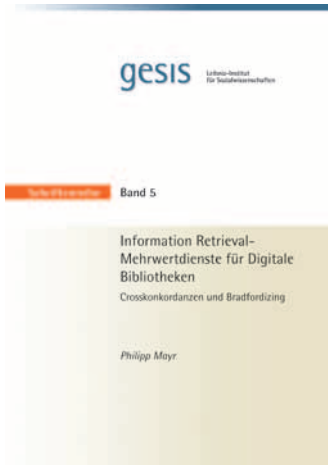


Information Retrieval-Mehrwertdienste für Digitale Bibliotheken: Crosskonkordanzen und Bradfordizing

Philipp Mayr – Bonn: GESIS – Leibniz-Institut für Sozialwissenschaften, 2010 (GESIS-Schriftenreihe; 5). 270 S.; ISBN 978-3-86819-012-0 ; 29,00 Euro



In den vergangenen Jahren haben sich auch die Anbieter von Fachdatenbanken verstärkt mit dem Thema Relevance Ranking beschäftigt bzw. beschäftigen müssen. Vor allem dem Druck durch die allgemeinen Suchmaschinen wie Google ist es geschuldet, dass Nutzer auch von diesen Informationssystemen eine nach angenommener Relevanz geordnete Trefferliste erwarten. Während es viele Versuche gab, allgemeine Rankingkriterien auf die Literaturnachweise der Fachdatenbanken anzuwenden, fehlt es bislang an Modellen, die Dokumente nach Kriterien sortieren, die den Ansprüchen der

Wissenschaft genügen bzw. einem Ranking, wie es Wissenschaftler selbst vornehmen würden, entsprechen. In diesem Bereich ist das Buch von Philipp Mayr angesiedelt, welches auf seiner 2008 an der Humboldt-Universität zu Berlin eingereichten Dissertation basiert.

Die Arbeit behandelt zwei separat voneinander stehende Bereiche: Einerseits geht es um sog. Crosskonkordanzen, also das Mapping von Thesauri, womit Recall bei Recherchen erhöht wird; andererseits um das Verfahren des Bradfordizing, welches Dokumente anhand ihrer Zugehörigkeit zu sog. Kernzeitschriften sortiert. Dabei ist das Ziel, „nach dem Re-Ranking mehr relevante Dokumente für ein Topic im ersten Drittel der Ergebnismenge zu listen.“ (S. 95)

Der Autor benennt als übergeordnetes Ziel seiner Arbeit, „die beiden Verfahren unabhängig voneinander zu beschreiben und über IR-Tests zu evaluieren. [...] Bradfordizing kann im weitesten Sinne auch als Kompensationsverfahren für die erweiterten Suchräume des vorherigen Mehrwertdienstes beschrieben werden. Im Fokus der Arbeit liegt aber die Evaluation der Einzelverfahren.“ (S. 18) Der Schwerpunkt (und wohl auch das Hauptinteresse des Autors) liegt allerdings beim Bradfordizing.

Der Zusammenhang der beiden Themen kann insofern hergestellt werden, dass nach dem Abschicken einer Suchanfrage durch die Verwendung der Crosskonkordanzen zunächst einmal die *Menge der potentiell relevanten Dokumente* erweitert wird, um sie schließlich durch das Bradfordizing so zu sortieren, dass im oberen Drittel der Ergebnisliste ein größerer *Anteil hoch relevanter Dokumente* gelistet wird.

Gegliedert ist die Arbeit in zehn Kapitel. Nach einer Einleitung, die die zentralen Themen der Arbeit benennt, wird zuerst ein Überblick über bibliographische Datenbanken, Portale und Suchmaschinen gegeben (Kap. 2). Es folgt ein Kapitel über Heterogenität und Information Retrieval, in welchem das Projekts KoMoHe¹, in dessen Rahmen die vorliegende Arbeit entstand, beschrieben wird. Im vierten Kapitel wird *Bradford's Law of Scattering*, welches die Grundlage für das Bradfordizing-Verfahren bildet, vorgestellt. Das fünfte Kapitel („Material und Methode“) beschreibt die Evaluationsverfahren, welche zum Test der Effektivität der beiden zentralen Verfahren des Projekts angewendet werden. Im sechsten Kapitel werden die Effekte der Crosskonkordanzen und des Bradfordizings beschrieben, während in Kapitel 7 dann die Ergebnisse der Retrievaltests vorgestellt werden. Im achten Kapitel werden die Ergebnisse diskutiert, Kapitel 9 bietet einen kurzen Ausblick und Kapitel 10 fasst die gewonnenen Erkenntnisse in Form von Thesen zusammen.

Beim Lesen des Buchs fällt auf, dass die beiden Themenbereiche stark getrennt voneinander behandelt werden. Darauf wird zwar schon in der Einleitung hingewiesen, es erschließt sich jedoch beim Lesen nicht, warum dies so sein muss. Einerseits wird ein Verfahren beschrieben, das die Treffermenge vergrößert, andererseits eines, das die Ergebnismenge zwar nicht verkleinert, jedoch (unter der vom Autor beschriebenen Annahme, dass ein Nutzer nur das erste Drittel der Ergebnismenge sichtet) einschränkt. Hier wäre es interessant gewesen zu erfah-

¹ <http://www.gesis.org/en/research/external-funding-projects/projektuebersicht-drittmittel/komohe/>

Gesamtlösungen für Digitalisierungsprojekte

ImageWare

MyBib eDoc® Workflow-System

BCS-2® Scansoftware

C-3 Periodikaerschließung

MyBib eRoom Elektronischer Lesesaal

SCANROBOT® Bookeye® Buchscanner

ren, inwieweit die Kombination beider Verfahren die Qualität der Treffermenge verbessern kann. Die vorgestellten Evaluierungen gehen darauf jedoch nicht ein, sondern beschränken sich auf die Effektivität der einzelnen Verfahren.

Ein großes Verdienst des Autors ist es, das Verfahren des Bradfordizing als Methode des Re-Rankings eingeführt zu haben. Grundsätzlich ist das Verfahren vielversprechend und sollte weiter untersucht werden. Der praktische Nutzen erscheint sehr hoch, allen Anbietern von Fachinformationsdiensten sei daher eine vertiefende Auseinandersetzung mit dem Verfahren nahe gelegt. Das Bradfordizing ist eine Möglichkeit, die Trefferlisten in Fachdatenbanken sinnvoll anzuordnen. Da die Nutzer solche gerankten Listen in Informationssystemen (zurecht) erwarten und die Standardverfahren sich auf bibliographische Daten nur sehr eingeschränkt anwenden lassen, sollten in der Praxis neue Verfahren implementiert und systematisch anhand des konkreten Anwendungsfalls implementiert werden. Hier sind die Anbieter von Fachdatenbanken aufgerufen, ihre Systeme entsprechend zu verbessern.

Inhaltlich ist Mayrs Buch damit auf jeden Fall ein Gewinn. Auch die Aufbereitung der Inhalte ist verständlich. Zu kritisieren ist allerdings die Zusammenstellung des Buchs: Es handelt sich hier nicht um einen Text „aus einem Guss“, vielmehr liest sich das Buch wie ein Sammlung aneinander gereihter Aufsätze aus unterschiedlichen Phasen des Projekts. Hierzu drei Beispiele: Der Überblick über die Datenbanken, Portale und Suchmaschinen wirkt veraltet und ist sicher einer frühen Phase des Projekts KoMoHe zuzuordnen (S.23ff.); die Kurzbeschreibungen der verwendeten Thesauri (S.72f.) stehen mitten im deutschen Text in englischer Sprache. Das neunte Kapitel („Ausblick“) liest sich wie ein Ausschnitt aus einem Projektantrag für ein Folgeprojekt. Eine solche Zusammenstellung trübt leider die verdienstvolle wissenschaftliche Arbeit des Autors.

Hingewiesen werden soll an dieser Stelle auch auf den inzwischen implementierten Prototypen², in dem neben den beiden beschriebenen Verfahren auch weitere im Ranking verschiedener sozialwissenschaftlicher Datenbanken ausprobiert werden können.

Dirk Lewandowski, Hamburg

Die Form der Unruhe

Tina Piazza, Stefan M. Seydel, rebell.tv AG
Bd. 1., Das Statement: die aktuelle Metamorphose der sozialen Frage: vom Buchdruck zum Computer; Collagierung eines Suchprozesses zum Umgang mit Informationen. Hamburg: Junius, 2009. 512 S.: überw. Ill.; 28 cm. ISBN 978-3-88506-456-5 kart.; 29,90 Euro, 48,90 sfr (freier Pr.)

Bd. 2., Die Praxis: vom Buchdruck zum Computer; Handlungsprinzipien zum Umgang mit Informationen auf der Höhe der Zeit. Hamburg: Junius, 2010. 192 S.: graph. Darst.; 15 cm. ISBN 978-3-88506-474-9 kart.; 7,90 Euro, ca. 12,50 sfr (freier Pr.)



Das erste, was auffällt, ja stört, ist die unruhige Form des Buches – der Bücher. Ein großer gewichtiger Band und ein kleines Heft, fast im Reclam-Format. Schließlich der Titel: erst die Untertitel der beiden Teilbände verraten, worum es geht: um „den Umgang mit Informationen auf der Höhe der Zeit“. Die zwei Autoren kommen eher aus einem wenig informationswissenschaftlichen Bereich, sie sind und verstehen sich als Sozialarbeiter. Sie arbeiten also am „sozialen Alltag“ und haben dies von 2004 bis 2009 recht erfolgreich mit Hilfe neuer Medien in ihrem Startup rebell.tv getan. Der erste Band „Das Statement“ dokumentiert die Arbeit am Zeitgeist per Videointerviews, die stets versuchten, den blinden Fleck der Fokussierung auf den offiziellen Diskurs zu meiden. Der Band ist geprägt vom Abdruck von 2375 Videostills aus den rebell.tv-Produktionen, die bis 2010 im Web zu sehen waren. Die Videos von rebell.tv sind in der dem Namen entsprechenden Filmsprache und Interviewtechnik am ehesten zu beschreiben als eine Mischung aus Georg Stefan Troller und Mario Sixtus. Sie spielen z.B. am Rande von Veranstaltungen wie re:publica oder Ars Electronica und erzählen von Begegnungen und Gesprächen mit („besonderen“) Menschen und in besonderen Situationen, Orten und Diskursen in Politik, Ökonomie, Kultur und Gesellschaft, speziell der *nächsten* Gesellschaft, die Dirk Baecker die Computergesellschaft

nennt. Prägend ist dabei das neue Medium Web TV, das die Möglichkeit eines schnellen, alternativen Journalismus bietet, den die Macher von rebell.tv, die sich eher als Initiatoren von Gesprächen verstehen denn als Journalisten, perfekt ausloten.

Die Videoproduktionen wurden begleitet und eingebettet von Magazinbeiträgen, „Statements“ und Podcasts, die sich aus der Arbeit an rebell.tv ergaben. Insgesamt 93 Themenfelder, wie sie Stefan Seydel nennt, wurden von Hanspeter Spörri essayistisch kommentiert und mit Verweisen (Hyper-Links) auf weitergehendes Material und die entsprechenden Videos versehen, und Tina Piazza entwickelt dazu jeweils ein an Schultz von Thun angelehntes „Werte und Entwicklungsquadrat“ etwa zum Thema „Bildersturm“ das Quadrat: (+) „Analoge Kommunikation“ / (-) „Prekäre Mehrdeutigkeit“ vs. (+) „Digitale Kommunikation“ / (-) „Dekontextualisierte Information“. Oder zum Thema „Philosophische Gleichzeitigkeit“ die viergliedrige Gegenüberstellung (+) „Mündlichkeit“ / (-) „Vergessen“ vs. (+) „Schriftlichkeit“ / (-) „Kontrollieren“.

Zur Einbettung der so in Bewegtbild, Stills, Grafik und Texten dokumentierten gesellschaftlichen Praxis von rebell.tv werden dem Band 1 eine Reihe von einleitenden Originalbeiträgen vorangestellt: neben einer Erläuterung der eigenen Methode von Tina Piazza und Stefan M. Seydel sind hier knappe aber prägnante Texte über rebell.tv von Franz Schultheis (Soziale Frage der Gegenwartsgesellschaft), Andrea Back (Publizieren und Kommunizieren in innovativer Genese), Dirk Baecker (Herausforderung der nächsten Gesellschaft), Heiko Kleve (Sozialarbeit der nächsten Gesellschaft) und Bazon Brock („Gib mal laut“) abgedruckt. Alles gewichtige Stimmen im aktuellen Diskurs. Die drei Teile bilden „den symbolischen Band 1“, ein „dickes, schweres Buch“ als „dem Königsweg der Dokumentation“, bereit, in Bibliotheken verwaltet zu werden, wo es neben anderen „eine große, unheimliche, lodernde ‚Unruhe‘“ darstellt. Es stellt den Versuch dar, die Tradition der Sozialen Arbeit mit der „Arbeit im Umgang mit Information“ zu kombinieren. Leitfragen dazu sind etwa: welche Bedeutung Information und Wissen für den Menschen haben oder was passiert, wenn sich das dominante Kommunikationsmedium ändert und die Informationen nunmehr statt zwischen Menschen durch vernetzte Computer vermittelt werden. Auch und gerade weil die Aussage nicht einfach sein kann, bietet das „Schaulager“ wie die Autoren es nennen, in der Konfrontation der beiden Perspektiven der Sozi-

2 <http://multiweb.gesis.org/irsa/IRMPPrototype>