

■ RELEVANCE RANKING IN DEN ONLINE-KATALOGEN DER „NÄCHSTEN GENERATION“

von Otto Oberhauser

Relevance Ranking in Online-Katalogen ist zwar kein neues Thema, doch liegt dazu nicht allzu viel Literatur vor, die das Prädikat „ernstzunehmen“ verdient. Dies ist zum einen darin begründet, dass das Interesse an der Ausgabe ranggereihter Ergebnislisten auf Seiten aller Beteiligter (Bibliothekare, Softwarehersteller, Benutzer) traditionell gering war. Zum anderen ging die seit einigen Jahren populär gewordene Kritik an den bestehenden OPACs vielfach von einer unzureichenden Wissensbasis aus und produzierte oft nur polemische oder emotional gefärbte Beiträge, die zum Thema *Ranking* wenig beitrugen.

Schon vor mehreren Jahren haben wir versucht, den Informationsstand und die Zukunftsperspektiven zu diesem Thema auszuloten (Oberhauser & Labner, 2003). Wir hatten damals festgestellt, dass es das Feature *Relevance Ranking* – im Gegensatz zur Meinung der bald darauf einsetzenden, aber eben unzureichend informierten OPAC-Kritik – sehr wohl in einer Reihe von OPACs in verschiedenen Ländern und von verschiedenen Herstellern gab, fallweise sogar als das voreingestellte Sortierkriterium. In den auf der Software „Aleph 500“ basierenden OPACs war das Feature damals ebenfalls vorhanden (Version 14.2), aber schlecht dokumentiert und mit schwerwiegenden Performanceproblemen behaftet. Wenige Anwenderbibliotheken setzten es in ihren Katalogen ein; im Österreichischen Verbundkatalog wurde es aber tatsächlich angeboten. Mit der nicht lange danach eingeführten Aleph-Version 16 kam allerdings das „Aus“, da die Herstellerfirma nunmehr (vermutlich aus Performancegründen) für die Ausgabe von Ergebnislisten ein Anzeige- und Sortierlimit eingeführt hatte. *Relevance Ranking* – technisch gesehen eine Sortieroption – war von diesem Limit insofern betroffen, als nur mehr maximal 1000 aus dem Gesamtergebnis auf nicht transparente Weise ausgewählte Treffer sortiert bzw. angezeigt werden konnten. Später liess der Hersteller zwar eine Parametrisierung auf höhere Trefferzahlen zu, doch de facto hätte im Verbundkatalog eine Erhöhung des Limits einen deutlichen Performanceverlust bedeutet und war somit nicht möglich. Angesichts der Absurdität dieser Situation – *Relevance Ranking* soll ja aus grossen Treffermengen die relevantesten Titel an die Spitze reihen, was aber bei einer zuvor durch ein solches Limit gehaltenen Ergebnisliste nicht möglich ist – musste das Feature aus dem

Katalog entfernt werden. Dies ist bis heute so, denn die Herstellerfirma hielt es nicht für nötig, diesbezüglich grundsätzliche Änderungen vorzunehmen. Dass die Konkurrenz seit Jahr und Tag – z.B. im OPAC des GBV – grosse und grösste Treffermengen blitzschnell zu ranken vermochte, war wohl ohne grosse Bedeutung, da die bibliothekarische Klientel dem Feature ohnedies zum überwiegenden Teil gleichgültig gegenüberstand.

Erst als vor einigen Jahren die oben angedeutete OPAC-Kritik einsetzte, wurde Relevance Ranking plötzlich zu einem Thema. „Experten“ nahmen die bestehenden Online-Kataloge und ihre Features kritisch ins Visier und „entdeckten“ Schwachstellen, die Personen mit offenen Augen und informationswissenschaftlichen Basiskenntnissen seit langem kannten. Jeder, der etwa mit Online- oder CD-ROM-Datenbanken zu tun gehabt hatte, wusste, dass bei den OPACs vieles im Argen lag, nur die Softwarehersteller und ihre Bibliothekskunden wollten dies lange nicht wahrhaben. In diese Kritik, die möglicherweise sogar von der nach jahrelangem Dornröschen-schlaf an einer Neubelebung des Geschäftes interessierten Softwareindustrie initiiert wurde, stimmten schliesslich einige Bibliothekare und zahlreiche Endbenutzer (auch Hochschullehrer) ein, unter letzteren peinlicherweise oft solche, deren Informationskompetenz sich im Eintippen einzelner Wörter in den Google-Suchschlitz erschöpfte. Immerhin aber spiegelte die Kritik das Unbehagen wider, das gerade dabei war, die bibliothekarische Welt zu durchdringen. Immer mehr war nämlich der – sogar empirisch untermauerte – Eindruck entstanden, dass sich in den Benutzergewohnheiten ein Wandel vollzog – weg von den Online-Katalogen und hin zu der scheinbar omnipotenten Suchmaschine Google, die angeblich so leicht zu bedienen sei. (Dass letzteres unrichtig ist – wirklich gut in Google zu recherchieren ist nämlich gar nicht einfach – war und ist freilich nur wenigen bewusst.) Die „Schuld“ für den offenkundigen Wandel des Benutzerverhaltens wurde spontan den Online-Katalogen gegeben und nicht etwa auch Faktoren wie

- der Trägheit und Beschränktheit vieler Endbenutzer, die sich willfährig von Google vorgaukeln lassen, dass dort die Suche einfach und umfassend sei;
- den unterschiedlichen Inhalten der Datenpools (Metadaten in den OPACs vs. vielfach direkt beziehbare Volltexte in Google);
- der seit Jahrzehnten verfehlten und auch durch die Automation nicht geänderten Bearbeitungspraxis der Bibliotheken (veraltete Regelwerke, keine Aufnahme analytischer Datensätze, unzureichende sachliche Erschliessung).

Kommen wir zurück zum Relevance Ranking. Für Google (und andere Suchmaschinen) sprach und spricht natürlich auch, dass diese die Suchresultate typischerweise nach Relevanz sortiert ausgeben, sodass sich mit einem Glück die relevanten Treffer auf der ersten und vielleicht zweiten Seite finden. Mehr anzusehen – so hat man in empirischen Untersuchungen herausgefunden – lässt die Trägheit vieler Benutzer heute nicht mehr zu. Relevance Ranking ist somit primär zu einem Bequemlichkeitsfaktor geworden, den viele herkömmliche OPACs, die die Ergebnisse alphabetisch oder chronologisch reihen, nicht bieten können. Für auf „Suchmaschinentechnologie“ basierende Rechercheinstrumente ist Relevance Ranking dagegen ein Standardfeature, vermutlich oft besser programmiert und implementiert als in der bisherigen Katalogsoftware (sofern dort überhaupt vorhanden).

Vielfach sieht die verunsicherte Bibliothekswelt gegenwärtig in der offenbar wie eine Blackbox hingenommenen „Suchmaschine“ den Rettungsanker, um in der Konkurrenz gegen den allmächtigen Koloss Google nicht unterzugehen. Allerdings ist in keiner Weise bewiesen, dass diese Strategie auch aufgehen wird (die obige Anmerkung über die traditionsverhafteten Bearbeitungspraktiken lässt vielmehr daran zweifeln). Aufgegangen ist da gegen die Taktik mancher Softwarelieferanten, die für ihre Suchmaschinen-Lösungen (die mitunter modifizierte Open-Source-Produkte sind) astronomische Preise verlangen und auch bezahlt bekommen. Wäre es nicht vielmehr die moralische Verpflichtung dieser Hersteller gewesen, neue Katalogsoftware im Rahmen der auch so schon hohen Anschaffungs-, Lizenz- und Wartungskosten für die bestehenden Bibliothekssysteme zu entwickeln und bereitzustellen?

Inzwischen sind an etlichen Einrichtungen auf „Suchmaschinentechnologie“ basierende OPACs in Betrieb, auch in Europa. Viele weitere sind in Vorbereitung und werden folgen. Oft werden sie – mit einem etwas unglücklichen, da zeitpunktabhängigen Ausdruck – als „Bibliothekskataloge der nächsten Generation“ (next-generation library catalogues) bezeichnet. Eine oft bessere Performance, ein meist moderneres Webdesign, eine adäquate Berücksichtigung von Texten aus der Kataloganreicherung sowie Standardfeatures wie Relevance Ranking und Facettennavigation zeichnen diese Systeme aus. Daneben werden häufig verschiedene Möglichkeiten der benutzerseitigen Eingabe von Bewertungen, Rezensionen, Schlagwörtern usw. angeboten, über deren Wert man verschiedener Meinung sein kann. Diese „Web 2.0“-Features werden in der (deutschen) Literatur vor allem von einer sich im Kreis zitierenden Lobby gefordert, wobei entsprechende Benutzererwartungen taxfrei unterstellt, keineswegs aber bewiesen werden. Die betreffenden Beiträge sind vielmehr durch Begeisterung und

etwas naive Sozialromantik charakterisiert. Demgegenüber attestieren die wenigen wirklichen Untersuchungen des Benutzerverhaltens meist nur Interesse an verbesserten Such- und Ausgabefunktionalitäten, weniger bzw. gar nicht dagegen an den „Web 2.0“-Features (vgl. z.B. Tam et al., 2009, sowie die dort zitierten Studien).

Relevance Ranking wird, so kann man diesen Studien entnehmen, somit auch benutzerseitig erwartet bzw. gewünscht, in der Regel wohl auch als die Standardsortierung der vom System ausgegebenen Ergebnislisten. Letzteres wird im Suchmaschinenkontext auch anwenderseitig nicht in Frage gestellt, ja nicht einmal mehr diskutiert (wie noch im Fall der herkömmlichen OPACs). Dass dabei kaum darüber nachgedacht wird, wie ein solches Relevance Ranking zustande kommt, kann hinsichtlich der Benutzer als ziemlich sicher gelten, hinsichtlich der Softwareanwender als recht wahrscheinlich. Die vielfach unreflektierte Übernahme der kurzsichtigen Maxime „Make It Like Google“ führt offenbar dazu, dass davon ausgegangen wird, die Blackbox „Suchmaschine“ könne dies quasi „von alleine“. Google beherrscht Relevance Ranking aber keineswegs „von alleine“, sondern mittels des raffinierten und seit seiner ersten Publikation vielfach (im Hintergrund) weiterentwickelten „PageRank“-Verfahrens, das allerdings für Web-Dokumente – und nicht für bibliothekarische Metadaten – massgeschneidert wurde. Für letztere sieht die Situation nämlich ganz anders aus.

Die traditionellen Gewichtungsverfahren der Information-Retrieval-Forschung beruhen darauf, dass die gefundenen Dokumente für die Ausgabe nach einem für sie ermittelten Gewicht absteigend gereiht werden, wobei die Gewichte auf der Basis der Häufigkeit des Auftretens der Begriffe in den Dokumenten (*term frequency* oder TF) bzw. der inversen Häufigkeit ihres Auftretens in der Datenbank (*inverse document frequency* oder IDF) berechnet werden. Dieser Ansatz gibt den Termen, die in einem Dokument oft, in der gesamten Kollektion jedoch selten auftreten, die höchsten Gewichte, und denjenigen Termen, die im Dokument selten, in der Kollektion jedoch häufig auftreten, die niedrigsten Gewichte.

Schon vor fast zwanzig Jahren wurde allerdings in einer Untersuchung mit MARC-Datensätzen festgestellt, dass das bewährte TFIDF-Verfahren in diesem bibliothekstypischen Kontext zu auffallend schlechten Ergebnissen führte (Larson, 1992). Dieser Befund wurde auf die Begrenztheit des Vocabulars zurückgeführt, zumal die verwendeten MARC-Datensätze nur Titel und Subject Headings (LCSH) enthielten. Im Gegensatz dazu hatte die traditionelle Information-Retrieval-Forschung, in der sich der TFIDF-Ansatz bewährt hatte, meist auf Datensätze zurückgegriffen, die neben dem Titel auch ein Abstract und oft eine grosse Zahl von Deskriptoren, wenn nicht

gleich auch Volltexte, aufwiesen. Einige Jahre später gab ein anderer Autor zu bedenken, dass „der Einsatz von z.B. Algorithmen zum Relevance-Ranking oder einer sog. Fuzzy-Suche (...) nur dann sinnvoll möglich [sei], wenn eine ausreichende Textbasis als Grundlage für die Rechenoperationen zur Verfügung steht.“ Und weiter: „Die Gestaltung fortschrittlicher Information-Retrieval-Systeme für Bibliotheken führt daher zunächst über die Veränderung und die Erweiterung des bibliothekarischen Dokuments.“ (Lepsky, 1996, 66). In jüngerer Vergangenheit verwiesen etwa Dellit & Boston (2007) und Beall (2008) darauf, dass die Metadaten der Kataloge meist nicht genug Text enthielten, um beim Text-Matching zu guten Ergebnissen zu führen. Und sogar jene Autorin, die durch ihre OPAC-Kritik im Vulgärgargon bekannt geworden war, musste inzwischen kleinlaut eingestehen, dass Relevance Ranking bibliothekarischer Daten keine so einfache Sache ist wie ursprünglich behauptet, sondern vielmehr „a complex soup“ (Schneider, 2007). Nachfolgend sollen daher drei rezente Beiträge vorgestellt werden, die sich ernsthaft mit der Frage auseinandersetzen, wie OPAC-Datensätze in Ergebnislisten – über ein einfaches TFIDF-Ranking hinausgehend – ranggereiht werden können. Dabei geht es einerseits um die Ausarbeitung feldspezifischer und zum Teil von bestimmten Vorbedingungen abhängigen Gewichtungseinstellungen, andererseits aber auch um die Einbeziehung weiterer Faktoren in den Rankingalgorithmus.

Der erste Artikel stammt von dem Hamburger Hochschullehrer Dirk Lewandowski (2009), der fünf Gruppen von Faktoren diskutiert, die für das Ranking im bibliothekarischen Kontext anwendbar sein können:

- *Text matching*: Hier wird die eben erwähnte Problematik des TFIDF-Rankings angesprochen und vorgeschlagen, nicht nur (wie bisher) eine unterschiedliche Gewichtung bestimmter Felder wie Titel oder Schlagwort vorzunehmen, sondern Datensatztypen mit unterschiedlich langen Volltextelementen separat zu ranken und erst danach in einer Ergebnisliste zu kombinieren. Die Frage ist freilich, wie das konkret geschehen soll.
- *Popularity*: Dies ist primär ein Aspekt für Kataloge lokaler Bibliotheken, wo Ausleihfrequenz, Zahl der vorhandenen Exemplare oder Häufigkeit des Downloadens von elektronischen Dokumenten Kriterien der Beliebtheit sind. Sofern Benutzerbewertungen vorliegen, könnten diese ebenfalls in das Ranking (auch bei überregionalen Katalogen) miteingehen. Als Basis für die Beliebtheitseinstufung kommen sowohl einzelne Titel als auch Gruppierungen von Titeln (z.B. Werke desselben Autors, Bücher desselben Verlages oder aus derselben Reihe) in Frage.

- *Freshness*: Lewandowski meint, dass die meisten herkömmlichen OPACs chronologisch sortieren, doch mindestens ebensoviele reihen ihren Output alphabetisch nach Autor/Titel. Aktualität ist sicherlich kein einfach zu gewichtendes Kriterium, da ihre Bedeutung fachgebietsspezifisch stark variieren mag. Als Masszahlen kommen sowohl Erscheinungsdatum als auch Zugangsdatum in Frage.
- *Locality*: Gemeint sind hier sowohl der physische Standort des Benutzers (z.B. die Nähe zu einer Zweigbibliothek) als auch die Verfügbarkeit eines Exemplars. Wenn sich Benutzer ausserhalb der Bibliothek aufhalten, mag ein elektronisches Exemplar interessanter als ein physisches sein, wogegen in der Bibliothek diese Einschränkung wegfällt.
- *Other ranking factors*: Hier werden noch Aspekte wie die Grösse oder der Typ der Publikation in Erwägung gezogen, aber auch Benutzergruppen (z.B. sind Lehrbücher für Studienanfänger relevanter als für Fortgeschrittene). Ein schwierig zu implementierender Faktor wäre sicherlich die Art der Anfrage selbst: Bei einer thematisch breiten Anfrage wäre wohl eine „gute Mischung“ von Ergebnissen (Nachschlagewerk, Lehrbuch, relevante Zeitschrift und/oder Datenbank) erwünscht, wogegen es bei spezifischeren Anfragen sinnvoll wäre, mehr Zeitschriftenartikel (und Konferenzvorträge) anzuzeigen.

Zu Recht argumentiert Lewandowski abschliessend, dass es die Aufgabe der Informationsfachleute sei, zu bestimmen, wie gute Resultate zu stande kommen. „We cannot expect computer scientists or information system vendors to do this work for us. We know our material best and should therefore best know how a good results set could be achieved.“ (ibid., 592). Diese Aussage bezieht sich wohl auf die mitunter überheblichen, zumal nichtssagenden Behauptungen mancher Softwarehersteller. Die Parametrisierung des Relevance Ranking sollte daher unter der Kontrolle der anwendenden Bibliotheken stattfinden – sich auf das gekaufte Softwareprodukt bzw. die Fähigkeit von dessen Hersteller, einen guten Ranking-Algorithmus zu liefern, zu verlassen, wird vermutlich nicht immer zu dem gewünschten Ergebnis führen.

Der zweite hier zu referierende Artikel soll zeigen, wie das in der Praxis funktioniert. Annette Langenstein und Leonhard Maylein, zwei Kollegen aus der UB Heidelberg, haben nämlich in verdienstvoller Weise offengelegt (2009), wie sie das Relevance Ranking ihres lokalen, auf der Open-Source-Suchmaschine Lucene basierenden Bibliothekskatalogs HEIDI parametrisiert haben, d.h. wie an welchen „Stellschrauben“ gedreht wurde.

Herausragend ist der Beitrag vor allem auch deshalb, da kaum jemals zuvor Ähnliches im Detail publiziert wurde. Insbesondere die Hersteller kommerzieller Software tendieren ja dazu, sich hinter nebulosen Statements zu verstecken oder zum Thema überhaupt zu schweigen. Langenstein & Maylein nennen dagegen nicht nur die konkret parametrisierten Faktoren, sondern auch die jeweils vergebenen Gewichte. Sie unterscheiden prinzipiell zwei Massnahmen zur Anpassung des Rankings, nämlich *statische*, die nur beim Neuaufbau eines Index verändert werden können, und *dynamische*, die jederzeit verändert werden können.

- *Statische Massnahmen* (greifen bereits während des Indexaufbaus und sind von der Suchanfrage selbst unabhängig):
 - Feldlänge: Lucene (eigentlich ja auch für Web-Dokumente gedacht) berücksichtigt in seiner Rankingformel die Länge der Felder, in denen die Suchbegriffe gefunden werden. In Heidelberg wurde entschieden, wegen der Struktur der bibliothekarischen Daten auf dieses Kriterium zu verzichten.
 - Dokumenten-Boost: Das System erlaubt es, einzelne Katalogisate höher oder niedriger zu gewichten. In Heidelberg wurde entschieden, Gesamtaufnahmen von Zeitschriften und Zeitungen höher zu gewichten als einzelne Bände, neuere Auflagen eines Titel höher als ältere sowie neuere Titel (in geringem Masse) höhere als ältere.
 - Feld-Boost: Auf die prinzipielle Möglichkeit, feldbezogene Boostfaktoren bereits bei der Indexierung einzusetzen, wurde verzichtet; statt dessen setzte man auf eine Feldgewichtung über die Suchanfrage).
- *Dynamische Massnahmen* (beeinflussen das Ranking erst bei der Suche):
 - Feld-Boost bei der Suche: Nachdem zunächst für die feldübergreifende Freitextsuche keine unterschiedliche Gewichtung der Ursprungsfelder vorgesehen worden war, wurde dies später gewünscht und nachträglich eingerichtet. Dabei wurden „Einworttitel“ mit einem höheren Gewicht, Volltexte aus Kataloganreicherungen mit einem niedrigeren Gewicht versehen. Auf eine unterschiedliche Gewichtung von Feldern wie Titel, Autor, Körperschaft, Schlagwörter usw. wurde weiterhin verzichtet.
 - Phrasensuche: Bei Sucheingaben mit mehreren Wörtern erfolgt für Treffer mit exakter Übereinstimmung der Phrase eine sehr starke und für solche mit teilweiser Übereinstimmung eine starke Höhergewichtung.

Langenstein & Maylein betonen, dass auch nach zwei Jahren ihre Parametrisierung des Relevance Rankings noch nicht abgeschlossen ist. Noch gar nicht implementiert wurden etwa die auf „Popularity“ basierenden Faktoren (Ausleihefrequenz, Verfügbarkeitsstatus, Exemplarzahl, Nutzerbewertungen bzw. statistisch ermittelte Empfehlungen). Außerdem könnten auch noch individuelle Nutzerprofile miteinbezogen werden. Ziel einer empirischen Untersuchung der Ranking-Ergebnisse soll es sein, auch auf dieser Basis eine weitere Optimierung vorzunehmen. Zu Vergleichszwecken zitieren die Autoren übrigens ein kurzes Webdokument der Australischen Nationalbibliothek, in dem diese die Ranking-Einstellungen für ihren ebenfalls mit Lucene betriebenen OPAC beschreibt (National Library of Australia, o.J.).

Dies leitet über zum dritten Beitrag (Dellit & Boston, 2007), in dem die Autoren über die Modernisierung des überregionalen Katalogs *Libraries Australia* berichten. Dabei ging es neben anderen Features auch um die Implementierung eines über das traditionelle TFIDF-Verfahren hinausreichenden Relevance Rankings. Dellit & Boston führen zunächst die Datenquellen an, die für ein solches Ranking in Frage kommen. Neben bibliographischen Feldern (insbesondere Schlagwort, Autor, Titel) werden genannt: Bestandsinformationen (Zahl der besitzenden Bibliotheken), mit Titeln korrelierte Leserdaten (z.B. Alter der Leser), Verfügbarkeitsinformation (für den Leser erreichbare lokale Bestände in Bibliotheken oder auch im Buchhandel), Information aus Zitatanalysen sowie benutzerseitig generierte Daten (Ausleihzahlen, Benutzer-Rezensionen, -Schlagwörter oder -Empfehlungen). Bislang hat man sich bei der Implementierung jedoch auf die Informationen aus den bibliographischen Datenfeldern und den Bestandangaben beschränkt. Diesem Verfahren wurden die folgenden Regeln zugrunde gelegt:

1. Übereinstimmungen (matches) in den Feldern Autor, Titel, Schlagwort, Format und Dokumententyp sind (ge-)wichtiger als solche im sonstigen Text.
2. Matches in mehreren der obigen Felder sind wichtiger als Matches in nur jeweils einem dieser Felder.
3. Die exakte Übereinstimmung (exact match) des Suchbegriffs bzw. der Suchbegriffe mit dem Inhalt des Datenfeldes Titel ist wesentlich wichtiger als die Übereinstimmung mit einem Teil dieses Datenfeldes (phrase match), die wiederum wichtiger ist als die Übereinstimmung auf Wortebene (word match), bei der alle Suchbegriffe im Datenfeld vorkommen, aber in beliebiger Reihenfolge.
4. Regel 3 gilt in modifizierter Weise für Autor bzw. Schlagwort. Bei Autorennamen gilt ein exact match nur als ein wenig besser als ein phrase

match, doch gelten beide als viel besser als ein *word match*. Im Falle der Schlagwörter werden die unterschiedlichen Grade der Übereinstimmung weniger stark differenziert gewichtet.

5. Übereinstimmung in den „main entry fields“ (MARC-Kategorien 245\$a,b und 100, d.h. Titel/Untertitel und erster Autor) ist wichtiger als Übereinstimmung in „additional entry fields“ oder „alternate titles“ (weitere Autoren und Titelformen).
6. Exakte Übereinstimmung mit der benutzerseitig eingegebenen Schreibweise von Suchbegriffen wird höher gewichtet als Übereinstimmung mit trunkierten oder vom System korrigierten Varianten von Suchbegriffen.
7. Wenn alles andere gleich ist, gelten „collection-level records“ (Bindeeinheiten und andere bibliographische Agglomerate) als relevanter als „non-collection level items“.
8. Werke mit grösseren Bestandszahlen in den Bibliotheken dieses überregionalen Katalogs gelten als relevanter als Werke mit geringeren Beständen.

Diese drei Beiträge zeigen, in welche Richtung die Entwicklung verläuft. In allen drei wird gezielt auf ein Ranking hingearbeitet (konzeptionell bzw. praktisch), das die im Kontext kurzer Metadaten-Records begrenzte Brauchbarkeit des einfachen TFIDF-Verfahrens zu überwinden sucht und das für Datenbanken mit typischen bibliothekarischen Datensätzen optimiert ist. Besonders die beiden aus der Praxis kommenden Berichte zeigen aber auch, dass hier eine grosse Vielfalt unterschiedlicher Ansätze möglich ist. Um es bildlich zu formulieren – es gibt zahlreiche „Stellschrauben“ und diese können auf vielerlei Art gedreht werden. Und manche Frage bleibt auch nach der Lektüre dieser Beiträge offen – was muss z.B. im Fall unvollständiger Daten sinnvollerweise geschehen (etwa beim Fehlen von Schlagwörtern in einer grossen Zahl von Datensätzen)? Die Erarbeitung komplexer Rankingverfahren ist typischerweise ein iterativer Prozess, der sich über einen grösseren Zeitraum hinziehen kann. Dazu gehören auch Tests der jeweils erreichten Güte des Rankings, die gar nicht so leicht zu realisieren sind. Der zuletzt referierte Beitrag aus Australien berichtet auch kurz über bereits getätigte Anstrengungen in dieser Richtung; der Beitrag aus Heidelberg stellt solche Benutzertests erst in Aussicht.

Zum Abschluss soll punktuell – d.h. anhand einer bekannten Suchaufgabe – überprüft werden, wie gut das Relevance Ranking in Suchmaschinen-OPACs derzeit funktioniert. Gerne wird ja als Vorteil von Google gegenüber konventionellen OPACs angeführt, dass die Suchmaschine es schafft,

die Webseiten zu den wichtigen Zeitschriften „Nature“ und „Science“ an die Spitze der Ergebnislisten zu reihen, während im traditionellen Online-Katalog die feldunspezifische Stichwortsuche mit derartigen Suchbegriffen üblicherweise zu keinen brauchbaren Ergebnissen führt. Insbesondere ist interessant, ob die eben vorgestellte Ranking-Einstellung im Heidelberger OPAC die beiden Zeitschriften an die Spitze reiht, ob die Australische Nationalbibliothek dies ebenfalls schafft und welche Ergebnisse zum Vergleich verschiedene andere „next-generation“ OPACs erbringen. Als zweites Beispiel wird die Wochenzeitung „Die Zeit“ herangezogen, bei der es sich zwar aus bibliothekarischer Sicht um einen „Einworttitel“ handelt, aus Benutzersicht möglicherweise aber nicht. Der Titel sollte daher idealerweise sowohl *ohne* führenden Artikel als auch *mit* diesem an die Spitze der Ergebnisliste sortiert werden. Die Tabelle im Anhang zeigt das Resultat dieses kleinen Tests.

Die Abfragen wurden Ende Januar 2010 durchgeführt. Im jeweiligen Katalog wurden – ungeachtet allfälliger sonstiger Suchmöglichkeiten – die Suchwörter „science“ und „nature“ im Suchschlitz für die feldunspezifische Suche eingegeben und in der resultierenden Ergebnisliste die Position des ersten relevanten Treffers ermittelt, wobei gedruckte und Online-Ausgaben der Periodika nicht unterschieden wurden. Je nach System wurden zwei oder drei Ergebnisseiten (30–50 Treffer) durchgesehen; wenn dabei das gesuchte Periodikum nicht aufschien, wurde ein „N“ in die Tabelle eingetragen. Im Falle der „Zeit“ wurden dagegen zwei Suchen – sowohl mit „zeit“ als auch mit „die zeit“ – durchgeführt.

Im Grossen und Ganzen erbrachte diese knappe Momentaufnahme ein überraschend positives Ergebnis. Was die Zeitschriften „Science“ und „Nature“ betrifft, so werden diese von der Mehrheit der hier getesteten Kataloge an den ersten Platz bzw. einen der ersten Plätze der Ergebnisliste gereiht. Bei der „Zeit“ muss hingegen ins Kalkül gezogen werden, dass vermutlich einige der untersuchten Bibliotheken diese deutsche Wochenzeitung nicht im Bestand haben (was nicht weiter überprüft werden konnte). Der Name des Blattes („Die Zeit“) erweist sich in manchen Katalogen als problematisch, sofern mit „zeit“ alleine gesucht wird. Gibt man hingegen „die zeit“ ein, qualifiziert sich der Datensatz dann oft doch an vorderer Stelle. Auffallend ist, dass genau dies bei den beiden deutschsprachigen Primo-Katalogen (Mannheim, Innsbruck) der Fall ist.

Im oben erwähnten Katalog der UB Heidelberg rangieren alle drei Periodika, wenn man sie als Einworttitel recherchiert, an der ersten Stelle der Ergebnisliste; die Eingabe von „die zeit“ ergibt die dritte Position. Im australischen Katalog führen „science“, „nature“ und „die zeit“ zur ersten

Stelle, „zeit“ als Einworttitel zur zweiten Position. Die geschilderten Bemühungen, ein gutes Ranking zu erreichen, haben sich offenbar gelohnt, zumindest was diesen punktuellen Test betrifft. Gute Ergebnisse erzielten aber auch andere Systeme, auch solche von kommerziellen Herstellern, die meist keine Einblicke in ihre Ranking-Algorithmen gestatten. Schwer erkläbar sind – ohne detaillierte Kenntnis des betreffenden Umfelds – die nicht zufriedenstellenden Resultate einiger Kataloge, insbesondere dann, wenn das dem Katalog zugrundeliegende Produkt in anderen Bibliotheken sehr gute Ergebnisse erbrachte (vgl. etwa Ottawa oder Oxford). Im Fall der Endeca-Vorzeigeanwendung NCSU (North Carolina) kann ebenfalls nicht beurteilt werden, warum diese keine guten Resultate erbringt, die dort parallel eingesetzte Suchmaschine „Summon“ dagegen schon.

Der hier beschriebene Test ist natürlich in keiner Weise erschöpfend oder repräsentativ. Dennoch gibt er, wie ich glaube, Anlass zu einiger Hoffnung. Er lässt vermuten, dass die „neuen“ OPACs – zumindest was das Relevance Ranking betrifft – auf dem Weg in die richtige Richtung sind. Wie gut es wirklich gelingen wird, die Rankingleistung von Suchmaschinen wie Google, die unter völlig anderen Voraussetzungen arbeiten, einzuhören, wird aber erst die Zukunft zeigen.

Anmerkung

Ich danke Martin Hekele, Markus Knitel und Josef Labner (alle OBVSG) für die Durchsicht des Manuskripts.

Dr. Otto Oberhauser MPhil MSc MLIS
Die Österreichische Bibliothekenverbund und Service GmbH
Brünnlbadgasse 17 / 2A, A-1090 Wien
<http://oco.vfi-online.org>

Zitierte Literatur

- Beall, J. (2008) The weakness of full-text searching. *Journal of academic librarianship*. 34 (5). 438–444.
- Dellit, A.; Boston, T. (2007) Relevance ranking of results from MARC-based catalogues: From guidelines to implementation exploiting structured metadata. Paper presented at *Information online 2007*, Sydney, NSW; 30th January – 1st February 2007. Online: <http://conferences.alia.org.au/online2007/Presentations/30Jan.B1.relevance.ranking.pdf> [30.12.2009]

- Langenstein, A.; Maylein, L. (2009) Relevanz-Ranking im OPAC der Universitätsbibliothek Heidelberg. *B.I.T. online*. 12 (4). 408–413.
- Larson, R. R. (1992) Experiments in automatic Library of Congress Classification. *Journal of the American Society for Information Science*. 43 (2). 130–148.
- Lepsky, K. (1996) Vom OPAC zum Hyperkatalog: Daten und Indexierung. In: Hermes, H.-J.; Wätjen, H.-J. (Hrsg.) *Erschliessen, Suchen, Finden: Vorträge aus den bibliothekarischen Arbeitsgruppen der 19. und 20. Jahrestagungen (Basel 1995 / Freiburg 1996) der Gesellschaft für Klassifikation*. Oldenburg: Bibliotheks- u. Informationssystem der Universität. 65–73.
- Lewandowski, D. (2009) Ranking library materials. *Library hi tech*. 27 (4). 584–593.
- National Library of Australia (o.J.) Set of rules for Lucene relevance ranking. Online: <http://li01.nla.gov.au/docs/LuceneRRNotes.html> [30.12.2009]
- Oberhauser, O.; Labner, J. (2003) Relevance Ranking in Online-Katalogen: Informationsstand und Perspektiven. *Mitteilungen der VÖB*. 56 (3/4). 49–63.
- Schneider, K. G. (2007) Relevance ranking and OPAC records. Online: <http://freerangelibrarian.com/2007/07/25/relevance-ranking-and-opac-records/> [30.12.2009]
- Tam, W.; Cox, A. M.; Bussey, A. (2009) Student user preferences for features of next-generation OPACs. *Program: Electronic library and information systems*. 43 (4) 2009. 349–374.

Institution	System	Web-Adresse	„Science“	„Nature“	„Zeit“	„Die Zeit“
<i>University of Chicago Library</i>	Aquabrowser	http://lens.lib.uchicago.edu/	2	11	19	3
<i>Edinburgh University Library</i>	Aquabrowser	http://aquabrowser.lib.ed.ac.uk/	1	11	N	N
<i>Elektronische Bibliothek SUB Bremen</i>	E-LIB	http://elib.suub.uni-bremen.de/	N	N	N	N
<i>Trinity College Library Dublin</i>	Encore	http://stella.catalogue.tcd.ie/iii/encore/home?lang=eng	2	2	N	N
<i>Hongkong Polytechnic University Library</i>	Encore	http://encore.lib.polyu.edu.hk/	1	1	N	N
<i>University of Toronto Libraries</i>	Endeca	http://discover.library.utoronto.ca/catalogue	1	2	N	13
<i>Florida State University</i>	Endeca	http://fsu.catalog.fcla.edu/	1	3	1	1
<i>McMaster University Libraries</i>	Endeca	http://libcat.mcmaster.ca/	6	29	1	17
<i>North Carolina State University Libraries</i>	Endeca	http://www.lib.ncsu.edu/catalog/	N	N	N	N
<i>North Carolina State University Libraries</i>	Summon (Beta)	http://www.lib.ncsu.edu/summon/	1	1	N	N
<i>University of Ottawa Library</i>	Endeca (Beta)	http://catalogue.bib.uottawa.ca/html/?language=en	N	N	8	1
<i>Universitätsbibliothek Heidelberg</i>	Heidi (Lucene)	http://katalog.ub.uni-heidelberg.de	1	1	1	3
<i>Australian National Library</i>	VuFind (Solr / Lucene)	http://catalogue.nla.gov.au/	1	1	2	1
<i>Wissenschaftspark Albert Einstein</i>	Albert (Lucene)	http://waesearch.kobv.de/	2	1	N	N
<i>Vanderbilt University J&S Heard Library</i>	Primo	http://discoverlibrary.vanderbilt.edu/	1	1	3	1
<i>Oxford University Library Services</i>	Primo	http://solo.ouls.ox.ac.uk	N	N	2	N
<i>Royal Library Copenhagen</i>	Primo	http://www.kb.dk/en/ext_links/REX	1	1	N	1
<i>Universitätsbibliothek Mannheim</i>	Primo	http://www.bib.uni-mannheim.de/133.html	1	1	25	3
<i>Universitäts- und Landesbibliothek Tirol (UB Innsbruck)</i>	Primo	http://search.obvsg.at/UILB	1	2	N	1
<i>British Library</i>	Primo	http://searchbeta.bl.uk/	1	2	N	6