, Auflage, Seitenzahl, Jahr der Veröffentlichung etc. Aber auch bibliotheksspezifische Daten wie Schlagworte bzw. Schlagwortketten, Sachgebiete in der Klassifikation der Bibliothek, denen das Buch

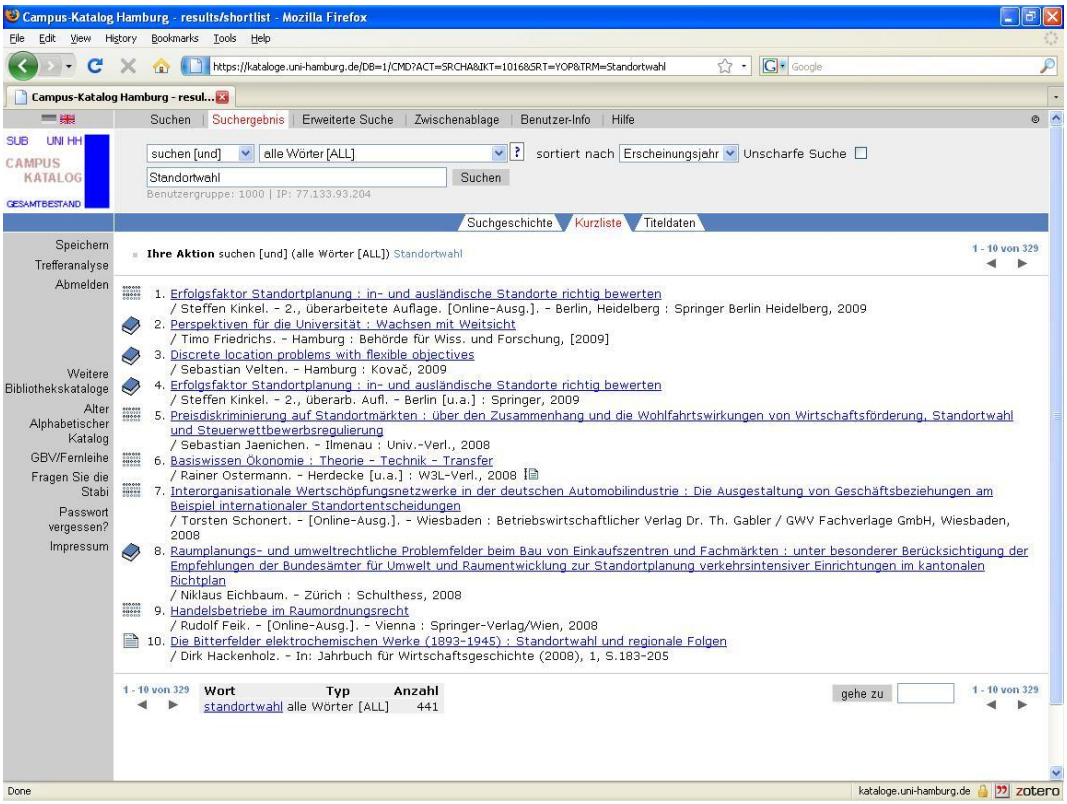


Abb. 6: Trefferliste im OPAC

zugeordnet wurde, die Signatur sowie Informationen zum Standort oder zu Ausleihmodalitäten und –Status werden meist angezeigt.

Die unter 2.1.1.3. schon angesprochenen *Catalogue-Enrichment-Features* Inhaltsverzeichnis, Titelblatt, Abstract, und Klappentext die bei der Darstellung von Büchern innerhalb der Book Search direkt in die entsprechende Seite eingebunden werden, stehen im OPAC lediglich als Link zur Verfügung. Die Links sind mit der URL (Uniform Resource Locator) zu der sie führen benannt. Eine Beschreibung, welcher Inhalt durch den Link abrufbar ist, steht in Klammern hinter dem Link. Hier werden z.T. englische Bezeichnungen wie *Publisher Description* oder *Contributors* verwendet, die vermutlich nicht jedem geläufig sind.

2.3. Catalogue Enrichment

Unter *Catalogue Enrichment* versteht man das Anreichern des Kataloges um zusätzliche Informationen über die bibliographischen Angaben hinaus. Dazu gehören Inhaltsverzeichnisse, Klappentexte, Bilder des Umschlags sowie Abstracts und Rezensionen

(vgl. LÖHRER 2007, S. 4). Diese sollen es dem Nutzer erleichtern, die Relevanz eines Buches zu bewerten.

Die Einbindung von Catalogue-Enrichment-Features in OPACs ist zunächst deshalb sinnvoll, weil sie es Nutzern erlaubt, sich bereits am OPAC ein präziseres Urteil über die im OPAC gefundene Quelle zu bilden. Dies ist besonders in Magazinbibliotheken relevant. Hier vergeht oft ein Tag, bis der Nutzer das bestellte Buch in der Hand hat und sich ein Urteil über dessen Nützlichkeit für ihn bilden kann. Wird er in die Lage versetzt, schon während der Suche die Relevanz beurteilen zu können, ist das sowohl für ihn ein deutlicher Vorteil, als auch für das Bibliothekspersonal, das entlastet wird (vgl. LÖHRER 2007, S. 6).

Besonders groß dürfte der durch Kataloganreicherung erreichte Vorteil bei Sammelwerken und Zeitschriften sein. Der Anteil von Sammelwerken am Bestand von Hochschulbibliotheken beträgt zwischen zehn und 20 Prozent (vgl. POULSEN 1996, S. 137). Diese waren bisher oft nicht tief erschlossen, d.h. es waren nur die bibliographischen Daten *des Werks* verfügbar, nicht jedoch Angaben zum Inhalt, zu den unselbständig erschienenen Texten *in den Werken* (vgl. LÖHRER 2007, S. 14). Können dem Nutzer nun mehr Zugriffspunkte (z.B. die Inhaltsverzeichnisse der einzelnen Zeitschriftenhefte oder die Titel der einzelnen Aufsätze in einem Sammelband) zugänglich gemacht werden, kann dieser sich nicht nur ein besseres Urteil bilden, sondern es erhöht sich auch die Wahrscheinlichkeit dafür, dass das Buch genutzt wird, da es öfter gefunden wird (vgl. MORRIS 2001).

Zum anderen ist die Kataloganreicherung auch deshalb sinnvoll, weil sie die Anwendung von Suchmaschinen-Technik in OPACs erlaubt. Zwar ist anzunehmen, dass in der Mehrheit der Bibliotheken derzeit noch nicht geplant wird, Suchmaschinentechnik zu verwenden, doch ebenso ist anzunehmen, dass früher oder später in der Mehrheit der (wissenschaftlichen) Bibliotheken Suchmaschinentechnik eingesetzt werden wird[9]. Möglichst viele inhaltsbeschreibende Texte in den OPAC zu integrieren ist von Vorteil, da so die Menge an indexierbarem Material erhöht wird. Dies ist nicht nur relevant weil die Inhalte so ausführlicher beschrieben werden, sondern auch, weil so eine größere

[9] Pilotprojekte gibt es derzeit z.B. in Bremen (elib genannt) und Karlsruhe (xOpac genannt). Eine Übersicht solcher Projekte findet sich im Blog des Beluga-Projektes an der SUB Hamburg.

Datengrundlage für einen Ranking-Algorithmus zur Verfügung steht (vgl. LEWANDOWSKI 2006). Je größer die Datengrundlage, die einem Ranking-Algorithmus zur Verfügung steht, desto präziser ist das Ranking.

In Deutschland arbeiten die regionalen Bibliotheksverbünde bei der Kataloganreicherung miteinander zusammen und kooperieren auch mit der Deutschen Nationalbibliothek und dem Börsenverein des deutschen Buchhandels. Die Abstimmung erfolgt unter anderem im Rahmen einer Arbeitsgruppe der Arbeitsgemeinschaft der Verbundsysteme (vgl. BEHRENS-NEUMANN 2008).Ziel der zahlreichen Kooperationen ist es, die Anzahl der Inhaltsverzeichnisse, die doppelt – also von zwei Bibliotheken – eingescannt werden, zu minimieren. Darüber hinaus soll erreicht werden, dass der Zugriff auf die Daten anderer Verbünde nicht nur in Form von Links erfolgen kann, sondern dass die Digitalisate auch zur Verfügung stehen, um im Volltext indexiert werden zu können. Für diese Arbeit sind die Kooperationen insoweit relevant, als dass nicht nur die Häufigkeit, mit der Kataloganreicherungs-Merkmale im OPAC der SUB Hamburg vorhanden sind, untersucht wird. Der OPAC der SUB kann als stellvertretend für die OPACs der Bibliotheken des GBV zum jetzigen Zeitpunkt betrachtet werden.

3. Ziele der Untersuchung

In dieser Untersuchung sollen Google Books und der OPAC der SUB Hamburg auf die Retrievalqualität hin verglichen werden, d.h. es wird die Qualität der Trefferbeschreibungen der Ergebnisse von Suchanfragen bewertet und verglichen. Bei dem durchgeführten Vergleich handelt es sich also um einen indirekten Vergleich, d.h. der Vergleich bezieht sich nicht auf die Systeme selbst, sondern auf die Ergebnisse die sie liefern (vgl. WOMSER-HACKER 2004, S. 227). Es wird von zwei verschiedenen Nutzungsmodellen, die jeweils das Verhalten bestimmter Nutzergruppen abbilden sollen, ausgegangen. Die Qualität wird durch eine Relevanzbewertung ermittelt, die durch Evaluatoren durchgeführt wird. Neben der Relevanz der Ergebnisse sollen weitere Aspekte der beiden Angebote bewertet werden, um eine umfassendere Beurteilung der Retrievalqualität der beiden Angebote zu erlauben. So soll die Frage beantwortet werden, ob Google Books qualitativ mit konventionellen Bibliotheken vergleichbar ist.

Der umfangreichste Teil der Studie betrifft die Untersuchung von thematischen Suchanfragen. Bei diesen ist der Nutzer nicht auf der Suche nach einem bestimmten Buch, sondern möchte sich einen Überblick über Bücher zu einem Thema verschaffen. Es werden 50 Suchanfragen in beiden Katalogen ausgeführt und jeweils die ersten zehn Treffer auf ihre Relevanz hin bewertet. Diese Relevanzuntersuchung wird jedoch nur für dem Nutzer direkt zugängliche Werke durchgeführt, denn bei dem ersten angenommen Nutzerverhalten in dieser Studie wird davon ausgegangen, dass Nutzer wenn sie nach einem Buch suchen, nicht nur Informationen darüber möchten, sondern auch den Zugriff darauf.

Deshalb liegt es nahe, hier ebenfalls zu ermitteln, wie hoch der Anteil an tatsächlich zugänglichen Büchern an der Gesamtheit der untersuchten Treffer ist. Für die Treffer, die bei der Suche im Bibliothekskatalog erzielt wurden, bedeutet das, den Anteil an Büchern zu ermitteln, die der Nutzer entleihen oder wenigstens in der Bibliothek einsehen kann (die also weder entliehen sind noch dem Nutzer durch Bearbeitung o.ä. nicht zur Verfügung stehen).

Als Entsprechung für dem Nutzer zugängliche Medien, werden in dieser Studie Medien angesehen, die in der Google Buchsuche entweder als Vollansicht oder als eingeschränkte Vorschau verfügbar sind. Bei Volltexten hängt es nur vom Inhalt des Buches

ab, ob der Informationsbedarf des Nutzers befriedigt werden kann. Bei Büchern, die in Form einer eingeschränkten Vorschau verfügbar sind, kann neben dem Inhalt des Buches auch die Art der Zugänglichkeit darüber entscheiden, ob der Nutzer seinen Informationsbedarf befriedigen kann. Da mehrere Seiten des Textzusammenhangs, um die gesuchten Textstellen herum, angezeigt werden können, besteht die Möglichkeit dazu. Es ist unbekannt, wie häufig eine eingeschränkte Vorschau tatsächlich ausreicht, um dem Nutzer die gesuchten Informationen zur Verfügung zu stellen. In dieser Untersuchung wird davon ausgegangen, dass der Informationsbedarf der Nutzer auch bei Büchern mit eingeschränkter Vorschau befriedigt werden kann.

Bei Google können Suchanfragen auf solche Treffer beschränkt werden, die in diesen Ansichten vorliegen. Um den Anteil der zugänglichen Bücher an der Gesamtheit der in der Buchsuche zur Verfügung gestellten Bücher zu ermitteln, wurden die thematischen Suchanfragen ein zweites Mal gestellt und alle Treffer zugelassen, unabhängig davon, auf welche Art sie zugänglich sind. Der Anteil der unmittelbar zugänglichen Werke in der Buchsuche ist auch unabhängig vom Vergleich mit dem OPAC interessant, da er bisher nur geschätzt worden ist.

Da sich das Verhalten bei der Informationssuche dahingehend verändert, das zunehmend auf Online-Quellen statt auf analoge Medien zugegriffen wird, soll dies im zweiten modellierten Nutzerverhalten der Studie berücksichtigt werden (vgl. KLATT 2001). Hier wird davon ausgegangen, dass Nutzer nicht bereit sind, den Weg in die Bibliothek zurückzulegen, um ihr Informationsbedürfnis zu befriedigen. Sie möchten also nicht nur Zugriff auf die Medien, wie im ersten Verhaltensmodell, sondern sie möchten Online-Zugriff. Es wird angenommen, dass sie sich mit direkt am heimischen PC verfügbaren Quellen zufriedengeben. Deshalb wird der Anteil an Treffern aus dem OPAC ermittelt, die online verfügbar sind. Zudem wird die durchschnittliche Relevanz dieser Quellen ermittelt.

Kostenpflichtige Datenbanken kommerzieller Anbieter wie z.B. LexisNexis oder SpringerLink können von Bibliotheksnutzern meist nur aus dem Rechner-Netz der Universität (oder sonstigen Institution) genutzt werden. Grund dafür ist, dass die Universitäten Abonnements für die Datenbanken erwerben. Technisch wird das umgesetzt, indem die IP-Adressen der Universität für die Benutzung der Datenbank

freigeschaltet werden. Eine Methode, den Nutzern auch von Rechnern außerhalb der Bibliothek bzw. des Universitätsnetzes Zugang zu den erworbenen Angeboten zu gewähren, ist die VPN-Technik (Virtual Private Network). Dabei kann z.B. der Privat-PC des Anwenders simulieren, Teil des Universitäts-Netzwerks zu sein, ohne dass dies der Fall ist. Für diese Studie ist diese Technik insofern relevant, als dass davon ausgegangen wird, dass die Nutzer beim zweiten gedachten Nutzerverhalten ihre Rechner entsprechend konfiguriert haben und alle im OPAC nachgewiesenen Quellen vom heimischen PC aus nutzen können.

Die bei thematischen Suchanfragen erzielten Ergebnisse sollen auf ihre Qualität hin untersucht werden, aber auch unter anderen Gesichtspunkten. So soll zum einen das Alter der Treffer ermittelt und verglichen werden. Die Hypothese ist, dass zahlreiche Volltexte der Buchsuche veraltet sind. Der Altersdurchschnitt müsste also höher sein, als bei den Treffern aus dem OPAC. Zum anderen soll die Häufigkeit ermittelt werden, mit der die Kataloganreicherungen Inhaltsverzeichnis, Abstract, Cover oder Klappentext (bei Google „Buchbeschreibung“ genannt) verfügbar sind, da diese dem Nutzer helfen können, sich ein präzises Bild von dem beschriebenen Treffer zu machen. Vermutlich ist die Diskrepanz zwischen einer Relevanzbewertung der Trefferbeschreibungen und einer Bewertung der Treffer selbst, bei dem Katalog am geringsten, der den Nutzern die meisten Informationen (und Kataloganreicherungselemente) anbietet. Dass es eine solche Diskrepanz zumindest bei einigen Internet-Suchmaschinen gibt, ist nachgewiesen (vgl. LEWANDOWSKI 2008a).

Weniger aufwändig als die Untersuchung von thematischen Suchanfragen ist die Untersuchung von 70 Anfragen, mit denen gezielt bestimmte Bücher gefunden werden sollen, (so genannte *Known-Item-Search*). Hier wird ermittelt, ob die gesuchten Bücher gefunden wurden, und auf welchem Platz der Trefferliste sie sich befanden. Deshalb ist der Einsatz von Evaluatoren nicht erforderlich.

4. Methodik

Die verwendete Methodik findet sonst im Bereich der Suchmaschinen-Forschung Anwendung. Dort ist es üblich, die Relevanz von Treffern, die von verschiedenen Suchmaschinen zu denselben Suchbegriffen gefunden wurden, miteinander zu vergleichen. So wird die Präzision der einzelnen Suchmaschinen ermittelt und die Suchmaschinen können miteinander verglichen werden. Die Suchanfragen werden meist aus Logfiles von Suchmaschinen ausgewählt.

Diese Methodik kann für den hier durchgeführten Vergleich übernommen werden, da auch OPACs Suchmaschinen sind. Bisher ist noch keine derartige Studie durchgeführt worden. Zwar werden im OPAC für die meisten Medien nur Meta-Informationen angezeigt, doch dies stellt kein Hindernis dar, denn hier werden auch bei der Bewertung der Relevanz der Treffer aus der Buchsuche nur die Meta-Informationen verwendet. Der Ansatz, Trefferbeschreibungen zu evaluieren wurde schon in LEWANDOWSKI 2008a verfolgt. Die hier untersuchten Systeme unterscheiden sich insofern voneinander, als dass der Bestand im OPAC ausschließlich über bibliographische Daten bzw. Meta-Informationen erschlossen ist, in der Buchsuche zum Teil darüber hinaus auch durch eine Volltextindexierung (vgl. LEWANDOWSKI 2005, S. 68 ff.).

Entwickelt wurde die Methodik zum Vergleich von Suchmaschinen im Jahr 2000 in einer Studie von Hawking et al. die auf einem Artikel von Gordon und Pathak aufbauten (vgl. HAWKING 2001). Dort wurde Methodik der TREC 8, die 1999 stattfand, für die Evaluation von öffentlichen Suchmaschinen im Internet modifiziert. TREC steht für Text REtrieval Conference. Diese Konferenzen werden seit 1992 in unregelmäßigen Abständen abgehalten. Sinn und Zweck der Konferenzen ist es, Methoden zur Evaluation von Text-Information-Retrieval-Systemen zu entwickeln. Die Teilnehmer sind dementsprechend Institutionen die solche Systeme betreiben (vgl. VOORHEES 2005, S. 3 ff.).

Da Hawking et al. die Methodik zur Untersuchung von Suchmaschinen im Internet begründeten, wird ihr Aufsatz im Folgenden bei der Diskussion der einzelnen Aspekte der Methodik berücksichtigt. Sie gaben in ihrem Aufsatz Anregungen für zukünftige Studien, die hier, soweit möglich, umgesetzt werden sollen (vgl. HAWKING 2001, S. 22 ff.). Tague-Sutcliffe verfasste 1992 einen Überblicksartikel über die Evaluation von

Information-Retrieval-Systemen im Allgemeinen, auch er wird im Folgenden berücksichtigt. In den folgenden Gliederungspunkten wird, über die beiden genannten Artikel hinaus, auf Veröffentlichungen eingegangen, die einzelne Aspekte detaillierter behandeln oder neueren Datums sind.

4.1. Auswahl der Bibliothek

Für die Untersuchung sollte eine Bibliothek bzw. ein Bibliothekssystem ausgewählt werden, das als stellvertretend für die Allgemeinheit der deutschen Hochschulbibliotheken angesehen werden kann. Es wurde davon ausgegangen, dass dies dann gewährleistet ist, wenn es sich um die Bibliothek einer Universität mit einem breiten Fächerprofil handelt. Dies schloss technische Universitäten, aber auch andere spezialisierte Hochschulen wie Kunsthochschulen aus. Die Eignung als „Stellvertreter" für die deutschen Hochschulen wird bei den meisten Universitätsbibliotheken insofern beeinträchtigt, als dass sie am Literaturerwerbungsprogramm der Deutschen Forschungsgemeinschaft (DFG) teilnehmen.

Die DFG fördert die Literaturerwerbung in der Bundesrepublik mit der Zielsetzung, jedes relevante Buch bundesweit mindestens einmal angeschafft zu haben. Dazu wurden sog. thematische Sondersammelgebiete gebildet und unter vierzig deutschen Hochschulbibliotheken aufgeteilt (vgl. DFG 2009, SUB 2006). Die Bibliotheken, die für ein Sondersammelgebiet zuständig sind, haben in diesem Bereich einen außergewöhnlich umfangreichen Bestand. Die ausgewählte Bibliothek soll stellvertretend für alle Bibliotheken stehen, weshalb keine Suchanfragen aus thematischen Bereichen ausgewählt werden sollen, in denen sie einen deutlich überdurchschnittlichen Bestand hat, wie dies bei den Sondersammelgebieten der Fall ist.

Für die Auswahl der Bibliothek wurden die Bestandszahlen von Bibliothekssystemen einiger großer deutscher Hochschulen verglichen (vgl. Tabelle 1). Zur Beurteilung der Größe wurde auf die Anzahl der an der Hochschule eingeschriebenen Studierenden zurückgegriffen (vgl. STATISTISCHES BUNDESAMT 2009, S. 30 ff.). Als Auswahl für diese Studie kamen Bibliotheken mit einem durchschnittlich großen Bestand und einem OPAC der Firma OCLC-PICA in Frage (siehe dazu 4.2.). Beim Vergleich wurden der Einfachheit

halber lediglich gedruckte Medien berücksichtigt, wie auf den Internetseiten der Bibliotheken angegeben.

Name der Hochschule	Bücher im Bestand der Hochschulbibliothek	Anzahl der Studierenden im Wintersemester 2008/2009
Freie Universität Berlin	8,5 Mio. Medien	30.439
Friedrich-Alexander-Universität Erlangen-	5,5 Mio. Medien	25.036
Ruhr-Universität Bochum	1,6 Mio. Medien	31.007
Universität zu Köln	3,7 Mio. Medien	41.765
Universität Duisburg-Essen	2,6 Mio. Bände	31.288
Christian-Albrechts-Universität Kiel	4,6 Mio. Bände	22.141
Justus-Liebig-Universität Gießen	3,8 Mio. Bände (Bücher und Medien)	22.974
Georg-August-Universität Göttingen	4 Mio. Bände	23.086
Universität Hamburg	3 Mio. Bände	36.112
Johannes Gutenberg Universität Mainz	2,8 Mio. Bücher	36.338
	Durchschnitt: 4 Mio.	

Tabelle 1: Bestandsgrößen von Hochschulbibliotheken. Quellen: INTERNETSEITEN DER BIBLIOTHEKEN

4.2. Auswahl des OPACs

Die Auswahl der zu untersuchenden Bibliothek erfolgte über das inhaltliche Profil hinaus auch nach der verwendeten Bibliotheks-Software. Diese verwaltet nicht nur Titelnachweise und Nutzerkonten sondern stellt auch die OPAC-Schnittstelle zur Verfügung (vgl. JELE 2001, S. 57 ff.). Da es mehrere Hersteller von Bibliothekssoftware gibt, nutzen deutsche Bibliotheken verschiedene OPAC-Produkte. Neben den Produkten der etablierten Hersteller gibt es auch Eigenentwicklungen ambitionierter Hochschulen, z.B. die Bremer elib oder den Karlsruher xOPAC (vgl. ELIB 2009; DIEROLF 2007, S. 130 ff.). Für einen umfassenden Vergleich zwischen Google Books und den wissenschaftlichen Bibliotheken in Deutschland als Ganzes, müssten alle relevanten Produkte der einzelnen Hersteller in den Vergleich mit einbezogen werden. Dies würde hier den Rahmen sprengen, deshalb beschränkt sich der Vergleich auf ein maßgebliches OPAC-System.

Die Bibliotheken der Bundesrepublik sind in geographisch begrenzte Verbünde zusammengeschlossen, in den meisten Fällen wird innerhalb eines Verbundes nur Software eines Herstellers genutzt. Es wird angenommen, dass die Bedeutung eines Bibliotheksverbundes mit der Größe des Verbundes zusammenhängt, und dass wiederum die Bedeutung der Software-Systeme davon abhängt in welchen Bibliotheken bzw. Verbünden sie eingesetzt werden. Der – nach Medieneinheiten – größte Verbund ist der Gemeinsame Bibliotheksverbund GBV (vgl. INTERNETSEITEN DER VERBÄNDE). Der GBV und der Verbund HeBIS (hessisches Bibliotheksinformationssystem) nutzen die Software der Firma OCLC-PICA. Sie kommt in mehr als 200 Hochschulbibliotheken zum Einsatz (vgl. OCLC 2009). Für den Vergleich in dieser Untersuchung soll deshalb das OPAC-System der Firma OCLC verwendet werden.

Die verbleibenden Verbünde sind das hbz (Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen), das die Software Aleph 500 des Herstellers Ex Libris nutzt. Diese Software bzw. deren Nachfolger MetaLib kommt auch im BVB (Bibliotheksverbund Bayern) sowie überwiegend auch im KoBV (Kooperativer Bibliotheksverbund Berlin-Brandenburg) zum Einsatz. Der SWB (Südwestdeutscher Bibliotheksverbund) nutzt sowohl die Software Horizon der Firma SirsiDynix als auch Software von PICA (vgl. OCLC 2004, BIEHLER 2005, S. 19).

4.3. Entscheidung für eine Bibliothek

Für die Untersuchung wurde die Staats- und Universitätsbibliothek Hamburg ausgewählt. Hier sind die gegebenen Bedingungen erfüllt, d.h. es wird Bibliotheks-Software vom Hersteller OCLC-PICA verwendet und die Größe des Bestandes entspricht in etwa dem Durchschnitt der betrachteten Bibliotheken. An der Bibliothek werden die folgenden Sondersammelgebiete betreut: Politik und Friedensforschung, Verwaltungs-wissenschaften, Spanien und Portugal, Indianer- und Eskimosprachen und –Kulturen, Küsten- und Hochseefischerei, Hamburg (vgl. SUB 2006).

4.4. Auswahl der Suchanfragen

Um bei der Untersuchung der Retrievalqualität möglichst aussagekräftige Ergebnisse zu erzielen, soll in der Untersuchung das reale Nutzerverhalten im OPAC so weit möglich abgebildet werden. Das Nutzerverhalten im OPAC dient deshalb als Orientierung, weil

untersucht werden soll, in wie weit die Buchsuche qualitativ dem Angebot entspricht, dass Nutzer beim Gebrauch des OPACs geboten bekommen. Allerdings schränkt die geringe Anzahl an Suchanfragen, die im Rahmen dieser Arbeit durchgeführt werden kann, die Möglichkeiten zur Simulation realen Verhaltens ein.

4.4.1. Literatur zum Nutzerverhalten im OPAC

Wertet man vorhandene Studien zum Verhalten von Nutzern beim Gebrauch von OPACs aus, so zeigt sich ein einheitliches Bild im Bezug auf die Verteilung der Suchanfragen auf die einzelnen Datenfelder wie z.B. Autor oder Titel. In Tabelle 2 sind die Werte einiger aktueller Studien, welche die entsprechende Verteilung ermittelten, dargestellt.

	Titel	**Autor**	**Schlagwort**	**Sonstige Felder**
HENNIES 2006[10]	53	40	17	5
REMUS 2002[11]	32,9	49,2	14,4	3,5
OBERMEIER 1999 [10]	43	45	14	11,6

Tabelle 2: Häufigkeit mit der die einzelnen Datenfelder zur Suche im OPAC genutzt werden

Wie schon erwähnt, ist es weder üblich noch sinnvoll, bei *Known-Item-Searches* die Relevanz der gesuchten Objekte zu bewerten. Einzig die Fragestellungen, ob das gesuchte Item gefunden wurde und manchmal auch, auf welcher Position es gefunden wurde sind von Bedeutung (vgl. HAWKING 2001, STOCK 2000). Als *Known-Item-Searches* sind Suchanfragen einzuordnen, bei denen nach Autoren oder Titel der gesuchten Werke gesucht wird. Solche Suchanfragen machen lt. der o.g. Studien mehr als 80% der Anfragen aus. Zwar können auch Suchanfragen nach Verleger, Signatur o.ä. eine *Known-Item-Anfrage* darstellen, doch diese werden nur sehr selten an OPACs gestellt. Vermutlich finden sie auch aus diesem Grund in der Literatur im Zusammenhang mit *Known-Item-Anfragen* wenig Beachtung (vgl. LEE 2006, S. 12). Da eine Relevanz-bewertung nur für thematische Suchanfragen sinnvoll ist, wurde entsprechend verfahren und in die Liste der Suchbegriffe nur thematische Suchanfragen aufgenommen. D.h. die Verteilung der Suchanfragen auf die einzelnen Datenfelder wurde im Set der verwendeten Suchanfragen nicht abgebildet.

[10] Die Summe der Zahlen ist größer als 100, da bei einigen Suchanfragen mehr als ein Datenfeld durchsucht wurde

[11] Zahlen beziehen sich auf die Standardsuche im Web-OPAC, vgl. REMUS 2002, S. 88

Hennies ermittelte ebenfalls die Häufigkeit, mit der Suchanfragen eine bestimmte Anzahl an Wörtern hatten. Diese Erkenntnisse können in dieser Studie verwendet werden und sind in nachfolgender Tabelle 3 aufgeführt.

Anzahl an Wörtern	**Anteil an Gesamtheit**
1	42 %
2	26,4 %
3	16,6 %
4	7,9 %
5 oder mehr	6,8 %

Tabelle 3: Worthäufigkeiten im OPAC

Weiterhin soll erwähnt werden, dass diese Studie versucht, die Retrievalqualität für den typischen Nutzer zu ermitteln. Dieser unterscheidet sich vom fortgeschrittenen Nutzer dadurch, dass er die Voreinstellungen der Internetseite, die er verwendet, nicht verändert. Hennies stellte fest, dass bei mehr als 85% der Suchanfragen die voreingestellten Optionen (Suchfelder, etc.) nicht verändert werden (vgl. HENNIES 2006, S. 7). Dies erklärt auch die seltene Nutzung der Expertensuche. In der Studie von Hennies wurden 1,6% der Suchanfragen von der entsprechenden Seite aus ausgeführt, in der Untersuchung von Remus waren es 3 % (vgl. ebd.; REMUS 2002, S. 111). Diese und weitere Studien zeigen auch, dass nur sehr wenige Nutzer mit Booleschen Operatoren arbeiten (vgl. MALLIARI 2007, S. 11). Für diese Studie bedeutet das, dass die Suchbegriffe in das entsprechende Feld auf der Startseite eingetippt wurden und nur wenige Änderungen an den Einstellungen vorgenommen wurden. Die hier erzielte Retrievalqualität liegt also vermutlich unter der optimalen Retrievalqualität, die durch die Verwendung der „Erweiterten Suche“ erzielt werden könnte.

Die einzigen Änderungen die gemacht wurden, waren nötig, um Vergleichbarkeit herzustellen. So wurde bei Google der Ansichtsmodus „eingeschränkte Vorschau und vollständige Ansicht“ ausgewählt und im OPAC die Sortierung nach Relevanz. Den Ansichtsmodus bei Google zu ändern, war nötig, um die Annahme, dass der Nutzer ausschließlich zugängliche Werke sucht, umzusetzen. Zudem war sie nötig, um die Vergleichbarkeit zwischen Google und dem OPAC herzustellen, denn in dem OPAC einer Bibliothek sind ausschließlich die dort zugänglichen Werke nachgewiesen.

4.4.2. Auswahl der Suchanfragen aus dem Logfile

Es wurden manuell thematische Suchanfragen aus einem Logfile des OPACs der Universitätsbibliothek Frankfurt am Main ausgewählt. Ein beispielhafter Auszug aus diesem Logfile findet sich im Anhang. Es wurden Suchanfragen, die an mehreren Tagen im Juni 2009 gestellt wurden, auf ihre Eignung hin untersucht. In Betracht gezogen wurden nur orthographisch korrekte Suchanfragen, so dass die Rechtsschreibkorrektur von Google, die mit den Worten „meinten Sie“ alternative Schreibweisen vorschlägt, keinen Einfluss auf die Ergebnisse hatte. Auch eine Liste der ausgewählten Suchanfragen findet sich im Anhang.

4.4.2.1 Worthäufigkeiten

Hennies berücksichtigte bei der von ihm durchgeführten Ermittlung von Worthäufigkeiten alle untersuchten Suchanfragen (zumindest berichtet er von keiner Einschränkung, vgl. HENNIES 2006, S.7). Hier sollen jedoch nur thematische Suchanfragen gestellt werden. Bei der Auswertung des Logfiles fiel auf, dass der Anteil an thematischen Suchanfragen mit drei oder mehr Wörtern äußerst gering war, und deutlich unter den von Hennies ermittelten Werten lag. Hier kann vermutet werden, das Suchanfragen mit einer hohen Anzahl an Wörtern vor allem Anfragen sind, bei denen der komplette Titel des gesuchten Buches eingegeben wird. Deshalb sollen weniger Suchanfragen mit drei oder vier Wörtern in die Liste der untersuchten Begriffe aufgenommen werden, als dies statistisch der Fall sein müsste. Thematische Suchanfragen mit fünf Wörtern wurden ignoriert, da im untersuchten Teil des Logfiles keine protokolliert war. Die folgende Tabelle zeigt in den ersten beiden Spalten die statistischen Häufigkeiten in dem von Hennies untersuchten Logfile und die Anzahl an Suchanfragen, die dem Wert in einem Set von 50 Suchanfragen entsprechen würde. In den rechten beiden Spalten ist die Anzahl der entsprechenden verwendeten Suchanfragen sowie deren prozentualer Anteil an der Gesamtheit der untersuchten Suchanfragen angegeben.

Anzahl an Wörtern	Anteil an Gesamtheit	Entspräche in Suchanfragen	Verwendete Anzahl Suchanfragen	Anteil an Gesamtheit
1	42 %	21	26	52 %
2	26,4 %	13	17	34%
3	16,6 %	8	5	10%
4	7,9 %	4	2	4 %
5 oder mehr	6,8 %	4	-	-

Tabelle 4: Worthäufigkeiten der ausgewählten Suchanfragen

4.4.2.2. thematische Suchanfragen

Bevor Suchbegriffe in das Set der geeigneten Anfragen aufgenommen wurden, wurde überprüft, ob sie mindestens zehn Treffer im OPAC produzierten und ob sie inhaltlich einem der DFG-Sondersammelgebiete der ausgewählten Bibliothek entsprachen. Durch die Vermeidung von Suchanfragen, die weniger Treffer lieferten, sollte eine bessere Vergleichbarkeit der ermittelten Daten gewährleistet werden. Auffällig war, dass viele der Suchanfragen, in denen vier Wörter verwendet wurden, so speziell waren, dass weniger als zehn Treffer produziert wurden.

Die mutmaßlich thematischen Suchbegriffe wurden daraufhin überprüft, ob es sich tatsächlich um thematische Anfragen handelte und nicht um *Known-Item*-Suchanfragen die auf ein bestimmtes Buch mit dem Suchbegriff als Titel zielten. Dies geschah durch Eingabe des Suchbegriffs in den Katalog der Deutschen Nationalbibliothek bzw. – bei englischsprachigen Suchanfragen – in den Katalog der Library of Congress. Es wurde davon ausgegangen, dass es kein Buch mit dem entsprechenden Titel gibt, wenn die Referenzbibliothek keines nachweist. Ein Beispiel für eine solche Suchanfrage ist z.B. „Drogenkonsum und Kriminalität“

Es wurde zudem eine Homonym-Kontrolle durchgeführt, d.h. Suchbegriffe, die mehrere Dinge bezeichnen, wurden von der Untersuchung ausgeschlossen. Dies sollte es den Evaluatoren erleichtern, ein Urteil zu fällen. Möglicherweise hat dieses Vorgehen das Ergebnis des Tests leicht positiv beeinflusst.

4.4.2.3. Known-Item-Suchanfragen

Bei der *Known-Item-Search* konnten Suchanfragen bei denen nur der Name eines Autors als Suchbegriff benutzt wird, nicht berücksichtigt werden. Der Grund dafür ist, dass die Beurteilung, ob eines der gefundenen Bücher das gesuchte Item ist, nicht von den Personen vorgenommen wurde, die die Suchanfragen ursprünglich stellten, sondern vom Verfasser dieser Arbeit. Stattdessen sind auch die *Known-Item-Searches* aus dem Logfile der Universitätsbibliothek Frankfurt am Main entnommen. Bei der Bewertung des gefunden Treffer ist also unklar, welcher Autor gemeint ist, falls es mehrere Autoren desselben Namens gibt. Zudem haben viele Autoren mehr als ein Buch geschrieben, auch in diesem Fall ist bei der Bewertung nicht klar, welches der gefundenen Bücher das gesuchte ist.

Bei der *Known-Item-Search* wurden insgesamt 70 Suchanfragen durchgeführt. Bei 35 wurde nach dem Titel eines Buches gesucht, bei 35 nach der Kombination von Titel und Autor. Die Anfragen wurden jeweils im Katalog der Deutschen Nationalbibliothek bzw. der Library of Congress daraufhin überprüft, ob sie tatsächlich – so wie es den Anschein hatte - geeignet waren, um ein bestimmtes Buch zu finden. Bei den Titel-Anfragen wurde, um eine Bewertung zu ermöglichen, zudem überprüft, ob nur ein Buch mit dem entsprechenden Titel erschienen ist.

4.4.2.4 Verteilung der Suchbegriffe auf Sprachen und Themen

Hawking et al. raten dazu, bei Untersuchungen von Suchmaschinen eine große Anzahl an Suchbegriffen zu verwenden. Dies konnte in dieser Studie geleistet werden. Sie regen weiterhin an: „Search topics should be representative of genuine user need.“ (HAWKING 2001, S. 25). Sie forderten also, dass die Suchbegriffe repräsentativ für echte Informationsbedarfe der Nutzer sein sollten. Dies wurde hier nicht geleistet, immerhin handelt es sich bei den hier verwendeten Suchbegriffen aber um echte, von Nutzern formulierte und an den OPAC gestellte Anfragen. Hawking et al. vertreten auch die Auffassung, dass in Studien verwendete Suchbegriffe die volle Spannweite an Informationsbedarfen repräsentieren sollten (ebd.). Mangels Daten, zu welchen Themen Nutzer in OPACs Anfragen stellen, kann keine Aussage darüber getroffen werden, in wie weit die 50 zufällig ausgewählten Suchbegriffe tatsächlich repräsentativ sind.

Um Suchanfragen inhaltlich bestimmten Gruppen zuordnen zu können, entwickelte Spink eine Klassifikation für Suchanfragen (vgl. SPINK 2001, S. 23). Diese wurde jedoch für Suchanfragen an öffentliche Internet-Suchmaschinen entwickelt und lässt sich nicht zur Klassifizierung von Suchanfragen an den OPAC einer wissenschaftlichen Bibliothek verwenden. Die Klassifikation enthält z.B. eine Klasse „Sex, Pornographie, Vorlieben“, für die im OPAC nicht mit Anfragen oder Treffern zu rechnen ist. Andererseits werden wissenschaftliche Suchanfragen nur in zwei Klassen unterteilt: „Gesundheit, Naturwissenschaften“ und „Bildung, Geisteswissenschaften“.

Von den 50 Suchbegriffen sind 40 in deutscher Sprache, neun sind in Deutschland im allgemeinen Sprachgebrauch übliche Anglizismen (wie z.B. rock’n’roll) und eine Anfrage ist in englischer Sprache (*molecular modelling*).

4.5. Datenerfassung und Anonymisierung bei der thematischen Suche

Hawking et al. legen nahe, dass in zukünftigen Studien die Bewertung der Treffer in anonymisierter Form erfolgen sollte (vgl. HAWKING 2001, S. 25). Damit soll vermutlich bezweckt werden, ein möglichst objektives Urteil über die Treffer, bzw. in dieser Studie über die Trefferbeschreibungen, zu erhalten. Der Gedanke, dass das vertraute Layout einer beliebten Suchmaschine oder aber die Position eines Suchergebnisses das Relevanzurteil beeinflussen, ist naheliegend. Eine Anonymisierung ist insbesondere dann wichtig, wenn es sich bei einer der untersuchten Suchmaschinen um Google handelt. Machill zeigte, dass Google ein sehr positives Image bei deutschen Nutzern hat: „ […] das Image eines sympathischen Dienstleisters, bei dem technisch ambitionierte, junge Mitarbeiter für optimale Leistung sorgen […] “ (MACHILL 2003, S. 188).

Die Anonymisierung der Treffer erfolgte mit Hilfe des Programms Zotero. Bei diesem Programm handelt es sich um ein Plugin[12] für das Browser-Programm Mozilla Firefox. Es stellt Funktionen zur Literaturverwaltung zur Verfügung und ermöglicht insbesondere den automatischen Import bibliographischer Daten von im Internet betrachteten Literatur- bzw. Quellenangaben in die Datenbank des Programms. Die in der Datenbank gespeicherten Katalogeinträge lassen sich in Teil-Sammlungen gliedern. Diese

[12] Ein Plugin ist ein Programm, das ein anderes Programm um zusätzliche Funktionen erweitert.

Sammlungen wiederum können in verschiedenen Datei-Formaten exportiert und so zur Nutzung mit anderen Programmen bereitgestellt werden.

Die ersten zehn Ergebnisse einer Suchanfrage in einem der Kataloge wurden in einer Sammlung gespeichert. Dabei fiel auf, dass die Trefferbeschreibungen von Google weniger bibliographische Informationen enthielten als diejenigen im OPAC. Der Großteil der Beschreibungen in Google enthielt die Informationen Titel, Autor, Verlag, ISBN und Seitenzahl. Bei Trefferbeschreibungen im OPAC wurden darüber hinaus in vielen Fällen Informationen zu Herausgebern, zu Zeitschriftenreihen oder zur Auflage automatisch in den Zotero-Datensatz exportiert. Die Untertitel mussten bei den Trefferbeschreibungen beider Angebote manuell ergänzt werden.

Die beiden Sammlungen, welche die Ergebnisse einer in beiden Katalogen gestellten Anfrage enthielten, wurden exportiert und in eine gemeinsame Datei kopiert. Durch dieses Verfahren veränderte sich das Layout der Informationen völlig. Eine Unterscheidung, aus welchem Angebot die Trefferbeschreibungen stammten war nur noch anhand von vergebenen Identifizierungsnummern möglich. Die Reihenfolge der Treffer wurde dadurch verändert, dass das Programm Zotero die Datensätze in einer Sammlung alphabetisch sortiert. Huang und Wang untersuchten, in wie weit die Relevanz von Dokumenten anders beurteilt wird, wenn das Dokument an anderer Stelle in der Trefferliste angezeigt wird (vgl. HUANG 2004). Sie empfehlen abschließend, Dokumente in eine zufällige Reihenfolge zu bringen, bevor sie bewertet werden, um zu verhindern, dass das Urteil durch die Platzierung beeinflusst wird. Zu diesem Schluss kommt auch Eisenberg (EISENBERG 1988, S. 384).

Als Ergebnis dieses Arbeitsschrittes waren die anonymisierten Treffer in einer Datei des Textverarbeitungsprogramms Microsoft Word (Version 2003) gespeichert (Auszug im Anhang). Die Beschreibungen waren nicht durchmischt, d.h. die ersten zehn Treffer waren Ergebnisse aus dem einen Katalog und die darauffolgenden zehn Treffer die Ergebnisse des anderen Katalogs. Um auszuschließen, dass die Ergebnisse dadurch verändert werden, wurde die Reihenfolge nach der Bearbeitung der Hälfte der Dateien verändert. In 25 Dokumenten folgen also die Ergebnisse der Buchsuche auf die des OPACs und in den verbleibenden 25 Dokumenten folgen die Ergebnisse aus dem OPAC auf die Ergebnisse aus der Buchsuche.

4.6. Datenerfassung bei der Known-Item-Search

Bei der Erfassung der Ergebnisse der Suche nach bekannten Dokumenten mussten einige Regeln definiert werden, um eine einheitliche Erfassung zu ermöglichen. So galt nur die neueste im jeweiligen Angebot vorhandene Auflage eines Treffers als das eine gewünschte Ergebnis. Dies entspricht bei Sach- und Fachliteratur vermutlich der Realität, für Belletristik erfolgt diese Vorgabe aus methodischen Gründen. Alle Suchanfragen wurden so interpretiert, dass sie gedruckte Werke finden sollten, d.h. im OPAC verfügbare Hörbücher wurden nicht berücksichtigt. Die thematischen Sondersammelgebiete der DFG wurden bei der *Known-Item-Search* nicht berücksichtigt. Die gesuchten Bücher wurden also nicht daraufhin kontrolliert, ob sie eines der Sondersammelgebiete behandeln, da dies zu aufwändig gewesen wäre.

Es wurden nur die ersten zehn Ergebnisse zu einer Suchanfrage berücksichtigt. Die Verwendung von *Cut-Off*-Werten, also die Beschränkung auf eine bestimmte Anzahl Treffer, ist durchaus üblich (vgl. LEWANDOWSKI 2005, S. 140). Nur die ersten zehn Treffer zu berücksichtigen ist angemessen, da es dem Nutzerverhalten in OPACs entspricht. Hennies stellte fest, dass nur 11% der Datensätze, die aufgerufen wurden, auf Rangplatz elf oder tiefer angezeigt wurden, und das obwohl der von ihm untersuchte OPAC 20 Treffer auf einer Seite darstellte (vgl. HENNIES 2006, S. 11). Es kann davon ausgegangen werden, dass es sich zusätzlich negativ auf die Nutzungshäufigkeit von Ergebnissen auswirkt, wenn diese auf Seiten platziert sind, die für den Nutzer nur durch weitere Klicks erreichbar sind.

4.7. Relevanzbewertung

Die Relevanzbewertung wurde von zehn Evaluatoren durchgeführt, die jeweils alle 20 Treffer von fünf Suchanfragen bewerteten. Als Evaluatoren fungierten Studierende diverser Studiengänge. Hawking et al. forderten dazu auf, Studien so zu gestalten, dass alle Bewertungen zu einer Suchanfrage von demselben Evaluatoren durchgeführt werden sollten (vgl. HAWKING 2001, S.25). Dies ist in dieser Studie geschehen. Hawking forderte weiter, dass die Bewertungen entweder von denjenigen Personen vorgenommen werden sollten, die den Informationsbedarf bzw. die Suchanfrage

formulierten, oder aber von unabhängigen Evaluatoren (ebd.). In dieser Studie wurde letzteres umgesetzt.

Hawking forderte allerdings auch, dass in dem Fall, dass Evaluatoren eingesetzt werden, um die Treffer zu Suchanfragen anderer Personen zu bewerten, der Informationsbedarf dieser Personen bestmöglich erfasst sein muss (ebd.). Hintergrund für diese Forderung ist die Tatsache, dass es Evaluatoren leichter fällt, Urteile zu fällen, wenn sie das tatsächliche Informationsbedürfnis kennen und nicht nur die formulierte Suchanfrage. Diese methodische Forderung konnte in dieser Untersuchung nicht erfüllt werden.

Die Word-Dateien, welche die anonymisierten Trefferbeschreibungen enthielten, wurden zusammen mit einem Anschreiben per Email an die Evaluatoren gesendet. In diesem Anschreiben waren die Skalen abgebildet und nur die Endpunkte benannt. Auf die Benennung von Zwischenkategorien wurde verzichtet. Die Endpunkte wurden mit „nicht relevant“ und „sehr relevant“ benannt. Das Konzept „Relevanz“ wurde in dem Anschreiben nicht erläutert, da es sich dabei laut Saracevic um ein intuitives Konzept handelt, das keiner Erklärung bedarf (vgl. SARACEVIC 2007, S.4). Die Juroren notierten daraufhin ihre Wertung in dem Word-Dokument.

4.7.1 Die verwendeten Skalen

In informationswissenschaftlichen Untersuchungen werden verschiedene Skalen verwendet, um die Relevanz von Dokumenten zu bewerten. Es kommen sowohl binäre Skalen zum Einsatz als auch Skalen die eine feinere Unterscheidung ermöglichen als „relevant“ oder „nicht relevant“ (vgl. TANG 1999). Bei Ordinal-Skalen, die auch als Kategorie-Skalen bezeichnet werden, liegt die höchste in der Literatur verwendete Anzahl an Kategorien bei elf. Ein noch genaueres Urteil ermöglichen Intervall-Skalen, bei denen innerhalb von Grenzwerten jeder Wert vergeben werden kann (z.B. Prozentwerte) (vgl. GREISDORF 2001).

Greisdorf und Spink untersuchten, welchen Einfluss die verwendete Skala auf die Relevanzurteile hatte, die von Juroren vergeben wurden (vgl. GREISDORF 2001). Es wurde geprüft, ob sich die Relevanzurteile für Dokumente veränderten, wenn den Juroren verschiedene Bewertungs-Skalen vorgelegt wurden. Greisdorf und Spink stellen fest, dass die Häufigkeit mit der Dokumente als relevant, irrelevant oder teilweise

relevant bewertet wurden, sich auch bei Verwendung unterschiedlicher Skalen nicht änderte. Deshalb kam die Studie zu dem Ergebnis, dass die Skala mit Hilfe derer die Juroren urteilen sollen, keinen wesentlichen Einfluss auf das Ergebnis der Beurteilung hat: „ […] it is becoming evident from the research that end-users are capable of using a variety of scales.“ (ebd, S. 11). Spink und Greisdorf schlossen deshalb „ […] that scale choice may not be a key issue in relevance assessments.“ (ebd.).

Tang et al. verfolgten einen anderen Ansatz: Sie untersuchten, ob Anwender bestimmte Skalen bevorzugten, um ein Relevanzurteil über bibliographische Datensätze zu fällen. Dazu ließen sie Juroren Dokumente auf unterschiedlichen Skalen bewerten. Die Juroren sollten zusätzlich zu jedem Urteil angeben, wie stark ihr Vertrauen in das gefällte Urteil war. Es zeigte sich, dass Evaluatoren am wenigsten Vertrauen in die von ihnen gefällten Urteile hatten, wenn eine binäre Skala verwendet worden waren. Die Skala, auf der die Juroren sich ihres Urteils am sichersten waren, war eine Ordinal-Skala mit sieben Ausprägungen (vgl. TANG 1999, S. 10). Dementsprechend wird hier eine solche Skala, mit den Werten von 0 bis 6, verwendet.

Auch Hawking et al. sprechen sich für die Verwendung einer nicht-binären Skala aus, um Ergebnisse informationeller Suchanfragen zu bewerten (vgl. HAWKING 2001, S. 24). Allerdings ergibt sich bei der Verwendung einer solchen Skala das Problem, dass das klassische Retrievalmaß Precision nicht ohne weiteres gebildet werden kann. Um die Precision (dazu mehr unter 5.1.) zu ermitteln, ist eine binäre Beurteilung der Treffer notwendig (vgl. BORLUND 2003, S.7). Viele Studien lösten dieses Problem, indem sie verschiedene Kategorien einer Skala in die Gruppen „relevant“ und „nicht relevant“ zusammenfassten. Beispielsweise wurden auf einer Skala mit 4 Ausprägungen die Kategorien 0 und 1 als nicht relevant betrachtet, die Kategorien 2 und 3 hingegen als relevant. Laut Eisenberg ist dieses Vorgehen jedoch problematisch, da Nutzer die Grenze zwischen relevant und nicht-relevant nicht bei dem Mittelwert der Skala ziehen, sondern bei einem tieferen Wert (vgl. EISENBERG 1988, S. 15). Lewandowski löste dieses Problem, indem er die Juroren, zusätzlich zu nicht-binären Wertungen, Bewertungen auf einer binären Skala abgeben ließ (vgl. LEWANDOWSKI 2008a). Dieses Vorgehen soll hier übernommen werden.

5. Ergebnisse

Zur Erfassung der Daten aus den Word-Dateien, in denen die Relevanzurteile gespeichert waren, wurde eine Excel-Datei verwendet. Die Auswertung erfolgte mittels Excel-Funktionen, wie z.B. Häufigkeit und Anzahl.

5.1. Subject Search

Das Retrievalmaß Precision gibt an, wie hoch der Anteil an relevanten Dokumenten an der Gesamtheit, der auf eine Suchanfrage hin zurückgelieferten Dokumente ist. Die Precision ermittelt „ […] die Fähigkeit eines Information-Retrieval-Systems, unerwünschte Ballast-Dokumente auszufiltern.“ (WOMSER-HACKER 2004, S. 229). Als Ballast-Dokumente werden nicht-relevante Dokumente bezeichnet. Wie auch bei der Ermittlung von Ergebnissen für die *Known-Item-Search* ist es üblich, die Precision nicht für alle zurückgelieferten Dokumente zu ermitteln, sondern bei der Untersuchung der Ergebnisse nach einer bestimmten Position der Ergebnis-Liste, dem *Cut-Off*-Wert, abzubrechen (vgl. LEWANDOWSKI 2005, S. 140). Ein häufig verwendeter *Cut-Off*-Wert ist zehn.

5.1.1. Micro Precision

> Bei der Mikromethode werden die Elementarparameter des Retrievalergebnisses (d.h. die Anzahl der relevanten und nicht-relevanten Antwortdokumente) über den gesamten Aufgabenbestand addiert; dann erfolgt auf dieser Basis die verallgemeinernde Berechnung des entsprechenden Bewertungsmaßes. (WOMSER-HACKER 2004, S. 231)

Das hier gewählte Bewertungsmaß ist die Precision. Es wurde die Anzahl der relevanten und nicht-relevanten Dokumente in der Menge der Ergebnisse aller durchgeführten Suchanfragen ermittelt. Als Teilergebnis wurde zudem der Anteil relevanter Ergebnisse an der Menge derjenigen Dokumente ermittelt, die – bei allen 50 Suchanfragen - jeweils an erster Stelle in der Ergebnisliste platziert waren. Auch für alle weiteren Plätze der Ergebnislisten wurde so verfahren, so dass der Recall-Precision-Graph die Precision bis zum jeweiligen Platz der Ergebnisliste angibt (siehe Diagramm 1). An Platz zehn schließlich finden sich die durchschnittlichen Precision-Werte aller untersuchten Treffer.

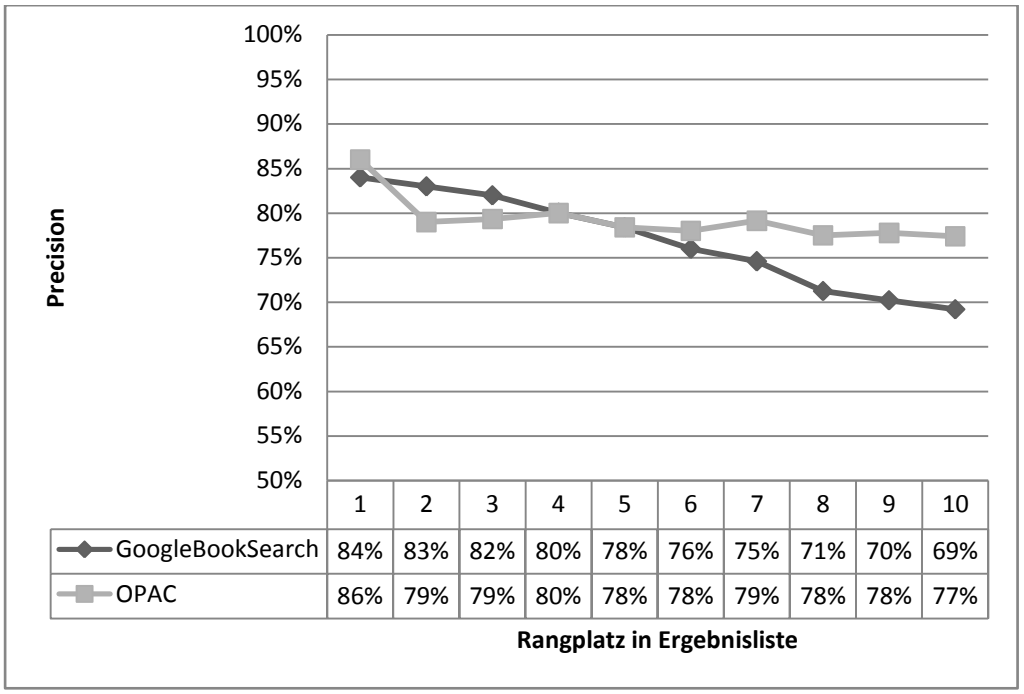


Diagramm 1: Recall-Precision-Graph

Die Betrachtung von Diagramm 1 zeigt ein uneinheitliches Bild. Auf dem ersten Platz erzielt der OPAC bessere Wertungen als die Buchsuche. Bis zu den Plätzen zwei und drei ist Google besser. Berücksichtigt man alle Ergebnisse bis zur vierten und fünften Anfrage, zeichnet sich kein Unterschied zwischen den Angeboten ab. Ab Platz sechs ist auf allen folgenden Plätzen das gleiche Bild zu finden: der OPAC hat eine höhere Präzision als die Buchsuche. Auffällig ist zudem, dass sich bei den Ergebnissen des OPACs kein Abnehmen der Relevanz zeigt, je mehr Plätze berücksichtigt werden. Dieses Verhalten ist beim Einsatz eines Ranking-Verfahrens zu erwarten und bei der Buchsuche deutlich sichtbar.

5.1.2. Makro-Precision

> Bei der Makromethode wird das gewählte Bewertungsmaß für jedes Retrievalergebnis (d.h. für jede Anfrage/Aufgabe) getrennt berechnet, anschließend werden die erzielten Werte arithmetisch gemittelt. (WOMSER-HACKER 2004, S. 231)

In der Tabelle 5 ist die durchschnittliche, also arithmetisch gemittelte, Precision der Ergebnisse aller Suchanfragen dargestellt.

	Google Books Durch-schnitt	Google Books Einge-schränkte Vorschau	Google Books Voll-ansicht	OPAC Durch-schnitt	OPAC online verfüg-bare Quellen	OPAC offline verfügbare Quellen
Durchschnitt-liche Precison	0,69	0,70	0,64	0,77	0,81	0,76

Tabelle 5: Durchschnittliche Precision der Ergebnisse aller Suchanfragen in beiden Angeboten, getrennt nach Anzeigemodus bzw. Art der Zugänglichkeit.

	Google Books Durchschnit t	Google Books Einge-schränkte Vorschau	Google Books Voll-ansicht	OPAC Durch-schnitt	OPAC online verfügba re Quellen	OPAC offline verfügbare Quellen
Durchschnitt-liche Relevanz auf Skala von 0 bis 6	3,18	3,21	2,93	3,64	3,69	3,57

Tabelle 6: Durchschnittliche Relevanz der Ergebnisse aller Suchanfragen getrennt nach Anzeigemodus bzw. Art der Zugänglichkeit.

Die Tabellen zeigen, dass sowohl die Precision- als auch die Relevanz-Ergebnisse des OPACs höher liegen als die Ergebnisse der Buchsuche. Eine detaillierte Betrachtung der Ergebnisse, getrennt nach der Art der Zugänglichkeit, zeigt ein einheitliches Bild. In beiden Arten der Zugänglichkeit wurden die Retrieval-Ergebnisse des OPACs besser bewertet als die Ergebnisse der Buchsuche. Die Bewertung der im OPAC nachgewiesenen Online-Quellen ist sehr gut, allerdings muss bedacht werden, dass sie auf nur 26 Dokumenten beruht (siehe 5.4.). Die Gruppe der bei Google im Volltext verfügbaren Bücher erzielt, unabhängig von der Bewertungsskala, die schwächsten Ergebnisse aller Gruppen.

	Google Books	OPAC
Bei Bewertung auf binärer Skala [A]	16	29
Bei Bewertung auf nicht-binärer Skala	18	32

Tabelle 7: Anzahl der Suchanfragen, bei denen die Angebote die besseren Ergebnisse hatten.
[A]: Zu 50 fehlende: Bei fünf Suchanfragen wurde in beiden Angeboten die gleiche Anzahl an Ergebnissen als relevant bewertet

Für die Darstellung der Ergebnisse in Tabelle 7 wurde die Anzahl der auf der binären Skala als relevant bewerteten Ergebnisse der einzelnen Suchanfragen ermittelt. Die Summen beider Angebote wurden verglichen. Dementsprechend wurden, bei der Betrachtung der Bewertungen auf der nicht-binären Skala, die Bewertungen der einzelnen Ergebnisse einer Suchanfrage addiert und die Summen miteinander verglichen. Es zeigt sich, unabhängig davon, auf welcher Skala bewertet wurde, dass der OPAC bei deutlich mehr Anfragen die besseren Ergebnisse liefert, als die Buchsuche dies tut.

5.1.3. Sonstige Darstellungen

In Diagramm 1 sind die Ergebnisse der von den Juroren auf der binären Skala durchgeführten Wertung dargestellt. Diagramm 2 zeigt, wie häufig die einzelnen Relevanzurteile der feineren Skala bei welchem Angebot abgegeben wurden.

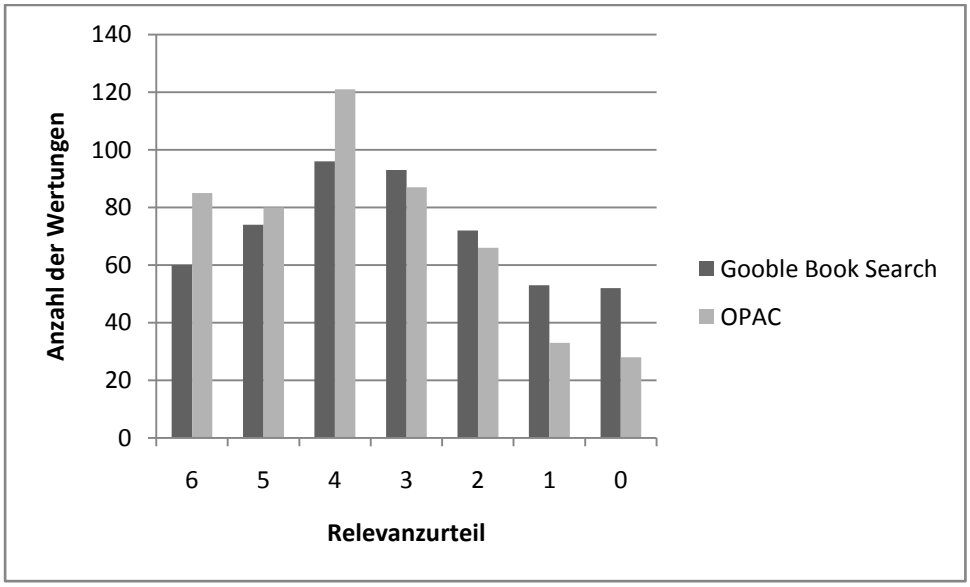


Diagramm 2: Häufigkeit der Relevanzurteile auf der nicht-binären Skala in absoluten Zahlen

Diagramm 2 veranschaulicht, dass die positiven Wertungen von 4 bis 6 für Ergebnisse des OPACs deutlich häufiger vergeben wurden, als für Ergebnisse der Buchsuche. Die schlechteren Wertungen von 0 bis 2 dagegen wurden für die Ergebnisse der Buchsuche häufiger vergeben. Auffällig ist zudem, dass deutlich weniger Ergebnisse des OPACs mit 0 bzw. 1 bewertet wurden.

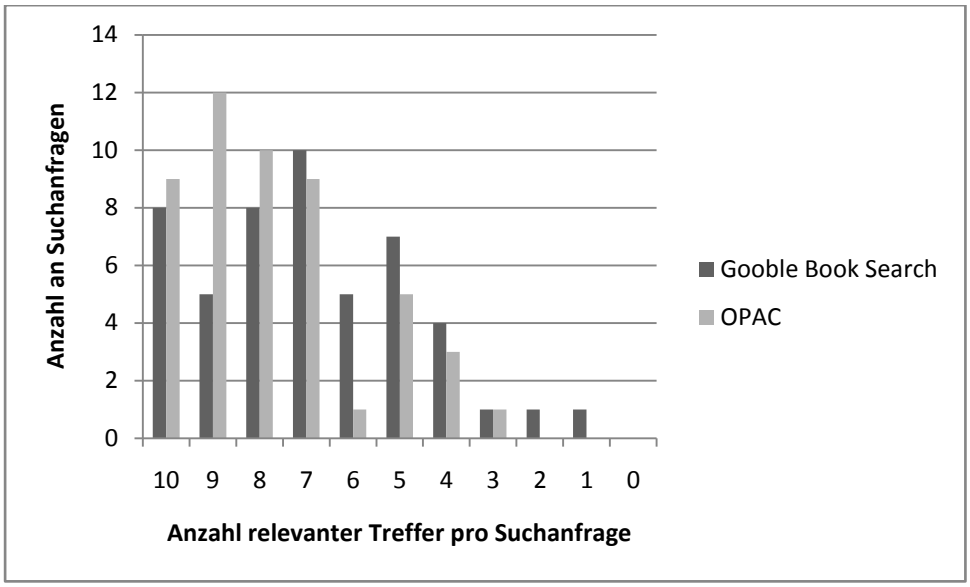
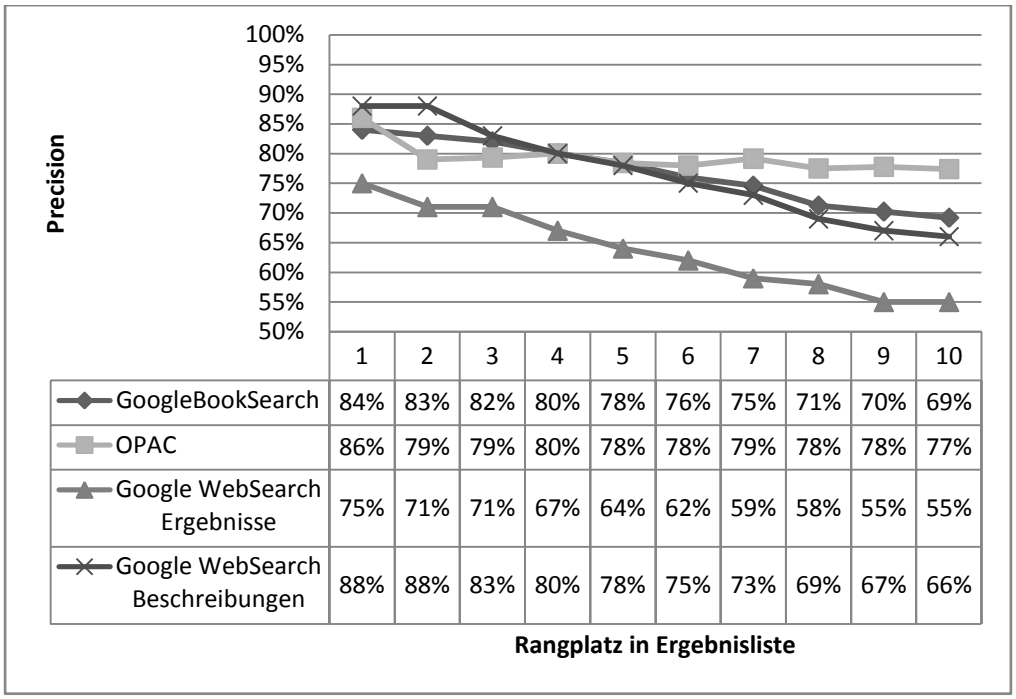


Diagramm 3: Anzahl an Suchanfragen mit n relevanten Treffern

Für Diagramm 3 wurde ermittelt, bei wie vielen Suchanfragen unter den Ergebnissen eine bestimmte Anzahl relevanter Dokumente war. Dazu wurde auf die binäre Relevanz-Bewertung zurückgegriffen. Es fällt auf, dass es im OPAC viele Suchanfragen gab, die eine hohe Anzahl (8-10) an relevanten Ergebnissen produzierte. Die Anzahl solcher Suchanfragen in der Buchsuche war geringer. Anfragen, die nur wenige (0-3) relevante Ergebnisse lieferten, waren in der Buchsuche häufiger als im OPAC.

5.2. Vergleich der Ergebnisse der thematischen Suche mit Ergebnissen einer anderen Studie

Lewandowski führte 2008 eine Untersuchung durch, in der er die Qualität der Retrieval-Ergebnisse von fünf Suchmaschinen ermittelte (vgl. LEWANDOWSKI 2008a). Er berücksichtigte sowohl die Qualität der Ergebnisse selbst, als auch die Qualität der Ergebnisbeschreibungen auf den Trefferlisten der Suchmaschinen. Ein Vergleich der Ergebnisse der vorliegenden Untersuchung mit den Ergebnissen seiner Studie ist insofern interessant, als dass die Ergebnisse der Google Buchsuche mit den Ergebnissen der Google Websuche verglichen werden können.



	1	2	3	4	5	6	7	8	9	10
GoogleBookSearch	84%	83%	82%	80%	78%	76%	75%	71%	70%	69%
OPAC	86%	79%	79%	80%	78%	78%	79%	78%	78%	77%
Google WebSearch Ergebnisse	75%	71%	71%	67%	64%	62%	59%	58%	55%	55%
Google WebSearch Beschreibungen	88%	88%	83%	80%	78%	75%	73%	69%	67%	66%

Diagramm 4: Precision-Ergebnisse dieser Studie im Vergleich mit Ergebnissen der Google Websuche aus LEWANDOWSKI 2008a.

Diagramm 4 zeigt die etwa parallel zueinander verlaufenden Graphen der Precision von Ergebnissen und Ergebnisbeschreibungen in der Websuche. Die Precision der Beschreibungen ist durchgängig etwa 12% höher als die der Ergebnisse. Die Beschreibungen erzielten in allen von Lewandowski untersuchten Suchmaschinen bessere Wertungen als die Ergebnisse selbst (vgl. LEWANDOWSKI 2008a, S. 14). Ob diese Situation auch bei der Buchsuche bzw. im OPAC auftritt ist unbekannt.

Auf dem Diagramm wird deutlich, dass die Precision der Google Buchsuche derjenigen der Trefferbeschreibungen der Google Websuche sehr ähnlich ist. Dies ist insofern passend, als das auch bei der Untersuchung der Buchsuche nur die Trefferbeschreibungen und nicht die Treffer selbst bewertet wurden. An Position 1 und 2 liegt die Precision der Trefferbeschreibungen der Websuche noch 5% bzw. 4% über der Precision der Buchsuche. Auf den folgenden Plätzen ist der Unterschied der Precision beider Gruppen aber nicht größer als 3%.

5.3. Known-Item-Search

	Google Books	OPAC
Alle Known-Item-Suchanfragen:		
durchschnittlicher Rang des Items	1,6	1,8
nicht gefundene Items:	7%	24%
auf 1 gerankt	73%	54%
Suchanfragen nach Titeln:		
durchschnittlicher Rang des Items	1,1	1,7
nicht gefundene Known Items:	3%	29%
auf 1 gerankt	94%	51%
Suchanfragen nach Titel und Autor:		
durchschnittlicher Rang des Items	2,2	2
nicht gefundene Known Items:	11%	20%
auf 1 gerankt	51%	57%

Tabelle 8: Ergebnisse der *Known-Item*-Suchanfragen.

Die Betrachtung der Ergebnisse der Suchanfragen nach bekannten Dokumenten zeigt zunächst, dass die Anzahl an Dokumenten, die im OPAC der SUB Hamburg nicht gefunden werden, deutlich höher ist, als die Anzahl an Dokumenten, die bei Google nicht gefunden werden. Dies verwundert angesichts des – Schätzungen zufolge – größeren Bestandes der Buchsuche nicht. Das Ranking scheint bei Google besser zu funktionieren. Der durchschnittliche Rangplatz der gefundenen Dokumente ist besser (näher an 1) und der Anteil der Dokumente, die tatsächlich auf dem ersten Platz der Ergebnisliste angezeigt wurden, ist höher.

Betrachtet man die Ergebnisse der *Known-Item-Search* detaillierter, nämlich die Ergebnisse der Suchanfragen nach Buchtiteln und der Suchanfragen nach Kombinationen von Buchtitel und Autorennamen getrennt, zeigt sich ein anderes Bild. Bei den Suchanfragen nach Titeln ist Google mit 94% auf dem ersten Platz angezeigten Treffern sehr gut, der OPAC mit beinahe einem Drittel an – bis zum zehnten Platz - nicht gefundenen Dokumenten eher enttäuschend. Bei den Suchanfragen nach einer Kombination aus Autor und Buchtitel sind die Ergebnisse gemischt, kein Angebot ist klar

besser als das andere. Die Buchsuche ist also nur bei Suchanfragen nach Titeln deutlich effizienter als der OPAC. Dieser bringt bei beiden Typen von Suchanfragen ähnliche Ergebnisse, Google beantwortet Titelanfragen wesentlich besser als Kombinationsanfragen. Allerdings umfasst die Datengrundlage bei dieser Betrachtungsweise nur 35 Suchanfragen je Typ.

Bei den Ergebnissen der Buchsuche fiel auf, dass bei solchen Suchanfragen, bei denen das gesuchte Item nicht an erster Position gezeigt wurde, stattdessen häufig solche Bücher an erster Stelle gezeigt wurden, in deren Literaturverzeichnis das gesuchte Buch erwähnt war. Möglicherweise wertet der Ranking-Algorithmus Quellenverzeichnisse höher, als dies angemessen wäre.

5.4. Arten der Zugänglichkeit

Bei der Google Book Search werden unterschiedliche Anzeigemodi dadurch gebildet, dass Bücher, die auf verschiedene Arten zugänglich sind, gemeinsam angezeigt werden. Wie häufig welche Art der Zugänglichkeit in welchem Ansichtsmodus ist, ist in den Diagrammen 5 und 6 dargestellt.

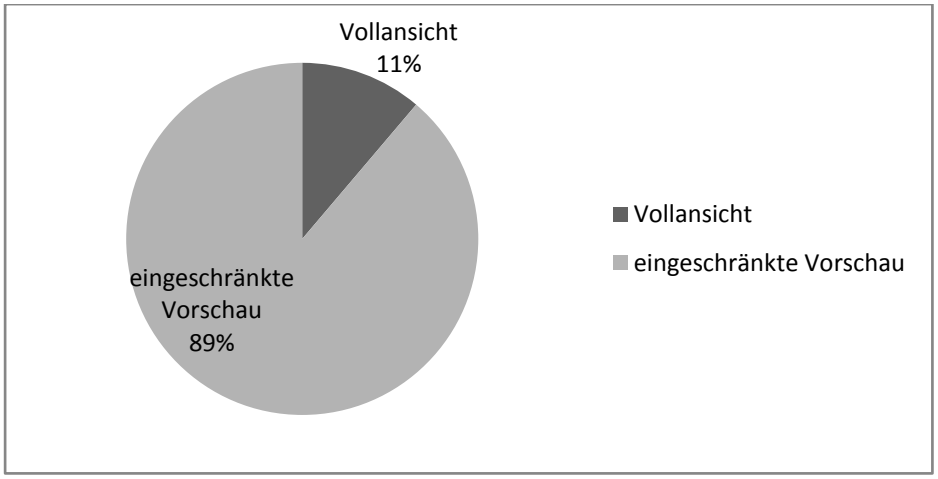


Diagramm 5: Zugänglichkeit der Bücher im Anzeigemodus „eingeschränkte Vorschau und Vollansicht“ (100% = 500 Treffer).

Diagramm 5 zeigt, dass in dem Anzeigemodus, der die Ergebnisse auf dem Nutzer tatsächlich zugängliche Bücher beschränkt, der Anteil an Büchern, die im Volltext zugänglich sind, sehr gering ist. Der Nutzer ist also darauf angewiesen seinen Informationsbedarf in Quellen zu befriedigen, zu denen er nur eingeschränkten Zugang hat.

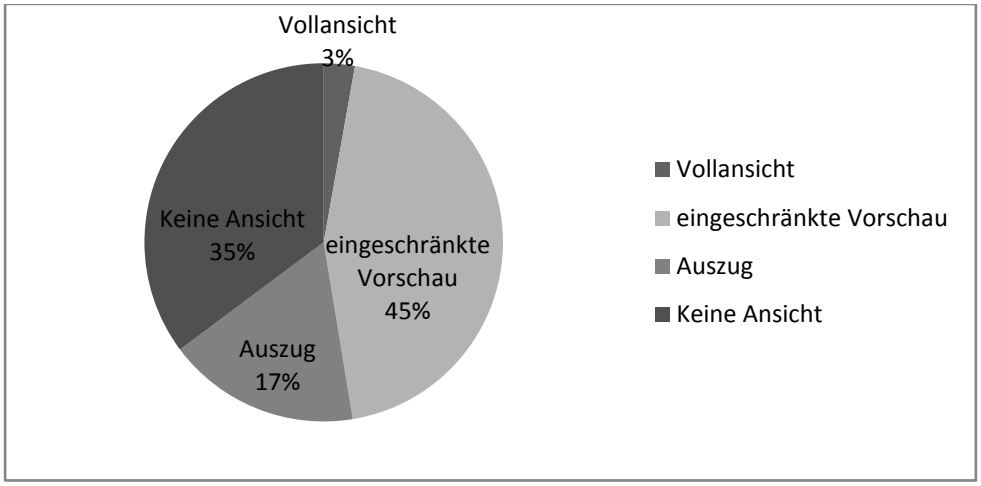


Diagramm 6: Zugänglichkeit der Bücher im Anzeigemodus alle Bücher (100% = 500 Treffer).

Diagramm 6 zeigt, dass auch im Anzeigemodus „Alle Bücher“ solche Bücher, die in eingeschränkter Vorschau zugänglich sind, am häufigsten vorkommen. Volltexte spielen mit nur 3 % der Bücher kaum eine Rolle. Mit 35% ist der Anteil der Bücher, die dem Nutzer in keiner Form zugänglich sind, ihm also keinen Informationsgewinn bringen, relativ hoch. Leider ist nicht bekannt, ob die Ergebnisse der Suchanfragen in ihrer Verteilung auf die Arten der Zugänglichkeit repräsentativ für den gesamten Daten-bestand der Google Buchsuche sind. Es ist möglich, dass der Ranking-Algorithmus Ergebnisse mit einer bestimmten Art der Zugänglichkeit bevorzugt, so dass diese in den ermittelten Ergebnissen überrepräsentiert sind.

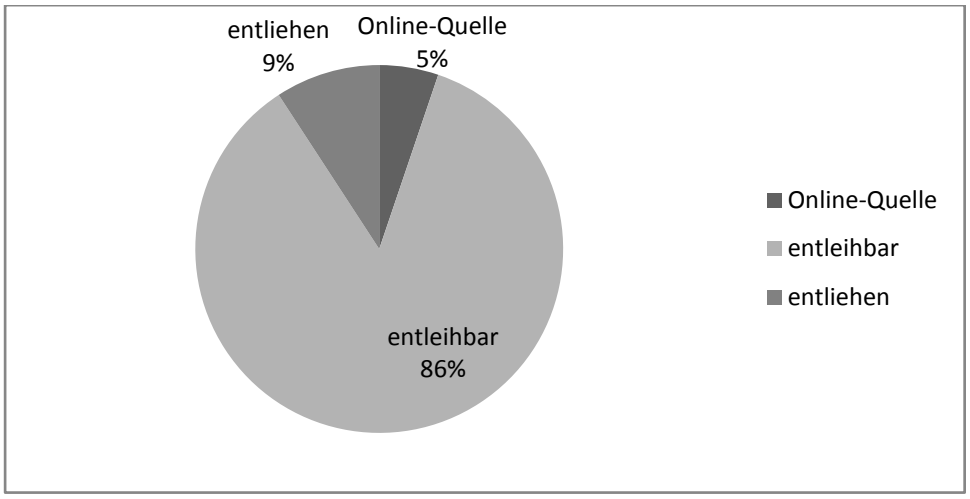


Diagramm 7: Verfügbarkeit der Ergebnisse im OPAC.

Die in Diagramm 7 dargestellte Absenzquote von 9% im Bestand der SUB Hamburg ist überraschend niedrig. Der Anteil an online verfügbaren Quellen ist mit 5% überraschend hoch.

5.5. Kataloganreicherungen

Die Häufigkeit, mit der die einzelnen Kataloganreicherungen vorhanden waren, wurde manuell ermittelt. Untersucht wurden in diesem Zusammenhang nur die Ergebnisse der thematischen Suche, die Ergebnisse der *Known-Item-Search* wurden nicht berücksichtigt. Bei der Untersuchung von Google wurden sowohl die Ergebnisse der Suche mit dem Anzeigemodus „Vollansicht und Eingeschränkte Vorschau“ untersucht, als auch die Ergebnisse der Suche im Anzeigemodus „Alle Bücher“. Es wurden also 1000 Treffer der Buchsuche untersucht sowie 500 Treffer aus dem OPAC der SUB Hamburg.

Es gibt in beiden Katalogen Inhaltsverzeichnisse und Rezensionen. Während die Buchbeschreibung in der Buchsuche (die meist durch Klappentexte gebildet wird) der *Publisher Description* im OPAC entspricht, gibt es für die Informationen über *Contributors* (also am Buch Mitwirkende), die im OPAC gelegentlich angezeigt werden, keine Entsprechung in der Buchsuche. Auch mit objektiver Zielsetzung verfasste Abstracts finden sich nur im OPAC, nicht in der Buchsuche. Bei Titelblättern ist das Gegenteil der Fall, diese finden sich nur in der Buchsuche und nicht im OPAC. Hier sei

noch einmal darauf hingewiesen, dass die Kataloganreicherungen bei Google direkt in die beschreibende Seite integriert werden, beim OPAC jedoch nur als Link zur Verfügung stehen.

	Google Books Durchschnitt	**Google Books eingeschränkte Vorschau und Vollansicht**	**Google Books Alle Bücher**	**OPAC**
Grundgesamtheit (Anzahl der Ergebnisse im Anzeigemodus)		500	500	500
Titelblatt	85%	99%	70%	
Inhaltsverzeichnis	79%	90%	67%	33%
Buchübersicht/ Publisher Description	16%	19%	12%	7%
Rezension	4%	4%	5%	1%
Contributors				2%
Abstract				0,2%

Tabelle 9:Verfügbarkeit der einzelnen Kataloganreicherungen in Prozent der Datensätze bei den einstellbaren Anzeigemodi von Google Books sowie im OPAC (die Anzahl der Rezensionen wurde nicht berücksichtigt).

Die Grundgesamtheit bei der Betrachtung der Durchschnittswerte der Book Search ist unbekannt, da es Überschneidungen zwischen den Ergebnissen der Suchanfragen im Anzeigemodus „Eingeschränkte Vorschau und Vollansicht“ und den Ergebnissen der Suchanfragen im Anzeigemodus „Alle Bücher“ gibt. Wie hoch der Anteil an Überschneidungen ist, ist unbekannt.

Tabelle 9 zeigt, dass die Ergebnisse beider Anzeigemodi von Google mehr Kataloganreicherungs-Objekte enthalten als die Ergebnisse des OPACs. Fast jedes

Ergebnis von Google im Anzeigemodus „Eingeschränkte Vorschau und Vollansicht“ enthält ein Titelblatt. Im OPAC ist der Anteil der Datensätze, die mit anderen Objekten als einem Inhaltsverzeichnis angereichert wurden, gering.

	Vollansicht	**Eingeschränkte Vorschau**	**Auszug**	**Keine Vorschau**
Grundgesamtheit (Anzahl der Ergebnisse mit Art der	70	667	87	176
Titelblatt	97%	99%	98%	18%
Inhaltsverzeichnis	70%	94%	90%	14%
Buchübersicht/ Klappentext	9%	20%	2%	7%
Rezension	1%	5%	2%	6%

Tabelle 10: Verfügbarkeit der einzelnen Kataloganreicherungen in Prozent der Datensätze für die einzelnen Arten der Zugänglichkeit bei Google Books (die Anzahl der Rezensionen wurde nicht berücksichtigt).

Die Betrachtung der Ergebnisse der Google Buchsuche in Tabelle 10 zeigt, dass Bücher, für die keine Vorschau verfügbar ist, deutlich schlechter mit Titelblättern und Inhaltsverzeichnissen erschlossen sind, als die anderen Bücher der Buchsuche. Klappentexte und Rezensionen treten bei diesen Büchern aber in ähnlicher Häufigkeit auf wie bei den anderen Büchern.

5.5.1. Rezensionen

Tabelle 9 berücksichtigt nicht die Anzahl der Rezensionen, sondern lediglich die Datensätze, für die mindestens eine Rezension vorlag. Untersucht man die Anzahl der Rezensionen genauer, zeigt sich, dass Treffer deren Beschreibung mindestens eine Rezension enthält, in der Buchsuche durchschnittlich zwei Rezensionen enthalten. Im OPAC war bei keiner Trefferbeschreibung mehr als eine Rezension enthalten, dementsprechend liegt die durchschnittliche Anzahl bei eins.

	Google Books Durchschnitt	Google Books eingeschränkte Vorschau und Vollansicht	Google Books Alle Bücher	OPAC
Anzahl der Datensätze mit Rezension	37	39	35	7
Anzahl der Rezensionen	73	63	83	7
Von Nutzern geschriebene Rezensionen	9	11	7	-
Fremdsprachliche Rezensionen	9	9	9	0
Durchschnittliche Anzahl an Rezensionen pro Datensatz mit min. 1 Rezension	2	2,4	1,6	1

Tabelle 11: Anzahl an Rezensionen bei den einstellbaren Anzeigemodi von Google Books sowie im OPAC.

Auch wenn die Tabelle zeigt, dass Rezensionen in Trefferbeschreibungen bei Google Books deutlich häufiger sind, als bei Trefferbeschreibungen im OPAC, sagt dies nichts über die Qualität der Inhaltsbeschreibung durch die Kataloge aus. Der höheren Anzahl an Rezensionen innerhalb der Buchsuche steht deren fragwürdige Qualität gegenüber. Bei Rezensionen, die im OPAC verlinkt worden sind, kann davon ausgegangen werden, dass sie gewissen qualitativen Ansprüchen genügen. Für die von Google verlinkten Rezensionen kann dies dagegen nicht angenommen werden, da die Verlinkung vermutlich automatisch geschieht. Unter den Quellen der Rezensionen finden sich neben Seiten, bei denen davon auszugehen ist, dass sie einen Qualitätsanspruch haben (z.B. die Homepages der kommerziellen Anbieter JSTOR und Wiley), Seiten unbekannter Qualität (z.B. das Blog von „Lady Taluka“[13] und der Blog-Anbieter Blogspot allgemein). Die Qualität der von Nutzern der Buchsuche geschriebenen Rezensionen ist ebenfalls nicht mit der Qualität von professionellen Rezensionen vergleichbar. Fremdsprachliche Rezensionen sind naturgemäß nur hilfreich, wenn der Nutzer der Sprache mächtig ist[14].

[13] ladytaluka.com

[14] Als „fremdsprachlich“ wurden Rezensionen eingeordnet, wenn sie in einer anderen Sprache als Englisch oder Deutsch verfasst waren.

Auf eine Darstellung der Verteilung der Häufigkeiten nach Arten der Zugänglichkeit (analog zu Tabelle 10) soll verzichtet werden, da die Grundgesamtheiten für Bücher, die in Auszügen oder in Vollansicht darstellbar sind, lediglich zwei bzw. eine Rezension umfassen und somit keine sinnvolle Analyse erlauben.

5.5.2. Inhaltsverzeichnisse

Die Inhaltsverzeichnisse bei Google sind von unterschiedlicher Qualität. Es ist anzunehmen, dass ein automatisiertes Texterkennungsverfahren eingesetzt wird, um Inhaltsverzeichnisse zu erkennen. Dieses ist noch nicht völlig ausgereift. Zwei unterschiedliche Probleme treten im Zusammenhang mit Inhaltsverzeichnissen häufig auf. Das erste sind Inhaltsverzeichnisse, die zwar die Seitenzahl angeben, auf der ein Kapitel beginnt, nicht jedoch die Kapitelüberschrift. Stattdessen sind Standard-Kapitelüberschriften wie „Kapitel 1“ oder „Abschnitt 1“ angegeben (vgl. Abb. 7). Ein solches Inhaltsverzeichnis hilft dem Nutzer nicht bei der inhaltlichen Orientierung im Text. Im Folgenden soll angenommen werden, dass dies die nützlichste Funktion von Inhaltsverzeichnissen für Nutzer von Internet-Katalogen ist. Die Navigationsfunktion, die Inhaltsverzeichnisse auch haben, baut auf der erschließenden Funktion von Inhaltsverzeichnissen auf.

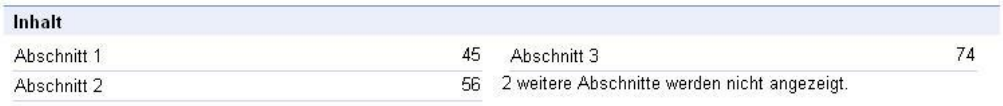


Abb. 7: Inhaltsverzeichnis mit Standard-Überschriften.

Unklar ist, warum in vielen dieser Inhaltsverzeichnisse nicht alle Kapitel auf diese Art enthalten sind. Häufig enden die Verzeichnisse mit dem Hinweis „X weitere Abschnitte werden nicht angezeigt“. Eine genauere Untersuchung der Ergebnisse, deren Beschreibung ein solches Inhaltsverzeichnis mit Standard-Überschriften enthält, zeigt, dass das Durchschnittsalter dieser Treffer mit 69 Jahren deutlich höher ist, als das Durchschnittsalter aller Treffer der Buchsuche (24 Jahre). Allerdings enthalten Bücher aus allen Jahrzehnten solche Inhaltsverzeichnisse. Es kann also nicht festgestellt werden,

dass vor allem alte Bücher derartige Inhaltsverzeichnisse enthalten. Zu bedenken ist weiterhin, dass diese Auswertung auf lediglich 40 Büchern basiert.

Das zweite, häufig auftretende Problem sind Inhaltsverzeichnisse, bei denen die Kapitelüberschriften in fehlerhafter Form ausgelesen wurden. Die automatische Texterkennung von Google scheint vor allem dann Probleme zu haben, wenn die Kapitelüberschrift einen Zeilenumbruch enthält. Als Überschriften werden dann oft sinnlose Reihen von Worten und Wortteilen angezeigt. Häufig sind auch Inhaltsverzeichnisse zu finden, in denen anstatt der korrekten Kapitelüberschriften die Namen der Autoren der einzelnen Kapitel als Kapitelüberschriften erkannt werden. Ein anderes Problem stellt die relativ kurze Zeilenlänge dar, die Google für Kapitelüberschriften vorgesehen hat. Viele wissenschaftliche Werke haben längere Kapitelüberschriften, diese werden nach maximal 60 Zeichen abgeschnitten. Auch bei all diesen Inhaltsverzeichnissen ist eine Orientierungsfunktion nicht gegeben.

	Google Books Durchschnitt	**Google Books eingeschränkte Vorschau und Vollansicht**	**Google Books Alle Bücher**	**OPAC**
Anteil der Datensätze mit Inhaltsverzeichnis	79 %	90 %	67 %	33 %
Aussagekräftige Inhaltsverzeichnisse [B]	63 %	69 %	54 %	33 %
Stark fehlerhafte Inhaltsverzeichnisse [B]	32 %	27 %	39 %	
Inhaltsverzeichnisse mit Standard-Überschriften [B]	5 %	4 %	7 %	

Tabelle 12: Übersicht über die Häufigkeit des Vorkommens von Inhaltsverzeichnissen sowie deren Qualität bei den einstellbaren Anzeigemodi von Google Books sowie im OPAC in Prozent der Datensätze ([B] 100 % = Anzahl aller Inhaltsverzeichnisse im jeweiligen Anzeigemodus).

	Vollansicht	Eingeschränkte Vorschau	Auszug	Keine Vorschau
Grundgesamtheit (Anzahl der Ergebnisse mit Art der Zugänglichkeit)	70	667	87	176
Anteil der Datensätze mit Inhaltsverzeichnis	70 %	94 %	90%	14 %
Aussagekräftige Inhaltsverzeichnisse [c]	47 %	73 %	0	28 %
Stark fehlerhafte Inhaltsverzeichnisse [c]	37 %	25 %	88 %	44 %
Inhaltsverzeichnisse mit Standard-Überschriften [c]	16 %	3 %	12 %	28 %

Tabelle 13. : Übersicht über die Häufigkeit des Vorkommens von Inhaltsverzeichnissen sowie deren Qualität für die einzelnen Arten von Zugänglichkeit bei Google Books in Prozent der Datensätze ([C] 100 % = Anzahl aller Inhaltsverzeichnisse bei Treffern mit der jeweiligen Art von Zugänglichkeit).

Bei der Betrachtung der Zahlen der Book Search muss bedacht werden, dass bei den meisten Büchern in der Buchsuche, denjenigen die als eingeschränkte Vorschau oder im Volltext betrachtet werden können, der Nutzer auf das Inhaltsverzeichnis im Text des Buches zurückgreifen kann, um sich über den Inhalt des Buches zu orientieren. Die hier ermittelten Zahlen beziehen sich nur auf die Inhaltsverzeichnisse, die von Google direkt in die Übersichts-Seite integriert wurden. Aus Tabelle 13 geht hervor, dass gerade bei den Büchern, bei denen Nutzer nicht auf den Text des Buches zugreifen können, nur selten Inhaltsverzeichnisse zur Verfügung stehen.

Als „aussagekräftige Inhaltsverzeichnisse“ wurden auch Inhaltsverzeichnisse mit kleineren Fehlern eingeordnet, solange sie ihre Funktion erfüllen, dem Nutzer eine inhaltliche Orientierung zu ermöglichen. Die Entscheidung, ob dies der Fall war, wurde nach subjektivem Empfinden getroffen. Bei einigen Ergebnissen wurde als Buchbeschreibung ein Inhaltsverzeichnis veröffentlicht, dies wurde bei der Datenerfassung als Inhaltsverzeichnis gewertet und nicht als Buchbeschreibung.

Die in Tabelle 12 ermittelten Zahlen belegen, dass die Ergebnisrepräsentation bei Google erheblich mehr Inhaltsverzeichnisse enthält als im OPAC. Zudem wird bezüglich des Anteils an Ergebnissen, deren Beschreibung ein Inhaltsverzeichnis enthält, ein

deutlicher Unterschied zwischen den beiden Anzeigemodi von Google Books sichtbar, beim Anzeigemodus „Eingeschränkte Vorschau und Vollansicht“ ist dieser Anteil 23 % höher. Dies liegt nahe, da die Bücher in diesem Modus nicht nur für den Nutzer besser zugänglich sind, sondern vermutlich auch für Google.

Tabelle 13 zeigt das Ergebnis der Untersuchung getrennt nach der Art der Zugänglichkeit. Es fällt auf, dass bei Büchern, für die keine Vorschau möglich ist, der Anteil an Inhaltsverzeichnissen besonders gering ist. Ursache dafür ist möglicherweise die Tatsache, dass Google bei diesen Büchern kein Volltext zur Verfügung steht. Weiterhin auffällig ist, dass bei Büchern, die im Auszug dargestellt werden, von 78 Inhaltsverzeichnissen kein einziges als aussagekräftig eingeordnet wurde. Allerdings ist zu bedenken, dass diese Beurteilung auf subjektivem Empfinden beruht.

5.6. Alter der Treffer

In Diagramm 8 ist zu erkennen, dass bei allen Angeboten die Hälfte, bzw. mehr als die Hälfte der Ergebnisse Bücher aus diesem Jahrzehnt waren. Die aktuellsten Treffer stellt Google im Anzeigemodus „eingeschränkte Vorschau und Vollansicht“ zur Verfügung. Im OPAC ist der Anteil an Treffern, deren Erscheinen weiter zurückliegt als 1970, sehr gering.

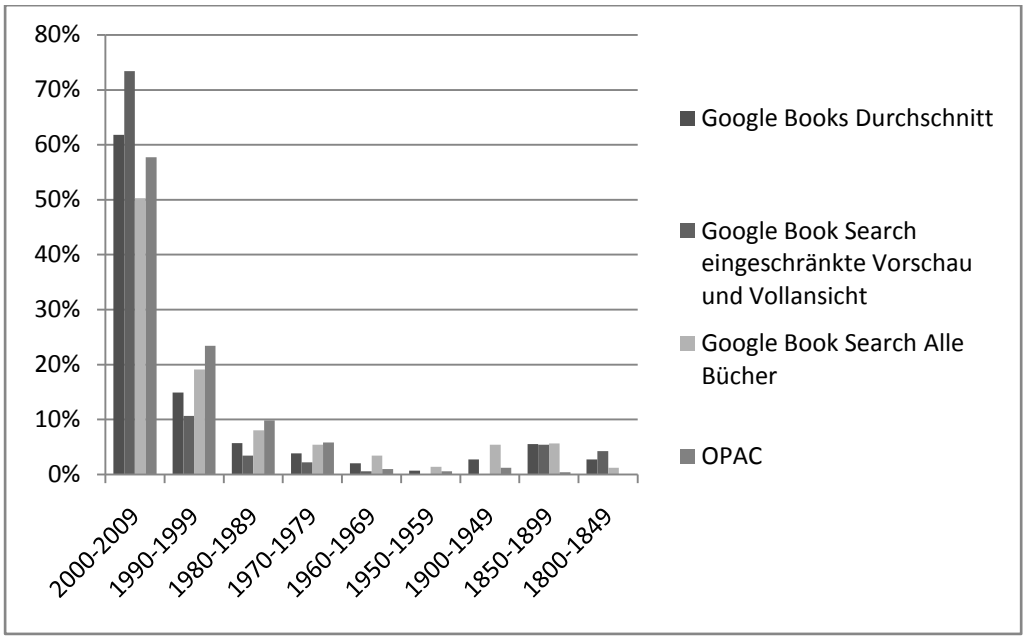


Diagramm 8: Verteilung der Treffer auf Zeitabschnitte nach Erscheinungsjahr, nach einzelnen Angeboten bzw. Anzeigemodi getrennt. Zu beachten ist, dass die Zeitabschnitte ungleichmäßig gebildet wurden (100 % = Anzahl aller untersuchten Treffer im jeweiligen Angebot (jeweils 500)).

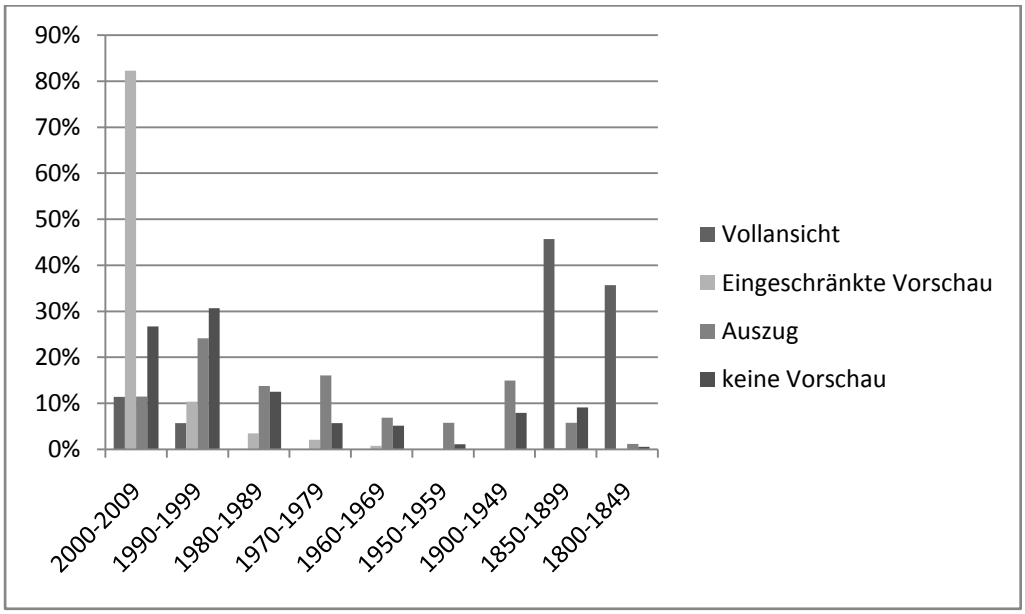


Diagramm 9: Verteilung der Treffer auf Zeitabschnitte nach Erscheinungsjahr, nach einzelnen Arten der Zugänglichkeit getrennt. Zu beachten ist, dass die Zeitabschnitte ungleichmäßig gebildet wurden (100 % = Anzahl aller Treffer mit der jeweiligen Art der Zugänglichkeit (siehe Tabelle 13)).

	Google Books Durchschnitt	Google Books eingeschränkte Vorschau und Vollansicht	Google Books Alle Bücher	OPAC
Grundgesamtheit		500	500	500
Durchschnittsalter (in Jahren)	23,9	22,3	25,5	12,3

Tabelle 14: Durchschnittliches Alter der Ergebnisse bei den einstellbaren Anzeigemodi von Google Books sowie im OPAC.

Tabelle 14 zeigt, dass Ergebnisse der Buchsuche durchschnittlich etwa doppelt so alt sind wie Ergebnisse des OPACs. Zwischen den einzelnen Anzeigemodi von Google Books besteht ein Unterschied von 3 Jahren.

	Vollansicht	Eingeschränkte Vorschau	Auszug	Keine Vorschau
Grundgesamtheit	70	667	87	176
Durchschnittsalter (in Jahren)	135,1	7,2	41,9	33,8

Tabelle 15: Durchschnittliches Alter der Ergebnisse für die einzelnen Arten von Zugänglichkeit bei Google Books

Tabelle 15 zeigt deutliche Unterschiede des Durchschnittsalters für die Ergebnisse der einzelnen Arten von Zugänglichkeit. Bei Büchern in eingeschränkter Vorschau beträgt das Durchschnittsalter nur 7,2 Jahre, liegt also noch 5 Jahre unter dem des OPACs. Die Ergebnisse in Vollansicht, die mit den vorgenannten Ergebnissen zum Anzeigemodus „Eingeschränkte Vorschau und Vollansicht“ kombiniert werden, sind deutlich schlechter, das Durchschnittsalter beträgt 135,1 Jahre. Die These, dass der Großteil an im Volltext verfügbaren Büchern sehr alt ist, kann also belegt werden.

6. Fazit

Die Untersuchung zeigte, dass die Retrievalergebnisse des OPACs der SUB Hamburg bei der thematischen Suche in beiden gedachten Nutzungssituationen besser sind als die der Google Book Search. In der ersten angenommen Nutzungssituation, in welcher der Nutzer bereit ist, in die Bibliothek zu fahren, um Bücher zu entleihen, liegt die Precision des OPACs um 8% über der Precision der Buchsuche. In der zweiten Nutzungssituation, die davon ausgeht, dass der Nutzer nur online verfügbare Quellen berücksichtigt, sogar um 12%. Der OPAC hatte für 58% bzw. 64% der Anfragen – je nach verwendeter Skala – die besseren Ergebnisse.

Bei der *Known-Item-Search* hat die Buchsuche die besseren Ergebnisse. Hier gibt es deutlich weniger Items, die durch die Suche nicht gefunden wurden, und die gefundenen Items sind häufiger an erster Position der Trefferliste. Bei Suchanfragen nach Titel waren die Ergebnisse Googles mit einigem Abstand besser als die des OPACs.

Vergleicht man die Qualität der Ergebnisse der Google Buchsuche mit derjenigen der Google Websuche, die in einer Studie von Lewandowski ermittelt wurde, zeigt sich eine erstaunliche Ähnlichkeit. Volltexte machen in der Buchsuche die Minderheit der Texte aus. Im Ansichtsmodus „Eingeschränkte Vorschau und Vollansicht“ sind 11% der Ergebnisse Volltexte, im Ansichtsmodus „Alle Bücher“ sind es nur 3%. Sieht man auch die Bücher in eingeschränkter Vorschau als zugänglich an, sind 52% der Ergebnisse von Suchanfragen im Ansichtsmodus „Alle Bücher“ zugänglich. In wie weit Bücher, die in eingeschränkter Vorschau verfügbar sind, tatsächlich geeignet sind, Informationsbedarfe zu befriedigen, ist, aufgrund des großen Anteils solcher Bücher am Bestand der Buchsuche, von erheblicher Bedeutung.

Kataloganreicherungen sind bei Google wesentlich häufiger als im OPAC, jedoch teilweise auch von fragwürdiger Qualität. Durchschnittlich enthalten 85% der Datensätze in der Buchsuche ein Titelblatt und 79% ein Inhaltsverzeichnis. Von den Inhaltsverzeichnissen besitzt aber mehr als ein Drittel wenig Informationswert für den Nutzer.

Bei beiden Angeboten macht Literatur aus diesem Jahrzehnt den Großteil der Ergebnisse aus - 58% im OPAC und 62% bei der Buchsuche. Das Durchschnittsalter der Treffer ist bei

Google mit 23,9 Jahren fast doppelt so hoch wie im OPAC. Der Altersdurchschnitt der Ergebnisse bei Google schwankt, je nach Art der Zugänglichkeit, stark. Im Volltext verfügbare Bücher sind durchschnittlich 135 Jahre alt.

Obwohl abschließend festgestellt werden kann, dass die Retrievalergebnisse der thematischen Suchanfragen bei Google schlechter bewertet wurden, ist die Google Buchsuche doch mit Sicherheit ein geeignetes Instrument zur Recherche. Speziell die Volltext-Indexierung ermöglicht es, zu anderen Rechercheergebnissen zu kommen, als im OPAC. Eine ergänzende Verwendung der Buchsuche bei thematischen Suchanfragen ist deshalb sinnvoll.

Die *Known-Item-Search* hat in dieser Studie wenig Raum eingenommen, stellt in der Realität jedoch die am häufigsten auftretende Art von Anfrage dar. Gerade bei diesen Anfragen erzielte die Google Book Search die besseren Ergebnisse. Lediglich fünf von 70 Items konnten nicht gefunden werden. Die Buchsuche ist also sehr gut geeignet um Anfragen nach bestimmten Dokumenten – in diesem Fall Büchern – auszuführen. Da beinahe die Hälfte der Ergebnisse der Book Search - 48% - für den Nutzer inhaltlich mehr oder weniger zugänglich ist, besteht für ihn durch die Buchsuche eine realistische Möglichkeit, von jedem PC mit Internet-Anschluss auf das Buch, das er sucht, zugreifen zu können.

Literaturverzeichnis

BAND 2005
BAND, Jonathan: Features - The Authors Guild vs. The Google Print Library Project Online-Ressource: Law and technology resources for legal professionals http://www.llrx.com/features/googleprint.htm Stand 15.10.2005 Abruf 27.08.2009

BAND 2006
BAND, Jonathan: The Google Library Project : Both sides of the story. In: Plagiary: Cross-Disciplinary Studies in Plagiarism, Fabrication, and Falsification, Jg. 2 (2007) 1. Band. Auch als Online-Ressource: http://www.plagiary.org/Google-Library-Project.pdf, Abruf 27.08.2009. Angaben zu Seitenzahlen bei Zitaten beziehen sich auf die Online-Resource.

BEHN 1995
BEHN, Inka: Der SISIS-OPAC in den Bibliotheken der Fachhochschule Hamburg : Entwicklung eines Benutzungsmanuals sowie eine Untersuchung zum studentischen Benutzungsverhalten. Hamburg : Fachhochschule Hamburg, 1995. - 88 S., Ill.

BEHRENS-NEUMANN 2008
BEHRENS-NEUMANN, Renate: Aus der 55. Sitzung der Arbeitsgemeinschaft der Verbundsysteme am 5. und 6. November 2008 in Frankfurt am Main. In: Bibliotheksdienst Jg. 43. (2009), 2. Heft

BORLUND 2003
BORLNUD, Pia: The Concept of Relevance in IR. In: JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, Jg. 54(2003) 10. Band

BRANTLEY 2009
BRANTLEY, Peter: The Orphan Monopoly
Online-Ressource: Blog von Peter Brantley
http://blogs.lib.berkeley.edu/shimenawa.php/2009/03/15/the-orphan-monopoly Stand 15.03.2009
Abruf 27.08.2009

CEYNOWA 2008
CEYNOWA, Klaus: Der „BSB-Google-Deal“ In: GRIEBEL, Rolf (Hrsg.): Bibliotheksmagazin : Mitteilungen aus den Staatsbibliotheken in Berlin und München, Jg. 2 (2008) 1. Heft.

COMPUTER UND RECHT 2006
Wissenschaftliche Buchgesellschaft vs. Google Book Search In: CR : Computer und Recht, Jg. 21 (2006) 8. Heft. Angaben zu Seitenzahlen bei Zitaten beziehen sich auf einen Ausdruck aus der Genios-Datenbank.

COMPUTERWOCHE 2007
Google dreht an vielen Rädern In: Computerwoche, Jg. 33 (2007) 4. Heft. Angaben zu Seitenzahlen bei Zitaten beziehen sich auf einen Ausdruck aus der Genios-Datenbank.

COMSCORE 2008
comScore Releases July 2008 German Search Rankings
Online-Ressource: Homepage der ComScore Inc.
http://www.comscore.com/Press_Events/Press_Releases/2008/09/Top_German_Search_Engines
Abruf: 27.08.2009

COPYRIGHT
Copyright Law of the United States of America and Related Laws Contained in Title 17 of the United States Code : § 107. Limitations on exclusive rights: Fair use. Auch als Online-Ressource: http://www.copyright.gov/title17/92chap1.html#107, Abruf 27.08.2009

DATEMA 2007
DATEMA, Jay: ALA 2007: Online Books, Copyright, and User Preferences
Online-Ressource: Blog von Jay Datema
http://www.bookism.org/open/2007/07/06/ala-2007-online-books-copyright-and-user-preferences/ Stand 06.07.2007
Abruf 27.08.2009

DATEMA 2007a
DATEMA, Jay: IDPF: Google and Harvard
Online-Ressource: Blog von Jay Datema
http://www.bookism.org/open/2007/05/09/idpf-google-and-harvard/ Stand 09.05.2007
Abruf 27.08.2009

DFG 2009
Förderprogramm: Überregionale Literaturversorgung
Online-Ressource: Homepage der DFG
http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/projektfoerderung/foerderziele/literaturerwerbung.html Stand 25.06.2009
Abruf 27.08.2009

DIEROLF 2007
DIEROLF, Uwe: Der Uni-Katalog in Karlsruhe : Ein Katalog im Wandel. In: BIT: Zeitschrift für Bibliothek, Information und Technologie, Jg. 10 (2007) 2. Heft

DRUMMOND 2005
Brief des Google-Vizepräsidenten David Drummond an das US Copyright Office
Online-Ressource: Homepage des United Staates Copyright Office
http://www.copyright.gov/orphan/comments/OW0681-Google.pdf
Abruf 27.08.2009

EICHHOLTZ 2002
EICHHOLTZ, Stephanie: Die US-amerikanische Class Action und ihre deutschen Funktionsäquivalente. Tübingen : Mohr Siebeck, 2002. -355 S.

EISENBERG 1988
EISENBERG, Michael B.: Measuring Relevance Judgements. In: Information Processing & Management, Jg. 24 (1988) Heft 4

ELIB 2009
Die elektronische Bibliothek Bremen…
Online-Ressource: Homepage der elektronischen Bibliothek Bremen
http://elib.suub.uni-bremen.de/frs_projekt_elib.html Stand 17.08.2009
Abruf 27.08.2009

FAZ FINANCE 2009
Kennzahlen zu Google Inc.
Online-Ressource: Homepage des Finanzportals der Frankfurter Allgemeinen Zeitung
http://www.fazfinance.net/d/WertpapierAttr.faz?id=kenn&ident=sto_US38259P5089
Stand 27.08.2009
Abruf 27.08.2009

FAZ FINANCE 2009a
Profil der Thyssen Krupp AG
Online-Ressource: Homepage des Finanzportals der Frankfurter Allgemeinen Zeitung
http://www.fazfinance.net/d/WertpapierAttr.faz?id=profile&ident=sto_DE0007500001
Stand 27.08.2009
Abruf 27.08.2009

FAZ FINANCE 2009b
Profil der Axel Springer AG
Online-Ressource: Homepage des Finanzportals der Frankfurter Allgemeinen Zeitung
http://www.fazfinance.net/d/WertpapierAttr.faz?id=profile&ident=sto_DE0005501357
Stand 27.08.2009
Abruf 27.08.2009

FEDERAL REGISTER 2005
Orphan Works. In: Federal Register January 26, Jg. 70 (2005) Heft 16. Auch als Online-Ressource: http://www.copyright.gov/fedreg/2005/70fr3739.html
Abruf 27.08.2009

FUNK 2008
FUNK, Victor: Wissen für Alle.
Online-Ressource: Homepage der Frankfurter Rundschau
http://www.fr-online.de/in_und_ausland/kultur_und_medien/medien/?em_cnt=1322607& Stand 21.04.2008
Abruf 27.08.2009

GBV-WIKI
Google Buchsuche
Online-Ressource: GBV-Verbund-WIKI
http://www.gbv.de/wikis/cls/Google_Buchsuche Stand 2.10.2008
Abruf 27.08.2009

GERICHTSURTEIL
Gerichtsurteil im Verfahren Kelly gegen Arriba Soft Corporation
Online-Ressource: Homepage von Dennis Karjala
http://homepages.law.asu.edu/~dkarjala/cyberlaw/KelllyvArriba(9C2003).htm
Abruf 27.08.2009

GOOGLE 2004
Google Checks Out Library Books
Online-Ressource: Google Press Center
http://www.google.com/press/pressrel/print_library.html
Abruf 27.08.2009

GOOGLE 2008
Authors, Publishers, and Google Reach Landmark Settlement
Online-Ressource: Google Press Center
http://www.google.com/intl/en/press/pressrel/20081027_booksearchagreement.html
Abruf 27.08.2009

GOOGLE 2009
Unternehmensbezogene Informationen – Google Management
Online-Ressource: Homepage von Google
http://www.google.de/intl/de/corporate/execs.html#larry
Abruf 27.08.2009

GOOGLE 2009a
Unternehmensbezogene Informationen - Unternehmensprofil
Online-Ressource: Homepage von Google
http://www.Google.de/intl/de/corporate/
Abruf 27.08.2009

GOOGLE 2009b
Vergleich (in Englisch)
Online-Ressource: Homepage mit Informationen zum Vergleich, der den Rechtsstreit zwischen Authors Guild und Google beendet
http://www.googlebooksettlement.com/r/view_settlement_agreement
Abruf 27.08.2009

GOOGLE 2009c
Startseite
Online-Ressource: Homepage mit Informationen zum Vergleich, der den Rechtsstreit zwischen Authors Guild und Google beendet
http://books.google.com/booksrightsholders/
Abruf 27.08.2009

GOOGLE 2009d
Mitteilung für vom Vergleich betroffene Autoren
Online-Ressource: Homepage mit Informationen zum Vergleich, der den Rechtsstreit zwischen Authors Guild und Google beendet
http://www.googlebooksettlement.com/intl/de/notice.html
Abruf 27.08.2009

GOOGLEBLOG 2004
Bookmark this Page
Online-Ressource: Blog von Google
http://googleblog.blogspot.com/2004/10/bookmark-this-site.html Stand 6.10.2004
Abruf 27.08.2009

GOOGLEBLOG 2004a
All booked up
Online-Ressource: Blog von Google
http://googleblog.blogspot.com/2004/12/all-booked-up.html Stand 14.12.2004
Abruf 27.08.2009

GOOGLEBLOG 2005
Google Print and the Authors Guild
Online-Ressource: Blog von Google
http://googleblog.blogspot.com/2005/09/google-print-and-authors-guild.html Stand 20.09.2005
Abruf 27.08.2009

GOOGLEBLOG 2005a
Making Books easier to find
Online-Ressource: Blog von Google
http://googleblog.blogspot.com/2005/08/making-books-easier-to-find.html Stand 12.08.2005
Abruf 27.08.2009

GOOGLEBLOG 2008
New Chapter for Google Book Search
Online-Ressource: Blog von Google
http://googleblog.blogspot.com/2008/10/new-chapter-for-google-book-search.html Stand 28.10.2008
Abruf 27.08.2009

GOOGLEBLOG 2008a
Book Info where you need it when you need it
Online-Ressource: Blog von Google
http://googleblog.blogspot.com/2008/03/book-info-where-you-need-it-when-you.html Stand 13.03.2008
Abruf 27.08.2009

GOOGLE BOOKS 2009
Bibliothekspartner
Online-Ressource: Homepage von Google Books
http://books.google.com/googlebooks/partners.html
Abruf 27.08.2009

GOOGLE BOOKS 2009a
Warum kann ich nicht das Vollständige Buch lesen?
Online-Ressource: Homepage von Google Books
http://books.Google.de/support/bin/answer.py?answer=43729&topic=9259&hl=de
Abruf 27.08.2009

GOOGLE BOOKS BLOG 2009
New Features on Google Books
Online-Ressource: Blog von Google Books
http://booksearch.blogspot.com/2009/06/new-features-on-google-books.html Stand 18.06.2009
Abruf 27.08.2009

GOOGLE BOOKS BLOG 2009a
Books Are Full of Visual Gems: 19th century fashion edition
Online-Ressource: Blog von Google Books
http://booksearch.blogspot.com/2009/07/books-are-full-of-visual-gems-19th.html Stand07.07.2009
Abruf 27.08.2009

GOOGLE BOOKS NUTZERBIBLIOTHEK
Nutzerbibliothek von Nutzer Alex
Online-Ressource: Webseite im Google-Books-Angebot
http://books.google.de/books?uid=16997521105364809234&source=gbs_reviews_v2_r Stand 27.08.2009
Abruf 27.08.2009

GOOGLE BOOKS REZENSION
Rezension von Nutzer Alex zum Buch Standortwahl der Unternehmen
Von Thomas Salmen in der Google Buchsuche
http://books.google.de/books?id=e43855az23QC&dq=Standortwahl&ei=JnSGSpWFApXszASbpIzhDQ
Abruf 27.08.2009

GOOGLE BOOKS STARTSEITE
Google Bücher Startseite
Online-Ressource: Webseite im Google-Books-Angebot
http://books.google.de/books
Abruf 27.08.2009

GREISDORF 2001
GREISDORF, Howard ; SPINK Amanda: Median Measure - An Approach to IR Systems evaluation. In: Information Processing and Management, Jg. 37 (2001) 6. Heft

HAWKING 2001
HAWKING, David ; CRASWELL, Nick ; BAILEY, Peter ; u.a.: Measuring Search Engine Quality. In: Journal of Information Retrieval, Jg. 4 (2001) Heft 1. Auch als Online-Ressource: http://david-hawking.net/pubs/hawking_ir01.pdf, Abruf 28.08.2009. Angaben zu Seitenzahlen bei Zitaten beziehen sich auf die Online-Ressource.

HEIDENREICH 2009
HEIDENREICH, Stephan: Für eine unendliche Bibliothek
Online-Resource: Homepage der Tageszeitung taz
http://www.taz.de/1/zukunft/konsum/artikel/1/fuer-eine-unendliche-bibliothek/ Stand 13.05.2009
Abruf 27.08.2009

HENNIES 2006
HENNIES, Markus: Clients information seeking behaviour: An OPAC transaction log analysis
Online-Ressource: Homepage zu Konferenzen der Australian Library and Information Association
http://conferences.alia.org.au/alia2006/Papers/Markus_Hennies.pdf Auch als Kurzfassung: http://conferences.alia.org.au/alia2006/presentations/Hennies.pdf
Abruf 27.08.2009

HIRTLE 2009
HIRTLE, Peter B.: Copyright Term and the Public Domain in the United States
Online Ressource: Copyright Information Center der Cornell University
http://www.copyright.cornell.edu/resources/publicdomain.cfm Stand 05.01.2009
Abruf 27.08.2009

HUANG 2004
HUANG, Mu-hsuan ; WANG, Hui-yu: The Influence of Document Presentation Order and Number of Documents Judged on Users Judgements of Relevance. In: Journal of the American Society for Information Science and Technology, Jg. 55 (2004) Heft 11

INTERNETSEITEN DER BIBLIOTHEKEN
Info-Flyer: Die UB stellt sich vor
Online Ressource: Homepage der Bibliothek der Freien Universität Berlin
http://www.ub.fu-berlin.de/service_neu/downloads/flyer-info1.pdf, S. 2
Abruf 27.08.2009

UB im Überblick
Online Ressource: Homepage der Bibliothek der Friedrich-Alexander-Universität Erlangen-Nürnberg
http://www.ub.uni-erlangen.de/
Abruf 27.08.2009

Der Bestand der Bibliothek in Zahlen
Online Ressource: Homepage der Bibliothek der Ruhr-Universität Bochum
http://www.ub.ruhr-uni-bochum.de/informationen/statistik.html
Abruf 27.08.2009

Jahresbericht 2007
Online Ressource: Homepage der Bibliothek der Universität zu Köln
http://www.ub.uni-koeln.de/e30/e12404/e12405/e15068/Jahresbericht2007_ger.pdf, S. 22
Abruf 27.08.2009

UB Zwischenbericht 2006
Online Ressource: Homepage der Bibliothek der Universität Duisburg-Essen

http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-14068/UB-Zwischenbericht_2006.pdf, S. 16
Abruf 27.08.2009

Das Bibliothekssystem der Universität Kiel
Online Ressource: Homepage der Bibliothek der Christian-Albrechts-Universität Kiel
http://www.ub.uni-kiel.de/ueber/profil/index.html
Abruf 27.08.2009

Wie viele Bücher und Medien besitzt das Bibliotheksystem?
Online Ressource: Homepage der Bibliothek der Justus-Liebig-Universität Gießen

http://www.ub.uni-giessen.de/faqs/faq_lang.php?id=257
Abruf 27.08.2009

Bestände und Besondere Aufgaben
Online Ressource: Homepage der Staats- und Universitätsbibliothek Göttingen

http://www.sub.uni-goettingen.de/index.html
Abruf 27.08.2009

Profil der Stabi
Online Ressource: Homepage der Staats- und Universitätsbibliothek Hamburg
http://www.sub.uni-hamburg.de/informationen/portrait/profil.html
Abruf 27.08.2009

Bibliotheksprofil 2009
Online Ressource: Homepage der Bibliothek der Johannes Gutenberg Universität Mainz
http://www.ub.uni-mainz.de/Dateien/Profil_2009.pdf, S. 7
Abruf 27.08.2009

INTERNETSEITEN DER VERBÄNDE

Über die Datenbank
Online Ressource: Gemeinsamer Verbundkatalog des GBV
http://gso.gbv.de/xslt/
Abruf 27.08.2009

KOBV-Index
Online Ressource: Homepage des KOBV
http://www.kobv.de/kobvindex.html
Abruf 27.08.2009

Hebis in Zahlen
Online Ressource: Homepage des Hebis
http://www.hebis.de/de/1ueber_uns/verbund/zahlen.php
Abruf 27.08.2009

Produkte des HBZ
Online Ressource: Broschüre des HBZ NRW
http://www.hbz-nrw.de/dokumentencenter/flyer/produkte.pdf
Abruf 27.08.2009

Willkommen beim Bibliotheksservice-Zentrum Baden-Württemberg (BSZ)
Online-Ressource: Startseite des Katalogs des SWB
http://www2.bsz-bw.de/cms/index_html
Abruf 27.08.2009

Statistik der Verbunddatenbank
Online-Ressource: Homepage des BVB
http://www.bib-bvb.de/bvb-statistik_2_08.htm
Abruf 27.08.2009

JELE 2001
JELE, Harald: Informationstechnologien in Bibliotheken. München [u.a.] : Oldenbourg, 2001. - 132 S., Ill., graph. Darst.

KLAGESCHRIFT
Klageschrift der Anwälte der Authors Guild im Rechtsstreit mit Google
Online-Ressource: Homepage der Authors Guild
http://www.authorsguild.org/advocacy/articles/settlement-resources.attachment/authors-guild-v-google/Authors%20Guild%20v%20Google%2009202005.pdf
Abruf 27.08.2009

KLATT 2001
KLATT, Rüdiger ; GAVRIILIDIS, Konstantin ; KLEINSIMLINGHAUS, Kirsten ; u.a.: Nutzung elektronischer wissenschaftlicher Information in der Hochschulausbildung : Barrieren und Potenziale der innovativen Mediennutzung im Lernalltag der Hochschulen (Kurzfassung)
Online-Ressource: Homepage des Projektes „Nutzung elektronischer wissenschaftlicher Informationen in der Hochschulausbildung“
http://www.stefi.de/download/kurzfas.pdf Stand Juni 2001
Abruf 27.08.2009

LEE 2006
LEE, Jin Ha ; RENEAR, Allen ; SMITH, Linda C.: Known-Item Search : Variations on a concept. In: GROVE, Andrew (Ed.): Proceedings 69th Annual Meeting of the American Society for Information Science and Technology (ASIST), Jg. 43 (2006) Band 1. Auch als Online-Ressource: http://eprints.rclis.org/8353/1/Lee_Known-Item.pdf Abruf 27.08.2009

LEWANDOWSKI 2005
LEWANDOWSKI, Dirk: Web Information Retrieval : Technologien zur Informationssuche im Internet. Frankfurt am Main : Dt. Ges. für Informationswiss. und Informationspraxis, 2005. – 248 S., graph. Darst. Auch als Online-Ressource: http://www.bui.haw-hamburg.de/fileadmin/user_upload/lewandowski/doc/Web_Information_Retrieval_Buch.pdf Abruf 28.08.2009

LEWANDOWSKI 2006
LEWANDOWSKI, Dirk: Suchmaschinen als Konkurrenten der Bibliothekskataloge : Wie Bibliotheken ihre Angebote durch Suchmaschinentechnologie attraktiver und durch Öffnung für die allgemeinen Suchmaschinen populärer machen können. In: Zeitschrift für Bibliothekswesen und Bibliographie, 53. Jg. (2006) Heft 2. Auch als Online-Ressource: http://www.bui.haw-hamburg.de/fileadmin/user_upload/lewandowski/doc/Suchmaschinen_Konkurrenten_Bibliothekskataloge_ZfBB2006.pdf Abruf 28.08.2009

LEWANDOWSKI 2006a
LEWANDOWSKI, Dirk: Bücher kostenlos zum Download. In: Password : das unabhängige Newsletter für Informationsprofis, Jg. 20 (2006) Heft 10. Auch als Online-Ressource: http://www.bui.haw-hamburg.de/fileadmin/user_upload/lewandowski/password/password10-2006.pdf Abruf 28.08.2009

LEWANDOWSKI 2007
LEWANDOWSKI, Dirk: Qualitätsmessung bei Suchmaschinen : System – und Nutzerbezogene Evaluationsmaße. In: Informatik-Spektrum, Jg. 30 (2007) Heft 3. Auch als Online-Ressource: http://www.bui.haw-hamburg.de/fileadmin/user_upload/lewandowski/doc/Qualitaets messung_bei_Suchmaschinen_Informatik_Spektrum_2007.pdf, Abruf 28.08.2009

LEWANDOWSKI 2008
LEWANDOWSKI, Dirk: Wie verändert die Einigung mit Verleger und Autoren die Buchsuche? In: Password : das unabhängige Newsletter für Informationsprofis, Jg. 22 (2008) Heft 12 Auch als Online-Ressource: http://www.bui.haw-hamburg.de/fileadmin/user_upload/lewandowski/password/password12-2008.pdf Abruf 28.08.2009

LEWANDOWSKI 2008a
LEWANDOWSKI, Dirk: Retrieval Effectiveness of Web Search Engines : Considering Results Descriptions. In: Journal of Documentation, Jg. 64 (2008) Heft 6.Auch als Online-Ressource: http://www.bui.haw-hamburg.de/fileadmin/user_upload/lewandowski/doc/retrieval_effectiveness_results_descriptions_JDoc2008.pdf Abruf 28.08.2009

LÖHRER 2007
LÖHRER, Sabina: Kataloganreicherung in Hochschulbibliotheken : State of the Art Übersicht und Aussichten für die Schweiz. Chur : Hochschule für Technik und Wirtschaft, 2007. – 124 S.

MACHILL 2003
MACHILL, Marcel: Wegweiser im Netz : Qualität und Nutzung von Suchmaschinen. Gütersloh : Verl. Bertelsmann-Stiftung, 2003. - 543 S., Ill., graph. Darst.

MILNE 2005
MILNE, Ronald: The Google Library Project at Oxford In: MILLER, William (Hrsg.): Libraries and Google. Binghamton : Haworth Information Press, 2005. – 240 S.

MORRIS 2001
MORRIS, Ruth C.: Online Table of Contents for Books : Effect on Usage. In: Bulletin of the Medical Library Association, Jg. 89 (2001) Heft 1. Auch als Online-Ressource http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=31701

OBERMEIER 1999
OBERMEIER, Franz: Schlagwortsuche in einem lokalen OPAC am Beispiel der Universitätsbibliothek Eichstätt : Benutzerforschung anhand von OPAC – Protokolldaten. In: Bibliotheksforum Bayern, Jg. 27 (1999) Heft 3

OCLC 2004
BSZ implementiert OCLC PICA'S zentrales Bibliothekssystem (CBS)
Online-Ressource: Homepage von OCLC Communications
http://www.oclcpica.org/?id=1432&ln=de&p=y Stand 21.12.2004
Abruf 28.08.2009

OCLC 2009
Lokales Bibliotheksverwaltungssystem mit integrierten Funktionen zum Katalogisieren und Bestellen
Online-Ressource: Homepage von OCLC Communications
http://www.oclc.org/de/de/lbs/default.htm
Abruf 28.08.2009

POULSEN 1996
POULSEN, Claus: Tables of Contents in Library Catalogs : A Quantitative Examination of Analytic Catalogs. In: Library Resources & Technical Services, Jg. 40 (1996) Heft 2

QUIRMBACH 2009
QUIRMBACH, Sonja: Universal Search In: LEWANDOWSKI, Dirk: Handbuch Internet-Suchmaschinen : Nutzerorientierung in Wissenschaft und Praxis. Heidelberg : Akad. Verl.-Ges., 2009. - VIII, 409 S., Ill., graph. Darst.

REMUS 2002
REMUS, Ingo: Benutzerverhalten in Online-Systemen. Potsdam : Fachhochschule Potsdam, 2002. – 134 S., graph Darst. Auch als Online Ressource: http://freenet-homepage.de/Remus/TLA.htm Abruf 28.08.2009

SARACEVIC 2007
SARACEVIC Tefko: Relevance : A Review of the Literature and a Framework for Thinking on the Notion in Information Science : Part II Nature and Manifestations of Relevance. In: Journal of the American Society for Information Science and Technology, Jg. 58 (2007) Heft 13

SPINK 2001
SPINK, Amanda ; WOLFRAM, Dietmar ; JANSEN, Major B.J. ; u.a.: Searching the Web : The Public and Their Queries. In: Journal of the American Society for Information Science and Technology, Jg. 52 (2001) Heft 3

STATISTISCHES BUNDESAMT 2009
Studierende an Hochschulen Wintersemester 2008/2009 - Vorbericht - Fachserie 11 Reihe 4.1 – 2009
Online-Ressource: https://www-ec.destatis.de/csp/shop/sfg/bpm.html.cms.cBroker.cls?cmspath=struktur,vollanzeige.csp&ID=1023717
Abruf 28.08.2009

STOCK 2000
STOCK, Mechthild ; STOCK, Wolfgang: Internetsuchwerkzeuge im Vergleich : Teil I : Retrievaltest mit Knwon Item Searches. In: Password : das unabhängige Newsletter für Informationsprofis, Jg. 14 (2000) Heft 11. Auch als Online-Ressource: http://www.phil-fak.uni-duesseldorf.de/infowiss/admin/public_dateien/files/1/1078739217password_1.pdf Abruf 28.08.2009

STOCK 2007
STOCK, Wolfgang: Information Retrieval : Informationen suchen und finden. München [u.a.] : Oldenbourg, 2007. – 599 S., Ill.

SUB 2006
Sondersammelgebiete
Online-Ressource: Homepage der Staats- und Universitätsbibliothek Hamburg http://www.sub.uni-hamburg.de/informationen/portrait/fachlich.html?aufruf Stand 11.10.2006
Abruf 28.08.2009

TAGUE-SUTCLIFFE 1992
TAGUE-SUTCLIFFE, Jean: The Pragmatics of Information Retrieval Experimentation, Revisited. In: Information Processing & Management, Jg. 28 (1992) Heft 4

TANG 1999
TANG, Rong ; SHAW, William M. ; VEVEA, Jack L.: Towards the Identification of the Optimal Number of Relevance Categories. In: Journal of the American Society for Information Science, Jg. 50 (1999) Heft 3

ULCA 2006
Vertrag zwischen der Bibliothek der Universität Kalifornien und Google
Online-Ressource: Homepage der California Digital Library
http://www.cdlib.org/news/ucgoogle_cooperative_agreement.pdf
Abruf 28.08.2009

VISE 2007
VISE, David A. ; MALSEED, Mark: Die Google-Story. Hamburg : Murmann, 2007. - 300 S., Ill.

VOORHEES 2005
VOORHEES, Ellen M. (Hrsg.) ; HARMAN, Donna K. (Hrsg.): TREC : experiment and evaluation in information retrieval. Cambridge, Mass. [u.a.] : MIT Press, 2005. - X, 462 S., graph. Darst.

WALKER 2008
WALKER, Julia: Da bleiben alle Wünsche offen
Online-Ressource: Homepage der Tageszeitung taz
http://www.taz.de/1/politik/schwerpunkt-ueberwachung/artikel/1/da-bleiben-alle-wuensche-offen/ Stand 15.07.2008
Abruf 28.08.2009

WOMSER-HACKER 2004
WOMSER-HACKER, Christa: Theorie des Information Retrieval III: Evaluierung. In: KUHLEN; Rainer (Hrsg.) ; SEEGER, Thomas (Hrsg.) ; STRAUCH, Dietmar (Hrsg.): Grundlagen der praktischen Information und Dokumentation. München : Saur, 2004. - XLI, 762 S., Ill., graph. Darst.

Anlagen
Auszug aus dem Logfile

neu Tue Jun 2 15	40	44 CEST 2009 all	schroeder
neu Tue Jun 2 15	40	47 CEST 2009 all	Ethnopolitische Konflikte
neu Tue Jun 2 15	40	48 CEST 2009 all	nancy fraser
neu Tue Jun 2 15	40	52 CEST 2009 name	Rizzini
neu Tue Jun 2 15	40	53 CEST 2009 all	Wertz Gegenstempel
neu Tue Jun 2 15	40	54 CEST 2009 all	schroeder joachim
neu Tue Jun 2 15	40	56 CEST 2009 all	Fiukowski
neu Tue Jun 2 15	40	58 CEST 2009 all	Cornelia Helfferich
neu Tue Jun 2 15	40	59 CEST 2009 all	Gegenstempel
neu Tue Jun 2 15	41	01 CEST 2009 all	Adoption? Schweden
neu Tue Jun 2 15	41	03 CEST 2009 all	Vladimir Holan
neu Tue Jun 2 15	41	12 CEST 2009 all	technische Schwingungslehre
neu Tue Jun 2 15	41	15 CEST 2009 all	Fiukowski
neu Tue Jun 2 15	41	15 CEST 2009 all	Lojo, MarÃa Rosa
neu Tue Jun 2 15	41	16 CEST 2009 all	mortimer
alt Tue Jun 2 15	41	16 CEST 2009 all	schroeder joachim
alt Tue Jun 2 15	41	18 CEST 2009 all	prs
neu Tue Jun 2 15	41	23 CEST 2009 all	Kirche Schweden
neu Tue Jun 2 15	41	24 CEST 2009 all	embruckenkurs
neu Tue Jun 2 15	41	24 CEST 2009 all	herlth, Alois
alt Tue Jun 2 15	41	26 CEST 2009 all	schwebende geschÃ¤fte
neu Tue Jun 2 15	41	29 CEST 2009 all	bruckenkurs
neu Tue Jun 2 15	41	34 CEST 2009 name	Stella tit
neu Tue Jun 2 15	41	34 CEST 2009 all	aemr
neu Tue Jun 2 15	41	35 CEST 2009 all	Reinicke Banking
neu Tue Jun 2 15	41	35 CEST 2009 all	em
alt Tue Jun 2 15	41	40 CEST 2009 all	dalai lama
neu Tue Jun 2 15	41	41 CEST 2009 all	diplomarbeiten
neu Tue Jun 2 15	41	42 CEST 2009 all	Reinicke
neu Tue Jun 2 15	41	51 CEST 2009 all	sonderpÃ¤dagogik
alt Tue Jun 2 15	41	52 CEST 2009 all	em
neu Tue Jun 2 15	41	54 CEST 2009 all	Reinicke, Wolfgang

Liste der Suchanfragen

Angola
balanced scorecard
Beleidigungen
bewegungsspiele
Christologie
Diskriminierung Integration
dokumentarfilm fernsehen
Drogenkonsum und Kriminalität
e learning in der betriebswirtschaft
einführung interkulturelle pädagogik
Elearning
ethnographie
ethnologie öffentlichkeit
fair value
fixed income
frankophonie
gothic kultur
Humankapital
jugendschutz fernsehen
kinderchirurgie
Kirchengeschichte
klima stadt
kommunikationstraining
Konzernabschlüsse
Kreuzzuege
kultur recht gesellschaft
kulturmanagement
lernumgebung
Marketingmix
medien in japan
molecular modeling
motorische entwicklung
neue soziale bewegungen
peer to peer netzwerk
Privatisierung
Projektmanagement
rechtsgeschichte des mittelalters
Rechtsphilosophie Hegel
rock'n'roll
selbstorganisiertes lernen
software wirtschaftlichkeit
Sozialkompetenz
sozialpsychologie
Sportpädagogik
Sprechakttheorie
standortwahl
straßenkinder
street art
vaterbild
Virtual Communities

Auszug aus einem Word-Dokument zur Bewertung

Die Bewertung befindet sich direkt hinter dem Titel.

Suchbegriff: **Sprechakttheorie**

Perspektiven auf Sprache : interdisziplinäre Beiträge zum Gedenken an Hans Hörmann N 4

Type	Book
Author	Hans-Georg Bosshardt
Publisher	Walter de Gruyter
Date	1986
Pages	326
ISBN	
Loc. in Archive	7
Date Added	Donnerstag, 30. Juli 2009 14:25:47
Modified	Donnerstag, 30. Juli 2009 14:29:51

Kommunikation und Grammatik : Texte, Aufgaben, Analysen J 3

Type	Book
Author	Gerd Fritz
Author	Manfred Muckenhaupt
Publisher	Gunter Narr Verlag
Date	1981
Pages	216
ISBN	3878081634, 9783878081630
Loc. in Archive	5
Date Added	Donnerstag, 30. Juli 2009 14:24:58
Modified	Donnerstag, 30. Juli 2009 14:29:08

Metaphertheorien :Typologie, Darstellung, Bibliographie J 5

Type	Book
Author	Eckard Rolf
Publisher	Walter de Gruyter
Date	2005
Pages	305
ISBN	3110183315, 9783110183313
Loc. in Archive	4
Date Added	Donnerstag, 30. Juli 2009 14:24:49
Modified	Donnerstag, 30. Juli 2009 14:28:59

Eidesstattliche Versicherung

Ich versichere, die vorliegende Arbeit selbständig ohne fremde Hilfe verfasst und keine anderen Quellen und Hilfsmittel als die angegebenen benutzt zu haben. Die aus anderen Werken wörtlich entnommenen Stellen oder dem Sinn nach entlehnten Passagen sind durch Quellenangabe kenntlich gemacht.

Ort, Datum Unterschrift