

INFORMATION RETRIEVAL IN REGENSBURG

Mark Rittberger
Universität Konstanz

Information Retrieval '93 Von der Modellierung zur Anwendung

1. Fachtagung der Gesellschaft für Informatik (GI) in Regensburg vom 13.- 15. September 1993

Zwei Jahre nach ihrer Gründung veranstaltete die Fachgruppe Information Retrieval der GI in Zusammenarbeit mit der Linguistischen Informationswissenschaft der Universität Regensburg ihre erste Fachtagung im oberpfälzischen Regensburg.

Etwa 100 Teilnehmer, vornehmlich aus dem deutschsprachigen Raum, fanden sich im mittelalterlichen Regensburg zusammen, um 17 Vorträge zu verfolgen. Gleich im voraus sei gesagt, daß sowohl die 3 eingeladenen Referenten als auch die 14 begutachteten Vortragenden überaus interessante Entwicklungen aus dem Bereich des Information Retrieval und angrenzender Gebiete, wie etwa Datenbanken, Benutzerschnittstellen oder Hypertext vorstellten. Karen Sparck Jones von der Cambridge University (GB) eröffnete die Tagung mit ihrem Vortrag "What might be in a summary?". Nachdem sie einen ausführlichen Bericht über Methoden zur Unterstützung von automatischen Zusammenfassungen gegeben hatte, beschrieb sie ihr eigenes Projektvorhaben. Sie teilte die für die Generierung von Zusammenfassungen notwendige Information in die drei Typen linguistische Information, Information über das

Weltwissen und kommunikative Information ein. Diese werden zusammen mit zwei Repräsentationsformaten zur Textrepräsentation für die Generierung von Zusammenfassungen verwendet.

Im zweiten eingeladenen Vortrag sprach Nicholas J. Belkin von der Rutgers University (USA) über "Interaction with Texts: Information Retrieval as Information Seeking Behaviour". Ausgehend vom Standard Information Retrieval Modell zeigte er, daß beim Information Retrieval als Suchvorgang das gesamte Verhalten eines Benutzers bei der Informationssuche berücksichtigt werden muß. Daraus folgt notwendigerweise, daß ein Information Retrieval System den Benutzer in allen Stufen des Information Retrieval Prozesses unterstützen muß. Der zentrale Prozeß dieser Unterstützung ist die Interaktion mit dem Nutzer. Die Interaktion spielt aber auch gegenüber anderen Prozessen, wie etwa Repräsentation oder Vergleich eine wichtige Rolle.

Als letzten eingeladenen Gast begrüßten die Teilnehmer Edward A. Fox von dem Virginia Tech (USA). Sein Vortrag "From Information Retrieval to Networked Multimedia Information Access" bezog sich vornehmlich auf die Chancen, die neue Informationsdienste, insbesondere solche, die über das Internet angeboten werden, für die Forschung im Information Retrieval bieten. Insbesondere beklagte er den Mangel an Aktivität im Bereich Multimedia bzw. Netzwerke und Information Retrieval. Er

rief die Forschungsgruppen im Bereich Information Retrieval auf, sich mehr in diesen Gebieten zu engagieren, da dort die Notwendigkeit für praktisch einsetzbare Information Retrieval Verfahren besonders hoch sei.

Parallel zu den in fünf Sitzungen aufgeteilten Vorträgen wurden verschiedene der in den Vorträgen diskutierten Systeme vorgeführt.

A. Burghardt, N. Fuhr, K. Grossjohann, U. Pfeifer, H. Spielmann und O. Stach begannen die erste Sitzung mit "NOSFERATU - ein integriertes Datenbank- und Information-Retrieval-System basierend auf dem Datenstrom-Paradigma". Sie stellten ein System vor, welches vage Anfragen und unsichere Repräsentationen in Datensammlungen berücksichtigt. Mit einem probabilistischen NF2-Modell versuchen sie bei großen Datenmengen den Zugriff auf Fakten- und Textinformation mit gerichteten, azyklischen Graphen zu verbessern.

H.-P. Frei und Y. Qiu berichteten über ein Retrieval-Experiment mit der Online-Datenbank INSPEC ("Effectiveness of Weighted searching in an Operational IR Environment"), in dem boole'sche und gewichtete Anfragen auf ihre Genauigkeit und Brauchbarkeit hin verglichen wurden.

Die zweite Sitzung wurde von J. Gu, U. Thiel und J. Zhao mit "Efficient Retrieval of Complex Objects: Query Processing in a hybrid DB and IR System" eröffnet. Dabei wurde die Verbindung von relationalen Datenbanksystemen und auf invertierten Dateien beruhenden Information Retrieval Systemen untersucht, um komplexe Objekte leichter suchbar zu machen.

Im Anschluß daran erklärte M. Hemmje "eine inhaltsorientierte, intuitive 3D Benutzerschnittstelle für Information-Retrieval-Systeme", bei der der Benutzer seine Suchterme auf der Oberfläche einer Kugel entsprechend ihrer Gewichtung zueinander bewegt. Abhängig von der Bewegung werden dann die Antworten als räumlich den Suchtermen am nächsten liegende Objekte

präsentiert.

R. Kuhlen und M. Hess beschreiben das "Passagen-Retrieval- auch eine Möglichkeit der automatischen Verknüpfung in Hypertexten". Die Autoren verglichen verschiedene Ähnlichkeitsmaße in Bezug auf ihre Effektivität bei der Verknüpfung von Textpassagen innerhalb eines aus dem Buchformat in Hypertextformat transformierten Buches.

Am zweiten Tag begannen M. Schmidt und U. Pfeifer die Vorträge zur dritten Sitzung mit einer Untersuchung zur "Eignung von Signaturbäumen für Best-Match Anfragen". Dabei zeigte sich, daß ein höhenbalancierter Signaturbaum als Zugriffspfad zur Beantwortung von Best-Match-Anfragen bei größeren Kollektionen nicht geeignet erscheint.

Im Beitrag von S. Meienberg ("Relevance Feedback by Relative Relevance Assessments") ging es um eine neue Relevance-Feedback-Methode zur Verbesserung der Frageformulierung. Deren Effektivität erwies sich im Vergleich zu anderen Verfahren als sehr hoch.

S. Roppel, C. Wolff und C. Womser-Hacker ("Intelligentes Faktenretrieval am Beispiel der Werkstoffinformation") stellten die im Projekt WING-IIR entwickelten Retrievalkomponenten für das graphische Retrieval von numerischer Information vor, welches Antworten über vages Wissen zu Materialeigenschaften präsentiert. Sie beschrieben ein Benutzermodell, mit dem die Benutzerschnittstelle in Abhängigkeit von den Benutzerinteressen vereinfacht wird.

Zu Beginn der vierten Sitzung knüpften J. Marx und M. Schudnagis an den vorangegangenen Vortrag an, indem sie die Entwicklung der Benutzerschnittstelle im Projekt WING-IIR erläuterten. Dabei wurden verschiedene Schnittstellenkonzepte in eine multimodale Datenbankoberfläche eingearbeitet.

A. Glöckner-Rist bezog sich in ihrem Vortrag "Suche und Du wirst finden: Die Formulierung von Suchproblemen und ihre

Transformation in Suchfragen " auf eine Versuchsanordnung, in der die Abhängigkeit der Suchfrageformulierung von den Retrievalkenntnissen der Benutzer bei Recherchen in Online-Datenbanken untersucht wurde. Die Effektivität wurde dabei sowohl durch die Inhalte der Problembeschreibung, als auch die Information Retrieval- und Problemkenntnisse der Rechercheure beeinflusst.

M. Herfurth, P. Mutschke und H.P. Ohly stellten "AKCESS: Konzept-orientiertes Retrieval mit bibliographischem Kontextwissen" vor. Sie nutzen wissenschaftsstrukturelle Information über Literatur, Forschungsprojekte und -institute sowie über Lehrveranstaltungen, um einschlägige Wissenschaftler in Bezug auf eine Fragestellung zu identifizieren.

Die Schlußsitzung wurde von R. Hammwöhner und M. Rittberger mit "KHS-ein offenes Hypertextsystem" eröffnet. Sie gaben zunächst einen Überblick über das Konstanzer-Hypertext-System (KHS) und die Aspekte der Offenheit, die das System auszeichnen. Anschließend beschrieben sie die Integration von e-mail und Online-Datenbanken in das KHS.

U. Kampffmeyer stellte "Multilinguale Dokumenten- Retrievalsysteme. Implementierung und Beispiele" vor. Er erläuterte die Notwendigkeit solcher multilingualer Retrievalsysteme, ehe er ihre Konzeption und die Arbeitsweise anhand eines Archivierungssystems für Faksimile-Dokumente und Dateien, einer Pressedokumentation in der Schweiz und einer Metadatenbank für die Vereinten Nationen beschrieb.

Ein weiterer Beitrag im Schnittfeld Hypertext und Information Retrieval von den Autoren M. Hofmann, E. Bartsch und S. Schmezko ("Klassendefinition und Anfrageformulierung in Informationssystemen durch graphische Interaktion") beschäftigte sich mit der graphischen Schnittstelle im Hypertextsystem CONCORDE, die zur Definition von Knoten- und Linktypen dient und die Abfrage von typisierten Strukturen

in einem Hypertext erlaubt.

Begleitet wurde die Tagung von einem Tutorial zum Thema Information Retrieval, welches von N. Fuhr, U. Pfeifer und G. Ruge angeboten wurde. Für Auflockerung sorgte ein kulturell und kulinarisch gelungener Abend, der mit einer Besichtigung der Handels- und Reichstagsstadt Regensburg begann und einem gemeinsamen bayerischen Abendessen in der Altstadt schloß.

Zusammenfassend läßt sich sagen, daß sowohl die Vielfalt und Breite der Themen, das große Interesse von Information Retrieval-Spezialisten aus Wissenschaft und Praxis, als auch die hervorragende Organisation zum Gelingen der Tagung beitrugen. Nach diesem Erfolg wird es daher auch eine zweite Information Retrieval Tagung im Mai 1995 in Konstanz geben.

Alle Beiträge sind im Tagungsband (Information Retrieval '93. Von der Anwendung zur Modellierung) herausgegeben von G. Knorz, J. Krause und C. Womser-Hacker im Universitätsverlag Konstanz erschienen.