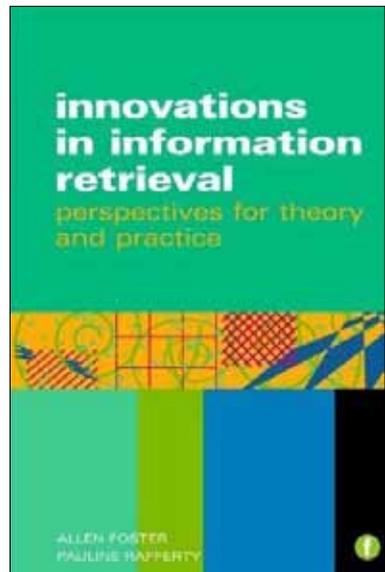


■ **Foster, A.; Rafferty, P. (Eds.): Innovations in Information Retrieval: Perspectives for Theory and Practice. London: Facet 2011. ISBN 978-1-85604-697-8. £ 44.95 / EUR 54.99**

Dieser mit 156 Seiten (inklusive Register) relativ schmale Band enthält sieben mit dem Gütesiegel „peer-reviewed“ versehene Beiträge namhafter Autoren zu „research fronts“ auf dem Gebiet des Information Retrieval (IR) – ein Begriff, der hier durchaus breit verstanden wird. Wie die Herausgeber Allen Foster und Pauline Rafferty – beide aus dem Department of Information Studies an der Aberystwyth University (Wales) – in ihrer Einleitung (p. xv–xviii) betonen, sind Theorie und Praxis der Wissensorganisation im Internet-Zeitalter nicht mehr nur die Domäne von Informationswissenschaftlern und Bibliotheksfachleuten, sondern auch von Informatikern, Semantic-Web-Entwicklern und Wissensmanagern aus den verschiedensten Institutionen; neben das wissenschaftliche Interesse am Objektbereich ist nun auch das kommerzielle getreten. Die Verarbeitung von Massendaten, die Beschäftigung mit komplexen Medien und die Erforschung der Möglichkeiten zur Einbeziehung der Rezipienten sind insbesondere die Aspekte, um die es heute geht. Im folgenden soll auf die drei Beiträge des Bandes, die mich thematisch besonders ansprechen, näher eingegangen werden.



David Bawden (Professor an der City University, London) betitelt seinen Beitrag „Encountering on the road to Serendip? Browsing in new information environments“ (p.1–22). Ihm geht es um die Frage, inwieweit sich die aus der Forschungstradition zu Informationssuche und Informationsverhalten bekannten Konzepte „browsing“, „serendipity“ (zufälliges Auffinden von Informationen), „information encountering“ und „literature discovery“ im Web-Kontext verändert haben und ob sie tatsächlich noch brauchbare Ansätze darstellen. Zu diesem Zweck wird im Beitrag die aktuelle Literatur seit 1990 abgehandelt. In früheren Zeiten, so Bawden, hatte Browsing (am Regal) keinen guten Ruf – es galt der Bibliothek als Indikator dafür, daß entweder ihre Kataloge schlecht oder die Benutzer unzureichend geschult waren. Heute sieht man das ganz anders: die Unmenge an Webdokumenten hat Browsing de facto zur vorherrschenden Methode der Informationssuche werden lassen. Dies diskutiert der Autor im ersten Abschnitt seines Artikels. Man unterscheidet verschiedene Formen bzw. Kategorien des Browsing (z.B. Marcia Bates in einem 4-Felder-Schema nach den Dimensionen „active/passive“ und „directed/undirected“, wobei die Kombination „active+directed“ dem konventionellen IR entspricht, die Kombination „active+undirected“ dem eigentlichen Browsing). Browsing ist auch eine wichtige Komponente vieler Modelle des Information Seeking Behaviour (z.B. von Wilson, Ellis, Foster, McKenzie, Chu, Makri & Warwick). In der Literatur wurden des weiteren Verhaltensmuster („patterns“) beim Browsing untersucht. Diese sind zwar im Prinzip auch im Print-Umfeld möglich, werden aber durch die digitalen Informationssysteme mehr unterstützt und durch diverse Features verbessert (Ergebnis-Ranking, Ergebnis-Clustering, Facettierte Navigation, automatisches Kategorisieren, Hyperlinks usw.). Studien des Suchverhaltens am Web haben aber auch neuartige Formen des Browsings enthüllt (etwa „bouncing and flicking“ – die sehr rasche Bewegung von Website zu Website, ohne genauere Überprüfung des Inhalts; weitere Schlagwörter in diesem Zusammenhang: „fast surfing“, „broad scanning“).

Im Abschnitt Encountering and serendipity behandelt Bawden das zufällige Auffinden von Information. Bekannt wurde in diesem Konnex der Begriff „information encountering“ (Sandra Erdelez), definiert als „serendipitous information acquisition“. Erdelez hat auch hierfür verschiedene Verhaltenskategorien erstellt und ein Indiz für das überraschenderweise nach wie vor eher negative Image des Browsings gefunden – sog. „super-encounters“ (also Personen, deren Suchverhalten ganz besonders darauf basierte), wollten nicht darüber sprechen, da sie selbst der Meinung waren, es handle sich nicht um eine „anerkannte“ Vorgangsweise. Interessan-

terweise präferierte diese Gruppe sogar gedruckte Quellen gegenüber der Web-Recherche. Eine ganze Reihe weiterer Untersuchungen in verschiedenen Kontexten ergänzt den Forschungskorpus zu „incidental information acquisition“ bzw. „finding without seeking“. Abzugrenzen davon ist hingegen „information discovery“, ein Begriff für die Verwendung von Prozeduren bzw. Algorithmen zur Auffindung von Analogien und Beziehungen in Informationssammlungen (v.a. im medizinischen Bereich).

Der dritte Abschnitt behandelt Persönlichkeitsmerkmale und individuelle Unterschiede im Zusammenhang mit Lernstilen und kognitiven Stilen. So fand man etwa eine besondere Neigung zu Serendipität bei Personen mit hoher Motivation, energischer Persönlichkeit und positiver Emotionalität, wogegen Stressanfälligkeit, Unsicherheit und geringe Motivation gegenteilig wirkten. Ein fünfstufiges Kategorienschema der „information attitude“ (Heinström) beinhaltet in drei dieser Kategorien eine Browsing-Präferenz. Für diese Studien wurden zwar durchaus anerkannte Persönlichkeitsmodelle und Lernstil-Konzeptionen herangezogen, doch blieben die meisten dieser Untersuchungen bislang eher „indikativ“ als „konklusiv“.

Abschließend widmet sich der Beitrag der Frage nach den Features von Informationssystemen, die das zufällige Auffinden von Informationen besonders unterstützen könnten. Dazu gibt es bislang nur einige wenige Studien; die bisherigen Aussagen und Vorstellungen klingen eher vage. Flexibilität und Einstellbarkeit hinsichtlich persönlicher Präferenzen und Vorkenntnisse werden hier genannt, etwa in Bezug auf Relevanzkriterien bzw. Genauigkeit/Unschärfe der Information. Daß in diesem Zusammenhang die bereits 1979 von Marcia Bates formulierten 17 Arten von „idea tactics“, deren einige in solche Systeme einzubauen wären, Erwähnung finden, ist angesichts der Qualität dieser Autorin wohl kein Zufall. Bawden selbst meint, derartige Features müßten folgendes zeigen können: (a) Beispiele dafür, was das Informationssystem beinhaltet, (b) die Struktur dessen, was das System ausmacht, und (c) was das System beinhaltet – bezogen auf das, was dem User bereits bekannt ist. Geeignete Suchfunktionen müßten etwa auf Ähnlichkeits- bzw. Unähnlichkeitsurteilen basieren. In Summe hat der Autor hier einen interessanten Beitrag vorgelegt, der gut lesbar ist und einen sehr brauchbaren Einstieg in die Thematik vermittelt.

Der zweite Artikel, „Classification revisited: A web of knowledge“ (p. 23–48), stammt von Aida Slavic (Chefredakteurin des UDC Konsortiums, Den Haag). Das sich [angeblich] entwickelnde Semantic Web wird, so Slavic, zum Teil durch Rückgriff auf die Semantik aus existierenden Wissensorganisationssystemen und Metadaten aufgebaut. Der Beitrag bezieht sich auf die Rolle, die einige weitverbreitete Universalklassifikationen, mit denen be-

reits sehr große Mengen bibliographischer Daten erschlossen worden sind, als sprachunabhängige Begriffsschemata zur Unterstützung von Menschen und Computer beim Strukturieren und Präsentieren von Informationen sowie beim Formulieren von Suchanfragen spielen können.

In einem relativ umfänglichen Einführungskapitel behandelt Slavic zunächst – in kompetenter und auch ansprechender Weise – die Grundlagen von Klassifikation und Klassifikationssystemen. Auf die Details (DDC, UDC, LCC, Hierarchien, Facetten usw.) braucht hier nicht näher eingegangen werden, denn das hier Präzentierte sollte eigentlich als Vorwissen für die Auseinandersetzung mit dem wirklichen Thema vorausgesetzt werden können.

Im zweiten Teil geht es dann um Klassifikationssysteme beim IR. Jedes Informationssystem mit klassifikatorischen Daten sollte nach Slavic eine systematische Ergebnisanzeige mit folgenden Möglichkeiten aufweisen: (a) semantische Sucherweiterung, und (b) Unterstützung von Disambiguierung bzw. Kontextualisierung von Suchbegriffen in entsprechenden Hierarchien. OPACs, so meint die Autorin, seien durch „very poor use of classification“ charakterisiert. Dies mag in vielen Fällen stimmen, kann aber dennoch nicht widerspruchlos hingenommen werden, wenn man sich selbst längere Zeit mit dem Einbau von klassifikatorischen Komponenten in Online-Kataloge beschäftigt hat. In einem Rückblick auf die Zeitspanne 1990–2010 behandeln drei Unterabschnitte folgende Anwendungsgebiete:

- *Fachliche Portale (subject gateways, hubs, portals)*: Während diese Anwendungen ursprünglich große Systeme wie DDC oder UDC bevorzugten, führte die fachliche Spezialisierung oft zur Aufgabe dieser Schemata, da für das Auffinden speziellerer Inhalte zu viele Hierarchiestufen benötigt wurden.
- *Automatisches Klassifizieren von Web-Ressourcen*: Leider werden hier nur das schon seit Jahren von der Bildfläche verschwundene deutsche System GERHARD (das die mehrsprachige UDC-Version der ETHZ anwandte) sowie das angestaubte System SCORPION von OCLC (das überdies nur registrierten Kunden zur Verfügung steht) näher erwähnt. Zumindest deutet die Autorin an, daß im kommerziellen Bereich die automatische facetiierte Kategorisierung (etwa von Produkten) an Raum gewonnen hat.
- *Klassifikation als Drehpunkt- (pivot) und Terminologiedienst*: Heute mehr als je geeignet sind Klassifikationssysteme als das zentrale Mapping-Rückgrat für die Verknüpfung von Vokabularien in verschiedenen Sprachen. In diesem Zusammenhang werden Projekte wie CORC (OCLC), die tschechische DDC-LCSH-Konkordanz, MSAC (Schlag-

wörter europäischer Nationalbibliotheken), HILT (UK), sowie European Library bzw. Europeana, angeführt.

Im abschließenden Abschnitt verweist die Autorin auf die Notwendigkeit maschinlesbarer und standardisierter Formate. Wichtig ist hier sicherlich die Feststellung, daß etwa MARC ausserhalb des Bibliothekswesens ohne viel Wert sei. In diesem Konnex spielen nicht nur nationale und internationale Normen eine Rolle, sondern auch Ansätze wie XML/RDF und Linked Data. Hinsichtlich der Nützlichkeit von Klassifikationssystemen für das Browsing in Informationssystemen gibt es, so die Autorin etwas resignativ, nur wenige Studien.

Der Beitrag ist zwar flüssig geschrieben und gut lesbar, m.E. aber dennoch eher enttäuschend. Zuviel Platz wird mit den einführenden Grundlagen verbraucht, die Darstellung der eigentlichen Thematik bleibt kursschissig und ist wohl auch nicht immer auf dem aktuellsten Stand.

Der dritte hier zu referierende Artikel behandelt „Digital information interaction as semantic navigation“ (p.117–134) und stammt von drei kanadischen Hochschullehrern (Richard Kopak, Luanne Freund, Heather L. O’Brien, alle von der University of British Columbia). Konkret geht es darum, wie sich Informationssuchende direkt mit Informationsobjekten auseinandersetzen bzw. diese manipulieren, wenn sie in digitalen Informationsräumen suchen und browsen. Dabei gilt den Autoren das von Ingwersen & Järvelin (2005) propagierte „retrieval in context“ als Perspektive, die das gesamte System, die affektiven, kognitiven und physischen Merkmale der Benutzer sowie das Umfeld, in der die Suche stattfindet, in Betracht zieht. Das Konzept der „semantischen Navigation“ soll dazu beitragen, IR-Systeme so zu gestalten, daß diese eine tiefere Interaktion mit einer Vielzahl von Informationsobjekten sowie deren Verwendung unterstützen können.

Vor etwa 10–15 Jahren wurde in den Arbeiten von Dillon und Vaughan impliziert, daß die navigatorischen Probleme der Benutzer nur lösbar wären, wenn die in der Information enthaltenen semantischen „cues“, die bei der Navigation helfen können, vom System ausreichend berücksichtigt würden. Der Begriff „semantic navigation“ meint, daß dies nur möglich ist, wenn bei den navigatorischen Optionen die Repräsentation der Bedeutung in bestimmter Weise sichtbar wird. Als die beiden Kontextfaktoren, die hier ganz besonders einflussreich sind, wurden „domain“ (Wissensgebiet) und „task“ (Aktivitäten zur Erreichung von Zielen) identifiziert.

Um dies in Informationssystemen zu bewirken, sehen die Autoren drei Ansatzpunkte: (1) die Benutzer durch „labelling“, „highlighting“ und

„mapping“ auf die in den Informationsobjekten enthaltenen Bedeutungen hinweisen; (2) die Benutzer mit Werkzeugen ausstatten, mit denen sie den Informationsraum examinieren und ihr eigenes Verständnis davon dokumentieren können, während sie darin navigieren (um so aus passiv Lernenden aktiv Lernende zu machen); (3) das Design der Informationsräume sollte die gesamte Benutzererfahrung weiter ausformen (durch interessante, relevante, ansprechende Inhalte und intuitive Technik). Vier spezifische Forschungsbereiche, die zur semantischen Navigation beitragen können, werden dargestellt:

- *Dokumententyp* (document genre): wurde bisher wenig berücksichtigt, findet nunmehr bei der facettierten Navigation zusehends Verwendung, das „genre“ symbolisiert, wann, warum und durch wen das Dokument erzeugt wurde und gibt eventuell Hinweise auf Struktur bzw. Inhaltsaufbereitung oder die Art der enthaltenen Botschaft. „Genre classification“ gilt als wichtige Herausforderung an die Forschung.
- *Annotation*: gemeint sind die Markierungen (Worte, Notizen) eines Nutzers zum Inhalt, die beim Lesen eines Dokuments gemacht werden. Sie dienen stets der „Signalisierung“ für die spätere Beachtung bzw. Interpretation. Im digitalen Umfeld können Benutzer so zusätzliche Information hinzufügen, die dem späteren Wiederauffinden zu Gute kommen kann – ihnen persönlich, aber eventuell auch anderen Nutzern. Zur zunehmenden Zahl von Annotationswerkzeugen in digitalen Umfeldern gibt es bisher aber noch wenig Forschung.
- *Linking*: nach Ansicht der Autoren lohnt es sich, Wege zu erforschen, um die Fähigkeit der (vordergründig) neutralen Links, auch Bedeutung zu tragen, evidenter zu machen. Dazu sind Tools nötig, die aus Lesern „Autoren“ machen, die während des Lesens Links generieren können. Links könnten darüber informieren, welche Beziehung zwischen Informationsobjekten besteht, quasi als Vorschau über das zu Erwartende. Sie könnten durch die Nutzer je nach ihrem Zweck benannt werden (z.B. „definiert“, „illustriert“, „fasst zusammen“). Techniken wie „mouseover“, „tooltips“ usw. könnten dies dann für spätere Leser visualisieren. Auch in diesem Bereich ist der Forschungsstand noch rudimentär.
- *Benutzerbeteiligung* (user engagement): Als Schlüssel zu der heute als besonders wichtig angesehenen „user experience“ (UX) gilt „engagement“, d.h. „the quality of being drawn in, captivated by and actively participating in an interaction“ (p. 128). Systeme mit dem

Prädikat „engaging“ haben folgende Gemeinsamkeiten: (1) sie sind fähig, das Interesse der Nutzer zu wecken und zu erhalten (durch Feedback und Kontrolle über die Interaktion), (2) sie stimulieren Exploration und Freude (durch interaktive und multimediale Features). In der semantischen Navigation haben sie das Potenzial, die Erfahrung der Benutzer zu bereichern, indem Informationsobjekte direkt und intuitiv manipulierbar gemacht und die Nutzer bei ihren Zielen/Aufgaben unterstützt werden, Dafür wäre es auch nötig, in Ergebnislisten und Dokumenten „cues“ zu plazieren, die die User in die Lage versetzen, das „Herausragende“ zu erkennen und tiefer in den Text „einzutauchen“. Kommerzielle Suchmaschinen gehen heute in diese Richtung, etwa durch die Wahlmöglichkeit zwischen Kategorien von Ergebnissen (Formen, Genres) oder Personalisierung der Benutzerschnittstelle. Im Hinblick auf die Forschung nennen die Autoren longitudinale Studien (wie verändert sich die Benutzererfahrung über die Zeit hinweg), Messprobleme und neue Hardware als hauptsächliche Felder.

Kopak, Freund & O'Brien haben hier einen überaus inspirierenden Beitrag vorgelegt, dessen Detailreichtum im Vorstehenden vielleicht nur grob wiedergegeben werden konnte. Ihre Conclusio lautet, daß es nicht nur darum gehen kann, daß Benutzer die Inhalte verstehen, sondern daß sie auch erkennen, was sie mit diesen Inhalten anfangen können.

Die übrigen vier Artikel dieses Readers befassen sich mit Belletristik-IR (Anat Vernitski & Pauline Rafferty), Musik-IR (Charlie Inskip), Social Tagging (Isabella Peters) sowie mit der webometrischen Evaluierung von Suchmaschinen (Mike Thelwall). Auch sie versprechen eine interessante Lektüre, sodaß der Band in informationswissenschaftlichen Sammlungen nicht fehlen sollte. Alle Aufsätze weisen übrigens recht ausführliche Verzeichnisse mit weiterführender Literatur auf.

Das Buch enthält neben einem brauchbaren (Sach-)Register ein Abbildungs- und Tabellenverzeichnis, einen Abschnitt mit kurzen Biographien der Verfasser sowie ein Begleitwort der südafrikanischen Professorin Ina Fourie. Die Machart des broschierten Bandes ist einfach, aber zweckmäßig. Der Preis liegt hingegen jenseits dessen, was ein 160-Seiten-Buch vernünftigerweise kosten darf. Außer Bibliotheken werden sich daher wohl kaum Käufer finden.

Otto Oberhauser, Wien