

Harald Killenberg/Rainer Kuhlen/Hans-Jürgen Manecke (Hg.)

# Wissensbasierte Informationssysteme und Informationsmanagement

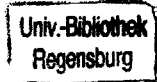
Proceedings des 2. Internationalen Symposiums  
für Informationswissenschaft (ISI '91)  
zusammen mit dem 17. Internationalen Kolloquium  
für Information und Dokumentation

Hochschulverband für Informationswissenschaft (HI) e. V. Konstanz;  
Institut für Informationswissenschaft der Technischen Hochschule  
Ilmenau

UNIVERSITÄTSVERLAG KONSTANZ

M116060

M/ST 271 K48 K9



46268987

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

**Wissensbasierte Informationssysteme und Informationsmanagement:** proceedings des 2. Internationalen Symposiums für Informationswissenschaft zusammen mit 17. Internationales Kolloquium für Information und Dokumentation / ISI '91, 4.11. - 7.11.1991. Harald Killenberg ... (Hg.). Hochschulverband für Informationswissenschaft (HI) e.V., Konstanz; Institut für Informationswissenschaft der Technischen Hochschule Ilmenau. - Konstanz : Univ.-Verl. Konstanz, 1991

(Schriften zur Informationswissenschaft ; Bd. 2)

ISBN 3-87940-412-7

NE: Killenberg, Harald (Hrsg.); Internationales Symposium für Informationswissenschaft <02, 1991, Ilmenau>; Internationales Kolloquium für Information und Dokumentation <17, 1991, Ilmenau>; GT

ISSN 0938-8710

ISBN 3-87940-412-7

© Universitätsverlag Konstanz GmbH, Konstanz  
Druck und Bindung: Siegl Druck GmbH, Friedrichshafen  
Einbandgestaltung: Riester & Sieber GmbH, Konstanz  
Papier: Chlorfrei gebleicht

## **Inhalt**

Vorwort 10

### **Plenarbeiträge**

Die EG-Politik im Bereich des Informationsmarktes  
Huber, W. 11

Zur Theorie informationeller Mehrwerte / Kuhlen, R. 26

Die vernetzte Weltgemeinschaft: Potentiale und Risiken  
Stroetmann, K.A. 477

### **Themenkreis 1:**

#### **Grundlagen, Verfahren und Systeme der Wissensverarbeitung**

#### **Theoretische Ansätze, Konzepte und Modelle**

Theorien epidemischer Prozesse in der Informations-  
wissenschaft / Wagner-Döbler, R. 40

Erfolgsfaktorenforschung in der Informationsvermittlung:  
Konzepte, Modelle, Ergebnisse / Herget, J. 52

#### **Repräsentation von Wissen**

Vom traditionellen Indexieren zum Wissenserwerb /  
Mater, E. 70

Zur Repräsentation von Programmierwissen / Dosch, W. 71

Beschreibung von Bürodokumenten durch kontextsensitive  
Merkmale / Holbein, R.; Teufel, B. 81

#### **Retrieval-, Hypertext- und wissensbasierte Informations- systeme**

Prototype for Integrating Probabilistic Fact and Text  
Retrieval / Fuhr, N.; Hoffmann, Th. 94

Navigationsstrategien in der Informationssuche - vom Information Retrieval zu Hypertext / Krüger, F.	104
Graphische Recherche und intelligente Hilfe als Komponenten eines multimodalen Werkstoffinformationssystems WING / Roppel, St.; Wolff, Ch.	115
Ein Verfahren zur automatischen Trunkierung beim Zugang textbezogenen Informationsbanken / Zimmermann, H.	125
Eine objektorientierte Sicht auf Expertensystem-Wissensbasen: auf dem Weg zu einer hypermedia-gestützten Dialogkomponente / Yetim, F.	145
Design-Vorschläge für Hypertext-Applikationen aufgrund von Feldversuchen / Schwarz, E. J.; Bertha, E.; Schmidt, S.	157
Zur Beziehung zwischen Entwicklungsaufwand und Leistungsfähigkeit eines Wörterbuches für die automatische Indexierung / Tzeras, K.	167
Hyper-Advokat - Konzeption und Realisierung eines juristischen Hypertextsystems / Michel, E.	182
1789 - Die Französische Revolution auf einem Hypermedia Informationssystem / Schröder, Thomas A.	197
Die Wissensrepräsentation in der Software Technologie auf der Basis des ereignisorientierten Ansatzes / Kulawiak, W.	207
Publikationsstrukturen im Ländervergleich / Bonitz, M.; Bruckner, E.; Scharnhorst, A.	218
<b>Informationswissenschaftliche Aus- und Fortbildung</b> Konzept der informationswissenschaftlichen Ausbildung in Ilmenau / Killenberg, H.	228

Informationswissenschaft in Saarbrücken - Stand und  
Perspektiven / Harms, I.; Schulz, A. 237

## **Themenkreis II:**

### **Entwicklungstendenzen des Informationsmarktes und Management von Informationen**

#### **Internationale, soziale, politische und ökonomische Aspekte**

Die Informationsgesellschaft: eine Utopie im Spiegel  
postmoderner Theorie? / Wawra, St. 247

Konsequenzen aus dem Paperwork Reduction Act /  
Weitzendorf, Th. 256

Wissenschaftliche Grundlagen einer Informationstech-  
nologie im Bereich des Managements / Anan'eva, T. A. 259

#### **Einsatzmöglichkeiten neuer Technologien**

Bereitstellung von Agrardatenbanken auf CD-ROM -  
ein Projekt für die neuen Bundesländer / Münch, E.;  
Thulke, M.; König, E. 274

Einsatz eines optischen Speichersystems in einer  
Wirtschaftspresdokumentation / Michelson, M. 286

Neue Informationstechniken am Richterarbeitsplatz  
auf der Ebene der Amtsgerichte / Scheuermann, F. 298

Information System for Educational Software: State-  
of-the-Art and Perspectives for its Development in  
Bulgaria / Tsankova, R.; Koleva, R. 305

WIDAB-Konzeption und Pilotrealisierung eines experimen-  
tellen Information-Retrieval-Systems / Englmeier, K. 307

AIR/PHYS - Automatische Indexierung in der Praxis /  
Schwantner, M. 320

## **Informationsorganisatorische Strukturen und Verfahren des Informationsmanagements**

- Moderne Technik und betriebliche Realität im Informationsmanagement am Beispiel eines Automobilkonzerns /  
Schael, F. 333
- Vom Fachinformationssystem Dieselmotoren zum  
Informationsmanagement / Hillebrand, V. 341
- Veränderungen des Aufgabenprofils einer Informationseinrichtung in einem Unternehmen des Schwermaschinenbaus /  
Kruske, H. 351
- Akzeptanzhürden spielerisch überwinden - "Informationsmanagement" in der berufsbegleitenden Weiterbildung.  
Ein Vermittlungskonzept dargestellt am Beispiel  
"Kosten und Leistungsrechnung für Informationsdienstleistungen" / Frerk, G.; Oßwald, A. 358
- ### **Forschungs- und Managementinformation**
- Das "Online-Unternehmensbild" anhand von Wirtschaftsdaten und informetrischen F&E-Indikatoren / Stock, W. G. 376
- Biotechnology Information Service and its Value for Research / Poetzsch, E. 387
- Komponenten von Informationsvermittlung und -consulting für die veränderten Informationsbedürfnisse von Forschung und Wirtschaft in den neuen Bundesländern /  
Kurby, St. 396
- Managementinformationen auf der Basis von statistischen Untersuchungen in Patentinformationsdiensten /  
Babst, P. 410

<b>Aspekte der Informationsökonomie und des europäischen Informationsmarktes</b>	
Elektronische Märkte -Ihre Entwicklung, Bedeutung und der Einsatz von Informations- und Kommunikationstechnologien / Himberger, A.; Zbornik, St.	433
Europäisches Informationsrecht: Neuere Entwicklungen / Goebel, J.W.; Burkert, H.	447
Der europäische Informationsmarkt: Entwicklungs- und informationspolitische Aspekte und Perspektiven / Schwuchow, W.; Stroetmann, K. A.	450
<b>Studentenbeiträge</b>	
Graphische Benutzeroberflächen: Methoden und Werkzeuge für Prototyping und Entwicklung / Jochum, B.	472
Ein Konzept für die Auftragsabwicklung in der Informationsvermittlung / Sprenger, A.	475
<b>Autorenverzeichnis</b>	494
<b>Stichwortverzeichnis</b>	496

## Vorwort

Das Institut für Informationswissenschaft der Technischen Hochschule Ilmenau und der Hochschulverband Informationswissenschaft (HI) e.V. Konstanz veranstalten gemeinsam das 17. Internationale Kolloquium für Information und Dokumentation in Verbindung mit dem 2. Internationalen Symposium für Informationswissenschaft (ISI'91) in Oberhof Thüringen. Mit dieser Tagung setzen wir bewahrenswerte Traditionen der Reihe der Kolloquien für Information und Dokumentation fort und verbinden sie mit den Intentionen der im Oktober 1990 mit ISI'90 in Konstanz ins Leben gerufenen Folge informationswissenschaftlicher Fachtagungen. Dabei sind die ISI-Tagungen von Anfang nicht auf Deutschland beschränkt; vielmehr sind - entsprechend den Zielsetzungen des HI - Österreich und die Schweiz daran beteiligt. Erweiterungen deuten sich mit der ISI'92 in Saarbrücken an. Mit der Vergabe der ISI-Konferenz nach Oberhof sollte deutlich gemacht werden, daß auch in der ostdeutschen Informationswissenschaft und -praxis die Selbständigkeit und der Anschluß an die internationalen Standards so schnell wie möglich erreicht werden müssen. Durch die gemeinsamen Anstrengungen von HI und Ilmenau sowie zahlreiche Solidaritätsbekundungen anderer Personen und Institutionen ist es heute realistisch, daß das Institut für Informationswissenschaft an der TH Ilmenau über eine Einbindung in die aufzubauende Wirtschaftsinformatik seine Arbeit wird dauerhaft fortsetzen können. Neue Studienangebote werden entwickelt. Das ist sicher eine Anerkennung der Bedeutung von professioneller Informationsverarbeitung für den Aufbau der Wirtschaft auch in den neuen Ländern.

Ziele dieser Tagung sind es, weitere Impulse zur theoretischen Formierung der Informationswissenschaft, auch unter dem Aspekt zukünftiger Ausbildung, zu geben und zugleich Mittler zu praktischen Anwendungen, insbesondere in der Unternehmensinformation, zu sein. Insofern besteht die Tagung aus zwei Hauptteilen: "Grundlagen, Verfahren und Systeme der Wissensverarbeitung" und "Entwicklungstendenzen des Informationsmarktes und des Management von Informationen". Der vorliegende Band zeigt deutlich, daß auch in diesem Jahr eine Vielzahl von Beiträgen mit gutem Niveau eingereicht worden sind, von denen die hier dokumentierten nach den Beratungen des Programmkomitees akzeptiert worden sind. Auch dieser Band wird hoffentlich in der Zukunft noch mehr Wissenschaftlerinnen und Wissenschaftler ermuntern, die ISI-Konferenzen als ihre Fachkonferenz anzunehmen, auf der sie ihre Forschungsergebnisse vor einem kompetenten Publikum diskutieren lassen können. Wir danken allen an der Auswahl und der Erarbeitung der Beiträge Beteiligten. Ebenso danken wir der GMD, daß sie auch dieses Jahr die Durchführung der ISI-Konferenz und damit auch diesen Proceedings-Band, der zur Konferenz vorliegt, finanziell unterstützt hat.

10/1991 H. Killenberg (Ilmenau) H.-J. Manecke (Ilmenau) R. Kuhlen (HI, Konstanz)



## Die EG-Politik im Bereich des Informationsmarktes

Wolfgang HUBER

### *I. Die Bedeutung der Information für die wirtschaftliche und soziale Entwicklung*

Information ist von strategischer Bedeutung für Wissenschaft und Technik, Industrie und Handel.

Die durch neue Technologien und Spezialisierung hervorgerufene fortschreitende wirtschaftliche Entwicklung, die das Aktionsfeld durch die Schaffung des europäischen Binnenmarktes von lokalen zu nationalen und internationalen Märkten verlagert hat, ruft einen immer steigenden Informationsbedarf hervor.

Information ist heute in der Wirtschaft allgegenwärtig, sei es, um nach besseren technischen Lösungen zu suchen, eine Produktionslinie optimal zu fahren, Absatzchancen zu ermitteln oder sei es, um Finanzmittel bestens zu plazieren. Von gut funktionierenden Informationssystemen hängt heute in hohem Masse die Wettbewerbsfähigkeit vieler Wirtschaftsbereiche ab.

Um die Flut von Informationen in den Griff zu bekommen, ist die Anwendung moderner, elektronischer Datenverarbeitungs- und Kommunikationstechniken nicht mehr wegzudenken.

Die Bereitstellung von Information in elektronischer Form entwickelt sich immer mehr zu einem eigenständigen Wirtschaftszweig. Der Umsatz der Online-Datenbanken und der Echtzeit-Informationsdienste weltweit dürfte 1990 schätzungsweise bei rund 10 Mrd. Ecu gelegen haben, wovon rund 3 Mrd. Ecu auf Europa entfielen. In Anbetracht einer jährlichen Wachstumsrate von mehr als 20% bei Online-Diensten und der Entwicklung des Marktes für neue Speicherungsformen maschinenlesbarer Information (z.B. CD-ROM) schätzen Beobachter, dass der Weltmarkt für elektronische Informationsprodukte und -dienste bis zum Jahr 2000 auf 100 Mrd. ECU wachsen wird.

## II. *Gegenwärtige Lage und voraussichtliche Entwicklung*

Die Position Europas auf dem Weltmarkt für elektronische Informationsdienste ist immer noch wesentlich schwächer, als sie sein sollte. So ist der Markt für Online-Informationendienste (ohne Videotex) nach Umsatz in Europa derzeit nur rund ein Drittel so gross wie in den USA. Die on-line über ASCII zugängliche Anzahl von unterschiedlichen Datenbanken in den USA ist mehr als doppelt so hoch wie in der Europäischen Gemeinschaft (1989: USA 2214; EG 1048). Auch produzierten in den letzten Jahren die Länder der Europäischen Gemeinschaft nur etwas mehr als halb so viele neue Datenbanken wie die USA. Dazu kommt, dass in den USA wesentlich mehr höherwertige (Fakten und Volltexte) und wesentlich umfangreichere Datenbanken entwickelt werden als in Europa. Ein weiterer Unterschied ist, dass in den USA die weitaus meisten Datenbanken von privatwirtschaftlichen Unternehmen entwickelt werden (83% im Jahre 1989), während sie in Europa mehrheitlich von Organisationen ohne Erwerbszweck erstellt werden (54% im Jahre 1989).

Der Informationsmarkt ist seiner Natur gemäss stark nach Sachbereichen, Benutzergruppen und Verbreitungsmedien segmentiert. So bestehen beispielsweise keine Verbindungen zwischen dem Markt für Echtzeit-Finanzinformationsdienste und dem Markt für Kreditauskunftsdienste, obwohl beide dem Sektor Bank- und Finanzwesen angehören.

Die Zersplitterung der Märkte infolge von sprachlichen, technischen und rechtlichen Barrieren ist ein schwerwiegender und allen elektronischen Informationsdiensten in Europa gemeinsamer Nachteil. Deshalb können die Informationsanbieter Kostendegressionsvorteile bei der Produktion und Vermarktung nicht hinreichend nutzen. Viele Marktsegmente erreichten nicht die kritische Schwelle, auf deren Grundlage ein gesundes Wachstum entwickelt und aufrechterhalten werden kann.

Es ist davon auszugehen, dass der durch den einheitlichen europäischen Binnenmarkt ermöglichte freie Verkehr von Waren, Dienstleistungen, Kapital und Personen die Nachfrage nach gemeinschaftsweiten Informationsdiensten ansteigen lassen wird.

Die europäischen Informationsanbieter werden jedoch Schwierigkeiten haben, diese Marktchancen voll zu nutzen. Und zwar aus folgenden Gründen:

Die Zahl der Teilnehmer am europäischen Informationsmarkt ist gross. Bis auf wenige grosse sind es kleinere Unternehmen, die oft mehr den nationalen Markt als den europäischen oder gar den Weltmarkt im Auge haben. Und wegen ihrer nationalen oder einsprachigen Ausrichtung fehlen ihnen die Mittel und Fähigkeiten, um auf einem wirklich gesamt-europäischen Markt tätig zu werden. Die wenigen grossen europäischen Anbieter, die die Mittel und Fähigkeiten besitzen, um transnational zu operieren, scheinen lieber in den USA zu investieren als in Europa. Demgegenüber sind amerikanische Unternehmen wegen der Chancen, die sie im einheitlichen europäischen Binnenmarkt sehen, jetzt stark daran interessiert, europäische Unternehmen aus allen Bereichen des Informations- und des Mediensektors zu übernehmen.

Mit der Liberalisierung der Telekommunikationsdienste, den Fortschritten in der optischen Speichertechnik und dem Erscheinen Japans als möglicher konkurrierender Anbieter moderner Informationsdienste stellen sich der europäischen Informationsindustrie weitere neue Herausforderungen.

Ein zusätzliches und schwerwiegendes Problem ist die regional unterschiedliche Versorgung mit und Nutzung von modernen Informationsdiensten in der Gemeinschaft. Dieses Ungleichgewicht kommt in dem Umstand zum Ausdruck, dass in einem Mitgliedstaat, genauer gesagt in Grossbritannien, ein Drittel aller Datenbanken in der Gemeinschaft hergestellt werden und auf dasselbe Land 3/4 aller Datenbankexporte aus der Gemeinschaft in die USA entfallen. In anderen Mitglied-

staaten dagegen werden nur wenige elektronische Informationsdienste angeboten, und daran hat sich in den letzten Jahren auch nicht sehr viel geändert. Diese unterschiedliche Struktur gefährdet aber das Zusammenwachsen der Länder der Europäischen Gemeinschaft und kann in einem einheitlichen Binnenmarkt zu Wettbewerbsverzerrungen führen.

### *III. Aktionen der Europäischen Gemeinschaft zur Förderung des Informationsmarktes*

Anfang der siebziger Jahre wurde die europäische Gemeinschaft erstmals im Bereich der wissenschaftlichen und technischen Information tätig.

Die ersten Aktionspläne für Information und Dokumentation über den Zeitraum 1975-83 zielten vorrangig auf die Entwicklung der grundlegenden Infrastruktur ab, die für den Zugang zu den in der Gemeinschaft vorhandenen Online-Datenbanken erforderlich war. Dieses Ziel wurde durch den Aufbau des Euronet-Diane Netzes erreicht, das seit 1985 durch den Verbund der nationalen Paketvermittlungsnetze ersetzt ist.

Später bemühte man sich im Rahmen eines fünfjährigen Aktionsprogramms für die Entwicklung des Fachinformationsmarktes (1984-1988) in erster Linie um die Verbesserung von Qualität und Vielfalt der in der Gemeinschaft erstellten Datenbanken und -Dienste.

Vor der Einführung des Euronet-Diane Netzes Ende der siebziger Jahre gestaltete sich der Umsatz des europäischen und des amerikanischen Online-Informationsmarktes im Verhältnis 1 zu 10. Dieser Abstand konnte immerhin auf 1 zu 3 reduziert werden.

Der immer noch beachtliche Grössenunterschied wird sich jedoch nur ganz allmählich weiter abbauen lassen.

Um einerseits der strategischen Bedeutung der Information und andererseits der möglichen raschen Entwicklung der Informationsindustrie als eigenem Wirtschaftssektor Rechnung zu tragen, wurde ein zweijähriger Aktionsplan (1989-1990) zur Schaffung eines Marktes für Informationsdienste aus der Taufe gehoben (\*). Dieses als IMPACT (Information Market Policy Actions) bekannte Programm war mit einem Budget von 36 Mio. ECU ( $\pm$  DM 72 Mio.) ausgestattet und hatte folgende Ziele:

- Schaffung eines Binnenmarktes für Informationsdienste bis Ende 1992;
- Anregung und Stärkung der Wettbewerbsfähigkeit europäischer Anbieter von Informationsdiensten;
- verstärkte Nutzung modernster Informationsdienste in der Gemeinschaft;
- Verstärkung der Bemühungen im Hinblick auf eine gemeinschaftsweit gleichmässige Entwicklung der Informationsdienste.

Im Interesse der Erreichung dieser Ziele wurden zwei sich ergänzende Ansätze gewählt:

- als "horizontale Massnahmen" ein ständiges Bemühen zur Verbesserung der Marktbedingungen und zur Förderung der Nutzung moderner Informationsdienste und
- als "vertikale Massnahmen" die Durchführung von Pilot- und Demonstrationsprojekten, die auf die Entwicklung des Marktes in Schlüsselbereichen einen katalytischen Effekt ausüben können.

#### Horizontale Massnahmen

Die im Rahmen von IMPACT angelaufenen horizontalen Massnahmen dienen dem Abbau von Hindernissen auf dem

---

(\*) Entscheidung des Rates 88/524/EWG, veröffentlicht im ABL. Nr. L 288 vom 21.10.1988

Markt und der verstärkten Nutzung von Informationsdiensten. Dazu gehören die Einsetzung eines Beobachtungsgremiums für den Informationsmarkt, die Untersuchung der rechtlichen und verwaltungstechnischen Hemmnisse und die Suche nach Möglichkeiten zu ihrer Überwindung, die Ausarbeitung von Leitlinien zur Verbesserung des Zusammenwirkens zwischen dem öffentlichen und dem privaten Sektor sowie Fördermassnahmen um Informationssysteme besser zu nutzen.

#### Beobachtungsgremium für den Informationsmarkt (IMO)

Das Beobachtungsgremium besteht aus einem Sekretariat bei der KEG in Luxemburg und einem Netz von Sachverständigen in den Mitgliedstaaten.

Seine Aufgabe ist, Daten über den Informationsmarkt zusammenzutragen und Analysen durchzuführen, um die Kommission bei ihren Entscheidungen zu unterstützen. Hauptuntersuchungsbereich ist der Markt für berufsbezogene Information, die online oder über maschinenlesbare Datenbanken verbreitet wird.

Das Beobachtungsgremium greift dabei voll auf die in den Mitgliedstaaten vorhandenen statistischen Daten zurück. Da diese nicht immer europaweit vergleichbar sind, ermutigt es Forschungseinrichtungen und Berufsverbände, die auf nationaler Ebene Daten erheben, zusammenzuarbeiten und die Definitionen und Methoden ihrer Erhebungen zu vereinheitlichen. Das Gremium trägt selbst zentral Daten zusammen und führt Studien durch, um die gravierendsten Lücken zu füllen und durch Festlegung des erforderlichen begrifflichen Rahmens zur Einbeziehung des Informationssektors in die amtlichen Statistiken beizutragen. Es erstellt zweimonatlich Syntheseberichte und jährlich für den Ministerrat und das Europäische Parlament einen Bericht über den Stand des Marktes. Diese Berichte werden den entsprechenden Berufsverbänden und Marktteilnehmern zugänglich gemacht.

## Rechtsbeirat für den "Informationsmarkt"

Ein Rechtsbeirat von hochrangigen Juristen, die auf Fragen des Informationsrechts in den Mitgliedstaaten spezialisiert sind, unterstützt die Kommission bei der Untersuchung der rechtlichen Hemmnisse und der Ausarbeitung von Vorschlägen zu ihrer Überwindung.

Sechs Themenkreise, die den Markt für Informationsdienste betreffen, sollen vorrangig geprüft und in Angriff genommen werden: geistiges Eigentum, Echtheit elektronischer Unterschriften, Computerbetrug, Haftung bei Informationsdiensten, Vertraulichkeit von Datenbankabfragen und Datenschutz.

Aufbauend auf den Empfehlungen des juristischen Beirates hat die Kommission der Europäischen Gemeinschaften am 18. Juli 1990 eine Richtlinie zum Schutz von Personen hinsichtlich der Verarbeitung personenbezogener Daten beschlossen, die zur Zeit dem Ministerrat zur Zustimmung vorliegt (KOM(90) 314 end. - SYN 287 und 288).

### Verbesserung des Zusammenwirkens zwischen öffentlichem und privatwirtschaftlichem Sektor

Dem öffentlichen Sektor kommt sowohl als bedeutendem Produzenten und Anbieter von Daten als auch als Benutzer von Informationsdiensten eine entscheidende Rolle für die Entwicklung des Informationsmarktes zu.

Um ein optimales Zusammenwirken zwischen den Massnahmen des öffentlichen und den Initiativen des privaten Sektors zu gewährleisten, hat die Kommission in Abstimmung mit den Beteiligten Leitlinien ausgearbeitet. (\*)

Diese Leitlinien, die keine rechtlich bindende Funktion haben, stellen eine Sammlung von allgemeinen Prinzipien

---

(\*) Leitlinien für ein besseres Zusammenwirken zwischen dem öffentlichen und dem privaten Sektor auf dem Informationsmarkt (GD XIII - Luxemburg)

und Empfehlungen dar, wie z.B. über die zur Verfügungstellung von öffentlichen Daten, einem fairen Wettbewerb und der finanziellen Unterstützung von Informationsdiensten.

#### Fördermassnahmen zur Nutzung von Informationsdiensten

Die Benutzer sehen sich einer Vielzahl von Datenbankdiensten und zugleich der Schwierigkeit ihres Zuganges gegenüber. Um die Benutzung europäischer Informationsdienste zu fördern, wurde von der Kommission der europäischen Gemeinschaften umfangreiches Informationsmaterial in allen Gemeinschaftssprachen ausgearbeitet.

Eine wichtige Rolle bei diesen Fördermassen spielt ECHO, der Host der Europäischen Kommission(\*\*). ECHO bietet ein vielsprachiges Help Desk zur Beantwortung von Benutzerfragen und mehrsprachige Verzeichnisse an, um den Benutzer durch den Informationsdschungel zu führen. Ferner bietet ECHO Informationsmaterial, eine Trainingsdiskette und Videofilme an, um den Interessenten in die Welt der elektronischen Information einzuführen.

ECHO erprobt auch innovative Systeme, um den Informationszugang zu erleichtern, wie z.B. Retrieval in natürlicher Sprache und Audiotex-Zugang zu Datenbanken mit Stimmeingabe und -ausgabe.

#### Vertikale Massnahmen

Um eine katalytische Wirkung auf die Schaffung einer neuen Generation benutzerfreundlicherer Informationsdienste auszuüben und die europaweite Zusammenarbeit zwischen den Marktteilnehmern zu fördern, hat die Kommission eine Reihe von Pilot- und Demonstrationsprojekten in Schlüsselbereichen gefördert.

---

(\*\*) ECHO Datenbasen und Dienste. Broschüre erhältlich bei: ECHO, B.P. 2373, L-1023 Luxemburg



Für diese Projekte wurden sechs vorrangige Bereiche ausgewählt: Touristik-Information, Patentinformation, Information über Strassengüterverkehr, Information über Normen, Bildbanken und intelligente Schnittstellen.

Auf die Ende 1988 und Anfang 1989 veröffentlichten Aufrufe zu Vorschlägen erfolgte eine starke Reaktion der Informationsindustrie. Es gingen 178 Vorschläge ein, an denen 730 Organisationen beteiligt waren. Mittel in Höhe von 160 Millionen ECU wurden als Beihilfen beantragt, denen jedoch nur verfügbare Haushaltsmittel in Höhe von 23 Millionen ECU gegenüberstanden. So konnte nur etwa jedes zehnte Vorhaben für eine Kofinanzierung ausgewählt werden.

Neunzehn Vorhaben sind angelaufen, an denen rund 100 Organisationen aus elf Mitgliedstaaten beteiligt sind. Die Gesamtkosten belaufen sich auf etwa 67 Millionen ECU, von denen rd. 30 % von der Kommission übernommen werden.

Die Projekte entfallen von der Anzahl her auf folgende Bereiche: Bildbanken: 5, Touristik-Information: 4, Patentinformation : 3, intelligente Schnittstellen: 3, Information über Normen: 2 und Information über Strassengüterverkehr: 2.

Darüber hinaus wurden zur Vorbereitung eines besonderen Aktionsplans für Bibliotheken aus IMPACT zwei Projekte speziell für Bibliotheken finanziert. Ferner wurden drei unter früheren Programmen angelaufene Aktivitäten unter IMPACT weitergeführt. Hier ging es um Videotext für die Landwirtschaft, Werkstoffdatenbanken in Europa und Tätigkeiten zur Schaffung einer effizienten Informationsinfrastruktur für Biotechnologie-Forschung und -Industrie.

#### IV. IMPACT 2

(Programmorschlag für die Jahre 1991-95 ) (\*)

Die bisher erzielten Ergebnisse zeigen, dass die Europäische Gemeinschaft für das Wachstum der europäischen Informationsindustrie und für die weitere Entwicklung des europäischen Informationsmarktes eine zentrale Rolle spielt. Angesichts der kurzen Laufzeit des bisherigen IMPACT-Programms und der recht begrenzten Mittel kamen einige der Projekte über die Anfangsphase nicht hinaus und müssen über die Laufzeit von IMPACT 1 hinaus weitergeführt werden, wenn sie voll zur Wirkung kommen sollen.

Die vorher genannten allgemeinen Ziele von IMPACT 1 bleiben für das neue Programm IMPACT 2 gültig, das eine Laufzeit von 5 Jahren (1991-1995) haben soll. Die gewonnenen Erfahrungen und die Veränderungen auf dem Informationsmarkt lassen jedoch eine gewisse Neuorientierung der Initiativen angeraten erscheinen.

Oberstes Ziel ist weiterhin die Integrierung des Marktes auf Gemeinschaftsebene. Dieses könnte innerhalb von 5 Jahren erreicht werden, wenn angemessene Budgetmittel bereitgestellt werden und die Aktionen der Europäischen Gemeinschaft durch nationale Massnahmen von den Mitgliedstaaten unterstützt werden.

##### **1. Schaffung eines europäischen Binnenmarktes für Informationsdienste**

Hinsichtlich der Schaffung eines europäischen Binnenmarktes für elektronische Informationsdienste wird die Verbesserung der Rahmenbedingungen für den Informationsmarkt das vorrangige Ziel sein. Nur durch Beseitigung der Barrieren, die den gemeinschaftlichen Informationsmarkt in kleine Einzelmärkte teilen, kann ein einheitlicher Markt entstehen, der der europäischen Informationsindustrie die Kostendegressionsvorteile

---

(\*) IMPACT 2 KOM (90) 570 end.

bietet, die sie benötigt, um auf dem Weltmarkt konkurrenzfähig zu sein.

Es ist daher beabsichtigt, die Tätigkeit des Beobachtungsgremiums für den Informationsmarkt fortzuführen und zu erweitern und die Verbreitung der Ergebnisse ihrer Untersuchungen zu verbessern. Vermehrte Anstrengungen sind zu richten auf die Gewinnung besserer Einblicke in Beschaffenheit, Struktur, Verhalten und Leistung von Informationsmarkt und Informationsindustrie sowie auf ergänzende Konzeptstudien und Prognosen, auf deren Grundlage vernünftige und wirksame Politiken und Massnahmen erarbeitet werden können.

Die Kommission wird sich ausserdem eingehend mit Rechtsfragen beschäftigen, die sich aus den Leitlinien für die Verbesserung des Zusammenwirkens zwischen öffentlichem und privatem Sektor ergeben, und geeignete Vorschläge für eine Harmonisierung der Vorschriften zur Vermarktung bestimmter Datenbestände vorbereiten. Weiter wird sie die Ausarbeitung von europäischen Verhaltenskodizes fördern und die Einhaltung von Regeln zum Schutz der Privatsphäre im Zusammenhang mit den elektronischen Informationsdiensten überwachen. Darüber hinaus werden gemeinschaftliche Leitlinien ausgearbeitet, die den Rahmen für vertragliche Vereinbarungen zwischen Marktteilnehmern bilden sollen, etwa in Fragen der Haftung des Herausgebers, Gewährleistung eines gleichbleibenden Dienstleistungsniveaus, Vertraulichkeit, geistiges Eigentum, Nutzung von Datenbanken und Rechte des Herausgebers.

Normen und Standards für die Strukturierung und Verarbeitung von Information sollen im Programm IMPACT 2 grösseres Gewicht bekommen. Vordringlich ist insbesondere die Vereinheitlichung von Format und Struktur aller Arten von Information, sowohl an den Systemschnittstellen als auch auf öffentlich vertriebenen Speichermedien, um den Austausch von Daten

und das Einbinden von vorhandenen Daten in neue Informationssysteme zu erleichtern. In erster Linie ist auf die Anwendung und Weiterentwicklung vorhandener Normen und Standards wie SGML, ODA und der Grafikschnittstellen-Formate hinzuwirken. Das ist durch Demonstrationsprojekte und direkte Unterstützung von Industrieübereinkommen zu fördern.

## **2. Stärkung der Wettbewerbsfähigkeit der europäischen Informationsdiensteanbieter**

Eine grosse Gefahr darf nicht verkannt werden: eine verbesserte Marktinfrastruktur öffnet den Markt für Informationsanbieter jeglicher Herkunft. Wenn nichts erfolgt, um auch die Qualität der Dienste europäischer Informationsanbieter zu verbessern, wird mit dem Binnenmarkt für Informationsdienste nur ein leichter Zugang zu aussereuropäischen Diensten geschaffen. Die Lage der europäischen Informationsindustrie würde sich dadurch in keiner Weise verbessern.

Die im Rahmen von IMPACT 1 durchgeführten Pilot- und Demonstrationsprojekte haben sich als wirksames Mittel zur Mobilisierung der Industrie und zur Anregung transnationaler Partnerschaften in der Gemeinschaft erwiesen. Solche Massnahmen zur Förderung der Zusammenarbeit sollen weitergeführt und intensiviert werden. Neben den derzeitigen Zuschüssen zu den Projekten werden zusätzliche Aktionen (Schaffung nationaler Anlaufstellen) erwogen, um kleine Unternehmen insbesondere in den benachteiligten Gebieten zu einer aktiven Beteiligung anzuregen.

Die Entwicklung von Gateways zwischen Hosts ist ebenfalls ein Mittel zur Ueberwindung der Zersplitterung des europäischen Marktes. Nach den Erkenntnissen des Beobachtungsgremiums werden aber die meisten Gateway-Vereinbarungen auf dem Markt spontan und zwischen Partnern im selben Land getroffen. Es wäre wünschenswert, die Entwicklung transeuropäischer

Gateways, aufbauend auf den Ergebnissen der Pilotprojekte und auf den Erfahrungen der Mitgliedstaaten, aktiver zu fördern.

Das Echo der Industrie auf die in der ersten Phase von IMPACT 1 ergangene Aufforderung zu Vorschlägen für Pilot- und Demonstrationsprojekte beweist, dass ein Bedarf nach Unterstützung von Informationsdiensten besteht, die wirtschaftlich noch nicht existenzfähig sind, es aber durch Stimulierungsmassnahmen werden können. Ihre Förderung ist besonders dort angebracht, wo die Mittel für die Anfangsinvestitionen nicht ohne weiteres auf dem Markt zu beschaffen sind.

Wie schon eingangs erwähnt, ist der elektronische Informationsmarkt recht heterogen. Zu ihm gehören alle Organisationen und Personen, die Teil der elektronischen Informationskette von der Erfassung über die Verarbeitung und Verbreitung bis zur Nutzung von Information sind. Diese werden von einer Vielzahl von Verbänden vertreten. Anregungen und Anstösse von seiten der EG sollten dazu führen, bei Projekten auf der Basis geteilter Kosten zusammenzuarbeiten. Im Rahmen einer solchen Zusammenarbeit könnten Vorhaben von allgemeinem Interesse finanziert werden.

### **3. Förderung der Verwendung fortschrittlicher Informationsdienste**

Die Aufnahme des Bereichs "intelligente Schnittstellen" in den Arbeitsplan von IMPACT 1 war ein erster Versuch, Entwicklungen zur Vereinfachung des Benutzerzugangs zu vorhandenen Informationsdiensten zu fördern. In diesem Zusammenhang ist es besonders wichtig, die Lücke zwischen der Forschung zur Entwicklung intelligenter Mensch-Maschine-Schnittstellen und der Umsetzung der Erkenntnisse in kommerzielle Anwendungen zu schliessen.

Das "Kiosk"-Konzept hat sich technisch und kommerziell als brauchbar für die Vereinfachung des Benutzerzugangs

zu einer Vielzahl weiträumig verteilter Videotex-Dienste erwiesen. Es wird daher vorgeschlagen, ein übergreifendes Projekt zur Erprobung der Funktionsfähigkeit von europäischen professionellen Kiosk-Dienstleistungen auszuarbeiten, der Klein- und Mittelunternehmen einfachen Zugang zu professionellen Audiotex-, Videotex- und ASCII-Informationsdiensten verschafft.

Wegen der wachsenden Bedeutung verschiedener Darstellungs- und Übertragungsformen wird IMPACT 2 auch in diesem Bereich aktiv werden. Gedacht ist an die Förderung von kombinierter Text-, Grafik-, Bild- und Toninformation sowie der Integration von Offline- und Online-Systemen.

Die Vielzahl der Sprachen behindert den Zugang zu den Informationsbeständen in der Gemeinschaft. Doch die Ueberwindung der Sprachbarrieren ist sehr kostenaufwendig, denn zusätzlich zur Uebersetzung erfordert sie meistens noch internationale Zusammenarbeit für die lokale Unterstützung der Benutzer in ihrer jeweiligen Sprache. Die finanzielle Unterstützung von Partnerschaften für Pilotanwendungen vorhandener mehrsprachiger Systeme durch die Gemeinschaft ist daher vorgesehen.

Mangelndes Wissen um den Wert von Information und über vorhandene Informationsdienste und deren Nutzung ist ein wesentliches Hemmnis für das Marktwachstum. Die von der Kommission im Rahmen von ECHO initiierten Sensibilisierungsmassnahmen sollen verstärkt werden.

#### **4. Stärkung der gemeinschaftlichen Anstrengungen im Hinblick auf eine gemeinschaftsweit gleichmässige Entwicklung der Informationsdienste**

Das Ziel der Schaffung eines Binnenmarktes für Informationsdienste wird so lange nicht voll erreicht, wie in einigen Regionen ein Markt von sehr bescheidenem Ausmass besteht. In den benachteiligten Regionen ist der Markt

für elektronische Informationsdienste so unterentwickelt, dass dort nur wenige Stellen von der Informationsmarktpolitik der Europäischen Gemeinschaft wirklich profitieren können. Daher müssen bestehende Programme für diese Regionen (Regionalförderung) auch für die Ziele von IMPACT genutzt werden.

Dabei sollte besonderes Gewicht auf Ausbildungsmassnahmen gelegt werden. Ziel ist die Ausbildung von Informationsfachleuten und die Unterstützung der Einrichtung von Fachzentren, die später an die Tätigkeit der Gemeinschaft anknüpfen könnten.

**Rainer Kuhlen**  
**Universität Konstanz - Informationswissenschaft**  
**Postfach 5560 - 7750 Konstanz**

**Zur Theorie informationeller Mehrwerte**

**Zusammenfassung.** Das Konzept informationeller Mehrwerte erweist sich als gleichermaßen produktiv für eine allgemeine informationswissenschaftliche Theorie und für die Einschätzung der Leistung und der Akzeptanz von elektronischen Mehrwertdiensten auf dem allgemeinen Informationsmarkt. Durch eine Literaturübersicht werden gegenwärtige Tendenzen der Mehrwertdiskussion herausgestellt, wobei vor allem die elektronischen Mehrwertdienste (VANS) auf der Basis elektronischer Kommunikationsnetze eine zentrale Rolle spielen. Über fünf Thesen wird diese Diskussion zusammengefaßt. Eine Systematisierung informationeller Mehrwertbildung wird über vier verschiedene Mehrwert(bildungs)typen unternommen: a) größerer Informationswert eines elektronischen Mediums gegenüber dem entsprechenden konventionellen; b) inhärenter Mehrwert bei bestehenden Produkten; c) Mehrwerteffekte durch Agglomeration; d) Mehrwerteffekte durch Integration. Der erste Typ wird mit einer detaillierten Diskussion informationeller Mehrwerte durch eine Dynamisierung linearer und nicht-linearer Metainformationen bei der Konversion von Text in Hypertext erläutert.

[Der folgende Text spiegelt den Vortrag nur sehr ansatzweise wider, da er als Vortragslinearisierung auf der Grundlage eines nicht-linearen Hypertextes entstehen wird. Verschiedene Hypertext-Mehrwerteffekte, z.B. über Animationen, "gezoomte" Bilder oder komplexe Graphiken können in der gedruckten Version nicht wiedergegeben werden. Ein Hinweis auf die sich anbahnende Aufgabe linear produzierter Proceedings? Interessenten (mit Zugriffe auf Toolbook unter Windows 3.0) können die entsprechende Hypertextbasis beim Autor gegen Selbstkosten anfordern.]

**Inhalt**

1. Zum Konzept des informationellen Mehrwerts
2. Zum Stand der Diskussion informationeller Mehrwerte
3. Eine Systematisierung der Typen informationeller Mehrwertbildung



## 1. Zum Konzept des informationellen Mehrwerts

Informationswissenschaft wird von mir zunehmend als Theorie zur Erzeugung informationeller Mehrwerte verstanden; oder anders: das Konzept des informationellen Mehrwertes erweist sich als zentral zur Einschätzung der Leistung und der Akzeptanz von elektronischen Mehrwertdiensten. Dies ist kein grundsätzlich neuer Ansatz. Robert S. Taylor hat in seinem Buch "Value-added processes in information systems" 1986 Wert und Mehrwert von Information zum systematischen Ansatzpunkt gewählt. Er hat dort vier verschiedene Interpretationen von "value-added" unterschieden (17ff passim):

"The first interpretation is the use of "value-added" or "added value" as used in the discipline of economics ... Basically, added value is the creation of wealth ...

The second interpretation of the concept is one more central...: What characteristics or attributes are added to the data and information items being processed that make them more useful (i.e. valuable, beneficial) to users, clients, customers, than they were at the start of the process? This interpretation of value goes beyond that of the economist, for in a sense we are asking precisely what values are indeed being added during the period of processing or producing, activities into which the producer invests capital, labor, and materials ...

The third possible interpretation of value has to do with the relationship of the customer or client to the information. What, for example, would a person in a given situation pay in dollars for the information? ...

A fourth interpretation of value also has to do with the client or user and with the context within which information is used ... Value, in this context, has to do with the benefit a user or an organization accrues as a result of information use (or even, dare we say, of nonuse)."

Der erste, der ökonomische Begriff von Mehrwert soll in diesem Beitrag nicht im Zentrum des Interesses stehen. Dennoch soll nicht vergessen werden, obgleich es als Bestandteil realer politischer Praxis schon sehr tiefe, kaum mehr erinnerbare Vergangenheit zu sein scheint, daß der Begriff des Mehrwerts eine kritische Funktion in der auf Marx zurückgehenden politischen Ökonomie bzw. in der allgemeinen Werttheorie gespielt hat. Die "Plusmacherei", wie sie auch von Marx genannt wird, wurde nicht nur von Marx, sondern durchaus auch von bürgerlichen Ökonomen als "Geheimnis der Kapitalverwertung und als das Grundgesetz der kapitalistischen Produktionsweise" angesehen (Schmidtgall 1980), wobei allerdings die letzteren kaum die politischen bzw. klassenkämpferischen Konsequenzen daraus gezogen haben. Mehrwerte entstehen im ökonomischen Sinne dadurch, daß durch die an den Kapitaleigner ausgeliehene und bezahlte Arbeitskraft ein höherer Wert produziert wird, als diese selbst verkörpert. Ein Beispiel aus unserem Gebiet: Ein Informationswissenschaftler oder wissenschaftlicher Dokumentar, der für einen Datenbasishersteller ein *Abstract* schreibt, verleiht dem Ausgangstext zusammen mit dem *Abstract* einen höheren Wert als der Wert der Arbeit des *Abstracting* für sich dar-

stellt. Der erzielte Mehrwert wird auf dem Markt durch entsprechende Preisbildung abgeschöpft.

Wir sehen schon an diesem einfachen Beispiel, daß in den Prozeß informationeller Mehrwertbildung neben der unmittelbaren Mehrwerterzeugung (hier durch das *Abstracting*) entscheidend Ergebnisse gesellschaftlich erzeugter und vor allem gesellschaftlich, d.h. durch die Allgemeinheit finanzierter Arbeit eingehen, hier also die mit *Abstracts* zu versendenden Ausgangsdokumente. Die Aneignung des monetär anrechenbaren und durch weitere Leistungen entstehenden Mehrwerts durch private Anbieter von Informationsprodukten und -dienstleistungen auf dem allgemeinen Informationsmarkt ist damit zumindest ambivalent (vgl. Kuhlen 1987a) - einerseits das Grundprinzip unserer Wirtschaftsordnung (Information als Ware bzw. freies Gut), andererseits der Anspruch aller gesellschaftlichen Gruppen (als Finanzierer der Produktion von Wissen) auf freien Zugriff zur gesellschaftlich produzierten Basisinformation (Information als öffentliches Gut - Anspruch auf eine informationelle Grundversorgung).

Verlassen wir dieses heute sehr weit und sehr abstrakt gewordene Feld und zitieren einen in diesem Zusammenhang sicherlich unverdächtigen Autor, nämlich Stonier in seinem Buch "The wealth of information" von 1983. Er versucht, den Begriff des Mehrwerts (*value-added*, auch: *surplus value*) auf jegliches Material oder jegliche Informationsdienstleistung anzuwenden, die so bearbeitet worden sind, "that the final product is more valuable than the original." (17). Entsprechend gilt in der Informationsökonomie als Grundprinzip, daß aus Objekten beliebiger Art nicht länger durch Arbeit, Kapital, Energie oder Grund und Boden allein, sondern zusätzlich und vor allem durch Information, d.h. näher durch Informationsarbeit, Mehrwertobjekte entstehen (vgl. Abb. 1).

Um noch einmal Stonier zu zitieren:

"In the agrarian economy, the economy was concerned primarily with producing enough food and the limiting factor usually was the availability of good land. In the industrial economy, the economy was geared to the production of goods and the limiting factor tended to be capital. In an information economy, the economy is concerned with the creating and application of information to make all forms of production more efficient and to create new wealth. The limiting factor tends to be existing knowledge" (21).

### **Abb. 1 Allgemeiner Prozeß der Erzeugung informationeller Mehrwert**

## **2. Zum Stand der Diskussion informationeller Mehrwerte**

Seit den frühen Arbeiten der Informationsökonomien hat der Mehrwertbegriff Konjunktur. Eine begrifflich sehr eng gehaltene Recherche in der informationswissenschaftlichen Datenbank INFODATA ergab für die letzten vier Jahre 45 einschlägige Dokumente. Ich fasse die Tendenzen kurz zusammen:

Die offensichtlich breiteste Verwendung findet der Mehrwert/*value-added*-Begriff im Telekommunikationsbereich. "Mehrwertdienste" ist eine direkte Übersetzung des amerikanischen "*value added (network) services* (VANS)". Eine begriffliche Klärung unternimmt Müller (1991), indem er Mehrwertdienste als "zusätzliche Wertschöpfung" "auf der Basis bestehender Fernmeldedienste" begreift: Die Wertschöpfung "kann von relativ einfachen Diensten, wie zum Beispiel Weck- und Ansagedienste, zu elektronischen Datenbanken und elektronischem Datenaustausch bis hin zu Systemüberwachung und -steuerung reichen" (aus dem Autorenreferat). Als Kennzeichen von Mehrwertdiensten gilt, daß sie über den bloßen Informationstransport hinausgehen und neben (temporären oder dauerhaften) Speicherungsmöglichkeiten vor allem auch die Verarbeitung von Benutzerinformationen im Netzwerk über verteilte Rechner anbieten (Frantzen/Trox 1988; vgl. auch Peuckert 1987 speziell mit Rücksicht auf Büroanwendungen). Besonders die Kombination bislang getrennter Bereiche (Daten, Texte, gesprochene Sprache, jetzt zunehmend mehr multimediales Material) bewirken ganz neue Mehrwerteffekte (Kuhlen 1991). Auf einige rechtliche Probleme im Zusammenhang von Mehrwertdiensten weist Reidenberg (1988) hin.

Besondere Mehrwerteffekte werden durch die sich entwickelnden Standards im OSI-Modell erwartet. Damit sind spezielle Kommunikationsdienstleistungen angesprochen, die innerhalb, aber vor allem oberhalb der 7 Schichten im allgemeinen OSI-Modell entwickelt und angeboten werden (Denenberg 1990), z.B. auf der Basis von X.400 (für elektronische Post) und FTAM (für den Datentransfer, z.B. das Osiris-F-Projekt der Deutschen Bank). Durch die Liberalisierung des Telekommunikationsbereichs (Entzerrung des staatlichen Monopols in den meisten Industrie-/Informationsländern) ergeben sich für elektronische Diensteanwender dadurch Wettbewerbsvorteile, daß sie über die angemieteten oder sogar selbst erstellten Netze Mehrwertdienste anbieten (Franz/Eberhard 1991). In unserer zweiten These zum Mehrwert weisen wir auf diese Wettbewerbsvorteile auf dem Informationsmarkt durch Mehrwertleistungen explizit hin (s. unten).

Auf der Grundlage dieser Mehrwertdienste sehen verschiedene Autoren auch Mehrwerteffekte im Kontext des *Informationsmanagement* bzw. in der Nutzung von Information für operative, aber vor allem strategische Zwecke (Beyenburg/Emde 1990; Kuhlen 1987b). Es ist dies also eine besondere Herausforderung an das strategische Informationsmanagement, integrierte Kommunikationssysteme aufzubauen und die entsprechenden Mehrwertdienste, also kommunikationstechnische Dienstleistungen (E-mail, allgemeiner Datenaustausch), für organisationelle Zwecke auszunutzen (Müller 1990; Huttel 1990). Gerade der Dienstleistungs-, vor allem der Bankbereich, nimmt hier bevorzugt die offensichtlichen Mehrwerteffekte wahr (Arnold 1990). Auf dem Informationsmarkt läßt sich hier eine Entwicklung feststellen, die auch für andere elektronische Dienst- und Vermittlungsleistungen gilt. Verschiedentlich wandeln sich frühe Benutzer von Mehrwertdiensten zu eigenen Anbietern von VANS. Dabei werden bislang von Kunden eindeutig Komplettlösungen auf der Basis international anerkannter Schnittstellen-Standards bevorzugt. Dies entspricht dem Prinzip des im Amerikanischen beliebten "One-Stop-Shopping-

Service", der auch schon im Agglomerationseffekt bei den Datenbankanbietern angewendet wird; s. unten (Diebold MR 1989).

Gegenüber dem florierenden amerikanischen, allgemein angelsächsischen Markt (Chang/Hitchcock 1989) kann man im deutschsprachigen Bereich eher noch Zurückhaltung feststellen, obgleich, wie Schlieker schon 1988 festgestellt hatte, sich im Prinzip ein riesiger Markt dadurch ergibt, daß elektronische Mehrwertdienste durchaus auch für kleine und mittlere Unternehmen in der Bundesrepublik attraktiv sein können (Schlieker 1988). Ein Überblick über Anwendungen von Mehrwertdiensten in Deutschland und Europa wird durch die Beiträge in Häusser (1991) gegeben; vgl. auch VANS 90 report (1990).

Je nach vertretenem Standpunkt wird auch ein engerer Begriff von Mehrwertdienst verwendet. So setzt von Wietersheim (1991) den VANS-Markt mit dem Online-Markt weitgehend gleich, eine Tendenz, die ja auch für das informationspolitische Verständnis des BMFT bezüglich Fachinformation allgemein zutrifft. Auf Mehrwerteffekte im Bibliothekswesen weisen Demo/MacClure (1988) allgemein und spezieller Laughlin (1989) für das Management durch einen entsprechenden Bibliotheksservice hin. Mehrwerteffekte durch intelligente "front ends" und Benutzerschnittstellen werden verschiedentlich im Kontext der allgemeinen *Gateway*-Diskussion behandelt, z.B. Holmes (1989). Auf Verbesserungen der Endbenutzersituation durch Mehrwertdienstleistungen weisen Burris/Molink (1991), mit Blick auf das Retrieval Tenopir et al. (1989) das "Downloading" von Datenbeständen Bauer (1989), auf Mehrwerteffekte durch CD-ROM-Nutzung Taylor (1989) und auf die allgemeinen pragmatischen Mehrwerteffekte Kuhlen (1989, 1990, 1991) hin.

Nun gehen uns in diesem Vortrag weniger die allgemeinen informationsökonomischen Gesichtspunkte der Mehrwerttheorie an (vgl. Repo 1987), sondern mehr die informationsmethodischen Gesichtspunkte. Wir wollen also hier nicht weiter behandeln, wie z.B. die landwirtschaftliche Produktion die professionelle Verwendung von Information (z.B. über Boden, Saatgut, Düngemittel, Weltpreisentwicklung) oder Industrieprodukte durch informationsgesteuerte Roboter oder die Verwaltung durch den Einsatz von Informationssystemen in ihrer Effizienz gesteigert und so Mehrwerte erzeugt werden (zu diesen Zusammenhängen vgl. Jonscher 1983), sondern uns in einem informationsmethodischen Ansatz darauf beschränken, darzustellen, wodurch informationelle Mehrwerte an sich entstehen. Oder anders formuliert: wenn wir es als die grundsätzliche Aufgabe von Informationsarbeit ansehen, aus existierenden Beständen von Wissen aktuelle, pragmatisch relevante Information zu erzeugen, dann könnte der Zweck der Erzeugung informationeller Mehrwerte darin bestehen, aus dem Rohprodukt Wissen ein handlungsrelevantes Produkt "Information" zu machen. Wir haben in unserem Beitrag für die ISI' 90-Tagung (Kuhlen 1990) diese Prozesse der Mehrwerterzeugung durch Differenzierungen im

Begriff der Informationsarbeit ausführlich behandelt und wollen uns in diesem Beitrag auf die Systematisierung der Mehrwerteffekte beschränken.

Fassen wir die bisherige Diskussion in fünf allgemeinen Thesen zusammen, so ergibt sich folgendes Bild:

**These 1:** Mehrwert bei neuen Produkten und Dienstleistungen des Informationsmarktes ist in erster Linie Ergebnis methodisch kontrollierter Informationsarbeit.

**These 2:** Neue Produkte und Dienstleistungen des Informationsmarktes werden nach dem Kriterium des informationellen Mehrwertes beurteilt, den sie für Benutzer erbringen. Auf dem internationalen Informationsmarkt erzielen die Anbieter von Informationsprodukten oder -dienstleistungen dann Wettbewerbsvorteile, wenn sie diese gegenüber Konkurrenten mit informationellen Mehrwerten ausstatten können.

**These 3:** Mehrwerteffekte bei elektronischen Kommunikationsdiensten entstehen dadurch, daß über Transport und Speicherung von Daten hinaus Verarbeitungsleistungen im Sinne der These 1 angeboten werden.

**These 4:** Mehrwerteffekte werden dann besonders erzielt, wenn mit den Verarbeitungsleistungen - das sind die Ergebnisse von Informationsarbeit - auf die pragmatischen Rahmenbedingungen individueller oder organisationaler Art eingegangen werden kann.

**These 5:** Mehrwerteffekte entstehen besonders durch die Kombination von Informationsdiensten, die außerhalb der elektronischen Kommunikation eher getrennt angeboten wurden.

### **3. Eine Systematisierung der Typen informationeller Mehrwertbildung**

Im weiteren Verlauf der Darstellung wird eine Systematisierung der verschiedenen Typen informationeller Mehrwertbildung aus informationswissenschaftlicher Sicht unternommen.

#### **Abb. 2 Typen informationeller Mehrwertbildung**

Der Mehrwert der elektronischen Version gegenüber der korrespondierenden soll hier nur an zwei Beispielen aus dem On-line- und dem Hypertextbereich verdeutlicht werden (vgl. Abb. 3-5)

#### **Abb.3 Beispiele für Mehrwertbildung des ersten Typs**

#### **Abb.4 Mehrwerteffekte des On-line-Retrieval**

#### **Abb. 5 Mehrwerteffekte von nicht-linearen Hypertexten gegenüber linearen Büchern**

Von den verschiedenen Möglichkeiten aus Typ 1, informationelle Mehrwerte zu erzeugen, soll hier noch ein spezieller Mehrwerteffekt eines nicht-linearen Mediums gegenüber linearen gezeigt werden, und zwar speziell unter dem Gesichtspunkt der Dynamisierung von Metainformationen. Mit "Metainformationen" werden solche Bestandteile von Informationssystemen bezeichnet werden, die indirekt informativ sind, d.h. die von Benutzern in der Regel nicht als Selbstzweck eingesehen werden, sondern die dazu dienen, Orientierung zu verschaffen. Sind solche Metainformationen in allen Informationssystemen nützliche Orientierungs- und Einstiegshilfen, so sind sie bei nicht-linearen Systemen wie Hypertext unverzichtbar.

Metainformationen in Büchern werden über den Übersichtszweck hinaus benutzt, um interessierende Passagen in Texten schnell zu finden oder wiederzufinden. Sie sind aufgrund der direkten Einstiegsmöglichkeit als nicht-lineare Mittel anzusehen. Gleichwohl werden sie in Texten meistens linear präsentiert. Verschiedentlich wird auch schon in sehr tief strukturierten Büchern versucht, Inhaltsverzeichnisse dynamisch aufzubauen, z.B. am Anfang eines Buches nur bis zu zwei Hierarchiestufen anzuzeigen, bei jedem Kapitel oder Abschnitt dann entsprechend tiefer. Diese Idee - und das ist typisch für eine hypertextspezifische Mehrwerterzeugung - kann bei Hypertext radikalisiert werden. Abb. 6 zeigt ein sogenanntes dynamisches Inhaltsverzeichnis, das durch "Aufzoomen" (vgl. Furnas 1986) der ersten Abstraktionsstufe der Gliederung entsteht. Die Hypertextforschung hat einige Evidenz dafür gefunden (vgl. Kuhlen 1991), daß flexible dynamisierte Inhaltsverzeichnisse für die Orientierung und den Einstieg besser geeignet sind als bloß alphabetische lineare Register. Für weitere flexibilisierte Metainformationen (z.B. Register, Graphische Browser) vgl. die Übersichten in Kuhlen (1991).

#### **Abb. 6 Dynamisches Inhaltsverzeichnis**

Nach unserer Einschätzung werden dynamische, Mehrwert erzeugende Flexibilisierungsleistungen auch unabhängig von der speziellen Hypertextentwicklung allgemein zur Ausstattung von elektronischen Informationssystemen gehören. Je komplexer die Nutzungsangebote von elektronischen Systemen werden, umso wichtiger werden flexibilisierte Metainformationen. Sorgen für die Produktion von Wissen die einzelnen Fachwissenschaften, so ist für die Konzeption und Realisierung von solchen flexibilisierten Metainformationen die Informationswissenschaft zuständig.

Der zweite Mehrwerteffekt - wir wollen ihn auch den *inhärenten Mehrwerteffekt* nennen - ergibt sich durch Verbesserung einzelner Komponenten bereits existierender Systeme oder durch eine systematische Verbesserung der Gesamtsystemleistung. Hierzu gibt es sehr viele Beispiele, z.B.

- natürlichsprachiger Zugriff zu einer ansonsten mit einer formalen Abfragesprache recherchierbaren Daten- oder On-line Informationsbank

- Komponente zur Erleichterung der Frageformulierung durch einen on-line verfügbaren Thesaurus
- Stufenanpassung der Retrievalmöglichkeiten an unterschiedliche Benutzer, z.B. die verschiedenen Ebenen des Retrieval bei der Gateway-Software von EASYNET.
- Verknüpfung einer bibliographischen (Referenz-) Informationsbank mit einer entsprechenden Volltextdatenbank, z.B. der BRS/LINK von BRS Information Technologies, mit dem aus der bibliographischen Datenbank MEDLINE direkt die fachlich entsprechende Volltextdatenbank *Comprehensive Core Medical Library* - CCML aufgerufen werden kann, ohne daß eine neue Recherche aktiviert werden muß.

Der dritte Mehrwerteffekt durch Agglomeration ist ebenfalls des längeren in der Informationspraxis bekannt. Abb. 7 führt zwei Beispiele an. Der Mehrwerteffekt für den Benutzer ergibt sich hier vor allem durch das "one-stop-shopping"-Prinzip, eines der Erklärungsmuster für den Erfolg großer Datenbankanbieter, wie z.B. DIALOG.

**Abb. 7 Zwei Beispiele für Mehrwerteffekte durch Agglomeration**

Die meisten Mehrwerteffekte werden ganz offensichtlich durch den vierten Typ erzielt. Gerade die Informatisierung der Informations- und Kommunikationsindustrie hat diese Effekte möglich gemacht. Abb. 8 führt nur einige Beispiele für diesen integrierenden Mehrwerteffekt an. Andere sind z.B. die Integration von Retrieval- und Browsing-Komponenten in offenen Hypertextsystemen oder die breite Palette des Angebots von multimedialen Produkten. Wir wollen aber betonen, daß die bloße integrierte Kombination technologisch bislang getrennter Teilbereiche kaum zu den erwünschten Effekten führen wird, sondern auf eine ebenfalls integrierte Methodologie, z.B. bezüglich der Benutzerschnittstelle, der Suchmöglichkeiten oder der Techniken der Wissensrepräsentation, angewiesen ist. Wie wir zu Anfang hingewiesen haben, hängt der Erfolg von Mehrwertprozessen in erster Linie von methodisch kontrollierter Informationsarbeit, weniger von der technischen Perfektion, ab.

**Abb. 8 Einige Beispiele für integrierte Mehrwerteffekte**

Mehrwerteffekte, gezeigt an dem kleinen Beispiel des dynamischen Inhaltsverzeichnisses aus der Hypertextwelt, aber auch in der auf alle Informationsprodukte und -dienstleistungen ausgedehnten Verallgemeinerung, können als Kriterium für Leistung, Erfolg und Akzeptanz von eben diesen Informationsprodukten und -dienstleistungen angesehen werden. Sie stehen damit im Zentrum der theoretischen und praktischen Aufmerksamkeit gleichermaßen.

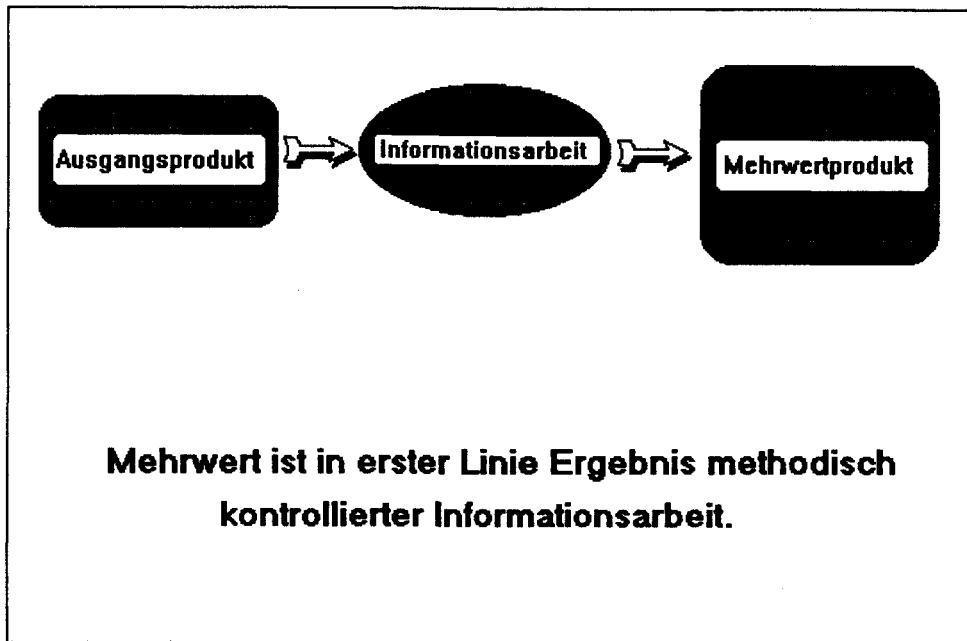


Abb. 1 Allgemeiner Prozeß der Erzeugung informationeller Mehrwert

## Typen informationeller Mehrwertbildung

- 1) Größerer Informationswert der elektronischen Version gegenüber der korrespondierenden konventionellen
- 2) Größerer Informationswert durch Verbesserung einzelner Komponenten eines elektronischen Produkts (oder einer Dienstleistung) oder durch Verbesserung der Gesamtleistung
- 3) Größerer Informationswert durch Agglomeration ursprünglich isolierter Leistungen oder Produkte
- 4) Kombination verschiedener Typen von Informationsprodukten oder -dienstleistungen

Abb. 2 Typen informationeller Mehrwertbildung



## Typen informationeller Mehrwertbildung

1) Größerer Informationswert der elektronischen Version gegenüber der korrespondierenden konventionellen

Beispiel1: Online-Informationsbanken gegenüber Katalogen von Bibliothekssystemen

Beispiel2: Nicht-linearer Hypertext gegenüber linearem Buch



Abb.3 Beispiele für Mehrwertbildung des ersten Typs

## Typen informationeller Mehrwertbildung

**wie zu messen?**

**Schnelligkeit**  
**Genauigkeit**  
**Vollständigkeit**  
**Aktualität**  
**Abdeckung**

sche

**Merkmale:**

**postkoordiniert vs. linearer Präkoordination**

**selektiv**

**schnell**

**aktuell**

**preiswert**

**zuverlässig**

Beispiel1: Online-Informationsbanken gegenüber Katalogen von Bibliothekssystemen

Beispiel2: Nicht-linearer Hypertext gegenüber l



Abb.4 Mehrwerteffekte des On-line-Retrieval

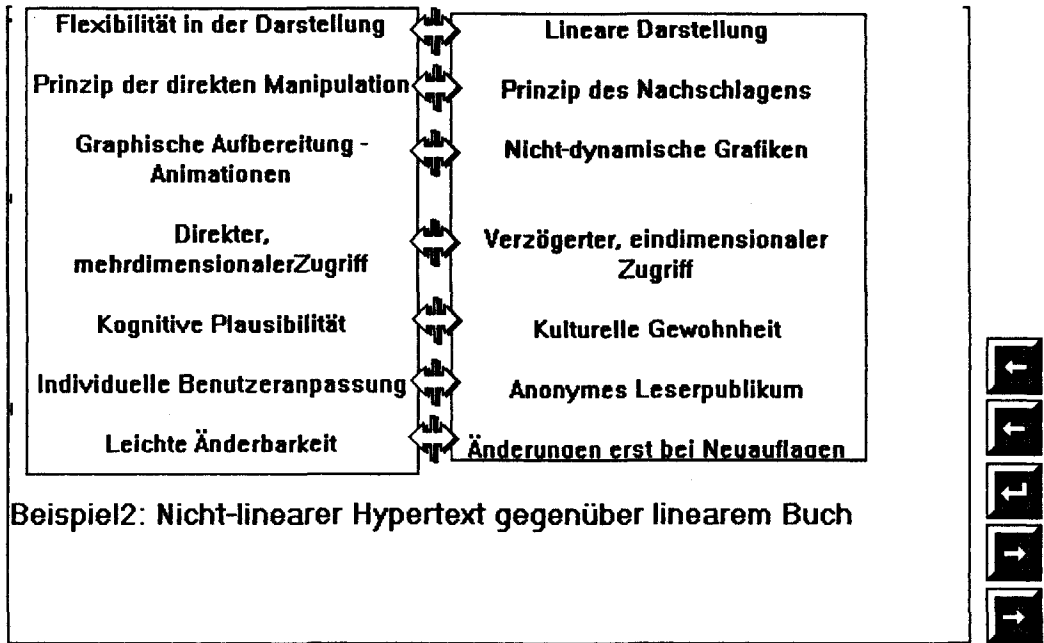


Abb. 5 Mehrwerteffekte von nicht-linearen Hypertexten gegenüber linearen Büchern

1.1 Thesen zum informationellen Mehrwert These 1 Mehrwert durch Informationsarbeit These 2 Mehrwert Nutzen-Kosten-Balance These 3 Marktvorteile durch Mehrwerte	2.1 Grundidee von Hypertext
1.2 Typen informationeller Mehrwertbildung elektronisch vs. konventionell inhärenter Mehrwert Agglomerationseffekt Mehrwert durch integrierende Kombination	2.1.1 Nicht-Linearität
1.3 Mehrwert durch Informationsarbeit	2.1.2 Animation zu Nicht-Linearität
<b>1. Zur Theorie informationeller Mehrwerte</b>	2.1.3 Kognitive Plausibilität
<b>2. Zur Hypertext-Methodologie</b>	2.2 Systemarchitektur
<b>3. Beispiele informationeller Mehrwerte durch Hypertext</b>	2.2.1 Hypertextbasis
<b>4. Konsequenzen für Informationswissenschaft und -praxis (Abschätzung informationeller Mehrwerte)</b>	2.2.2 Authoring/Konversion
	2.2.3 Browsing/Suche
	2.3 Browsing vs. Matching
	2.4 Formen des Browsing
	2.4.1 Gerichtetes "Browsing" mit Mitnahmeeffekt
	2.4.2 Gerichtetes "Browsing" mit "Serendipity"-Effekt
	2.4.3 Ungerichtetes Browsing
	2.4.2 Assoziatives Browsing

Abb. 6 Dynamisches Inhaltsverzeichnis

## Beispiele für Agglomerationsmehrwerte

- a) Universalhosts: viele Datenbanken mit einer Datenverwaltung, einer homogenen Retrievalsprache, einheitlichem Kundenservice, z.B. durchsichtige Kostenabrechnung, effizientes Marketing
- b) Gateways (in den verschiedenen Organisationsformen und Mehrwert erzeugenden Dienstleistungen)



Abb. 7 Zwei Beispiele für Mehrwerteffekte durch Agglomeration

## Beispiele für integrierte Mehrwerte

- a) Integrierte Systeme (Textverarbeitung, Datenbank, Graphik, Spreadsheet, Kommunikation), z.B. Lotus 1.2.3, Symphony, KnowledgeMan
- b) Verteilte (graphisch unterstützte) Dokumentverwaltungs- / Datenbanksysteme, z.B. LOTUS Notes
- c) Lokales (offenes) Hypertextsystem mit (einbettendem) Zugriff auf externe Software, z.B.
  - "warm links" zu Spreadsheetprogrammen,
  - "communication links" zu "email"-/Konferenzsystemen,
  - "query links" zu externen Datenbanken oder Online-Informationenbanken



Abb. 8 Einige Beispiele für integrierte Mehrwerteffekte

## Literaturhinweise

- Arnold, F. (ed.) (1990):** Electronic Banking: Mehrwertdienste für das Finanzmanagement (Symposium 1-3): *Telekommunikation in Europa: Quo vadis? Europäische Kongressmesse für Technische Kommunikation*. Online 90. Velbert: Online 1990.
- Bauer, P. (1989):** Veredelung eines Rohstoffs. Wie man Recherche-Ergebnisse inhouse verarbeiten kann. *Cogito. Neue Wege zum Wissen der Welt, Informationen wirtschaftlich nutzen* 5, 1989, 3, 26-32.
- Beyenburg, R.; Emde, P. (1990):** Interaktion im globalen Netz. net. *Zeitschrift für angewandte Telekommunikation* 44, 1990, 1/2, 16-18, 20-21.
- Burris, R.A.; Molinek, F.R. (1991):** Establishing and managing a successful end-user search service in a large special library. *Online. The magazine of online information systems* 15, 1991, 2, 36-39.
- Chang, C.; Hitchcock, D. (1989):** The VANS handbook. Pinner, GB: Blenheim Online: 1989.
- Demo, T.L.; MacClure, C.R. (1988):** Information and referral in the academic library. Lessons in attitude and service from the public library. *The reference librarian* 21, 1988, 95-108.
- Denenberg, R. (1990):** A special issue on open systems interconnection. *Library Hi Tech* 8, 1990, 4, 7-144 [dieses Sonderheft enthält 11 Beiträge].
- Diebold MR (1989):** Konjunktur für Mehrwertdienste. *Diebold Management Report. Analysen und Meldungen zu aktuellen Fragen der Informationsverarbeitung* 1989, 4, 4-10.
- Frantzen, V.; Trox, R. (1988):** Mehr Effizienz durch Mehrwertdienste. Mehrwertdienste der Text-, Daten- und Sprachkommunikation in öffentlichen Fernmeldenetzen. *Datenschutz und Datensicherung (DUD): Recht und Sicherheit der Informations- und Kommunikationssysteme* 1988, 7, 329-336.
- Franz, W.; Eberhardt, R. (1991):** Mehrwertdienste und deren Umfeld: *Nachrichtentechnische Zeitschrift. Informationstechnik und Telematik für Experten* 44, 1991, 7, 472-475.
- Furnas, G.W. (1986):** Generalized fisheye views; in: *Proceedings of the ACM CHI '86 Conference on Human Factors in Computing Systems*. Boston, Mass. 13.-17. April 1986. New York: ACM, 16-23.
- Häusser, E. (ed.) (1991):** Informations- und Mehrwertdienste in Deutschland und Europa (Symposium 5-1): *Europäischer Congress für Mehrwertdienste, Informations-Datenbanken und CD-ROM. Europäische Congressmesse für Technische Kommunikation*. Online 91. Velbert: Online 1991.
- Holmes, P.L. (1989):** Intelligent front end interfaces. *Online information 89. 13th international online meeting*. London 12.-14. Dez. 1989. London: Learned Information 1989, 587-596.
- Huttel, G. (1990):** Unternehmenskommunikation ohne Grenzen - Value Added Networks; in: *Wettbewerbsfaktor Informationsmanagement. Herausforderung für Marketing und Vertrieb*. Tagung Köln 31.5.-1.6.1991. Düsseldorf: VDI-Verlag 1990.
- Jonscher, C. (1983):** Information resources and economic productivity. *Information Economics and Politics*, Vol. 1, 1983, 13-35.-
- Kuhlen, R. (1987a):** Information in der informierten Gesellschaft - Politische, ökonomische und technische Rahmenbedingungen von Informations- und Dokumentationsprogrammen. *Gewerkschaftliche Monatshefte* 38, 1987, H.6, S.337-352.-
- Kuhlen, R. (1987b):** Informationsmanagement der Zukunft. ik Report 1, 1987, 10-17.
- Kuhlen, R. (1989):** Pragmatischer Mehrwert von Information. Sprachspiele mit informationswissenschaftlichen Grundbegriffen. Bericht 1/89. Universität Konstanz, Informationswissenschaft, Okt. 1989 [engl. Version Computer and the Humanities 1991].
- Kuhlen, R. (1990):** Zum Stand pragmatischer Forschung in der Informationswissenschaft; in: Herget, J.; Kuhlen, R. (eds.): *Pragmatische Aspekte beim Entwurf und Betrieb von Informationssystemen. Proceedings des 1. Internationalen Symposiums für Informationswissenschaft*. Konstanz: Universitätsverlag Konstanz 1990, 13-18.
- Kuhlen, R. (1991):** Hypertext - ein nicht-lineares Medium zwischen Buch und Wissensbank. Heidelberg: Springer 1991.
- Laughlin, K. (1989):** The corporate library of the future; in: Lee, J.H. (ed.): *Online searching. The basics, settings, and management*. Englewood, CO: Libraries Unlimited 1989, 82-89.
- Müller, J. (1991):** Mehrwertdienste. Neue Dienstleistungen für Unternehmen. *Computer und Recht. Forum für die Praxis des Rechts der Datenverarbeitung, Information und Automation* 7, 1991, 1, 45-48.
- Peuckert, H. (1987):** Value added services für das Büro. *Proceedings GI - 17. Jahrestagung Computerintegrierter Arbeitsplatz im Büro*. Informatik-Fachberichte 156. Berlin: Springer 1987, 135-149
- Reidenberg, J. (1988):** Information property. Some intellectual property aspects of the global information economy. *Information age* 10, 1988, 1, 3-12.

- Repo, A.J.** (1987): Economics of information. *Annual review of information science and technology* Vol. 22, 1987, 3-35.
- Schlieker, J.** (1988): Entwicklung von Nachfrage und Angebot für Mehrwertdienste für kleine und mittlere Unternehmen sowie Handwerksbetriebe. Karlsruhe:1988.
- Schmidtgall, H.** 1980: Artikel Mehrwert, in: *Historisches Wörterbuch der Philosophie* Bd.5. Spalte 2011. Basel/Stuttgart: Schwabe & Co 1989.
- Stonier, T.** (1983): The wealth of information. A profile of the post-industrial economy. London: Thames Methuen 1983.
- Taylor, H.** (1989): Science citation index print and CD. The best of both worlds from ISI; in: Helal, A.H.; Weiss, J.W.: *The impact of CD-ROM on library operations and universal availability of information. 11th International Essen Symposium*. Essen: Veröffentlichungen der Universitätsbibliothek Essen 11, 1989, 45-61.
- Taylor, R.S.** (1986): Value-added processes in information systems. Norwood, NJ: Ablex 1986.
- Tenopir et al.** (1989) [Tenopir, C.; Nahl-Jakobovitz, D.; Howard, D.L.]: Magazines online. Users and uses of full text; in: Katzer, J. et al. (eds): *Managing information and technology Proceedings ASIS '89*. Washington, DC: Learned Information etc. 1989, 172-176.
- VANS 90 report** (1990): Value added networks and services. Starnberg: 1990.
- von Wietersheim, B.** (1991): Der VANS-Markt im Überblick: neue Zahlen zur Nutzung von elektronischen Informationsdiensten in der Bundesrepublik; in: Häusser 1991.

# Theorien epidemischer Prozesse in der Informationswissenschaft

Roland Wagner-Döbler

Die epidemiologische Analyse der Verteilung und Ausbreitung von Informationen, ein in Deutschland kaum rezipierter Ansatz am Schnittpunkt von Informationswissenschaft und Wissenschaftsforschung, wird erläutert. Da dieser Ansatz in seiner modernen Form keine Vorentscheidungen über Mechanismen der Ausbreitung von Informationen impliziert, eignet er sich auch als Instrument für die systematische Analyse und Modellierung wissenschaftlicher Disziplinen als Informationsströme.

The epidemiological analysis of the distribution and diffusion of information is presented, a concept at the intersection of information science and the science of science, which is almost unknown in Germany. Not implicating any predecision of the mechanisms of information transfer, this concept is appropriate to an instrument of systematic analysis and modelling of scientific areas, understood as information streams.

Es ist unumstritten, daß informationswissenschaftliche Forschung auch der Informations- und Dokumentationspraxis zugute kommen soll. Es wird aber auch kaum bestritten werden, daß **der informationswissenschaftlichen Forschung wiederum die Entwicklung theoretischer Modelle zugute kommt**. Mir scheint jedoch keineswegs, daß die Informationswissenschaft an einem Überangebot fruchtbarer Modelle leidet. Immer noch beherrschend sind, soweit ich sehe, Arbeiten, in denen bestimmte Verteilungsregularitäten modelliert werden, wobei einerseits eine Vielzahl empirischer Einzelergebnisse vorliegen, andererseits Fortschritte in der Vereinheitlichung dieser Modelle erzielt wurden. In Entwicklungen wie Co-word-analysis in der Szientometrie oder in quasi-ethnologischen Analysen des Forschungsprozesses kann ich nur wenig informationswissenschaftliche relevanten theoretischen Gehalt erkennen. In Deutschland wiederum scheint mir nicht einmal die Rezeption einschlägiger Forschungen anderer Länder besonders intensiv betrieben zu werden.

Ich möchte im folgenden einen in Deutschland wenig be-

kannten neueren theoretischen Ansatz vorstellen, der am Schnittpunkt zwischen Informationswissenschaft und Wissenschaftsforschung liegt, nämlich die **epidemiologische Analyse wissenschaftlicher Entwicklungen**.

"Epidemiologie kann als die Wissenschaft definiert werden, deren Gegenstand die Verteilung und Ausbreitungsweisen von Krankheiten in menschlichen Bevölkerungen ist."/1/ Diese Gegenstandsbeschreibung entstammt der einzigen deutschsprachigen Einführung in die Epidemiologie. Ersetzt man den Begriff "Krankheit" durch "Information", stößt man offenbar in eine zentrale Fragestellung sowohl von Informationswissenschaft als auch Wissenschaftsforschung vor, wenn zusätzlich "menschliche Bevölkerung" durch den - zugegebenermaßen etwas technischen - Ausdruck "wissenschaftliche Populationen" ausgetauscht wird. Unter "wissenschaftlichen Populationen" wollen wir durch bestimmte, beispielsweise disziplinäre Merkmale zu kennzeichnende Gruppen von Wissenschaftlern verstehen. Die Definition lautet dann: **"Epidemiologie kann als die Wissenschaft definiert werden, deren Gegenstand die Verteilung und Ausbreitungsweisen von Information in wissenschaftlichen Populationen ist."** Den Begriff "Information" kann man auf unterschiedliche Weise operationalisieren. Arbeiten Wissenschaftler beispielsweise mit einer bestimmten labortechnischen Methode, kann man annehmen, daß sie über diese Technik "informiert" sind. Dasselbe kann man annehmen, wenn sie etwa ein bestimmtes mathematisches Verfahren anwenden, wenn sie eine bestimmte experimentelle Methode verwenden oder empirische Daten einer Sozialerhebung verwenden. Da wir im epidemiologischen Zusammenhang über Psychologie des Wissenserwerbs und über die Vorgänge des Informierens keine Aussagen machen wollen, beschränken wir uns auf einfache Feststellungen wie: "Auf den Wissenschaftler X trifft das Merkmal M zum Zeitpunkt t zu. M bedeutet: Benützung der Information I durch X." Diese Charakterisierung der Information I kann auch auf andere Weise geschehen, wobei allerdings der Begriff "Benützung" nicht mehr ganz so passend erscheint, dann nämlich, wenn wir den Begriff Information in Verbindung mit wissenschaftlichen Konzepten oder Ideen bringen. Wir können beispielsweise sagen: X hat das Merkmal M, und M ist hier die Eigenschaft, von einer

bestimmten wissenschaftlichen Idee befallen zu sein oder von ihr Gebrauch zu machen. Auf jeden, der - grob gesprochen - im Rahmen einer bestimmten wissenschaftlichen Idee arbeitet, treffe dieses Merkmal zu. Wir können selbstverständlich dabei versuchen, den oder die Urheber einer wissenschaftlichen Idee, mit der sich Wissenschaftler beschäftigen, zu identifizieren.

Es gibt verschiedene Wege, die soeben diskutierten Merkmale  $M$  festzustellen. Eine Möglichkeit ist es, die Wissenschaftler zu befragen. Wir können auch versuchen, ihren Texten oder den bibliographischen Daten dieser Texte passende Merkmale zu entnehmen, nachdem die Arbeitsergebnisse von Wissenschaftlern in der Grundlagenforschung nun einmal weitgehend in der Fachliteratur ihren Niederschlag finden. Dann können wir allerdings nur mit Angaben rechnen, die im Rahmen des formalen wissenschaftlichen Kommunikationssystems als mitteilbar oder mitteilenswert gelten.

Wissenschaftsforscher verschaffen sich einen Überblick über eine Disziplin, indem sie die Merkmalszuschreibung noch gröber vornehmen. Das Merkmal besteht dann beispielsweise darin, daß ein Forscher sich mit dem mathematischen Spezialgebiet der nicht-klassischen Logik beschäftigt. Welche Ideen und Konzepte er dabei im einzelnen verwendet, bleibt zunächst offen. Hinsichtlich des damit verbundenen Informationsvorgangs bleibt lediglich die recht vage Annahme, daß dieser Forscher über das Spezialgebiet der nicht-klassischen Logik soweit informiert ist, daß er auf diesem Gebiet arbeiten kann.

Betrachtet man solche Merkmale als Verteilung über eine Gesamtheit von Forschern zum Zeitpunkt  $t$ , stellt man fest, daß sich innerhalb der Gesamtpopulation der, sagen wir, auf dem Gebiet der Mathematik Arbeitenden, sich  $x$  Prozent auf dem Gebiet der nichtklassischen Logik befinden,  $y$  Prozent auf dem Gebiet der klassischen,  $z$  Prozent auf dem Gebiet der Beweistheorie usw.

Merkmale dieser Grobheit lassen sich ohne weiteres einschlägigen Bibliographien oder auch Datenbanken entnehmen.

**Betrachtet man die Merkmale mit Hilfe geeigneter Indikatoren im Zeitablauf, kann man die Ausbreitungsweise einer Forschungsrichtung verfolgen:** Man sieht beispielsweise, daß die nichtklassische Logik bereits einmal vor der Jahrhundert-



wende ein nicht seltenes Arbeitsgebiet innerhalb der mathematischen Logik war, dann aber bis etwa 1925 fast keinerlei Beachtung mehr fand, gefolgt von einem phänomenalen Aufstieg in weniger als 15 Jahren. Die Mengenlehre hingegen erlebte seit ihrem explosionsartigen Aufschwung mit einem Höhepunkt um 1910 einen stetigen Rückgang auf einen eher den anderen logischen Arbeitsgebieten entsprechenden Anteil an der mathematischen Logik.

Interessiert man sich mehr für die Entwicklung eines einzelnen Gebiets, wird man beispielsweise auf Trendänderungen der Anzahl der Forscher dieses Gebiets achten. Zu diesem Zweck könnte man die jährliche Änderung der Anzahl der Forscher ermitteln.

Ich habe nun implizit bereits einige wichtige Grundbegriffe der Epidemiologie eingeführt: **Population**, **Prävalenzrate**: den Anteil der von einer "Infektion" Befallenen innerhalb einer Gesamtpopulation zum Zeitpunkt  $t$ , und **Inzidenz**: die Fälle von "Neuansteckungen" innerhalb eines bestimmten Zeitintervalls.

Die oben beschriebenen Beobachtungen über die Entwicklung von Forscher-Populationen kann man auch ohne das Instrumentarium der Epidemiologie machen. Und in der Tat kennen wir eine Vielzahl ohne diesen Hintergrund zustandegekommener Beobachtungen. Welche zusätzlichen Gesichtspunkte verspricht hier die Epidemiologie? Sie macht es sich zur Aufgabe, Eigenschaften von Forscherpopulationen in ein quantitatives Modell zu fassen. Erst dann kann man versuchen, etwaige Abhängigkeiten zwischen Teilepidemien zu ermitteln, zeitliche Verlaufsmuster zu entdecken, Schwellenwerte zu berechnen und mit Prognosen und Ex-post-ante-Prognosen zu experimentieren. Ohne Zweifel können epidemiologische und nicht-epidemiologische Untersuchungen ineinander übergehen. Aber epidemiologische Forscher versuchen häufiger, von statistischen Zustandserhebungen zu Modellbildungen vorzustoßen, welche die Herausarbeitung **grundlegender und allgemeiner Prinzipien** zum Ziel haben.

Vor allem im englischsprachigen Raum wurden die mathematischen Grundlagen epidemiologischer Modelle ausgebaut. Aber schon mit Hilfe einfacher mathematischer Verfahren und einfacher Grundbegriffe lassen sich wissenschaftliche Entwick-

lungen epidemiologisch in grober Form modellieren. Diese Modelle gestatten dann auch die Simulation in Rechnern - die einzige, quasi-experimentelle Methode, die bleibt, wenn man mit seinem Untersuchungsobjekt keine Realexperimente durchführen kann. Dies trifft auf die wissenschaftliche "Evolution" kaum weniger zu als auf viele epidemiologische Problemlagen in der Medizin und wird dann besonders wichtig, wenn man nicht nur einen einzigen epidemischen Prozeß verfolgt, sondern das Zusammenspiel mehrerer Prozesse, beispielsweise wissenschaftlicher Teilgebiete, untersuchen will. Oftmals sind in der Tradition der Pionierarbeiten von D. Price Wachstumsprozesse einzelner wissenschaftlicher Disziplinen isoliert dargestellt worden; Aufgabe einer epidemiologischen Analyse wäre es, die **Entwicklung wissenschaftlicher Teilgebiete in ihrem wechselseitigen Zusammenhang** quantitativ zu untersuchen.

William Goffman, der amerikanische Pionier der Anwendung epidemiologischer Methoden in der Informationswissenschaft, hat sich stark von der Idee der Ansteckung leiten lassen, die ja schließlich auch die Wurzel der medizinischen Erforschung von Epidemien war./2/ Die moderne Epidemiologie umfaßt jedoch auch die statistische Beschreibung der Verteilung von solchen Krankheiten auf Bevölkerungsgruppen, die nicht auf Ansteckung beruhen, sondern auf andere Weise verursacht werden. Zu dieser Verbreiterung der Basis der Epidemiologie hat sicherlich auch der Zwang beigetragen, "krankmachende" Faktoren in einer zunehmend belastenden Umwelt aufzuspüren./3/ Mit dieser Feststellung soll einem ersten Einwand begegnet werden, der lauten könnte, daß die Verbreitung wissenschaftlicher Ideen nicht als "Ansteckung" betrachtet werden könne. Wissenschaftliche Ideen verbreiteten sich durch Überzeugung anderer Wissenschaftler, sie verschwänden, weil sie zu keinen fruchtbaren Theorien oder erfolgversprechenden Anwendungen führten; kurzum, es handele sich um Ergebnisse einer argumentativen Auseinandersetzung, die sich nicht mit dem mechanischen Vorgang einer Ansteckung beschreiben ließen. Diejenigen Wissenschaftstheoretiker, die dazu tendieren, den Vorgang der Verbreitung wissenschaftlicher Ideen mit den Augen eines Massenpsychologen zu sehen, würden möglicherweise gerade an dieser "mechanistischen" Sichtweise Gefallen finden./4/

Eine epidemiologische Analyse darf jedoch meiner Ansicht weder den ersten noch den zweiten Standpunkt voraussetzen. **Apriorische Aussagen über den Mechanismus der Übertragung von Informationen oder wissenschaftlichen Ideen sind nicht gerechtfertigt**, gleich ob es sich um psychologische oder um wissenschaftstheoretische oder -soziologische Thesen handelt. Dies ist der ganze "mechanistische" Kern des epidemiologischen Ansatzes. In epidemiologischer Betrachtung unterscheidet sich die Durchsetzung einer rationalen Überzeugung zunächst einmal nicht vom Aufkommen wissenschaftlicher Moden, deren quantitative Untersuchung übrigens der Botaniker J. F. Cole schon 1916 empfahl./5/ Erst eine detaillierte Beschreibung des Befundes kann Hinweise auf kausale Strukturen und wissenschaftstheoretische Deutungsmöglichkeiten geben.

Die exakte Beschreibung von Verteilung und Ausbreitung von Krankheiten wird **deskriptive** Epidemiologie genannt, im Unterschied zur **explanatorischen** (auch "analytisch" genannt). Soweit ich sehe, stehen Informationswissenschaft und Wissenschaftsforschung schon auf dem Feld der deskriptiven Epidemiologie die Türen noch weit offen, umso mehr, wenn es um explanatorische Epidemiologie geht.

Daß empirisch erhaltbare, informationswissenschaftlich-epidemiologische Befunde **wissenschaftstheoretische Deutungsversuche in Frage stellen können**, sei an einem Beispiel demonstriert.

Behauptet ein Wissenschaftstheoretiker, eine Theorie würde fallengelassen, wenn sie falsifiziert worden sei, ein anderer, wenn die "Gründungsväter" der Theorie ausstürben, ein dritter, wenn der "Problemvorrat" erschöpft sei /6/, sind Merkmale festzulegen, die Wissenschaftler identifizierbar machen, die sich mit einer bestimmten Theorie beschäftigen. Stellt sich heraus, daß die Beschäftigung mit der Theorie nicht abnimmt oder gar aufhört, obwohl sich (mindestens) ein Ereignis ausfindig machen läßt, das "Falsifikation" genannt werden kann, muß die erste These falsch sein. Stellt er fest, daß eine Theorie ihre Gründer überlebt und von folgenden Wissenschaftlergenerationen in unverminderter Intensität weitergeführt wird, muß die zweite These falsch sein. Stellt er fest, daß die Zahl der Probleme, vor die eine Theorie stellt oder die

sie aufstellt, wächst, die Zahl der Wissenschaftler jedoch, die sich damit beschäftigen, sinkt, ist auch die dritte These falsch.

Bei der Modellierung und Quantifizierung dieser Vorgänge kann der Epidemiologe auf zahlreiche Methoden und Verfahrensweisen der **Bibliometrie und Szientometrie** zurückgreifen, insbesondere, wenn er sich im Bereich der sogenannten deskriptiven Epidemiologie befindet. Er kann versuchen, das disziplinäre Geschehen im Zeitablauf auf verschiedenen Aggregationsebenen epidemiologisch zu beschreiben: Beispielsweise die Prävalenz der mathematischen Logik innerhalb der Mathematik; die Prävalenzen der Teilgebiete der mathematischen Logik; und innerhalb dieser Teilgebiete wiederum einzelne Arbeitsgebiete, die Maria Nowakowska "Mikroparadigmen" nennt.

Informationswissenschaftler dürften besonders am Aspekt des Informationstransfers interessiert sein - ein in Forschungs- und Technologiepolitik vielgehörtes Schlagwort. Darunter wird meistens der Vorgang verstanden, daß wissenschaftliche Informationen technologisch nutzbar gemacht werden; dies kann beispielsweise durch persönliche Kontakte zwischen Wissenschaftlern und Ingenieuren geschehen. Es kann aber auch im Rückgriff auf das formale Fachkommunikationssystem geschehen: Eine Informationsvermittlungsstelle kann versuchen, in der Masse einschlägiger Fachaufsätze Arbeiten zu finden, die beispielsweise eine Wissenslücke über das Belastungsverhalten eines neuartigen Materials zu schließen helfen.

Insgesamt ist darüber, wie wissenschaftliche Informationen in anwendungsnahe Bereiche diffundieren, wenig bekannt. Technologietransfereinrichtungen müssen sich hier wohl mehr auf Faustregeln verlassen als sich auf fundiertes Wissen stützen zu können. Es liegt nahe, den Prozeß einmal mit **epidemiologischen Methoden** nachzuzeichnen und damit informationswissenschaftliche Aussagen zu einer Diskussion beizusteuern, die von der Behauptung reicht, wissenschaftliche Erkenntnisse seien für die technologische Entwicklung weitgehend unbedeutend und unerheblich /7/, bis zur Behauptung, **jede** technologische Entwicklung ruhe heute auf wissenschaftlichen Fundamenten /8/. Beide Behauptungen können gewichtige Argumente

anführen; aber nur systematische empirische Informationen können hier wirklich weiterhelfen.

Man könnte diese "Wissenschaftsbindung" technologischer Entwicklungen im Spiegel der Fachkommunikation analysieren. Will man beispielsweise feststellen, wie häufig Erkenntnisse der Mathematik in anderen Wissenschaftsbereichen verwendet werden, läßt sich die Zitierung mathematischer Arbeiten als grober quantitativer Indikator verwenden; man erhält ein Bild davon, in welchem Ausmaß Erkenntnisse, die an den Forschungsfronten der Mathematik gewonnen werden, für andere Wissenschaften von Interesse sind. Die Einschränkung auf Forschungsfront ist nötig, weil die Quellen älterer und allgemein bekannter mathematischer Anwendungen nur noch selten ausdrücklich angegeben werden./9/ Verläßt man bei der Beschreibung dieses Prozesses die deskriptive Epidemiologie und geht von bestimmten Übertragungstheorien aus, hat man es einer **speziellen Diffusionstheorie zu tun, mit der die Ausbreitung wissenschaftlicher Erkenntnisse erklärt werden soll**. Diffusionstheorien werden häufiger mit der Ausbreitung technischer Neuerungen als mit der Ausbreitung wissenschaftlicher Erkenntnisse in Verbindung gebracht. Aber gerade dieser zuletzt genannte Vorgang fällt auch in das Arbeitsgebiet von Informationswissenschaftlern. Es gibt - freilich noch zaghafte - amerikanische Ansätze, die Goffmansche epidemische Theorie auf diesem Feld fruchtbar zu machen./10/ Hier sind nicht nur theoretische Überlegungen weiterzuführen, sondern auch szientometrische und bibliometrische Methoden in stärkerem Maße nutzbar zu machen. Um nur ein Beispiel zu nennen: Bei der epidemiologischen Untersuchung der Verbreitungskanäle wissenschaftlicher Erkenntnisse sollte man versuchen, den "impact factor" eines speziellen Kanals zu gewichten - analog den Methoden, die unter Zuhilfenahme des Science Citation Index entwickelt wurden.

Es gibt noch weitere Wege, Erkenntnisse über Informationstransfer zu gewinnen. Eine der möglichen Methoden, der wiederum Goffman viele wichtige Impulse gegeben hat, ist die Analyse der sogenannten Mobilität von Forschern. Anders als in der Soziologie wird in diesem Zusammenhang darunter nicht die Bewegung im geographischen, sondern die Bewegung - wenn man so will - im kognitiven Raum der Wissenschaften verstanden. Wenn

ein Forscher X sein Forschungsgebiet  $G_1$  verläßt und sich einem anderen Forschungsgebiet  $G_2$  zuwendet, faßt man dies unter dem Begriff der **Mobilität** zusammen. Möchte man eher den Informationsaspekt betonen, kann man von folgender Vermutung ausgehen: Der Wechsel von Forschungsgebieten ist nicht beliebig. Er wird einerseits von einer Art fachlichen Nähe abhängen, andererseits von der Möglichkeit, auf dem alten Betätigungsgebiet gewonnene Erkenntnisse im neuen Bereich anzuwenden. In diesem Fall liegt ein klassischer Fall von personalem Informationstransfer vor, der ja letztlich Grundlage eines jeden Bildungssystems ist. Es gibt eine große Anzahl soziologischer Arbeiten über die institutionelle und geographische Mobilität von Forschern, jedoch nur sehr wenige Arbeiten, die sich mit dieser **kognitiven** Mobilität befassen. Goffman hat hier die Anwendbarkeit der mathematischen Theorie finiter Markov-Ketten demonstriert. Hierbei geht es kurz gesagt darum, die Häufigkeiten aller logisch möglichen Übergänge innerhalb eines disziplinären Systems quantitativ zu beschreiben. Das Resultat sind - gegebenenfalls unter Hinzuziehung weiterer Kennziffern dieses Systems - stochastische Aussagen beispielsweise über die Rückkehrwahrscheinlichkeit von Forschern in ihr ehemaliges Arbeitsgebiet.

Mobilitätsanalysen spielen auch bei der epidemiologischen Analyse eine wichtige Rolle: Bemerkte man den Aufstieg einer Forschungsrichtung, ist herauszufinden, woher sich die Forscher rekrutieren, die zu diesem Aufschwung beitragen; Abwanderung enttäuschter Wissenschaftler aus "absteigenden" Disziplinen ist ebenso möglich wie die Zuwanderung von Neulingen, ein Vorgang, den man gern mit dem Vorgang der "Schulenburg" in Verbindung bringt.

Die Goffmanschen Arbeiten haben eine Reihe von Arbeiten angeregt, in denen versucht wird, andere als die von ihm behandelte biologischen und mathematischen Disziplinen einer epidemiologischen Analyse zu unterziehen. Meistens handelt es sich um einzelne kleine wissenschaftliche Teilgebiete. So sehr dies auch das Arsenal einzelner empirischer Beobachtungen vergrößert, einer epidemiologischen Modellbildung ist damit nicht in selbem Maße gedient. Auf dieser Seite geschieht wohl auch deshalb weniger, weil dies mit aufwendigen Datenerfas-

sungs- und Auswertungsarbeiten verbunden ist; Online- oder CD-ROM-Datenbanken lassen sich für diese Zwecke unter anderem wegen ihrer Kosten und wegen derzeit fehlender Eignung der Retrieval-Software nur bedingt gebrauchen.

Dabei hat Goffman, anders als viele Wissenschaftsforscher, das Risiko auf sich genommen, aus seinen epidemiologischen Analysen **nachprüfbare Prognosen** abzuleiten. Gerade diese empirische Nachprüfbarkeit schien uns ein großes Verdienst seines theoretischen Beitrags darzustellen. Wir - d.h. Professor Jan Berg vom Institut für Philosophie der Technischen Universität München und ich - haben uns vorgenommen, die von Goffman durchgeführten Prognosen so weit wie möglich nachzuprüfen sowie mit eigenen Prognosen oder Ex-post-ante-Prognosen zu konfrontieren. Selbstverständlich wäre nicht nur ein positives, sondern auch ein negatives Resultat von Nutzen. Stellte sich heraus, daß mit den vorhandenen epidemiologischen Modellen keine Prognosen möglich sind, ergäben sich Hinweise auf ihre Unvollständigkeit oder Hinweise auf prinzipielle Grenzen der Prognostizierbarkeit wissenschaftlicher Entwicklungen - oder auch von "Informationsströmen".

Ich habe bisher vor allem **Argumente für die Anwendbarkeit der Epidemiologie auf Gegenstände der Informationswissenschaft und Wissenschaftsforschung** anführen wollen und auf die detailliertere Darstellung des mathematischen und begrifflichen Apparats ebenso verzichtet wie auf die Darstellung der zahlreichen empirischen Befunde, die Goffman, seine Mitarbeiter und andere erarbeitet haben, sowie erster eigener Ergebnisse auf dem Gebiet der Mathematik. Ebensowenig konnte ich an dieser Stelle ausführlicher auf Erhebungsverfahren und die Bildung geeigneter - vor allem bibliometrischer - Indikatoren eingehen.

Diesen Lücken zum Trotz hoffe ich, Ihnen meine Sicht des möglichen Standorts epidemiologischer Theorien in Informationswissenschaft und Wissenschaftsforschung angedeutet zu haben. Ich würde mich allerdings weniger freuen, wenn Sie von diesen Ideen auf "mechanische" Weise angesteckt würden, als wenn Sie sie der informationswissenschaftlichen Diskussion für würdig hielten.

## Literatur

1. Rainer Frentzel-Beyme: Einführung in die Epidemiologie, Darmstadt 1985, S. 1.
2. Siehe zusammenfassend v.a.: William Goffman; Kenneth S. Warren: Scientific information systems and the principle of selectivity, New York 1980. Goffman: A mathematical method for analyzing the growth of a scientific discipline, in: Journal of the Association for Computing Machinery, Vol. 18, 1971, S. 173-185.
3. Einen Einstieg auch in die moderne Literatur bietet Eugene Garfield: Mapping the world of epidemiology. P. 1. The disease detectives, in: Current Contents, 1988, No. 35, S. 3-11. Ein guter Kurzüberblick über den epidemiologischen Ansatz in der Informationswissenschaft findet sich in Garfield: The epidemiology of knowledge and the spread of scientific information, in: Current Contents, 1980, No. 35, S. 5-10. (Beides wiederabgedruckt in Garfield: Essays of an information scientist.)
4. Manfred Kochen nannte Goffmans Ansatz einmal eine "mechanische" Sichtweise - siehe: Kochen: Models of scientific output, in: Toward a metric of science, ed.: Yekuda Elkana, New York 1978, S. 97-136.
5. F. J. Cole; Nellie B. Eales: The history of comparative anatomy. P. 1. A statistical analysis of the literature, in: Science Progress, Vol. 11, 1917, S. 578-596.
6. Dies letztere vermutet Maria Nowakowska: Epidemical spread of scientific objects: An attempt of empirical approach to some problems of meta-science, in: Theory and Decision, Vol. 3, 1973, S. 262-297. Siehe auch Nowakowska: Microparadigms, simulations, prognoses. Further remarks on the theory of epidemical spread of scientific objects, in: Problems of the Science of Science, 1972/3, S. 97-112.
7. David Collingridge: Incremental decision making in technological innovation: What role for science?, in: Science, Technology & Human Values, Vol. 14, 1989, S. 141-162.
8. Siehe z.B. Hendrik Casimir: Haphazard reality. Half a century of science, New York 1983, S. 297 u. passim.
9. In Westdeutschland beschäftigt sich eine Arbeitsgruppe um Hariolf Grupp, Thomas Reiss und Ulrich Schmoch am Fraunhofer-Institut für Systemtechnik und Innovationsforschung in Karlsruhe mit solchen Fragen vor allem aus der Perspektive von Forschungs- und Technologiepolitik. Unter anderem wird am einen Indikator für die Wissenschaftsnähe von Technologien gearbeitet, der auf der Zitierung wissenschaftlicher Literatur in Patentedokumenten verschiedener Patentklassen beruht.



10. Siehe z.B.: George A. Barnett; Edward L. Fink; Mary B. Debus: A mathematical model of academic citation age, in: Communication Research, Vol. 16, 1989, S. 510-531. Yves F. LeCoadic: Modelling the communication, distribution, transmission or transfer of scientific information, in: Journal of Information Science, Vol. 13, 1987, S. 143-148.

Dr. Roland Wagner-Döbler  
Institut für Philosophie  
Technische Universität München  
Lothstr. 17  
8000 München 2

# **Erfolgsfaktorenforschung in der Informationsvermittlung: Konzepte, Modelle, Ergebnisse.**

Josef Herget

## **Inhalt**

1. Einführung
2. Erfolgsfaktorenforschung
3. Erfolgsfaktorenforschung in der Informationsvermittlung
  - 3.1. Fragestellungen
  - 3.2. Modelle
  - 3.3. Untersuchungsdesign
4. Ergebnisse
  - 4.1. Methodik der Untersuchung
  - 4.2. Strukturdaten
  - 4.3. Erfolgsfaktoren
5. Ausblick

## *Zusammenfassung*

In diesem Beitrag wird das Konzept der Erfolgsfaktorenforschung für das Gebiet der Informationsvermittlung entwickelt und fruchtbar gemacht.

Bezugnehmend auf ein entwickeltes Informationsvermittlungs-Modell (bestehend aus den Objektbereichen Anbieter, Nachfrager und Rahmenbedingungen) wird dann eine Forschungsmethodik zur Erfassung von Erfolgsfaktoren in der Informationsvermittlung vorgestellt.

Im einzelnen werden methodologische Grundüberlegungen und das Untersuchungskonzept (bestehend aus "breiteren" empirischen Ansätzen, Einzelfallstudien in ausgewählten Informationsvermittlungsstellen (IVS) und experimentellen Untersuchungen zur Kausalanalyse) dargestellt.

Ergebnisse einer schriftlichen Befragung von 1200 Leitern von Informationsvermittlungsstellen zur Beurteilung der Wichtigkeit von Erfolgsfaktoren schließen die Ausführungen ab.

## *Abstract*

The nature and usefulness of success factor research in general, and in relation to the information and library area in particular, is explained. It is suggested that an understanding of the factors of importance to success in the Information Resource Unit environment can be beneficial to planning and management. Results are presented of a survey which generated some evidence of success factors generic to the information and library environment.

## 1. Einführung

Zahlreiche empirische Studien zur Nutzungssituation externer Informationsressourcen in der betrieblichen Praxis konstatieren eine defizitäre und ungenügende Inanspruchnahme von am Informationsmarkt und insbesondere am Online-Markt angebotenen Informationsdiensten und -produkten (vgl. z.B. Ahrend 1990, Herget 1990a, Krups 1985, Reinhard 1987). Neben der Beschreibung existenter Hemmnisse und Barrieren der Informationsnutzung (vgl. hierzu Herget 1990b) ist es jedoch von besonderem Interesse, diejenigen Faktoren zu identifizieren, die - positiv gesprochen - erfolgreiche Transferprozesse zwischen Informationsanbietern und Informationsnachfragern ermöglichen und fördern.

In der informationswissenschaftlichen Forschung ist diesen Interaktionen - im Unterschied zur Forschung auf dem Gebiet der direkten Interaktion in Information-Retrieval-Prozessen zwischen Informationssystemen und Rechercheuren im Rahmen der Mensch-Computer-Interaktion (Belkin/Vickery 1985, Efthimiadis/Robertson 1989, Gloeckner-Rist/Lehmler/Wettler 1989) - bislang nur wenig Aufmerksamkeit geschenkt worden. Ihre Betrachtung erstreckte sich allenfalls auf den Kontext des Marketings (vgl. z.B. Müller/Winand 1988) und der Benutzerforschung (vgl. z.B. Wind 1980).<sup>1</sup>

Im Gegensatz hierzu beschäftigt sich die Managementliteratur seit einigen Jahren intensiv mit der Frage, welche Faktoren das erfolgreiche Überleben von Unternehmen im Markt gewährleisten und fördern. Diese Forschungsrichtung wird als Erfolgsfaktorenforschung bezeichnet. Ihr Ziel ist es, vor allem solche Faktoren zu identifizieren, die gestaltbar beeinflusst werden können.

Die folgenden Ausführungen übertragen und wenden das Konzept der Erfolgsfaktorenforschung auf das Gebiet der Informationsvermittlung an.

## 2. Erfolgsfaktorenforschung

Die Erfolgsfaktorenforschung erfreut sich einer guten Konjunktur. Zahlreiche, insbesondere aus betriebswirtschaftlich-/management-orientierter Sicht motivierte Untersuchungen, widmen sich der Erforschung derjenigen Faktoren, die den Erfolg einer Person, einer Abteilung, eines Unternehmens, von Branchen oder gar ganzen Volkswirtschaften bedingen. Als einige neuere Beispiele der in- und ausländischen Literatur sollen Diegruber (1991), Fritz (1990), Gemünden (1990), Geschka (1989), Kaufmann (1990), Knoll (1989), Kuipers (1990), Lehner (1990), Magal/-Carr/Watson (1988), Mair (1989), Nagel (1988), Patt (1990), Peters/Waterman

---

1 Anfang Mai 1991 wurden Online-Recherchen in verschiedenen informationswissenschaftlichen (INFODATA, LISA) und wirtschaftswissenschaftlichen bibliographischen Datenbanken (ABI/INFORM, BLISS, ECONOMIC LITERATURE INDEX) durchgeführt, die für das Thema "Erfolgsfaktoren und Informationsvermittlung" nur zu einem (jedoch nicht relevanten) Treffer führten. Es besteht also offensichtlich Forschungsbedarf.

(1984), Poth (1989), Porter (1990), Trommsdorff (1990), Unterkofler (1989) und Wohlgemuth (1989) genannt werden.

Was sind Erfolgsfaktoren?

In der Literatur findet sich eine Vielzahl von unterschiedlichen Begriffen für "Erfolgsfaktoren" (z.B. Erfolgspotentiale, strategische Erfolgspotentiale, kritische (Erfolgs-)Faktoren, strategische (Erfolgs-)Faktoren, strategische Faktoren, strategische Schlüsselfaktoren, Erfolgsdeterminanten, Einfluß- und Bestimmungsgrößen des Erfolgs, Unternehmensgrundsätze, Erfolgs- oder Führungsprinzipien, unternehmerische Grundtugenden, Handlungsmaximen, Erfolgsregeln etc.; siehe Mair (1989, S. 15), Patt (1990, S. 6)). Diese Begriffe deutlich voneinander abzugrenzen, ist schwierig (Mair 1989, S. 15), jedoch auch nicht notwendig, da sie zumeist undefiniert und als Synonyma verwendet werden.

Die Ursprünge des Begriffs Erfolgsfaktor (success factor) stammen bereits aus den 60er Jahren und wurden zum ersten Mal von Daniel (1961) für den Bereich Manager-Information angewandt. Erneut aufgegriffen wurde dieses Konzept<sup>2</sup> Ende der 70er Jahre von Rockart (1979), der es um das Adjektiv kritisch (critical) erweiterte. Danach lassen sich drei bis sechs branchenspezifische Erfolgsfaktoren identifizieren, deren "Funktionieren" für den Unternehmenserfolg entscheidend ist. In jüngster Zeit wurde das Konzept über die Begrenzung auf branchenbezogene Faktoren hinaus auch auf organisationsspezifische, funktionsspezifische und individuelle Faktoren ausgeweitet.<sup>3</sup>

*Als Erfolgsfaktoren werden im folgenden endogene oder exogene Einflußgrößen bezeichnet, deren Performanz für den Erfolg der betrachteten Systeme oder ihrer Elemente bestimmend ist.*<sup>4</sup>

### **3. Erfolgsfaktorenforschung in der Informationsvermittlung**

#### **3.1. Fragestellungen**

Eine Analyse von Erfolgsfaktoren innerhalb des komplexen Gebildes der Informationsvermittlung erfordert zunächst eine Identifikation und Systematisierung unterschiedlicher Vermittlungssituationen. Eine produktive Einteilung ist - trotz der allen Dichotomien inhärenten Unzulänglichkeiten - zunächst eine Unterscheidung der Informationsanbieter in direkte (d. h. der Informationsnutzer kommuniziert direkt mit Informationssystemen) und indirekte (d. h. ein Informationsvermittler

---

2 siehe zu den Anwendungsbereichen z.B. Leidecker/Bruno (1984, S. 23) und Hoffmann (1986, S. 832 f.)

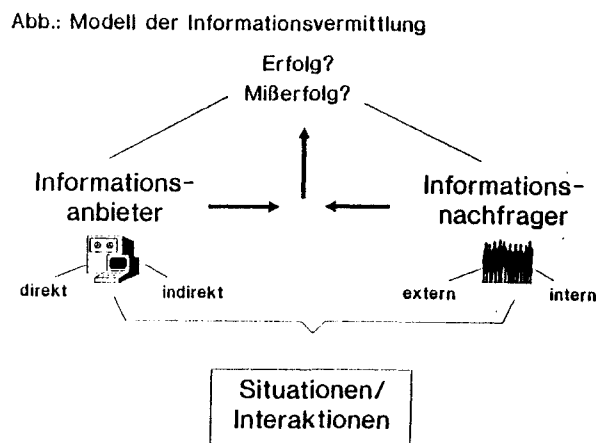
3 Zum Literaturstand der identifizierten Erfolgsfaktoren vgl. Herget/Hensler (1991a).

4 Unter Erfolg wird ein intendiert positives Ergebnis von gestaltenden Maßnahmen auf ein betrachtetes System verstanden.

liefert dem Informationsnutzer das nachgefragte Informationsprodukt). Die Informationsnutzer können nach dem Kriterium des systemischen Bezugs zum Informationsanbieter klassifiziert werden. Interne Informationsnachfrager befinden sich in der gleichen Organisation wie der Informationsanbieter (z. B. innerbetriebliche Informationsvermittlung); externe Informationsnachfrager stellen für den Informationsanbieter exogene Marktpartner dar. Diese verschiedenen Ausprägungen können in der Realität auch kombiniert auftreten. Die dargestellten Anbieter-Nachfrager-Relationen erfordern zur Vollziehung von Transaktionen verschiedener Interaktionsprozesse.

Die beschriebenen Beziehungen zwischen Informationsanbietern und -nachfragern - modellhaft dargestellt in der folgenden Abbildung - sollen den Ausgangspunkt für die Ermittlung von Erfolgsfaktoren der Informationsvermittlung darstellen.

Abb. 1 : Modell der Informationsvermittlung



## Beziehungen

- a) direkter Informationsanbieter - interner Informationsnachfrager  
(Beispiel: Nutzung innerbetrieblicher Datenbanken)
- b) direkter Informationsanbieter - externer Informationsnachfrager  
(Beispiel: Online-Nutzung bei Hosts durch Unternehmen mit eigenem Online-Anschluß)
- c) indirekter Informationsanbieter - interner Informationsnachfrager  
(Beispiel: innerbetriebliche Informationsvermittlung)
- d) indirekter Informationsanbieter - externer Informationsnachfrager  
(Beispiel: Informationsbroker und Unternehmen).

Ausgehend von diesem Modell sollen die drei Objektbereiche Anbieter, Nachfrager und Interaktion (Situationen der Rahmenbedingungen) analysiert werden.

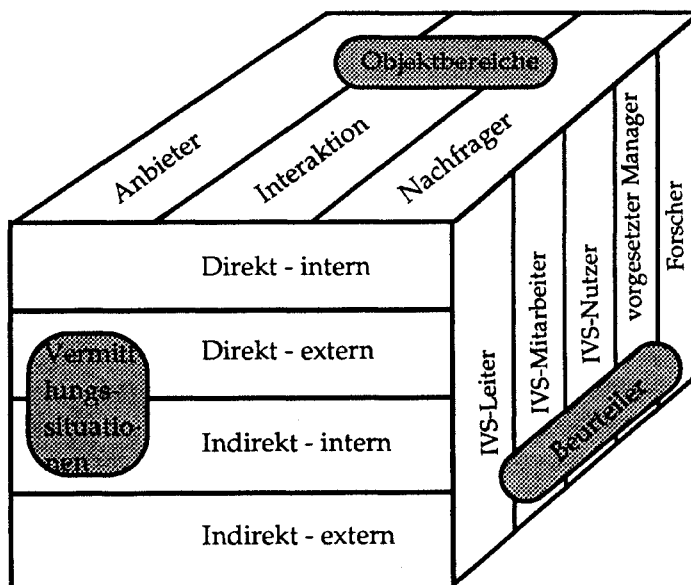
Die forschungsleitenden Fragestellungen sind insbesondere:

1. Wie lassen sich Erfolg bzw. Mißerfolg "der" Informationsvermittlung definieren und messen?
2. Was zeichnet "erfolgreiche" Informationsanbieter aus?
3. Was zeichnet "erfolgreiche" Informationsnachfrager aus?
4. Welche Interaktionen ermöglichen eine "erfolgreiche" Informationsvermittlung?

### 3.2. Modelle

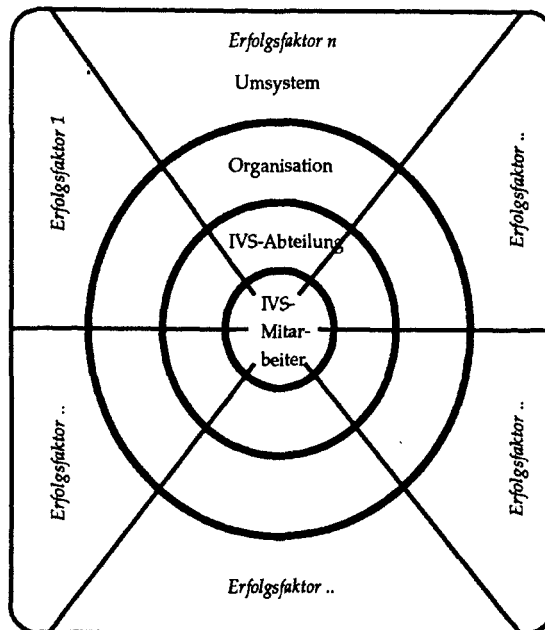
Innerhalb dieses Forschungsprojektes kann die Erforschung der Erfolgsfaktoren nach Vermittlungssituation, Objektbereich und den Beurteilenden unterschieden werden. Die möglichen Kombinationen sind in der nachfolgenden Abbildung dargestellt:

Abb. 2: Typologie der Untersuchungssituationen



Bei der Analyse von Erfolgsfaktoren sind ebenso mehrere Betrachtungsebenen möglich. So können Erfolgsfaktoren auf der Ebene IVS-Mitarbeiter, IVS-Abteilung, Organisation oder im Umsystem angesiedelt sein. Von dieser Einbettung hängt dann auch der Handlungsspielraum zur Beeinflussung der identifizierten Erfolgsfaktoren ab. Dieser Zusammenhang ist in der nachfolgenden Abbildung wiedergegeben:

Abb 3: Meta-Modell zu Erfolgsfaktoren der Informationsvermittlung



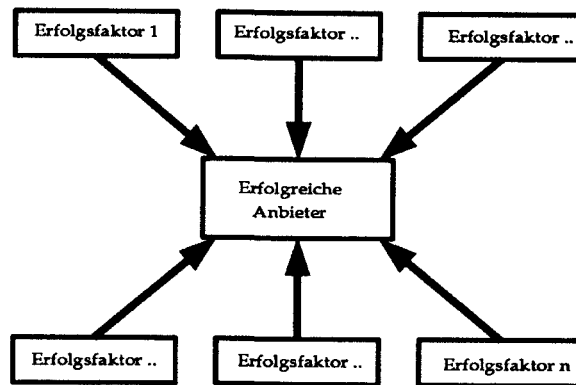
Für die drei Objektbereiche der Informationsvermittlung - Anbieter, Nachfrager und Interaktion - lassen sich unter Berücksichtigung der Vermittlungssituation zunächst jeweils eigene Modelle entwickeln.<sup>5</sup>

Am Beispiel der Anbieter läßt sich die Struktur solcher Modelle wie folgt darstellen:<sup>6</sup>

5 In einem späteren Stadium sollen diese zu *einem* Modell integriert werden.

6 Zur konkreten Ausgestaltung eines solchen Anbietermodells siehe das Beispiel in Herget (1991b).

Abb. 4: Modell: Erfolgsfaktoren der Anbieter



Die obige Modelldarstellung - die vor allem einer schärferen Konturierung dient - unterstellt prima vista einen monokausalen Zusammenhang, demzufolge es genügte, nur "an einem Erfolgsfaktor zu drehen", dann würde sich schon die erwünschte positive Wirkung einstellen. So einfach stellt sich die Realität aber nicht dar. Die einzelnen Erfolgsfaktoren stehen in mannigfaltigen interdependenten und vernetzten Wechselbeziehungen zueinander und können sich wechselseitig verstärken, aber auch vermindern. Darüberhinaus ist aber auch keine Linearität dieser gegenseitigen Einflußbeziehungen gegeben, so daß beispielsweise eine sich zunächst verstärkende Beziehung ab einem bestimmten Ausprägungsgrad in der Wirkung sogar "umkippen" kann. Ebenso ist der Wirkungsmechanismus zeitlich variabel. Die praktische Folge ist, daß manche "Regelung" eines Faktors sofort (kurzfristig), oder mit gewisser Verzögerung (mittel- bis langfristig) wirkt. Mit dieser Betrachtung schließen wir uns der systemorientierten Sichtweise an, wie sie beispielsweise von Gomez/Probst (1988), Ulrich/Probst (1990) und Vester (1980) vertreten wird.<sup>7</sup>

### 3.3. Untersuchungsdesign

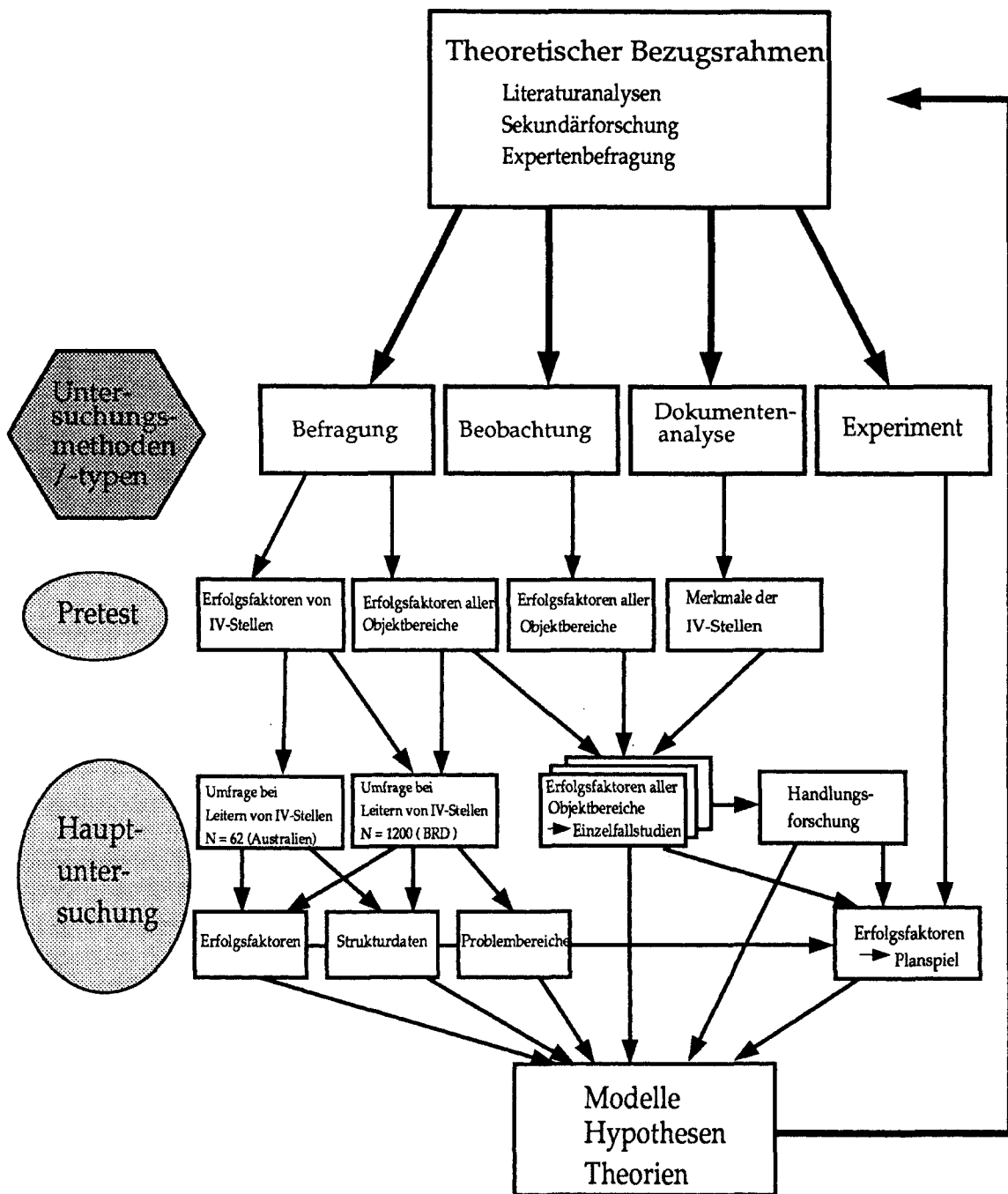
Der mehrteiligen Hauptuntersuchung<sup>8</sup> sind eine explorative und eine Pretest-Phase vorgeschaltet. Die verschiedenen Phasen sind in der Abbildung im Gesamtüberblick dargestellt:

<sup>7</sup> Vergleiche hierzu ausführlich Herget (1991a).

<sup>8</sup> Zur ausführlichen Darstellung des Untersuchungskonzeptes siehe Herget/Hensler (1991b). Dieses ambitionöse Forschungsprogramm ist auf mehrere Jahre angelegt und bedarf zur vollständigen Durchführung umfassende personelle und finanzielle Ressourcen, die zum jetzigen Zeitpunkt nicht gesichert sind.



Abb. 5: Untersuchungsdesign



Im folgenden werden einige Ergebnisse einer schriftlichen Befragung von Leitern von IV-Stellen zur Bedeutung von Erfolgsfaktoren in ihrem Bereich vorgestellt.

## 4. Ergebnisse

### 4.1. Methodik der Untersuchung

#### Zur Erhebung

Es wurde ein siebenseitiger halbstandardisierter Fragebogen mit integriertem Anschreiben entwickelt, der sich an die Leiter von IV-Stellen richtete. Die Antworten konnten - falls gewünscht - anonym erfolgen.

Der Fragebogen setzte sich aus vier Teilen zusammen:

- Erfolgsfaktoren von IV-Stellen
- Problembereiche von IV-Stellen
- Wirtschaftlichkeitsnachweis (nur unselbständige IV-Stellen)
- Strukturdaten (Angaben zur Person des Leiters und zur IV-Stelle).

Der Fragebogen wurde Ende Juni 1991 an 1200 rechtlich selbständige und unselbständige IV-Stellen in der Bundesrepublik Deutschland (einschließlich der neuen Bundesländer) verschickt.<sup>9</sup>

Insgesamt wurden 411 Fragebogen zurückgesandt. Dies entspricht einer Rücklaufquote von 34.25 %. Von diesen 411 Fragebogen waren 386 auswertbar. Der Rücklauf kann angesichts des Umfangs des Fragebogens und der teilweise nur unzulänglichen Qualität des verfügbaren Adreßmaterials als sehr zufriedenstellend bezeichnet werden.

#### Auswertung

Die Auswertung des umfangreichen Datenbestandes (über 60000 Datenelemente) erfolgt mit dem PC-Statistikprogrammpaket CSS. Als statistische Verfahren sollen dabei hauptsächlich univariate Statistik (Berechnung von diversen Mittelwerten und Streuungsmaßen), bivariate Statistik (z.B. Korrelationsrechnung und Erstellen von Kreuztabellen) sowie auch multivariate Statistik (z.B. Faktoren-, Cluster- und Varianzanalyse) zur Anwendung kommen.

---

9 Bei den durchzuführenden Einzelfallanalysen werden die verschiedenen Vermittlungssituationen erfaßt und jeweils mehrere IVS (multiple-case-design) untersucht. Bei diesem Forschungsansatz wird es möglich, sämtliche Objektbereiche der Informationsvermittlung in ihrer Ganzheitlichkeit zu erforschen.

## 4.2. Ausgewählte Strukturdaten

1) Von den antwortenden IV-Stellen sind

- rechtlich selbständig: 48 (12,63 %)
- rechtlich unselbständig: 332 (87,37 %)
- Summe: 380 (100 %) <sup>10</sup>

2) Die Organisation ist zuzuordnen

- öffentlich-rechtlich: 177 (48,23 %)
- privatwirtschaftlich: 154 (41,96 %)
- Mischform: 36 (9,81 %)
- Summe: 367 (100 %)

3) Nach der Gewinnerzielungsabsicht sind

- gewinnorientiert: 125 (35,82 %)
- nicht gewinnorientiert: 224 (64,18 %)
- Summe: 349 (100 %)

4) Die IV-Stellen beschäftigen im Durchschnitt 8,6 Mitarbeiter (Vollzeit- und Teilzeitbeschäftigte) (n = 377)

5) Die IV-Stellen weisen ein Durchschnittsalter von 17,6 Jahren auf (n = 373).

## 4.3. Erfolgsfaktoren

42 alphabetisch sortierte Erfolgsfaktoren<sup>11</sup> mußten auf einer siebenstufigen Intervallskala (von 1 = "nicht wichtig" bis 7 = "sehr wichtig") beurteilt werden.<sup>12</sup> Als Erfolgsfaktoren waren folgende vorgegeben (in den nächsten Abbildungen werden aus Übersichtlichkeitsgründen nur noch die Abkürzungen (z. B. EF 1 usw.) verwendet:

---

10 Nicht alle Fragen wurden beantwortet, deshalb auch hier die Abweichung von N=386.

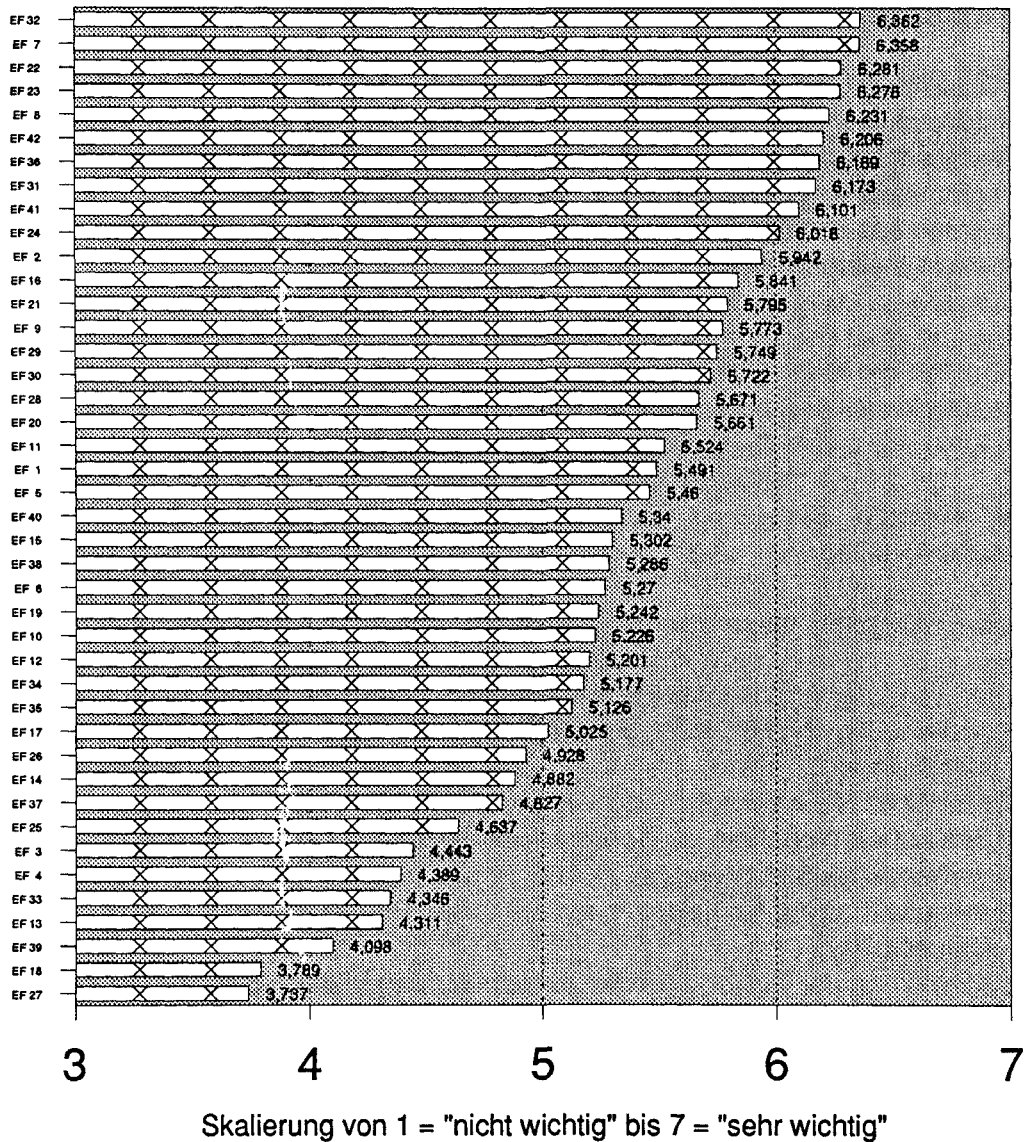
11 Anzumerken ist, daß es sich hier nicht um 42 Erfolgsfaktoren im eigenständigen Sinne handelt, sondern vielmehr um Sub-Erfolgsfaktoren oder Indikatoren, die zu Erfolgsfaktoren aggregiert werden können (vgl. hierzu Herget 1991a).

12 Die (ergänzbare) Liste der Erfolgsfaktoren war für alle IV-Stellen einheitlich. Sollten etwaige Erfolgsfaktoren für eine bestimmte IV-Stelle nicht zutreffen, konnten diese mit "n/a" für "nicht anwendbar" gekennzeichnet werden.

- EF 1: aktive Versorgung der Nutzer mit Informationsdiensten (z.B. Eigeninitiativen, Impulsinformationen)
- EF 2: Aktualität der internen Informationsbestände in der IVS
- EF 3: Anwendung fortgeschrittener Management- und Planungsmethoden in der IVS
- EF 4: Anwendung von Methoden zur Benutzerforschung, um Nutzerbedürfnisse zu identifizieren
- EF 5: Atmosphäre in der IVS (Geräuschpegel, Freundlichkeit der Mitarbeiter, ...)
- EF 6: Aufbau und Betrieb von internen Datenbanken
- EF 7: Aufgeschlossenheit der IVS-Mitarbeiter Nutzer-/Kundenwünschen gegenüber
- EF 8: Bekanntheitsgrad der IVS
- EF 9: Bereitstellung eines umfassenden und vielseitigen Dienstleistungs- und Produktangebotes
- EF 10: Bereitstellung von individualisierten Diensten und Produkten (z.B. individuelle Profildienste)
- EF 11: Einbindung der IVS in die Informationsstrategie der Organisation
- EF 12: eindeutige Funktion (Rolle, Leitbild) der IVS
- EF 13: Einsatz von Verfahren zur Benutzerrückmeldung
- EF 14: formulierte Ziele und Strategien in der IVS
- EF 15: formale Aufbereitung und Präsentation von Informationsprodukten (Format, Layout, Lesbarkeit usw.)
- EF 16: Image der IVS und ihrer Mitarbeiter
- EF 17: inhaltliche Veredelung von Diensten und Produkten (value added)
- EF 18: Inhouse-Training für Nutzer im selbständigen Recherchieren in externen Datenbanken
- EF 19: Innovationsfähigkeit (neue Dienste/Produkte) in der IVS
- EF 20: jederzeitige Ansprech- und Erreichbarkeit der IVS-Mitarbeiter
- EF 21: Kenntnis der Tätigkeit und Aufgabenbereiche der Nutzer/Kunden
- EF 22: Kompetenz und Qualifikation der IVS-Mitarbeiter (IuD-Gebiet)
- EF 23: Kompetenz und Qualifikation der IVS-Mitarbeiter (Kommunikationsfähigkeit)
- EF 24: Kompetenz und Qualifikation der IVS-Mitarbeiter (Sach-/Fachgebiet)
- EF 25: Kooperation mit anderen IVS
- EF 26: Marketing für IVS-Dienste und -Produkte
- EF 27: Mitwirkung der Nutzer-/Kunden bei der Gestaltung der IVS-Geschäftspolitik
- EF 28: Nutzer-/Kunden- und Serviceorientierung der IVS-Mitarbeiter
- EF 29: Nutzung moderner Technologie in der IVS
- EF 30: physische Zugänglichkeit zu Ressourcen in der IVS (Bücher, Berichte, Zeitschriften etc.)
- EF 31: Qualität der Unterstützung und Beratung für die Nutzer-/Kunden
- EF 32: Qualität von Informationsdiensten und -produkten (Aktualität, Genauigkeit, Vollständigkeit etc.)
- EF 33: räumliche Nähe der IVS zu den Nutzern/Kunden
- EF 34: regelmäßige Kommunikation zwischen IVS und Fachabteilungen
- EF 35: regelmäßige Kommunikation zwischen IVS und vorgesetztem Management
- EF 36: Schnelligkeit der Informationsversorgung
- EF 37: Schulung der Nutzer in der Nutzung des Dienstleistungs- und Produktangebotes der IVS (z.B. Einführungsveranstaltungen)
- EF 38: Unterstützung durch das Topmanagement
- EF 39: vorhandenes Statistikwesen (Nutzer, Profile, ...)
- EF 40: wirtschaftliche (kostengünstige) Leistungserbringung
- EF 41: Zufriedenheit und Motivation der Mitarbeiter
- EF 42: Zugang zu Datenbanken (Online und CD-ROM)

In der folgenden Abbildung sind die Mittelwerte der Erfolgsfaktoren für alle IVS in absteigender Folge dargestellt:

Abb. 6: Erfolgsfaktoren (alle IVS)



Eine Interpretation der Ergebnisse auf dieser generellen Ebene muß vorerst noch vorsichtig formuliert werden. Notwendig ist eine differenzierte Betrachtung, die nach Größe, Gewinnerzielungsabsichten, Rechtsform, Einbindung in Organisationen usw. unterscheidet.

Bereits auf dieser Ebene kann jedoch festgestellt werden, daß die Antwortenden das Spektrum der Antwortmöglichkeiten gut ausgeschöpft haben. Den vorgelegten Erfolgsfaktoren wird zum großen Teil (31 Werte liegen über 5,0) eine wichtige Bedeutung zugesprochen. Auch ist die Differenz zwischen dem höchsten (6,362) und dem niedrigsten Mittelwert (3,737) sehr deutlich.

Als wichtigste zehn Erfolgsfaktoren werden folgende genannt:

- 1) Qualität von Informationsdiensten und -produkten (Aktualität, Genauigkeit, Vollständigkeit etc.)
- 2) Aufgeschlossenheit der IVS-Mitarbeiter Nutzer-/Kundenwünschen gegenüber
- 3) Kompetenz und Qualifikation der IVS-Mitarbeiter (IuD-Gebiet)
- 4) Kompetenz und Qualifikation der IVS-Mitarbeiter (Kommunikationsfähigkeit)
- 5) Bekanntheitsgrad der IVS
- 6) Zugang zu Datenbanken (Online und CD-ROM)
- 7) Schnelligkeit der Informationsversorgung
- 8) Qualität der Unterstützung und Beratung für die Nutzer-/Kunden
- 9) Zufriedenheit und Motivation der Mitarbeiter
- 10) Kompetenz und Qualifikation der IVS-Mitarbeiter (Sach-/Fachgebiet)

Die generelle Bedeutung der Erfolgsfaktoren "Qualität von Produkten und Diensten", "Mitarbeiter" und "Kundenorientierung" wird hier sofort ersichtlich.

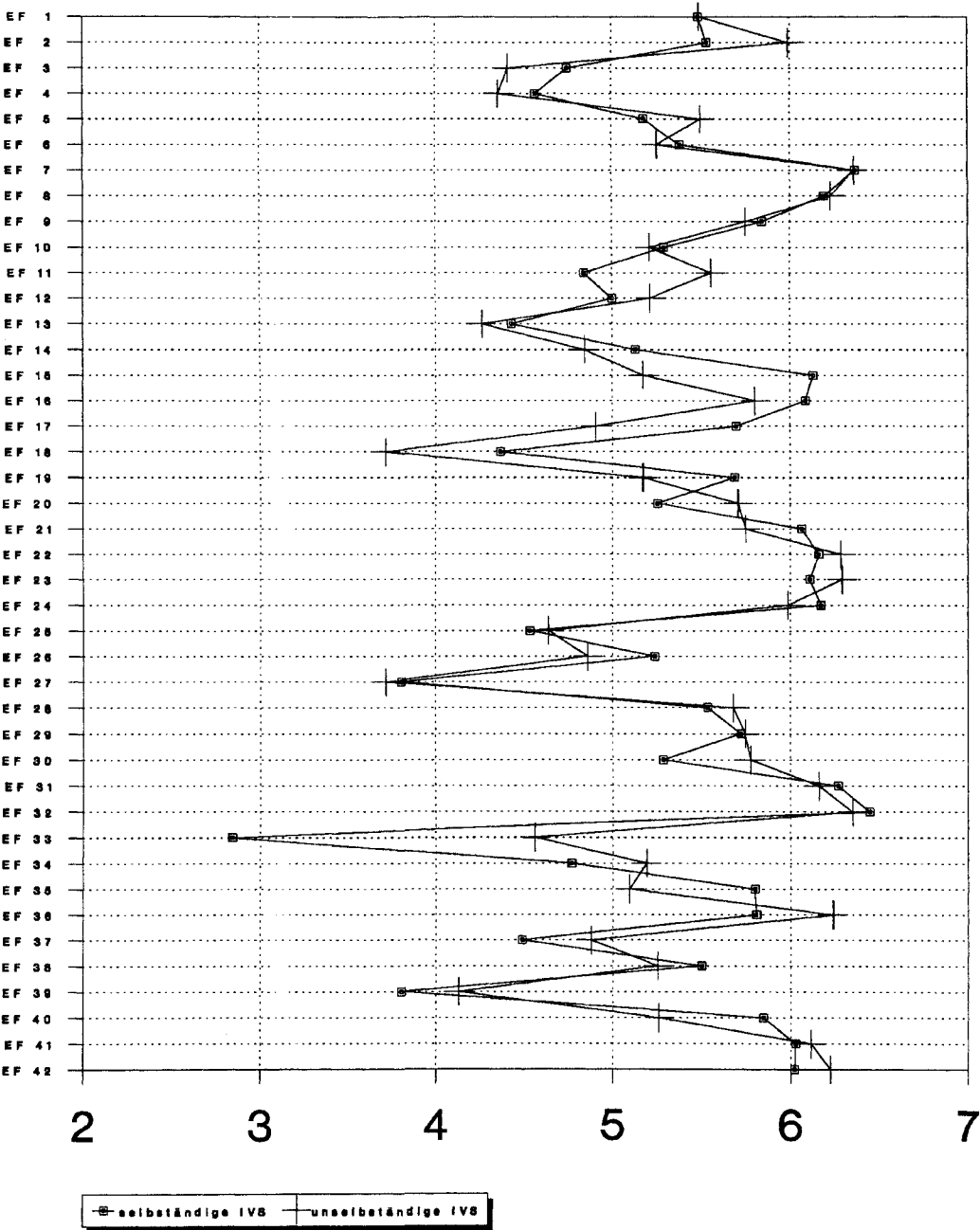
Als am wenigsten relevant wurden eingestuft:

- 36) Anwendung fortgeschrittener Management- und Planungsmethoden in der IVS
- 37) Anwendung von Methoden zur Benutzerforschung, um Nutzerbedürfnisse zu identifizieren
- 38) räumliche Nähe der IVS zu den Nutzern/Kunden
- 39) Einsatz von Verfahren zur Benutzerrückmeldung
- 40) vorhandenes Statistikwesen (Nutzer, Profile, ...)
- 41) Inhouse-Training für Nutzer im selbständigen Recherchieren in externen Datenbanken
- 42) Mitwirkung der Nutzer-/Kunden bei der Gestaltung der IVS-Geschäftspolitik.

Dem Einsatz von Management- und Planungsinstrumenten zur besseren Eruiierung der Nutzer und ihrer Bedürfnisse und der Einbindung der Nutzer bei der Strategieentwicklung wird überraschenderweise als Erfolgsfaktoren wenig Bedeutung beigemessen. Dies könnte auch als ein Hinweis auf mangelnde Professionalisierung beim Einsatz von Führungsinstrumenten in den Informationsvermittlungsstellen gedeutet werden.

Die nächste Abbildung zeigt die Mittelwerte der Erfolgsfaktoren differenziert nach rechtlich selbständigen und unselbständigen IVS:

Abb. 7: Erfolgsfaktoren (rechtlich selbständige und unselbständige IVS)



Überraschend ist hier die weitgehende Übereinstimmung bei der Bewertung von Erfolgsfaktoren zwischen selbständigen (extern orientierten) und unselbständigen (intern orientierten) IVS. Interessant ist aber die unterschiedliche Bewertung des Aspektes der formalen Aufbereitung und Präsentation von Informationsprodukten (EF 15), die bei selbständigen, also vor allem mit externen Nutzern befaßten IVS, eine viel höhere Priorität genießt. Ebenso verhält es sich bei der inhaltlichen Veredelung von Diensten und Produkten (EF 17). Diese Value-added-Aspekte werden von den selbständigen Informationsvermittlern als für die erfolgreiche Informationsvermittlung weitaus bedeutsamer eingeschätzt.

Weitere Analysen und Interpretationen können dem ausführlichen Forschungsbericht (Herget/Hensler 1991c) entnommen werden.

## 5. Ausblick

Die Ergebnisse der Erfolgsfaktorenforschung leisten neben theoretischem Erkenntnisgewinn vor allem auch in praxeologischer Hinsicht einen wichtigen Beitrag.

Die Kenntnis von Erfolgsfaktoren der Informationsvermittlung

- gibt Hinweise auf notwendige Ausbildungs- und Weiterbildungsmaßnahmen in der Ausbildung von Informationsspezialisten;
- zeigt kontextspezifische Erfolgskonstellationen und prototypische Modelle von leistungsfähigen Informationsvermittlungsstellen;
- bietet ein Instrumentarium zur Analyse und Evaluation von Informationsorganisationen und hilft somit notwendige Optimierungsstrategien zu entwickeln (Audit-Konzept).

Aufgrund des kontingenten Wirkungsgeflechts innerhalb dessen sich die Informationsvermittlungsstellen bewegen, dürfen allerdings keine allgemeingültigen Erkenntnisse mit Gültigkeitscharakter für alle verschiedenen Ausprägungsarten erwartet werden. Wohl aber begründete Annahmen über erfolgversprechende Strategien zur Optimierung der Informationsvermittlung.

Zusammenfassend betrachtet leistet die Erfolgsfaktorenforschung in der Informationswissenschaft einen Beitrag zur weiteren Optimierung der Informationsvermittlung mit dem Ziel, die gesellschaftliche Informationssituation von Individuen und Organisationen zu verbessern.



## Literatur

Ahrend, W.M. Die Nützung von externen Informationsdatenbanken durch kleine und mittlere Unternehmen: Ergebnisse einer empirischen Untersuchung. In: J. Herget/ R. Kuhlen (Hg.) 1990.

Belkin, N.J./Vickery, A. Interaction in information systems: a review of research from document retrieval to knowledge-based systems. London: British Library, 1985.

Daniel, D.R. Management information crisis. Harvard Business Review, 1961, 39(5), 111-121.

Diegruber, J. Erfolgsfaktoren nationaler europäischer Linienluftverkehrsgesellschaften im Markt der 90er Jahre. Dissertation. St. Gallen 1991.

Efthimiadis, E.N./Robertson, S.E. Feedback and interaction in information retrieval. In: C. Oppenheim/C. L. Citroen/J.-M. Griffith. (Hg.) Perspectives in information management 1. London: Butterworths, 1989.

Fritz, W. Marketing - ein Schlüsselfaktor des Unternehmenserfolgs? Eine kritische Analyse vor dem Hintergrund der empirischen Erfolgsfaktorenforschung. Marketing, 1990, 12, 91-110.

Gemünden, H.G. Erfolgsfaktoren des Projektmanagements - eine kritische Bestandsaufnahme der empirischen Untersuchungen. Projekt Management Nr. 1&2/1990, S. 4-15.

Geschka, H. Voraussetzungen für erfolgreiche Innovationen - Beachtung von Hindernissen und Erfolgsfaktoren bei der Innovationsplanung. In: H. Corsten (Hg.) Die Gestaltung von Innovationsprozessen. Hindernisse und Erfolgsfaktoren im Organisations-, Finanz- und Informationsbereich. Berlin: Schmidt, 1989.

Gloekner-Rist, A./Lehmler, W./Wettler, M. Endnutzerrecherchen in Bibliographischen CD-ROM-Datenbanken. Nachrichten für Dokumentation, 1989, 40, 151-159.

Gomez, P./Probst, G. Vernetztes Denken im Management, Die Orientierung Nr. 89, Bern: Schweizerische Volksbank, 1987.

Herget, J. Externalisierung der Informationsmanagement-Funktion in mittelständischen Unternehmen. Ergebnisse einer explorativen Studie zur Nutzung von Informationsprodukten aus Online-Datenbanken. In: J. Herget/R. Kuhlen (Hg.) 1990. (= Herget 1990a)

Herget, J. Online-Datenbanken als modernes Informationsmedium. Zur Situation der Informationsvermittlung bei mittelständischen Unternehmen. Dissertation. Konstanz 1990. (= Herget 1990b)

Herget, J. Erfolgsfaktoren der Informationsvermittlung. Teil III: Theoretischer Bezugsrahmen. Universität Konstanz, Informationswissenschaft, Oktober 1991. (= Herget 1991a)

Herget, J. Erfolgsfaktoren beim Management von Informationsvermittlungsstellen. Ergebnisse einer empirischen Studie. In: Deutsche Gesellschaft für Dokumentation e.V. (Hg.) Deutscher Dokumentartag 1991, 30. September bis 2. Oktober, Ulm: "Information und Dokumentation in den 90er Jahren - Neue Herausforderung, neue Technologien". München: Saur, 1991. (= Herget 1991b)

Herget, J./Hensler, S. Erfolgsfaktoren der Informationsvermittlung. Teil 1: Theoretische Grundlagen und methodische Konzepte. Universität Konstanz, Informationswissenschaft, Bericht 4/91, Juni 1991. (= Herget/Hensler 1991a)

Herget, J./Hensler, S. Erfolgsfaktoren der Informationsvermittlung. Teil 2: Modelle, Methoden und Instrumente. Universität Konstanz, Informationswissenschaft, Bericht 5/91, September 1991. (= Herget/Hensler 1991b)

Herget, J./Hensler, S. Erfolgsfaktoren der Informationsvermittlung. Teil 5: Ergebnisse einer schriftlichen Befragung von Leitern von Informationsvermittlungsstellen. Universität Konstanz, Informationswissenschaft, Oktober 1991. (= Herget/Hensler 1991c)

Herget, J./Kuhlen, R. (Hg.) Pragmatische Aspekte beim Entwurf und Betrieb von Informationssystemen. Proceedings des 1. Internationalen Symposiums für Informationswissenschaft, Universität Konstanz, 17.-19. Oktober 1990. Konstanz: Universitätsverlag, 1990.

Hoffmann, F. Kritische Erfolgsfaktoren - Erfahrungen in großen und mittelständischen Unternehmungen. Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung, 1986, 38, 831-843.

Kaufmann, T. Kauf und Verkauf von Unternehmungen. Eine Analyse qualitativer Erfolgsfaktoren. Dissertation. St. Gallen. 1990.

Knoll, G. Strategische Erfolgsfaktoren in mittelständischen Tiefbauunternehmen. Eine empirische Untersuchung. Spardorf: Wilfer, 1989.

Krups, M. Marketing innovativer Dienstleistungen am Beispiel elektronischer Wirtschaftsinformationsdienste. Frankfurt am Main: Lang, 1985.

Kuipers, M. Erfolgsfaktoren der Unternehmensgründung. Eine Untersuchung erfolgreicher und erfolgloser Unternehmensgründer in der Schweiz. Dissertation. St. Gallen. 1990.

Lehner, F. Die Erfolgsfaktoren-Analyse als Instrument des Informationsmanagements - Erfahrungen bei der praktischen Anwendung. In: J. Herget/R. Kuhlen (Hg.) 1990.

Leidecker, J.K./Bruno, A.V. Identifying and using critical success factors. Long Range Planning, 1984, 17, 23-32.

Magal, S.R./Carr, H.H./Watson, H.J. Critical success factors for information center managers. MIS Quarterly, 1988, 12, 412-425.

Mair, S. Erfolgsfaktoren der mittelständischen Druckindustrie. Dissertation. St. Gallen, 1989.

Müller, D./Winand, U. Fachinformations-Marketing. Sankt Augustin: GMD, 1988.

Nagel, K. Die 6 Erfolgsfaktoren des Unternehmens. Strategie, Organisation, Mitarbeiter, Führungssystem, Informationssystem, Kundennähe. Landsberg am Lech: Moderne Industrie, 1988, 2. Auflage.

Patt, P.-J. Strategische Erfolgsfaktoren im Einzelhandel. Eine empirische Analyse am Beispiel des Bekleidungsfachhandels. Frankfurt am Main: Lang, 1990.

Peters, T.J./Waterman, R.H., Jr. Auf der Suche nach Spitzenleistungen: Was man von den bestgeführten US-Unternehmen lernen kann. Landsberg am Lech: Moderne Industrie, 1984, 6. Auflage.

Poth, L.G. Erfolgsfaktoren von Marktführern - Ergebnisse einer Studie im Fachbereich Wirtschaft der Fachhochschule Düsseldorf. In: K. Apitz/W. Gäbler/B.F. Holz/L.G. Poth/B.J. Wieczorek (Hg.) Erfolgsfaktoren von Marktführern. Landsberg am Lech: Moderne Industrie, 1989.

Porter, M.E. The competitive advantage of nations & their firms. New York: Free Pr., 1990.

Reinhard, M. Wirtschaftsinformation in der Bundesrepublik Deutschland - Angebot und Nutzung aus der Sicht der Nachfrager. München: Ifo, 1987

Rockart, J.F. Chief executives define their own data needs. Harvard Business Review, 1979, 57(2), 81-93.

Trommsdorff, V. Erfolgsfaktorenforschung, Produktinnovation und Schnittstelle Marketing - F&E. Technische Universität Berlin 1990.

Ulrich, H./Probst, G. Anleitung zum ganzheitlichen Denken und Handeln. Bern: Haupt, 1990.

Unterkofler, G. Erfolgsfaktoren innovativer Unternehmensgründungen: ein gestaltungsorientierter Lösungsansatz betriebswirtschaftlicher Gründungsprobleme. Frankfurt am Main: Lang, 1989.

Vester, F. Sensitivitätsmodell, Frankfurt am Main, 1980.

Wind, T. Benutzer und Informationseinrichtung. Dissertation. Kassel, 1979.

Wohlgemuth, A.C. Unternehmungsd Diagnose in Schweizer Unternehmungen: Untersuchungen zum Erfolg mit besonderer Berücksichtigung des Humanpotentials. Bern: Lang, 1989.

Dr. Josef Herget  
Universität Konstanz  
Informationswissenschaft  
Postfach 5560  
W-7750 Konstanz

## Zur Abhängigkeit der Ergebnisse maschineller Indexierung vom verwendeten Begriffssystem

Erich Mater, Berlin

Ziel dieses Beitrages ist es zu zeigen, in wie hohem Maße die Ergebnisse automatischer Inhaltserschließung vom verwendeten Bezugssystem abhängen.

Als Beispiel dafür dient die deutschsprachige Fassung eines in der Internat. Classification (15, 1988; 3, p. 125) publizierten Textes über menschliche und maschinelle Verfahren bei der Inhaltserschließung.

Die Methode beruht auf Selektion und Zählung von Textwörtern. Sie ergeben bei unterschiedlichen Bezugssystemen verschiedene Relevanzwerte hinsichtlich Vollständigkeit und Genauigkeit. Mit solchen frequenzstatistischen Verfahren lassen sich Suchmerkmale gewinnen, die für einen Retrievalprozeß nützlich sind. Dabei sollen hier nur verschiedene Methoden zur Gewinnung derartiger Suchmerkmale vorgeführt werden, ohne diese generell zu bewerten.

Die theoretische Basis solcher Methoden soll dem Verständnis für höher entwickelte Verfahren der Wissensgewinnung dienen. Dagegen werden die erforderlichen Rechnerprozeduren auf allgemeinverständlichem Niveau gehalten, damit der Vorrang der geistigen vor den technischen Prozessen deutlich hervortritt.

Die Ergebnisse derartiger Indexierverfahren dürften sowohl theoretisch wie praktisch von Interesse sein: praktisch, weil man sich unter verschiedenen Varianten diejenige auswählen kann, die in Bezug auf eine gegebene Aufgabe die besten Ergebnisse bringt; theoretisch, weil strukturverarbeitende Prozesse im Logikteil eines Computers auf der semantischen Ebene menschlichen Wissens über Benennung-Bezeichnung-Bedeutung-Begriff zu unterschiedlichem Genauigkeitsgrad von Wissen führen.

# Zur Repräsentation von Programmierwissen

Walter Dosch

**Zusammenfassung.** Die schrittweise Entwicklung von Programmen aus ihrer formalen Spezifikation stützt sich auf regelorientierte, semantiktreue Übergänge. Die Transformationsregeln repräsentieren grundlegendes Programmierwissen in schematischer Weise. Für eine erfolgreiche Programmkonstruktion sind die Entwicklungsstrategien und Anwendungskontexte der Regeln wichtig.

**Abstract.** The stepwise development of programs from their formal specification is based on rule-oriented, semantics-preserving transformations. The transformation rules represent fundamental programming knowledge in a schematic way. For the successful construction of programs the development strategies and application contexts of the rules are important.

## 1 Einführung

Programmieren bedeutet im weitesten Sinn, zu einem gegebenen Problem einen effizienten Algorithmus zu entwerfen; neben allgemeinen Qualitätsmerkmalen wie Modularität, Anpaßbarkeit, Wartbarkeit und Wiederverwendbarkeit ist Korrektheit ein zentrales Anliegen. Dabei heißt ein Algorithmus korrekt, wenn er die vorgegebene Aufgabenstellung erfüllt.

Aus der meist informellen, vagen Aufgabenbeschreibung entsteht beim Erarbeiten der *Anforderungsdefinition* zunächst eine formale *Problemspezifikation* ("was ist zu tun"). Die Implementierung der im allgemeinen nichtoperativen Spezifikation folgt der Methode der schrittweisen Verfeinerung (/16/): durch *Entwurfsentscheidungen* wird die operative Durchführung detailliert, Programmieren wird zum Entwicklungsprozeß (/2/). Bei "freien" Übergängen zwischen aufeinander folgenden Programmfassungen ist für die Korrektheit des entstehenden Programms allerdings die Verifikation der Einzelschritte oder — ungleich schwerer — des Gesamtschritts nötig.

Im Gegensatz zu einer nachträglichen Verifikation setzt sich zunehmend die Auffassung durch, Programmierung und Verifikation methodisch zu verbinden, also Programmwurf und Korrektheitsnachweis simultan durchzuführen. Das entstehende Programm ist dann per Konstruktion korrekt (/14/). Vorteilhafterweise stützt man sich dazu auf formale, regelorientierte Übergänge zwischen Programmen, wobei die Semantiktreue der Regeln ein- für allemal bewiesen wird (/13/). Damit wird Programmieren zur *Programmkonstruktion* (/7/), also zur Ableitung des Programms aus seiner Spezifikation nach formalen Regeln (/8/).

Die Transformationsregeln repräsentieren Programmierwissen in schematischer Weise, indem sie vom Anwendungsbereich des Einzelprogramms abstrahieren. Sie sind deshalb auf weite Programmklassen anwendbar und können als allgemeine konzeptuelle Basis der

Programmierung dienen. Die Regeln codifizieren Programmierwissen; erst durch diesen Formalisierungsgrad wird eine im Gegensatz zu vorhandenen Systemen nicht syntaktisch, sondern semantisch orientierte Rechnerunterstützung der Programmierung möglich. Die Entwicklungsgeschichte eines Programms dient zudem als Dokumentation, um die Wartbarkeit bei sich ändernder Spezifikation oder Zielmaschine zu unterstützen.

## 2 Grundlagen

Programmierwissen über die Gleichwertigkeit von Programm(teil)en kann man durch Transformationsregeln repräsentieren.

### 2.1 Programme und Programmschemata

Die kontextfreie Syntax einer Programmiersprache wird durch eine endliche *Signatur*  $\Sigma = (\mathcal{S}, \mathcal{F})$  beschrieben, deren Sorten  $\mathcal{S}$  die syntaktischen Einheiten wie **identifier**, **expression** oder **statement** und deren Funktionen  $\mathcal{F}$  die Sprachkonstrukte wie

$$\text{assign: identifier} \times \text{expression} \rightarrow \text{statement}$$

sind (compare /12/). Die Grundtermalgebra  $W_\Sigma = (W_\Sigma^s)_{s \in \mathcal{S}}$  umfaßt Grundterme, die Programme in *abstrakter Syntax* darstellen, zum Beispiel

$$x := 1; y := 2 \hat{=} \text{semi}(\text{assign}(x, 1), \text{assign}(y, 2)) \in W_\Sigma^{\text{statement}}.$$

Über einer stetigen  $\Sigma$ -Algebra  $\mathcal{A}$ , einem semantischen Modell der Sprache, wird die *denotationelle Semantik* kompositional durch eine Interpretationsfunktion  $\llbracket \cdot \rrbracket: W_\Sigma \times \mathcal{E} \rightarrow \mathcal{A}$  festgelegt, wobei die Umgebung  $\mathcal{E}$  die Belegung von Bezeichnern festhält. Zwei Programm(teil)e  $p_1, p_2 \in W_\Sigma^s$  heißen (*semantisch*) *äquivalent*, in Zeichen  $p_1 \approx p_2$ , falls  $\llbracket p_1 \rrbracket(\rho) = \llbracket p_2 \rrbracket(\rho)$  für alle  $\rho \in \mathcal{E}$  gilt.

Man abstrahiert vom konkreten Programm zur Klasse syntaktisch ähnlicher Programme durch eine abzählbare Familie  $X = (X^s)_{s \in \mathcal{S}}$  von *Schemaparametern*, die für Programmteile der jeweiligen syntaktischen Einheit stehen. Die Termalgebra  $W_\Sigma(X)$  umfaßt dann die Familie der *Programmschemata*. Bei der *Instantiierung*  $p[\sigma]$  eines Programms  $p$  unter einer *Substitution*  $\sigma: X \rightarrow W_\Sigma(X)$  werden alle im Programm  $p$  vorkommenden Schemaparameter simultan durch Programme bzw. Programmschemata ersetzt.

### 2.2 Transformationsregeln und ihre Anwendung

Häufig unterliegen Transformationen Einschränkungen. Als *Anwendbarkeitsbedingungen* dienen Horn-Klauseln, deren atomare Teilformeln aus Anwendbarkeitsprädikaten aufgebaut sind. Dazu zählen neben statisch überprüfbaren syntaktischen Anwendbarkeitsprädikaten wie  $x \in \text{free}(e)$  oder  $x \in \text{bound}(e)$  zur Prüfung, ob ein Bezeichner frei bzw. gebunden in einem Ausdruck  $e$  vorkommt, auch semantische Anwendbarkeitsprädikate wie  $\text{defined}(e)$  oder  $\text{strict}(x, e)$  zur Prüfung, ob ein Ausdruck  $e$  definiert bzw. strikt im Bezeichner  $x$  ist. Die Verifikation semantischer Anwendbarkeitsbedingungen kann komplexe

Beweise, wie etwa den Terminierungsnachweis von Rechenvorschriften oder Prozeduren erfordern.

Eine *Transformationsregel*  $cond \Rightarrow in \approx out$  besteht aus einem *Eingabeschema*  $in$ , einem *Ausgabeschema*  $out$  und einer Anwendbarkeitsbedingung  $cond$ , wobei  $in, out \in W_\Sigma(X)$  Programmschemata gleicher Sorte sind.

Wendet man eine Transformationsregel auf einen Programmteil  $p$  an, so führt die Unifikation des Eingabeschemas  $in$  mit  $p$  zu einer Substitution  $\sigma: X \rightarrow W_\Sigma$ ; Ergebnis ist das instantiierte Ausgabeschema  $out[\sigma]$ , falls die instantiierte Anwendbarkeitsbedingung  $cond[\sigma]$  erfüllt ist.

Zum Beispiel führt die Anwendung der Regel von der *Negation der Bedingung*

$$\text{if } c \text{ then } e_1 \text{ else } e_2 \text{ fi} \approx \text{if } \text{not}(c) \text{ then } e_2 \text{ else } e_1 \text{ fi}$$

auf den bedingten Ausdruck  $\text{if } x \geq 0 \text{ then } x \text{ else } -x \text{ fi}$  durch die Substitution  $\{c \mapsto x \geq 0, e_1 \mapsto x, e_2 \mapsto -x\}$  zum Ergebnis

$$\text{if } \text{not}(x \geq 0) \text{ then } -x \text{ else } x \text{ fi}.$$

## 2.3 Gültigkeit von Regeln

Eine Transformation  $cond \Rightarrow in \approx out$  von Programmteilen heißt *gültig*, wenn das Ein- und Ausgabeprogramm unter allen Belegungen  $\rho \in \mathcal{E}$ , welche die Anwendbarkeitsbedingung erfüllen, äquivalent sind: wenn  $\models_\rho cond$ , dann  $\llbracket in \rrbracket(\rho) = \llbracket out \rrbracket(\rho)$ . Für das Definiertheits- bzw. Striktheitsprädikat legt man fest:

$$\begin{aligned} \models_\rho \text{defined}(e) & \text{ gdw. } \llbracket e \rrbracket(\rho) \neq \perp \\ \models_\rho \text{strict}(x, e) & \text{ gdw. (wenn } \rho(x) = \perp, \text{ dann } \llbracket e \rrbracket(\rho) = \perp) \end{aligned}$$

Eine Transformationsregel  $cond \Rightarrow in \approx out$  heißt *gültig*, wenn sie für alle Substitutionen  $\sigma: X \rightarrow W_\Sigma$  eine gültige Transformation  $cond[\sigma] \Rightarrow in[\sigma] \approx out[\sigma]$  ergibt.

## 2.4 Sprache

Die Gültigkeit von Transformationsregeln hängt wesentlich von der Semantik der verwendeten Programmiersprache ab. Zum Beispiel gilt die  $\beta$ -Konversion

$$((\text{t } x)\text{r}: e_1)(e_2) \approx e_1[e_2 \text{ for } x]$$

bei einer *call-by-name* Semantik; eine *call-by-value* Semantik erfordert jedoch die Einschränkung  $\neg \text{defined}(e_2) \Rightarrow \text{strict}(x, e_1)$ . Bei vielen existierenden Programmiersprachen führen zudem Mängel im Sprachentwurf, wie zum Beispiel Seiteneffekte und Sprünge in Ausdrücken, zu einer komplexen Semantik. Dies erschwert Formulierung und Gültigkeitsnachweis von Transformationsregeln, da die zugrunde liegende "Algebra der Programme" (/1/) wenig Regularität zeigt.

Zur Unterstützung der transformationellen Programmierung wurde die in dieser Arbeit benutzte Breitbandsprache CIP-L (/4/) entworfen, welche die verschiedenen Abstraktionsebenen der Programmierung in einem kohärenten semantischen Rahmen integriert. Auch sind ihre Sprachkonzepte wie zum Beispiel erratischer Nichtdeterminismus oder algebraische Spezifikationen mit loser Semantik auf die Bedürfnisse der Programmentwicklung ausgerichtet. Die Regeln dieser Arbeit setzen deterministisches CIP-L voraus.

### 3 Klassifikation von Transformationsregeln

Die Transformationsregeln lassen sich in einem Transformationskatalog unter verschiedenen Gesichtspunkten klassifizieren.

#### 3.1 Sprachebenen

Die Programmentwicklung überspannt das weite Feld von problemorientierter Spezifikation zu maschinenorientierter Implementierung oder Hardwarerealisierungen (/9/). Entsprechend teilt man Transformationen nach der *Sprachebene* ein, auf der die Programmumformung stattfindet.

Auf der *Spezifikationsebene* wird das Ein-/Ausgabeverhalten eines Algorithmus durch seine Eigenschaften definiert. Kennzeichnende Sprachelemente sind prädikatenlogische Formeln mit All- und Existenzquantor, Auswahl- und Kennzeichnungsoperator. Die Umformung von Spezifikationen bereitet den Übergang zur algorithmischen Formulierung vor. Zum Beispiel dient die *Adjunktion einer Tautologie*

$$\text{defined}(c_1) \Rightarrow c \approx c \text{ and } (c_1 \text{ or } \text{not}(c_1))$$

zur Einführung einer Fallunterscheidung.

Die *funktionale Programmierung* (/10/) unterstützt die problemorientierte Formulierung von Algorithmen auf einer abstrakten Ebene. Kennzeichnende Sprachelemente sind Funktionsanwendung, Abstraktion, Fallunterscheidung und Rekursion. Die zugehörigen Regeln formen applikative Ausdrücke um, zum Beispiel durch Vereinfachen des Konditionals oder Expandieren und Komprimieren von Rechenvorschriften:

$$\begin{aligned} f \notin \text{free}(e_2) \wedge \text{defined}(e_2) \Rightarrow \\ \text{begin funct } f \equiv (\text{t } x) \text{ r: } e_1; f(e_2) \text{ end } \approx \\ \text{begin funct } f \equiv (\text{t } x) \text{ r: } e_1; e_1[e_2 \text{ for } x] \text{ end} \end{aligned}$$

Die *prozedurale Ebene* der Programmierung dient zur Implementierung von Algorithmen auf der VON NEUMANN Maschine. Kennzeichnende Sprachkonstrukte sind Zuweisungen, bedingte Anweisungen, Prozeduren und Schleifen. Die zugehörigen Regeln beschäftigen sich vornehmlich mit der Umformung von Anweisungen; ein Beispiel ist das *Strecken der while-Schleife*:

$$\text{while } c \text{ do } s \text{ od } \approx \text{if } c \text{ then } s; \text{ while } c \text{ do } s \text{ od else skip fi}$$

Die *maschinenorientierte Ebene* ist bei der Kontrollstruktur mit dem Übergang zu Sequenzen einfacher Anweisungen und (bedingter) Sprungbefehle, bei der Datenstruktur zum homogenen, adressierten, wortorientierten Speicher verbunden. Hier fallen zahlreiche Regeln zur Optimierung von Anweisungs- oder Sprungsequenzen an, wie zum Beispiel

$$x \notin \text{free}(e) \Rightarrow x := e; y := e \approx x := e; y := x.$$



## 3.2 Übergänge zwischen Sprachebenen

Neben den Transformationen *innerhalb* einer Sprachebene sind auch die Transformationen *zwischen* zwei Sprachebenen wichtig.

Beim Übergang von logischen Spezifikationen zu funktionalen Programmen werden wesentliche Entwurfsentscheidungen für ein algorithmisches Prinzip getroffen. Zum Beispiel kann ein *Existenzquantor* über einem endlichen Bereich durch Ausschöpfung algorithmisiert werden:

$$\text{defined}(c) \wedge \mathbf{m} = m_1 | m_2 | \dots | m_k \Rightarrow (\exists \mathbf{m} x: c) \approx c[m_1 \text{ for } x] \vee \dots \vee c[m_k \text{ for } x]$$

Der Übergang von der funktionalen zur prozeduralen Ebene geht mit der Einführung von Programmvariablen einher. Wichtigste Regel ist die *Überführung repetitiver Rechenvorschriften in iterative Form*:

$$\begin{aligned} f \notin \text{free}(c) \cup \text{free}(e_1) \cup \text{free}(e_2) \wedge \text{new}(X) &\Rightarrow \\ \text{funct } f &\equiv (\mathbf{t} x) \mathbf{r}: \text{ if } c \text{ then } e_1 \text{ else } f(e_2) \text{ fi } \approx \\ \text{funct } f &\equiv (\mathbf{t} X) \mathbf{r}: \text{ begin var } \mathbf{t} x := X; \text{ while } \text{not}(c) \text{ do } x := e_2 \text{ od}; e_1 \text{ end} \end{aligned}$$

Der Schritt zu einem maschinennahen Programm wird meist schematisch von einem Übersetzer vollzogen, der bei der Codeerzeugung Regeln wie das *Aufbrechen der while-Schleife* verwendet:

$$\text{new}(m) \Rightarrow \text{while } c \text{ do } s \text{ od} \approx \text{begin } m: \text{ if } c \text{ then } s; \text{ goto } m \text{ else skip fi end}$$

## 3.3 Kontroll- und Datenstruktur

Transformationsregeln für die *Kontrollstruktur* ändern den Ablauf eines Programms. Dazu zählen vielfältige Umformungen rekursiver Rechenvorschriften wie etwa die *Umklammerungsregel*

$$\begin{aligned} \text{free}(c) \cup \text{free}(e_1) \cup \text{free}(e_2) \cup \text{free}(e_3) \subseteq \{x\} \wedge \text{monoid}(\mathbf{m}, \circ, n) &\Rightarrow \\ \text{funct } f &\equiv (\mathbf{m} x) \mathbf{m}: \text{ if } c \text{ then } e_1 \text{ else } e_2 \circ f(e_3) \text{ fi } \approx \\ \text{funct } f &\equiv (\mathbf{m} x) \mathbf{m}: g(n, x) \text{ where} \\ \text{funct } g &\equiv (\mathbf{m} r, \mathbf{m} x) \mathbf{m}: \text{ if } c \text{ then } r \circ e_1 \text{ else } g(r \circ e_2, e_3) \text{ fi} \end{aligned}$$

Transformationen der *Datenstruktur* spiegeln hingegen die algebraischen Eigenschaften der Objektstruktur wider, wie zum Beispiel die *Zugriffsgesetze für Stapel*:

$$\begin{aligned} \text{defined}(e_2) &\Rightarrow \text{first}(\text{prefix}(e_1, e_2)) \approx e_1 \\ \text{defined}(e_1) &\Rightarrow \text{rest}(\text{prefix}(e_1, e_2)) \approx e_2 \end{aligned}$$

Bei komplexen Übergängen müssen Transformationen für Kontroll- und Datenstruktur zuweilen auch verschmolzen werden (*“joint development”*).

## 3.4 Elementare und abgeleitete Regeln

Eine andere Einteilung unterscheidet elementare von abgeleiteten Regeln. *Elementare Transformationsregeln* wie das Expandieren einer Rekursion

$$(\text{fix } x: e) \approx e[(\text{fix } x: e) \text{ for } x]$$

geben grundlegende Eigenschaften einzelner Sprachkonstrukte wider; ihre Gültigkeit läßt sich unmittelbar aus der Semantik nachrechnen:

$$\begin{aligned}
& \llbracket e[(\mathbf{fix} \ x: e) \mathbf{for} \ x] \rrbracket(\rho) && \{\text{Substitutionssatz}\} \\
= & \llbracket e \rrbracket(\rho[x/\llbracket \mathbf{fix} \ x: e \rrbracket(\rho)]) && \{\text{Es sei } \llbracket \mathbf{fix} \ x: e \rrbracket(\rho) = \mu(g), \text{ wobei } g(d) = \llbracket e \rrbracket(\rho[x/d])\} \\
= & g(\mu(g)) && \{\text{Fixpunkteigenschaft}\} \\
= & \mu(g) && \{\text{Definition von } g\} \\
= & \llbracket \mathbf{fix} \ x: e \rrbracket(\rho) && \diamond
\end{aligned}$$

Bei einer modularen Sprachbeschreibung kommen *definierende Transformationen* hinzu, mit denen die erweiterten Sprachkonstrukte auf eine bereits definierte Kernsprache zurückgeführt werden. Dazu zählen nicht nur die Einführung syntaktischer Varianten wie der *sequentielle Disjunktion*

$$c_1 \text{ orseq } c_2 \approx \text{if } c_1 \text{ then true else } c_2 \text{ fi,}$$

sondern auch die Definition neuer Sprachkonstrukte wie zum Beispiel die Zurückführung von Abschnitten auf Abstraktion und Fixpunktbildung:

$$\mathbf{begin} \ t \ x \equiv e_1; e_2 \ \mathbf{end} \approx ((t \ x) \mathbf{r}: e_2)(\mathbf{fix} \ x: e_1).$$

Weitere elementare Transformationen sind durch die Regeln über Prädikate und Mengen sowie durch die Axiome der Datenstruktur gegeben.

Im Gegensatz zu diesen elementaren Regeln stellen *abgeleitete Transformationen* kompakte Übergänge bereit, zum Beispiel zur Abkürzung häufig auftretender Sequenzen von Einzelschritten. Der Gültigkeitsbeweis erfordert im allgemeinen Induktionsmethoden wie Fixpunktinduktion, Parameterinduktion oder Kontextinduktion. Auch bei Datenstrukturen können aus den elementaren algebraischen Eigenschaften durch Induktionsverfahren wie strukturelle Induktion weitere Gesetze abgeleitet werden. Zum Beispiel läßt sich aus den beiden Zugriffsgesetzen für Stapel die *Dekompositionsregel*

$$i\text{empty}(s) \approx \text{false} \Rightarrow s \approx \text{prefix}(\text{first}(s), \text{rest}(s))$$

ableiten, die als Grundlage für den Algorithmenentwurf nach dem Strategie "Teile und Herrsche" dient.

## 4 Anwendungswissen

Für die erfolgreiche und zielgerichtete Programmentwicklung ist der *Kontext* wichtig, in dem eine Transformationsregel angewandt wird. Er beschreibt charakteristische Situationen, in denen die Anwendung der Transformation zu einer Verbesserung des Programms führt, etwa durch Vereinfachung des Ablaufs oder Verringerung des Speicherplatzbedarfs.

Während die lokale Umformung von Programmen durch Regeln angemessen repräsentiert wird, gehört das Anwendungswissen zu einem *Methodenkatalog*. Das Anwendungswissen zielt auf die Erkennung von Klassen algorithmischer Situationen ab, in denen eine bestimmte Regelanwendung erfolgreich ist. Es tritt in zwei Richtungen auf: zu einer Regel gehören ihre charakteristischen Anwendungskontexte, umgekehrt müssen in einem Programm die erfolgreich anzuwendenden Regeln erkannt werden. Der Unterschied zwischen

Regel- und Anwendungswissen wird an zwei Beispielen erläutert.

### 1. Beispiel Mit Hilfe des *Transitivgesetzes* ganzer Zahlen

$$\forall \text{int } x, y, z: x \leq y \wedge y \leq z \Rightarrow x \leq z$$

kann die Spezifikation zur Prüfung, ob ein Feld ganzer Zahlen sortiert ist,

$$\text{funct } \textit{issorted} \equiv ([1:n] \text{ array } a) \text{ bool: } \forall i, j \in [1, n]: i \leq j \text{ impl } a[i] \leq a[j]$$

so umgeformt werden, daß statt  $\mathcal{O}(n^2)$  nur die  $\mathcal{O}(n)$  Vergleiche unmittelbar benachbarter Elemente anfallen:

$$\text{funct } \textit{issorted} \equiv ([1:n] \text{ array } a) \text{ bool: } \forall i \in [1, n-1]: a[i] \leq a[i+1]$$

Die Regel selbst gibt eine algebraische Eigenschaft der Relation  $\leq$  auf ganzen Zahlen wider; in dieser Anwendung erlaubt sie algorithmisch die Lokalisierung von Vergleichsoperationen.

### 2. Beispiel Mit Hilfe der *Sequentialisierung der Disjunktion*

$$\textit{defined}(c_2) \Rightarrow c_1 \text{ or } c_2 \approx c_1 \text{ orseq } c_2$$

kann die Rechenvorschrift zur Prüfung, ob ein Element  $x$  in einer Sequenz  $s$  auftritt,

$$\begin{aligned} \text{funct } \textit{contains} \equiv (\text{sequ } s, \text{ int } x) \text{ bool:} \\ \text{if } \textit{isempty}(s) \text{ then } \textit{false} \text{ else } \textit{first}(s) = x \text{ or } \textit{contains}(\textit{rest}(s), x) \text{ fi} \end{aligned}$$

in repetitive Form überführt werden; zugleich wird für  $x \in s$  vorzeitige Terminierung erreicht:

$$\begin{aligned} \text{funct } \textit{contains} \equiv (\text{sequ } s, \text{ int } x) \text{ bool:} \\ \text{if } \textit{isempty}(s) \text{ then } \textit{false} \text{ else } \textit{first}(s) = x \text{ orseq } \textit{contains}(\textit{rest}(s), x) \text{ fi} \end{aligned}$$

Die Regel selbst beschreibt nur die in den folgenden Tafeln hervorgehobene Wertverlaufsgleichheit zwischen Disjunktion und sequentieller Disjunktion:

or	$\perp$	$T$	$F$
$\perp$	$\perp$	$\perp$	$\perp$
$T$	$\perp$	<b><math>T</math></b>	<b><math>T</math></b>
$F$	$\perp$	$T$	<b><math>F</math></b>

orseq	$\perp$	$T$	$F$
$\perp$	$\perp$	<b><math>\perp</math></b>	<b><math>\perp</math></b>
$T$	$T$	<b><math>T</math></b>	<b><math>T</math></b>
$F$	$\perp$	$T$	<b><math>F</math></b>

In obigem Kontext ändert ihre Anwendung den Rekursionstyp und verkürzt den Berechnungsablauf.

## 5 Strategiewissen

Das lokale Umformen eines Programms ist wenig aussichtsreich, wenn nicht die Gesamtentwicklung — einer Strategie folgend — auf ein Ziel ausgerichtet ist.

*Entwicklungsziele* bei Kontrollstrukturen sind die Vereinfachung des Berechnungsablaufs und die Verringerung des Berechnungsaufwands, zum Beispiel die Überführung exhaustiver Verfahren in zielgerichtete Algorithmen, die Vereinfachung der allgemeinen Rekursion zur repetitiven Rekursion oder das Vermeiden wiederholter (Teil-)Berechnungen. Entwicklungsziele bei Datenstrukturen sind die effiziente Realisierung der Grundoperationen und die kompakte Repräsentation von Daten, zum Beispiel durch gemeinsame Benutzung von Teilobjekten (“*sharing*”) oder durch selektive Änderung (“*selective updating*”).

Zur Erreichung dieser Ziele haben sich verschiedene *Entwicklungsstrategien* herauskristallisiert: Die Strategie “Teile und Herrsche” führt zu rekursiven Algorithmen, indem zwischen dem einfachen, unmittelbar zu lösenden Fall und dem zusammengesetzten Fall unterschieden wird, dessen Lösung sich aus der Kombination der Teillösungen ergibt. Die Strategien “Funktionskombination” und “Funktionskomposition” verschmelzen ähnliche, parallel bzw. sequentiell ablaufende Teilberechnungen zu einer Gesamtberechnung. Die Entwicklungsstrategie “Formales Differenzieren” verringert den Berechnungsaufwand, indem statt der wiederholten Auswertung eines zusammengesetzten Ausdrucks dessen Wert auf einem Parameter mitgeführt und der neue Wert inkrementell aus dem vorausgehenden berechnet wird. Die Entwicklungsstrategie “Tabellieren” vermeidet die wiederholte Berechnung einer Funktion für das gleiche Argument, indem die errechneten Funktionswerte in einer geeigneten Datenstruktur gespeichert werden.

Das Programmierwissen über Entwicklungsstrategien hat also die erfolgversprechende Verknüpfung von Einzelschritten zum Erreichen des Ziels zum Inhalt. Die ursprünglich zur Beschreibung von Strategien vorgeschlagenen regulären Ausdrücke über Regeln sind dazu nur ein erster Schritt.

Als Beispiel wird die Auswertung eines Polynoms  $p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$  mit ganzzahligen Koeffizienten behandelt, die als Sequenz  $a = \langle a_n, a_{n-1}, \dots, a_0 \rangle$  vorliegen. Die spezifikationsnahe, nach der Strategie “Teile und Herrsche” entwickelte Lösung

```

funct eval  $\equiv$  (sequ a, int x) int:
    if isempty(a) then 0 else first(a)  $\times$  x|a|-1 + eval(rest(a), x) fi

```

ist nicht repetitiv und berechnet die Potenzen  $x^i$  getrennt. Ziel ist deshalb eine repetitive, die getrennte Berechnung der Potenzen  $x^i$  vermeidende Rechenvorschrift. Ausgehend von der Verallgemeinerung

```

funct horner  $\equiv$  (sequ a, int x, s) int: s  $\times$  x|a| + eval(a, x)

```

erhält man nach der Strategie “Expandieren, Umformen und Komprimieren” durch die Ableitung ( $a \neq \langle \rangle$ )

$horner(a, x, s)$	{Expandieren von <i>horner</i> }
$\approx s \times x^{ a } + eval(a, x)$	{Expandieren von <i>eval</i> }
$\approx s \times x^{ a } + first(a) \times x^{ a -1} + eval(rest(a), x)$	{Algebra}
$\approx (s \times x + first(a)) \times x^{ a -1} + eval(rest(a), x)$	{Komprimieren von <i>horner</i> }
$\approx horner(rest(a), x, s \times x + first(a))$	

die repetitive Rechenvorschrift (“Horner-Schema”)

```

funct horner  $\equiv$  (sequ a, int x, s) int:
    if isempty(a) then s else horner(rest(a), x, s  $\times$  x + first(a)) fi;

```

die Einbettung lautet  $eval(a, x) = horner(a, x, 0)$ .

Der Schlüssel zum erfolgreichen Anwenden der Transformationsstrategie "Expandieren, Umformen, Komprimieren" liegt im Finden einer geeigneten *Einbettung*. Dabei verallgemeinert man eine Aufgabenstellung zu einem algorithmisch leichter lösbaren Problem, das die ursprüngliche Aufgabe als Spezialisierung enthält. Als Generalisierung wird häufig der Definitionsbereich von Rechenvorschriften durch Parametrisierung von Konstanten erweitert oder eine Datenstruktur auf Sequenzen  $x \mapsto \langle x \rangle$  oder Mengen  $x \mapsto \{x\}$  ausgeweitet.

## 6 Schlußbemerkung

Erst eine Formalisierung von Programmierwissen ermöglicht eine Rechnerunterstützung der Programmierung (/15/). Bei der Technik der transformationellen Programmierung (für einen Überblick siehe /11/) fallen dem Rechner die "mechanischen" Aufgaben zu; dazu gehören die Anwendung von Regeln auf Programme, die Verifikation syntaktischer Anwendbarkeitsbedingungen, die Bereitstellung einer Transformationsdatenbank sowie die Buchführung über den Entwicklungsprozeß, insbesondere zum Rücksetzen bei Sackgassen. Vom einem fortgeschrittenen Transformationssystem (/5/) erwartet man zudem Unterstützung bei der Verifikation semantischer Anwendbarkeitsbedingungen und die Auswertung intelligenter Anfragen an die Transformationsdatenbank, bis hin zum Aufzeigen von Transformationsstrategien.

Dem Programmierer verbleiben die "kreativen" Tätigkeiten in der Programmierung. Dazu zählen die Formalisierung der Problemstellung, die Auswahl von Transformationsregel und Anwendungsstelle, die Suche nach erfolgversprechenden Entwicklungsstrategien, die Ableitung neuer Transformationsregeln sowie die Formalisierung immer mächtigerer Strategien.

Gegenüber dem "direkten Herunterprogrammieren" ist die Ableitung eines Programms aus seiner Spezifikation aufwendig; auch ist sie für größere Programme ohne Systemunterstützung kaum praktikabel. Allerdings erhält man ein nachweislich korrektes Programm nicht umsonst. Wie könnte ein Informatiker die Korrektheit seines "Produkts" Algorithmus gegenüber der Spezifikation als "Kontrakt zum Auftraggeber" (/3/) besser belegen als durch die *konstruktive Entwicklung* des Programms aus seiner Aufgabenstellung?

**Danksagung.** Die Arbeit stützt sich auf vielfältige Ergebnisse des Münchner CIP-Projekts (computer-aided, intuition-guided programming); als aktuelle Überblicksarbeit sei /6/ empfohlen.

### Literatur

- /1/ Backus, J.: Can Programming Be Liberated from the von Neumann Style? A Functional Style and its Algebra of Programs. Communications ACM 21, 613-641 (1978)
- /2/ Bauer, F.L.: Programming as an Evolutionary Process. Proceedings Second International Conference on Software Engineering. San Francisco 1976, S. 223- 234

- /3/ Bauer, F.L.: Programming As Fulfillment of a Contract. In: Henderson, P. (ed.): Infotech State of the Art Report 9:6. Maidenhead: Pergamon Infotech 1981, S. 165–174
- /4/ Bauer, F.L., Berghammer, R., Broy, M., Dosch, W., Geiselbrechtinger, F., Gnatz, R., Hangel, E., Hesse, W., Krieg-Brückner, B., Laut, A., Matzner, T., Möller, B., Nickl, F., Partsch, H., Pepper, P., Samelson, K. †, Wirsing, M., Wössner, H.: The Munich project CIP. Volume I: The Wide Spectrum Language CIP-L. Lecture Notes in Computer Science 183. Berlin: Springer 1985
- /5/ Bauer, F.L., Ehler, H., Horsch, A., Möller, B., Partsch, H., Paukner, O., Pepper, P.: The Munich Project CIP. Volume II: The Transformation System CIP-S. Lecture Notes in Computer Science 292. Berlin: Springer 1987
- /6/ Bauer, F.L., Möller, B., Partsch, H., Pepper, P.: Formal Program Construction By Transformations — Computer-Aided, Intuition-Guided Programming. IEEE Transactions on Software Engineering 15:2, 165–180 (1989)
- /7/ Broy, M.: Algebraic Methods for Program Construction: The Project CIP. In: Pepper, P. (ed.): Program Transformations and Programming Environments. NATO ASI Series F: Computer and System Sciences, Vol. 8. Berlin: Springer 1984, S. 199–222
- /8/ Broy, M., Pepper, P.: Program Development as a Formal Activity. IEEE Transactions on Software Engineering, vol SE-7, 10–22 (1982)
- /9/ Delgado Kloos, C., Dosch, W.: Transformational Development of Circuit Descriptions For Binary Adders. In: Broy, M., Wirsing, M. (Hrsg.): Methodik des Programmierens — Eine Festschrift zu Ehren von F.L. Bauer. Passau: Fakultät für Mathematik und Informatik, Bericht MIP 8915, 1989, S. 99–117. Auch: Lecture Notes in Computer Science. Berlin: Springer 1991 (to appear)
- /10/ Dosch, W.: Functional Programming. In: De Blasi, M., Luque, E., Scerri, E. (eds.): Education and Application of Computer Technology. Proc. Third Biennial Meeting of the CMU on Microcomputers and Their Applications. Bari: Edizione Fratelli Laterza 1990, S. 457–469
- /11/ Feather, M.S.: A Survey and Classification of Some Program Transformation Approaches and Techniques. In: Meertens, L.G.L.T. (ed.): Program Specification and Transformation. Amsterdam: North-Holland 1987, S. 165–198
- /12/ Goguen, J.A., Thatcher, J.W., Wagner, E.G., Wright, J.B.: Initial Algebra Semantics and Continuous Algebras. Journal ACM 24:1, 68–95 (1977)
- /13/ Partsch, H.: Program Specification and Program Transformation — A Formal Approach to Software Development. Texts and Monographs in Computer Science. Berlin: Springer 1990
- /14/ Partsch, H., Möller, M.: Konstruktion korrekter Programme durch Transformation. Informatik-Spektrum 10, 309–323 (1987)
- /15/ Partsch, H., Steinbrüggen, R.: Program Transformation Systems. ACM Computing Surveys 15:3, 199–236 (1983)
- /16/ Wirth, N.: Program Development by Step-wise Refinement. Communications ACM 14, 221–227 (1971)

Prof. Dr. Walter Dosch	Telefon (0821) 528–2170
Institut für Mathematik	Telefax (0821) 528–2200
Universität Augsburg	Telex 53830
Universitätsstraße 2	e-mail dosch@uni-augsburg.de
D–8900 Augsburg	

# Beschreibung von Bürodokumenten durch kontextsensitive Merkmale

*Ralph Holbein, Bernd Teufel*

*Bereich Büroautomation*

*FAW Ulm*

*Postfach 2060*

*D-7900 Ulm*

**Zusammenfassung.** In diesem Beitrag wird ausgehend von einer Anforderungsanalyse ein neuer Ansatz entwickelt, der eine Unterstützung der Integration der unterschiedlichen Anwendungsbereiche im Büro erlaubt. Der Ansatz repräsentiert ein Beschreibungsverfahren für Dokumente, welches entsprechende Zusatzinformationen - abgesehen vom eigentlichen Dokumentinhalt und den herkömmlichen administrativen Merkmalen - in die verschiedenen Anwendungen einbringt und so bei der Bearbeitung und Verwaltung von Dokumenten unterstützend wirkt. Das entwickelte Beschreibungsverfahren wird in seinen Grundzügen beschrieben.

**Abstract:** In this paper a new approach - based on a requirements analysis - is introduced supporting the integration of the different fields of application in an office. The approach represents a method of description for documents which provides appropriate information - additionally to the actual content of the document and the usual administrative features - for the various applications, thus, supporting processing and management of documents. The fundamentals of the developed method of description is introduced.

## 1. Bürokommunikation

Umfangreiche Erfahrungen aus dem Bereich Industrie und Verwaltung zeigen klar eine Verlagerung der derzeitigen Planungen, und zwar weg von eindeutig abgrenzbaren Einzelproblemen, wie Textverarbeitung, isolierter PC-Einsatz oder individuelle Datenbankanwendungen, und hin zu einer integrierten Lösung, d.h. zu einer durchgängigen Unterstützung von komplexen arbeitsplatzübergreifenden Prozessen. Eine solche Unterstützung muß natürlich für den Benutzer transparent sein. Hier setzt ein, was in dem vorgestellten Kontext unter einer integrierten Bürokommunikation verstanden werden soll: die auf allen Ebenen durchgängige Unterstützung von Tätigkeiten in einer Organisation. Dabei muß sich die Integration auf folgende Bereiche erstrecken: die organisatorische Integration, die Anwendungsintegration, die technische Integration.

Die *organisatorische Integration* steht für die *ganzheitliche Bearbeitung von Bürovorgängen*. Hier ist es unabdingbar, daß Wissen über die Organisationsstruktur wie auch beispielsweise Organisationsrichtlinien einfließt. Bürovorgänge zeichnen sich durch ihren Prozesscharakter aus, worin sich die Abfolge von Teilaufgaben innerhalb des Vorganges widerspiegelt. Die einzelnen Teilaufgaben benötigen Informationen und erzeugen in der Regel auch Informationen, die größtenteils durch Dokumente repräsentiert werden. Damit ist die durchgängige - und transparente - Verfügbarkeit von Information in und über Dokumente hier zentral.

Die *Anwendungsintegration* steht für die *Durchgängigkeit der einzelnen Anwendungen*, d.h. daß

unterschiedliche Anwendungen aufeinander abgestimmt sind. Funktionale Integration, im Sinne der Verfügbarkeit von entsprechenden Schnittstellen zur Kooperation und Kommunikation der verschiedenen Dienste bzw. Funktionen, wie auch eine integrierte Benutzeroberfläche sind hier zu nennen. Mit letzterem ist gemeint, daß die Gesamtheit der Bürofunktionen über eine standardisierte Benutzeroberfläche, d.h. über eine nach einheitlichen Prinzipien aufgebaute Bedienschnittstelle verfügbar ist. Die gleichartige Bedienlogik wie auch die bekannten Ansätze der objektorientierten Oberflächen sind hier Stichworte.

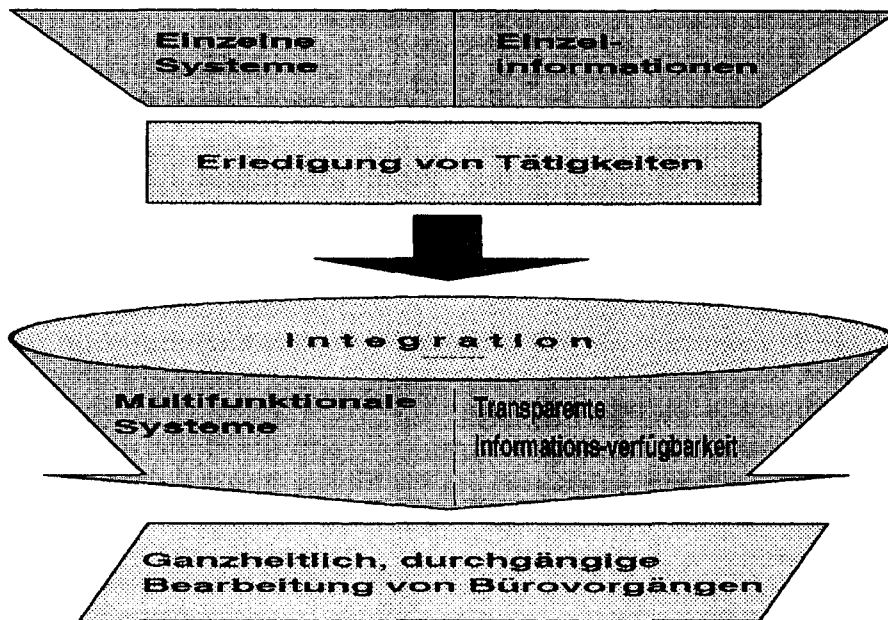


Abb. 1: Integrierte Bürokommunikation

Hinter der *technischen Integration* schließlich verbergen sich *Basiselemente* wie Netzwerke, LAN, Bridges, Gateways bis hin zu Betriebssystemen und integrierten Datenbasen, auf die der Zugriff aus allen Anwendungen oder Funktionsmoduln möglich sein muß.

Es zeigt sich klar ein Trend, wie er in Abbildung 1 zusammengefasst ist: weg von der Sicht der Einzelsysteme und Einzelinformationen und der mehr oder weniger separaten Erledigung von Tätigkeiten, hin zu der ganzheitlich durchgängigen Bearbeitung von Bürovorgängen. Dabei greift hier der Integrationsgedanke in dem vorgenannten Sinne durch die organisatorische Integration, die Anwendungsintegration und die technische Integration.

## 2. Schwerpunkte der Automation

Unter der Integrationsprämisse wie sie in Kapitel 2 dargestellt wurde können drei Schwerpunkte der Automation im Einsatzfeld Büro ausgemacht werden: die integrierte Vorgangsbearbeitung, die Unterstützung der Kommunikation und vor allem die Unterstützung in der Dokumentenverwaltung.

Diese Aufzählung ist nicht nur von akademischem Wert, sondern bestätigt sich in realen Umgebungen. Bei der Vorgangsbearbeitung ist es natürlich von entscheidender Bedeutung, Büro-



vorgänge zu koordinieren und zu kontrollieren. Es wurde bereits darauf hingewiesen, daß Büro- vorgänge sich in Teilaufgaben untergliedern lassen, wobei hierbei nicht immer nur an starre Abhängigkeiten zu denken ist, sondern es muß die Möglichkeit einer kontextsensitiven, d.h. situationsabhängigen Unterstützung kooperativer Abläufe geboten werden, wie es beispielsweise im Rahmen der FBS-Philosophie vorgeschlagen ist [KARA 89]. Es muß die Möglichkeit bestehen, abhängig von der momentanen Situation und den bereits durchgeführten Tätigkeiten bzw. Entscheidungen den Verlauf innerhalb eines Vorganges in einem bestimmten Rahmen alternativ fortzuführen. Da, wie ebenfalls bereits erwähnt, bei der Ausführung von Aufgaben innerhalb eines Vorganges Informationen in Form der Dokumentenverarbeitung verbraucht und/oder erzeugt werden, müssen Dokumente über ausreichendes Wissen verfügen, um den Vorgangsablauf zu unterstützen. Hierzu wird die im folgenden vorgestellte erweiterte Dokumentenbeschreibung einen entsprechenden Beitrag leisten.

Zur Kommunikationsunterstützung: Büroarbeit bedeutet in aller Regel kooperative Arbeit. Kooperation an sich steht schon für ein Höchstmaß an Kommunikation und Austausch von Informationen. Das heißt, daß hier eine entsprechende Unterstützung geboten werden muß. Dies kann im einfachsten Fall die automatische Adressierung und Verteilung von Information bedeuten und kann bis zur Unterstützung bei Abstimmungsproblemen gehen.

Die Dokumentenverwaltung schließlich ist von zentraler Bedeutung. Dokumente sind, wie Krönert es formuliert, die Hauptinformationsträger in Unternehmen oder Organisationen [KRÖN 88], was ja auch bei den zuvor gemachten Ausführungen deutlich wurde: die dokumentenorientierte Vorgangssteuerung wie auch der dokumentenbasierte Informationsaustausch. Unter der Dokumentenverwaltung ist einzuordnen: die Ablage von Dokumenten (d.h. den "darin enthaltenen" bzw. den "daran angehängten" Informationen), die Formulierung von Informationsbedürfnissen, die Verfügbarmachung von Information oder Dokumenten.

Natürlich ist mit der Ablage von Dokumenten auch der Aspekt der Indexierung zu verbinden. Dies wird für die Forschung um so wichtiger, als neue Ansätze im Hypermedia-Bereich auch neue Ansätze zur Dokumentenbeschreibung erfordern. Sehr wichtig ist auch die Betrachtung der Formulierung von Informationsbedürfnissen. Hier ist es wichtig, daß in Zukunft der Informationssuchende besser als bisher unterstützt wird, z.B. durch graphische Möglichkeiten, oder die Möglichkeit in Informationsstrukturen zu navigieren. Hierbei sind die Untersuchungen von Schimdt zu berücksichtigen, wo eine Dreiteilung der Fragestruktur im Bürobereich aufgezeigt wurde [SCHM 91]: exakte Fragen, teilweise unscharfe Fragen und unscharfe Fragen.

Schließlich muß ein entsprechend spezifiziertes Informationsbedürfnis auch die Verfügbarmachung, d.h. das Retrieval, von Information bzw. Dokumenten nach sich ziehen. Neben neuen Retrieval-Ansätzen im Multimedia-Bereich (siehe z.B. [FREI 91]) können hier auch neue Ansätze im Textbereich, wie z.B. Informationsspuren [TEUF 90], zum tragen kommen.

Das hier betrachtete Konzept *integrierte Bürosysteme* stellt sich damit wie in Abbildung 2 skizziert dar.

Werden die verschiedenen Bereiche in einer Unternehmung betrachtet, so ist bekannt, daß die Tätigkeiten innerhalb der Bereiche, wie auch die Vorgänge über die Bereichsgrenzen hinweg, Dokumente involvieren, d.h. erzeugen und verbrauchen. Eine geeignete Dokumentenbeschrei-

bung, die nicht nur zuvorgenannte Retrievalfunktionen unterstützt, sondern auch als eine Art Vorgangstrigger wirken kann (durch die Bereitstellung ablaufsteuerender Information), ist damit als *ein* Integrationsfaktor für ein integriertes Bürokommunikationssystem auffaßbar.

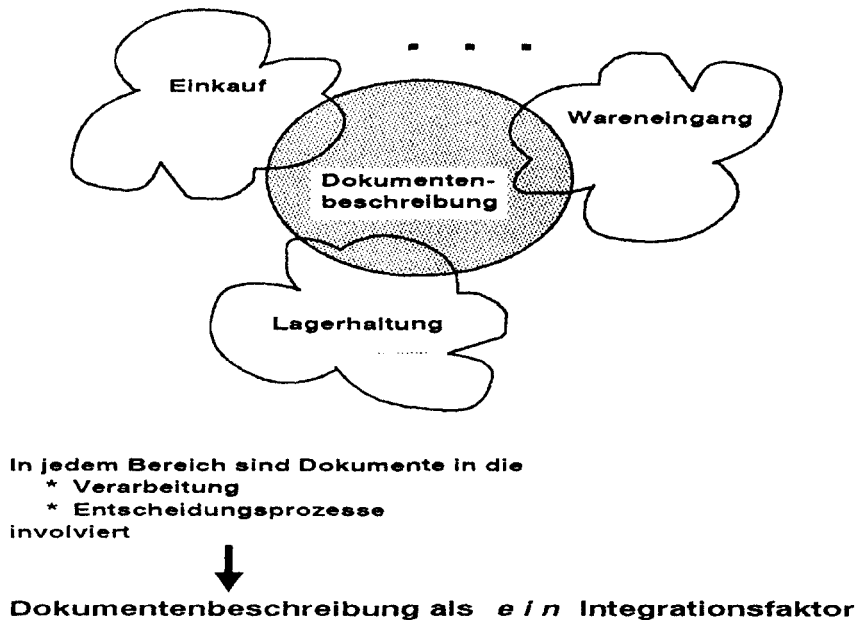


Abb. 2: Konzept integrierte Bürosysteme.

Offensichtlich kann hier nicht nur die logische Ebene der organisatorischen Einheiten betrachtet werden, sondern es müssen auch beteiligte Basissysteme wie ein Kommunikationssystem, ein Vorgangssystem oder ein Filing und Retrieval-System mit einbezogen werden. Jedes dieser Systeme nutzt entsprechende Merkmale aus der Dokumentenbeschreibung, um die gewünschten Aktivitäten durchzuführen. Das Kommunikationssystem kann von "Weiterleitungsmerkmalen" profitieren, während das Vorgangssystem z.B. von Merkmalen profitiert, die historisches Wissen darstellen [HOLB 91]. Also auch auf dieser Ebene kann die Dokumentenbeschreibung als einer der Integrationsfaktoren in einem Bürokommunikationssystem verstanden werden.

### 3. Dokumente

In den vorhergehenden Kapiteln hat der Begriff des Dokumentes bereits breite Anwendung gefunden. Es ist daher notwendig hier zu einer Präzisierung des Begriffs zu gelangen.

Wird von einer neutralen, d.h. vorgangsunabhängigen Objektdefinition beim Entwurf eines unternehmensweiten Datenmodells ausgegangen, so ergeben sich die Entitäten entsprechend den in der Unternehmung auftretenden Begrifflichkeiten (z.B. Kunde, Rechnung, etc.). Diese sind Informationsobjekte, allgemein eine Sache, ein Ort, eine Person, ein Ereignis oder ein Konzept über welches das Unternehmen Informationen speichert [NOBS 89]. Wird allerdings nun die Dynamik des Unternehmens bezgl. dem Erreichen von unternehmerischen Zielen betrachtet, so werden Aufgabenkomplexe erkennbar, die nach bestimmten Regeln zusammenspielen, d.h. Vorgänge. Diese Aufgaben und Vorgänge bilden nun den Kontext und die Basis für das Entste-

hen von Dokumenten. D.h. bestimmte Objekte, welche zunächst vorgangsunabhängig definiert sind, werden innerhalb eines realen Kontextes betrachtet und erhalten dadurch einen Rahmen, in welchem ihnen konkrete Bedeutungen und Beziehungen zukommen. Mit Dokument sind jetzt informationelle Einheiten im Verständnis von Schriftstücken gemeint, die in einem derartigen Kontext als vollständig relevante Einheiten auftauchen - grundsätzlich: "... als Einheit übermittelbare Informationsmenge zur menschlichen Wahrnehmung" [BORM 90]. Die prinzipielle Aufteilung dieser informationellen Einheiten, dieser Dokumente, ist durch die Beschreibung, den Index, und den eigentlichen Dokumenteninhalt gegeben.

### 3.1 Büروفunktionen an Dokumenten

Die Hauptfrage, die sich nun bezüglich Dokumenten stellt, betrifft die Verwendung dieser informationellen Einheiten in einer Organisation. Bei der Analyse der Bearbeitungsarten, welche zur Aufgabenerfüllung an Dokumenten durchgeführt werden können, sind die Gruppen Verteilung, Bearbeitung und schließlich die Ablage (bzw. Filing und Retrieval, F&R) erkennbar (siehe Abb. 3).

		F & R	Verteilung	Bearbeitung				
				1	2	3	4	5
Ablage Zwischenablage Suche		*						
		*	*					
		*						
				=> Folgen ergeben sich (Bearbeitungsschritt hat pragmatischen Charakter, der über den eigentlichen Vorgangsablauf hinausgeht)				
Eingang Vorlage Weiterleiten Ausgang			*					
			*					
			*					
			*					
Erstellen Überarbeiten Ergänzen								*
								*
								*
Einsehen	"Antrag"			Z				
	"Bestät."				Z	Z	++	*

1. Genehmigen 2. Bestätigen 3. Kenntnisnahme  
4. Prüfen 5. Allgemein

Abb. 3: Büروفunktionen an Dokumenten.

Ein Dokument gelangt in die Organisation, es muß in der Poststelle verteilt werden, es muß bearbeitet werden, wozu dann beispielsweise auch schon die Registration des Dokumentes zählt, es muß dann eventuell mehrmals weitergeleitet bzw. weiterverteilt werden, jeweils mit entsprechenden nachfolgenden Bearbeitungsschritten, bevor es schließlich abgelegt wird, um zu einem späteren Zeitpunkt eventuell wieder gesucht werden zu können. Bei den Untersuchungen diesbezüglich am Forschungsinstitut für anwendungsorientierte Wissensverarbeitung (FAW Ulm) haben sich als Bearbeitungsarten Genehmigen, Bestätigen, Kenntnisnahme und Prüfen als zu-

sammengesetzte Formen (einsehen in Verbindung mit abzeichnen bzw. überprüfen) sowie allgemeine Bearbeitungsarten (atomar: erstellen, überarbeiten, ergänzen, einsehen) ergeben.

Beispielsweise muß ein Dienstreiseantrag genehmigt werden, während ein Bewerbungsschreiben bestätigt wird (im Sinne der Eingangsbestätigung), gleiches gilt z.B. für Aufträge. Eine Kolloquiumsankündigung beispielsweise dient nur der Kenntnisnahme, während Lieferscheine oder Rechnungen geprüft werden müssen. Im Filing und Retrieval-Bereich ist die eigentliche Ablage, sowie die Zwischenablage und damit verbunden auch die Wiedervorlage als Verteilungsaspekt, sowie die Suche nach Dokumenten angesiedelt. In den Bereich der Verteilung fällt der Eingang, die Vorlage, das Weiterleiten, sowie der Ausgang von Dokumenten.

### 3.2 Funktionale Dokumentklassifikation

Neben der Bestimmung der Bürofunktionen an Dokumenten, die im übrigen nicht FAW-spezifisch sind, ist es erforderlich, eine Grobklassifikation der Dokumente durchzuführen. Durch eine solche Klassifikation wird eine Unabhängigkeit von der konkret betrachteten Organisation - im hier vorliegenden Falle also vom FAW - erreicht. Das FAW als Forschungsinstitut hat natürlich einen gewissermaßen eingeschränkten Dokumentenvorrat, was sich auch in der funktionalen Dokumentklassifikation niederschlägt; jedoch kann die gewählte Klassifikation leicht für andere Organisationen erweitert werden.

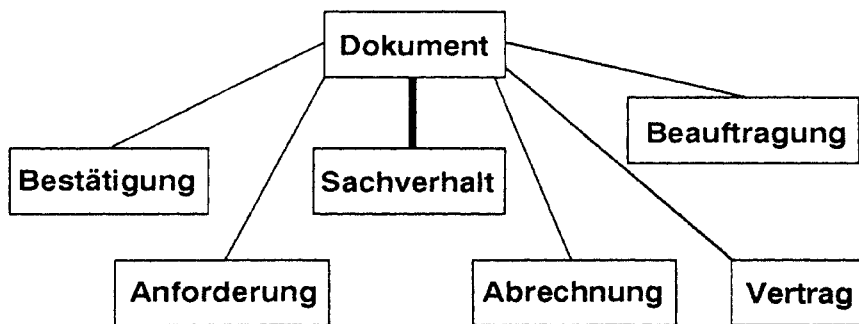


Abb. 4. Funktionale Dokumentklassifikation.

Die hier gewählte Klassifikation (siehe Abb. 4) orientiert sich an Ähnlichkeiten im ablauftechnischen Umfeld der Dokumente, d.h. es wird mehr oder weniger die Dokumentverwendungsart zur Klassifikation herangezogen. Bei der Erhebung am FAW sind 60 verschiedene Dokumente bestimmt und dann an Hand funktionaler Kriterien klassifiziert worden.

Am FAW haben sich 6 Hauptklassen herausgebildet: Bestätigungen, wie Eingangsbestätigungen oder Teilnahmebestätigungen; Anforderungen, wie z.B. für Sachmittel in Form von Anträgen für Hardware- oder Software-Beschaffung; Abrechnungen, wie beispielsweise Reisekostenabrechnungen, Vortragshonorare etc.; Beauftragungen, wie etwa Buchbestellungen in der Bibliothek oder generell Aufträge für eine bestimmte Leistungserstellung; Verträge, wie Personalverträge, Versicherungen oder Projektverträge; Mitteilungen, wie Fachinformationen oder allgemeine Mitteilungen in Form von Umläufen, Vortragsankündigungen oder Workshop-Programme. Abb. 5 zeigt die Modellierung der Dokumentenklassen auf dem Rechner.

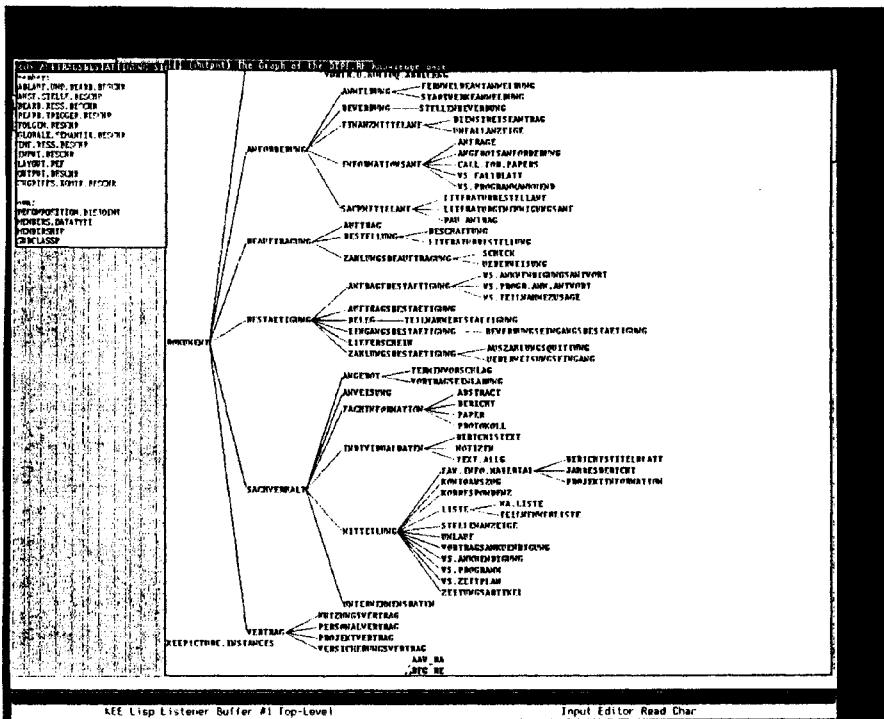


Abb. 5: Modellierung der Dokumentenklassen.

#### 4. Merkmalsgruppen

Im vorhergehenden Abschnitt wurden die Dokumente klassifiziert und die Bürofunktionen in Verbindung mit Dokumenten bestimmt. In diesem Abschnitt steht nun die Indexierung der Dokumente (in Form der Zuordnung von Merkmalen zu Dokumenten) im Mittelpunkt.

Was hier nicht weiter verfolgt werden soll, sind Beschreibungsmöglichkeiten, wie sie beispielsweise aus dem klassischen Information Retrieval bekannt sind. Dies sind also z.B. die Zuordnung von Termen - seien sie nun gewichtet oder nicht - zu den Dokumenten (vergl. [SALT 83], [SALT 89], [VANR 79]), oder aber auch die Verwendung neuartiger Verfahren, die auf Informationsspuren basieren (vergl. [TEUF 89]). Außerdem sollen die Möglichkeiten, wie sie in der ODA-Norm [ISO 89] vorgesehen sind, also beispielsweise die Verwendung von administrativen Merkmalen, wie Erstellungsdatum, Autor, etc., als gegebene Basis betrachtet werden.

Die am FAW durchgeführten Untersuchungen basierten konkret auf der Frage: Gibt es Merkmale oder Merkmalsgruppen außer den bekannten IR-Merkmalen oder den administrativen ODA-Merkmalen, die zu einer Beschreibung der Dokumente herangezogen werden können? Ist also eine erweiterte Dokumentenbeschreibung möglich?

Dabei nützt es zu untersuchen, für welche Bereiche Merkmale als erweiterte Dokumentenbeschreibung überhaupt nutzbar sein können. Also sollen sie nur ein etwaiges Retrieval von Dokumenten unterstützen oder aber können diese Merkmale weitere Unterstützung bieten, also z.B. Vorgänge unterstützen oder deren Ablauf beeinflussen. Genau auf letzteren Punkt soll hier abgezielt werden, d.h. nicht nur die Unterstützung des Retrievals, sondern auch die Unterstüt-

zung (in Form eines Informationsträgers) für den Vorgangsablauf. Es wurden dazu die in Abbildung 6 gezeigten Merkmalsgruppen definiert.

1. Informationelle Ressourcen  
(im Dokument enthaltene Informationsquellen)
2. Sicherheitskonzept  
(Zugriffskontrolle)
3. Datenimport  
(konkreter Import von Daten)
4. Datenexport  
(konkreter Export von Daten)
5. Globale Semantik  
(z.B. die Historie betreffend)
6. Zuständigkeit  
(ausführende Stelle)
7. Bearbeitungsressourcen  
(Hilfsmittel, Kompetenzen)
8. Vorgangsablauf und Bearbeitungseigenschaften  
(Trigger für folgende Bearbeitungsschritte)
9. Zielrelevante Folgen  
(Folgen, die bzgl. der Bearbeitung zu beachten sind)

Abb. 6: Merkmalsgruppen.

Informationelle Ressourcen beschreiben die Quellen der im Dokument enthaltenen bzw. mit dem Dokument in Zusammenhang stehenden Informationen. Dies wird besonders bei Dokumenten relevant, die bestätigt oder genehmigt werden müssen, oder aber für Dokumente, welche als Entscheidungsgrundlage dienen. Eine Zugriffskontrolle für Dokumente muß natürlich vorhanden sein, um beispielsweise bereits genehmigte Dokumente nicht mehr schreibend verändern zu können. Datenimport oder Informationsimport liegt dann vor, wenn Daten aus einer konkret anzugebenden Quelle auf Grund bestimmter kontextabhängiger Zusammenhänge importiert werden. Der Unterschied zur informationellen Ressource ist also offensichtlich: hier liegt eine 1:1 Abbildung von Daten aus einem Dokument in ein anderes vor, während eine informationelle Ressource die Quelle darstellt, welche für das Zustandekommen von dargestellten Informationen die Basis ist. Diese Information muß aber nicht 1:1 abgelidet sein. Der Datenexport ist analog zu verstehen.

Die Merkmalsgruppe Globale Semantik beschreibt einen vom funktionalen Kontext eher unabhängigen Einsatzzweck eines Dokuments. Damit sind nicht die Eigenschaften gemeint, die durch die funktionale Dokumentklassifikation bereits abgedeckt sind, sondern z.B. die Herstellung des Versicherungsschutzes für den Reisenden bei der Genehmigung einer Dienstreise. In



## 5. Beschreibung einer Dokumentenklasse

Die im vorigen Abschnitt festgelegten Merkmale bilden nun die Basis für die erweiterte Dokumentenbeschreibung. Diese kann wie in Abbildung 8 gezeigt dargestellt werden.

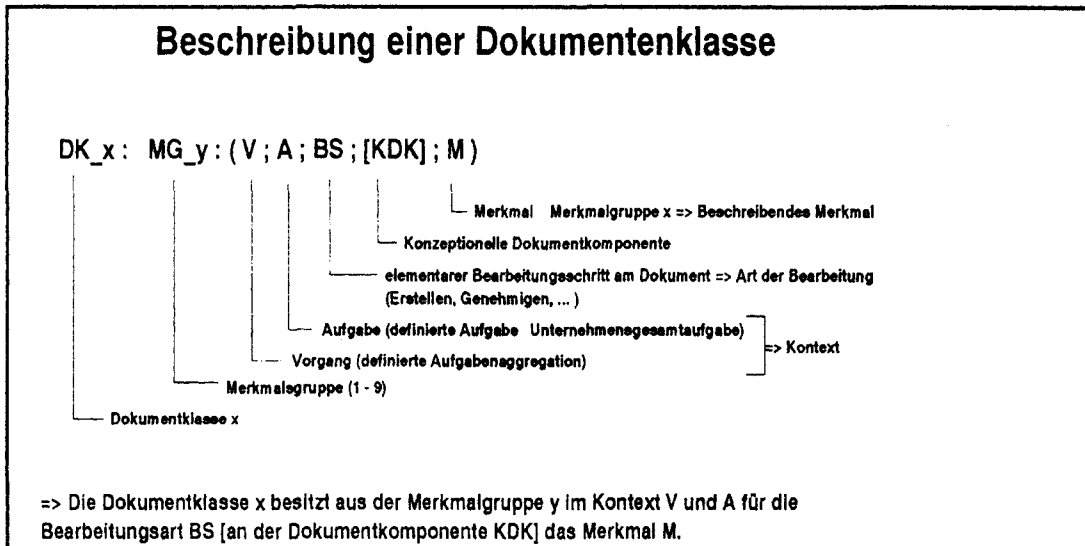


Abb. 8: Beschreibung einer Dokumentenklasse.

Einer Dokumentenklasse  $x$  und damit einer konkreten Instanz, d.h. einem Dokument dieser Klasse, sind Merkmale aus den verschiedenen Merkmalsgruppen zugeordnet sowie der Rahmen in dem diese Merkmale gültig sind. Dieser Rahmen besteht zum einen aus dem Vorgang (V) und der Aufgabe (A), wodurch der Kontext angegeben ist, des weiteren wird die Bearbeitungsart (BS), also "Erstellen", "Genehmigen" etc. angegeben. Optional ist es möglich den Gültigkeitsrahmen des Merkmals noch weiter zu präzisieren. Dies wird durch die Angabe einer konzeptionellen Dokumentkomponente (KDK) erreicht, die das Merkmal für die angegebenen V, A, BS auf einen als konzeptionellen Teil des Dokuments verstandenen Bereich bezieht. In Abbildung 9 ist zur Verdeutlichung ein Beispiel gegeben, wobei nicht für sämtliche 9 Merkmalsgruppen sinnvolle Merkmale bezüglich derselben V, A, BS angegeben werden können.

Im Bereich der Dokumentenklasse "Abrechnungen" wird das Reisekostenabrechnungsf formular betrachtet. Für die einzelnen Merkmalsgruppen ist der Kontext durch den Vorgang "Dienstreisen" und die Aufgaben "Erstattungsbearbeitung", "Zahlungsbetrag bestimmen" und "Auszahlung veranlassen" gegeben. Die Bearbeitungsschritte sind "Erstellen", "Ergänzen" und "Genehmigen"; die konzeptionelle Dokumentenkomponenten verweisen auf "Zahlungsempfänger" und "Auszahlungsbetrag". Die konkrete Merkmalsausprägung bei der informationellen Ressource wie auch beim Datenimport ist der Dienstreiseantrag - also sozusagen das Vorgängerdokument. Jedoch referenziert in der Merkmalsgruppe 1 der "Dienstreiseantrag" die Quelle der Genehmigung für die Dienstreise, während in der Merkmalsgruppe 3 aus dem Dienstreiseantrag konkrete Daten wie z.B. Antragsteller, Vergütungsgruppe, Projektzugehörigkeit oder Reiseziel übernommen werden. In der Merkmalsgruppe 8 wird in diesem Kontext z.B. auf ein Kostenabrechnungsregelwerk verwiesen, welches die Bestimmung der Kilometerpauschale bei Reisen mit dem privaten PKW erlaubt.



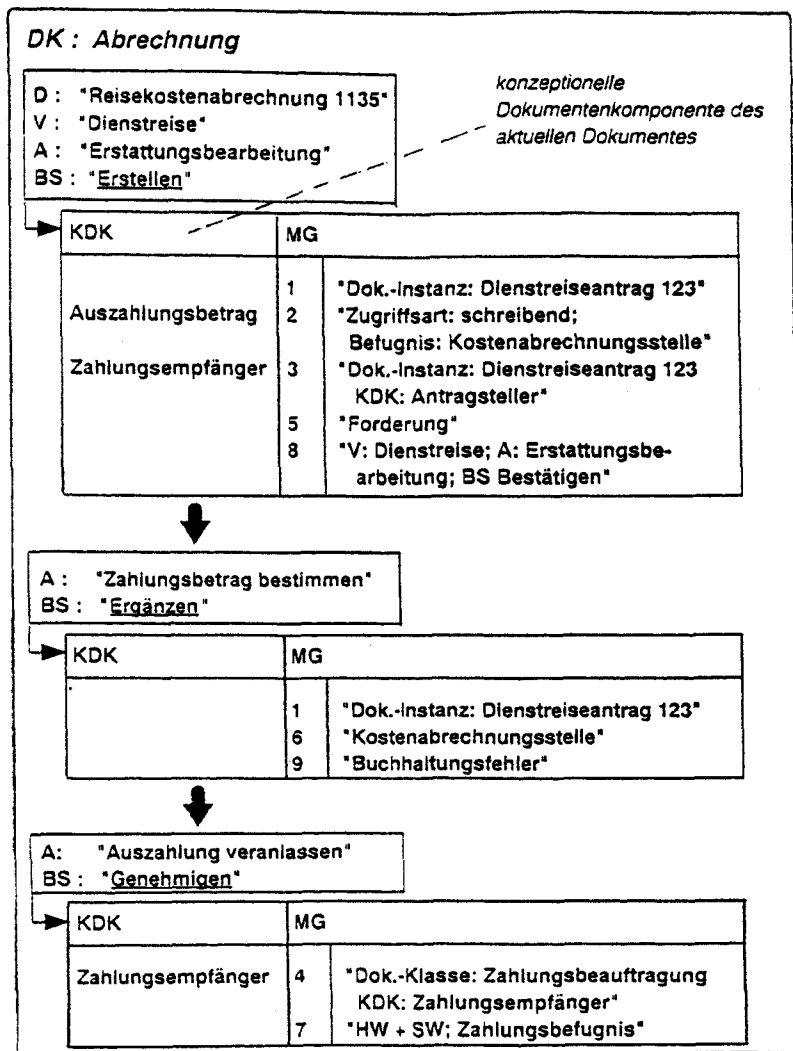


Abb. 9: Dokumentenbeschreibung

Werden die Merkmalsgruppen unter systemtechnischen Gesichtspunkten betrachtet, so zeigt sich, daß hier in der Regel eine referentielle Beschreibung vorliegt. Entweder wird eine Referenz auf den Dokumentenvorrat, d.h. auf konkrete Dokumente, gegeben (dies betrifft beispielsweise die Merkmalsgruppen 1, 3 und 4), oder aber es sind Verweise auf die Organisationsstruktur, die Funktionsstruktur, ein Vorgangsmo­dell bzw. Richtlinien und Regeln vorhanden. Es liegt also eine Situation vor, wie in Abbildung 10 gezeigt.

Das verwendete Dokumentenmodell enthält entsprechend der Darstellung in Abbildung 10 Beschreibungselemente z.B. gemäß ODA sowie eventuell Elemente, wie sie aus der klassischen Inhalt­erschließung fürs Information Retrieval bekannt sind. Ergänzt werden diese Beschreibungselemente durch organisatorische Merkmale, wodurch eine erweiterte Dokumenten­beschreibung vorliegt. Diese organisatorischen Merkmale referenzieren neben dem Dokumenten­vorrat einer Organisation deren Funktionsstruktur und Organisationsstruktur sowie die Vor­gangsbeschreibung der Organisation. In der Funktionsstruktur werden die in einer Unterneh-

mung auftretenden Tätigkeiten strukturiert dargestellt (Funktionenbaum). Die Organisationsstruktur zeigt die organisatorischen Abhängigkeiten der Unternehmung oder Organisation (Organigramm), während die Vorgangsbeschreibung die Struktur von Vorgängen unter Berücksichtigung von Organisationsrichtlinien und -Regeln darstellt. Zusammen stellen die Funktionsstruktur, die Organisationsstruktur und die Vorgangsbeschreibung das Wissen über die Organisation und deren Funktionalität dar.

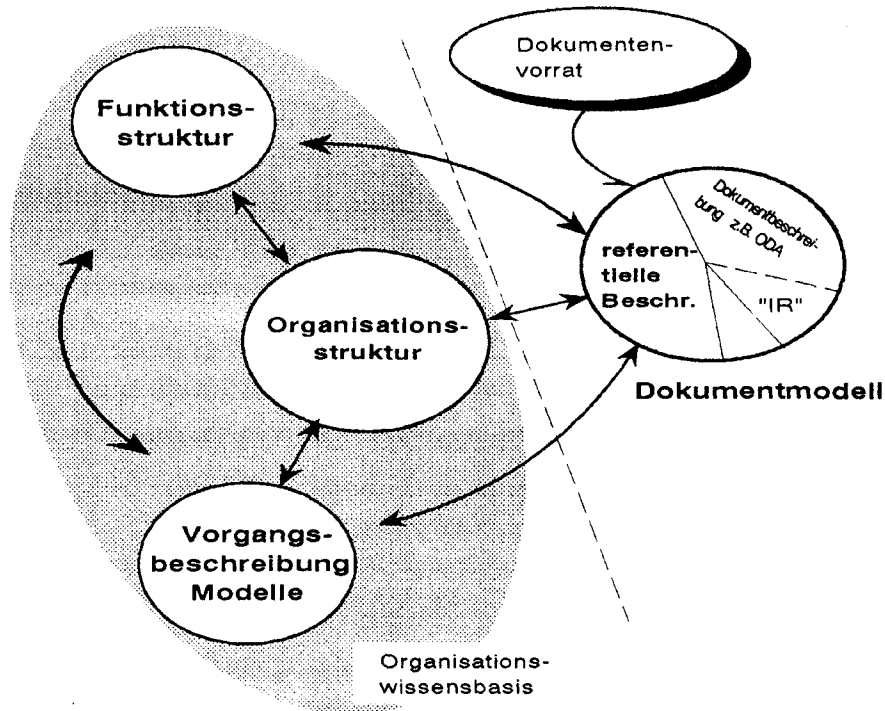


Abb. 10: Organisatorische Merkmale ergänzen die Dokumentenbeschreibung.

## 6. Schlußbemerkung

Durch die Umsetzung der entwickelten Konzepte innerhalb eines Prototyping, zur erweiterten Dokumentbeschreibung, konnten verschiedene Anwendungsmöglichkeiten, für ein Kontext-Retrieval von Dokumenten, und auch zur Unterstützung von Vorgangsabläufen durch diese Beschreibung, verdeutlicht werden. Es hat sich gezeigt, daß die einem Dokument (Instanz) konkret zugewiesenen Merkmale, bestens als Protokoll für eine Ablaufanalyse geeignet sind, wodurch Schwachstellen in der Ablauforganisation einer Unternehmung festgestellt werden können.

Grundsätzlich stellt der verfolgte Ansatz einen Versuch dar, Bürodokumente möglichst stark mit der Organisationsstruktur (Aufbau- und Ablauforganisation) zu verbinden, um diese mehrwerterzeugend in die funktionalen Komponenten zur Bearbeitung und Verwaltung von Dokumenten im Bürobereich, integrieren zu können. Die konkrete Untersuchung der zugrundegelegten Gedanken, sowie die dabei gewonnen Erkenntnisse, wurden interessiert aufgenommen und verdienen durchaus der Berücksichtigung für die Betrachtung rechnergestützter Erfassung und Beschreibung von Dokumenten.

Für eine Anwendung des entwickelten Ansatzes und der hierfür notwendigen vollständigen Modellierung der beschriebenen Systemkomponenten wurde allerdings das folgende Problem offensichtlich:

Bei einer (vorauszusetzenden) Dynamik innerhalb der Organisationsstruktur und der Komplexität der Modellierung für den dargestellten Ansatz sind sehr hohe Anforderungen an eine entsprechende Modellierungskomponente zu stellen. Dies bedeutet, daß ein Systemadministrator für die Organisationsmodellierung durch intelligente Werkzeuge unterstützt werden muß, um beliebige Strukturveränderungen in einer konsistenten Art und Weise durchführen zu können.

## Literatur

- [BORM 90] U. Bormann, C. Bormann. Offene Dokumentbearbeitung - Status und Weiterentwicklung. *Informationstechnik* 32, 1990.
- [FREI 91] H. P. Frei, P. Schäuble. Designing a Hypermedia Information System. *Proceedings DEXA '91*, Springer-Verlag, Wien, 1991.
- [HOLB 91] R. Holbein. Beschreibungsverfahren für Bürodokumente durch kontextsensitive Merkmale. Diplomarbeit, Universität Konstanz, 1991.
- [ISO 89] ISO 8613. *Information Processing - Text and Office Systems - Office Document Architecture (ODA) and Interchange Formats*. International Organization for Standardization, Geneva, 1989.
- [KARA 89] D. Karagiannis. Flexible Bürosysteme (FBS): Architektur und Einsatzmöglichkeiten. In: S. Fuhrmann, T. Pietsch (eds.). *Praktische Anwendungen moderner Bürotechnologien*. Band 12, Erich Schmitt Verlag, Berlin, 1989.
- [KRÖN 88] G. Krönert. Genormte Austauschformate für Dokumente. *Informatik-Spektrum* 11(2), 71 - 84, 1988.
- [NOBS 89] A. Nobs. Modernes Informationsmanagement im Dienst strategischer Unternehmensführung. *Computer Magazin Wissen, Sonderheft 101*, 1989.
- [SALT 83] G. Salton, M. J. McGill. *Introduction to Modern Information Retrieval*. McGraw-Hill, Tokyo, 1983.
- [SALT 89] G. Salton. *Automatic Text Processing*. Addison Wesley, Reading, 1989.
- [SCHM 91] S. Schmidt. Modellierung von Büroinformationssystemen - ein organisationsorientierter Ansatz. Dissertation, Universität Zürich, 1991.
- [TEUF 89] B. Teufel. *Informationsspuren zum numerischen und graphischen Vergleich von reduzierten natürlichsprachlichen Texten*. vdf-Verlag, Zürich, 1989.
- [TEUF 90] B. Teufel. Informationsspuren - Perspektive für Textvergleich und Visualisierung. *Proceedings ISI '90*, Konstanz, 1990.
- [VANR 79] C. J. van Rijsbergen. *Information Retrieval*. Butterworth, London, 1979.

# A Prototype for Integrating Probabilistic Fact and Text Retrieval

Norbert Fuhr

Thorsten Hoffmann

## Zusammenfassung

Wir stellen einen Prototypen für ein Informationssystem vor, das Text- und Faktenretrieval integriert. Eine Anfrage besteht aus einer Menge von Bedingungen, wobei eine Bedingung sich entweder auf den Text oder einen Attributwert eines Datenbankobjektes bezieht. Bedingungen können sowohl bezüglich der Anfrage als auch eines Objektes gewichtet sein, woraus das System als Antwort auf eine Anfrage ein Ranking der Datenbankobjekte berechnet. Als Benutzerschnittstelle verwendet das System einen Menüorientierten Browser.

## Abstract

We describe a prototype for an information system that integrates text and fact retrieval. A query is a set of conditions which relate either to the text or the attribute values of a database object. Conditions may be assigned weights w.r.t. the query as well as to an object. These weights form the basis for a ranking of the database objects w.r.t. the query. As user interface, the system provides a menu-oriented browser.

## 1 Extending text retrieval methods for coping with factual data

In the past, research in the field of information retrieval (IR) has focused on the problem of text retrieval. It has been shown that Boolean retrieval yields a rather poor retrieval quality, and the complexity of the user interface prevents many potential users from using this kind of retrieval systems. On the other hand, systems based on the vector space model [Salton 71] or probabilistic retrieval [Rijsbergen 79] improve retrieval effectiveness by producing a ranked list of answers instead of a set only as in Boolean retrieval. In both models, a query is a set of terms (without logical operators, as in Boolean retrieval). Each term can be given a search term weight which reflects its importance with respect to the query. By means of relevance feedback, better values for these weights can be estimated, thus yielding significant improvements of retrieval quality [Robertson et al. 81]. A term also can be assigned a so-called-indexing weight with respect to a document, thus discriminating terms of different importance in a document [Salton & Buckley 88]

[Fuhr & Buckley 91]. Experiments have shown that this kind of weighted indexing outperforms binary indexing for most applications. Besides the benefits in terms of retrieval quality, the non-Boolean models also improve the user-friendliness of the retrieval system. Much of the complexity of the Boolean model is reduced by omitting Boolean operators and treating queries as a set of terms only. Furthermore, ranking allows a user to select any number of documents that he wants to see, while with Boolean systems, he has to transform the query formulation until the appropriate size of the output is achieved.

However, in real IR applications, the objects to be retrieved hardly ever consist of text only. For example, document databases offered by public hosts also provide bibliographic information. In office information systems, documents also comprise a number of attributes besides the text itself. Queries posed to these systems frequently relate to the attribute values of the documents, e.g. the name of the author (or sender of a letter) or the publication (or mailing) date. With regard to these requirements, text retrieval methods only offer a partial solution. So concepts for combining text and fact retrieval have to be devised.

The approach described in [Raghavan et al. 86] combines probabilistic text retrieval methods with Boolean fact retrieval. This leads to a rather complex retrieval procedure, which we deem not to be adequate for the task of integrating text and fact retrieval. The major drawback of this approach, however, is the rigid treatment of query conditions relating to attributes. We think that for most of the applications discussed here, it is not appropriate to apply Boolean logic for fact retrieval. Instead, conditions specified by a user should be regarded as being vague in most cases. For example, in an office information system, when a user seeks for a letter that was sent last month, then there is a certain probability that the letter may date one or two weeks earlier. With proper names, there is a similar problem, since the user may have difficulties to give the correct spelling (or the name in the document may be misspelled). For this reason, in case there is no answer that fulfills all the criteria specified in the query, the system also should search for objects that are close to these criteria. As a result, the system should present a ranking of objects, from which the user may select the one he is looking for. In comparison to this approach, today's systems would just return a null answer, thus forcing the user to reformulate his query by loosening some of the criteria, waiting for the answer, and possibly repeating these steps until the desired object is retrieved.

The discussion above has shown that for many applications, fact retrieval should be handled in a similar way as text retrieval. So we seek for an approach that integrates text and fact retrieval in a way that both conditions relating to text or facts can be regarded as being vague. So far, two proposals for solving this problem have been published:

- In [Croft & Krovetz 88], an office information system allowing vague fact and text retrieval is proposed. Whereas probabilistic indexing and search term weighting is applied for the text part of the query, the theory of endorsements is used for conditions relating to facts. However, there is no obvious way how the results of the two parts of the query can be combined for producing a single ranking.
- A more advanced system for retrieval of multimedia data (including facts) is described in [Rabitti & Savino 90]. This system uses probabilistic retrieval for text as well as for facts. Both kinds of conditions can be assigned weights with respect to the query. In addition, text terms can be given probabilistic index terms weights w.r.t. an object, whereas conditions relating to attributes are either true or false for

a specific object. So conditions relating to facts are not regarded as being vague in this approach. In order to represent vague fact conditions, a user can split the condition into several conditions with different attribute values and assign different weights to these conditions. As a result, this system yields a probabilistic ranking of objects w.r.t. a query. This ranking can be improved further by means of relevance feedback.

Both approaches suffer from the fact that they treat text and facts differently. In this paper, we show how a probabilistic text retrieval model can be combined with a probabilistic model for vague fact retrieval. This combination treats texts and facts in a uniform way. The new model can be applied in many areas where texts and facts are stored together, and where at least parts of the queries relating to facts should be interpreted as vague conditions (e.g. uncertainty about dates, correct spelling of names, similarity of objects like products or software modules).

In the following section, we first describe the combined model. Then we present a prototype information retrieval system that implements this model. We show that the user interface of this system is more user-friendly than those of today's systems, while the effectiveness of the system is improved at the same time.

## 2 A unified model for text and fact retrieval

Here we give a brief description of the unified model. The details of the fact retrieval model are described in [Fuhr 90], whereas the text retrieval model is presented in [Fuhr 89] and [Fuhr & Buckley 91].

In our approach, a database is a set of objects  $O$ . For simplicity, we assume here that a vague query always relates to all elements of  $O$ . In [Fuhr 91], we present a more detailed model, where a query may also comprise a Boolean part which yields a preselection of objects from  $O$ . Furthermore, of course, there may be different classes of objects (e.g. relations in a relational database) to which different queries may relate to, but we will restrict to the simple case in the following.

A vague query  $q_k$  is a set  $q_k^c = \{c_i\}$  of vague conditions  $c_i$ , where a vague condition may be either a text condition or a fact condition (see below). For an object  $o_m$ , the indexing weight  $u_{im}$  gives the weight of object  $o_m$  with respect to condition  $c_i$ . In [Fuhr 90], we discuss how this weight can be interpreted as the probability  $P(C|c_i, o_m)$  that a random user will judge object  $o_m$  to be a "correct" (or acceptable) answer to condition  $c_i$ . Furthermore, it is shown how this probability can be estimated by using relevance feedback data that has been collected from queries previously submitted to the system. In this paper, we will not go into the details of estimating probabilistic indexing weights from feedback data., we will just assume that there is a so-called indexing function  $e(c_i, o_m)$  that computes the indexing weights  $u_{im}$ .

A user may additionally assign weights to the conditions of this query, in order to denote their different importance w.r.t. his need. For this purpose, let  $w_{ik}$  give the weight of condition  $c_i$  in query  $q_k$ .

Based on the indexing weights  $u_{im}$  and the query conditions weights  $w_{ik}$ , we can compute the relevance status value for a query-object pair  $(q_k, o_m)$  by means of a so-called retrieval

function  $\varrho(q_k, o_m)$ , and then the objects can be ranked according to decreasing relevance status values. Based on a decision-theoretic model [Wong & Yao 90], the retrieval function yields

$$\varrho_{lin}(q_k, o_m) = \sum_{c_i \in q_k^c} w_{ik} \cdot u_{im}. \quad (1)$$

If relevance feedback data is available, the initial ranking can be further improved. Assume that  $O_k^R$  denotes the set of objects judged relevant by the user so far, then we have to estimate the average indexing weight for each condition  $c_i$  for the relevant objects as well as for all objects, that is

$$r_{ik} \approx \frac{1}{|O_k^R|} \cdot \sum_{o_m \in O_k^R} u_{im} \quad \text{and} \quad s_{ik} \approx \frac{1}{|O|} \cdot \sum_{o_m \in O} u_{im}$$

With these parameters, we can apply the RPI model [Fuhr 89], which gives us the retrieval function

$$\varrho(q_k, o_m) = \prod_{c_i \in q_k^c} \left[ \left( \frac{r_{ik}(1 - s_{ik})}{s_{ik}(1 - r_{ik})} - 1 \right) \cdot u_{im} + 1 \right] \quad (2)$$

Now we describe the process of computing the indexing weights, which is different for text and fact conditions.

For fact conditions, we assume that there is a set of attributes  $A = \{a_1, \dots, a_n\}$ , and that each object  $o_m$  is represented as a tuple of values  $t_m(a_1, \dots, a_n)$  for these attributes. For vague fact conditions, we introduce the concept of vague predicates, which may be either unary predicates (like “high”, “low”, “some”) or binary predicates (like “similar to” or vague interpretations of standard binary predicates like  $<$  or  $>$ ). Let  $F_i^1$  denote the set of unary predicates applicable to attribute  $a_i$  and  $F_i^2$  the corresponding set of binary predicates. Then a fact condition  $c_i$  can have one of the two forms

- $(a_i, f_i)$  with  $a_i \in A$  and  $f_i \in F_i^1$ ,
- $(a_i, f_i, d_i)$  with  $a_i \in A$ ,  $f_i \in F_i^2$  and  $v_i \in D_i$ , where  $D_i$  is the domain of attribute  $a_i$  and  $v_i$  is called the comparison value.

For the computation of the indexing weight of a condition-object pair  $(c_i, o_m)$ , we first collect information about this pair in a so-called relevance description  $x(c_i, o_m)$ . A relevance description contains values of properties of the relationship between condition  $c_i$  and attribute value  $t_m(a_i)$  (e.g. for binary conditions and numerical attribute values the relative difference between  $v_i$  and  $t_m(a_i)$ ). In a second step, a so-called indexing function  $e(x(c_i, o_m))$  is developed which maps relevance descriptions onto probabilistic indexing weights  $u_{im}$  which are estimates of the probability  $P(C|x(c_i, o_m))$ . The indexing function can be developed from a learning sample with relevance descriptions and correctness decisions which are derived from relevance judgements of previous uses of the system. In this paper, we will only describe the predefined indexing functions which have to be supplied

when a new database is set up for which no feedback information is available yet (see next section).

A text condition is a term  $t_i$ , which may be either a word or a noun phrase. Let  $d_m$  denote the textual part of object  $o_m$ , then we seek for a weighting of  $t_i$  w.r.t.  $d_m$ . Like for facts, we first have to construct a relevance description  $x(t_i, d_m)$ , which forms the basis for the computation of the indexing weight  $e(x(t_i, d_m))$  by means of the indexing function  $e$ . An important element of the relevance description for text conditions is the so-called form of occurrence of term  $t_i$  in document  $d_m$ . This concept covers the significance of  $t_i$  w.r.t.  $d_m$  (e.g. the within-document frequency or the parts of the document in which  $t_i$  occurs) as well as the certainty of identification of  $t_i$  in  $d_m$  (e.g. by means of different kinds of stemming or the syntactic structure of a noun phrase). Furthermore, additional information about  $t_i$  (e.g. its document frequency) and  $d_m$  (e.g. the length of the text) may be included in  $x(t_i, d_m)$  (for further details, see [Fuhr & Buckley 91]). Then an indexing function is developed in the same way as for fact conditions.

### 3 An application example

In order to illustrate the concepts outlined in the previous section, we present an application of integrated text and fact retrieval here. First, the database and the indexing functions for the different attributes are described, followed by the presentation of the PFTR system developed for this application.

As database, we have used a number of documents from the NTIS database which contains references to research reports of scientific projects in the USA. An example document is shown in figure 1.

From the different categories of these documents, the title, the summary, the controlled terms (CT) and the text of the classification codes (CC) can be accessed via text retrieval. In order to compute the indexing weights for the textual part, the SMART indexing procedures have been applied ([Salton & Buckley 88]). These procedures have the advantage that no relevance feedback information is required for indexing. However, as shown in [Fuhr & Buckley 91], when feedback data is available, a better retrieval quality can be achieved with probabilistic indexing methods.

For the other categories of the document as well as for the classification codes itself, vague fact retrieval methods are supplied. Except for the attribute 'date' (see below), string distance measures are used as relevance descriptions within the indexing functions of the different attributes. Given a distance measure  $\mu(v, s)$  for two strings  $v$  and  $s$ , the indexing function yields

$$e_\mu(x(c_i, o_m)) = 1 - \mu(v_i, t_m(a_i))$$

(remember that  $v_i$  is the comparison value specified in the query condition and  $t_m(a_i)$  is the corresponding attribute value of object  $o_m$ ). The distance measure actually used depends on the attribute. In the following, we describe the different measures along with the attributes for which they are applied.

Here we assume that a string  $s$  is a sequence of characters  $s_1, s_2, \dots, s_{l(s)}$ , where  $l(s)$  denotes the length (the number of characters) of  $s$ . In some of our definitions given below, we will also refer to the characters  $s_0$  and  $s_{l(s)+1}$ , which we assume to be blank.



AN 87(01):411 NTIS Order Number : AD-A172 502/7/XAD  
 TI Controlling Inference. (Doctoral thesis)  
 AU Smith, David E.  
 CS Stanford Univ., CA. Dept. of Computer Science  
 NC Contract : N00014-81-K-0004  
 NR AD-A172 502/7/XAD; STAN-CS-86-1107  
 197 p. NTIS Prices : MF A01  
 Availability : Microfiche copies only.  
 PD 860400  
 LA English CY United States  
 OS GRA&I8701  
 AB Effective control of inference is a fundamental problem in Artificial Intelligence. Unguided inference leads to a combinatorial explosion of facts or subgoals for even simple domains. To overcome this problem, expert systems have used powerful domain-dependent control...  
 CC 95F Bionics and artificial intelligence  
 CT \*Artificial intelligence; Data bases; Decision making; Global; Information exchange; Problem solving; Efficiency  
 \*Inference; Control; Theses; Expert systems  
 UT NTISDODXA

Figure 1: Example NTIS document

A well-known comparison method for strings (especially for names) is the trigram method. The set of trigrams  $T(s)$  of a string  $s$  is the set of all substrings of length three of  $s$ , that is

$$T(s) = \{(s_i, s_{i+1}, s_{i+2}) | 0 \leq i \leq l(s) - 1\}.$$

In order to measure the deviation of a string  $s$  from a specified string  $v$ , the proportion of those trigrams of  $v$  that occur also in  $s$  is computed:

$$\mu_T(v, s) = \frac{|T(v) \cap T(s)|}{|T(v)|}.$$

In our application, the trigram method is used for the attribute "Author".

For the comparison of classification codes, we use the prefix method, where the number of matching characters at the beginning of the two strings is divided by the length of the comparison string:

$$\mu_P(v, s) = \frac{1}{l(v)} \cdot \max_j (\forall_i 0 \leq i \leq j \Rightarrow v_i = s_i).$$

For the attributes "Number of Report" and "Other Sources", the proportion of matching characters is computed:

$$\mu_M(v, s) = \frac{|\{i | v_i = s_i\}|}{\min(l(v), l(s))}.$$

In the case of attributes that may consist of several words, a word-by-word comparison method is used. The set of words  $W(s)$  of a string  $s$  is defined as

$$W(s) = \{s_i, \dots, s_j | 0 \leq i < j \leq l(s) + 1 \wedge s_i = ' ' \wedge s_j = ' ' \wedge \forall_k i < k < j \Rightarrow s_k \neq ' '\}.$$

Then the proportion of matching words of the comparison string is computed as the value of the distance measure

$$\mu_M(v, s) = \frac{|W(v) \cap W(s)|}{|W(s)|}.$$

We use this method for the attribute "Corporate Source".

For dates, we have defined a function  $k(d)$  that gives us the number of days between the current date and the date  $d$ . As relevance description, we compute the relative difference between the comparison value  $v_i$  and the attribute value  $t_m(a_i)$  according to the formula

$$x(c_i, o_m) = \frac{k(v_i) - k(t_m(a_i))}{k(v_i)}$$

The final indexing function depends on the predicate  $f_i$ :

$$\begin{aligned} f_i = '=': & \quad e(x) = 1 - \min(1, |x|) \\ f_i = '>': & \quad e(x) = \begin{cases} 0 & \text{if } x < -0.75 \\ 1 & \text{if } x > 0 \\ 1 + x/0.75 & \text{otherwise} \end{cases} \\ f_i = '<': & \quad e(x) = \begin{cases} 0 & \text{if } x > 0.75 \\ 1 & \text{if } x < 0 \\ 1 - x/0.75 & \text{otherwise} \end{cases} \end{aligned}$$

We have described these indexing functions in detail because there are many similar applications where attributes of this kind occur, so that these relevance descriptions and indexing functions may be appropriate. Of course, the functions shown here should be regarded as prototypes only. For a specific applications, different coefficients in these formulas may be preferable. However, as soon as some feedback data is available, the specific parameters of the functions should be estimated via probabilistic indexing as described in [Fuhr 90].

## 4 The PFTR system

Now we describe a prototype system for probabilistic fact and text retrieval for the application presented in the previous section. This prototype serves mainly for the illustration of the user interface and the basic retrieval functions, whereas certain information system aspects (like e.g. the management of large collections of objects) have not been considered yet. Figure 2 shows the user interface of the PFTR system.

The windows in the upper row serve for the formulation of vague attributes to which queries may relate to ("Basic Index" = text conditions). After selecting an attribute with the mouse, the next window shows the predicates that are applicable for this attribute. Having chosen a predicate, the user may enter a comparison value and a query condition weight within the two small windows below the predicate window. Then the rightmost window shows the complete vague query constructed so far. With the operation "remove", single conditions can be removed from the query formulation. When the user has completed his query formulation, he issues an "update" command which tells the system that the windows in the bottom row should be updated such that they present the answers to the query formulated in the top half. In the leftmost window of the bottom row, the

QueryBrowser		Query:
Author Classification Codes Corporate Source Number of Report Other Sources Publication Date Basic Index -----	= < > -----  0786  1	1 Author            ■ Smith 1 Classification Codes ■ 12A 1 Other Sources    ■ GRA&17507 1 Publication Date   ■ 0786 1 Basic Index       ■ nat Inform
Controlling Inference. (Doctoral thesis) Probabilistic Analysis of Algorithms for NP-Complete Expert System for Predicting Component Kill Probabil Calculus of Uncertainty in Artificial Intelligence and Models of Cognitive Behavior in Nuclear Power Plant Models of Cognitive Behavior in Nuclear Power Plant Probabilistic Analysis of Algorithms for NP-Complete Multi-Disciplinary Techniques for Understanding Time Technology Transfer and Artificial Intelligence. (Tech Diagnosing Cognitive Errors: Statistical Pattern Class Selected Papers from the Annual Meeting of the Brit Artificial Intelligence and Its Importance in Education Toward a Theory of Plan Recognition NESTOR: A Computer-Based Medical Diagnostic Aid Minimal Cut-Set Methodology for Artificial Intelligen Automatic Probabilistic Knowledge Acquisition from D Intelligence Algorithm Methodology 1. (Final rep.) SPIRIT: An Evolutionarily Designed Intelligent Tutoring Model-Based Probabilistic Reasoning for Electronics Structure for Generation and Control of Intelligent B	Controlling Inference. (Doct Expert System for Predict Models of Cognitive Behav	RBV: 1.50734 Author: Franco, John Classification Codes: 12A Mathematics and statistics 72 Mathematical Sciences 62 Computers, Control, and information Theory Corporate Source: Indiana Univ. at Bloomington. Dept. of Computer Science Sponsor : Air Force Office of Scientific Research, Boeing AFB, DC Number of Report: AD-A168 617/9/KAD; AFOSR-TR-86-0310 8 p. NTIS Prices : PC A02/MF A01 Other Sources: GRA&18920 Publication Date: 86 12 12 Summary: The goal of this research is to develop and analyze algorithms which can, in some practical sense, solve certain NP-complete problems efficiently. By solve we mean determine whether a solution to a given instance of an NP-complete problem exists where, for the problems we have considered, a solution is an assignment of values to a list of variables which cause some predicate to be true. We do not consider actually finding solutions when they exist since doing so adds unnecessary complexity to the statement of the algorithms: the algorithms we consider can all be modified to find solutions without significantly altering performance. NP-complete problems are found in Cryptology, Operations Research, Artificial Intelligence, Computer System Design and many other areas. There is no known algorithm for an NP-complete problem which runs in time bounded by a polynomial on the length of the input (polynomial time) in the worst case nor is

Figure 2: User interface of the PFTR system

ranked list of items is shown by displaying the titles. By selecting one of these titles, the complete object is shown in the rightmost window. Furthermore, a user may mark some of the objects a being relevant to his query. In turn, these titles are shown in the middle window. When the user wants to see the effect of the relevance feedback information given to the system, he again issues the "update" command; in turn, the system computes the new ranking and displays the new list in the leftmost window. The user may also decide to modify his query by adding or removing conditions. After finishing the modification of the query, the "update" command produces the new ranking.

The basic idea behind the "update" command is to provide a generic operation for performing the retrieval step. In principle, a new ranked list of objects could be computed after each modification of the query or after ranking an object as being relevant. However, an implementation of an instant re-ranking would be rather inefficient and also confusing for the user (he would have to wait for the result, and furthermore he may not want instant changes of the ranked list). For this reason, the system indicates that an update operation would be possible by displaying a grey frame around the whole PFTR browser. So the user knows that the result displayed in the bottom half does not reflect the current state of the query and/or the feedback judgements, and that an update command would produce the latest state of the search. This strategy is rather different from that of today's

systems, where user commands correspond one-to-one to system operations, and where additional commands are necessary after a retrieve command in order to show elements of the answer.

## 5 Open Problems

The PFTR system is a first step towards probabilistic retrieval systems for both text and facts. In our application, we have chosen a database with uniformly structured objects. This fact makes the formulation of queries rather easy. In many real applications, however, the user has to deal with rather different kinds of objects, where he either relates to a single type of objects or to a type hierarchy (like in object-oriented databases). In these situations, an approach like in the RABBIT system [Williams 84] would be suitable for query formulation.

A second important problem is the further development of the PFTR prototype towards an information system that can handle large collections of objects. Currently, all objects are held in main memory, and the retrieval status values of all objects are computed for a query. In large information system, only a certain proportion of all objects should be considered w.r.t. a query, namely those that will form the head of the ranked list. For this purpose, special access paths and the corresponding ranking algorithms have to be developed.

## References

- Croft, W.;** **Krovetz, R.** (1988). Interactive Retrieval of Office Documents. In: *Proceedings of the Conference on Office Information Systems*. ACM, New York.
- Fuhr, N.;** **Buckley, C.** (1991). A Probabilistic Learning Approach for Document Indexing. *ACM Transactions on Information Systems* 9(2).
- Fuhr, N.** (1989). Models for Retrieval with Probabilistic Indexing. *Information Processing and Management* 25(1), pages 55–72.
- Fuhr, N.** (1990). A Probabilistic Framework for Vague Queries and Imprecise Information in Databases. In: McLeod, D.; Sacks-Davis, R.; Schek, H. (eds.): *Proceedings of the 16th International Conference on Very Large Databases*, pages 696–707. Morgan Kaufman, Los Altos, Cal.
- Fuhr, N.** (1991). *Probabilistic Retrieval for Imprecise Queries*. Report DV II 91-3, TH Darmstadt, FB Informatik, Datenverwaltungssysteme II.
- Rabitti, F.;** **Savino, P.** (1990). Retrieval of Multimedia Documents by Imprecise Query Specification. In: Bancilhon, F.; Thanos, C.; Tsichritzis, D. (eds.): *Advances in Database Technology - EDBT '90*, pages 203–218. Springer, Berlin et al.
- Raghavan, V.;** **Saxton, L.;** **Wong, S.;** **Ting, S.** (1986). A Unified Architecture for the Integration of Data Base Management and Information Retrieval Systems. In: Kugler, H.-J. (ed.): *Information Processing 86*, pages 1049–1054. Elsevier, Amsterdam.

- van Rijsbergen, C.** (1979). *Information Retrieval*. Butterworths, London, 2. edition.
- Robertson, S.; Van Rijsbergen, C.; Porter, M.** (1981). Probabilistic Models of Indexing and Searching. In: Oddy, R.; Robertson, S.; Van Rijsbergen, C.; Williams, P. (eds.): *Information Retrieval Research*, pages 35–56. Butterworths, London.
- Salton, G.; Buckley, C.** (1988). Term Weighting Approaches in Automatic Text Retrieval. *Information Processing and Management* 24(5), pages 513–523.
- Salton, G. (ed.)** (1971). *The SMART Retrieval System - Experiments in Automatic Document Processing*. Prentice Hall, Englewood Cliffs, New Jersey.
- Williams, M.** (1984). What makes RABBIT run? *International Journal on Man-Machine Studies* 21, pages 333–352.
- Wong, S.; Yao, Y.** (1990). Query Formulation in Linear Retrieval Models. *Journal of the American Society for Information Science* 41(5), pages 334–341.

**Author's address:**

Prof. Dr. Norbert Fuhr  
Universität Dortmund  
Informatik VI  
Postfach 500500  
W-4600 Dortmund  
E-mail: fuhr@bonny.informatik.uni-dortmund.de

# Navigationsstrategien in der Informationssuche -vom Information Retrieval zu Hypertext-

Frank Krüger, M.A.

## **Zusammenfassung**

Dieser Beitrag beschäftigt sich mit einer Erweiterung der Suchstrategien in elektronischen Informationssystemen. Zu der zielgerichteten Suche, typisch für das Information Retrieval (IR), kommen ungerichtete und teilweise auch passive Suchparadigmen hinzu, wie sie in Hypertext (HT) anwendbar sind. Ausgehend von dem Paradoxon des IR und dem Desorientierungsproblem in Hypertext, werden die Navigationsstrategien eines um die Methoden des IR erweiterten HT-Modells in drei Hauptgruppen eingeteilt: Die erste Gruppe der strukturorientierten Strategien umfaßt die bekannten Strategien der Hypertextnavigation (z.B. das Browsing); die zweite Gruppe sind die inhaltsorientierten Strategien, die auf der Anwendung von IR-Methoden des exakten bzw. partiellen Abgleichs auf Hypertext beruhen; die dritte Gruppe der dynamischen Strategien vereint schließlich die Gruppen der struktur- und inhaltsorientierte Strategien und führt zu einem dynamischen HT, wo die Struktur des Hyperdokumentes von verschiedenen, erst während einer Sitzung bestimmten Bedingungen abhängen kann. Den Schluß bilden einige allgemeine Anmerkungen zum Design von HT.

## **Abstract**

This article deals with an expansion of the range of search strategies in electronic information systems. The bestknown strategy, used in Information Retrieval (IR), is the specific search. But there are more, undirected and partly passive strategies, which are available in hypertext. The paradox of IR and the problem of disorientation in hypertext are discussed. In an expanded model of hypertext, which is enriched with methods from the IR, three major groups of navigation strategies are identified:

- 1) the group of structure-oriented strategies, which are the well-known methods of hypertext navigation (e.g. browsing),
- 2) the group of content-oriented strategies, based on the application of IR methods of exact resp. partial matching in hypertext, and
- 3) the group of dynamic strategies, which is a combination of the two former groups. Based on a dynamic hypertext model, the structure of the hyperdocument is defined only during a user session. Finally, some general remarks on the design of hypertext are made.

## 1. Einleitung

Auf dem Gebiet der rechnergestützten Informationssysteme zeichnet sich in den letzten Jahren eine tiefgreifende Änderung ab. Während bisher nur die zielgerichtete Suche in Texten von computergestützten Information Retrieval Systemen unterstützt wurden, bietet das Hypertext-Konzept nun auch die Möglichkeit in elektronischen Systemen nach anderen Paradigmen der Informationssuche vorzugehen (vgl. Tabelle 1).

Ziel dieses Beitrages ist es zu zeigen, wie sich die bisher getrennten Forschungs- und Anwendungsgebiete von Hypertext und Information Retrieval (IR) gegenseitig ergänzen können. Dazu sollen zunächst kurz die jeweils "klassischen" Konzepte definiert und ihre Probleme aufgezeigt werden. Den Hauptteil dieses Beitrags stellt eine Systematisierung der Vorgehensweisen bei der Informationssuche dar, die als Navigationsstrategien bezeichnet werden. Außerdem wird anhand von Beispielen eine Zuordnung zu entsprechenden Orientierungshilfen vorgenommen.

## 2. Die "klassischen" Konzepte

### 2.1. Information Retrieval

Der Gegenstand des Information Retrieval ist die "Repräsentation, Speicherung, Organisation [von] und der Zugriff auf Information" (/1/, S. 1), wobei die Information meist in Texten enthalten und damit (im Ggs. zu den Datenbanksystemen) nur schwach strukturiert ist.

Das dem IR zugrundeliegende Matching-Paradigma vergleicht eine Suchanfrage, die das akute Informationsbedürfnis durch Suchbegriffe beschreibt, mit der Dokumentenmenge, die durch Indexbegriffe repräsentiert wird. Dieser Abgleich kann entweder deterministisch (exact match) oder mit Ähnlichkeitsmaßen (partial match) bestimmt werden. Nach /2/ muß der Nutzer, um potentiell relevante Dokumente zu finden, daher

- a) sein Problem in explizite Beschreibungen umsetzen, die während der Suche ergänzt und weiterentwickelt werden, und
- b) die verschiedenen Suchbegriffe und ihre Ergebnisse protokollieren und verwalten.

Dieses Paradigma führt zu dem Problem oder Paradoxon des IR, nämlich "the need to describe that which you do not know in order to find it" (/2/, S. 221), mit mindestens zwei Lösungsaspekten:

- 1) Beschreibung des Unbekannten durch das Bekannte (nach /3/ eine "active directed" Strategie der Informationssuche), oder
- 2) Erkennen des Unbekannten, wenn es auftaucht.

## 2.1. Hypertext

Der Begriff "Hypertext" (auch als nichtlinearer Text bezeichnet) wird von seinem Schöpfer, dem Software-Pionier Theodor H. Nelson wie folgt definiert: "Hypertext is the combination of natural-language text with the computer's capacities for interactive, branching or dynamic display, when explicitly used as a medium " (/4/, S. 195).

Lynch unterscheidet in /5/ aus informationswissenschaftlicher Sicht zwei Ansätze von Hypertext: Einerseits kann "one think of hypertext as a way of providing a user interface to a database that contains a number of interrelated data objects" (/5/, S. 265). Der andere Ansatz definiert Hypertext als "a means of creating and organizing information as a series of explicitly interrelated segments" (/5/, S. 266).

Zwischen beiden Ansätzen besteht natürlich ein enger Zusammenhang, denn jede Datenmenge mit Hypertext-Struktur läßt sich durch eine Hypertext-Oberfläche darstellen und zugänglich machen. Zusätzlich ist z.B. die Hypertext-Oberfläche auch für andere Organisationsformen interessant, wie für Wissensbasen von Expertensystemen, wo die Verknüpfungen dynamisch berechnet oder abgeleitet werden.

Grundlegende Merkmale von Hypertext sind also die Modularisierung (in Knoten) und Assoziierung (durch die die Knoten verknüpfenden Kanten). Als weitere Merkmale können mit Hilfe des elektronischen Mediums noch die Dialogisierung (z.B. durch Briefe zwischen Autor und Leser) und die Individualisierung (z.B. durch Raum für individuell abgespeicherte Notizen) hinzutreten /7/.

Im allgemeinen werden in Hypertext zwei große Problembereiche identifiziert (Literatur vgl. /7/): Desorientierung ("Lost in Hyperspace") und die kognitive Überlastung eines Nutzers als Leser oder Autor.

Das Problem der Desorientierung besteht in der Gefahr, daß der Nutzer sich in der Masse der mit Information gefüllten und durch vielfache Kanten verbundenen Knoten verliert, da ihm grundsätzlich räumliche Anhaltspunkte jeder Art fehlen.

Das für die hier behandelte Thematik der Informationssuche zentrale Problem der Navigationsstrategien, ist eng mit dem Orientierungsproblem verbunden. Jede zusätzlich angebotene Strategie birgt nämlich die Gefahr der Desorientierung von Benutzern in sich. Das Problem der Desorientierung läßt sich außerdem in folgende Teilprobleme aufteilen /6/:

- Der Nutzer versteht die ihm angebotenen Navigationsmöglichkeiten nicht.
- Dem Nutzer ist die Modellierung der Information unverständlich.
- Der Nutzer weiß nicht, wie er im Hyperdokument zu der Stelle gelangen kann, wo er hin will (bzw. er weiß nicht, wo er sich im Dokument befindet).



Auf das erste Teilproblem wird im folgenden näher eingegangen, während die beiden anderen Teilprobleme durch geeignete Orientierungshilfen zu lösen sind, die hier weitgehend ausgeklammert sind (eine ausführliche Behandlung der Orientierungshilfen findet sich in /7/).

Der Übergang zum anderen grundsätzlichen Problem von Hypertext, der kognitiven Überlastung des Nutzers, ist fließend. Dieses Problem wird hier jedoch nicht gesondert behandelt.

### **2.3. Problematisierung und Lösungsvorschlag**

Bei dem Vergleich der Vorteile und der Probleme von Hypertext und IR läßt sich somit zweierlei erkennen. Einerseits kann in der Frage der Suchstrategien das dem Hypertext zugrundegelegte Browsing als prinzipielles Navigationsprinzip durch Strukturierungs- und Datenmodelle aus der IR-Forschung und -praxis angereichert werden, was zu neuen Navigationsstrategien führt.

Andererseits bietet Hypertext jedoch als eine (zunächst) neue Generation von Benutzeroberflächen für IR-Systeme eine (gegenüber beiden isoliert gesehenen Konzepten) verbesserte Orientierung für den Benutzer. Beide Bereiche sind nach Anwendungen und Problembereichen differenziert zu bewerten (vgl. /7/).

### **3. Systematik der Navigationsstrategien**

Wie aus Tabelle 1 hervorgeht, können grob insgesamt vier Such- oder Navigationsstrategien unterschieden werden. Das aktive, gerichtete Searching - hierunter fallen die IR-Methoden des exakten und partiellen Abgleichs - bildet in der hier aufgestellten Systematik (siehe dazu Abbildung 1) die Gruppe der inhaltsorientierten Strategien. Die drei anderen Strategien - Browsing, Scanning und Wandering - werden dagegen in der Gruppe der strukturorientierten Navigationsstrategien zusammengefaßt. Sie sind für die Suche nach dem vorherrschenden Hypertext-Modell kennzeichnend. Eine dritte Gruppe, die als dynamische Navigationsstrategien bezeichnet wird, mischt beide zuerst genannten Gruppen von Navigationsstrategien (und damit auch die vier Formen der Informationssuche). Außerdem erweitert sie das bisher zugrundeliegende statische Hypertextmodell um eine dynamische - also nicht vorherbestimmte, sondern auf verschiedene, wechselhafte Faktoren begründete - Informationsstruktur.

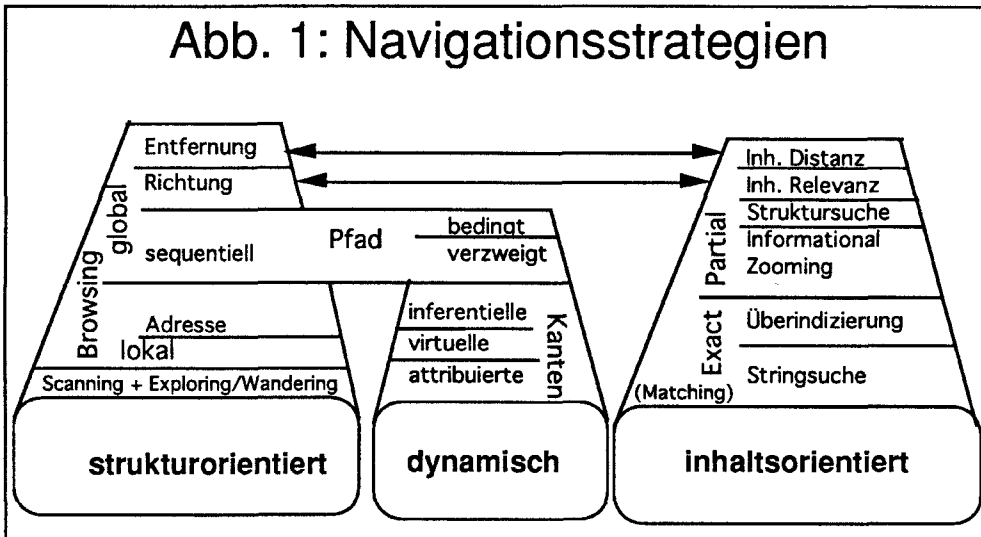
Jeder dieser Gruppen von Navigationsstrategien und teilweise sogar auch einzelnen Strategien lassen sich Orientierungshilfen zuordnen. Auf diese Zuordnung wird hier jedoch aus Platzgründen verzichtet. Für eine ausführliche Beschreibung vgl. /7/.

**Tabelle 1 Formen der Informationssuche**

	gerichtet "specific search" /8/	ungerichtet
aktiv	Searching: "I'm looking for info on..."	Browsing "purposeful search" [i.MCALESEE 1988;]
passiv	Scanning: "I'm keeping my eyes open for..."	Wandering: "Open to whatever impinges on my awareness"

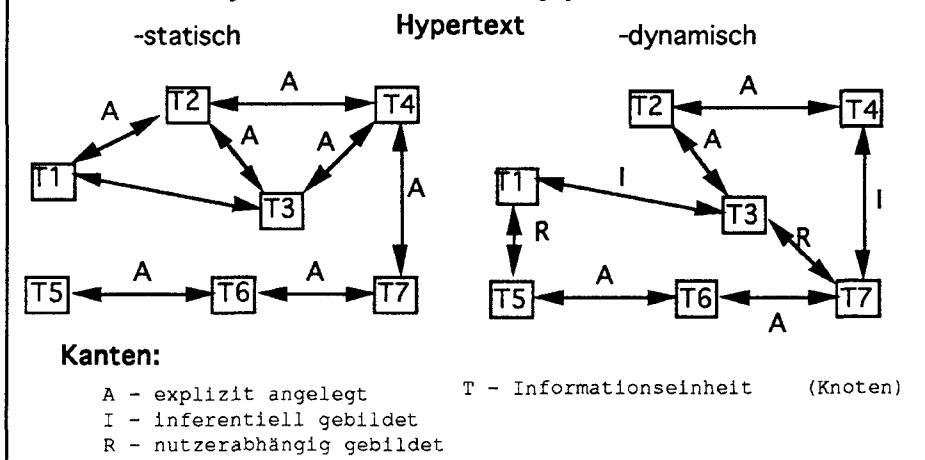
nach: /3/, S. 92, Fig.2

**Abb. 1: Navigationsstrategien**



aus: /7/

**Abb. 2: Dynamischer Hypertext**



nach: /9/, S. 260, Fig. 5.19

### **3.1. Strukturorientierte Strategien**

Unter Browsing soll hier das Paradigma der aktiven, ungerichteten Suche (vgl. Tabelle 1) durch ein Hyperdokument verstanden werden. Auch die beiden passiven Paradigmen der Informationssuche, also Scanning und Wandering (letzteres wird auch als Exploring bezeichnet), werden hier zu den strukturorientierten Navigationsstrategien gezählt. Dem Scanning entspricht etwa das Überfliegen einer Zeitung, während das Exploring z.B. für einen ersten Überblick über ein Sachgebiet verwendet werden kann. Der Schwerpunkt aller strukturorientierten Navigationsstrategien liegt auf dem Kontext der Information.

Grundsätzlich dienen graphische Orientierungshilfen dem strukturorientierten Navigieren. Sie können je nach Komplexitätsgrad und Umfang eines Hyperdokumentes linear als (intellektuell erstellte) Landkarte oder als dreidimensionale Fisheye-Views bzw. durch andere Methoden hierarchischer Darstellung dargestellt werden (vgl. /7/).

### **3.2. Inhaltsorientierte Strategien**

Bei dieser Technik beschreibt der Nutzer seine Problemstellung in der Art, wie es R. Hjerpe als die erste Lösung zu dem Paradoxon des IR anbietet: "Describing the unknown by the known" (/2/, S. 321).

Dabei ist das Unbekannte oft nicht so sehr die Fragestellung selbst, sondern - wie etwa bei Literaturdatenbanken - die spezielle oder weiterführende Information, wie sie in den Referenzen (hier Literaturangaben) vermutet wird. Eine komplette Übersicht über die in diesem Zusammenhang relevanten IR-Methoden gibt /7/. Hier sollen nur die wichtigsten Ansätze auf ihre Anwendbarkeit in Hypertext hin untersucht werden.

Eine Verbesserung der reinen Stringsuche ist die sog. Überindizierung durch Schlagwortvergabe oder Synonymkontrolle. Auch Thesauri, wo Begriffsbeziehungen in einem eigenen Begriffsnetzwerk kontrolliert werden, sind Hilfsmittel für die Stringsuche.

Der Nachteil dieser Verfahren liegt in dem hohen Arbeitsaufwand zur Erstellung und Pflege des sog. kontrollierten Vokabulars und der relativen Inflexibilität gegenüber neuen Sachverhalten in sich schnell verändernden Wissensgebieten. Diese Probleme geht das automatische Indexieren an, indem es sog. unkontrolliertes Vokabular verwendet, das den jeweiligen Texten entnommen ist.

Die inhaltliche Relevanz, "in gewissem Sinne das inhaltliche Gegenstück zur Richtungsstrategie" (/10/, S. 7), bestimmt eine inhaltliche Richtung durch unterschiedliche Gewichtungen. Die Relevanz von Texten kann auf unterschiedliche Arten bestimmt werden: Entweder intellektuell (Bedeutung des Autors, Originalität des Inhalts) oder maschinell (z.B. Zentralität ermittelt durch Zitatanalyse). Frisse führt in /11/, S. 250ff, aus, wie die von

Salton entwickelten statistischen Maße /1/, sog. intrinsic und extrinsic weights, Hyperdokumente inhaltlich besser zugänglich machen können. Die inhaltliche Distanz dagegen beruht auf probabilistischen Ähnlichkeitsfunktionen /1/ und ist vergleichbar mit der strukturorientierten Entfernungsstrategie. Voraussetzung ist allerdings eine gewisse inhaltliche Regelmäßigkeit.

Diesen beiden Verfahren des exakten Abgleichs stehen eine Reihe von "Partial Matching"-Techniken gegenüber, die sich auf den semantischen Inhalt der Termini beziehen.

Das Informational Zooming, "die Präsentation eines Textes auf verschiedenen, automatisch generierten Abstraktionsebenen von der Übersicht bis ins Detail" (/10/, S. 7), setzt ein sog. "Semantic Parsing" voraus, d.h. eine Extraktion von Konzepten aus einem Volltext. Diese Technik wird in /12/ als vertikale Navigation, d.h. als Navigation zwischen Einheiten unterschiedlichen Abstraktionsgrades, dem Browsing als horizontales Navigieren gegenübergestellt.

Dagegen beruht die Struktursuche, wie die zuerst genannten Techniken, auf dem Abgleich (engl. Matching) von Suchanfrage und gespeicherten Dokumente, berücksichtigt aber explizit die Struktur der Dokumente. Eine einfache Anfrage in der Struktursuche könnte etwa in einer juristischen Datenbank so aussehen: Suche nach Urteilen des Bundesverfassungsgerichts, in denen ein bestimmter Gesetzesparagraf zitiert wird. Damit rückt die Struktursuche in die Nähe des Faktenretrieval.

Bei der Implementierung konkreter Systeme werden oft mehrere inhaltsorientierte Strategien bzw. Verfahren zur Verfügung gestellt. SuperBook /13/, S.190ff, verwendet z.B. nicht nur Verfahren der Überindizierung (insbesondere die Lemmatisierung, d.h. Wortformenreduktion) und die Boole'schen Verknüpfung, sondern verbindet diese auch mit strukturellen Angaben. So werden etwa Informationen über die Häufigkeit des gesuchten Begriffs nach Kapiteln aufgeschlüsselt. Ferner ist eine Synonymkontrolle implementiert, deren Einträge der Nutzer auch selbst ergänzen kann. Als Orientierungshilfe dient hier die Listenauswahl. Mayes favorisiert in /14/ für sein Lernsystem StrathTutor eine auf Suchanfragen basierende Navigation, wobei die Orientierung allein durch die Bildung einer konzeptionellen Karte erfolgt. Davon erhofft er sich einen zusätzlichen Lernerfolg.

Die Verfahren des Partial Matching können konzeptuell durch Filter verdeutlicht werden, wie etwa bei dem System SemNet. Grundlage ist entweder eine hierarchische Clusterung der Wissensmenge mit Hilfe von Ähnlichkeitsberechnungen, eine Metastruktur aus einer "continuous set of graphic representations that differ only in size" (/15/, S. 221) oder eine unterschiedlich hohe Auflösung von wichtigen gegenüber unbedeutenden Eigenschaften. Kanten werden nur gezeigt, wenn der Ursprungsknoten nicht eindeutig ist und wenn bei großen Datenmengen die durch sie ausgedrückten semantischen Relationen von momentanem Interesse sind.

### 3.3. Dynamische Strategien

Die den struktur- sowie inhaltsorientierten Navigationsstrategien zugrundeliegende statische Natur von Hypertext-Netzen - Änderungen sind oft nur in einem expliziten Editiermodus möglich - ist insbesondere beim Aufbau von Hyperdokumenten nachteilig. Durch eine dynamische Definition der Hypertextstrukturen läßt sich dagegen die Flexibilität und Nutzbarkeit des Systems wesentlich erhöhen. Nachteilig ist jedoch der wesentlich erhöhte Rechenaufwand, insbesondere während der Nutzung des Hypertextsystems, was nicht zu verlängerten Antwortzeiten oder einer eingeschränkten Benutzeroberfläche führen sollte.

Eine Dynamik wird durch die dynamische Zuordnung von Attributen zu Kanten bzw. der Erzeugung der Kanten selbst während einer Benutzersitzung erreicht (vgl. Abb. 2). Beispiele für dynamisch berechnete Attribute sind die Häufigkeit der Benutzung einer Kante oder eine individuelle Gewichtung durch den Nutzer. Ebenso können Kanten nur virtuell, d.h. benutzer- oder sitzungsspezifisch angelegt sein. Eine dritte Möglichkeit sind inferentielle Kanten, die auf einer (automatischen) Auswertung des Benutzerverhaltens basieren.

Basis der Berechnungen für (dynamisch) attribuierte Kanten sind inhaltliche oder strukturelle Gewichte, also entweder das intrinsic (Eigengewicht oder Originalität) bzw. extrinsic (Beziehungsgewicht oder Zitierhäufigkeit) weight eines Konzepts (inhaltlich) oder der Vernetzungsgrad (engl. adjacency), die Schlüsselfunktion (engl. betweenness) und die Distanz zum aktuellen Konzept (strukturell). Besondere Vorteile ergeben sich bei einer dynamischen Neuberechnung der Konzeptkarte jeweils zum Zeitpunkt ihres Aufrufs /10/, S. 11, was jedoch einen hohen Aufwand an Rechenzeit erfordert.

Dagegen setzt /5/, S. 267ff, Kanten in direkte Beziehung zum traditionellen Indexieren, indem sie definiert, daß alle Knotenpaare mit dem gleichen Indexterm durch nur eine Kante verknüpft sind. Eine Verbesserung dieser Methode durch die Möglichkeit der Gewichtung sieht sie im Einsatz der automatischen Indexierungsverfahren. In großen Datenbanken, so Lynch weiter, dürften nur diejenigen Kanten angeboten werden, die verschiedene Selektionskriterien erfüllen, so z.B. Kanten mit Gewichten über einem bestimmten Schwellenwert, mit einem bestimmten Typus oder mit spezifischen Attributen. Mit einem Relevanzfilter dieser Art ließen sich - allerdings wiederum mit einem hohem Aufwand an Rechenzeit - Kanten mit Relevanzwerten unter dem (ggf. benutzerdefinierten) Schwellenwert ausblenden. Eine weitere Abstufung der Kanten läßt sich durch Zuordnung verschiedener Farbbereiche erreichen.

Auch eine erweiterte Form der Pfade läßt sich den dynamischen Strategien zuordnen; im Unterschied zur strukturorientierten Pfadstrategie wird der Pfad hier aber als eine selbständige Einheit mit der Möglichkeit zur Ablauf- und Präsentationssteuerung betrachtet /10/, S. 9. Pfade sind somit eine Abfolge von Knoten (oder "entries" bzw. "Tour Stops"), die

durch Kanten untereinander verbunden sind und an die auch Programme "angehängt" sein können um Aktivitäten verschiedener Art, z.B. Animationen oder Berechnungen, auf Anforderung ablaufen zu lassen. Die Pfade bilden eine (grundsätzlich lineare) Metaebene über dem eigentlichen Material (das in Knoten organisiert ist) und sind unabhängig davon zu sehen. Diese Möglichkeiten sind bei dem System Trellis bis zu einem auf Petri-Netzen basierenden Modell ausgearbeitet (vgl. /16/).

Pfadmechanismen sind in vielen Systemen konzipiert bzw. implementiert. "The HitchHikers Guide" /17/ - um nur ein Beispiel zu geben - nennt die Pfade "Coach Tours" im Rahmen einer "Travel Holiday" Metapher. In dieser Lernumgebung können die Pfade an jedem "Tour Stop" - das sind multimediale Knoten - verlassen werden, um in das freie Browsing überzugehen. Später können sie dann von dem zuletzt besuchten "Tour Stop" aus wieder aufgenommen werden. In WITEC und dem Flusser-Projekt bildet der Basistext bzw. Vortragstext eine Art roten Faden durch das jeweilige System und damit auch einen - vom Autoren jedoch explizit vorgegeben - Pfad. In dieser linearen Abfolge von Knoten ist die eigentliche Grundinformation enthalten, während alle anderen Knoten ergänzende und vertiefende Informationen enthalten.

#### **4. Fazit**

Es lassen sich zu der Problematik der Navigation und Orientierung in Hypertext folgende Punkte festhalten:

- Eine universelle Festlegung, welche Navigationsstrategien bzw. Orientierungshilfen "gut" oder "schlecht" sind, ist nicht möglich, da die Entscheidung aufgabenspezifisch und im jeweiligen Zusammenhang zu fällen ist. Unterschiedliche Anforderungen ergeben sich sowohl aus dem Themengebiet (z.B. Geschichte vs. Biologie), der Informationsart (Forschungsarbeit vs. Lehrbücher) und der Zielgruppe bzw. deren tägliche Aufgaben (Studenten vs. Sachbearbeiter).
- Es ist der Tradeoff zwischen Komplexität - bzw. einer darauf beruhenden Überlastung des Systems und/oder der kognitiven Überbeanspruchung des Nutzers - und der erstrebten Flexibilität zu beachten. Daher ist der Nutzen einer Systemkomponente (z.B. Stringsuche) gegenüber dem durch ihre Implementierung bedingten Aufwand und der zwangsläufigen Mehrbelastung des Nutzers abzuwägen.
- Es müssen Werkzeuge zur Verfügung gestellt werden, die den Aufbau von Orientierungshilfen und deren Pflege erleichtern. Hier ist insbesondere der Einsatz der automatischen Indexierung zur Unterstützung der Texterschließung erfolgversprechend.

Besonders lohnend für weitere Arbeiten sind - ggf. elektronisch vorliegende - juristische Texte bzw. Quellen, da diese bereits eine relativ ausgeprägte Hypertextstruktur aufweisen ("law is hypertext by nature"). Daher sollte sich mit solchem Material relativ schnell ein ausreichend großes Hyperdokument erzeugen lassen, an dem verschiedene Navigationsstrategien und Orientierungshilfen getestet werden können.

## Literatur

- / 1 / Salton, G. u. McGill, M.J.: "Information Retrieval - Grundlegendes für Informationswissenschaftler". Hamburg: McGraw-Hill 1987.
- / 2 / Hjerppe, R.: "Project HYPERCATalog: Visions and Preliminary Conceptions of an Extended and Enhanced Catalog". In: Brookes, B.C. (Hrsg.), "Intelligent Information Systems for the Information Society", Elsevier: Amsterdam 1986, S. 211-232.
- / 3 / Bates, M.C.: "An Exploratory Paradigm for Online Information Systems". In: Brookes, B.C. (Hrsg.), "Intelligent Information Systems for the Information Society". Amsterdam: Elsevier 1986, S. 91-99.
- / 4 / Nelson, T.H.: "Getting it Out of Our System". In: Schecter, G. (Hrsg.), "Information Retrieval: A Critical View". Thompson Book Co. 1967, S. 191-210.
- / 5 / Lynch, C.A.: "Hypertext, Large Databases and Relational Database Management Systems". In: National Online Meeting. Learned Information:1989, S. 265-270.
- / 6 / Brown, P.J.: "Hypertext: The Way Forward". In: Vliet, J.C. van (Hrsg.), "Proceedings of the International Conference on Electronic Publishing, Document Manipulation and Typography". Cambridge University Press: Cambridge
- / 7 / Krüger, F.: "Navigation und Orientierung in Hypertext". Magisterarbeit an der FR Informationswissenschaft, Univ. d. Saarlandes: Saarbrücken 1990.
- / 8 / McAleese, R.: "Navigation and Browsing in Hypertext". In: McAleese, R. (Hrsg.), "Hypertext: Theory into Practice" Norwood, NJ: Ablex 1989, S. 6-44.
- / 9 / Parsaye, K. e.a.: "Intelligent Databases". Wiley: 1989.
- / 1 0 / Kluge G.: "Orientierung und Navigation in Hypertext-Netzwerken". In: Seminar-Unterlagen des Hauptseminars "Hypermedia" des Instituts für Telematik, Universität Karlsruhe am CEC Karlsruhe, Februar 1990 .
- / 1 1 / Frisse M. u. Cousins, S.B.: "Information Retrieval from Hypertext: Update on the Dynamic Medical Handbook Project". In: Hypertext '89 Proceedings, S. 199-212.
- / 1 2 / Hahn, U., e.a.: "Inhaltsorientierte Navigation in Automatisch Generierten Hypertext-Basen". In: Gloor, P.A. u. Streitz, N.A. (Hrsg.), "Hypertext und Hypermedia". Springer: Heidelberg 1990, S. 205-219.
- / 1 3 / Remde J.R, Gomez, L.M. u. Landauer, T.K.: "SuperBook: An automated tool for information exploration- Hypertext?". In: Hypertext '87 Papers, S. 175-88.
- / 1 4 / Mayes, T. , Kibby, M. u. Anderson, T.: "Learning about Learning from Hypertext". In: D.H. Jonassen u.a. (Hrsg.): Hypertext/Hypermedia for Learning, Kapitel 13. Springer 1990

- / 1 5 / Fairchild K.M., Poltrock, S.E. u. Furnas, G.W.: "SemNet - Three-dimensional Graphic Representations of Large Knowledge Bases". In: Guindon, R. (Hrsg.), "Cognitive Science and its Applications for Human-Computer Interaction". Erlbaum: Hillsdale, NJ 1988, S. 201-233.
- / 1 6 / Furuta, R. u. Stotts, P.D.: "Programmable Browsing Semantics in Trellis". In: Hypertext '89 Proceedings, S. 27-42.
- / 1 7 / Hammond, N. u. Allinson, L.: "The Travel Metaphor as Design Principle and Training Aid for Navigating Around Complex Systems". In: Diaper, D. u. Winder, R. (Hrsg.). "People and Computers III: Proc. of the Third Conference of the BCS HCI Specialist Group". Cambridge University Press: Cambridge 1987, S. 75-90.

**Anschrift des Autors:**

Frank Krüger, M.A.

Lebacher Str. 70

W-6600 Saarbrücken



# **Graphische Recherche und intelligente Hilfe als Komponenten des multimodalen Werkstoffinformationssystems WING**

Stephan Roppel, Christian Wolff

## **Zusammenfassung**

Ziel des Projektes WING-IIR ist die empirisch fundierte Entwicklung einer multimodalen Benutzerschnittstelle für auf relationalen Datenbanken basierende Werkstoffinformationssysteme. Dabei stehen Erkenntnisse über eine ideale Mischform der beiden "natürlichen" Anfragemodi, des natürlichsprachigen und des graphisch-direktmanipulativen, im Mittelpunkt des Interesses, ebenso wie die Integration von intelligenten Hilfe- und Retrievalkomponenten. Als neuartiges Retrievalwerkzeug wird die interaktive Manipulation von Graphen, die Werkstoffeigenschaften repräsentieren, vorgeschlagen.

## **Abstract**

The goal of the project WING-IIR is the empirically motivated development of a multimodal user-interface for relational materials information systems. WING-IIR focuses on a synergetic blending of the two "natural" modes of interaction: natural language and direct manipulation. In addition intelligent help and retrieval components form part of the information system. As a new retrieval-tool the interactive manipulation of graphs representing materials properties is suggested.

## **1 Einführung**

Ziel des vom BMWI geförderten Projektes WING-IIR (Förderkennzeichen: Wi 71250) der Linguistischen Informationswissenschaft Regensburg (LIR) ist die Entwicklung einer multimodalen Benutzerschnittstelle für Werkstoffinformationssysteme. Dabei steht die empirische Ermittlung einer für das gewählte Anwendungsgebiet - Faktenretrieval in relationalen Werkstoffdatenbanken - idealen Mischform der beiden "natürlichen" Interaktionsmodi, des natürlichsprachigen und des graphisch-direktmanipulativen, im Mittelpunkt. Parallel zur Entwicklung der eigentlichen Retrievaloberfläche, die u.a. auch Hypertext- und Query By Example-Elemente enthält, werden ein wissensbasiertes adaptives Hilfesystem und IIR-Komponenten integriert.

Der von WING-IIR verfolgte Entwicklungsansatz basiert auf dem Verfahren des rapid prototyping, der wiederholten Verbesserung vorläufiger Prototypen anhand der Ergebnisse von Benutzertests. Da als Grundlage der WING-Entwicklungen die Werkstoffdatenbank des Industriepartners mtu fungiert, geben Benutzertests mit Werkstoffexperten der mtu fortlaufend Aufschlüsse über die realen Anforderungen an Benutzeroberfläche und Funktionalität eines Werkstoffinformationssystems (cf. Gehrman/Marx 1991, Marx 1991, Schudnagis 1991B), die wiederum in die nächste Version des Prototypen eingehen.

## **2 Theoretische Grundannahmen**

Die WING zugrundeliegende Forschungshypothese ist, daß zur Gestaltung einer optimalen Benutzerschnittstelle für den Bereich Werkstoffinformation verschiedene "natürliche" Interaktionsmodi miteinander verbunden werden müssen. Diese Hypothese gründet in der Tatsache, daß die kognitiven Konzepte von Werkstoffexperten auf verschiedene Modi verteilt sind. So stellen sich etwa Werkstoffexperten bei parametrisierten Materialeigenschaften Funktionsgraphen vor, denken also graphisch-visuell, während globalere Werkstoffcharakteristika eher begrifflich repräsentiert sind (cf. Wolff 1991).

Der resultierende theoretische Anspruch, verschiedene Interaktionstechniken für unterschiedliche Anwendungssituationen auf der Benutzeroberfläche gleichzeitig verfügbar zu machen, bedeutet jedoch nicht, daß ein genaues Abbild der "multimodalen" zwischenmenschlichen Kommunikation angestrebt wird. Vielmehr sollen die technik- und anwendungsspezifischen Stärken und Schwächen der jeweiligen Interaktionsformen in variierenden Dialogsituationen ermittelt werden, um dann in einer integrierten Schnittstelle ein optimales Zusammenspiel verschiedener Modi zu erreichen (cf. Krause et al. 1990, wo die einzelnen Grundtypen näher erörtert werden). Einzelheiten der Verteilung der jeweiligen Modi innerhalb der multimodalen Schnittstelle waren anfangs völlig offen und ließen sich nur empirisch ermitteln.

Dazu wurde zuerst eine Serie von Benutzertests - bis zu 10 Versuchspersonen arbeiteten unter Videoaufzeichnung mit Vor- und Nachinterview - für die verschiedenen Modi jeweils unabhängig durchgeführt. Getestet wurden der direktmanipulative und der natürlichsprachlicher Zugang, hierarchische Suchpfade sowie Hypertext. Die Testergebnisse bestätigten die Annahme, daß verschiedene Dialogsituationen bzw. unterschiedliche Anfragebedürfnisse jeweils andere Interaktionswerkzeuge verlangen. Damit lag die empirische Grundlage für die Gestaltung einer ersten gemischten Version des Prototypen vor.

### 3 Struktur der multimodalen Benutzerschnittstelle

Das Konzept für die Kombination verschiedener Zugriffsmodi umfaßt zwei Phasen: In einem ersten Schritt wurde analytisch eine Leistungsübersicht der verschiedenen Zugriffswege erstellt, d.h. für jeden Zugriffsweg wurde ermittelt, welche prinzipiellen Informationsbedürfnisse mit ihm erfüllt werden können.

In einem zweiten Schritt wurden für die einzelnen Recherchetypen die Ergebnisse der Benutzertests herangezogen, um zu ermitteln, welcher Zugriffsweg für ein bestimmtes Informationsbedürfnis (z.B. Suche nach Werkstoffen anhand bestimmter Kriterien, Vergleich mehrerer Werkstoffe) am geeignetsten erscheint.

Entsprechend dieser Bewertung lassen sich in den multimodalen Prototyp folgende Komponenten bzw. Leistungsmerkmale der Einzelsysteme übernehmen:

Der **kognitiv-graphische Zugangsweg WING-KOGRA** (cf. Wolff 1991 und s.u.) dient als prinzipielle Grundlage des multimodalen Gesamtsystems, da er von den getesteten Zugangswegen die leistungsfähigste Anpassung an die verschiedenen Recherchebedürfnisse der Werkstoffexperten darstellt und gleichzeitig durch die Einbettung in eine direktmanipulative Arbeitsumgebung sehr benutzerfreundlich ist. Der funktionale Schwerpunkt liegt insbesondere auf der Faktenrecherche, d.h. den nach kognitiven Gesichtspunkten strukturierten Recherchemöglichkeiten für numerische Werkstoffdaten.

Der **Hypertextzugang** (cf. Schudnagis 1991a) zur mtu-Datenbank ist nicht als globaler Zugriffsmodus, sondern nur für bestimmte Informationstypen und Recheresituation integriert. Hypertext findet insbesondere an solchen Stellen Verwendung, wo das Ziel der (nicht-typisierten) Links eindeutig vorhersehbar ist, d.h. der Benutzer eine hohe Interpretationswahrscheinlichkeit für das Ergebnis des jeweils nächsten Interaktionsschrittes hat.

Dies gilt insbesondere für textuelle Daten (Werkstoffbeschreibungen als Lang- und Kurztexe in Datentabellen) sowie das sog. "Datenblatt", aus dem heraus verschiedene Links sowohl auf textuelle als auch auf numerische Information angeboten werden.

Als globales Ordnungsprinzip für das Werkstoffinformationssystem scheidet der Hypertextmodus aus, da sich in den Benutzertests zeigte, daß für konkrete Recherchevorhaben zu wenig Führung durch das System geboten wurde. Insbesondere für die Datenrecherche mit ihren klar umrissenen Retrievalstrategien ist das für Hypertext charakteristische assoziative Vorgehen mit der relativ hohen Gefahr, sich im Hypertext-Netz zu verirren ("to be lost in hyperspace", cf. Schudnagis 1991b), weniger geeignet.

Der Zugriff durch **Query By Example (QBE)** (cf. Lickleder 1990) schied nach informellen Tests schon frühzeitig als genereller Zugangsmodus aus, da er sich aufgrund der prinzipiell erforderlichen Anlehnung an die abstrakte Tabellenstruktur der Datenbank schlecht als Einstiegsmodus eignet. Das Prinzip QBE findet aber als eingebetteter Modus im integrierten Prototyp Verwendung: QBE bietet sich **im Anschluß** an eine (in einem anderen Modus) durchgeführte Recherche zur Weiterarbeit an, da der Benutzer aus einer Ausgabetabelle sofort eine neue Anfrage zusammensetzen kann, indem er die Ausgabetabelle als Anfragemaske ausfüllt. Problematisch bleibt auch für diesen Fall die SQL-bedingte Verknüpfung mehrerer Datentabellen in einer Anfrage (join).

Der **hierarchische Zugang** modelliert die auch bei mtu als mainframe-Lösung implementierte Datenbankschnittstelle mit den Mitteln einer graphischen Benutzerumgebung. Seine streng vorgegliederte Eingabeabfolge geht ebenfalls partiell in die integrierte Systemlösung ein: Zum einen wird die von Versuchspersonen positiv beurteilte starke Systemführung in die Anfragemasken des kognitiv-graphischen Modus übernommen, so daß ein Höchstmaß sinnvoller Systemführung angeboten werden kann, ohne dabei die Inflexibilität der rein hierarchischen Struktur zu übernehmen. Zum anderen wird die Anfrage- und Anzeigezeile des hierarchischen Modus als Status- und Übersichtsinformation in die multimodale Shell übernommen und repräsentiert den jeweiligen Zustand des Anfrageaufbaus.

Für die Integration der **natürlichsprachlichen Anfragekomponente** sind mehrere Anwendungsbereiche vorgesehen, für präzisere Aussagen müssen jedoch Tests mit dem an die Anwendungsdomäne angepaßten Parser, der zur Zeit noch nicht vorliegt, abgewartet werden.

Die Anpassung des Parsers, die unter Berücksichtigung der "computer talk"-Hypothese vorgenommen wird, d.h. der Annahme, daß die Mensch-Maschine-Interaktion keine 1:1-Abbildung der menschlichen Kommunikation darstellt, soll folgende Einsatzgebiete der natürlichsprachlichen Anfrage ermöglichen:

- a) **Natürlichsprachlicher Zugang** als Ersteinstieg, um v.a. ungeübten Benutzern die Formulierung ihrer Recherche zu ermöglichen (Laienmodus). Unter kognitiven Gesichtspunkten liegt es nahe, daß zumindest die anfängliche Problemformulierung sprachlich vorgenommen wird. Daher dürfte sich der natürlichsprachliche Zugangsmodus gut als Einstiegsmöglichkeit eignen, auch wenn der Benutzer nach dem Einstieg in einem anderen Modus fortfahren sollte.
- b) **Natürlichsprachlicher Zugriff** als Alternative zu anderen Rechercheformen (mit Umstiegsmöglichkeit). Dabei sind zwei Lösungen denkbar:

- b1) An bestimmten, geeigneten Punkten innerhalb anderer Modi wird natürlichsprachlicher Zugang zur Fortführung einer direkt-manipulativ aufgebauten Recherche eingesetzt. Die mtu-Tests haben ergeben, daß bestimmte Teile der Recherche (z.B. Nebenbedingungen der Profilsuche) von den Versuchspersonen spontan natürlichsprachlich formuliert wurden.
- b2) Natürlichsprachlicher Zugang existiert als an jedem Punkt einer Recherche angebotene Überstiegsmöglichkeit. Diese Konzeption stellt die vollständige Erweiterung des Vorschlags von b1) dar. Hier ist dem Benutzer der Wechsel in den natürlichsprachlichen Modus prinzipiell zu jedem Zeitpunkt möglich. Tests müssen zeigen, ob Lösung b1) ausreichend ist, oder ob eine Umstiegsmöglichkeit immer vorzusehen ist.
- c) Natürlichsprachliche Eingabe als ständig verfügbarer Hilfeanforderungsmodus. Für diese Aufgabe müßte allerdings das Lexikon des Parsers um systembezogene Terminologie erweitert werden. Dieser Vorschlag beruht auf der Beobachtung, daß Benutzer bei Bedienungsschwierigkeiten gerne auf eine sprachliche Formulierung ihres Problems zurückgreifen, so daß sich der natürlichsprachliche Zugang als (zusätzliche) Schnittstelle zum Hilfesystem anbietet.

Für die einzelnen Werkstoffen zugeordneten Langtexte ist neben der Erschließung durch Hypertext ergänzend der Zugang durch **Dokumentretrieval mit boolschen Operatoren** (auf der Basis kommerzieller Software) vorgesehen; dieser Zugang bietet die Möglichkeit, nach Schlüsselbegriffen im Volltext zu recherchieren und ist vor allem als Erstzugriff zur Textselektion gedacht. Innerhalb der auf eine Anfrage gefundenen Texte kann mit Hypertextlinks navigiert werden.

Die empirische Grundlage der Konzeption des multimodalen Prototypen sind Benutzertests der jeweils parallel entwickelten Modi. Neue Erkenntnisse zu einer verbesserten Integration der verschiedenen Modi, die die jeweiligen Nachteile eliminiert, Übergänge von einem Modus zum anderen erleichtert und ihre Stärken synergetisch nutzt, werden durch weitere Prototyping-Zyklen gewonnen werden. Im Folgenden sollen exemplarisch zwei Komponenten des integrierten Prototypen, die graphisch-direktmanipulative Rechercheform sowie die Hilfekomponente WINGHELP, näher erläutert werden.

#### **4 Kognitiv motivierte Recherchetypen und graphisches Retrieval (KOGRA)**

Der Zugangsweg KOGRA hatte sich zunächst auf die Modellierung typischer Problemstellungen im Werkstoffaktenretrieval im Kontext einer graphischen Benutzerschnittstelle beschränkt.

Basierend auf einer Analyse charakteristischer Probleme von Werkstoffexperten wurde der Problembereich unter Beachtung softwareergonomischer Kriterien mit folgenden Suchtypen modelliert:

- Suche nach Daten zu einem Werkstoff,
- Vergleich mehrerer Werkstoffe,
- Suche nach Werkstoffen, die ein gegebenes Anforderungsprofil erfüllen,
- Umschlüsselung von Werkstoffnamen
- Recherche über ein Datenblatt, das globale Werkstoffinformation enthält.

Für jeden dieser Suchtypen wurden eine bzw. mehrere Abfragemasken in einer direktmanipulativen Arbeitsumgebung entworfen.

Benutzertests bestätigten die Eignung der kognitiv motivierten Suchtypengliederung für die Werkstoffdatenrecherche, insbesondere im Vergleich zu hierarchischen Suchpfaden. Dabei war als Ausgabemodus zunächst nur textuelle und numerische Präsentation von Werkstoffinformation implementiert.

Sowohl die Benutzertests als auch eine Analyse der "traditionellen" Informationsmedien im Werkstoffbereich (Datenhandbücher, Werkstoffblätter etc.) zeigten aber, daß die graphische Aufbereitung numerischer Ergebnisse eine zentrale Rolle bei der Interpretation und der Bewertung der Rechercheergebnisse spielt. Die weitaus überwiegende Zahl der in der Datenbank des Anwendungspartners mtu enthaltenen Werkstoffkennwerte verlangt nach graphischer Aufbereitung, um eine adäquate Interpretation zu ermöglichen.

Neben der Implementierung einer kennwertspezifischen, statischen Standardgraphik, wie sie auch kommerzielle Anwendungsprogramme anbieten, steht damit für die nächste Version des multimodalen Prototypen eine dynamisierte Ausgabe- und Abfragegraphik im Mittelpunkt.

Ausgehend von der zentralen Bedeutung graphischer Information bei der Interpretation von Werkstoffdaten wird die Rolle von Graphik als reinem Ausgabemodus erweitert: die Graphik wird dynamisch manipulierbar und ersetzt bzw. ergänzt die formal- bzw. natürlichsprachliche Abfrageformulierung. Anhand einer Ausgabegraphik zu einem bestimmten Werkstoff kann der Benutzer durch direktmanipulatives Verändern der Graphik selbst eine neues, bildlich repräsentiertes Abfrageschema erstellen. Dabei gibt es innerhalb des Konzeptes manipulierbarer Graphik eine Reihe von Gestaltungsalternativen, die jeweils in unterschiedlichem Maße systemseitige Intelligenz voraussetzen: im einfachsten Fall gibt der Benutzer lediglich Bereiche an, in denen gesuchte Werkstoffe eine Ausgabegraphik "übertreffen" oder "unterbieten" sollen; die Recherche gibt dem Benutzer dann alle entsprechenden Werkstoffe bzw. die passenden Graphiken aus.

Im Falle der Veränderung einer Kurve selbst muß das System Wissen zu den einzelnen Kenn-

werten (graphische bzw. rechnerische Aufbereitung von Werkstoffdaten) besitzen, um dem Benutzer einen adäquaten Rahmen seiner Manipulationsmöglichkeiten bieten, bzw. offensichtlich "unsinnige" Kurvenveränderungen frühzeitig abfangen zu können.

Mit diesem Konzept einer dynamisierten Ausgabe- und Abfragegraphik läßt sich ein einheitlicher Arbeitsmodus für den Benutzer erreichen; ob der kognitiv sinnvolle Ausgabemodus Graphik sich in vollem Maße auch als Abfragemodus erfolgreich einsetzen läßt, müssen erneute Benutzertests zeigen. Diese Tests werden auch Aufschluß darüber geben, mit welchen Elementen anderer Modi (natürlichsprachlicher Zugang, Hypertext, Query By Example etc.) graphisches Retrieval kombiniert werden sollte.

## **5 Die aktiv-adaptive Hilfefunktion WINGHELP**

Mit der Entwicklung benutzerfreundlicher, an den kognitiven Konzepten der Anwender orientierter Schnittstellen, entfällt nicht die Notwendigkeit, dem Benutzer in Problemsituationen oder bei Fehlern, die in Systemen entsprechender Komplexität unvermeidlich sind, Hilfe zur Verfügung zu stellen. Deshalb erfolgte parallel zur Entwicklung der Retrievaloberfläche die Integration eines aktiven, adaptiven Hilfesystems (AAHS), WINGHELP, das den Benutzer in bestimmten Situationen mit Vorschlägen zur Weiterarbeit unterstützt (cf. Obermaier/Roppel 1991).

Als Basis für jede intelligente Hilfefunktion, die den Benutzer kontextspezifisch und adaptiv unterstützt, ist ein herkömmliches, mit Register und feststehenden Hilfetexten versehenes System unabdinglich. Hilfe zu aktuellen Bildschirmobjekten, verschiedene Stufen der Ausführlichkeit der Hilfetexte, Hypertextverbindungen und Browsing-Sequenzen zeichnen die entsprechende Komponente von WINGHELP aus. Erst die Koppelung mit einer auf Planerkennung aufbauenden aktiven und adaptiven Komponente macht jedoch die Unterstützung des Benutzers auch in Situationen, in denen dieser seine Probleme nicht selbst erkennt oder benennen kann, möglich.

Während der Arbeit mit dem Werkstoffinformationssystem beobachtet der WINGHELP-Planerkennner die Aktionen des Benutzers und ordnet sie vorher festgelegten Benutzerplänen zu. Diese Pläne bilden die funktionale Struktur der Benutzeroberfläche ab und sind zusätzlich durch Auswertung der Benutzertests empirisch abgesichert. Die aktuellen Eingaben des Benutzers werden dahingehend bewertet, ob sie einen Plan optimal, fehlerhaft oder umständlich erfüllen. Wenn nötig, werden gleichzeitig falsche Konzepte über die Funktionsweise des Systems festgestellt, die mit Hilfmeldungen korrigiert werden können.

WINGHELP unterscheidet zwei Fälle, bei denen es in Aktion tritt: Fordert der Benutzer durch die Wahl einer Hilfetaste während der Arbeit Unterstützung, so bietet WINGHELP bei erfolgrei-

cher Planerkennung Hilfe zu genau dem erkannten Plan, in der Regel also Information zur erfolgreichen Fortsetzung einer Aktion. Ist zum Zeitpunkt der Hilfeanforderung kein Plan erkannt, wird das allgemeine Befehls-Register von WINGHELP angeboten, das zu einem hypertext-basierten Hilfesystem überleitet. Die augenblickliche Struktur der WING-Benutzeroberfläche - typische Anfragebedürfnisse sind als Funktionskomplexe modelliert - erleichtert das frühe Erkennen von Benutzerplänen, und die Hilfe zur Fortsetzung eines begonnenen Plans scheint nach Auswertung der Benutzertests die am häufigsten auftretende Unterstützungssituation zu sein.

Werden die Benutzereingaben jedoch als fehlerhaft oder offensichtlich umständlich bewertet, so schaltet sich WINGHELP ohne Benutzeraufforderung automatisch mit einer Hilfemeldung in den Dialog ein. Sie enthält den Auslöser der Hilfeleistung sowie Alternativ-Vorschläge, um damit zukünftigen Fehlern oder Umständlichkeiten vorzubeugen.

In beiden Fällen, Hilfe nach Benutzeranforderung bzw. automatische Reaktion von WINGHELP, kann der Benutzer erweiterte Information zu den vorgeschlagenen Aktionen durch kontextsensitiven Überstieg in das Hypertext-Hilfesystem erhalten.

Zusätzlich sind auch für solche Probleme, die sich nicht auf die Interaktionsebene, sondern auf die eigentlichen Rechercheergebnisse beziehen, intelligente Unterstützungsmechanismen, die unter den Begriff Intelligentes Information Retrieval subsumiert werden können, notwendig. Diese Form der Unterstützung betrifft z.B. Situationen, in denen der Benutzer zu einer bestimmten Problemformulierung keine Ergebnisse in der Datenbasis recherchieren kann. Eine intelligente Recherchehilfe sollte, basierend auf Werkstoffwissen und Wissen zu Datenbankanhalten, Vorschläge zu Rechercheumformulierung oder alternativen Suchstrategien enthalten.

Die Erweiterung der Wissensbasis von WINGHELP, die sich im Moment lediglich auf Wissen zur Bedienung der Benutzeroberfläche bezieht, um Wissen zu Datenbankanhalten und Frageformulierungen und die zusätzliche Einbindung von werkstoffbezogenem Expertenwissen bietet für die Zukunft die Möglichkeit, auch IIR-Komponenten in den WINGHELP-Ansatz zu integrieren.

## **6 Stand der Implementierung**

Die Datenbankgrundlage von WING-IIR bildet eine strukturell vollständige Untermenge der Datenbank des Anwendungspartners mtu, erweitert um exemplarische Text-Datensätze (Datenblatt und Langtexte). Es handelt sich um eine relationale Datenbank im SQL-Standard, die auf verschiedenen Systemplattformen vorliegt (SQLBase, Paradox und Oracle unter MS-DOS, Oracle unter UNIX).



Mit Ausnahme der auf SUN-Workstations implementierten Parsingkomponente dienen PCs als Entwicklungsplattform, wobei die verschiedenen Zugriffsmodi zunächst als Einzelsysteme wie folgt realisiert wurden:

Der Parser der natürlichsprachlichen Komponente wurde innerhalb des Projektverbunds durch CIS (München) als DCG-System in Prolog-II entwickelt und wird derzeit an die inhaltlichen und sprachlichen Anforderungen der Werkstoffdomäne angepaßt.

Der hierarchische sowie der kognitiv-graphische Zugangsweg wurden mit dem Software Development Toolkit von MS-Windows 3 (Windows SDK) implementiert.

Der Hypertext-Modus wurde mit dem Applikationsgenerator TOOLBOOK, ebenfalls unter MS-Windows 3, entwickelt.

Die QBE-Komponente wurde mit dem Datenbanksystem PARADOX realisiert und liegt gleichzeitig, in exemplarischer Form, als MS-Windows-fähige SQLWindows-Applikation vor.

Der erste multimodale Prototyp kombiniert zunächst die beschriebenen Unterausschnitte der einzelnen Zugriffsmodi, wobei die technische Integration unter einer MS-Windows basierten Shell schrittweise verfeinert wird. Die natürlichsprachliche Komponente muß für den multimodalen Prototyp zunächst noch simuliert werden, da das Anforderungsprofil für die Werkstoffdomäne noch nicht umgesetzt wurde (cf. Marx 1991).

## **7 Fazit**

Mit der kognitiv motivierten und empirisch abgesicherten Integration verschiedener Recherchemodi sowie deren Ergänzung durch intelligente Hilfefunktionen liegt ein fundierter Entwurf für eine multimodale Werkstoffdatenbankschnittstelle vor, der den kognitiven und inhaltlichen Anforderungen einer komplexen Anwendungsdomäne gerecht wird.

## **Literaturangaben**

Arents, Hans; Bogaerts, Walter (1989). Use of Hypertext for Corrosion Analysis. In: Jovanovic, A., Kussmaul, A.C et al. (eds.). Expert Systems in Structural Safety Assessment. Lecture Notes in Engineering, Vol. 53. Berlin, Heidelberg, New York. S. 375-389.

Barbieri, Gilberto; Colazzo, Luigi; Molinari, Andrea (1991). Relational Back-Ends in the Management of Large Hypertexts. In: Maurer, H. (ed.). Hypertext/Hypermedia '91. Berlin et al. Springer, S.34-51 (= IFB 276).

Gehrmann, Uli; Marx, Jutta(1991). Die erste Serie von Benutzertests des ersten WING-IIR Prototypen. Universität Regensburg, Linguistische Informationswissenschaft, Projekt WING-IIR, Arbeitsbericht 13, Mai 1991.

- Hanne, K.H., Hoepelman, J.P. (1988). Man Computer-Interfaces Combining Graphic- and Natural Language-Interaction. In: Bullinger et al. (eds.). EURINFO '88. Proceedings of the First European Conference on Information Technology for Organisational Systems. Athens, Greece, Amsterdam et al. S. 410-415.
- Krause, Jürgen et al. (1990). The research prototype of a multimodal materials information system, comprising natural language-, graphical/direct manipulation- and knowledge based components. In: Herget, J., Kuhlen R. (eds.) Pragmatische Aspekte beim Entwurf und Betrieb von Informationssystemen. Proceedings des 1. Internationalen Symposiums für Informationswissenschaft, 329-338.
- Krause, Jürgen (1991). The Combination of two "natural" modes of human-computer interaction: natural language and graphics. In: AAAI (ed.) Workshop Notes from the Ninth National Conference on Artificial Intelligence (AAAI-91). Workshop on Intelligent Multimedia Interfaces, Anaheim/CA., July, 15, 1991, 47-52
- Lickleder, Gerhard (1990). Softwareergonomische Überlegungen zur Abfragesprache Query-by-Example auf der Basis der Implementierung eines Werkstoffdatenbanksystems mit PARADOX. Universität Regensburg, Linguistische Informationswissenschaft, Projekt WING-IIR, Arbeitsbericht 19, Juli 1991.
- Lutz, Jutta (1990). Natürliche Sprache als Abfrageform von Werkstoffdatenbanken. Universität Regensburg, Linguistische Informationswissenschaft, Projekt WING-IIR, Arbeitsbericht 2, 1990.
- Marx, Jutta (1991). Benutzertests zur natürlichsprachlichen Komponente von WING-IIR. Testaufbau, Ergebnisse und Anforderungsprofil für den Parser. Universität Regensburg, Linguistische Informationswissenschaft, Projekt WING-IIR, Arbeitsbericht 2, 1990.
- Obermaier, Christoph; Roppel, Stephan (1991). WINGHELP. Eine aktive adaptive Hilfefunktion für den Prototypen WING eines multimodalen Werkstoffinformationssystems. Universität Regensburg, Linguistische Informationswissenschaft, Projekt WING-IIR, Arbeitsbericht 16, Juli 1991.
- Schudnagis, Monika (1991A). Der Hypertextzugang zur mtu-Werkstoffdatenbank. Universität Regensburg, Linguistische Informationswissenschaft, Projekt WING-IIR, Arbeitsbericht 15, Juli 1991.
- Schudnagis, Monika (1991B). Die Benutzertests des Hypertextzugangs zur mtu-Werkstoffdatenbank. Universität Regensburg, Linguistische Informationswissenschaft, Projekt WING-IIR, Arbeitsbericht 20, (erscheint).
- Wolff, Christian (1991). Der Prototyp 2 des kognitiv graphischen Zugangswegs von WING-IIR. Universität Regensburg, Linguistische Informationswissenschaft, Projekt WING-IIR, Arbeitsbericht 17, Juli 1991.

#### **Adresse der Autoren:**

Stephan Roppel, M.A.  
Christian Wolff, M.A.

Universität Regensburg  
FG Linguistische Informationswissenschaft (LIR)  
Postfach 397  
W-8400 Regensburg  
Tel: (0941) 943-3588  
FAX: (0941) 943-2305  
email: wolff@vax1.rz.uni-regensburg.dbp.de

# **Automatische Trunkierung beim Zugang zu textbezogenen Informationsbanken**

**Harald H. Zimmermann, Saarbrücken**

## **Zusammenfassung**

Die Trunkierung stellt bei der Suche mit Stich- und Schlagwörtern in elektronischen Textablagen, in Textdatenbanken bzw. Textfeldern bibliographischer Datenbanken (Titeln, Abstracts ...) - ob online oder lokal - ein wichtiges Mittel dar, das Resultat einer Recherche zu verbessern.

Ziel der vorgestellten Verfahren ist es, dem Benutzer eine klar definierte Menge von Varianten der Trunkierung bereitzustellen, die den Möglichkeiten seines Retrievalsystems gerecht wird.

Behandelt (und differenziert) werden u.a. die natürliche Trunkierung, die derivative Trunkierung, und die Integration von Wortrelationierungen mit einer Anzahl von Sonderformen.

Ein Ziel der Untersuchung ist es, über die Differenzierung der Möglichkeiten zur einer inhaltlichen Präzisierung dieses Themenkomplexes beizutragen, um der Informationstechnik eine Grundlage für sachgerechte Implementierungen zu geben.

Die Verfahren zur automatischen Trunkierung werden beispielhaft beschrieben.

## **Abstract**

For the retrieval with catchwords or keywords in electronic text archives, textual data bases and/or textual records in bibliographic data bases (titles, abstracts etc.), either online or local, TRUNCATION is an important tool to improve the results of a research process.

The aim of the procedures presented is to give the user a precisely defined set of TRUNCATION variants which fits the possibilities of the used retrieval system.

Natural TRUNCATION, derivative TRUNCATION and the integration of word RELATIONS with special variants are described.

One objective of the study is to contribute to a more precise contentual description of the theme to get a constructive basis for software implementation in information industry.

The main procedures and techniques of automatic truncation are described.

## 1. Aufgabenstellung

Unter einer textbezogenen Informationsbank wird eine Datenbank bzw. ein Archivierungssystem verstanden, die Texte in einer natürlichen Sprache enthält. Hierzu rechnen Volltextdatenbanken und Datenbanken mit Textelementen (z.B. mit Titeln, Abstracts). Der Zugang zu diesen Daten kann im Dialog mit einem Informationssystem über Datenfernübertragung (online) oder auch lokal (z.B. CD-ROM) erfolgen.

Als wesentliche Eigenschaft des Zugangssystems wird die Möglichkeit vorausgesetzt, eine Freitextsuche (Textretrieval) mit Bezug zu den Wörtern durchzuführen, die in den Texten der Datenbank vorkommen.

Es wird angenommen, daß das Retrievalsystem über die Möglichkeit verfügt, Textelemente trunkiert zu suchen. Unter einer **TRUNKIERUNG** wird im folgenden der Vorgang verstanden, eine oder mehrere Zeichenketten (Suchstrings) zu einem Flexionsparadigma eines Wortes zu bilden und dabei anzugeben, wie viele Zeichen rechts davon (d.h. am Ende der vorgegebenen Zeichenkette) jeweils noch zugelassen sein sollen.

Dieses Verfahren wird in praktisch allen existierenden Textdatenbanken bzw. bibliographischen Datenbanken technisch unterstützt, ergänzt um Möglichkeiten zu Abstandsangaben (bei Wortfolgen). Die Praxis zeigt, daß entweder aus Bequemlichkeit oder aus mangelndem Problembewußtsein - z.T. auch aus technischer Unkenntnis - diese Funktionen nur unzureichend genutzt werden. Die beschriebenen Funktionen sind in erster Linie zum Einsatz beim Nutzer gedacht (Implementierung auf PC / Workstation), können jedoch im Prinzip auch im Retrievalsystem selbst implementiert werden.

Es gibt verschiedene Möglichkeiten der Trunkierung. Ausgangspunkt der Betrachtung ist eine Trunkierung, die auf den "natürlichen" Stammformen (d.h. ausgewählten Stellvertretern für Wortformen innerhalb einer Wortklasse) aufbaut. Diese Trunkierung wird natürlichsprachige Trunkierung bzw. kurz **natürliche Trunkierung** genannt.

(1) Beispiele für eine **natürliche Trunkierung** (\$ = Trunkierungszeichen, Ziffer = Trunkierungslänge; + = weiterer - alternativer - Suchstring = Alternativstring; OR-Funktion beim Retrieval):

<u>Trunkierung</u>	<u>Flexionsparadigma</u>
Gebäude \$1	Gebäude, Gebäuden, Gebäudes
Werkzeug \$2	Werkzeug, Werkzeuge, Werkzeugs, Werkzeugen
Stall \$2 + Ställe \$1	Stall, Stalls, Stalles, Stalle, Ställe,
Ställen	
Thema \$2 + Themen	Thema, Themas, Themata, Themen
machine \$1	machine, machines
child \$5	child, child's, children, children's
mouse \$2 + mice	mouse, mouse's, mice
baby \$2 + babies	baby, baby's, babies
maison \$1	maison, maisons
bateau \$1	bateau, bateaux
cheval + chevaux	cheval, chevaux

Das Beispiel zeigt, daß es bei der natürlichen Trunkierung häufig nicht genügt, nur einen einzigen Suchstring anzugeben, um alle Flexionsformen zu erreichen. Dies zeigt sich besonders bei den sog. "starken" Verben (es sei hier einmal außer Acht gelassen, inwieweit Verben in realen Retrievalsituationen überhaupt eine Rolle spielen; vgl. dazu Abschnitt 3 (1)):

(2) Beispiele für Verbtrunkierungen (Auszüge):

vergeh \$5 + verging \$2,	vergehe, vergeht, vergehende, ...
+ vergang \$5	vergangen, vergangene, ...
seh \$5 + sah \$2	sehe, sehen, seht, ..., sah, sahst, ...
+ sieh \$2 + säh \$3	sieh, sieht, siehst, sähest, ...

In der Praxis des Retrieval ist man bemüht, - wenn möglich - nur eine einzigste Zeichenkette (Suchstring) zu benutzen. Dabei spielt die natürliche morphologische Trunkierung - die sich beispielsweise an Ablaut- oder Umlautregeln orientiert - nur eine sekundäre Rolle. Weit eher sucht man ein Optimum aus der möglichst geringen Anzahl der Suchstrings und einer möglichst geringen Zuordnung ungewollter Treffer (Überidentifikation) zu ermitteln. Bei den o.a. Verben kann man es kaum riskieren, eine Kette VERG zugrunde zu legen, da man bei einer Trunkierung von der an sich notwendigen Länge 8 zu viele irrelevante Treffer erhalten würde. (etwa auch: vergaß, vergelten, ...)

Demgegenüber gibt es v.a. im nominalen Bereich - d.h. bei Substantiven und Adjektiven - genügend Fälle (und gute Gründe), die eine solche Strategie rechtfertigen. Diese Trunkierung soll im folgenden formale Trunkierung heißen:

(3) Beispiele für formale Trunkierungen:

Gymnasi \$2	Gymnasium, Gymnasien
Them \$3	Thema, Themen, Themata
galax \$3	galaxy, galaxies
matri \$3	matrix, matrices
chev \$3	cheval, chevaux
blanc \$3	blanc, blancs, blanche, blanches

Mit Bezug auf die Praxis ist davon auszugehen, daß eine wortklassenbezogene Trunkierung (wie in den Beispielen (1) - (3)) beim Retrieval selten ausreicht. Vielmehr sind zumindest Derivationen (Wortableitungen) zu berücksichtigen, die vielfach ohne Bedeutungsänderung die Trefferquote erhöhen können (zum Problem der Mehrdeutigkeit vgl. Abschnitt 3.2). Hierzu gehören im Deutschen beispielsweise Suffigierungen auf -ung oder -heit, im Deutschen, Englischen und Französischen Suffigierungen auf -ion. Eine Trunkierung unter Berücksichtigung dieser Fragestellung heißt im folgenden derivative Trunkierung.

Um Fehler zu vermeiden, sollte bei der Generierung von Suffixen (optional) eine lexikalische Kontrolle durchgeführt werden können.

(4) Beispiele für derivative Trunkierung:

untersuch \$5	untersuchen, untersucht, ...,
+ Untersuchung \$2	Untersuchung; Untersuchungen
lieblos \$4	lieblos, liebloses, ..., lieblosesten
+ Lieblosigkeit \$2	Lieblosigkeit, Lieblosigkeiten
hopeful \$2	hopeful, hopefully
+ helpfulness	helpfulness
inform \$5	informe, informant, ... informaient
+ information \$1	information, informations

Aufgabe und Ziel der automatischen Trunkierung ist es, Funktionen für die entsprechenden Formen bereitzustellen.

Als zusätzliche Anforderungen sollen erörtert werden:

- (5) Möglichkeiten der Selektion und v.a. Tilgung von Wortformen - v.a. für Trivialwörter
- (6) Einbeziehung von Synonymen und anderen Begriffsbeziehungen
- (7) Behandlung von Mehrdeutigkeiten während des Generierungsprozesses.

## 2. Verfahrensweise

### 2.1 Die elektronischen Wörterbücher

Es wird nach den Erfahrungen der letzten Jahre mit rein formalen Verfahren nicht als sinnvoll - v.a. nicht als zukunftssträftig - angesehen, Trunkierungen rein graphematisch (d.h. unter Auswertung von Buchstabenfolgen) durchzuführen. Grundlage der folgenden Überlegungen zur automatischen Trunkierung sind daher **elektronische Wörterbücher**, die folgende Eigenschaften aufweisen:

#### (A) Flexionsmorphologisches Wörterbuch

Das Flexionsmorphologische Wörterbuch (im folgenden auch Identifikationswörterbuch genannt) umfaßt wortklassenspezifische Identifikationswortlaute mit Angaben zu den zugehörigen (zugelassenen) Endungen sowie Verweise auf sog. Grundformen. Eine **Grundform** ist eine ausgezeichnete (= ausgewählte) Flexionsform eines flexionsmorphologischen Paradigmas im o.a. Sinne. Die Grundform bildet einerseits den Schlüssel zu den Einträgen des Relationenwörterbuchs (vgl. B), andererseits verknüpft sie mehrere Identifikationswortlaute im flexionsmorphologischen Wörterbuch. Eine Suche im Flexionsmorphologischen Wörterbuch hat zum Ziel, alle möglichen Grundformen zu einer (vom Benutzer bzw. System vorgegebenen) Wortform zu ermitteln. Das Flexionsmorphologische Wörterbuch wird bei der Trunkierung einerseits zur Identifikation und Grundformenermittlung herangezogen, andererseits aber auch bei der Ermittlung der Restzeichenketten (Endungen) genutzt. Auf eine detaillierte Beschreibung der Strategien muß hier aus Platzgründen verzichtet werden.

#### (B) Relationenwörterbuch

Ein weiteres wichtiges lexikalisches Instrument ist ein Wörterbuch, das - ähnlich zu einem "klassischen" Thesaurus - begriffliche (semantisch-syntaktische) Beziehungen zwischen Wörtern (Grundformen ...) aufführt. Es soll im folgenden Relationenwörterbuch heißen. Die Relationen sind typisiert. Folgende Typen sind im vorliegenden Zusammenhang von Bedeutung, zusätzlich wird angenommen, daß die Wortklassen (Adjektiv, Sub-



stantiv, Verb ...) markiert sind:

- (1) **Relation Grundform / Alternativsuchstring**  
gehen / ging; (ge)gangen; Haus / Häuser; Thema / Themen; The-  
mata  
mouse / mice; gentleman / gentlemen  
cheval / chevaux; monsieur / messieurs
- (2) **Relation Alternativsuchstring / Grundform** (Inversion von 1)  
ging / gehen; gangen / gehen; Häuser / Haus; Themen / Thema  
mice / mouse; chevaux / cheval
- (3) **Derivationsrelation** (Sub-Typisierung anhand der Wortklassenken-  
nungen)  
lieblos / Lieblosigkeit; Lieblosigkeit / lieblos  
ahnen / Ahnung; Ahnung / ahnen  
hopeful / helpfulness; helpfulness / hopeful  
admirer / admiration; admiration / admirer
- (4) **Dekompositionsrelation** (Auszug)  
Haustür / Tür; verstehen / stehen (!); Armbanduhr / Uhr  
underestimate / estimate; intercontinental / continental  
disproportion / proportion

## 2.2 Die Verfahrensschritte

Im folgenden werden mögliche Verfahrensschritte beispielhaft beschrieben. Ein entsprechendes Verfahren ist in Entwicklung und weitgehend abgeschlossen.

## **(1) Das Identifikationsverfahren**

Auf eine detaillierte Beschreibung muß hier verzichtet werden. Das Identifikationsverfahren sollte in jedem Falle liefern:

- (a) alle möglichen Grundformen und
- (b) - soweit die Identifikation über Derivation oder Dekomposition erfolgte - die "letzten" Bestandteile eines Kompositums
- (c) die Wortklasse(n)

## **(2) Die Ermittlung von Alternativstrings (= Alternativsuchstrings)**

Je (möglicher) Wortklasse anschließend mit Hilfe der Relationierungen des Typs Grundform / Alternativstring (wortklassenbezogen) alle möglichen Identifikationsstrings ermittelt. Details können hier aus Platzgründen nicht vorgestellt werden.

## **(3) Flexionsformengenerierung und Trunkierung**

Ausgehend von den identifizierten (alternativen) Zeichenketten werden die möglichen Flexionsformen generiert. Dies erfolgt beispielsweise unter erneutem Zugriff auf das Flexionsmorphologische Wörterbuch, das schon zur Identifikation einer gültigen Wortform entsprechende Markierungen enthalten muß.

Die erzeugten Flexionsformen werden (unter Bewahrung der Informationen des Identifikationsstamms) physisch erzeugt und ggf. kumuliert. In dieser Phase können z.B. Restriktionen über Wortklassen oder auch anhand des Selektionswörterbuchs berücksichtigt werden. Auch hier muß auf eine detaillierte Beschreibung verzichtet werden.

## **(4) Derivationsgenerierung und derivative Trunkierung**

Die Derivationsgenerierung wird i.d.R. vor der Ermittlung der Alternativstrings angesprochen und setzt dabei eine entsprechende Parametrierung (Option: "Derivationen einbeziehen") voraus. Es gibt dabei folgende Teilverfahren (die sich auch ergänzen können):

- (a) Die Derivationen werden als Alternativstämme **lexikalisch** ermittelt ("lexikalische Derivationen einbeziehen"); dies setzt eine entsprechende lexikalische Vorbereitung voraus;
- (b) Die Derivationen werden als "Endungsketten" **graphematisch** ermittelt ("graphematische Derivationen einbeziehen"); dies erlaubt eine stärkeres Eingehen auf Nutzerwünsche, ist aber semantisch "riskant" (Beispiel: Mehrdeutigkeit zwischen dem Suffix "ei" und dem Substantiv "Ei" in Zusammensetzungen im Deutschen);
- (c) Die Derivationen werden nur dann als "Endungsketten" graphematisch ermittelt, wenn das Wort selbst nicht lexikalisch identifiziert wurde.

#### (5) Trunkierung von Komposita und von Mehrwortbegriffen

Im Information Retrieval spielt die Erkennung von Mehrwortbegriffen und die Zerlegung von (Augenblicks-) Komposita (im Deutschen) eine wichtige Rolle. Einerseits steigt die **Precision** und es wird damit weniger Ballast geliefert, wenn der Abstand, in dem derartige Wörter im Text stehen dürfen, eingegrenzt werden kann. Andererseits steigt v.a. der **Recall**, wenn neben einer Wortzusammensetzung auch die relevanten Wortteile in analoger Form bereitgestellt werden (Beispiele: Kanzlerreise -> Reise / Kanzler; Datenschutzbeauftragter -> Datenschutz / Beauftragter ...). Die Trunkierung von Wortzusammensetzungen stellt jedoch ein besonderes Problem dar.

Bei der Darstellung und damit der Identifikation und Generierung muß unterschieden werden zwischen Wortbildungen, die graphematisch zu einer **kontinuierlichen Zeichenkette ohne Wortzwischenraum** führen (die "klassischen" Komposita wie Sonnenuntergang, Armbanduhr, im folgenden **Typ 1**), und solchen Mehrwortbegriffen, die entweder in kontinuierlicher Folge (wie Französische Revolution, juristische Person) oder aber auch (v.a. im Deutschen) diskontinuierlich auftreten (Beispiele: finden ... statt -> stattfinden; nehmen ... ein Bad -> ein Bad nehmen), im folgenden **Typ 2**.

Mehrwortbegriffe vom Typ 2 sind morphologisch / syntaktisch analog den Einzelwörtern in einen Satz bzw. eine Phrase eingebettet und verhalten sich (etwa in der Wortstellung) wie sonstige Satzelemente. Man kann hierbei weiter unterscheiden nach solchen festen Wendungen, die positionsbezogen und z.T. auch flexionsbezogen nicht variiert werden können, ohne die Bedeutung zu "wechseln" (Beispiel: kalter Kaffee = Spezi; kalter und abgestandener Kaffee = Kaffee, der kalt ... ist), und solchen, die in begrenztem Umfang im Kontext variiert werden können (Beispiel: natürliche und juristische Person).

Für die Trunkierungsproblematik bedeutet dies zweierlei:

- (a) Wortzusammensetzungen müssen zunächst einmal als solche identifiziert werden, auch dann, wenn sie beispielsweise flektiert sind (Beispiel: juristischer Personen -> juristische Person; hielt ... den Atem an -> den Atem anhalten).
- (b) Mehrwortbegriffe vom Typ 2 müssen zur Identifikation lexikalisiert sein, Mehrwortbegriffe vom Typ 1 müssen dann lexikalisiert sein, wenn darauf Restriktionen wirken sollen.

Auf Verfahren und Probleme der Identifikation kann hier nicht weiter eingegangen werden. Im allgemeinen wird man heute im Zweifelsfall eine Lexikalisierung vornehmen, um überhaupt eine Identifikationsmöglichkeit zu haben. Ansätze, aufgrund lexikalisch-syntaktischer Analysen generell typische Mehrwortgruppen zum Retrieval bereitzustellen, haben bisher nicht den gewünschten Erfolg gebracht (vgl. Keil / Kroupa / Zimmermann 1983 und die Evaluierung in Krause et al. 1983) und verlangen zur Kontrolle in jedem Falle auch "Lexikalisierungen" (beispielsweise, um bei einer Wortfolge wie öffentliche Stellen und Kirchen nicht fälschlich öffentliche Kirche zu erzeugen).

Die Einbeziehung von Wortzusammensetzungen erfordert eine Erweiterung der Ergebnisdarstellung. Dementsprechend werden neben dem Trunkierungszeichen (hier "\$") und der zugehörigen Angabe der Trunkierungslänge sog. Positions- und Abstandszeichen mit einer Abstandsangabe (gerechnet in

Wortabständen) eingeführt. Zudem erhalten die in einer Suchanfrage auftretenden bzw. in den Ergebnissen dargestellten Elemente (= Wörter) eine Wortnummer (LWn, wobei n eine Zahl ist). Es bedeuten weiterhin (bezogen auf die Reihenfolge im Satz, wobei "links" als "vor" und "rechts" als "nach" gewertet wird):

< = "vor; <> = "vor" oder "nach"

Das Beispiel juristische Person führt also im Falle der (lexikalischen) Identifikation zu folgenden Angaben

LW1 = juristische Person ... (+ Merkmale, Übersetzungen usf.)

LW2 = juristisch ... (+ Merkmale ...)

LW3 = Person ... (+ Merkmale ...)

und dem (zusätzlichen) Hinweis: LW2 < LW3;

das Beispiel Angst haben führt zu folgenden Angaben

LW1 = Angst haben (...)

LW2 = Angst (...)

LW3 = haben (... ggf. "gestoppt")

und dem zusätzlichen Hinweis: LW2 <> LW3;

das Beispiel Datenschutz (nicht aber Haustür oder Herrenhemd) führt (wenn über eine Synonymie-Relation "lexikalisiert") zu folgenden Angaben (Auszug):

LW1 = Datenschutz (...)

LW2 = Schutz von Daten (...)

LW3 = Schutz (...)

LW4 = Daten + Datum (...)

und dem zusätzlichen Hinweis: LW3 < LW4.

Liegt eine lexikalische Identifikation nicht vor, so kann nach entsprechender flexionsmorphologischer oder aber auch struktureller Analyse (Identifikation) eine analoge Kennung vergeben werden, die formal jedoch entsprechend (mit F bei der Positionsangabe bzw. FW bei der Wortnumerierung) markiert sein kann.

Bei Komposita, die nicht lexikalisiert sind (wobei die Überprüfung der Lexikalisierung entweder anhand einer Relationierung oder aber einer Übersetzung erfolgen muß), führt dies im Falle der flexionsmorphologischen Identifikation zu folgenden Angaben:

FW1 Kanzlerreise (...)

LW2 Kanzler (...)

LW3 Reise (...)

und dem zusätzlichen Hinweis: LW2 <>F      LW3.

Das Beispiel öffentlicher Haushalt führt im Falle der (satzanalytischen) Identifikation (d.h., wenn nicht bereits lexikalisiert) zu folgenden Angaben

LW1 = öffentlich ... (+ Merkmale, z.B. Wortklasse ...)

LW2 = Haushalt ... (+ Merkmale ...)

und dem (zusätzlichen) Hinweis: LW1 <F LW2;

### 3. Zusatzfunktionen

Die im folgenden aus Platzgründen nur andeutungsweise beschriebenen Zusatzfunktionen stellen Optionen dar, die es erlauben, die Ergebnisse einzuschränken oder auch auszuweiten.

#### 3.1 Selektion

Eine Einschränkung (Selektion) kann beispielsweise dann sinnvoll sein, wenn sicher ist, daß es genügend Treffer in der Datenbank gibt und es nur um einen ersten Einstieg geht. Sie kann ferner nützlich sein, wenn durch zu viele Alternativstämme (etwa bei unregelmäßigen Verben) die Anfragestruktur sehr komplex wird, ohne daß sich die Anzahl der Treffer signifikant erhöht, demgegenüber aber die Möglichkeit der Überidentifikation (d.h. beispielsweise der Treffer, die über die Trunkierung erzielt werden, ohne daß die intendierten realen Endungen vorliegen) wächst.

### 3.2 Wortrelationierung

Unter Wortrelationierung wird - wie bereits eingangs angesprochen - eine Beziehung zwischen zwei Wörtern verstanden, die über die paradigmatische flexionsmorphologische Beziehung von Wortformen (Wortbeugung, -flexion) hinausreicht. Beziehungen zwischen Haus, Hauses, Häuser fallen beispielsweise unter den Bereich der paradigmatischen Beziehung, "Haus" - "hausen" nicht mehr.

Die derivative Trunkierung (vgl. oben) kann bereits als ein erster Anwendungsfall der Wortrelationierung gesehen werden. Sie ist jedoch nur eine Möglichkeit der Relationierung und wurde im Grunde nur in die Standardfunktionen einbezogen, weil sie inhaltlich-morphologisch in den meisten Sprachen den Flexionsformen am nächsten steht.

Bei der folgenden Beschreibung bleibt ein wichtiges Problem weitgehend außer Betracht: die semantische (und z.T. auch die funktionale) Mehrdeutigkeit. Zwar lassen sich in der Phase der Relationierung selbst noch Differenzierungen vornehmen, v.a. auch in der Interaktion Nutzer / System (durch Rückfragen), doch muß man - wie generell bei der Suche in Textdatenbanken) damit rechnen, daß ein formaler Treffer aufgrund der semantischen Mehrdeutigkeit der Zeichenkette inhaltlich - aus Benutzersicht - nicht relevant ist. Beispiel (Deutsch): Bank - Banken / Bänke. Als Beispiele für funktionale Mehrdeutigkeit kann im Deutschen "sondern" (Verb / Konjunktion) und "Ehe" (Konjunktion / Substantiv) genannt werden.

Der Nutzer ist hier in einem Dilemma: Das System bietet ihm ggf. eine allerdings nur partiell wirksame Hilfe (etwa den Ausschluß von Bänke), doch bleibt beim Retrieval ein Risiko. Zusammenhänge mit dem Retrieval (etwa in Hinsicht auf Einschränkungen auf Deskriptorfelder) können hier nur angedeutet werden.

Im folgenden werden nur einige wichtige Relationierungen aufgeführt:

**(1) Synonym- und Quasisynonymrelationierung, Schreibvarianten**

Hierunter fallen die strengen Synonyme wie Sonnabend / Samstag, aber auch Varianten wie Brötchen / Semmel oder Schreibvarianten wie Cuba / Kuba. Eine Erläuterung dieser Differenzierung erübrigt sich, die Nützlichkeit ist auch offenbar.

**(2) Dekomposition**

Die Problematik liegt hierbei weniger im formalen Bereich als in der praktischen Umsetzung im Hinblick auf ein Retrieval. Dies hängt damit zusammen, daß einerseits die Wortzusammensetzungen im oben definierten erweiterten Sinn häufig einen Bedeutungswandel mitmachen. Bei vielen gängigen Wörtern ist dies bekannt: Haustür ist eben nicht das gleiche wie Tür des Hauses usf. Aus Systemsicht ist jedoch nicht immer zu entscheiden, inwieweit ein identifiziertes, aber (noch nicht) lexikalisiertes "Kompositum" vom Nutzer auch in seinen Bestandteilen berücksichtigt werden soll.

Komplex ist die Situation bei Komposita (Deutsch), die nicht bereits lexikalisiert sind, aber durch eine Wortidentifikation in möglichen Bestandteilen identifiziert werden. Einerseits gibt es Probleme der Wortklassenzuordnung (Beispiel: Suchfunktion -> suchen / Suche), andererseits treten - wenn auch selten - Zerlegungsprobleme auf (klassisches Beispiel: Druckerzeugnis: Druck-Erzeugnis / Drucker-Zeugnis)

**(3) Übersetzungsrelation**

Es gibt v.a. bibliographische Datenbanken, in denen Titel verschiedener (natürlicher) Sprachen unübersetzt auftreten (man denke an die OPACs in Bibliotheken). Generell besteht daneben (bei der Suchanfrage) das Problem der Übersetzung eines Wortes in die Textsprache der Datenbank (etwa bei der Suche in Abstracts).



Je nach Situationszusammenhang und technischer Möglichkeit kann entweder die Konsultation eines elektronischen Übersetzungswörterbuchs vorangeschaltet oder aber in den Suchprozeß integriert sein. In jedem Falle muß diese Funktion vor der eigentlichen Trunkierungsphase ablaufen.

#### **(4) Weitere Relationen**

Soweit ein lexikalischer Zusammenhang zwischen einem **Fachthesaurus** einer Datenbank technisch hergestellt ist, kann beispielsweise auch der Verweis auf eine **Vorzugsbenennung**, einen Ober- und Unterbegriff, eine Zuordnung Wort - Klassifikation usw. genutzt werden. Derartige Verknüpfungen setzen jedoch in der Regel eine vorherige (und kontinuierliche) Zusammenarbeit zwischen einem Datenbankhersteller und dem Nutzer bzw. Anbieter des Trunkierungsverfahrens voraus.

### **3.3 Weitere Sonderfunktionen**

Die folgenden Funktionen gehen kurz auf eher **technische Sonderfälle** ein. Damit sollen aus Benutzersicht "lästige", aber letztlich triviale Manipulationen am Zeicheninventar systemseitig unterstützt werden.

#### **(1) Groß- / Kleinschreibung**

Bei der Differenzierung sollte (frühzeitig) die Frage der Berücksichtigung der Unterscheidung von Groß- und Kleinbuchstaben entschieden sein. Dies ist (wiederum) besonders beim Deutschen relevant, da hierdurch gelegentlich Homographen disambiguiert werden oder aber zusammenfallen. Beispiele: **Laut / Laute / laut - laute** ...; **Ehe / ehe**.

Wenn das Retrievalsystem (und die entsprechende Datenbank) selbst nicht weiter unterscheidet, muß diese Problembehandlung in den Trunkierungsmechanismus eingehen.

## (2) Umlaut- und Akzentbehandlung (Zeichensatz)

In vielen IR-Systemen sind Umlaute und "ß" aufgelöst bzw. bleiben Akzente unberücksichtigt. Dies hat einerseits Einfluß auf die Ergebnisdarstellung (äußere Form), kann sich aber auch auf die Trunkierung auswirken (andere Suchstrings). Man kann auch daran denken, Jokerzeichen einzuführen, um solche Positionen (im Wortinnern) zu markieren und ggf. auch zu differenzieren.

Beispiele (Jokerzeichen = # = 1 oder 2 Zeichen; % = 1 Zeichen):

Deutsch M#tter \$1; Stra#e \$1 (für Mutter / Mütter - Muetter; Straße / Strasse)

Franz.: gar%on \$1; ma%tre \$1 (für garçon; maître)

Im Regelfall wird man jedoch auf einen "aufgelösten" Alternativstring abbilden, da die Datenbank selbst nur die aufgelöste Form kennt, also:

Deutsch: Mutter \$1 + Muetter \$1; Strasse \$1

Franz.: garçon \$1; maitre \$1

Berücksichtigt werden muß ggf. die Frage der Versalienschreibung, so daß aus den o.a. Beispielen schließlich MUTTER \$1 + MUETTER \$1 wird.

## (3) Maximaltrunkierung (Maximaltrunkierungslänge)

Hierbei handelt es sich um eine formale Beschränkung, die der Nutzer vorgeben kann. In diesem Falle wird eine Längenangabe für den Fall, daß ihr Wert größer als der Wert der Maximaltrunkierung ist, auf den Wert der Maximaltrunkierung gesetzt. Dies ist zwar eher ein formaler Trick, vorwiegend um Ballast (Überindexierung) zu vermeiden, doch gibt es gute Argumente dafür, solch eine technische Funktionalität zuzulassen. In jedem Falle wird nämlich dann ein "realer" (korrekter) Längenwert zurückgegeben, wenn die natürlichsprachig bedingte Längenangabe kleiner oder gleich der Maxi-

mallänge ist.

### 3.4 Zusatzfunktionen

Als Zusatzfunktionen werden solche Funktionen gekennzeichnet, bei denen nicht eine Trunkierung angestrebt wird, sondern Ergebnisse und Teilergebnisse zu anderen Zwecken genutzt werden sollen. Hierzu rechnen - ohne daß dies hier näher ausgeführt werden kann - folgende Funktionen:

- **Grundformenermittlung**
- **Flexionsformenerzeugung**
- **Homographentypisierung**
- **Pluralstammerzeugung**

Die Funktion der Pluralstammerzeugung sollte bei der Trunkierung alternativ angeboten werden, da beispielsweise viele Datenbanken (bzw. Thesauri und Register) nicht die Singularform, sondern die Pluralform als Deskriptor angeben.

- **Übersetzungsfunktion**
- **Relationsfunktion**

### 4. Beispiel einer technischen Implementierung

Im folgenden wird der Rahmen für die technische Implementierung beschrieben. Diese Funktionen werden z.Z. im Rahmen zweier Forschungsprojekte entwickelt und getestet. Alle Funktionen sind bzw. werden in "C" programmiert:

- Im Forschungsprojekt "Intelligente Interfaces zu Textdatenbanken" (MITI/IMIS), das im Rahmen des Förderprogramms IMPACT der EG im Aufbau ist, wird eine UNIX-Applikation der Trunkierung realisiert. Das System selbst (und auch die Funktionen) wird dabei für SCO-UNIX unter Open DeskTop mit X-Windows-Basis implementiert.
- Im EUREKA-Projekt EURO-TRIANGLE (Nr. 257) sind entsprechende Funktionen unter MS-DOS in Entwicklung. Vorgesehen ist CommonView und MS-Windows als Rahmen.

Die Basisfunktionen werden dabei möglichst umgebungsunabhängig realisiert. Dieses Prinzip hat sich inzwischen schon bei Verfahren zur automatischen Silbentrennung, Rechtschreibkontrolle und Wortübersetzungshilfe bewährt.

Die gesamten beschriebenen (Trunkierungs-)Funktionen stellen eine Bündelung von Teilfunktionen dar. Das Trunkierungsbündel selbst wird über Parameter gesteuert, die letztendlich die Strategie (Vorgehensweise) bestimmen. Auf eine ausführliche Beschreibung muß hier verzichtet werden (vgl. ZIMMERMANN 1991).

## 5. Ausblick

Es hat sich in den Jahren der systematischen Entwicklung dieser Funktionalität immer wieder gezeigt (und dies ist sicherlich jedermann von vorneherein klar), daß eine allgemeine systematische Entwicklung solcher Funktionen nie völlig ausreicht, sondern erst mit der Anwendung und einer entsprechenden Anwendungsorientierung marktfähig wird. Es geht in Zukunft nicht (oder weniger) um die Frage der technischen Funktionalität (und Bereitstellung), sondern um die Adaption an konkrete Anwendungsfälle. Derartige Anwendungen orientieren sich in erster Linie auf solche Bereiche, wo UNIX- und MS-DOS-Umgebungen vorliegen (also weniger auf den Mainframe-Bereich).

Inzwischen bieten sich genügend Möglichkeiten an, derartige "höherwertige" Verfahren in der Praxis einzusetzen. Die Rahmenbedingungen sind inzwischen ausreichend (Speicherplatz, Rechengeschwindigkeit), die Schwierigkeiten liegen inzwischen eher in Details (Wörterbuchanpassungen, v.a. für Übersetzungen) als im Grundsätzlichen, auch wenn die Beispiele, genauer gesagt: die "Gegenbeispiele" auch in Zukunft noch zahlreich sein werden. Dabei muß explizit gesagt werden, daß mit den vorgestellten Verfahren keine perfekten Lösungen möglich sind (auch wenn man eine möglichst große Perfektion anstrebt), sondern im wesentlichen (nur) eine Entlastung des Anwenders von "primitiven" Handlungen, wie sie die Trunkierungsvarianten insgesamt darstellen, angestrebt wird.

Wer sich die Mühe gemacht hat, die gesamte Studie im Detail zu verfolgen, der wird vielleicht über die Komplexität erstaunt sein, die hinter einer brauchbaren Lösung der Trunkierungsproblematik steckt. Dies ist offenbar auch der eigentliche Grund dafür, daß es bisher v.a. für stark komponierende Sprache offenbar noch keine Verfahren gibt, die derartige Hilfen anbieten. Mit den Funktionen und auf der Basis großer elektronisch verfügbarer lexikalischer Datenmengen, wie sie hier vorgestellt wurden, sollte diese Funktionalität aber bald zum "Standard" aller Information Retrievalsysteme werden können.

## 6. Literatur

**Dietrich**, Rainer (1973): Automatische Textwörterbücher. Studien zur maschinellen Lemmatisierung verbaler Wortformen des Deutschen. Tübingen.

**Keil**, Gerald; **Kroupa**, Edith; **Zimmermann**, Harald H. (1983): CTX - Ein Verfahren zur computergestützten Texterschließung. Universität Saarbrücken.

**Krause**, J. et al. (1983): EVAL. Zur Evaluierung informationslinguistischer Komponenten von Informationssystemen. Universität Regensburg.

**Kuhlen**, Rainer (1977): Experimentelle Morphologie in der Informationswissenschaft. München.

**Panyr**, Jiri, **Zimmermann**, Harald H. (1989): Information Retrieval: Überblick über aktive Systeme und Entwicklungstendenzen. In: Computational Linguistics / Computerlinguistik (ed. I. Batori, W. Lenders, W. Putschke), S. 696 - 708.

**Zimmermann**, Harald H. (1979): Ansätze einer realistischen automatischen Indexierung unter Verwendung linguistischer Verfahren. In: Kuhlen: Datenbanken, Datenbasen, Netzwerke. Praxis des Information Retrieval Bd. 2, Frankfurt, S. 311 - 338.

**Zimmermann**, Harald H. (1982): Maschinelle Verfahren zur Erschließung von Textkorpora. In: Sprachen und Computer (ed. H. Fix, A. Rothkegel, E. Stegentritt), Saarbrücken.

**Zimmermann**, Harald H. (1983): Multifunctional Dictionaries. In: The possibilities and limits of the computer in producing and publishing dictionaries (ed. A. Zampolli, A. Capelli) Pisa (= Linguistica Computazionale III)., S. 279 - 288.

**Zimmermann**, Harald H. (1991): Ein Verfahren zur automatischen Trunkierung beim Zugang zu textbezogenen Informationsbanken (Langfassung, mit Applikationsbeschreibung und technischem Anhang; dt. und englisch; Manuskript, Saarbrücken).

# **Eine objekt-orientierte Sicht auf Expertensystem-Wissensbasen: auf dem Weg zu einer hypermedia-gestützten Dialogkomponente**

**Fahri Yetim**

Universität Konstanz

FG Informationswissenschaft

Postfach 5560, D-7750 Konstanz

## **Zusammenfassung**

In dieser Arbeit werden einige Aspekte der Entwicklung einer hypermedia-gestützten Dialogkomponente zu einem Expertensystem vorgestellt. Die objektorientierte Betrachtung der Wissensbasis-Elemente aus der Sicht der Dialogkomponente und deren Behandlung durch den Benutzer bei der Durchführung von entscheidungsrelevanten Aktivitäten, wie die Berechnungen und Präsentationen von Ergebnissen, bilden den Schwerpunkt der Darstellung.

## **Abstract**

In this paper some aspects of the development of a hypermedia-supported dialog component to an expert system are presented. The main emphasis is on the object-oriented view into the elements of the expert system from the dialog component point of view, as well as on the treatment of these elements by the user during the activities relevant to the decision, such as computation and presentation of results.

## **1. Einleitung**

Leistungssteigerungen bei der Hardware und ausgereifere Werkzeuge in der Systemsoftware sind die Ursachen dafür, daß sich seit einigen Jahren neue Möglichkeiten insbesondere für eine benutzerorientierte Gestaltung der Mensch-Computer-Schnittstelle anbieten. Hypermedia sind solche neuen Medien, die sich sehr rasch entwickeln und zu den Hoffnungsträgern der Informationstechnologie gehören. Die Vorteile dieser Medien liegen vor allem in ihren sehr komfortablen Benutzeroberflächen und in der Flexibilität ihrer Nutzung. Daher werden sie vom Prinzip her als humane Technik angesehen (Kuhlen 1991a,b). Eine Verbindung der Hypermedia-Konzepte mit den existierenden wissensbasierten Systemen, wie etwa Expertensystemen, verspricht demnach eine Erhöhung der Akzeptanz solcher Systeme (vgl. Yetim/Dambon 1991).

Einige Arbeiten zur Integration von Hypermedia-Methodologien<sup>1</sup> und Expertensystemansätzen wurden im Umfeld des wissensbasierten Kreditabsicherungssystems

---

<sup>1</sup> Die Bezeichnungen 'Hypermedia' und 'Hypertext' werden in dieser Arbeit synonym verwendet.

WISKREDAS bereits durchgeführt (Dambon/Yetim 1990, Yetim 1991a,b). Die Notwendigkeit dieser Integration leitet sich unmittelbar aus der Notwendigkeit der systematischen Einbeziehung des Nutzers bei einer kooperativen Problembearbeitung ab, worauf die Konzeption des Systems WISKREDAS basiert. Das entspricht der Idee, die Bearbeitung gutstrukturierter, komplexer, formalisierbarer Problemaspekte dem Rechner zu übertragen und die schwachstrukturierten Anteile, die globale und intuitive Prozesse erfordern, durch den Menschen bearbeiten zu lassen und gezielt zu unterstützen.

Die Leistung der hierfür zu konzipierenden hypermedia-gestützten Dialog-Komponente, die auf einem existierenden Expertensystem aufbaut und eine solche kooperative Problembearbeitung unterstützen soll, besteht somit sowohl in der geeigneten Aufbereitung und Präsentation der systemseitig durchgeführten Aktionen (Berechnungen) und ihrer Ergebnisse als auch darin, den Benutzer dabei zu unterstützen, eigene Entscheidungsvorschläge zu machen und deren Konsequenzen flexibler zu prüfen. Um dieser Anforderung an die Benutzer- und Aufgabenorientiertheit gerecht zu werden und eine flexible Nutzung des Systems zu ermöglichen, lassen sich die dem Benutzer in unterschiedlichen Tätigkeitsräumen (Fenstern) visualisierten Basis-Elemente des Systems (Konzepte, Daten etc.) als Objekte behandeln. Diese objekt-orientierte Sicht auf die Elemente des Expertensystems, die als eine Alternative oder Ergänzung zu der Inferenzsicht flexible Dialoge erlauben soll, kommt zugleich dem Hypermedia-Konzept entgegen, da ihm die objekt-orientierte Philosophie zugrundeliegt.

Um diese Aspekte hier ausführlicher zu behandeln, wird zunächst die Systemumgebung hinsichtlich der Funktion als kooperative Entscheidungsunterstützung vorgestellt und daraus die Anforderungen an die Dialogkomponente abgeleitet. Die objektorientierte Betrachtung der Wissensbasis-Elemente aus der Sicht der Dialogkomponente und deren Behandlung durch den Benutzer bei der Durchführung von entscheidungsrelevanten Aktivitäten, wie die Berechnungen und Präsentationen von Ergebnissen, bilden den Schwerpunkt des weiteren Inhaltes dieses Beitrages.

## **2. Das System WISKREDAS zur kooperativen Entscheidungsunterstützung**

Das System WISKREDAS<sup>2</sup> (vgl. Kuhlen et al. 1990) dient der Unterstützung eines Kreditsachbearbeiters bei der Beurteilung von Bürgschaftsanträgen und entstand nach eingehenden empirischen Untersuchungen in einer halbstaatlichen deutschen Bank (Wolf 1988a). Die Unterstützung besteht im wesentlichen darin, daß das System das Wissen, das zur Bearbeitung eines Kreditantrags eines Existenzgründers benötigt wird,

---

<sup>2</sup> Dieses wissensbasierte Kreditabsicherungssystem wurde im Rahmen des an der Universität Konstanz angesiedelten Sonderforschungsbereichs 221, "Verwaltung im Wandel", entwickelt.



verwaltet, zum Teil automatisch einholt, dieses nach Regeln, die auf dem Fachwissen des Sachbearbeiters beruhen, verarbeitet und schließlich, wenn die informationelle Basis als ausreichend eingeschätzt wird, einen Entscheidungsvorschlag macht, der im Ergebnis eindeutig sein muß: Empfehlung auf Ablehnung oder Übernahme der beantragten Bürgschaft.

Die Entscheidung zur Ablehnung bzw. Bewilligung eines Antrags basiert auf der Einschätzung der Erfolgsaussichten der zu gründenden Firma, konkret auf der Prognose, ob dem Antragsteller nach Abzug aller Ausgaben und Verpflichtungen genügend Mittel für den Lebensunterhalt verbleiben. Die Entscheidung stützt sich nicht nur auf wirtschaftliche Berechnungen, sondern auch auf Erfahrungen des Sachbearbeiters und auf die Gutachten von verschiedenen externen Experten (z.B. von Kammern, Verbänden usw.), die zu den wichtigsten Projektparametern Stellung nehmen.

### **2.1. Zur Ablaufstruktur des systemunterstützten Entscheidungsprozesses<sup>3</sup>**

Bei der ersten Bearbeitung eines Falls werden die Entscheidungsparameter vom Sachbearbeiter gesetzt, indem festgelegt wird, welchen Reingewinn der Antragsteller erzielen muß, um seinen Verpflichtungen nachzukommen, und mit welcher Sicherheit diese Vorhersage mindestens gemacht werden muß, um eine Entscheidung zuzulassen. Durch verschiedene Maßnahmen wird versucht, die Unsicherheit, die den Berechnungen des zu erwartenden Reingewinns anhaftet, zu reduzieren: (a) durch die Aufspaltung der Variablen in Teilvariablen und durch Zuführen weiterer Informationen (z.B. die Meinungen von Experten). Darüberhinaus werden widersprüchliche Informationen oder solche, bei denen man von einer gewissen Streuung der Werte ausgehen muß, durch die Berechnung verschiedener Varianten (z.B. optimistisch, pessimistisch und realistisch) und durch die Berücksichtigung von Sicherheitsfaktoren, die angeben, mit welcher Wahrscheinlichkeit jeder dieser Werte für den vorliegenden Fall voraussichtlich zutreffend sein wird, zu einem Gesamtbild aggregiert. Über die alternative Berücksichtigung einzelner oder mehrerer widersprüchlicher Angaben bei den Berechnungen hinaus ist vielmehr gerade die gleichzeitige, simultane Analyse aller für einen Fall vorhandenen Informationen - insbesondere der widersprüchlichen - interessant.

Da die Lösungsalternativen schon vorgegeben sind, kann hier keine Unterstützung zur Generierung neuer Möglichkeiten geboten werden, jedoch können alternative Wege der Informationsbeschaffung und -verarbeitung aufgezeigt werden. Die Entscheidungsfindung selbst wird zum einen durch einen vom System errechneten eigenen Vorschlag unterstützt; zum anderen dadurch, daß die für einen Fall relevanten Informationen so aufbereitet werden, daß der Sachbearbeiter durch einen vertieften Einblick und umfassende Variationsmöglichkeiten bei seiner eigenen Entscheidungsfindung weitge-

---

<sup>3</sup> Eine ausführlichere Beschreibung des Entscheidungsablaufs ist Wolf 1988b zu entnehmen.

hend unterstützt wird. Unterschiedliche Informationstypen müssen dabei berücksichtigt werden und in einer Entscheidungssituation in den Entscheidungsprozeß einfließen. Hier kommt daher sowohl der Wissensrepräsentation als auch der Präsentation eine große Bedeutung zu, deren Realisierung durch die dem Hypermedia-Konzept inhärente Flexibilität unterstützt werden kann.

## **2.2. Zur Wissensrepräsentation**

Zur flexiblen und dynamischen Verwaltung des Wissens, das u.a. als Grundlage für die Berechnungen dient, wurde die Frame-Repräsentation<sup>4</sup> gewählt (vgl. Dambon 1988). Die gesamte Framestruktur, der sog. Makroframe, wurde zur Integration von Hypertext-Eigenschaften um die entsprechenden Slots und Relationen erweitert (vgl. Dambon/Yetim 1990, Yetim 1991b). Den Makroframe kann man sich als Netzwerk vorstellen, dessen Knoten Konzepte (wiederum Frames) und dessen Verknüpfungen Beziehungen zwischen den Konzepten repräsentieren. Jeder Frame besteht wieder aus einer Menge von Eigenschaften (Slots) und zu jeder dieser Eigenschaften existieren einerseits Vorschriften, welche zulässige Einträge (Constraints, Defaults, Ranges, bedingte Prozeduraufrufe, Berechnungsformeln etc.) spezifizieren, andererseits existieren statische Texte (z.B. Definitionen, Fehlermeldungen etc.). Im Makroframe sind also sowohl die hierarchischen, strukturellen Beziehungen zwischen den einzelnen Berechnungsvariablen, als auch die Algorithmen (Regeln), nach denen sich sehr spezifische, differenzierte Informationen zu globaleren Variablen aggregieren lassen (prozedurale Aspekte), enthalten. Bei der Bearbeitung eines konkreten Falls werden aus diesem Makroframe Einzelfälle generiert. Diese stellen konkrete Ausprägungen (Instanzen) des Makroframes, d.h. sie enthalten nur konkrete, fallspezifische Slot-Einträge (Daten).

## **2.3. Anforderungen an die Dialog-Komponente**

Um die angestrebten wesentlichen Systemleistung zu erreichen, nämlich die Kooperation und Unterstützung intuitiver Prozesse, kommt der Dialogkomponente eine besondere Bedeutung zu. Bei der Entwicklung einer solchen Komponente, die ein hohes Maß an Flexibilität erfordert, reicht die Benutzerorientiertheit als Leitlinie allein nicht aus, denn nicht nur die Art der Gestaltung der Benutzungsoberfläche, die sich an den Grenzen und Fähigkeiten menschlicher Informationsverarbeitung orientiert, ist von großem Nutzen, vielmehr spielt die Funktionalität, die der Benutzer bei der Erledigung seiner Aufgaben bzw. der Lösung seiner Probleme benötigt, eine wesentliche Rolle (vgl. Streiz 1990). Damit ist die Aufgabenzentriertheit als weitere wichtige Entwicklungsleitlinie angesprochen, was wiederum bedeutet, daß die hier zu unterstützende Tätigkeit, nämlich die Antragsbearbeitung, im Lichte der durchgeführten empirischen

---

<sup>4</sup> Es besteht eine Ähnlichkeit zwischen dem 'Objekt'- und dem 'Frame'-Begriff (vgl. Stefik/Bobrow 1986).

Untersuchungen (vgl. Wolf 1988a), d.h. der arbeitsplatzbezogenen Analyse der Entscheidungsprozesse eines Sachbearbeiters bei der Beurteilung von Bürgerschaftsanträgen, zu berücksichtigen ist. Dies entspricht auch der Forderung von Don Norman: "interact with the task, not with the computer"<sup>5</sup>. Als ein weiterer entscheidender Faktor bei der Entwicklung dieser Dialogkomponente wird das theoretische Entscheidungsmodell von Wolf 1988b, welches auf der Basis der o.g. empirischen Untersuchungen entstanden ist, in Betracht gezogen. Schließlich kommt bei der Entwicklung die Technologiebewußtheit durch den Einsatz der Hypermedia-Technologie zum Tragen.

Auf der Grundlage der gleichzeitigen Berücksichtigung o.g. Leitgedanken wird die Entwicklung einer Dialogkomponente angestrebt, die dem Sachbearbeiter die Möglichkeit einräumt, in allen Stufen einer Entscheidung einzugreifen und sowohl vom System konsultiert zu werden als auch das System zu konsultieren. Von hoher Bedeutung ist daher die Flexibilität, die es dem Sachbearbeiter erlaubt, individuelles Vorgehen bei der Sachbearbeitung zu entwickeln. Damit sind die wohl allgemein akzeptierten Unterstützungsleistungen von Hypermedia durch ihre Navigationsmöglichkeiten angesprochen, die sich aber nicht nur auf statische Wissensbasis (statische Programmabläufe, Standard-erklärungen) beschränken, sondern sich auch auf dynamische Abläufe, wie die situationsspezifischen Eingaben vom Sachbearbeiter und die darauf erfolgten dynamischen Reaktionen vom System, beziehen. Dadurch bietet sich die Möglichkeit, alle Zwischenzustände, die während der Berechnung erreicht werden, zu überprüfen, wie auch die Angemessenheit der angewandten Regeln sowie die der Entscheidung zugrundeliegenden Parameter nachzuprüfen. Diese setzen wiederum die Möglichkeit der Einfügung der Zwischenschritte im Arbeitslauf voraus, ohne dabei den Verlust von Daten zu riskieren.

Um zu gewährleisten, daß der Sachbearbeiter die Systemleistungen in seine eigenen Entscheidungsprozesse integrieren kann, kommt der Präsentation der Ergebnisse große Bedeutung zu, die durch die graphische Aufbereitung der Informationen so unterstützt werden kann, daß sie vom Nutzer problemlos weiterverarbeitet werden können. Auch nach der Entscheidung eines Antrags auf Bürgerschaftsübernahme besteht ein erheblicher Erklärungsbedarf. Der positive oder negative Entscheid soll begründet werden. Retrospektiv könnten verschiedene Varianten oder ähnliche Fälle miteinander verglichen werden. In diesem Fall kann ein differenziertes Erklärungsangebot den Entscheidungsprozeß von der Eingabe bis zu einer retrospektiven Analyse begleiten.

Wie eingangs angedeutet, sind in dieser Arbeit nicht die Aspekte, wie das System selbständig seine Arbeiten erledigt, von vorrangigem Interesse, vielmehr die (teilweise) Beantwortung der Frage, welchen Einfluß der Benutzer auf die Systementscheidungen

---

<sup>5</sup> vgl. hierzu das von Rheingold 1990 durchgeführte Interview mit Don Norman.

hat und wie groß dieser ist, ferner welche Funktionalität die entsprechende Benutzungsoberfläche anbietet, um dem Sachbearbeiter bei der Antragsbearbeitung flexible Unterstützung zu leisten.

### **3. Objekte der Benutzungsoberfläche**

Die Dialogkomponente stellt dem Benutzer unterschiedliche Räume (spaces) zur Verfügung, die verschiedene Aktivitäten des Benutzers bei der Entscheidungsfindung unterstützen<sup>6</sup>. Diese Räume dienen zur Eingabe von Falldaten (Falleingabe), zu Analysen und Vergleichen von Fällen (Fallanalyse), zur Visualisierung der statischen und dynamischen Daten bzw. Ergebniszustände (Präsentationsraum), zur Durchführung von Berechnungen (Berechnungsraum), zur Erstellung des Sachberichts mit der Begründung der Ablehnung bzw. Bewilligung eines Antrags (Begründungsraum), zur Darstellung von Wissensbasisobjekten des Systems (Objektraum) und zur Beschaffung von Informationen aus internen und externen Datenbanken (Ressourcenraum). Darüberhinaus können verschiedene Browser die Arbeitstätigkeit des Benutzers unterstützen: so kann z.B. der retrospektive Browser die Visualisierung der Informationen, die sich auf die Aktivitäten in einer kompletten Arbeitssitzung beziehen, einen direkten Zugriff auf die gesuchte Information ermöglichen, die der Benutzer eventuell bei weiteren Berechnungen berücksichtigen möchte. Der Frage-Browser visualisiert Fragen, die der Benutzer im Dialog beispielsweise zur Erklärungszwecken stellen kann. Die Trennung dieser Aktivitäten bietet nicht nur eine verbesserte Unterstützung der einzelnen Tätigkeiten des Benutzers bei der Problemlösung, sondern auch die Möglichkeit der gleichzeitigen Betrachtung von dynamischen Daten. So können z.B. in einem Raum Berechnungen bzw. Aktionen ausgeführt werden, während in einem anderen ebenso aktiven Raum, wie dem Präsentationsraum, die Ergebnisse oder Veränderungen in Entscheidungszonen simultan beobachtet werden (vgl. Kapitel 3.3).

#### **3.1. Der Objektraum**

Wie in Abb. 1. dargestellt, werden im Objektraum die Objekte der Wissensbasis dem Benutzer zugänglich gemacht, und es wird ihm ermöglicht, mit diesen Objekten bestimmte Aktionen in der Wissensbasis und/oder auf dem Bildschirm hervorzurufen. Obwohl das gesamte statische Anwendungswissen in WISKREDAS in Frame-Strukturen mit prozeduralen Aspekten modelliert ist, wollen wir unter einem

---

<sup>6</sup> Die Verwendung der 'Räume' entspricht in Teilen der Idee, die sich im Konzept des Autorensystems SEPIA (Streitz 1990) und bei der Verwendung der 'Rooms'-Metapher von Card/Henderson 1987 findet. In WISKREDAS entsprechen diese Räume zum Teil den Komponenten des Systems, die von den einzelnen Autoren auf einem UNIX-Rechner unter Verwendung von PROLOG realisiert sind (vgl. Dambon 1988, Glasen 1990, Thost 1990, Yetim 1990). Weitere Realisierungsaspekte sind in Blumer 1990 und Banoss 1990 sowie Heizmann 1990 dokumentiert.

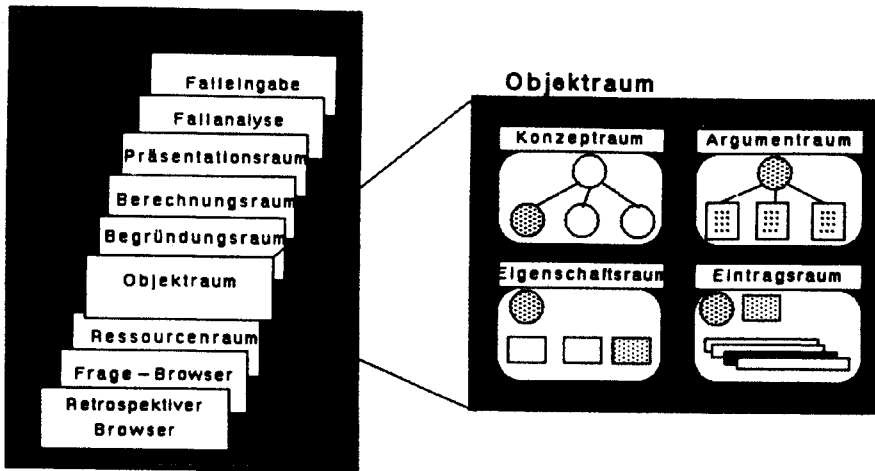


Abb. 1: Die Architektur der Benutzeroberfläche

Wissensbasisobjekt nicht nur Frames, wie es überwiegend in der Literatur üblich ist, als Ganzes verstehen, sondern auch Eigenschaften (Slots) und Slotspezifikationen sowie Slot-Einträge, wie sie alle in der PROLOG-Notation als Prädikate (PROLOG-Fakten und -Regeln) realisiert sind und direkt angesprochen werden können (vgl. hierzu Yetim 1991b)<sup>7</sup>. Dazu gehören auch nicht-formalisierte Texte (die Gutachten der Experten). Entsprechend werden folgende Unterräume des Objektraums unterschieden, wobei jedem Raum Funktionen zugeordnet sind, die der Wiedergabe und Veranschaulichung der Informationen dienen, jedoch in Abb. 1 nicht dargestellt sind:

- **Der Konzeptraum** enthält die Konzepte in der Wissensbasis. Die Darstellung der Konzepte können so erfolgen, daß die hierarchischen oder strukturellen Beziehungen zwischen den einzelnen Konzepten erkennbar werden. Wenn ein Konzept ausgewählt wird, so erscheinen die Slots des gewählten Konzepts im Eigenschaftsraum (vorausgesetzt, daß dieser Raum aktiv ist). Da es sich hier um einen aufgabenorientierten Dialog handelt, bietet dieser Raum die Möglichkeit, ebenso auf Konzepte Bezug zu nehmen bzw. sie direkt anzusprechen, die für den Benutzer bei einer Aufgabenbewältigung nicht sichtbar sind.

- **Der Eigenschaftsraum** enthält die Eigenschaften (Slots) eines selektierten Konzepts. Die Aktivierung eines bestimmten Slot kann entsprechend den Slot-Spezifikationen unterschiedliche Aktionen ausführen: von der Darstellung von Einträgen im Eintragsraum (s.u.) bis zum Ausführen von Berechnungen sowie zur Ausgabe von Erklärungstexten. Bei der Eingabe können beispielsweise die Werte eines Slot auf Korrektheit bzw. Zulässigkeit hin geprüft und vom System kritisiert werden. Existiert ein Slot mit der Teilvariablenliste, so hat dieser eine weitere wichtige Funktion: wenn im Laufe der Antragsbearbeitung zu einem Konzept (einer Variable) weitere

<sup>7</sup> In der Arbeit von Yetim 1991b wurden dieser Objekte als Hypertext-Knoten und die unterschiedlichen Beziehungen zwischen ihnen als Hypertext-Verknüpfungen betrachtet.

Informationen gesucht werden, werden Konzepte, die in der Teilvariablenliste enthalten sind, aktiviert und zeigen somit auf, wo weitere Informationen beschafft werden sollen.

- **Der Eintragsraum** enthält die Einträge (konkrete Daten) zum selektierten Slot eines Konzeptes. Jeder einzelne als Tupel repräsentierte Slot eintrag besteht außer dem Wert selbst aus weiteren Elementen, wie dem Sicherheitsfaktor, der Ressourcenbezeichnung, dem Variantennamen etc. Die dem Eintragsraum zugeordneten Funktionen erlauben die sortierte Darstellung der Einträge nach verschiedenen Gesichtspunkten, wie nach chronologischen oder Variantennamen etc., und ermöglichen auf dieser Weise einen vereinfachten Zugriff zu den gewünschten Einträgen: das Sortieren nach Variantennamen würde beispielsweise den Sachbearbeiter bei der Berechnung mit verschiedenen Varianten (s. Kapitel 3.2.) gezielt unterstützen.

- **Der Argumentraum** dient zur Darstellung von Expertenmeinungen zu bestimmten Konzepten. Diese in nicht-formalisierter Form vorliegenden Texte sind als Hypertext-Knoten organisiert. Zwischen diesen Knoten und dem entsprechenden Konzept bestehen Argumentationsverknüpfungen, die den Inhalt dieser Knoten charakterisieren. Dementsprechend kann in diesem Raum die Darstellung der Expertenaussagen über Konzepte nach verschiedenen Gesichtspunkten, wie nach unterstützenden und/oder nicht-unterstützenden sowie neutralen Argumenten erfolgen, um dadurch die intuitive Einschätzung des Benutzers bei der Entscheidungsfindung zu unterstützen.

Diese Form der Darstellung der Wissensbasisobjekte des Systems erlaubt es, sowohl dem Systementwickler das Verhalten des Systems und somit die Wissensbasis flexibel zu prüfen, als auch dem Sachbearbeiter bei der Problemlösung durch das Anwählen der gewünschten Objekte das Verhalten derselben in einem anderen Aktivitätsraum, wie z.B. im Berechnungsraum, zu beeinflussen.

### **3.2. Der Berechnungsraum**

Zur Durchführung der entscheidungsrelevanten Berechnungsaktivitäten dient der Berechnungsraum, in dem das Berechnungsmodell visualisiert wird, so daß dadurch einerseits die systemseitig durchgeführten Berechnungen transparent gemacht werden, andererseits dem Benutzer die Möglichkeit eingeräumt wird, eigene Berechnungen durchzuführen. Die Visualisierung des Berechnungsmodells kann in zwei parallelen Ansichten erfolgen: einer verkleinerten Anzeige, die die komplette Struktur auf einem Blick erfaßbar macht, und einer vergrößerten Ansicht, die nur einen Ausschnitt der Struktur zeigt und sowohl das Erkennen detaillierter Information als auch die Durchführung verschiedener Berechnungsvarianten erlaubt.

Die Ablaufstruktur der Berechnung weist gewisse Ähnlichkeit mit der Struktur eines UND/ODER-Graphen, dessen Darstellung auf der Philosophie der Reduktion

(der Zerlegung) eines Problems in Unterprobleme basiert, die voneinander unabhängig gelöst werden können. Der in Abb. 2 dargestellte Graph, der im folgenden aufgrund seiner Baumstruktur 'Berechnungsbaum' genannt wird, besteht nur aus UND-Verzweigungen, da er eine bestimmte Berechnungsvariante visualisiert und alle seiner Elemente berücksichtigt werden müssen, während die nicht präsentierten ODER-Verzweigungen die verschiedenen Berechnungsvarianten darstellen. Die Knoten des Baumes enthalten Daten (Werte), die in die Berechnung eingehen, während die Verknüpfungen formelmäßige Abhängigkeiten zwischen diesen Knoten darstellen. Entsprechend der gewählten Berechnungsvariante können die Werte der Knoten sowohl vom System automatisch berechnet werden, falls solche Soteinträge existieren, die für diese Berechnungsvariante vorgesehen sind, als auch durch den Benutzer direkt eingegeben bzw. aus dem Eintragsraum (vgl. Kap. 3.1.) ausgewählt werden. Die für die aktuelle Berechnungsvariante noch unbekanntene Werte sind in der Abbildung jeweils mit Fragezeichen gekennzeichnet.

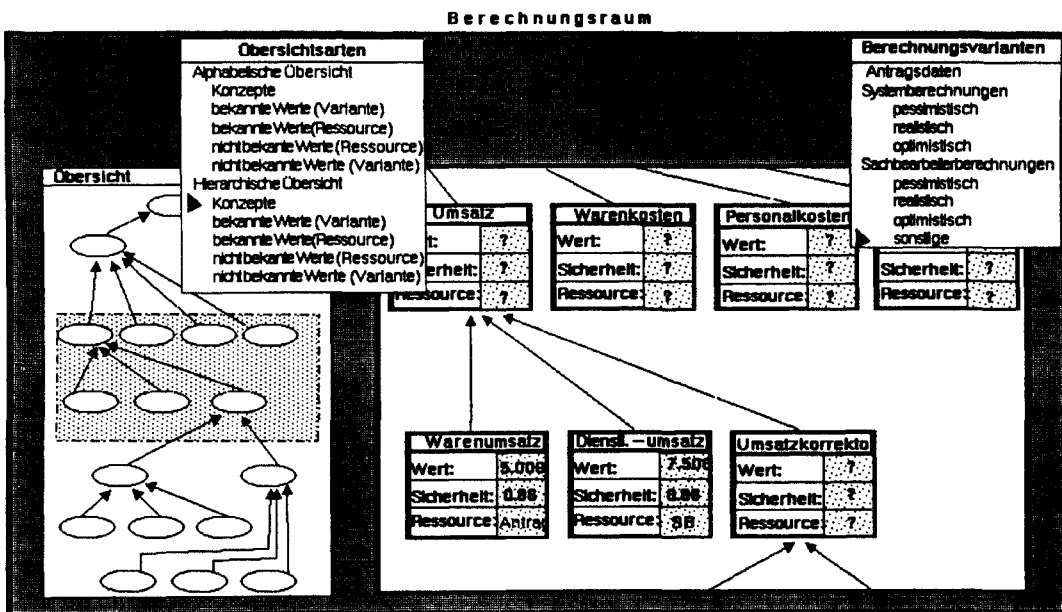


Abb. 2: Der Berechnungsraum (schematisch)

Führt das System eine Berechnungsvariante automatisch aus, so versucht es unter Aktivierung der den Knoten zugeordneten Berechnungsprozeduren Daten von unten nach oben zu aggregieren, in dem der Wert eines Knotens aus unmittelbar nachfolgenden Knoten berechnet wird. Bei den Knoten, für die keine Informationen existieren bzw. sich berechnen lassen, kann angezeigt werden, wo für sie Informationen geholt werden können. Zur Berechnung des Wertes eines Knoten ist jedoch die Berücksichtigung aller unmittelbar nachfolgenden Knoten dann nicht zwingend erforderlich, wenn für diesen Knoten bereits ein oder mehrere Werte in der Wissensbasis existieren oder wenn die Eingabe des Wertes vom Benutzer direkt erfolgen soll. In diesen Fällen hat

der Benutzer die Möglichkeit, auf jeder Ebene des Berechnungsbaumes die nächst höhere Ebene anzusteuern: er bestimmt die Vorbedingungen oder Daten, die auf der nächsten Ebene zu berücksichtigen sind. Dadurch wird dem Benutzer die Möglichkeit eingeräumt, die während der Berechnung erreichten Zwischen-Zustände zu überprüfen, und gegebenenfalls in den Entscheidungsprozeß einzugreifen. Auf jeder Zwischen-Ebene können beispielsweise verschiedene Ergebnisse im Präsentationsraum (vgl. Kapitel 3.3) zusammengestellt werden, um die Werte für die nächst höhere Ebene zu bestimmen, wie auch die Angemessenheit der angewandten Regeln nachgeprüft sowie Erklärungsfragen zu selektierten Knoten gestellt werden (vgl. auch Yetim 1991a).

Durch die Visualisierung des Berechnungsbaumes werden implizit die zielgerichteten Erklärungshilfen geliefert; und zwar dadurch, daß die allgemeine Information über den als nächstes erwarteten Schritt auf dem Weg zum Ziel vermittelt wird und somit auch die oft notwendige Frage 'Warum muß ich das machen?' sich an dieser Stelle erübrigt<sup>8</sup>.

Da im Berechnungsraum Lösungen produziert werden, die auf viele Varianten basieren, ist eine kritische Betrachtung der für den aktuellen Fall durchgeführten Berechnungen und Abweichungen und einen ebenso kritischen Vergleich dieser Falldaten (Fallanalyse) mit ähnlichen bereits bearbeiteten Fällen von großer Bedeutung. Erst die simultane Betrachtung der im Präsentationsraum visualisierten Ergebnisse der Lösungen und der Vergleiche führt zu einer endgültigen Entscheidung.

### 3.3. Der Präsentationsraum

Wie eingangs erwähnt, ist eine der wesentlichen Leistungen des Systems seine Funktion hinsichtlich der Informationsaufbereitung für die intuitiven Einschätzungen des Sachbearbeiters. Diese Aufbereitung macht die vorhandenen Informationen in einer Form verfügbar, in der sie in Entscheidungsprozessen effizient und effektiv wirksam werden können. Der Präsentationsraum dient zur Aufbereitung und Präsentation von Informationen und erlaubt u.a. die *gleichzeitige Betrachtung verschiedener Berechnungsvarianten*, welche dann schon einen detaillierteren Einblick in die Datengrundlage verschafft. Die in Abb. 3. entsprechend dem Vorschlag von Wolf 1988b erfolgte schematische Darstellung, beschränkt sich nur auf einige Präsentationsmöglichkeiten, welche die im vorigen Kapitel thematisierten Berechnungen betreffen<sup>9</sup>.

Dabei handelt es sich in Abb 3. konkret um die

(a) tabellarische Darstellung der Werte zu einem Konzept: dadurch wird auch die Möglichkeit gegeben, daß der Sachbearbeiter nach seinem Wunsch einen oder mehrere

---

<sup>8</sup> Darauf weisen auch Rada/Barlow 1988 in ihrer Arbeit hin, in der sie die Schwächen der von existierenden Expertensystemen produzierten Erklärungen aufzeigen und für eine hypertextartige Darstellung von Regeln des Systems plädieren.

<sup>9</sup> Für weitere Möglichkeiten wird auf Yetim 1991b verwiesen.



Werte aus dieser Tabelle für weitere Berechnungen auswählt.

(b) Darstellung der Beurteilungszonen, in denen die Berechnungsergebnisse liegen müssen, damit entweder eine ablehnende oder eine befürwortende Entscheidung zu zulassen, und der Verteilung der Einzelwerte, um einen visuellen Eindruck zu geben.

(c) Darstellung der Verschiebung in den Beurteilungszonen beispielsweise nach der Herabsetzung der Anforderungen an die Beurteilungssicherheit.

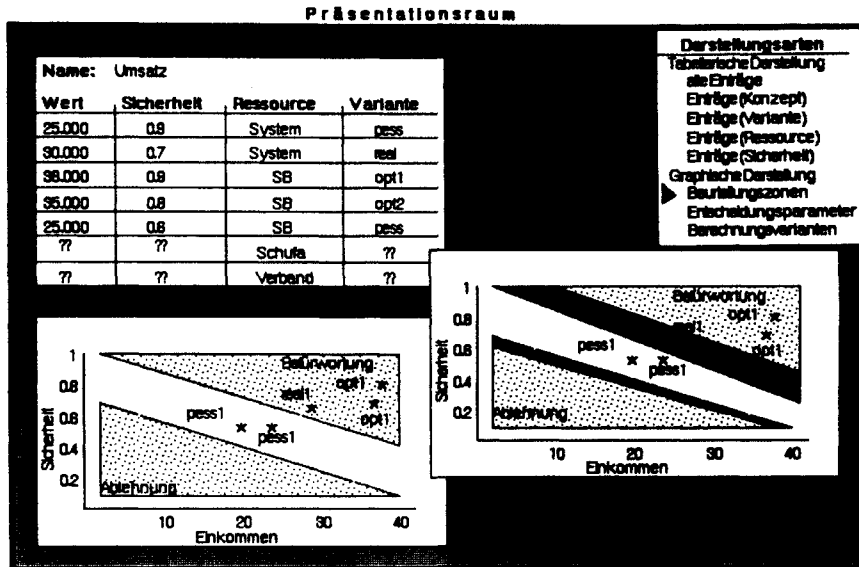


Abb. 3: Der Präsentationsraum (schematisch)

#### 4. Abschlußbemerkung

Durch die Entwicklung einer hypermedia-gestützten Dialogkomponente wird eine Verbesserung der Systemleistung erhofft, die vor allem auch in der Kooperation mit dem Benutzer und in der Unterstützung intuitiver Prozesse des Benutzers besteht. Durch diese Arbeit werden Aufgaben, die das System WISKREDAS mit seinen Schlußfolgerungen nicht beherrscht, auf den Benutzer verlagert.

#### Literatur

- Banos, W. (1990): Textgenerierung mit natürlichsprachlichen Merkmalen am Beispiel des Sachberichts für WISKREDAS; (Studienarbeit) SFB 221/B3.
- Blumer, H. (1990): Eine Hypertext-Erklärungskomponente zu einem Expertensystem. Universität Konstanz, Informationswissenschaft (Diplomarbeit)
- Card, S.K. ; Henderson, A. (1987): A multiple, virtual-workspace interface to support user task switching. In: Carroll, J.M. /Tanner, P.P (eds.): Proceedings of the CHI and GI'87 Conference on Human Factors in Computing Systems. Toronto/New York: ACM, 53-59.
- Dambon, P. (1988): Wissensbasierte Verwaltung eines administrativen Fallbearbeitungsvorgangs: Wissensrepräsentation in der "Fallbasis" von WISKREDAS. Universität Konstanz, Bericht SFB 221/B3-3/88.

- Dambon, P. ; Yetim, F. 1990: Integration of Hypertext into a Decision Support System. In: Herget, J.; Kuhlen, R. (eds.): *Pragmatische Aspekte beim Entwurf und Betrieb von Informationssystemen. Proceedings des 1. internationalen Symposiums für Informationswissenschaft, Universität Konstanz, Informationswissenschaft, 17-19 Oktober 1990, 64-77.*
- Glaser, F. (1990): WIREMAN: Ein wissensbasiertes System zur Erarbeitung von Information aus Datenbanken für die Kreditwürdigkeitsprüfung von Unternehmensgründungen. In: Herget, J.; Kuhlen, R. (eds.): *Pragmatische Aspekte beim Entwurf und Betrieb von Informationssystemen. Konstanz: Universitätsverlag, 99-120.*
- Heizmann, G. (1990): Grafische Bedienoberfläche für das System WISKREDAS (Softwaredokumentation); SFB 221/B3.
- Kuhlen, R.; Dambon, P.; Glaser, F.; Thost, M.; Wolf, M.; Yetim, F. (1990): WISKREDAS: Ein wissensbasiertes System zur Unterstützung der Kreditwürdigkeitsprüfung von Unternehmensgründungen, Bericht SFB 221/B3-10/90.
- Kuhlen, R. (1991a): Praktische und theoretische Visionen von Hypertext. In: *Nachrichten für Dokumentation, 42, 222-224.*
- Kuhlen, R. (1991b): *Hypertext - ein nicht-lineares Medium zwischen Buch und Wissensbank. Berlin et al: Springer Verlag.*
- Rada, R.; Barlow, J. (1988): Expert Systems and Hypertext. In: *The Knowledge Engineering Review, Vol. 3, 1988, 285-301.*
- Rheingold, H. (1990): An Interview with Don Norman. In: Laurel, B. (ed.): *The Art of Human-Computer Interface Design. Massachusetts et al: Addison-Wesley, 5-10.*
- Stefik, M.; Bobrow, D.G. (1986): Object Oriented Programming: Themes and Variations. In: *The AI Magazine, Vol. 6, No. 4, 40-62.*
- Streitz, N. (1990): Hypertext: Ein innovatives Medium zur Kommunikation von Wissen. In: Gloor, P./ Streitz, N. (eds.): *Hypertext und Hypermedia, Berlin et al.: Springer, 1990, 10-27.*
- Thost, M. (1990): *Informationsquellenmodelle. Ein Konzept zur wissensbasierten Bewertung der Glaubwürdigkeit von Meinungen. Konstanz: Martin Thost Verlag.*
- Yetim, F. (1990): Die Implementation der Entscheidungskomponente für WISKREDAS. Universität Konstanz, Informationswissenschaft, SFB 221/B3, (Softwaredokumentation).
- Yetim, F. (1991a): Eine Hypertext-Komponente zu einem Expertensystem: Benutzerfragen für Erklärungsdialoge. In: Maurer, H. (ed.): *Hypertext/Hypermedia '91. Berlin et al: Springer, 1991, 286-298.*
- Yetim, F. (1991b): Integration von Expertensystem- und Hypertext-Techniken am Beispiel des Systems WISKREDAS. Workshop 'Hypersystem-Konzepte in Medien und kultureller Produktion', Lüneburg, 15.-17. Juli 1991.
- Yetim, F. ; Dambon, P. (1991): Can Hypermedia improve the Acceptance of Knowledge-Based Systems?. In: Bullinger, H.-J. (ed.): *Proceedings of the 4th International Conference on Human-Computer-Interaction, Stuttgart, Sept. 1991*

Fahri Yetim, Dipl. Inf. (FH), Dipl. Inf.-Wiss.  
 FG Informationswissenschaft, Universität Konstanz  
 Postfach 5560, D-7750 Konstanz

## Design-Vorschläge für Hypertextapplikationen aufgrund von Feldversuchen

Erich J. Schwarz, Eva Bertha, Sabine Schmidt

Aus den praktischen Erfahrungen bei der Erstellung von Hypertextapplikationen und den Ergebnissen von Feldversuchen zur Systemakzeptanz der Benutzer wurden Design-Vorschläge entwickelt. Diese betreffen Benutzerklassen, Systembedienung, Struktur des Systems und Hilfefunktionen. Die Vorschläge zur Konzeption neuer Hypertextsysteme lauten:

- \* Funktionstasten müssen aus einer Kombination von Bild und Text mit klar verständlichen Metaphern bestehen.
- \* Für Studenten als Systembenutzer können identische Navigationselemente verwendet werden.
- \* Die Orientierung im System kann durch eine "stackmap" verbessert werden. Das Abweichen von völlig vernetzten Strukturen mildert das Problem "lost in hyperspace".
- \* Hilfefunktionen müssen sorgfältig geplant und aussagekräftig gestaltet werden.

Guidelines designing a hypertext system were developed by practical experience and assessment tests. These proposals deal with user groups, browsing and navigation through the system, structure of the system and help functions. The guidelines for designing hypertext systems are:

- \* The metaphors for icons must consist of pictorial and verbal elements together.
- \* Identical browsing facilities and navigation instruments can be used for different groups of students
- \* The orientation of the users within the system can be improved with a stackmap. The network structure of the system should be combined with an underlying hierarchy.
- \* Help functions must be planned very carefully and be easily recognizeable by the users.

## **1. Einleitung:**

Um die Einsatzmöglichkeiten von Hypermedien zu untersuchen, wurden mehrere Applikationen für den Bildungs- und Kulturbereich entwickelt, wobei jeweils unterschiedliche Schwerpunkte gesetzt wurden.

In einem Studienauskunftssystem für die Universität Graz werden die Probleme der Datenmanipulation im System HyperCard behandelt. Da sich die Daten in dieser Applikation laufend ändern, etwa Ort und Zeit von Lehrveranstaltungen, wird an diesem Beispiel versucht Datenkonsistenz und -integrität in einer HyperCard-Applikation zu gewährleisten. Die Eingabe- und Korrekturfunktionen werden über benutzerfreundliche Masken realisiert.

Anhand eines multimedialen Museumsführer werden die Möglichkeiten einer automationsunterstützten Applikationserstellung analysiert. Der Unterschied zu den übrigen Hypertextsystemen besteht darin, daß diese Applikation dem Museumsadministrator als Werkzeug dienen soll, um ein den Bedürfnissen des Endbenutzers entsprechendes Informationssystem zu erstellen. Durch Standardisieren der Schnittstellen wird versucht, die Übernahme von unterschiedlichen Daten (Text, Bild und Ton) zu erleichtern.

An einem Bibliotheksauskunftssystem für die Universitätsbibliothek der Technischen Universität Graz wurde die Akzeptanz der Benutzer in zwei Feldversuchen getestet. Im ersten Test wurde die Version 0.x dieser Applikation von 78 Personen bewertet /1/. Die Testpersonen - größtenteils Studenten der Technischen Universität - benutzten das in der Bibliothek aufgestellte System freiwillig. Untersucht wurden die Benutzerklassen, die Hardware- und Systembedienung, die Benutzeroberfläche, der Inhalt des Systems, das Medium Hypertext und das Navigieren im System /2/. Die Ergebnisse dienten als Grundlage für eine Verbesserung des Systems (Version 1.0) /3/. In einem neuerlichen Benutzertest bewerteten 63 Testpersonen sowohl die modifizierte als auch die unveränderte Applikation. Der Fragebogen für den zweiten Test wurde lediglich leicht verändert, um

die Ergebnisse mit der ersten Testreihe besser vergleichen zu können. Da beim ersten Feldversuch über 90 % der Systembenutzer Studenten waren, wurden für den zweiten nur mehr Studenten der Technischen Universität und der Karl-Franzens Universität Graz (etwa zu gleichen Teilen) als Testpersonen herangezogen. Die Studenten wurden in Lehrveranstaltungen des Institutes ersucht, freiwillig an den Tests teilzunehmen. Die Ergebnisse der beiden Feldversuche unterscheiden sich in bezug auf die Version 0.x nur unwesentlich.

Aus den Erfahrungen bei der Erstellung der erwähnten Hypertextapplikationen und den Ergebnissen der Feldversuche wurden Design-Vorschläge für künftige Systeme entwickelt.

## **2. Design-Vorschläge**

### **2.1. Benutzerklassen**

Beim zweiten Feldversuch wurden die einzelnen Benutzergruppen mit Hilfe des Chiquadrat-Unabhängigkeitstests auf Unterschiede bei der Systembewertung untersucht, um allgemein gültige Designvorschläge machen zu können. Die Testpersonen mußten Fragen betreffend die **Systembedienung** (Hilfetaste, Volltextsuche, Kontrollpaneel, Dimensionsselektor, Suchvorgang und Indexfenster) und die **Systemeigenschaften** (Originalität, Übersichtlichkeit, Verständlichkeit, Sinnhaftigkeit und Geschwindigkeit) mit einer vierstufigen Skala bewerten. Die Nullhypothese lautete: die Benotung der einzelnen Fragen und die unterschiedlichen Benutzerklassen sind voneinander unabhängig. Bei allen Fragestellungen wird von einem Signifikanzniveau von  $\alpha = 0,05$  ausgegangen /4/.

Die Testpersonen wurden in folgende Benutzerklassen unterteilt:

- \* Frauen versus Männer
- \* Programmierer versus Computer-Anwender
- \* Techniker versus Nicht-Techniker
- \* Nicht-Apple Anwender versus Apple Anwender

Im Gegensatz zu Baird /5/ können keine zuverlässigen Aussagen über das Verhalten von Personen mit und ohne Computererfahrung getroffen werden, weil die Benutzerklasse der Computerunerfahrenen lediglich 9 Personen umfaßt /6/.

#### **Geschlecht :**

Ein interessantes Ergebnis brachte der Vergleich zwischen Frauen und Männern. In bezug auf die Systembedienung konnte mit Ausnahme der Beurteilung der Indexfenster kein Unterschied zwischen den Geschlechtern festgestellt werden. Die Systemeigenschaften hingegen beurteilten die weiblichen Testpersonen in drei Fragestellungen weniger kritisch als ihre männlichen Kollegen. So fanden die Frauen die Hypertextapplikation origineller, sinnvoller und übersichtlicher als die Männer. An dieser Stelle ist anzumerken, daß die Gruppe der Frauen verhältnismäßig klein war (15 Frauen versus 48 Männer).

#### **EDV-Erfahrung:**

Der Vergleich zwischen Programmierern und reinen Computeranwendern zeigte, daß bei der Systembedienung mit Ausnahme der Beurteilung der Hilfetaste keine signifikanten Unterschiede existieren. Bei den Fragestellungen Verständlichkeit, Originalität und Sinnhaftigkeit des Systems kann die Nullhypothese verworfen werden: Programmierer urteilen in diesen Bereichen kritischer als bloße Computeranwender.

#### **Fachrichtung:**

Es bestehen keine signifikanten Unterschiede hinsichtlich der Systembedienung zwischen Studenten aus technischen und nicht-technischen Fachrichtungen. Bei der Beurteilung der Systemeigenschaften treten lediglich bei der Bewertung der Originalität signifikante Unterschiede auf. Techniker empfinden das System weniger originell als Nicht-Techniker.

#### **Betriebssysteme:**

Die Erfahrung mit Apple-Computern war das Merkmal, das hier zur Gruppenbildung herangezogen wurde. Da bei keiner Fragestellung signifikante Unterschiede festgestellt werden können, wird die Hypothese angenommen, daß der "apple-erfahrene" Benutzer Hyper-

textapplikationen nicht anders beurteilt als der "apple-unerfahrene".

Bei genauer Betrachtung der Ergebnisse sieht man, daß kaum signifikante Unterschiede bei der Beurteilung der Systembedienung existieren. Aus diesem Grund können für unterschiedliche Benutzergruppen von Hypertextsystemen gleiche Bedienungselemente verwendet werden. Die Systemakzeptanz einer Hypertextapplikation wird aber nicht nur von der Systembedienung sondern auch von Faktoren wie Originalität oder Übersichtlichkeit beeinflusst.

	Die getestete Hypertextapplikation ist			
	sinnvoll	originell	verständlich	übersichtlich
Mann	<b>1,59</b>	<b>2,18</b>	1,88	<b>1,73</b>
Frau	<b>1,33</b>	<b>1,47</b>	1,67	<b>1,33</b>
Programmierer	<b>1,68</b>	<b>2,58</b>	<b>1,95</b>	1,63
Anwender	<b>1,42</b>	<b>1,78</b>	<b>1,78</b>	1,64
Techniker	1,71	<b>2,29</b>	1,93	1,68
Nicht Techn.	1,40	<b>1,83</b>	1,77	1,63

Tab.1: Mittelwerte im Vergleich (vierteilige Skala)  
 Fettgedruckte Werte geben jene Gruppen an, bei denen der  
 Chiquadrat-Unabhängigkeitstest signifikante Unterschiede zeigt.

Wie Tabelle 1 zu entnehmen ist, existieren besonders kritische Benutzergruppen. Plant man die Erstellung einer Hypertextanwendung für eine solche Gruppe, müssen bereits beim Systemdesign diese Faktoren berücksichtigt werden.

## 2.2 Piktogramme

Bei der Version 0.x hat sich die Navigation durch das System als Schwachstelle erwiesen. So wurden die Navigationsinstrumente "Dimensionsselektor" und "Kontrollpaneel" von den Testpersonen schlecht beurteilt. Eine wesentliche Ursache dafür dürften die schwer verständlichen Metaphern bei der verwendeten Symbolik gewesen sein /7/. Daher wurden die zum Teil sehr abstrakten Piktogramme des Dimensionsselektors der Version 0.x durch reali-

stischere Darstellungen ersetzt (Abb. 1). Um das Verständnis zu erhöhen, wurden die Symbole zusätzlich mit Schrift versehen.

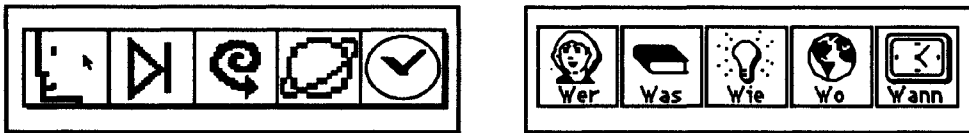


Abb.1: Dimensionssselektor der Versionen 0.x bzw. 1.0

Der zweite Feldversuch zeigt, daß die Funktionstasten für den gelegentlichen Benutzer verständlicher werden, wenn diese durch eine Kombination aus Bild und Text anstelle abstrakter Symbole dargestellt werden. Der Dimensionssselektor der Version 0.x wurde auf einer vierteiligen Skala mit 2,12 (Mittelwert), der geänderte mit 1,25 bewertet. Dieses Ergebnis bestätigt auch die experimentellen Untersuchungen von Guastello. Er verifizierte die Hypothese, daß "mixed modality metaphors are rated as more meaningful than icons that utilize verbal or pictorial elements only" /8/.

Das Kontrollpaneel (Abb.2) erhielt trotz Änderungen keine besseren Bewertungen. Sowohl die ursprüngliche als auch die modifizierte Version wurde mit einem Mittelwert von 1,97 bewertet. Mögliche Ursachen dafür könnten die komplexe Struktur der getesteten Applikation und die Start-Taste sein.

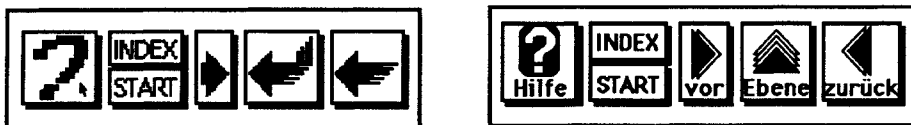


Abb.2: Kontrollpaneel der Versionen 0.x bzw. 1.0

So gelangt der Benutzer bei Betätigung der Start-Taste nicht direkt zur Systemübersichtskarte sondern an den Programmbeginn und muß jedesmal den nicht überspringbaren Programmvorspann ablaufen lassen. Aus diesem Grund ist bei der Neuüberarbeitung dieser Applikation eine Funktionstaste vorgesehen, die eine direkte Verbindung zur Übersichtskarte herstellt.

Für den gelegentlichen Benutzer ist auch die stark vernetzte Struktur dieses Hypertextapplikation schwer verständlich. Da die



wenigen Hierarchien kaum ausgeprägt sind, war die Bedeutung der Taste "Ebene" für viele Testpersonen nicht ersichtlich.

### 2.3 Systemstruktur

Obwohl der Dimensionsselektors der Version 1.x positiver bewertet wurde als der der Version 0.x, erfolgte keine Verbesserung bei der Bewertung der Frage, ob der Tester das Gefühl hat, "Herr der Lage" zu sein. Durch dieses Ergebnis wird die Annahme nicht bestätigt, daß durch leichter verständliche Navigationselemente die Orientierung im System verbessert wird /2, S.58/. Um diese Probleme des Benutzers zu lösen, die Conklin mit "lost in hyperspace" bezeichnet, müssen daher weitere Maßnahmen getroffen werden /9/. Trotz der Möglichkeit ein vollständiges und komplexes Netzwerk mit einem derartigen System aufzubauen, sollte auf hierarchische Strukturen nicht völlig verzichtet werden. Dadurch wird ein Hypertextsystem nicht nur für den Benutzer übersichtlicher, sondern ermöglicht auch dem Systemadministrator leichter Änderungen bzw. Erweiterungen durchzuführen. Conklin schlägt zur Unterstützung der Orientierung im System einen "graphical browser" vor, der die visuellen Fähigkeiten des Menschen ausnützt.

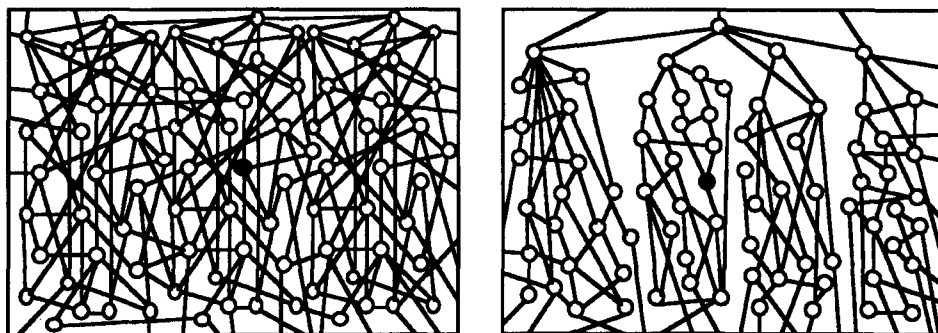


Abb.3: Stackmap eines völlig vernetztes Systems

Stackmap eines leicht strukturierten Systems

Mit Hilfe einer zwei- bis dreidimensionalen "stackmap" kann der Benutzer die momentane Position im System leichter erkennen, wie Abbildung 3 zu entnehmen ist.

## **2.4. Hilfe im System**

Das getestete Hypertextsystem bietet zwei Hilfefunktionen an. Das Kontrollpaneel (siehe Abb.2) enthält eine Hilfetaste, die jederzeit aktiviert werden kann und dem Benutzer Bedienungshinweise liefert. Mehr als 90% der Testpersonen bewerteten diese Systemhilfe positiv. Außerdem ermöglicht diese Applikation dem Benutzer direkt von der Übersichtskarte in eine "Systemeinführung" zu gelangen. Trotz positiver Bewertung der beiden Hilfefunktionen wurde deren Erweiterung und Verbesserung häufig von den Testpersonen gefordert. Es sollte daher bei der Neukonzeption von Hypertextsystemen besonders auf die sorgfältige und umfangreiche Gestaltung von Hilfefunktionen bedacht genommen werden. Auch die Implementierung einer "Systemdemo" sollte in Betracht gezogen werden.

## **3. Resümee**

Die umfangreichen Benutzerbefragungen brachten neue Erkenntnisse, die bei der Planung und Erstellung von Hypertextapplikationen berücksichtigt werden sollten:

- Funktionstasten werden für den gelegentlichen Benutzer verständlicher, wenn sie durch eine Kombination aus Bild und Text anstelle abstrakter Symbole allein dargestellt werden.
- Da die Elemente der Systembedienung von den unterschiedlichen Benutzergruppen ähnlich beurteilt werden, können für Studenten als Systembenutzer identische Navigationselemente verwendet werden.
- Orientierungsprobleme der Benutzer sind häufig nicht auf schlechte Navigationsinstrumente sondern auf die Struktur der Applikation zurückzuführen. Es müssen Methoden gefunden werden, die dem Benutzer erlauben, seinen Standort in der Systemstruktur zu erkennen. Die Implementierung einer "Stack-

map" ist bei Neukonzeptionen anzuraten. Durch eine Kombination aus vernetzter und hierarischer Struktur wird die Übersichtlichkeit der Applikation sowohl für den Benutzer als auch für den Systemadministrator erhöht.

- Auch bei der Erstellung von verhältnismäßig kleinen Hypertextsystemen ist auf die sorgfältige und umfangreiche Gestaltung der Hilfefunktion zu achten. Die Implementierung einer "Systemdemo" sollte in Betracht gezogen werden.

Bei der praktischen Erstellung von Hypertextapplikationen sind noch weitere Problemfelder aufgetreten. So sind Änderungen und Verbesserungen schon bei kleinen, nicht klar strukturierten Applikationen mit großem Aufwand verbunden. Auch ist das Verwalten der Daten in einem externen Datenbanksystem oder die Standardisierung der Schnittstellen zu Peripheriegeräten noch nicht zufriedenstellend gelöst. Durch Lösung dieser Problemfelder kann sich die Einsatzmöglichkeit von Hypertextsystemen wesentlich erhöhen.

### **Anerkennung**

Die Autoren bedanken sich bei Univ. Prof. Dr. Wolf Rauch und bei Univ. Prof. Dr. Hans Günther Kopetzky für die Unterstützungen und Anregungen, die zur Abfassung dieses Beitrages geführt haben.

### **Literatur**

/1/ Kappel, Herbert J.: Test eines HyperCard-Informationssystemes. 1991. Diplomarbeit am Institut für Informationswissenschaft, Universität Graz.

/2/ Bertha, Eva; Schwarz, Erich J.: Vom Prototyp zum Produkt Empirische Untersuchungen zum Akzeptanzverhalten einer Hypertextapplikation. In: Hypertext/Hypermedia '91. 27-28.Mai 1991 in Graz. Hrsg. H.Maurer. Berlin: Springer 1991, (Informatik-Fachberichte 276) S.52-62

- /3/ Bagary, Gudrun E.: Von der Laborversion zum fertigen Produkt. Weiterentwicklung einer HyperCard Applikation. 1991. Diplomarb. am Institut für Informationswissenschaft, Universität Graz. (In Vorbereitung)
- /4/ Bohley, Peter: Statistik. 2.Aufl. München: Oldenbourg 1987. ISBN 3-486-20323-1
- /5/ Baird, Patricia; Percival, Marc: Glasgow Online. Database Development using Apple's HyperCard. In: Hypertext theory into practice. Ed. by R. MacAleese. Oxford: Blackwell 1989, 75-92
- /6/ Schmidt, Sabine; Testen von Hypertext-Applikationen, Diplomarb. am Institut für Informationswissenschaft, Universität Graz (In Vorbereitung)
- /7/ Staufer, Michael J.: Piktogramme für Computer. Kognitive Verarbeitung, Methoden zur Produktion und Evaluation. Berlin: de Gruyter 1987. ISBN 3-11-010917-4
- /8/ Guastello, Stephen J.; Traut, Mary; Korienek, Gene: Verbal versus pictorial representations of objects in a human-computer interface. Int. J. Man Machine Studies, 31 (1989) 1, 99-120
- /9/ Conklin, Jeff: Hypertext. An Introduction and Survey. IEEE Computer 20 (1987) 9, 17-41

#### **Autoren**

Dipl.-Ing. Eva Bertha, Universitätsassistent  
cand.rer.soc.oec. Sabine Schmidt, Diplomandin

Beide:

Institut für Informationswissenschaft  
Karl-Franzens-Universität Graz  
Strassoldogasse 10  
A-8010 Graz, Austria

Dipl.-Ing. Erich Schwarz, Universitätsassistent  
Institut für Innovationsmanagement  
Karl-Franzens-Universität Graz  
Johann-Fuxgasse 32  
A-8010 Graz, Austria

# Zur Beziehung zwischen Entwicklungsaufwand und Leistungsfähigkeit eines Wörterbuches für die automatische Indexierung

Kostas Tzeras, Peter Reiss

## Zusammenfassung

Für die automatische Indexierung mit einem vorgegebenen Deskriptorensystem wird ein Wörterbuch benötigt, dessen gewichtete Relationen jeweils einen Term mit einem Deskriptor verbinden. Diese Relationen werden größtenteils durch eine aufwendige statistische Auswertung von Referatetexten und intellektueller Indexierung einer Dokumentmenge gewonnen. Dabei ist es von großem Interesse, den für eine bestimmte Indexierungsqualität erforderlichen Entwicklungsaufwand eines Wörterbuches im voraus abschätzen zu können. In der vorliegenden Arbeit werden von H. Hüther theoretisch hergeleitete Schätzverfahren verwendet, um die Beziehung zwischen der Anzahl der zur Entwicklung benutzten Dokumente und der Größe und Leistungsfähigkeit des Wörterbuches experimentell zu untersuchen.

## Abstract

For the automatic indexing with a controlled vocabulary a dictionary is needed, which contains relationships between terms from the text and descriptors from the controlled vocabulary. When the dictionary is derived from a collection of documents, the question of the interrelationship between the collection size, the dictionary size and the dictionary performance raises. In this paper we use formulas derived by H. Hüther for estimating the interrelationship between these parameters.

## 1 Einleitung

Die Automatisierung der Indexierung ist in der Retrievalpraxis bisher auf wenige Fälle beschränkt, was in erster Linie auf das Fehlen der dafür benötigten Wörterbücher zurückzuführen ist. Die Entwicklung eines Wörterbuches für die automatische Indexierung erfordert einen großen Aufwand, der in vielen Fällen erst durch den Einsatz automatischer Verfahren zur Extraktion des Wörterbuchwissens aus einer Menge von Texten bewältigt werden kann.

Da globale Wörterbücher, die in vielen Fachgebieten einsatzfähig wären, praktisch nicht erstellbar sind, hängt die Anwendung eines Indexierungssystems in einem Fachgebiet im wesentlichen von der Entwicklung eines fachspezifischen Wörterbuches ab. Dabei ist es von großem Interesse, den für eine bestimmte Indexierungsqualität erforderlichen Aufwand im voraus abschätzen zu können. Gesucht ist also eine Beziehung zwischen der Anzahl der zur Entwicklung benutzten Texte und der Größe und Leistungsfähigkeit des Wörterbuches.

Derartige Beziehungen wurden von Hüther ([Hüther 89]) für die automatische Indexierung mit einem vorgegebenen Deskriptorensystem untersucht und formal beschrieben. Weiterhin konnte die Zuverlässigkeit eines von Hüther vorgeschlagenen Verfahrens zur Abschätzung der Größe eines Wörterbuches in Abhängigkeit von der Anzahl der zugrundeliegenden Texte nachgewiesen werden ([Tzeras 91]).

In dieser Arbeit haben wir das letztgenannte Verfahren angewendet, um Wörterbuchgrößen für verschiedene Anzahlen von Ausgangstexten abzuschätzen. Weiterhin haben wir die Vollständigkeit des Wörterbuches mit Hilfe von in [Hüther 90] hergeleiteten Erwartungswerten für verschiedene Wörterbuchgrößen geschätzt. Analog dazu wurden für die jeweiligen Wörterbuchgrößen an Precision und Recall orientierte Maße ermittelt, die eine erste Annäherung für die Qualität der Indexierungsergebnisse darstellen. Um diese Ergebnisse möglichst gut abzusichern, haben wir diejenigen Schätzwerte, für die empirische Werte verfügbar sind, mit diesen verglichen und auf Genauigkeit untersucht.

In Abschnitt 2 wird die automatische Indexierung näher konkretisiert und der im Fachgebiet Datenverwaltungssysteme II der Technischen Hochschule Darmstadt entwickelte Lösungsansatz erläutert. Dabei steht die Rolle des Wörterbuches innerhalb eines automatischen Indexierungssystems im Vordergrund, wie es für den Zweck dieser Arbeit notwendig ist. Eine detaillierte Darstellung des Indexierungsansatzes ist in [Lustig 86a] und [Fuhr et al. 91] nachzulesen.

Abschnitt 3 enthält eine Herleitung der verwendeten Erwartungswerte und die Definition der benötigten Qualitätsmaße. Die experimentellen Ergebnisse werden dann in Abschnitt 4 vorgestellt und diskutiert. Dabei werden die Daten aus der Produktion des Wörterbuches PHYS/PILOT, das als Bestandteil des Indexierungssystems AIR/PHYS<sup>1</sup> aus 392 000 Referatetexten entwickelt wurde, für die Berechnung und teilweise für die Evaluierung der Erwartungswerte verwendet.

In Abschnitt 5 wird die Vorgehensweise zur weiteren Untersuchung der Zusammenhänge zwischen Entwicklungsaufwand und Leistungsfähigkeit eines Wörterbuches für die automatische Indexierung skizziert.

## 2 Das Indexierungswörterbuch

Die vorliegende Arbeit geht von einem vorgegebenen Deskriptorensystem aus. Das bedeutet, daß die Indexierung nicht „am Text kleben“ darf; sie muß begriffsorientiert sein. Weiterhin wird angenommen, daß nicht vollständige Dokumente, sondern nur Referate (Abstracts) einschließlich der Titel indexiert werden und daß diese in englischer Sprache vorliegen<sup>2</sup>.

Für die begriffsorientierte Indexierung muß ein Indexierungssystem die wichtigsten Stellen eines Textes herausfinden und durch Deskriptoren darstellen. Die Entscheidung über die Zuteilung eines Deskriptors zu einem Text kann in der automatischen Indexierung grundsätzlich nur aufgrund solcher Texteingenschaften getroffen werden, die rein formal verifizierbar sind.

Dementsprechend wird im Rahmen des zugrundeliegenden Darmstädter Indexierungsansatzes (DIA) der Indexierungsvorgang in zwei Schritte aufgeteilt: einen Beschreibungs- und einen Entscheidungsschritt (siehe Abb. 1 und [Lustig 86b]).

Im Beschreibungsschritt werden zunächst alle in einem Dokument  $d$  enthaltenen Eigenschaften bestimmt, die für die Zuteilung bzw. Nicht-Zuteilung des Deskriptors  $s$  relevant sein können. Dies geschieht durch die Beschreibungsfunktion  $x$ , die das Paar  $(d, s)$  in die Relevanzbeschreibung  $x(d, s)$  abbildet. Die Relevanzbeschreibung bildet die Entscheidungsgrundlage für die Indexierungsfunktion  $a$ , die jeder Relevanzbeschreibung  $x(d, s)$  ein Indexierungsgewicht  $g = a(x(d, s))$  zuordnet, das die Wahrscheinlichkeit approximiert, daß der Indexierungsfachmann den Deskriptor  $s$  dem Dokument  $d$  zuteilen würde, falls die Relevanzbeschreibung  $x(d, s)$  vorliegt. Die von der Indexierungsfunktion gelieferten Gewichte können durch die Einführung eines Schwellenwertes  $g_0$  auf die Gewichte 0 (Nicht-Zuteilung) bzw. 1 (Zuteilung) abgebildet werden.

Das Indexierungswörterbuch spielt bei der Erstellung von Relevanzbeschreibungen innerhalb des

<sup>1</sup>Das Indexierungssystem AIR/PHYS wird seit 1985 für die Inputproduktion der Physik-Datenbank des Fachinformationszentrums Karlsruhe eingesetzt (siehe den Beitrag von M. Schwantner in diesem Tagungsband oder [Fuhr et al. 91] und [Lustig 86a]).

<sup>2</sup>Im folgenden ist daher außer bei der intellektuellen Indexierung mit „Text“ bzw. „Text eines Dokuments“ bzw. „Dokument“ stets der Referatetext einschließlich des Titels gemeint.

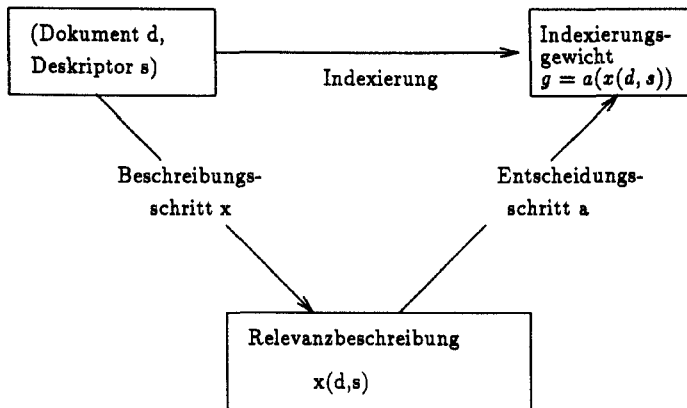


Abbildung 1: Aufteilung des Indexierungsvorganges in einen Beschreibungs- und einen Entscheidungsschritt

Beschreibungsschrittes eine zentrale Rolle. Es verbindet möglichst viele Terms (einfache oder zusammengesetzte Fachausdrücke, Formeln, fachspezifische Abkürzungen) des Anwendungsgebietes durch Relationen mit Deskriptoren. Während rein semantische Relationen<sup>3</sup> nur einen kleinen Teil des Wörterbuches ausmachen, spielt die statistische Relation  $Z$  sowohl in quantitativer als auch in qualitativer Hinsicht die entscheidende Rolle.

Die Relation  $Z$  zwischen einem Term  $t$  und einem Deskriptor  $s$  beruht auf dem Assoziationsfaktor  $z$  ([Lustig 69]). Bei Vorliegen einer sehr großen intellektuell indextierten Dokumentmenge wird  $z$  (im folgenden  $z$ -Wert genannt) wie folgt berechnet:

$$z(t, s, U) = \frac{h(t, s, U)}{f(t, U)}$$

Dabei bezeichnet

- $U$  eine Menge von Formen des Vorkommens<sup>4</sup>,
- $f(t, U)$  die Anzahl der Dokumente, in denen der Term  $t$  in eine zu  $U$  gehörenden Form des Vorkommens auftritt (Dokumenthäufigkeit), und
- $h(t, s, U)$  die Anzahl derjenigen unter diesen Dokumenten, denen der Deskriptor  $s$  intellektuell zugeteilt wurde (Paarhäufigkeit).

Der  $z$ -Wert approximiert die Wahrscheinlichkeit, daß beim zugrundeliegenden Deskriptorensystem ein Indexierungsfachmann den Deskriptor  $s$  einem Dokument zuteilen würde, falls der Term  $t$  in einer zu  $U$  gehörenden Form des Vorkommens im Referat des Dokuments identifiziert wird.

Ist eine formal verifizierbare Definition der erlaubten Formen des Vorkommens gegeben, so kann die Eigenschaft „ein Term  $t$  kommt in einem Dokument vor“ algorithmisch identifiziert und die Relation

<sup>3</sup>z.B. Synonym, Ober-, Unterbegriff, verwandter Begriff, USE-, SEE-Relation u.ä.

<sup>4</sup>Beispiele für Formen des Vorkommens sind bei Einzelwörtern Übereinstimmung zwischen Term im Text und Wörterbucheintrag in der Grund- bzw. Stammform; bei Mehrwortgruppen Grund- bzw. Stammformübereinstimmung und eine maximale Anzahl von anderen Wörtern zwischen den Komponenten der Mehrwortgruppen.

$Z$  aus einer Menge intellektuell indexierter Dokumente automatisch gewonnen werden. Prinzipiell ist eine stets exakte Identifizierung von Terms durch linguistische Verfahren möglich, in der Praxis enthält jedoch die Menge  $U$  nur solche Formen des Vorkommens, die mit Hilfe einfacher, meist heuristischer Verfahren erkannt werden können.

Im folgenden wird das Argument  $U$  nicht weiter in den Bezeichnungen mitgeführt, da sich die Betrachtungen auf die für den jeweiligen Wörterbuchaufbau zugrundegelegte Menge  $U$  beziehen.

Aus praktischen und statistischen Gründen werden nur solche  $z$ -Werte in das Indexierungswörterbuch übernommen, die mit geeigneten Konstanten  $c_1$  und  $c_2$  die Bedingungen  $z(t, s) \geq c_1$  und  $h(t, s) \geq c_2$  erfüllen.

Das soeben beschriebene Wörterbuch wird bei einem neu zu indexierenden Text folgendermaßen für die Erstellung von Relevanzbeschreibungen verwendet: Das Indexierungssystem findet einen Hinweis  $t \rightarrow s$  von einem Term  $t$  auf einen Deskriptor  $s$  genau dann, wenn es den Term  $t$  als im Text vorkommend erkennt und eine Relation zwischen  $t$  und  $s$  im Wörterbuch feststellt.

Ein Hinweis wird durch verschiedene Eigenschaften beschrieben, wie z.B. die Form des Vorkommens des Terms  $t$ , Angaben über die Relation zwischen  $t$  und  $s$  und weitere Informationen über  $s$  und  $t$ . Nach Erstellung der Hinweise von Terms des Textes auf eine Menge von Deskriptoren werden die Relevanzbeschreibungen gebildet. Die Relevanzbeschreibung  $x(d, s)$  eines Deskriptors  $s$  bezüglich eines Dokuments  $d$  enthält alle Hinweise von Terms des Textes auf den Deskriptor  $s$  und auch Eigenschaften, die nicht einem einzelnen Hinweis zuzuordnen sind<sup>5</sup>.

Die für die Entscheidung erforderliche Bewertung der Relevanzbeschreibung erfolgt nun im Entscheidungsschritt durch die eingangs beschriebene Indexierungsfunktion  $a$ . Für die Entwicklung von Indexierungsfunktionen wird eine Stichprobe von intellektuell indexierten Dokumenten verwendet, auf die ein probabilistischer Klassifikationsalgorithmus angewendet wird. Verschiedene Algorithmen für die Klassifikation von Objekten auf der Basis von Stichproben korrekter Entscheidungen wurden in der Literatur vorgeschlagen. In Zusammenhang mit der automatischen Indexierung wurden bisher einige Ansätze experimentell untersucht und evaluiert (vgl. [Beinke-Geiser et al. 86], [Knorz 83], [Pfeifer 90], [Faißt 90]).

Die Aufteilung des Indexierungsvorganges in Beschreibungs- und Entscheidungsschritt ermöglicht eine differenzierte Beurteilung der Leistungsfähigkeit eines Indexierungswörterbuches. Diese kann zum einen aufgrund der Ergebnisse des Beschreibungsschrittes (Wörterbuchqualität), zum anderen aufgrund der Ergebnisse des darauf aufbauenden Entscheidungsschrittes (Indexierungsqualität) beurteilt werden. Dazu müssen die Begriffe Wörterbuchqualität und Indexierungsqualität definiert und mit geeigneten Maßen quantifizierbar gemacht werden (vgl. auch [Schwantner 91]).

Die Verwendung von Maßen, die auf den Ergebnissen des Beschreibungsschrittes beruhen, hat den Vorteil, daß der Einfluß des Wörterbuches isoliert vom Einfluß der Indexierungsfunktion betrachtet werden kann. Sie ermöglicht eine Beurteilung der Entscheidungsgrundlage für die Indexierungsfunktion und liefert obere bzw. untere Grenzen für die erreichbare Indexierungsqualität.

Beurteilt man ein Wörterbuch aufgrund der Indexierungsergebnisse, so lassen sich die Aussagen nicht ohne weiteres auf ein System übertragen, in dem eine andere Indexierungsfunktion angewendet wird, die Qualitätsbeurteilung ist dafür umso genauer.

Generell kann man sagen, daß ein Zusammenhang zwischen der Menge der zum Wörterbuchaufbau verwendeten Texte, der Wörterbuchgröße und der Wörterbuchqualität existiert. Weiterhin existiert ein Zusammenhang zwischen der Wörterbuchqualität und den Indexierungsergebnissen, der jedoch ausschließlich durch die jeweils verwendete Indexierungsfunktion bestimmt wird.

Für die Untersuchung der Beziehung zwischen Entwicklungsaufwand und Leistungsfähigkeit eines Indexierungswörterbuches bietet sich daher die Untersuchung der gegenseitigen Abhängigkeiten dieser vier Parameter an. Um diese Abhängigkeiten experimentell untersuchen zu können, wäre es denkbar, verschiedene Wörterbuchvarianten aufzubauen und deren Größe, sowie Wörterbuch- und Indexie-

<sup>5</sup>z.B. Länge des Referatetextes von  $d$ , Anzahl der Hinweise von Terms des Textes auf den Deskriptor  $s$ , usw.



rungsqualität zu bestimmen. Da diese Vorgehensweise aus Aufwandsgründen nicht in Frage kommt (ein Wörterbuchaufbau kostet einige hundert CPU-Stunden), muß stattdessen auf Schätzverfahren zurückgegriffen werden.

### 3 Schätzverfahren für die Wörterbuchgröße und -qualität

Die in Abschnitt 2 vorgestellte Relation  $Z$  macht sowohl in qualitativer als auch in quantitativer Hinsicht den wichtigsten Bestandteil<sup>6</sup> eines Indexierungswörterbuches aus und trägt somit entscheidend zur Qualität der Indexierung bei. Daher sollen im folgenden die übrigen Relationen vernachlässigt werden, d.h. man betrachtet nur die Relation  $Z$ . Die automatische Gewinnung der Relation  $Z$  ermöglicht eine formale Beschreibung der Abhängigkeiten zwischen der Anzahl der zugrundeliegenden Texte, der Größe und der Qualität der daraus berechneten Relation  $Z$ .

#### 3.1 Ein Schätzverfahren für die Größe der Relation $Z$

In [Hüther 89] werden verschiedene Möglichkeiten zur Schätzung der Größe der Relation  $Z$  in Abhängigkeit vom Entwicklungsaufwand vorgeschlagen. Dabei wird die Größe der Relation  $Z$  an der Anzahl der verschiedenen Term-Deskriptor-Paare  $(t, s)$  gemessen, während der Entwicklungsaufwand durch die Anzahl der zugrundeliegenden Dokumente und den in Abschnitt 2 eingeführten Schwellenwerten  $c_1$  und  $c_2$  beschrieben wird.

Eine Möglichkeit, die Anzahl der verschiedenen Term-Deskriptor-Paare der Relation  $Z$  abzuschätzen, besteht darin, aufgrund einer kleinen Stichprobe von Dokumenten eine Teilmenge der Relation  $Z$  zu bilden, um dann mittels einer Hochrechnung einen Erwartungswert für die Größe der gesamten Relation  $Z$  zu erhalten.

Im folgenden wird mit  $(t, s)$  ein Term-Deskriptor-Paar bezeichnet, wobei  $t$  den Term und  $s$  den Deskriptor repräsentiert. Für ein Dokument wird die Bezeichnung  $d$  verwendet. Die Anzahl der zur Entwicklung des Wörterbuches zur Verfügung stehenden Dokumente (Korpus) wird mit  $n$  bezeichnet, während die Anzahl der Dokumente einer Stichprobe aus dem Korpus mit  $n_s$  bezeichnet wird.

Weiterhin wird mit  $((t, s), d)$  das Vorkommen eines Paares  $(t, s)$  in einem Dokument  $d$  bezeichnet, wobei „Vorkommen“ in dem Sinne gemeint ist, daß der Term  $t$  im Referatetext von  $d$  vorkommt und der Deskriptor  $s$  dem Dokument intellektuell zugeteilt ist.  $((t, s), d)$  wird im folgenden auch als Paartoken bezeichnet.

$D(n)$  bezeichnet die Menge der Dokumente des Korpus<sup>7</sup>,  $D(n_s)$  die Menge der Dokumente der Stichprobe, wobei  $D(n_s)$  eine Teilmenge von  $D(n)$  ist. Für einen Term  $t$  bezeichnet  $D(t, n)$  die Menge der Dokumente von  $D(n)$ , in denen  $t$  vorkommt. Analog bezeichnet  $D(s, n)$  die Menge der Dokumente von  $D(n)$ , denen der Deskriptor  $s$  zugeteilt wurde, und  $D(t, s, n)$  die Menge der Dokumente von  $D(n)$ , in denen  $(t, s)$  vorkommt.

Weiterhin wird mit  $Z(n)$  die Menge der Term-Deskriptor-Paare, aus denen die aus  $D(n)$  gewonnene Relation  $Z$  besteht, bezeichnet:

$$Z(n) = \{(t, s) | z(t, s) \geq c_1, h(t, s) \geq c_2\}.$$

Mit  $Z(h, f, n)$  wird die Teilmenge von  $Z(n)$  bezeichnet, deren Paare  $(t, s)$  eine Paarhäufigkeit  $h(t, s) = h$  und eine Dokumenthäufigkeit  $f(t) = f$  haben:

$$Z(h, f, n) = \{(t, s) \in Z(n) | h(t, s) = h, f(t) = f\}.$$

<sup>6</sup>Im Wörterbuch PHYS/PILOT sind z.B. von insgesamt 878 674 Einträgen 805 853 aufgrund der Relation  $Z$  gebildet worden.

<sup>7</sup>Dabei wird angenommen, daß  $D(n)$  repräsentativ für das jeweilige Fachgebiet ist.

Aus der Menge  $Z(h, f, n)$  und der Stichprobe  $D(n_s)$  wird die Menge der Paartoken

$$(Z(h, f, n), D(n_s)) = \{(t, s, d) | (t, s) \in Z(h, f, n), d \in D(t, s, n) \cap D(n_s)\}$$

gebildet.

Als Erwartungswert  $E(|(Z(h, f, n), D(n_s))|)$  für die Anzahl der Paartoken einer zufällig gewählten Stichprobe  $D(n_s)$  erhält man:

$$E(|(Z(h, f, n), D(n_s))|) = |(Z(h, f, n), D(n))| \cdot \frac{n_s}{n}. \quad (1)$$

Aus (1) ergibt sich folgender Schätzwert für  $|Z(h, f, n)|$ :

$$E^*(|Z(h, f, n)|) = \frac{n}{n_s} \frac{1}{h} |(Z(h, f, n), D(n_s))|. \quad (2)$$

(2) ermöglicht Schätzungen für die Anzahl der  $z$ -Werte mit einer bestimmten Paarhäufigkeit  $h$  und Dokumenthäufigkeit  $f$  bezüglich des gesamten Korpus. Durch Summation über alle möglichen Werte von  $h$  und  $f$ , wobei nur solche  $h$ - und  $f$ -Werte betrachtet werden, die mit den Konstanten  $c_1$  und  $c_2$  die Bedingungen  $h \geq c_2$  und  $\frac{h}{f} \geq c_1$  erfüllen, erhält man einen Schätzwert für die Größe der Relation  $Z$ :

$$E^*(|Z(n)|) = \frac{n}{n_s} \sum_{h \geq c_2} \frac{1}{h} \sum_{h \leq f \leq \frac{h}{c_1}} |(Z(h, f, n), D(n_s))|. \quad (3)$$

Als Schätzwert für die Anzahl der Term-Deskriptor-Paare mit einer bestimmten Paarhäufigkeit erhält man

$$E^*(|Z(h, n)|) = \frac{n}{n_s} \frac{1}{h} \sum_{h \leq f \leq \frac{h}{c_1}} |(Z(h, f, n), D(n_s))|. \quad (4)$$

Eine systematische Herleitung der durch die Gleichungen (2) bis (4) gegebenen Schätzwerte ist in [Hüther 89] nachzulesen, während in [Tzeras 91] die Zuverlässigkeit des Schätzwertes  $E^*(|Z(n)|)$  empirisch nachgewiesen werden konnte.

### 3.2 Schätzung einfacher Maße für die Qualität der Relation $Z$

Die in [Hüther 90] hergeleiteten Erwartungswerte ermöglichen die Schätzung einfacher Maße, welche für die Beschreibung der Wörterbuchqualität verwendet werden können. In diesem Abschnitt werden die grundlegenden Erwartungswerte erläutert, die für die Definition der Qualitätsmaße in Abschnitt 3.3 benötigt werden.

Mit  $p_i$  wird im folgenden die Wahrscheinlichkeit  $P(t_i)$  bezeichnet, daß der Term  $t_i$  in einem zufällig gewählten Dokument vorkommt.  $p_{ij}$  sei die Wahrscheinlichkeit  $P(s_j | t_i)$ , daß der Deskriptor  $s_j$  einem zufällig gewählten Dokument intellektuell zugeteilt ist, unter der Voraussetzung, daß in diesem Dokument der Term  $t_i$  vorkommt.

Für die Wahrscheinlichkeit  $P((t_i, s_j) \in Z(h, f, n))$ , daß das Term-Deskriptor-Paar  $(t_i, s_j)$  in einer Menge von  $n$  Dokumenten eine Paarhäufigkeit  $h(t_i, s_j) = h$  und eine Dokumenthäufigkeit  $f(t_i) = f$  hat, erhält man:

$$P((t_i, s_j) \in Z(h, f, n)) = \binom{n}{f} p_i^f (1 - p_i)^{n-f} \binom{f}{h} p_{ij}^h (1 - p_{ij})^{f-h}.$$

Damit kann die Wahrscheinlichkeit  $P((t_i, s_j) \in Z(h \rightarrow h+1, f \rightarrow f+1, n \rightarrow n+1))$  berechnet werden, daß das Term-Deskriptor-Paar  $(t_i, s_j)$  mit  $h(t_i, s_j) = h$  und  $f(t_i) = f$  in  $n$  Dokumenten, die Parameter  $h(t_i, s_j) = h+1$  und  $f(t_i) = f+1$  in  $n+1$  Dokumenten hat:

$$P((t_i, s_j) \in Z(h \rightarrow h+1, f \rightarrow f+1, n \rightarrow n+1)) = \binom{n}{f} p_i^f (1-p_i)^{n-f} \binom{f}{h} p_{ij}^h (1-p_{ij})^{f-h} \cdot p_i \cdot p_{ij}$$

Dies kann transformiert werden in

$$P((t_i, s_j) \in Z(h \rightarrow h+1, f \rightarrow f+1, n \rightarrow n+1)) = \frac{h+1}{n+1} \binom{n+1}{f+1} p_i^{f+1} (1-p_i)^{n+1-(f+1)} \binom{f+1}{h+1} p_{ij}^{h+1} (1-p_{ij})^{f+1-(h+1)}.$$

Der Erwartungswert  $E(|Z(h, f, n)|)$  für die Anzahl der Term-Deskriptor-Paare  $(t, s)$  mit  $h(t, s) = h$  und  $f(t) = f$  in  $n$  Dokumenten ist gegeben mit

$$E(|Z(h, f, n)|) = \sum_{i \geq 1, j \geq 1} P((t_i, s_j) \in Z(h, f, n)).$$

Damit erhalten wir den Erwartungswert  $E(|Z(h \rightarrow h+1, f \rightarrow f+1, n \rightarrow n+1)|)$  für die Anzahl von Paaren  $(t, s)$  mit  $h(t, s) = h$  und  $f(t) = f$  in  $n$  Dokumenten und  $h(t, s) = h+1$  und  $f(t) = f+1$  in  $n+1$  Dokumenten:

$$E(|Z(h \rightarrow h+1, f \rightarrow f+1, n \rightarrow n+1)|) = \sum_{i \geq 1, j \geq 1} P((t_i, s_j) \in Z(h \rightarrow h+1, f \rightarrow f+1, n \rightarrow n+1)).$$

Dies ergibt folgende Gleichung:

$$E(|Z(h \rightarrow h+1, f \rightarrow f+1, n \rightarrow n+1)|) = \frac{h+1}{n+1} E(|Z(h+1, f+1, n+1)|). \quad (5)$$

Analog dazu erhalten wir den Erwartungswert  $E(|Z(h \rightarrow h, f \rightarrow f+1, n \rightarrow n+1)|)$  für die Anzahl von Paaren  $(t, s)$  mit  $h(t, s) = h$  und  $f(t) = f$  in  $n$  Dokumenten und  $h(t, s) = h$  und  $f(t) = f+1$  in  $n+1$  Dokumenten:

$$E(|Z(h \rightarrow h, f \rightarrow f+1, n \rightarrow n+1)|) = \frac{f+1-h}{n+1} E(|Z(h, f+1, n+1)|). \quad (6)$$

Ist eine  $(h, f)$ -Häufigkeitsverteilung einer aus  $n$  Dokumenten gewonnenen Relation  $Z$  gegeben, wobei  $\#(i, j, n)$  die Anzahl der Paare  $(t, s)$  mit  $h(t, s) = i$  und  $f(t) = j$  bezeichnet, so lassen sich die Erwartungswerte der rechten Seiten in (5) bzw. (6) durch die Approximationen

$$\frac{E(|Z(h+1, f+1, n+1)|)}{n+1} \approx \frac{\#(h+1, f+1, n)}{n}$$

und

$$\frac{E(|Z(h, f+1, n+1)|)}{n+1} \approx \frac{\#(h, f+1, n)}{n}$$

abschätzen, so daß sich folgende Gleichungen ergeben:

$$E(|Z(h \rightarrow h+1, f \rightarrow f+1, n \rightarrow n+1)|) = \frac{h+1}{n} \#(h+1, f+1, n) \quad (7)$$

und

$$E(|Z(h \rightarrow h, f \rightarrow f + 1, n \rightarrow n + 1)|) = \frac{f + 1 - h}{n} \#(h, f + 1, n). \quad (8)$$

Gleichung (7) liefert einen Erwartungswert für die Anzahl der Paare  $(t, s)$ , die im  $n+1$ -ten Dokument gebildet werden<sup>8</sup> und eine bestimmte Paar- und Dokumenthäufigkeit besitzen.

Gleichung (8) liefert einen Erwartungswert für die Anzahl der Paare  $(t, s)$  mit einer bestimmten Paar- und Dokumenthäufigkeit, wobei Term  $t$  im  $n + 1$ -ten Dokument vorkommt, ohne daß der Deskriptor  $s$  diesem Dokument intellektuell zugeteilt ist.

### 3.3 Definition der Qualitätsmaße

Als ein Vollständigkeitsmaß der Relation  $Z$  bietet sich die erwartete Anzahl der in einem aus  $n$  Dokumenten gewonnenen Wörterbuch erfaßten Paare  $(t, s)$  an, deren Term  $t$  im  $n+1$ -ten Dokument vorkommt und deren Deskriptor  $s$  diesem Dokument intellektuell zugeteilt ist.

Dieses Maß kann als die Anzahl der richtig gebildeten Hinweise bei einem neu zu indexierenden Dokument interpretiert werden. Dabei ist „richtig“ in dem Sinne gemeint, daß von Terms des Textes auf solche Deskriptoren hingewiesen wird, die auch der Indexierungsfachmann dem Dokument zuteilen würde.

Ein Wörterbuch ist somit umso vollständiger, je größer die erwartete Anzahl der richtig gebildeten Hinweise bei einem neu zu indexierenden Dokument ist. Mit Hilfe des Erwartungswertes  $E(|Z(h \rightarrow h + 1, f \rightarrow f + 1, n \rightarrow n + 1)|)$  aus der Gleichung (7) kann dieses Vollständigkeitsmaß folgendermaßen definiert werden:

$$E(\#RICHTIG) = \sum_{(h \geq c_2) \wedge ((h/f) \geq c_1)} E(|Z(h \rightarrow h + 1, f \rightarrow f + 1, n \rightarrow n + 1)|). \quad (9)$$

Aus (9) und (7) ergibt sich:

$$E(\#RICHTIG) = \frac{1}{n} \sum_{(h \geq c_2) \wedge ((h/f) \geq c_1)} (h + 1) \#(h + 1, f + 1, n). \quad (10)$$

Analog zu  $E(\#RICHTIG)$  kann ein Erwartungswert für die Anzahl der falsch gebildeten Hinweise definiert werden. Das sind Hinweise von Terms eines neu zu indexierenden Textes auf Deskriptoren, die der Indexierungsfachmann dem Dokument nicht zuteilen würde:

$$E(\#FALSCH) = \sum_{(h \geq c_2) \wedge ((h/f) \geq c_1)} E(|Z(h, f \rightarrow f + 1, n \rightarrow n + 1)|). \quad (11)$$

Aus (11) und (8) ergibt sich:

$$E(\#FALSCH) = \frac{1}{n} \sum_{(h \geq c_2) \wedge ((h/f) \geq c_1)} (f + 1 - h) \#(h, f + 1, n). \quad (12)$$

Die zusätzliche Berücksichtigung von  $E(\#FALSCH)$  ist bei der Beurteilung der Wörterbuchqualität erforderlich; je kleiner die erwartete Anzahl der falsch gebildeten Hinweise bei einem neu zu indexierenden Dokument ist, desto genauer ist das Wörterbuch.

Das Verhältnis zwischen richtig und falsch gebildeten Hinweisen kann für die Schätzung von unteren bzw. oberen Grenzen der erreichbaren Indexierungsqualität ausgenutzt werden, wenn diese als Übereinstimmung mit der intellektuellen Indexierung aufgefaßt und durch die folgenden Maße dargestellt wird:

<sup>8</sup>d.h. der Term  $t$  kommt im Dokument vor und der Deskriptor  $s$  ist dem Dokument intellektuell zugeteilt.

Für eine Menge von Dokumenten bezeichne  $AUT$  die Menge der automatischen und  $INT$  die Menge der intellektuellen Deskriptorzuteilungen. Damit lassen sich die Maße

$$p = \frac{|AUT \cap INT|}{|AUT|} \quad \text{und} \quad r = \frac{|AUT \cap INT|}{|INT|}$$

definieren. Zur Abschätzung der erreichbaren Indexierungsqualität betrachten wir eine elementare Indexierungsfunktion  $a_1$ , bei der ein Deskriptor  $s$  genau dann einem Dokument zugeteilt wird, wenn mindestens ein Hinweis vom Text des Dokuments auf  $s$  gebildet wird.

Mit Hilfe von  $E(\#RICHTIG)$  und  $E(\#FALSCH)$  lassen sich in diesem Falle für die den Maßen entsprechenden Werte  $p_{a_1}$  bzw.  $r_{a_1}$  folgende untere und obere Grenzen definieren, wobei  $E(\#DESK)$  die erwartete Anzahl intellektuell zugeteilter Deskriptoren pro Dokument (Indexierungstiefe) bezeichnet:

$$p_{a_1} = \frac{\min(1, E(\#RICHTIG))}{\min(1, E(\#RICHTIG)) + E(\#FALSCH)}$$

und

$$\bar{p}_{a_1} = \frac{\min(E(\#RICHTIG), E(\#DESK))}{\min(E(\#RICHTIG), E(\#DESK)) + \min(1, E(\#FALSCH))}$$

Analog dazu gilt für  $r_{a_1}$ :

$$r_{a_1} = \frac{\min(1, E(\#RICHTIG))}{E(\#DESK)} \quad \text{und} \quad \bar{r}_{a_1} = \frac{\min(E(\#RICHTIG), E(\#DESK))}{E(\#DESK)}$$

Die Maße  $p_{a_1}$  und  $\bar{r}_{a_1}$  sind unabhängig von einer bestimmten Indexierungsfunktion zu interpretieren und ermöglichen somit eine generelle Beurteilung der Entscheidungsgrundlage für die Indexierungsfunktion. Für die Werte  $p_{a_1}$  bzw.  $r_{a_1}$  als auch für die Werte  $p_a$  bzw.  $r_a$  nach der Anwendung einer beliebigen Indexierungsfunktion  $a$  gilt nämlich:

$$p_{a_1} \leq p_a \leq p_a \quad \text{und} \quad \bar{r}_{a_1} \geq r_a \geq r_a$$

## 4 Experimentelle Ergebnisse

### 4.1 Aufbau der Experimente

Für die experimentelle Untersuchung der Beziehung zwischen Entwicklungsaufwand, Größe und Vollständigkeit eines Indexierungswörterbuches wurden bei jeweils unterschiedlichen Parametern der Wörterbuchgewinnung die durch die Gleichungen (3) und (10) gegebenen Schätzwerte  $E^*(|Z(n)|)$  und  $E(\#RICHTIG)$  berechnet. Außerdem wurde stets der durch die Gleichung (12) gegebene Schätzwert  $E(\#FALSCH)$  berechnet, um den Effekt der falsch gebildeten Hinweise bei der jeweiligen Vollständigkeit mitzuberechnen. Weiterhin wurden als eine erste Annäherung an die Indexierungsqualität die Werte  $p_{a_1}$  und  $\bar{r}_{a_1}$  für einige interessante Wörterbuchvarianten ermittelt.

Um möglichst aussagekräftige Ergebnisse zu erhalten, wurden insgesamt 48 Wörterbuchvarianten untersucht. Eine Wörterbuchvariante ist durch die Korpusgröße (98 000, 196 000, 294 000 oder 392 000 Dokumente) und die Schwellenwerte  $c_1$  (0.1, 0.2, 0.3, 0.4, oder 0.5) und  $c_2$  (3, 4, 5, 6 oder 7) festgelegt. Für die Berechnung der Schätzwerte standen 392 000 intellektuell indizierte Dokumente aus der Produktion des Wörterbuches PHYS/PILOT zur Verfügung. Die Schätzwerte konnten daher teilweise durch Vergleich mit Werten aus PHYS/PILOT evaluiert werden.

Die durch Gleichung (3) gegebenen Schätzwerte wurden aufgrund von Stichproben mit 1600 Elementen berechnet. Die zur Berechnung von  $E(\#RICHTIG)$  bzw.  $E(\#FALSCH)$  benötigten Anzahlen  $\#(h+1, f+1, n)$  bzw.  $\#(h, f+1, n)$  wurden durch die Schätzwerte  $E^*(|Z(h+1, f+1, n)|)$

bzw.  $E^*(|Z(h, f + 1, n)|)$  gemäß Gleichung (2) angenähert. Weiterhin wurde die zur Berechnung von  $\bar{r}_{a_1}$  benötigte Indexierungstiefe  $E(\#DESK)$  durch die mittlere Anzahl intellektuell zugeteilter Deskriptoren aus den obengenannten 392 000 Dokumenten<sup>9</sup> angenähert.

Es wurden ausschließlich Schätzungen bezüglich einer Relation  $Z$  mit einem Einzelwort in Grundform als Term durchgeführt. Auf Schätzungen mit Mehrwortgruppen als Terms wurde vorerst verzichtet.

## 4.2 Beziehung zwischen Wörterbuchgröße und Vollständigkeit

In Tabelle 1 wird die Wörterbuchgröße und Vollständigkeit für verschiedene Korpusgrößen und Aufnahmekriterien  $c_1$  und  $c_2$  dargestellt. Die Einträge sind für jede Korpusgröße nach fallender Wörterbuchgröße angeordnet.

Die Größe des Wörterbuches wird dabei durch die erwartete Anzahl der darin erfaßten Term-Deskriptor-Paare  $(t, s)$  ausgedrückt, während die Vollständigkeit durch die erwartete Anzahl der richtig gebildeten Hinweise bei einem neu zu indexierenden Dokument beschrieben wird.

Tabelle 1: Beziehung zwischen Entwicklungsaufwand, Wörterbuchgröße und Vollständigkeit

Korpusgröße	Größe	$c_1 = 0.1$ $c_2 = 3$	Größe	$c_1 = 0.1$ $c_2 = 4$	Größe	$c_1 = 0.2$ $c_2 = 3$	Größe	$c_1 = 0.1$ $c_2 = 5$
98 000	122 390	13.09	74 514	12.67	64 877	6.51	53 869	11.93
196 000	211 066	13.15	130 065	12.54	120 928	6.69	93 451	12.08
294 000	310 715	13.78	195 844	13.12	179 065	6.78	138 608	12.68
392 000	398 550	14.20	250 370	13.59	229 023	7.03	179 742	13.15

Korpusgröße	Größe	$c_1 = 0.3$ $c_2 = 3$	Größe	$c_1 = 0.1$ $c_2 = 6$	Größe	$c_1 = 0.2$ $c_2 = 4$	Größe	$c_1 = 0.1$ $c_2 = 7$
98 000	43 043	4.23	42 356	11.53	36 585	6.16	34 720	11.23
196 000	80 445	4.37	72 772	11.74	69 349	6.35	60 076	11.46
294 000	119 750	4.33	107 908	12.33	107 541	6.38	88 238	12.04
392 000	153 984	4.57	140 199	12.81	137 287	6.69	115 781	12.53

Korpusgröße	Größe	$c_1 = 0.4$ $c_2 = 3$	Größe	$c_1 = 0.2$ $c_2 = 5$	Größe	$c_1 = 0.5$ $c_2 = 3$	Größe	$c_1 = 0.3$ $c_2 = 4$
98 000	30 716	2.63	25 507	5.91	23 966	1.86	22 978	3.99
196 000	58 267	2.78	49 155	6.09	45 986	2.00	44 627	4.14
294 000	83 210	2.77	71 779	6.13	62 458	1.95	68 590	4.07
392 000	110 805	2.93	96 722	6.44	88 561	2.06	89 209	4.35

Man erkennt, daß mit steigender Korpusgröße die Wörterbuchgröße fast linear mitwächst, während die Vollständigkeit in etwa konstant bleibt. Dies ist ein interessantes Ergebnis, zumal auch die Anzahl der falsch gebildeten Hinweise sich mit steigender Korpusgröße nicht wesentlich verändert, wie aus Tabelle 2 hervorgeht.

Eine Erhöhung des  $c_2$ -Wertes bei gleichbleibendem  $c_1$ -Wert bewirkt ebenfalls eine deutliche Verringerung der Wörterbuchgröße, während die Vollständigkeit dabei nur geringfügig abnimmt (siehe z.B.

<sup>9</sup>8.72 Deskriptoren pro Dokument.

Parameterkombinationen  $c_2 = 3, 4, 5, 6, 7$  bei  $c_1 = 0.1$ ). Die Anzahl der falsch gebildeten Hinweise aus Tabelle 2 geht bei wachsendem  $c_2$ -Wert auch nur leicht zurück.

Tabelle 2: Erwartete Anzahl falsch gebildeter Hinweise

Korpus	Größe	$c_1 = 0.1$ $c_2 = 3$	Größe	$c_1 = 0.1$ $c_2 = 4$	Größe	$c_1 = 0.2$ $c_2 = 3$	Größe	$c_1 = 0.1$ $c_2 = 5$
98 000	122 390	64.18	74 514	58.68	64 877	14.93	53 869	55.18
196 000	211 066	61.32	130 065	56.86	120 928	14.65	93 451	53.87
294 000	310 715	65.40	195 844	60.91	179 065	15.10	138 608	58.01
392 000	398 550	66.03	250 370	61.90	229 023	15.20	179 742	59.11

Korpus	Größe	$c_1 = 0.3$ $c_2 = 3$	Größe	$c_1 = 0.1$ $c_2 = 6$	Größe	$c_1 = 0.2$ $c_2 = 4$	Größe	$c_1 = 0.1$ $c_2 = 7$
98 000	43 043	6.24	42 356	52.69	36 585	13.24	34 720	50.57
196 000	80 445	6.11	72 772	51.81	69 349	13.25	60 076	50.13
294 000	119 750	6.21	107 908	55.89	107 541	13.71	88 238	54.16
392 000	153 984	6.29	140 199	57.19	137 287	13.97	115 781	55.80

Korpus	Größe	$c_1 = 0.4$ $c_2 = 3$	Größe	$c_1 = 0.2$ $c_2 = 5$	Größe	$c_1 = 0.5$ $c_2 = 3$	Größe	$c_1 = 0.3$ $c_2 = 4$
98 000	30 716	2.69	25 507	12.33	23 966	1.49	22 978	5.49
196 000	58 267	2.64	49 155	12.46	45 986	1.45	44 627	5.48
294 000	83 210	2.71	71 779	12.78	62 458	1.45	68 590	5.53
392 000	110 805	2.74	96 722	13.19	88 561	1.48	89 209	5.74

Mit steigendem  $c_1$ -Wert nimmt bei einem festen  $c_2$ -Wert sowohl die Wörterbuchgröße als auch die Vollständigkeit deutlich ab (siehe z.B. Parameterkombinationen  $c_1 = 0.1, 0.2, 0.3, 0.4, 0.5$  bei  $c_2 = 3$ ). Die Verringerung der Wörterbuchgröße ist bei einer Erhöhung des  $c_1$ -Wertes stärker als bei einer Erhöhung des  $c_2$ -Wertes.

Diese Ergebnisse deuten auf die Wichtigkeit des Parameters  $c_1$  bezüglich der Vollständigkeit eines Wörterbuches hin. Die Vollständigkeit wird weitgehend durch die Wahl von  $c_1$  bestimmt, während die Parameter  $c_2$  und Korpusgröße eine untergeordnete Rolle spielen.

Bei allen vier untersuchten Korpusgrößen ergibt die Sortierung nach fallenden Wörterbuchgrößen die gleiche Reihenfolge der Parameterkombinationen für die Aufnahmekriterien  $c_1$  und  $c_2$ . Diese Parameterkombinationen können als unterschiedlich strenge Kriterien interpretiert werden, welche bei gegebener Korpusgröße die Wörterbuchgröße bestimmen.

Für eine beliebige Korpusgröße hat ein Wörterbuch mit den Parametern  $c_1 = 0.1, c_2 = 3$  den größten Umfang, gefolgt von einem Wörterbuch mit  $c_1 = 0.1, c_2 = 4$ , gefolgt von einem mit  $c_1 = 0.2, c_2 = 3$ , usw.

Wenn also ein Wörterbuch mit  $c_1 = 0.1, c_2 = 3$  verringert werden soll, so kommt als nächster Parameter  $c_1 = 0.1, c_2 = 4$  in Frage. Das Wörterbuch schrumpft dabei erheblich, während die Anzahl richtiger Hinweise fast konstant bleibt. Mit dem nächststrengeren Kriterium  $c_1 = 0.2, c_2 = 3$  schrumpft das Wörterbuch weiter, die Anzahl richtiger Hinweise fällt jedoch auf die Hälfte, während die Anzahl falscher Hinweise auf fast ein viertel zurückgeht. Dies sollte in der Praxis berücksichtigt

werden, denn mit dem nächststrengeren Kriterium  $c_1 = 0.1, c_2 = 5$  schrumpft das Wörterbuch weiter, während die Anzahl richtiger und falscher Hinweise wieder fast auf den alten Wert steigt. Es ist also sorgfältig zu überlegen, wie man die Parameter  $c_1$  und  $c_2$  wählt, um bei einer vorgegebenen Wörterbuchgröße ein Wörterbuch zu erhalten, welches die erwünschte Vollständigkeit besitzt und eine möglichst kleine Anzahl von falschen Hinweisen liefert.

Die Tabellen 1 und 2 geben nützliche Informationen für die Optimierung eines Wörterbuches. Wollte man z.B. anstelle des durch die Parameterkombination  $c_1 = 0.3, c_2 = 3$  bei einer Korpusgröße von 392 000 Dokumenten gegebenen Wörterbuches mit einer Größe von 153 984 nur ein halb so großes Wörterbuch produzieren, so wäre die sicherste Alternative eine Halbierung der Korpusgröße und eine Beibehaltung von  $c_1 = 0.3, c_2 = 3$ . Das somit erhaltene Wörterbuch wäre fast genauso vollständig, ohne daß es wesentlich mehr falsche Hinweise liefern würde.

### 4.3 Grenzen für die Indexierungsqualität

In Tabelle 3 werden für einige praxisrelevante Wörterbuchvarianten erste Annäherungen für die erreichbare Indexierungsqualität angegeben. Die Einträge können so interpretiert werden, daß die Werte  $p$  bzw.  $r$  unabhängig von der jeweils verwendeten Indexierungsfunktion im schlechtesten bzw. im günstigsten Fall  $\underline{p}_{a_1}$  bzw.  $\bar{r}_{a_1}$  betragen werden.

Tabelle 3: Untere und obere Grenzen der Indexierungsqualität

Korpus	$c_1 = 0.2, c_2 = 3$		$c_1 = 0.2, c_2 = 4$		$c_1 = 0.2, c_2 = 5$	
	$\underline{p}_{a_1}$	$\bar{r}_{a_1}$	$\underline{p}_{a_1}$	$\bar{r}_{a_1}$	$\underline{p}_{a_1}$	$\bar{r}_{a_1}$
98 000	0.0628	0.7471	0.0702	0.7062	0.0750	0.6781
196 000	0.0639	0.7679	0.0702	0.7280	0.0743	0.6989
294 000	0.0621	0.7773	0.0680	0.7319	0.0725	0.7027
392 000	0.0617	0.8063	0.0668	0.7674	0.0705	0.7384

Korpus	$c_1 = 0.3, c_2 = 3$		$c_1 = 0.3, c_2 = 4$		$c_1 = 0.3, c_2 = 5$	
	$\underline{p}_{a_1}$	$\bar{r}_{a_1}$	$\underline{p}_{a_1}$	$\bar{r}_{a_1}$	$\underline{p}_{a_1}$	$\bar{r}_{a_1}$
98 000	0.1382	0.4857	0.1540	0.4571	0.1629	0.4388
196 000	0.1406	0.5017	0.1542	0.4751	0.1630	0.4559
294 000	0.1387	0.4967	0.1532	0.4673	0.1634	0.4491
392 000	0.1373	0.5247	0.1483	0.4990	0.1559	0.4786

Korpus	$c_1 = 0.4, c_2 = 3$		$c_1 = 0.4, c_2 = 4$		$c_1 = 0.4, c_2 = 5$	
	$\underline{p}_{a_1}$	$\bar{r}_{a_1}$	$\underline{p}_{a_1}$	$\bar{r}_{a_1}$	$\underline{p}_{a_1}$	$\bar{r}_{a_1}$
98 000	0.2710	0.3021	0.2990	0.2851	0.3199	0.2713
196 000	0.2749	0.3189	0.3016	0.3019	0.3208	0.2875
294 000	0.2698	0.3176	0.2973	0.2975	0.3169	0.2849
392 000	0.2676	0.3366	0.2886	0.3176	0.3022	0.3039

Bei steigender Wörterbuchgröße erhöht sich die obere Grenze  $\bar{r}_{a_1}$  für den Wert  $r$ , während sich die



untere Grenze  $p_{a_1}$  für den Wert  $p$  verringert.

Eine Erhöhung des  $c_2$ -Wertes bei gleichbleibendem  $c_1$ -Wert bewirkt eine geringfügige Erhöhung von  $p_{a_1}$  und eine Verringerung von  $r_{a_1}$ .

Der  $c_1$ -Wert bestimmt weitgehend die Größenordnung von  $p_{a_1}$  und  $r_{a_1}$ . Eine Erhöhung des  $c_1$ -Wertes von 0.2 auf 0.3 bzw. von 0.3 auf 0.4 bewirkt eine Verdopplung von  $p_{a_1}$  und eine deutliche Verringerung von  $r_{a_1}$ .

Tabelle 4: Werte  $p_{a_1}$  und  $r_{a_1}$  für eine mit bereits entwickelten Wörterbüchern indizierte Stichprobe von 1000 Dokumenten

Korpus	$c_1 = 0.3, c_2 = 3$		$c_1 = 0.3, c_2 = 4$		$c_1 = 0.3, c_2 = 5$	
	$p_{a_1}$	$r_{a_1}$	$p_{a_1}$	$r_{a_1}$	$p_{a_1}$	$r_{a_1}$
392 000	0.3587	0.3509	0.3802	0.3394	0.3900	0.3323

Korpus	$c_1 = 0.4, c_2 = 3$		$c_1 = 0.4, c_2 = 4$		$c_1 = 0.4, c_2 = 5$	
	$p_{a_1}$	$r_{a_1}$	$p_{a_1}$	$r_{a_1}$	$p_{a_1}$	$r_{a_1}$
392 000	0.4574	0.2508	0.4886	0.2401	0.5048	0.2332

Die Grenzwerte aus Tabelle 3 liefern wichtige Informationen für die Verbesserung der automatischen Indexierung. So können z.B. bestimmte Wörterbuchvarianten aufgrund der mit bereits entwickelten Wörterbüchern erreichten Indexierungsqualität zukünftig von vornherein ausgeschlossen werden, andere jedoch als interessant für eine nähere Untersuchung angesehen werden.

Allerdings liefern die in Tabelle 3 dargestellten Grenzwerte nur grobe Schätzungen im Vergleich zu den in Tabelle 4 aufgelisteten Grenzwerten  $p_{a_1}$  und  $r_{a_1}$ . Letztere müssen jedoch bisher aufgrund der Indexierung einer Dokumentstichprobe berechnet werden, die wiederum die Existenz eines geeigneten Wörterbuches voraussetzt.

#### 4.4 Zuverlässigkeit der Schätzwerte

Um eine globale Aussage über die Zuverlässigkeit der Schätzungen zu machen, vergleichen wir in Tabelle 5 die Anzahl der  $Z$ -Relationen aus den Schätzungen mit den korrekten Werten aus dem Wörterbuch PHYS/PILOT, das aus einem Korpus von 392 000 Dokumenten gewonnen wurde. Angegeben sind die genaue und die geschätzte Anzahl der  $Z$ -Relationen mit dem Parameter  $c_1 = 0.3$ , sowie die prozentuale Abweichung des genauen Wertes vom jeweiligen Schätzwert.

Um eine Aussage über die Zuverlässigkeit der Schätzwerte  $E(\#RICHTIG)$  und  $E(\#FALSCH)$  zu erhalten, werden diese in Tabelle 6 mit empirischen Werten  $\#RICHTIG$  und  $\#FALSCH$  verglichen, die sich bei der Indexierung einer Stichprobe von 1000 intellektuell indizierten Dokumente aufgrund verschiedener Wörterbücher ergeben haben.

Die Abweichungen der betrachteten Schätzwerte liegen bei keiner der betrachteten Schätzungen über die 12%-Grenze, sodaß die Zuverlässigkeit der Schätzwerte zunächst als vertretbar beurteilt werden kann. Weitere Vergleiche mit empirischen Werten, die hier aus Platzgründen nicht dargestellt werden können, deuten ebenfalls auf akzeptable Schätzwerte hin.

Tabelle 5: Genauigkeit der Schätzwerte  $E^*(|Z(n)|)$

	$ Z(n) $	$E^*( Z(n) )$	Abweichung
$c_2 = 3$	153984	159930	-3.7%
$c_2 = 4$	89209	88197	+1.1%
$c_2 = 5$	61771	60097	+2.8%
$c_2 = 6$	45755	45814	-0.1%
$c_2 = 7$	36433	36634	-0.5%

Tabelle 6: Genauigkeit von  $E(\#RICHTIG)$  und  $E(\#FALSCH)$

n=392 000	$E(\#RICHTIG)$	$\#RICHTIG$	Abweichung
$c_1 = 0.3, c_2 = 3$	4.574	4.727	-3.34%
$c_1 = 0.3, c_2 = 4$	4.350	4.481	-3.01%
$c_1 = 0.3, c_2 = 5$	4.172	4.312	-3.36%
$c_1 = 0.3, c_2 = 6$	4.051	4.195	-3.55%
$c_1 = 0.3, c_2 = 7$	3.947	4.092	-3.67%

n=392 000	$E(\#FALSCH)$	$\#FALSCH$	Abweichung
$c_1 = 0.3, c_2 = 3$	6.285	7.009	-11.52%
$c_1 = 0.3, c_2 = 4$	5.743	6.165	-7.34%
$c_1 = 0.3, c_2 = 5$	5.414	5.830	-7.68%
$c_1 = 0.3, c_2 = 6$	5.170	5.532	-7.02%
$c_1 = 0.3, c_2 = 7$	4.995	5.342	-6.94%

## 5 Ausblick

Mit den Schätzwerten  $E^*(|Z(n)|)$ ,  $E(\#RICHTIG)$ ,  $E(\#FALSCH)$  konnten Aussagen über die Vollständigkeit von 48 Wörterbüchern gemacht werden, ohne diese real aufzubauen. Die verwendeten Maße sind einfach und ermöglichen nur eine grobe Abschätzung der Vollständigkeit eines Wörterbuches. Für eine genauere Beurteilung der Vollständigkeit wären auch andere Maße denkbar, wie z.B.  $r_{a_1}$ . Bis jetzt wurden dafür noch keine Schätzverfahren entwickelt.

Die Grenzwerte  $\underline{p}_{a_1}$  und  $\bar{r}_{a_1}$  für die Indexierungsqualität können als eine erste Annäherung für  $p_{a_1}$  und  $r_{a_1}$  verwendet werden. Der Vergleich der Schätzwerte mit empirischen Werten deutet auf eine vertretbare Genauigkeit der Schätzungen hin und unterstützt somit die Ergebnisse.

Für die weitere Untersuchung der Beziehung zwischen Entwicklungsaufwand und Leistungsfähigkeit eines Wörterbuches für die automatische Indexierung wäre ein Schätzverfahren für  $p_{a_1}$  und  $r_{a_1}$  in Abhängigkeit von den Parametern beim Wörterbuchaufbau sehr wichtig. Man könnte dabei experimentell untersuchen, wie sich die Hinweise auf Relevanzbeschreibungen verteilen und einen Erwartungswert für  $p_{a_1}$  und  $r_{a_1}$  in Abhängigkeit von  $E(\#RICHTIG)$  und  $E(\#FALSCH)$  herleiten. Weiterhin könnte man  $p_{a_1}$  und  $r_{a_1}$  direkt aufgrund von Häufigkeitsverteilungen der Paare  $(t, s)$

analog zu den in Abschnitt 3.2 hergeleiteten Schätzwerten abschätzen.

## Literatur

- Beinke-Geiser, U.; Lustig, G.; Putze-Meier, G. (1986). Indexieren mit dem System DAISY. In: Lustig, G. (ed.): *Automatische Indexierung zwischen Forschung und Anwendung*, S. 73–97. Olms, Hildesheim.
- Faißt, S. (1990). *Entwicklung von Indexierungsfunktionen auf der Basis probabilistischer Entscheidungsbäume*. Diplomarbeit, TH Darmstadt, FB Informatik, Datenverwaltungssysteme II. (In preparation).
- Fuhr, N.; Hartmann, S.; Knorz, G.; Lustig, G.; Schwantner, M.; Tzeras, K. (1991). AIR/X - a Rule-Based Multistage Indexing System for Large Subject Fields. In: *Proceedings of the RIAO'91, Barcelona, Spain, April 2-5, 1991*.
- Hüther, H. (1989). *Wachstumsfunktionen in der automatischen Indexierung*. Dissertation, TH Darmstadt, Fachbereich Informatik.
- Hüther, H. (1990). On the Interrelationship of Dictionary Size and Completeness. In: Vidick, J.-L. (ed.): *Proceedings of the 13th International Conference on Research and Development in Information Retrieval*, S. 313–326. ACM, New York.
- Knorz, G. (1983). *Automatisches Indexieren als Erkennen abstrakter Objekte*. Niemeyer, Tübingen.
- Lustig, G. (1969). Die automatische Zuteilung von Schlagwörtern des EURATOM-Thesaurus. *Neue Technik 11*, S. 247–256.
- Lustig, G. (ed.) (1986a). *Automatische Indexierung zwischen Forschung und Anwendung*. Olms, Hildesheim.
- Lustig, G. (1986b). Eine anwendungsorientierte Konzeption der automatischen Indexierung. In: Lustig, G. (ed.): *Automatische Indexierung zwischen Forschung und Anwendung*, S. 1–12. Olms, Hildesheim.
- Pfeifer, U. (1990). *Entwicklung log-linearer und linear-iterativer Indexierungsfunktionen*. Diplomarbeit, TH Darmstadt, FB Informatik, Datenverwaltungssysteme II.
- Schwantner, M. (1991) *Aufbau und Pflege eines Wörterbuches für die Automatische Indexierung*. Dissertation, TH Darmstadt, Fachbereich Informatik.
- Tzeras, K. (1991). *Zur Aufwandsabschätzung bei der Entwicklung eines Indexierungswörterbuches*. Erscheint im Tagungsband des „Workshop Information Retrieval“ (Darmstadt, 24. und 25. Juni 1991, GMD-IPSI).

Dipl. Wirtsch. Inf. Tzeras, Konstantinos,  
Cand. Inf. Reiss, Peter  
Technische Hochschule Darmstadt, Fachbereich Informatik  
Fachgebiet Datenverwaltungssysteme II  
Alexanderstr. 6,  
W-6100 Darmstadt,  
Germany  
E-Mail: kostas@dvs2.informatik.th-darmstadt.de

# HyperAdvokat

## Konzeption und Realisierung eines juristischen Hypertextsystems

Michel, Engelbert

### Abstract:

Der Prototyp eines juristischen Informationssystems wird vorgestellt. Zunächst wird die Eignung von juristischen Kommentaren für die Darstellung als Hypertext untersucht. Hypertext ermöglicht die Integration verschiedener Dokumentarten und ist geeignet, den Informationsbedarf von Juristen zu decken. Danach werden die wichtigsten Eigenschaften und Funktionen von HyperAdvokat beschrieben und die Voraussetzungen für die Realisierung umfangreicher Informationssysteme untersucht.

The prototyp of a law information system is presented. First the suitability of law commentaries for hypertext is examined. Hypertext allows the integration of different document types. It is a tool to satisfy the jurist's need for information. Then the important properties and functions from HyperAdvokat are presented and the suitability for the realization of big systems is shown.

### 1. Eignung juristischer Texte für Hypertext

Traditionell erfolgt die Erarbeitung, Speicherung und Darstellung von Wissen in Form von linearen Texten wie Artikeln, Aufsätzen Büchern usw. Die modernen Informations- und Kommunikationstechnologien eröffnen mit Hypertext eine neue Form der Textverarbeitung, die die lineare Beschränkung erkömmlicher Medien aufhebt, und multimediale Erweiterungsmöglichkeiten bietet. Hypertext ist ein Medium für die nicht-lineare Darstellung von Wissen und die Erarbeitung von Information. "Die Grundidee von Hypertext besteht darin, daß informationelle Einheiten, in denen Objekte und Vorgänge des gewählten Weltausschnittes auf textuelle, graphische oder audiovisuelle Weise dargestellt werden können, flexibel über Verknüpfungen manipuliert werden können. Die Manipulation besteht im wesentlichen darin, daß informationelle Einheiten oder Teile in ihnen als Ausgangspunkt einer Verknüpfung direkt aktiviert werden können, z.B. dadurch, daß über einen "Mouse"-Klick eine attraktive, d.h. weitere Informationen stimulierende Stelle (ein "button") angewählt wird. Das System wird der aktivierten Verknüpfung nachgehen und sollte sofort den Endpunkt der Verknüpfung, eine andere informationelle Einheit oder ein bestimmter Punkt in ihr am Bildschirm anzeigen, entweder dadurch, daß die neue Einheit in die alte eingeblendet wird oder diese ersetzt oder beide parallel in unterschiedlichen Fenstern angezeigt werden."<sup>1</sup>

Hypertext ist keine grundsätzlich neue Idee, sondern erhält neuen Auftrieb durch die Entwicklung leistungsfähiger Computersysteme, die erstmals die Möglichkeiten bieten, Hypertext in befriedigender Weise zu realisieren. Auch traditionelle lineare Medien wie Bücher enthalten viele nicht-lineare Strukturen wie Verweise, Inhaltsverzeichnisse, Glossarien etc. Es gibt weiterhin viele, zum Teil schon sehr alte Dokumente, wie z.B. das Talmud oder die Schriften Aristoteles, die als manuelle Hypertexte

bezeichnet werden könnten. In ihnen machen die Referenzierungen zu anderen Teilen des Dokumentes oder zu anderen Dokumenten ein Hauptteil des Werkes aus.<sup>2</sup> Dies gilt besonders für die Textsorte, mit der wir uns im folgenden beschäftigen.

Juristische Kommentare sind als hochkomprimierte Wissensstrukturen mit streng paragrafenorientierter Gliederung und Einteilung in kohärente und kohäsiv geschlossene Blöcke gut für die Darstellung als Hypertext geeignet.

Beispiel § 535 ordnet sich wie folgt in die Hierarchie des Bürgerlichen Gesetzbuches ein:

1. Buch: Recht der Schuldverhältnisse
2. Abschnitt. Einzelne Schuldverhältnisse
3. Titel: Miete. Pacht

Der Gesetzestext hat folgenden Wortlaut:

**§ 535 Vertragliche Hauptpflichten.** Durch den Mietvertrag wird der Vermieter verpflichtet, dem Mieter den Gebrauch der vermieteten Sache während der Mietzeit zu gewähren. Der Mieter ist verpflichtet, dem Vermieter den vereinbarten Mietzins zu entrichten.

Der Kommentar, hier die Randnummer 1, hat folgenden Wortlaut (die Verweise wurden vom Autor durch Fettdruck hervorgehoben):

1) Mietvertrag (Begründung und Rechtsnatur: **Einführung 1 vor § 535**. a) Abschluß erfolgt gemäß §§ 145 ff. Erforderlich ist Einigung über Mietgegenstand, -zeit und -preis, sowie darüber, daß die Überlassung zum Gebrauch erfolgt; nicht notwendig ist in der Regel Einigung über den konkreten Umfang des Gebrauchs (herrschende Meinung; vgl **Leenen MDR 80, 353**). Nur einseitige Gebrauchsanmaßung ohne Zustimmung des Vermieters ist kein Mietvertrag. Zusage im Rahmen von Vertragsverhandlungen kann Haftung aus c.i.c. (vgl **§ 276 Anm 6**) begründen (**LG Mannheim ZMR 71, 133**), falsche Selbstauskunft des Mieters die Anfechtung (hierzu **Hille WoM 84, 292 mwN**; **Schmid DWW 85, 302**). Form des Mietvertrages: Grundsätzlich frei; Ausnahme: § 566. Stillschweigender Abschluß eines Mietvertrages ist möglich, wenn eine Gemesch (§ 741) oder eine GesHand einem ihrer Mitglieder eine Sache, insbesondere Räume, gegen wiederkehrendes Entgelt überläßt (**BGH WM 69, 298**; vgl. aber **Einführung 2 f**). Eintritt des Ehegatten oder anderer Familienangehöriger des verstorbenen Mieters in den Mietvertrag ist nach § 596 a möglich. Begründung eines Mietvertrages durch richterliche Gestaltung ist bei Ehescheidung möglich (§ 5 **HausratsVO**). Besteht zwischen den Vertragsparteien Einverständnis mit vorzeitigem Einzug des Mieters, so ist ein vereinbarter vorzeitiger Vertragsbeginn (§ 305) mit allen Rechten und Pflichten anzunehmen.

Allein in dieser Randnummer befinden sich 14 Verweise auf andere Textstellen, davon

5 Verweise auf andere §§ des BGB;

3 Verweise auf Kommentarstellen des BGB;

3 Verweise auf Literatur;

2 Verweise auf Gerichtsurteile;

1 Verweis auf eine andere Gesetzesnorm.

Die Möglichkeit von Hypertext, diese Verweise als Ausgangspunkte von Verknüpfungen zu verwenden und die dazugehörigen externen Dokumente in die Hypertext-Basis einzubinden oder über externe Links zu integrieren, stellt einen hohen informationellen Mehrwert dar.

Ein Problem bei der Konvertierung besteht in der Notwendigkeit, die Formulierungen juristischer Texte unverändert zu übernehmen. Juristische Texte, insbesondere Kommentare, sind hochkomprimierte Wissenseinheiten, die auf geringem Raum eine Fülle von Informationen anbieten ohne die in anderen Texten üblichen Redundanzen und Füllmittel zu benutzen. Konvertierung bedeutet also: Übertragung 1:1 plus Erstellung der Verknüpfungen. Die Segmentierung und Relationierung muß deshalb im wesentlichen über die formalen Texteingenschaften wie Randnummern und Absätze erfolgen. Um längere Texte, z.B. Gerichtsurteile unverändert darstellen zu können, muß auf "Scroll-Felder" zurückgegriffen werden. Das Karten-Paradigma ist in diesem Fall nicht zur Darstellung geeignet.

## 2. Darstellung von HyperAdvokat

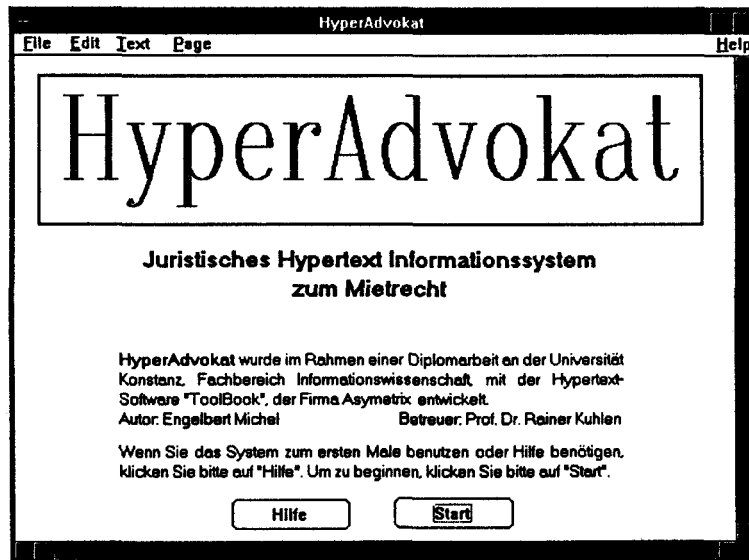


Abb.1 HyperAdvokat Eröffnungsbild

HyperAdvokat ist der Prototyp eines juristischen Hypertext Informationssystems zum Mietrecht (Abb. 1). Als Datenbasis wurde die Einführung zum Mietrecht (Einführung vor § 535) aus dem Bürgerlichen Gesetzbuch (BGB) verwendet. Diese Einführung besteht aus einem Inhaltsverzeichnis und einem Kommentartext, der in 99 Randnummern untergliedert ist. Die betreffenden Gesetzestexte wurden ebenfalls aufgenommen. Das Inhaltsverzeichnis des BGB wurde integriert, um den Einstieg in die Hypertextbasis auch über diesen Weg zu gewährleisten. Das Sachverzeichnis wurde in Auszügen implementiert, um einen Zugang über ein alphabetisches Register zu ermöglichen.

HyperAdvokat ermöglicht dem Juristen, seinen gewohnten Arbeitsstil beizubehalten. Die Software "ToolBook", die zur Erstellung von HyperAdvokat verwendet wurde, arbeitet mit der Buchmetapher. Dateien werden in ToolBook als Bücher und ein Bildschirm als Seite bezeichnet. Bei der Erstellung von HyperAdvokat wurde diese Vorgabe bewußt aufgegriffen und als Stilmittel eingesetzt, um dem Benutzer ein vertrautes Erscheinungsbild zu vermitteln. So befindet sich auf der ersten Bildschirmseite, zu der der Benutzer nach dem Start des Programms gelangt, eine graphische Darstellung der sechs Bücher des BGB (vgl. Abb. 2). Dieses gewohnte Erscheinungsbild soll Vorbehalte gegen das neue Medium *elektronischer Kommentar* gering halten. Gleichzeitig sollen die informationellen Mehrwerte des elektronischen Mediums, wie Suchfunktionen, Einstieg über ein Aktenzeichen oder eine Fundstelle, die Bereitschaft wecken, stärker damit zu arbeiten und auch andere elektronische Medien zu nutzen.

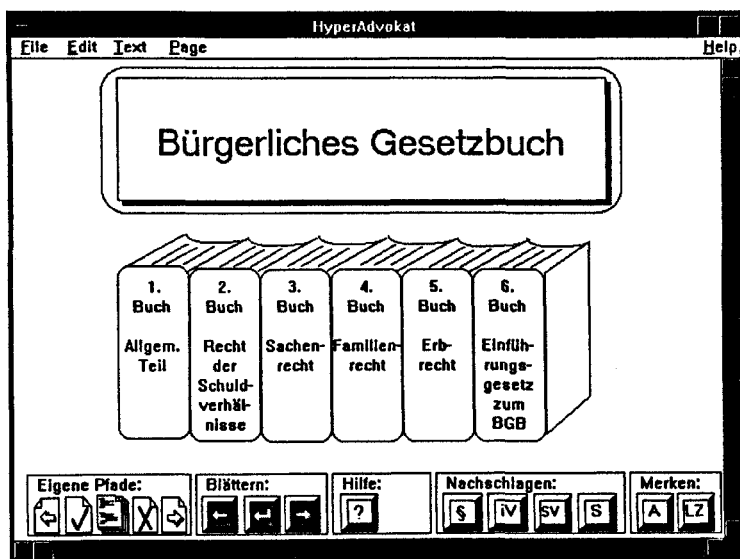


Abb. 2 Titelseite

Die einzelnen Funktionen von HyperAdvokat sind im Hinblick auf die traditionellen Arbeitsweisen von Juristen angelegt. So ist es möglich, ein Lesezeichen auf eine bestimmte Seite zu setzen und zu jeder beliebigen Textstelle Anmerkungen anzubringen. Die Nachschlagefunktionen des Papiermediums werden unterstützt und durch die Möglichkeiten des elektronischen Mediums ergänzt. Der Zugang zum Text ist sowohl über das Inhalts- und Sachverzeichnis als auch über eine bestimmte Gesetzesstelle

oder Zitierung (z.B. vgl § 535 Rn 12) möglich.

Neu hinzu kommt der Zugang über die Suche nach einer bestimmten Fundstelle oder nach dem Aktenzeichen eines Gerichtsurteils. Einen weiteren informationellen Mehrwert schafft die Möglichkeit, den Benutzer eigene Pfade definieren zu lassen. Mit dieser Funktion können die Seiten von HyperAdvokat gekennzeichnet werden, die einmal zur Lösung eines Falles herangezogen wurden. Dadurch kann bei ähnlichen Fällen der Subsumtionsvorgang leicht rekonstruiert werden.

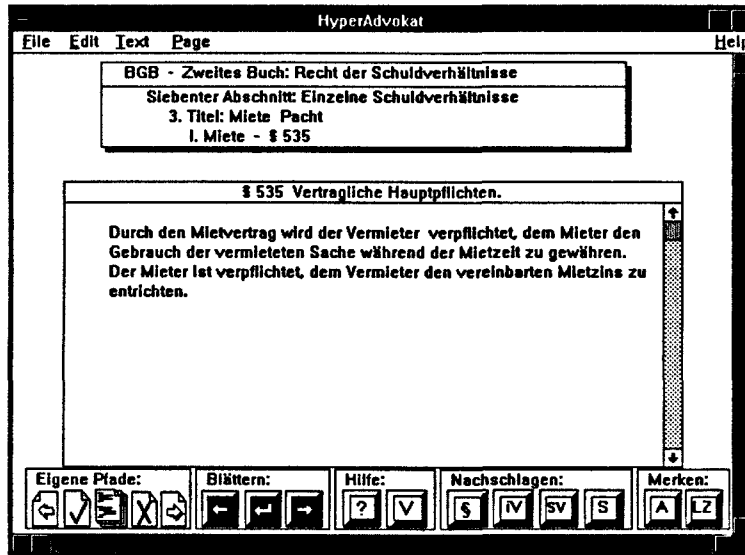


Abb. 3 Seite mit Gesetzestext

HyperAdvokat bietet den direkten Zugriff auf alle im Kommentar angegebenen Verweise. Nach der vollständigen Implementierung kann zu jedem textinternen Verweis das entsprechende Dokument am Bildschirm dargestellt werden, sei es ein Gesetzestext (vgl. Abb 3) ein Gerichtsurteil (vgl. Abb 4), eine andere Kommentarstelle oder ein Literaturhinweis.

Da HyperAdvokat nicht nur dem erfahrenen sondern auch dem gelegentlichen Benutzer seinen vollen Funktionsumfang zur Verfügung stellen soll, ist es notwendig, die Bedienung möglichst so zu gestalten, daß der Benutzer intuitiv die Bedeutung der Buttons erschließen und mit einem Mausklick aktivieren kann. Um eine Vertrautheit mit der Benutzeroberfläche zu erreichen, müssen die einzelnen Bildschirmseiten einheitlich gestaltet sein, d.h. die Buttons und Textfelder erscheinen immer an der gleichen Stelle. Bei Bedienungsproblemen soll ein Hilfesystem die aktuelle Fragestellung sofort beantworten. Von jeder HyperAdvokat-Seite aus kann Hilfe mit der rechten Maustaste aktiviert werden.



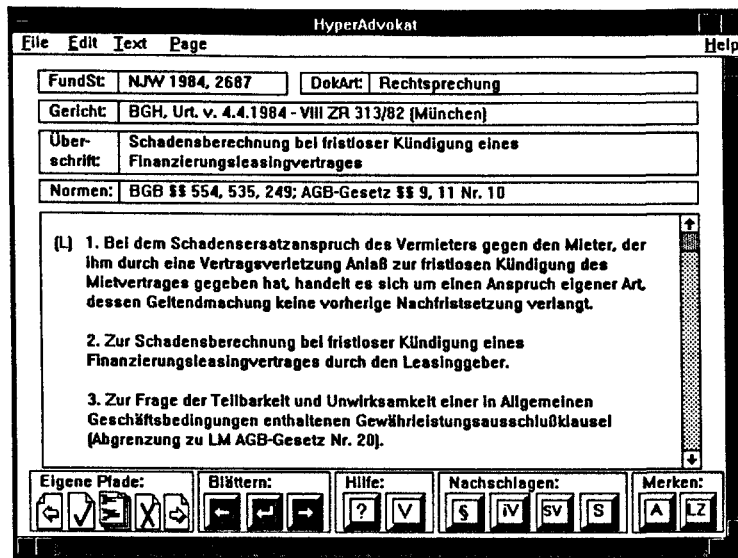


Abb. 4 Externes Dokument BGH-Urteil

Bei einem Kommentar ist es wichtig, daß sehr schnell auf jede beliebige Textstelle zugegriffen werden kann. Die Anzahl der Mausklicks, die benötigt wird, um auf eine bestimmte Stelle zugreifen zu können, muß so gering wie möglich gehalten werden. Außerdem muß es möglich sein, von jeder Seite aus auf jede andere Seite von HyperAdvokat zuzugreifen.

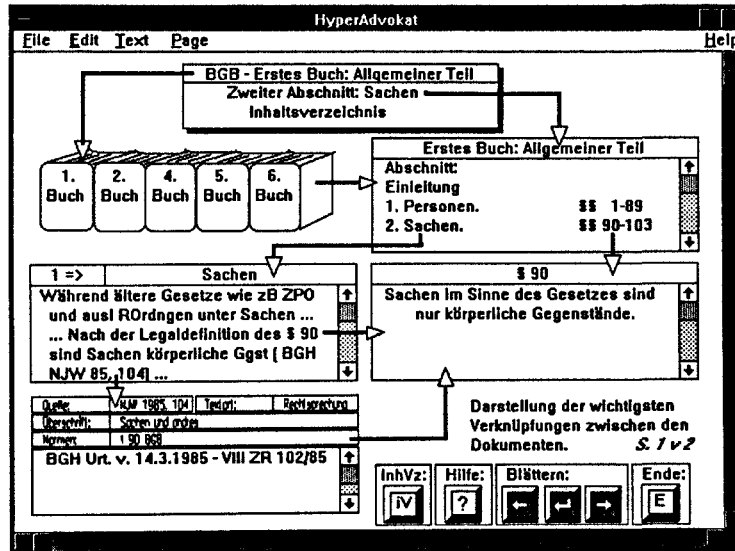


Abb. 5 Die wichtigsten Verknüpfungen in Hyperadvokat

### **3. Verknüpfungen der informationellen Einheiten**

Wesentliches Merkmal eines Hypertext-Systems sind seine Verknüpfungen und die damit einhergehenden Möglichkeiten der Navigation und des Browsing. Die in HyperAdvokat wichtigsten Verknüpfungen sind in der Abbildung 5 dargestellt. Sie sind nicht weiter etikettiert oder typisiert, ihre Semantik ist aus der direkten Funktionalität ersichtlich.

Um dem Benutzer die Möglichkeit zu geben, auf jede Seite in HyperAdvokat zuzugreifen, müßte eine Vernetzung von jeder Seite mit jeder Seite erfolgen. Dies ist über feste Links praktisch nicht realisierbar und würde zweifellos sehr schnell zu einem chaotischen Browsen mit totalem Orientierungsverlust führen. Deshalb wurden nur semantisch begründeten Verknüpfungen *fest verdrahtet*, d.h. allen im Text vorhandenen expliziten Verweisen kann per Mausklick gefolgt werden. Dazu wird die Verweisstelle im Text als sogenanntes *Hotword* definiert und als Ausgangspunkt der Verknüpfung benutzt. Der Zugriff auf eine beliebige andere Seite erfolgt mit einem Zwischenschritt über die Nachschlagefunktion.

### **4. Navigationsmöglichkeiten in HyperAdvokat**

Es bestehen folgende Möglichkeiten der Navigation in HyperAdvokat:

- 4.1 Blättern und Backtrackfunktion
- 4.2 Sprung zur Titelseite mit dem Inhaltsverzeichnis des BGB und zu den Inhaltsverzeichnissen der verschiedenen Büchern und Ebenen mittels Kopfzeile
- 4.3 Sprung zu bestimmten Textstellen mittels Nachschlagefunktion
- 4.4 Sprung zu einer bestimmten Textstelle mittels Sachverzeichnis
- 4.5 Navigation mit Hilfe der Hotwords
- 4.6 Inhaltsverzeichnisse

#### **4.1 Blättern in HyperAdvokat und Backtrackfunktion**

Das Blättern in HyperAdvokat erfolgt über die üblichen Navigationsbuttons im Feld "Blättern". Die Seiten von HyperAdvokat wurden so angeordnet, daß der Benutzer die Möglichkeit hat wie in einem Buch zu blättern. So kann er zuerst die Inhaltsverzeichnisse durchwandern, gelangt dann zu den Paragraphen und anschließend in den Kommentar. Vor diesen Seiten - für den Benutzer jedoch nicht erkennbar - befindet sich das Hilfesystem, das ebenfalls vom Inhaltsverzeichnis des Hilfesystems her durchblättert werden kann. Die externen Dokumente sind im Anschluß an den Kommentar angeordnet. Sie können ebenfalls durchblättert werden, befinden sich aber in ungeordneter Reihenfolge, sodaß reines Blättern keinen zusätzlichen Informationsgewinn bringt.

#### **4.2 Sprung zur Titelseite und zu den Inhaltsverzeichnissen der verschiedenen Bücher und Ebenen**

Mit Ausnahme der Hilfsseiten ist es möglich von jeder Seite aus zur Titelseite mit den sechs Büchern

des BGB zurückzukehren. Dies geschieht mit Hilfe des Buttons "IV" im Feld Nachschlagen. Damit ist gewährleistet, daß ein Leser, der dem Phänomen des "Lost in Hyperspace"<sup>3</sup> erlegen ist, jederzeit zu einem ihm bekannten Ausgangspunkt gelangen kann.

Mit Hilfe der Kopfzeile kann aus einer Kommentarstelle heraus jederzeit zu den darüberliegenden Inhaltsverzeichnissen gesprungen werden.

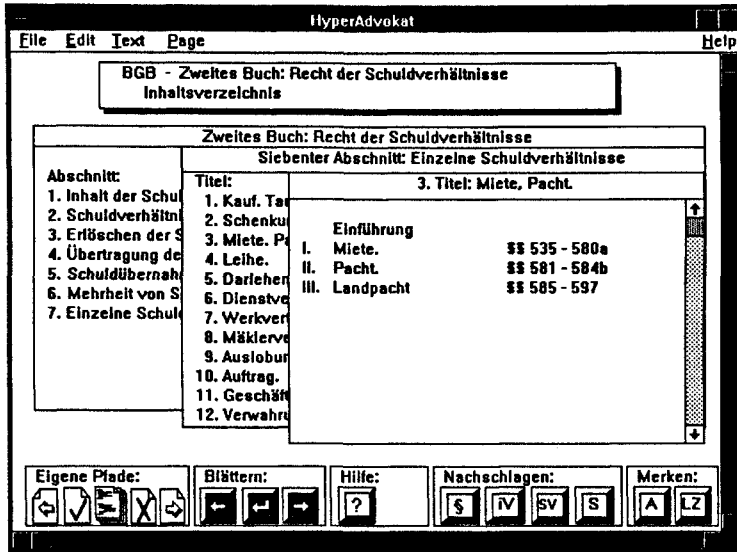


Abb. 6 Inhaltsverzeichnis mit Unterebenen

Ein Mausklick auf BGB in der ersten Zeile führt zur Titelseite mit dem Inhaltsverzeichnis des BGB, ein Klick in eine der darunterliegenden Zeilen zu dem Inhaltsverzeichnis der jeweiligen Hierarchiestufe. Die entsprechenden Unterebenen werden dabei automatisch eingeblendet (vgl. Abb. 6).

#### 4.3 Sprung zu einer bestimmten Textstelle mittels Nachschlagefunktion

Die Nachschlagefunktion wird mit dem Button "\$" im Feld "Nachschlagen" aktiviert. Dies führt zur einer Seite auf der drei Felder mit verschiedenen Nachschlagefunktionen untergebracht sind (Abb. 7). Es gibt folgende Möglichkeiten:

- 1) einen bestimmten Gesetzestext oder eine bestimmte Kommentarstelle direkt anzuspringen;
- 2) eine Fundstelle eines externen Dokumentes zu suchen;
- 3) das Aktenzeichen einer bestimmten Gerichtsentscheidung nachzuschlagen.

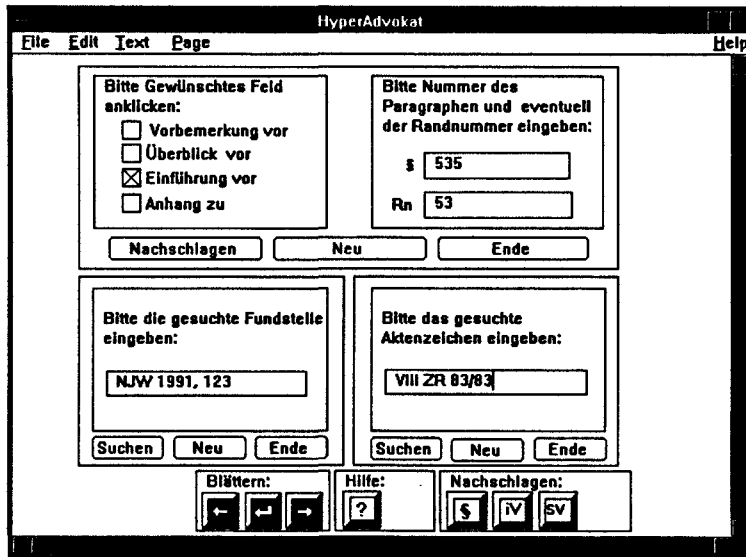


Abb. 7 Nachschlagefunktionen

#### 4.4 Suchen im Sachverzeichnis

Beim Sachverzeichnis handelt es sich um ein lexikographisch geordnetes Verzeichnis, bei dem die zusammengehörigen Sachverhalte systematisch zusammengestellt sind und neben einem schnellen Zugriff auch die Möglichkeit zum konzeptuellen Browsing ermöglichen. Die Suche im Sachverzeichnis kann von jeder Seite aus mit Hilfe des Button "SV" aktiviert werden.

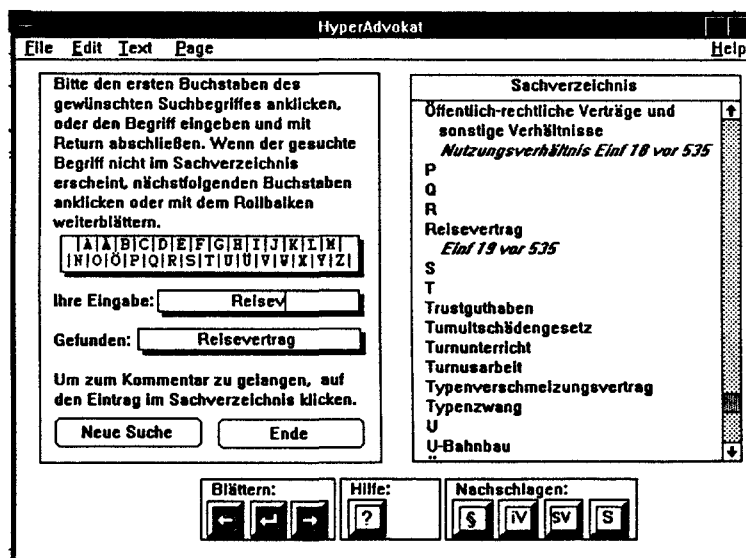


Abb. 8 Sachverzeichnis

Die Seite mit dem Sachverzeichnis (Abb. 8) besteht aus zwei Feldern. Das linke Feld enthält die Eingabefelder, das rechte Feld das eigentliche Sachverzeichnis. Der Benutzer hat verschiedene Mög-

lichkeiten, den gewünschten Begriff zu suchen. Er kann entweder im Alphabet die entsprechenden Buchstaben anklicken, den Begriff direkt eingeben oder im Sachverzeichnis mit Hilfe des Rollbalkens zu dem Begriff rollen.

## 5. Beschreibung der Funktionen

### 5.1 Definition eigener Pfade

Die Reihenfolge, in der die einzelnen Wissenseinheiten angeboten werden, ist ein wichtiger Faktor, der den Benutzern das Verständnis des angebotenen Materials erleichtern soll. Bei HyperAdvokat, einem *Nachschlagewerk*, ist die Reihenfolge, in der die informationellen Einheiten angeboten werden, dem Leser bekannt. Abhängig von dem zu subsumierenden Sachverhalt sind jeweils andere Einheiten wichtig, auf die der Benutzer schnell zugreifen können muß. Wichtig ist die Möglichkeit, Teilschritte die schon einmal zur Lösung eines Falles führten, speichern zu können.

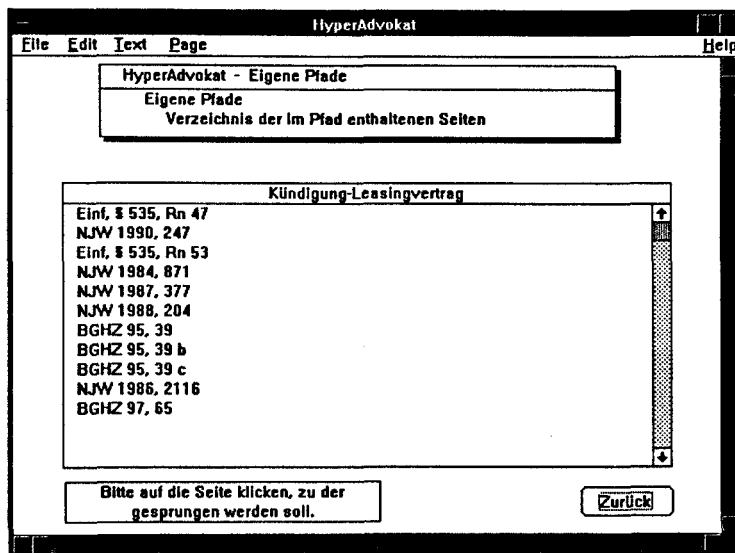


Abb. 9 Eigener Pfad "Kündigung-Leasingvertrag"

Jeder Benutzer von HyperAdvokat kann alle Seiten, die er benutzt hat und die für ihn relevant waren, problemlos markieren und bei Bedarf immer wieder aufrufen (vgl. Abb. 9). Diese benutzererstellten Verknüpfungen einzelner informationeller Einheiten werden hier als *eigene Pfade* bezeichnet. Denkbar und sinnvoll erscheinen auch *autorendefinierte Pfade* die Teilmengen aus den HyperAdvokat-Einheiten zusammenfassen. Sie können z.B. die für einen häufig wiederkehrenden Sachverhalt relevanten Einheiten repräsentieren.

Die Erstellung eigener Pfade und die Navigation im Pfad geschieht mit Hilfe der fünf Buttons im Feld "Eigene Pfade". Der mittlere Button dient den grundlegenden Funktionen, *Erstellung, Bearbeitung oder Löschung* eines Pfades. Die restlichen Buttons dienen der *Aufnahme* und *Streichung* von Seiten sowie dem *Vorwärts- und Rückwärtsblättern im Pfad*.

Der Pfad kann jederzeit durch das Anlicken eines anderen Buttons verlassen werden. Bestehende Pfade können außerdem durch die Aufnahme neuer Seiten erweitert werden. Die neuen Seiten werden an der Stelle in den Pfad eingefügt, an welcher der Pfad verlassen wurde.

## 5.2 Hilfefunktion

Die Benutzung von HyperAdvokat soll möglichst intuitiv und ohne große Einarbeitung, bzw. langes Studium der Dokumentation möglich sein. Dazu ist es erforderlich dem Benutzer an jedem Punkt seiner Navigation in HyperAdvokat kontextsensitive Hilfestellung anzubieten. Die Hilfefunktion umfasst die Möglichkeiten:

- 1) Zu jedem Objekt auf dem Bildschirm eine Kurzinformation aufzurufen;
- 2) Einen ausführlicheren Hilfetext anzeigen zu lassen.

Da die Navigation in HyperAdvokat ausschließlich mit der linken Maustaste erfolgt, konnte die rechte Maustaste mit der Hilfefunktion belegt werden. Jedes Objekt auf dem Bildschirm kann mit der rechten Maustaste angeklickt werden, um einen kurzen Hilfetext aufzurufen. Das daraufhin erscheinende Dialogfenster mit der Kurzinformation enthält zwei Buttons. Der erste dient der Beendigung der Kurzhilfe, der zweite führt zur ausführlicheren Hilfe.

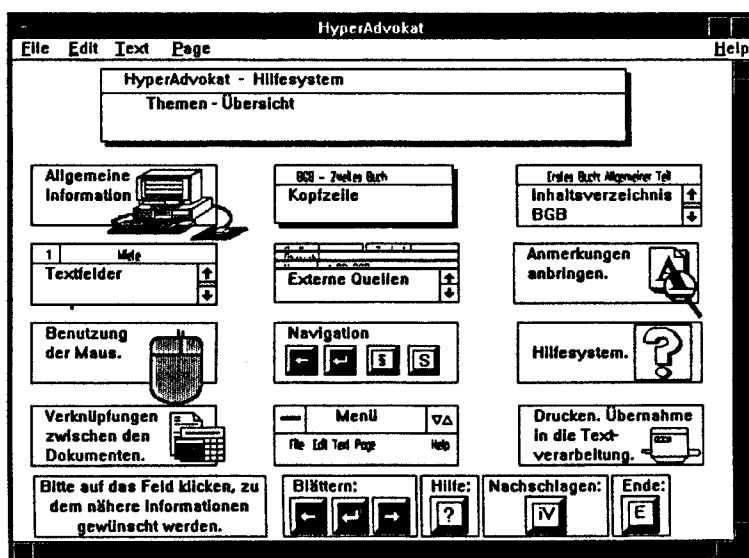


Abb. 10 Übersicht des Hilfesystems

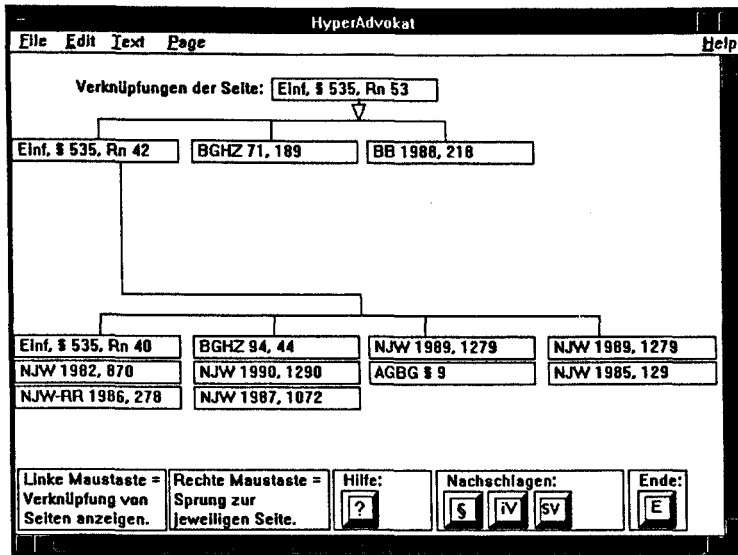


Abb. 11 Beispielseite aus dem Hilfesystem

### 5.3 Annotationen

Jeder Benutzer kann zu jeder beliebigen Textstelle im Kommentar, Gesetzestext oder den externen Dokumenten Anmerkungen anbringen. Diese werden in einem gesonderten ToolBook auf der Festplatte gespeichert. Für verschiedene Benutzer von HyperAdvokat können jeweils eigene Anmerkungsdateien eingerichtet werden, nicht zuletzt auch deshalb, um die Vertraulichkeit zu sichern.

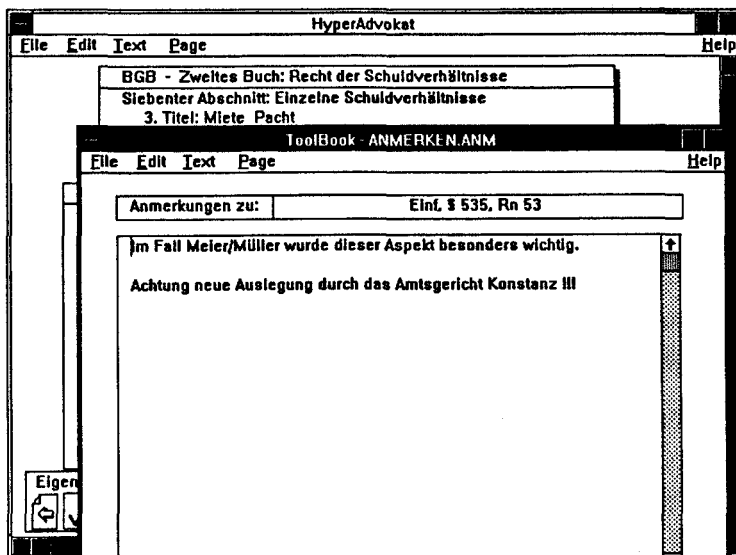


Abb. 12 Beispielseite aus dem Anmerkungs-ToolBook

Die Anmerkungsdatei bleibt als zweites Fenster aktiv (Abb. 12). Im Verlauf der Arbeit mit HyperAdvokat kann deshalb schnell zwischen den Fenstern gewechselt werden. Erst zum Ende der Sitzung wird die

Anmerkungsdatei geschlossen.

#### **5.4 Lesezeichen**

Jede Textseite von HyperAdvokat kann mit einem Lesezeichen versehen werden. Sobald ein Lesezeichen gesetzt wurde, kann diese Seite von jeder anderen Seite aus angesprungen werden.

#### **5.5 Suchfunktion**

Der Button "S" aktiviert eine Suchfunktion, mit der man bei längeren Dokumenten nach bestimmten Worten oder Teilen daraus suchen kann. Die Suchfunktion ist auf die Felder der jeweiligen HyperAdvokat Seite beschränkt. Wird der Eintrag in das Suchfeld mit einem "?" trunziert, wird die Suche als Stringsuche durchgeführt, das heißt jedes Vorkommen dieser Buchstabenfolge wird angezeigt, unabhängig davon, ob diese ein vollständiges Wort darstellt, oder ob es sich um Teile eines Wortes handelt.

### **6. Automatische Konvertierung**

Die automatische Konversion eines gedruckten Kommentars in Hypertext und der dafür erforderliche Aufwand werden von folgenden Faktoren beeinflusst:

#### **1. Grundstruktur von HyperAdvokat:**

- \* Die Grundstruktur von HyperAdvokat liegt vor und kann nun mit weiterem Text gefüllt werden. Die Funktionen sind so ausgelegt, daß sie problemlos auf größere Daten angewendet werden können.

#### **2. Verfügbarkeit des Ausgangstextes des BGB:**

- \* Liegt dieser Text maschinenlesbar vor?
- \* Müssen Steuer- und Sonderzeichen bereinigt werden?
- \* Soll eine Auflösung der Abkürzungen erfolgen?

Vorbereitung für den Import in HyperAdvokat:

- \* Wird ein automatisches Verfahren entwickelt, das den Text editiert?
- \* Muß die Editierung manuell vorgenommen werden?

#### **3. Externe Dokumente:**

- \* Liegen diese maschinenlesbar vor?
- \* Müssen sie von Steuer- und Sonderzeichen bereinigt werden?
- \* Müssen Texte eingescannt werden?
- \* Wie fehlerbehaftet ist die OCR-Software?
- \* Wie hoch ist der Korrekturaufwand?
- \* Müssen Texte manuell eingegeben werden?
- \* Vorbereitung für den Import in HyperAdvokat:
- \* Wird ein automatisches Verfahren für die Texteditierung entwickelt?



- \* Muß die Editierung manuell vorgenommen werden?

#### 4. Automatisierbarkeit der Relationierung:

- \* Müssen Verknüpfungen manuell erstellt werden?
- \* Können Verfahren entwickelt werden, die Ausgangspunkte von Links automatisch erkennen und diese *fest verdrahten*?
- \* Können dynamische Links (durch Anklicken einer Textstelle wird vom System ein Link automatisch erstellt) implementiert werden?

#### 5. Hypertextgerechte Überarbeitung des Kommentars:

- \* Ist es möglich, den Autoren des Kommentars die Überarbeitung nach bestimmten Regeln abzuverlangen, z.B. einheitliche Zitierung und Behandlung von Verweisen, Aufbereitung der zitierten externen Dokumente?

### Ausblick

Die Realisierung der vollständigen Überführung des BGB in Hypertext ist technisch möglich. Ob es innerhalb ökonomisch vertretbarer Grenzen erfolgen kann, hängt wesentlich von der Automatisierbarkeit einzelner Überführungsschritte ab. Die streng hierarchische Gliederung und die strikten Regeln, nach denen der Kommentar aufgebaut ist, begünstigt automatische Verfahren. Diese können mit einem vertretbaren Aufwand erstellt werden. An der Universität Konstanz wurde von Anton Rieser<sup>4</sup> ein Programm zur automatischen Konvertierung eines juristischen Textes in Hypertext erstellt. Diese HyperCard-Anwendung benutzt Textoberflächenmerkmale für die Segmentierung und Relationierung der einzelnen Abschnitte.

Hypertext und juristische Texte sind eine sinnvolle Verbindung. Ob in nächster Zeit Anwendungen vergleichbarer Art auf den Markt kommen werden hängt von verschiedenen Faktoren ab, z.B.

- \* Akzeptanz bei Juristen
- \* Mut eines Verlages dieses neue Medium zu verwenden
- \* Robuste Software bei großen "Text"mengen
- \* Lösung für die Konversionsprobleme
- \* Pflege und Updating
- \* Kosten
- \* Entwicklung von geeigneten Speichermedien

---

#### Literatur:

<sup>1</sup> Kuhlen, R.: Hypertext - ein nicht-lineares Medium zwischen Buch und Wissensbank; Springer-Verlag, Hamburg 1991

- 
- 2 Conklin, J.; Hypertext: An Introduction and Survey; Reprint from Computer Magazine Sept. 1987, S 17-41
  - 3 Edwards, D.M.; Hardman, L.; "Lost in Hyperspace": Cognitive mapping and navigation in a hypertext environment; in McAleese, R (Hrsg.) Hypertext: theory into practice; London 1989
  - 4 Rieser, Anton; Automatische Konversion von Text in Hypertext durch Ausnutzung formaler Texteingenschaften; Diplomarbeit; Universität Konstanz 1991

Autor: Dipl. Inf.wiss. Engelbert Michel;  
Universität Konstanz,  
Informationswissenschaft,  
Universitätsstr. 10,  
7750 Konstanz

Thomas A. Schröder

**Abstract**

In diesem Text wird ein Informationssystem, daß auf dem Konzept von Hypertext und Hypermedia basiert, über die Französische Revolution vorgestellt. Diese Applikation erfordert IBM kompatible PCs und ist in der Software ToolBook 1.5, einem "Software Construction Set for WINDOWS" realisiert. Ziel dieser experimentellen Applikation ist es, Strukturierungskriterien wie Verknüpfungsmöglichkeiten einer historischen Hypertextbasis zu gewinnen. Ein Beispiel für die Navigation innerhalb der bisher modellierten Teile von 1789 demonstriert deren Benutzeroberfläche, Flexibilität und Funktionalität.

*This paper explores how a Information System, based on the concepts of Hypertext and Hypermedia, can be used in the historical science, specially for the epoch of French Revolution. The Information system is beeing developed on IBM-PCs, using ToolBook 1.5, a "Software construction Set for WINDOWS". The Intention of this experimental Application is to get a struktur and a linking concept for a historical hypertextbase. One navigational example that have been created in 1789, describes design issues, flexibility and functionality.*

**Historische Wissenschaft und Hypertext**

Die Erkenntnisse der historischen Forschung basieren, wie bei jeder Geisteswissenschaft überwiegend auf Textanalyse. Dabei wird der Historiker mit einer Vielzahl unterschiedlicher Texte konfrontiert. Primäres Interesse finden Texte, die unter der spezifischen historischen Fragestellung zu Originalquellen werden können, sekundäres Interesse Literatur historischer und anderer wissenschaftlicher Diszipline. Zu einer historischen Quelle kann im Prinzip jeder Text werden, in der Praxis sind es meist Akten und Urkunden, Briefe, Monographien, Redemanuskripte und Berichte.

Der Historiker setzt für die Beschaffung und Analyse der Quellen bisher auf die klassischen geisteswissenschaftlichen Methoden. Da bereits Computer und PCs bei Historikern verbreitet sind, werden von diesen für die Quellenanalyse auf der Software-Seite relationale Datenbanken eingesetzt/1/.

Vorteile des Einsatzes relationaler Datenbanken für die Auswertung großen Quellenmaterials sind die unterschiedlichsten Retrievalformulierungen, die schnell abgearbeitet werden können. Retrieval mit einer hohen Präzision setzt aber eine adäquate Quellen-Repräsentation in der Datenbank voraus. Der Text einer weitgehend unformatierten Quelle muß kategorisiert und verschlüsselt werden, um der Definition einer Datenbank mit Datenfeldern und Datenlängen zu entsprechen. Um das Retrieval selbst durchzuführen muß der Historiker wenigstens in Grundzügen mit der Theorie der Relationen vertraut sein.

Das Hypertext-Konzept, in den historischen Wissenschaften bisher nur im geringen Umfang zur Kenntnis genommen/2/, zeichnet sich dagegen durch Quellennähe aus. Eine Quelle kann in einem Hypertextsystem direkt und ohne Codierung repräsentiert werden, wobei die Struktur der Quelle erhalten bleibt. Dem Benutzer kann angeboten werden selbst die Repräsentationsart zu wählen, z.B. zwischen einer Faksimile-Ausgabe oder einer wissenschaftlichen Edition mit kritischem Apparat. Der Quellentext kann zusätzlich nach externen, historischen Kriterien strukturiert oder analysiert werden. Zudem kann die einzelne Quelle mit anderen, direkt auf sie bezogenen (zeitgenössische Texte wie wissenschaftliche Analysen) oder semantisch assoziierten Quellen verknüpft werden.

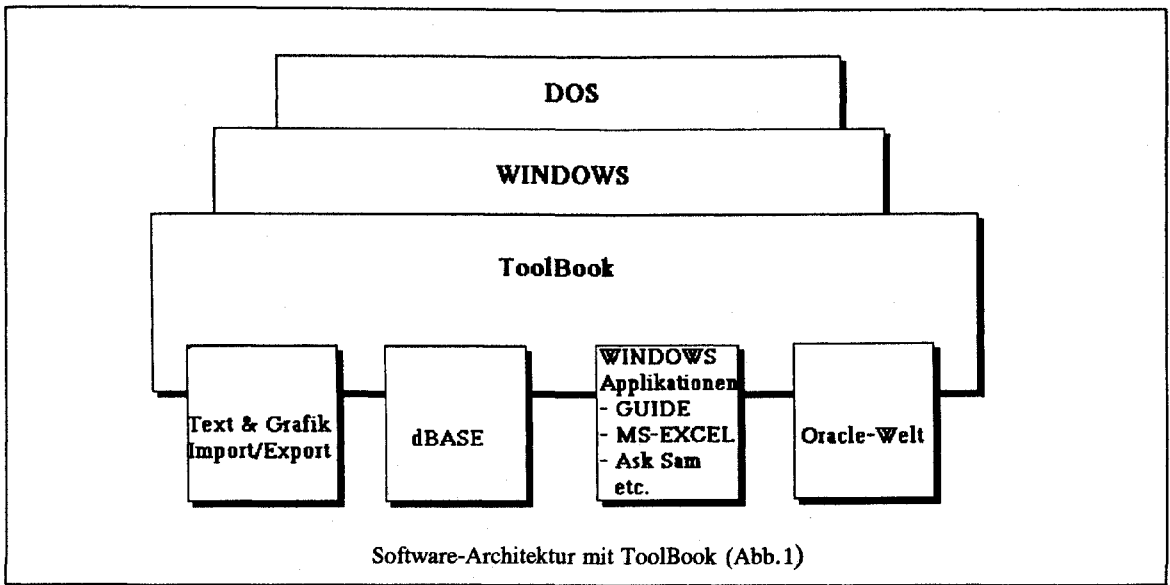
Aufgrund der grafischen Benutzeroberfläche eines Hypertextsystems kann die Navigation bzw. gezieltes Retrieval des Historikers durch das System so unterstützt werden, daß der Historiker keine speziellen Hypertext-Kenntnisse zu haben braucht. Das Hypertext-Konzept wird für den Historiker auch deswegen interessant, weil in Hypertext und dann Hypermedia auch nicht-textuelle Informationen (Illustrationen, Statistiken, Photos, Grafiken, Reden und Musik sowie Film) sowie Faktendaten (etwa Klimaangaben, Währungskurse, Bevölkerungszahlen) in das Informationssystem integriert werden können. Quellenarten, die in zunehmenden Maße das Interesse der historischen Forschung finden/3/.

Ziel dieser experimentellen Applikation ist es, durch die Modulation eines begrenzten historischen Themas (Deutsche Revolutionsreisende nach Paris 1789-1798) Strukturierungskriterien für eine historische Hypertextbasis über die Französische Revolution ("1789") und Vernetzungsmöglichkeiten unterschiedlicher Quellen zu finden.

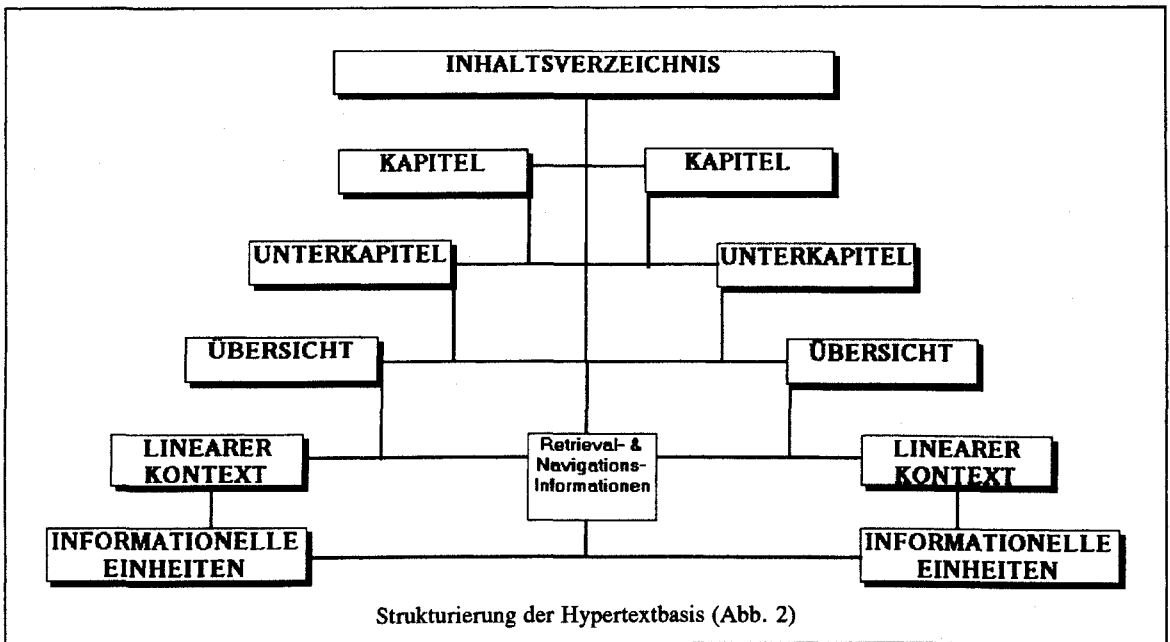
### **Hypermedia Informationssystem 1789**

Für die Realisierung des Informationssystems 1789 wird die kommerzielle Software ToolBook 1.5 der Firma Asymetrix/4/ als ein "front-end" System eingesetzt. ToolBook gilt als ein "Software Construction Set for WINDOWS", weil mit ToolBook komfortable Applikationen unter WINDOWS 3.0 aufgebaut werden können, ohne daß der Designer spezielle WINDOWS-Programmierkenntnisse zu haben braucht. Für den Einsatz von ToolBook spricht, daß innerhalb von ToolBook a) Datenbank-, Textverarbeitungs- und Tabellenkalkulationsfunktionen möglich sind, b) Volltexte abgebildet werden können, die c) mit Datensätzen auf einer (grafisch komfortabel gestaltbaren) Benutzeroberfläche kombiniert werden können und d) die Programmiersprache OpenScript, die an den Paradigmen der objektorientierten Programmiersprachen orientiert und leicht lernbar ist. Die e) Integrationsfähigkeit von ToolBook erlaubt es, Programme unter MS-WINDOWS, dBASE und der ORACLE Welt anzusprechen. Über die Software von ORACLE wird ToolBook SQL-fähig und kann online konsultiert werden. Durch f) Import Funktionen können Text- oder Grafikdateien eingelesen werden bzw. auch exportiert werden.

Für die Applikation 1789 ist eine Software-Architektur möglich wie sie Abbildung 1 zeigt. Gegenwärtig beschränkt sich 1789 auf Funktionen innerhalb von ToolBook.

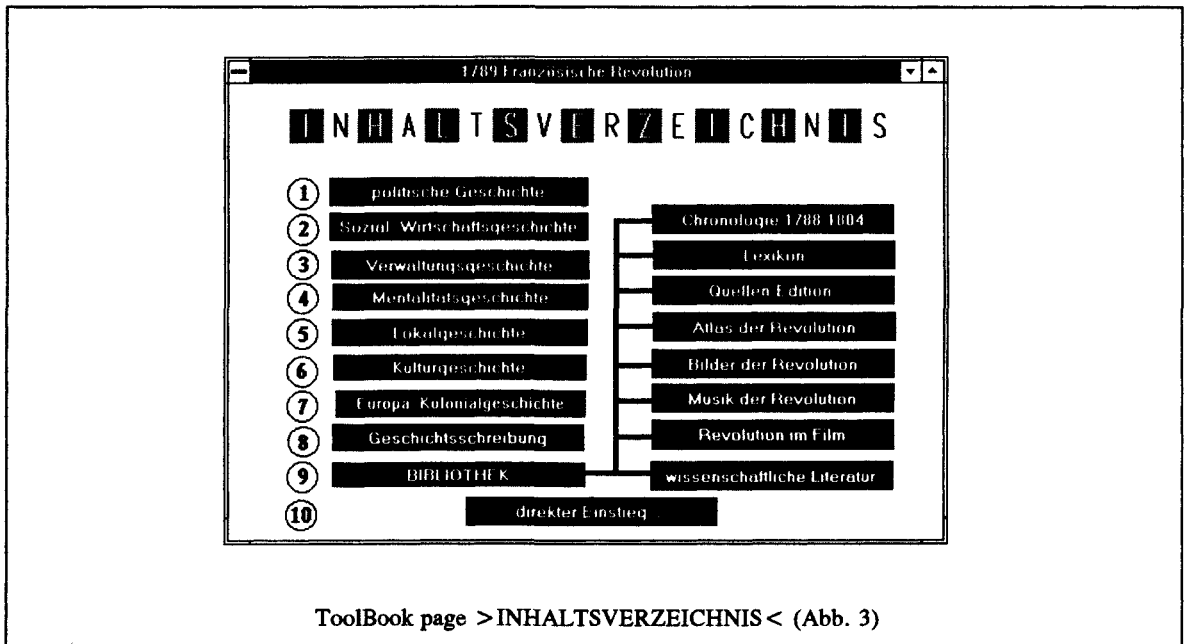


Die Strukturierung der Hypertextbasis orientiert sich primär an konventionellen und hierarchischen Prinzipien wie sie Abbildung 2 zeigt. So soll der Einstieg des Historikers in die Hypertextbasis durch die Präsentation historischer Fachtermini erleichtert werden. Um aber semantisch assoziierte Themen mit anderen Informationen bzw. Texten zu verknüpfen, müssen die Quellentexte in semantisch geschlossene Fragmente transferiert werden. In Anlehnung an Kuhlen/5/ verwenden wir für diese Texte den Begriff der "informationellen Einheit".



## Navigationsbeispiel

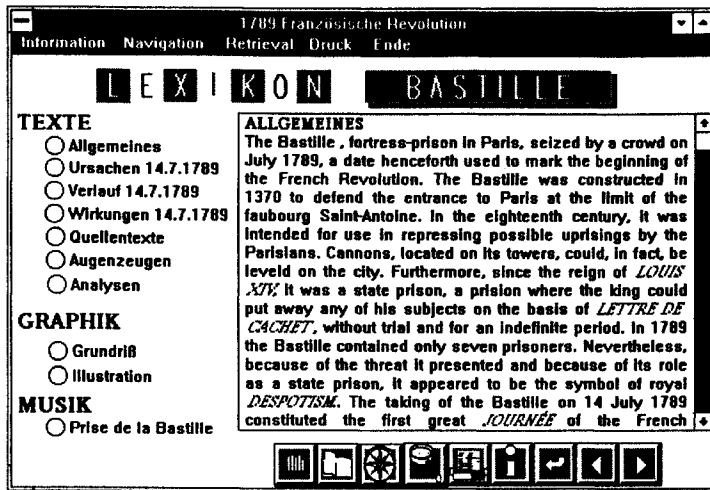
Die Applikation **1789** startet durch Anklicken des entsprechenden Icons im WINDOWS 3.0 Programm-Manager. Der Benutzer wird dann automatisch zur ToolBook page **>INHALTSVERZEICHNIS<** geführt.



Das Konzept für die strukturelle Modellierung der Hypertextbasis **1789** orientiert sich an den historischen Teildisziplinen bzw. der Klassifikation der "Bibliographie Annuel"/6/ und fügt für zentrale Informationen die Option **>BIBLIOTHEK<** hinzu. Diese grobe und ambivalente Klassifikation kann, durch Hypertext-Verknüpfungen, bereichert werden.

Der Einstieg in die Hypertextbasis erfolgt einerseits vertikal durch die Auswahl einer von neun Kapitelüberschriften wie sie Abbildung 3 zeigt. Jede historische Teildisziplin im **>INHALTSVERZEICHNIS<** ist als Button realisiert, der eine Graphik mit weiteren Unterteilungen eröffnet, deren Anklicken zu einer neuen page führt. Durch die Aufblend- und Abblendtechnik von Feldern/Buttons kann von dieser ToolBook page aus maximal zweiundsechzig Optionen der Einstieg gewählt werden. Andererseits ist es möglich durch die Option **>DIREKTER EINSTIEG<** das hierarchische Menüsystem zu umgehen um gezielt zu einem Begriff oder Thema zu navigieren.

Wird zum Beispiel der Button **>BIBLIOTHEK<** gewählt und dort die Option **>CHRONOLOGIE<**, gelangt man zu einer Übersicht (Zeitstrahl) über die Jahre 1788-1804. Wird das Jahr 1789 angeklickt, erscheint eine Chronologie nach Monaten. Interessiert man sich für den Beginn der Revolution und die Ereignisse im Juli 1789, besonders den 14. Juli und klickt auf **>Sturm auf die Bastille<** gelangt man zum **>LEXIKON Bastille<** und es erscheint die page wie sie Abbildung 4 zeigt.



ToolBook page >LEXIKON Bastille< (Abb. 4)

Jede ToolBook page innerhalb der Hypertextbasis ist im Prinzip gleich aufgebaut. Elemente jeder page sind:

- \* MS-WINDOWS Menüleiste, die definiert ist mit Pull-Down-Menüs INFORMATION, NAVIGATION, RETRIEVAL, DRUCK und ENDE. Wo nicht unbedingt notwendig, kann die Menüleiste mit den Pull-Down-Menüs abgeblendet werden.

- \* kontextunabhängige Buttons auf der unteren Bildschirmseite mit den Funktionen <Inhaltsverzeichnis> <Kapitelübersicht> <bisheriger Suchweg> <Retrieval> <Druck> <Information> <vorherige Seite> <Seite zurück> und <nächste Seite>; die Redundanz von manchen Funktionen (als Pull-Down-Menü und Button) soll die Benutzerorientierung erhöhen.

- \* eine Beschriftung wie >LEXIKON Bastille<; durch die lineare Speicherung der ToolBook pages (jede ToolBook page ist strukturell ein Datensatz, jedes Objekt ein Datenfeld) kann über den Button <Seite vor> bzw. <Seite zurück> linear geblättert werden wie hier zu den Lexikon-Artikel >LEXIKON Campe, Johann Heinrich< oder >LEXIKON Assignaten<; so weiß der Benutzer immer, in welchem Teil der Hypertextbasis er sich gegenwärtig befindet.

- \* der stringenten background-Gestaltung, wie überhaupt die Gestaltung mit Grafiken und Farben sparsam und kontextspezifisch sein muß (Dies kann in den Abbildungen nur angedeutet werden); die erläuternde Namensgebung jeder page ist wichtig für die Navigationsfunktion "bisheriger Suchweg"; standartmäßig speichert ToolBook die Namen der letzten einhundert pages, die der Benutzer passiert hat; durch Anklicken des Namens kann direkt dorthin gesprungen werden.

- \* einem Textfeld mit Informationen über den Begriff oder Namen; links vom Textfeld befinden sich kontextspezifische Buttons, die zu informationellen Einheiten springen oder von denen andere Funktionen gestartet werden können.

Die ToolBook page >LEXIKON Bastille< ist der Ausgangspunkt für folgende Zielpunkte/7/: Durch Anklicken des Buttons >Grundriß< kann man eine Karte der Bastille aufrufen und Erklärungen für bestimmte Gebäude erhalten. Man ist dann im Bereich >ATLAS< der Hypertextbasis und kann zum Beispiel zu einem Stadtplan von Paris navigieren, oder auch einfach zum LEXIKON zurückkehren.

Durch Anklicken des Buttons >La Prise de la Bastille< wird zur ToolBook page >MUSIK Bastille< gesprungen und dort kann das Lied "La Prise des la Bastille"(von 1789) gehört oder nur der Liedtext studiert werden. Durch einfaches Seitenblättern kann man zur page >MUSIK La Marseillaise< oder >MUSIK La Carmagnole< gelangen und Liedtext studieren wie Musik hören. Von dort könnte man zum Lexikon-Artikel, der sich grundsätzlich über die Musik der Revolution äußert, navigieren.

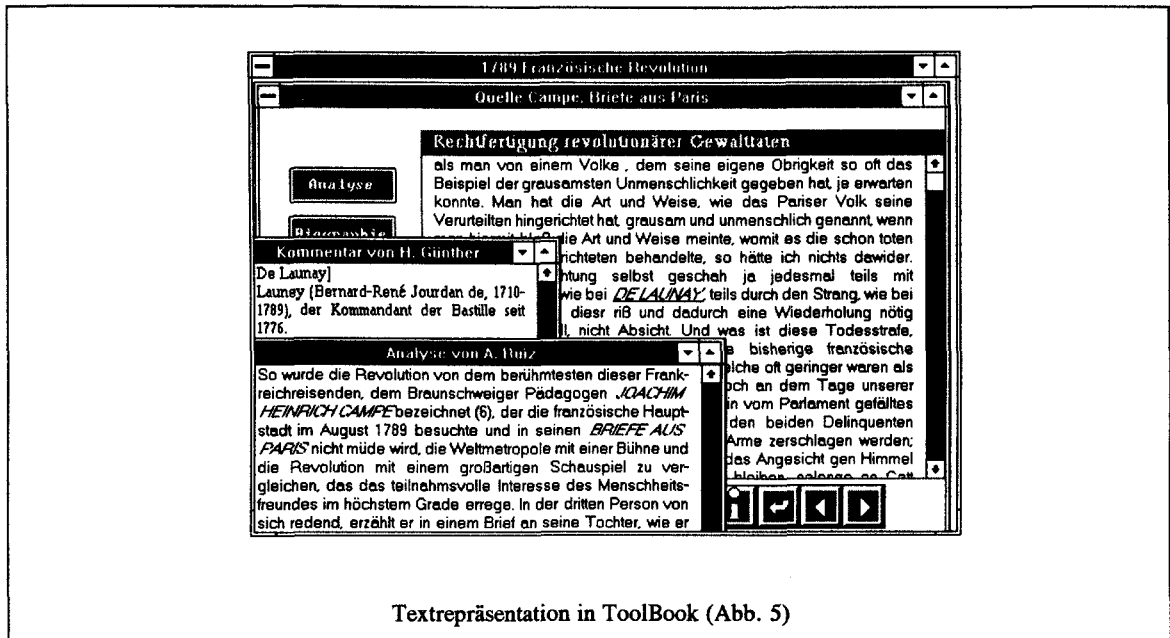
Durch die backtrack-Funktion gelangt man von >MUSIK La Prise de la Bastille< wieder zum >LEXIKON Bastille<. Interessiert man sich für den erläuternden Text >Verlauf des 14.Juli< wird zu der entsprechenden Passage des Textes geblättert. Im Text findet man die Aussage des Autors/8/, daß ausländische Beobachter mit Schrecken die Greuelthaten des 14.Juli wahrgenommen haben. Der Autor weist auf eine Stelle bei J.H.Campe, "Briefe aus Paris zur Zeit der Revolution geschrieben" (1790) Seite 56-57 hin. Klickt man auf diese Fußnote, wird der Benutzer in einem Dialogfenster gefragt, ob er direkt diese Quelle lesen oder eine Übersicht über ausländische Beobachter dieses Tages haben möchte. Wählt man nun in der >ÜBERSICHT ausländische Beobachter< die Option >Deutsche Beobachter< gelangt man zur ToolBook page >ÜBERSICHT Deutsche Revolutionsreisende<. Klickt man dort beispielsweise auf den Namen J.H.Campe kann man wieder in das >LEXIKON< zum Artikel über Johann Heinrich Campe springen. Von dort kommt man weiter in die >QUELLEN-EDITION Campe, Briefe aus Paris<. Auf der ToolBook page >ÜBERSICHT Deutsche Revolutionsreisende< kann man aber auch eine ToolBook-Datei, die ausschließlic als Datenbank definiert ist, laden und auf die Frage "Welche Deutschen Revolutionsreisende nach Paris der Jahre 1789 bis 1792 waren von Beruf Schriftsteller und kamen aus Norddeutschland"/9/ eine präzise Antwort bekommen.

Aufgrund des Dynamischen Datenaustausches (DDE) unter MS-WINDOWS ist es möglich, Texte außerhalb von ToolBook in GUIDE 3.0 zu realisieren und von ToolBook aus auf diese zu zugreifen sowie die GUIDE Funktionen zu nutzen. Diese Konnexion war sinnvoll, da die ältere Version von ToolBook in der Fähigkeit zum Textmanagement eingeschränkt war. Vergleiche der Quellen-Repräsentation in ToolBook (Version 1.5) und GUIDE 3.0 haben gezeigt, daß diese Konnexion nun keine Vorteile mehr bringt. Die Funktionen der GUIDE Buttons (Note, Expansion, Referenz, Command) können in ToolBook 1.5 mit mehr Flexibilität und Evidenz realisiert werden und eine einheitliche Benutzeroberfläche bleibt erhalten.

Um den Text zu lesen wird durch Aufruf der ToolBook page >QUELLEN-EDITION Campe, Briefe aus Paris< automatisch ein Button aktiviert und die Datei/10/ legt sich als neues Fenster vor die ToolBook Hauptdatei. In den "Briefen aus Paris" sind besondere Formulierungen Campes als "Hotwords" definiert, zu denen Anmerkungen einer wissenschaftlichen Edition/11/ in



einem weiteren Fenster geladen werden können. Parallel kann auch eine historische Analyse/12/ der "Briefe aus Paris" in einem vierten Fenster geladen werden. Diese Multi-Window-Technik zeigt Abbildung 5.



Hätte der Benutzer aber im oben erwähnten Dialogfenster statt >ÜBERSICHT Deutsche Revolutionsreisende< die Option >QUELLE< gewählt, so hätte er die Datei >QUELLEN-EDITION Campe, Briefe aus Paris< direkt laden können und wäre zu der entsprechenden Textpassage (S. 56-57) gelangt. Im Text ist festzustellen, daß J.H.Campe keineswegs revolutionäre Gewalttaten verurteilt, sondern sie eher aufgrund seiner Vorstellung des "guten Volkes" rechtfertigt. Werden die Textdateien verlassen, schließen sich die Fenster und die ToolBook Hauptdatei ist wieder aktiviert.

Mit einem einfachen Programm, geschrieben in der Programmiersprache von ToolBook, OpenScript, ist es möglich, das Vorkommen eines Begriffes in der Hypertextbasis festzustellen, diesen als Hotword zu definieren und dem Hotword ein Programm (Script) zu zuordnen, daß eine Verknüpfung (link) zu einer gewünschten page setzt.

Das folgende Programm sucht nach dem Begriff >Bastille< und definiert jeden gefundenen Begriff als ein Hotword, daß eine Verknüpfung zu der page >LEXIKON Bastille< anbietet.

```

to handle buttonUp
  system begriff
  set sysError to null
  go to first page
  ask "Welcher Begriff soll verknüpft werden? Achtung: Das Hotword mit dem gewünschten Script muß im Zwischenpeicher sein"
  put it into begriff
  search for begriff
  while "Not Found" is not SysError
    request "Soll hier gelinkt werden?" with "Ok" or Cancel"

```

```

If it is OK
  send transparent
  send paste -- aus dem WINDOWS Zwischenspeicher wird das Hotword mit Programm eingefügt
  search again
end
If it is Cancel -- nicht jeder vorkommende Begriff soll verknüpft werden !
search again
end
end
IF "not Found" is in sysError
  request "Alle Begriffe sind verknüpft !" with "OK"
  IF it is "Ok"
    break to system
  end
end
end
If it is "Cancel"
  break to system
end
end buttonUp

```

Die Verwendung der Zwischenablage von MS-WINDOWS kann die Programmstruktur vereinfachen. Ein Hotword mit folgenden Programm wird für jeden gefundenen Begriff von obigen Programm mit dem Befehl "paste" eingefügt.

```

to handle buttonUp
  request "Wollen Sie eine Erläuterung zur Bastille im LEXIKON erhalten? Von dort können weitere Verknüpfungen (Texte, Karten, Musik etc) aufgerufen werden." with "Ok" or "Cancel"
  If it is "Ok"
    go to page "LexikonBastille" of this book
  end
  If it is "Cancel"
    break to system
  end
end buttonUp

```

## Ausblick

Das Hypermedia Informationssystem 1789 befindet sich in einer experimentellen Phase. Problem jeder Applikation, die sich hauptsächlich auf Texte stützt, ist der Input, die Fragmentierung der Texte in informationelle Einheiten und ihre Vernetzung. Bisher muß diese Modulation überwiegend manuell erfolgen. Für die Hyper-Texte oder informationellen Einheiten können die Autoren auch nicht mehr eindeutig bestimmt werden, so daß die juristische Seite des Copyright problematisch ist. Da ein Thesaurus nicht vorhanden ist, bleibt eine Terminologiekontrolle sowie ein Abgleich der Suchbegriffe in unterschiedlichen Sprachen offen.

Die Modulation eines begrenzten Themas zeigt, eine komplexe historische Epoche wie die Französische Revolution kann als ein Hypermedia Informationssystem realisiert werden. Die bisherigen Möglichkeiten von 1789 eröffnen mit einer komfortablen Benutzeroberfläche, dem Retrieval in Volltexten wie Datensätzen sowie der Integration unterschiedlicher Quellen und Informationen eine umfassende und effiziente Möglichkeit der historischen Analyse.

## Anmerkungen

- 1 In Deutschland hat sich besonders Manfred Thaller vom Max-Planck Institut für Geschichte mit dem Thema Geschichte und Informationswissenschaft auseinandergesetzt (Thaller (1988))
- 2 Die wenigen anderen Projekte, die ein historisches Thema in Hypertext modellieren sind dokumentiert: [für Altertumswissenschaft, griechische Philologie] in Mylonas/Heath (1990), für [Mediaevistik, Ebstorfer Weltkarte auf Hypertext] Warnke (1990); verbreiteter sind Hypertext-Anwendungen im Museumsbereich (e.g. Nielsen (1990, 43-72), N.N.(1990))
- 3 Ein klassisches Werk dieser Forschungsrichtung, in Deutschland Mentalitätsgeschichte genannt, stammt von dem französischen Historiker Fernand Braudel (Braudel, F.(1949)). Über ein geographisches und historisches Subjekt, das Mittelmeer, werden zur Analyse alle Arten von Informationen herangezogen, so daß unter anderem Klimadaten und Wirtschaftsdaten in Abhängigkeit gesetzt werden können.
- 4 Informationen zu ToolBook: Hersteller ist die Asymetrix Corporation, 110-110th Avenue N.E.Suite 717, P.O.Box 40419, Bellevue, Washington 98004-0419, Phone: 001-206-637-1600. Distributor für Deutschland ist die ADI Software GmbH in Karlsruhe. Auf dem deutschen Software-Markt kostet ToolBook 1.5 zwischen 1000 bis 1500 DM und ist in einer englischen, deutschen, französischen, italienischen oder schwedischen Version erhältlich. Als Hardware muß ein IBM-kompatibler PC mit minimal 4 MB RAM, 386SX Prozessor (16 Mhz) und ein VGA Monitor sowie eine Mouse eingesetzt werden. ToolBook und die Applikation brauchen wenigstens 20 MB auf der Festplatte.
- 5 Kuhlen (1991)
- 6 Nach Bibliographie Annuelle (1988)
- 7 Informationswissenschaftlich gesprochen hat dieser node ein n:m Verhältnis.
- 8 Lüsebrink (1989)
- 9 Hypertext-Elemente in der Datenbank-Datei können erklären, welche berufliche Tätigkeit als "Schriftsteller" und welche Städte als "Norddeutschland" kategorisiert werden.
- 10 Französische Revolution, Bd.1 (1985)
- 11 Französische Revolution, Bd 4 (1985)
- 12 Ruiz (1982)

## Literatur

- Bibliographie Annuelle (1988): Bibliographie Annuelle de L'Histoire de France du cinquième siècle à 1958, Année 1988. Paris 1989.
- Braudel, F.(1949): La Méditerranée et le monde méditerranéen à l'époque de Phillippe II. Paris 1949
- Französische Revolution, Bd.1 (1985): Berichte und Deutungen deutscher Schriftsteller und Historiker (hg. v. H. Günther), Bd 1 Augenzeugen, Frankfurt 1985, S. 9-102
- Französische Revolution, Bd.4 (1985): Berichte und Deutungen deutscher Schriftsteller und Historiker (hg.v. H. Günther), Bd 4 Kommentare, Frankfurt 1985, S. 1253-1284
- Kuhlen, R. (1991): Hypertext. Ein nicht-lineares Medium zwischen Buch und Wissensbank. Berlin 1991
- Lüsebrink, H.J.(1989): Die »Bastille«. Zur Symbolgeschichte von Herrschaft und Freiheit, Frankfurt 1989, S. 55
- Mylonas, E. / Heath, S. (1990): Hypertext from the Data Point of View: Paths and Links in the Perseus Project, in: Hypertext: Concepts, Systems, Applications (hg. v. A. Rizk). Cambridge 1990, S. 324-336
- Nielsen, J. (1990): Hypertext & Hypermedia. Boston 1991
- N.N (1990): Das Europäische Museums-Netzwerk. AGE 7, Frühjahr 1990, S.12-15

Pfeiffer, St. (1990): Der PC als Multi-Media-Maschine? PC Praxis 1990, 12, 34-37

Ruiz, A. (1982): Deutsche Reisebeschreibungen über Frankreich im Zeitalter der Französischen Revolution (1789-1799) in: Reiseberichte als Quellen europäischer Kulturgeschichte, (hg.v. A. Maczak; J.Teiteberg), Wolfenbüttler Forschungen 21, Wolfenbüttel 1982

Tai, Th. (1990): Daten-Lotsen: Plus 2.0, ToolBook, Guide 3.0 Hyperpad. CHIP 1990, 11, 247-254

Thaller, M. (1988): Gibt es eine fachspezifische Datenverarbeitung in den historischen Wissenschaften; in: Geschichtswissenschaft und elektronische Datenverarbeitung, Wiesbaden 1988, S. 45-83

von Hoven, Ch. (1990): Hypercards für WINDOWS. TOOL, 1990, 12, 28-30

Warnke, M. (1990): Thema ist die ganze Welt: Hypertext im Museum, in: Hypertext und Hypermedia (hg. v. P.A. Gloor; N.A. Streitz), Berlin 1990, S. 268-278

## **Verfasser**

Schröder, Thomas A.

Informationswissenschaft /

Forschungsabteilung für philosophische Information und Dokumentation

Historisches Seminar / Lehrstuhl für Frühe Neuzeit

Universität Düsseldorf

Universitätsstraße 1 (Geb. 23.21.04)

D - 4000 Düsseldorf 1

# Die Wissensrepräsentation in der Software Technologie auf der Basis des ereignisorientierten Ansatzes

## Zur Integration der strukturierten Analyse und des Entity-Relationship-Modelling

Autor: Herr Kulawiak, Frankfurt

### Zusammenfassung / summary:

Der ereignisorientierte Ansatz als Schritt-für-Schritt-Verfahren wird so weit operationalisiert, daß er als Algorithmus für ein wissensbasiertes Werkzeug dienen kann. Voraussetzung für den gewünschten Leistungsumfang des Werkzeugs ist die Integration der strukturierten Analyse (SA) und dem Entity-Relationship-Modelling (ERM). Durch das Werkzeug werden aufgrund von Konfigurationsdaten Funktions-, Datenmodelle und die Testfälle des Systems erzeugt.

The event orientated approach as a step-by-step-procedure is operationalized to the extend that it can serve as algorithm for a knowlegde based tool. Prerequisite for the desired performance capacity of the tool is the integration of Structured Analysis (SA) and Entity-Relationship-Modelling (ERM). Based on configuration data the tool produces funktion- , data diagrams and test cases of the system.

### 1. Einleitung

Der ereignisorientierte Ansatz (essentielle Methode) wird soweit spezifiziert und formalisiert, daß er programmtechnisch zur automatischen Erzeugung von Dokumenten der Systementwicklung (Funktionsmodell, Datenmodell, Testfälle) eingesetzt werden kann. Die Ausführungen gliedern sich in zwei Hauptteile:

Die Logik der Vorgehensweise, die sich am ereignisorientierten Ansatz orientiert, wird in Hinblick auf einen formalisierbarer Algorithmus dargestellt. In bestimmten Regeln und Verfahren liegen daher Erweiterungen und Abweichungen gegenüber den publizierten Konzepten /1/ vor. Dabei ist zu beachten, daß die Regeln und Verfahren dieser Vorgehensweise auch ohne maschinelle Unterstützung anwendbar sind.

Der Leistungsumfang und der Aufbau des mit Mitteln der Wissensrepräsentation arbeitenden Tools (Arbeitstitel: ISE (Integrierte System Entwicklung)) werden dargestellt. Die Logik des Werkzeugs, das sich im Entwicklungsstadium eines Prototyps befindet, basiert auf der beschriebenen Vorgehensweise.

### 2. Zielsetzung der Vorgehensweise

Die Konzeption zielt auf die Erhöhung der Produktivität und Qualität der Software-Entwicklung, wobei folgende Aspekte in diesem Zusammenhang eine besondere Bedeutung besitzen:

Ein Mittel zur Erreichung dieser Ziele ist die Gestaltung von Teilergebnissen einer Projektphase in einer Weise, daß sie für darauf folgende Phasen eines Projekts bzw. für andere Projekte wiederverwendbar sind. Diese **Wiederverwendbarkeit** wird in verschiedener Hinsicht erreicht:

Alle Ergebnisse der Systementwicklung werden in einer Weise strukturiert, daß es möglich ist, sie in späteren Projektphasen wiederzuverwenden (Wiederverwendung innerhalb des Projekts). Das setzt eine straffe Zielorientierung der konkreten Projektarbeit voraus. Die Darstellung der Ergebnisse, die Definition der Begriffe muß sich an ihrer Zweckmäßigkeit zur Erreichung des Projektziels messen lassen. Gerade in der Projektpraxis werden immer wieder gleiche oder logisch abhängige Sachverhalte in verschiedener Form und unabhängig voneinander formuliert bzw. modelliert. Somit werden unnötige Aufwendungen verursacht. Diese Redundanz der Entwicklungsarbeit wird nicht zuletzt durch das Nebeneinander von ERM und SA bewirkt, indem in der Projektpraxis häufig das Daten- und Funktionsmodell organisatorisch und zeitlich getrennt erarbeitet werden.

Deshalb zielt das vorgestellte Verfahren auf die Integration des **Entity-Relation-Modelling (ERM)** und der **strukturierten Analyse (SA)**, damit die Konsistenz der verschiedenen Entwicklungsergebnisse gewährleistet ist.

Die Wiederverwendbarkeit über ein konkretes Projekt hinaus kann durch die Kategorisierung von Sachverhalten erreicht werden, indem funktionale bzw. attributive Gleichartigkeiten bzw. Ähnlichkeiten verschiedener Bereiche auf einem passender Abstraktionsniveau modelliert werden, um eine **redundanzarme Gesamtentwicklungsarbeit** zu erreichen.

Die **interpersonelle Transparenz und Reproduzierbarkeit** von Entwicklungsergebnissen muß vorhanden sein, um eine effektive Projektkommunikation und eine gute Wartbarkeit des software-Systems sicherstellen zu können. Dies wird durch das **operationalisierbare Vorgehensmodell** und dessen dokumentierten Ableitungen unterstützt.

Eine effiziente Wartbarkeit eines Software-Systems ist ein wichtiges Qualitätsmerkmal für das erarbeitete Werk. Eine durchgängige Modularisierung, eine strukturelle Unabhängigkeit der verschiedenen Systemkomponenten erhöhen die Qualität der erstellten Software beträchtlich.

### **3. Vorgehen /2/**

#### **3.1 Geschäftereignisse**

Die Organisation und Dynamik betriebswirtschaftlicher Systeme (sowohl der marktorientierten als auch der infrastrukturellen Bereiche) ist durch die Außenwelt bestimmt. Die Systemaußenwelt wird durch die als relevant erachteten Geschäftereignisse und die externen Informationsanforderungen repräsentiert, die struktur- und aufgabenbestimmend sind.

Geschäftsereignisse sind die externen Anlässe, die Vorgänge im System anstoßen. In freiem Text beschrieben sind sie die **Kommunikationsschnittstelle** zu den entsprechenden Ansprechpartnern in den Fachorganisationen, auf deren Basis sich strukturierte Interviews führen lassen. Da aufgrund des hier beschriebenen Vorgehens alle Systemergebnisse ableitbar sind, können auch die Zusammenhänge anschaulich und transparent erörtert werden.

Systematisch können Geschäftsereignisse folgendermaßen unterschieden werden:

- operationelle Geschäftsereignisse** (Aufträge, Einzahlungen)
- Geschäftsereignisse aus Überwachungsgründen** (Protokollisten)
- Geschäftsereignisse aus Bewertungssicht** (Kostenvoranschläge)
- Informationsanforderungen externer Stellen** (Meldungen).

**Informationsanforderungen** werden mit dieser Interpretation als Geschäftsereignis schon frühzeitig in die Analyse einbezogen. Dies hat den Vorteil, daß der Datenbedarf für Informationen frühzeitig untersucht wird und über die Funktionsanalyse im Datenmodell eingebracht werden kann. Der Zusammenhang zwischen Informationen und den zugrunde liegenden Daten werden im Informations-Modell sichtbar gemacht /3/.

### 3.2 Ereignisse / Reaktionen

Geschäftsereignisse werden von Systemen in der Form von **externen Ereignissen** wahrgenommen. Dabei kann ein Ereignis als Nachricht oder als Datenkonstellation (bei internen und Schwellenwert-Ereignissen) von dem entsprechenden System verarbeitet werden.

Die Ereignisse stoßen Vorgänge an, die von **Prozessen bzw. Funktionen** /4/ bearbeitet werden müssen. Diese Prozesse erzeugen verschiedene **Reaktionen**. Diese können folgendermaßen unterschieden werden:

Sie sind an Terminatoren / externen Partner als Informationen oder Nachrichten gerichtet und verlassen das System.

Sie stoßen als **interne Ereignisse** andere Funktionsbereiche an. Diese internen Ereignisse können in unterschiedlicher Weise diesen Anstoß vollziehen:

Ein Ereignis kann mehrere Funktionsbereiche adressieren (Möglichkeit der Parallelverarbeitung).

Verschiedene Ereignisse können den gleichen Funktionsbereich anstoßen (Zusammenführung von Ereignissen).

Ein Ereignis stößt unter bestimmten Bedingungen unterschiedliche Funktionsbereiche an (Differenzierung des Ereignisses).

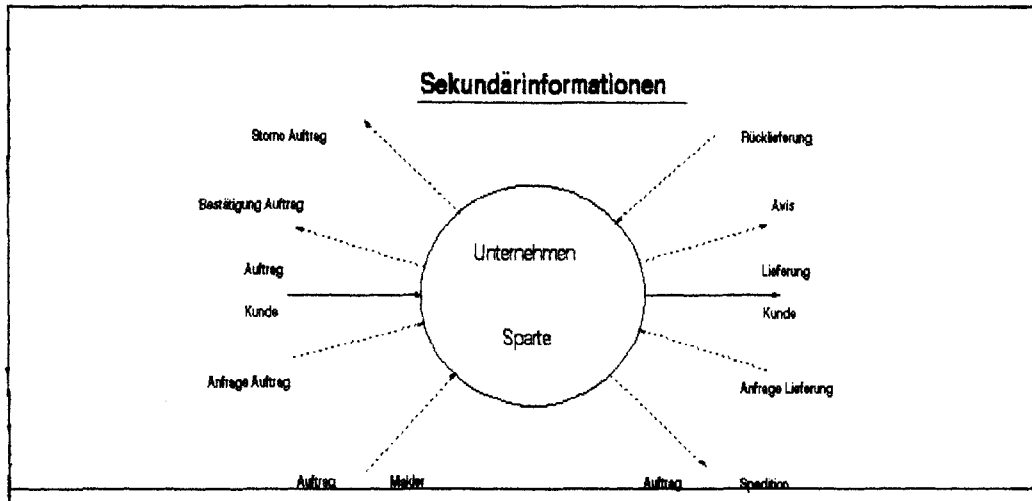
Ein Vorgang stellt sich somit als verkettete Ereignis- / Prozessstruktur dar, deren Grenze dadurch gekennzeichnet ist, daß

nur noch Reaktionen von dem Prozess erzeugt werden, die an externe Partner gerichtet sind.

Neben den externen Ereignissen und den internen Ereignissen können noch folgende zusätzliche Ereignistypen definiert werden:

Die **Schwellenwert-Ereignisse** (aufgrund erreichter Termine oder Limite (Verkaufslimit, Mindestbestellmenge u.ä.).

Die **abgeleiteten Ereignisse**, die aufgrund von logischen oder auf Erfahrung basierenden Überlegungen in Abhängigkeit der zuvor genannten (Primär-)Ereignissen definiert werden (siehe folgende Zeichnung).



Alle Funktionen werden daraufhin untersucht, welche Daten zur Erledigung der funktionalen Erfordernisse aus der Verarbeitung der Ereignisse und der Erzeugung der Reaktionen notwendig sind.

Die von den Ereignissen angestoßenen Prozesse sind die **Verarbeitungsprozesse**. Diese Prozesse sind die Top-Funktionen des Funktionsmodells.

### 3.3 Funktionsmodell

Ein **Funktionsmodell** (siehe folgende Zeichnung) hat innerhalb der beschriebenen Vorgehensweise folgende Funktionen :

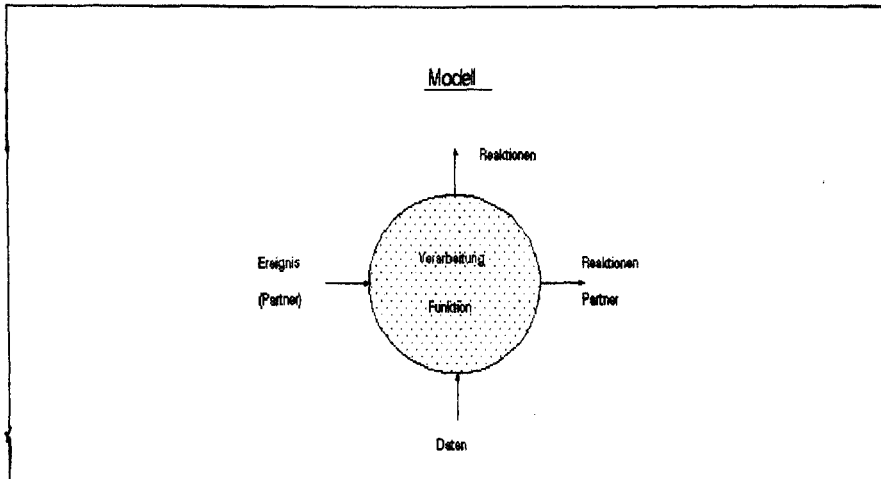
Es ist die Basis für die Programm- bzw. Modulstruktur des Systems. Alle Ergebnisinformationen, die für die Programmstruktur zu gebrauchen sind, sollten abgelegt werden (Wiederverwendbarkeit).

Es ist die Hauptdokumentations- und Kommunikationsmittel für die Systementwicklung.

Es ist die Hauptdatenbasis für die Ableitung der Testfälle (Wiederverwendbarkeit).



An der Struktur des Funktionsmodell sollte sich die Gliederung der sprachlichen Darstellung des System orientieren.



Aufgrund der vorliegenden Informationen können folgende funktionalen Ableitungen zur Erzeugung einer Funktionsstruktur getroffen werden:

**Empfangs- / Lesefunktionen oder Prüffunktionen**, die der spontanen Hülle zugeordnet sind, können auf der Basis der Ereignisnachricht abgeleitet werden. Dabei können unterschiedliche Funktionsstrukturen entwickelt werden, abhängig davon, ob die Ereignisse Nachrichten- oder Datencharakter besitzen. Analoges gilt auch für die Ableitung der Reaktionsfunktionen.

Ebenso können für alle Reaktionen nacheinander die **Ausgabenerzeugungsfunktionen** logisch abgeleitet werden und im Funktionsmodell gespeichert werden. Dabei ist wie bei der Funktionsdifferenzierung zwischen ODER- und UND-Reaktionen zu unterscheiden, die ihre Entsprechung in der abgeleiteten Funktionshierarchie finden muß.

Die für den Verarbeitungsprozess benötigten Daten müssen durch **Datenbeschaffungsfunktionen** bereitgestellt werden. Logisch kann daraus aber auch abgeleitet werden, daß benötigte Daten im System gespeichert werden müssen. Damit werden auch die **Datenspeicherfunktionen** ableitbar. Die für die Verarbeitung benötigten Daten sind bei einem hinreichenden Detaillierungsgrad als Attribute einer Entität zu benennen. Diese Analyse ist die Hauptbasis der Erarbeitung des Datenmodells (siehe unten).

Die Verarbeitungsprozesse, die durch die Ereignisse angestoßen werden, können weiter differenziert werden. Dabei wird in zweierlei Hinsicht die **Funktionsdifferenzierung** vorangetrieben:

Die Steuerattribute des Ereignisses führen zu alternativen Prozessen, die entweder Menue- oder Entscheidungscharakter haben können.

Diese Prozesse werden verrichtungs- bzw. objektorientiert analysiert und differenziert, wobei ablaforientierte Erkenntnisse festgehalten werden sollten. Die Analyse der ereignisverarbeitenden Prozesse führt zur Aufdeckung funktionaler und attributiver Redundanzen in den verschiedenen Funktionsbereichen, die durch Modulbildungen abgefangen werden können.

Die Differenzierung der Funktionen kann weitere Erkenntnisse über deren Datenbedarf ergeben, so daß wiederum funktionale Ableitungen vorgenommen können (siehe oben).

Eine solche Differenzierung innerhalb des Funktionsmodells muß auf einer Funktionsebene klar definiert sein. Es gibt folgende **Ebenenkategorien im Funktionsmodell:**

Die Menue-Ebene sind alternative Ausführungspfade innerhalb des Systems, die durch externe Anstöße gestartet werden. Externe Ereignisse (aufgrund von Geschäftsereignissen) stoßen daher immer Menue-Funktionen an. Formal sind diese Funktionen (gegenüber der Entscheidungsebene) durch das Fehlen von Regeln gekennzeichnet. Innerhalb der Funktionsdifferenzierung sind Menuefunktionen dann zweckmäßig, wenn Systementscheidungen aufgrund von Daten gefällt werden (z.B. vom Benutzer), die der Systemwelt nicht zur Verfügung stehen.

In der Entscheidungsebene (ODER-Ebene) werden Funktionen über Regelbedingungen angestoßen. Diese Regeln werden der Abhängigkeitsbeziehung zwischen den Entscheidungsfunktionen und der entsprechenden Kopffunktion zugeordnet. Die Zweckmäßigkeit dieser Ausgestaltung der Regeln liegt darin begründet, daß damit die testrelevante Systemstruktur formalisiert gespeichert wird.

Die Ablauf-Ebene (UND-EBENE) stellt die sequentielle Ausführung eines Teils des Vorgangs auf der entsprechenden Differenzierungsebene dar. Einer bestimmten Funktion kann eine Regel zugeordnet werden, wenn sie innerhalb des Ablaufs nur bedingt ausgeführt wird.

Es gibt - wie oben ausgeführt - damit zwei Möglichkeiten, die **funktionale Differenzierung** eines Systems zu erreichen:

Die verrichtungs- bzw. objektorientierte funktionale Differenzierung, die in einem Funktionsbaum unmittelbar ablesbar ist.

Die über interne Ereignisse verkettete bzw. angestoßene Funktionsbereiche (Modelle). Die Struktur eines so modellierten System erinnert an ein Netzwerk von (biologischen) Zellen, die über Nachrichten (Transmitterstoffe) kommunizieren.

Beide Möglichkeiten sind ergänzende Alternativen. Innerhalb der Modelle werden die Verarbeitungsfunktionen, die interne Ereignisse bearbeiten, funktionell differenziert, so daß das Gesamtsystem immer als Funktionsmodell abgebildet werden kann.

Die Wahl einer bestimmten Alternative bei der Darstellung eines Sachverhalts sollte aufgrund der konkreten Situation entschieden werden. Mit beiden Alternativen ist jedenfalls eine modulare Systemstruktur erarbeitbar und abbildbar.

Die Funktionsdifferenzierung sollte soweit wie möglich und sinnvoll heruntergebrochen werden. Die Zweckmäßigkeit der Detaillierung ergibt sich aus der Funktion des Funktionsmodell innerhalb der Systementwicklung und sollte so gestaltet sein, daß das Funktionsmodell eine geeignete Datenbasis für die Gestaltung der nachfolgenden Systementwicklungs-Ergebnisse dienen kann. (siehe oben).

### 3.4 Datenmodell

Das logische Datenmodell ist die Systemdarstellung aus Sicht der Daten und muß als solches in engem Zusammenhang mit dem Funktionsmodell stehen. Eine parallele, verzahnte Erarbeitung ermöglicht, den gleichen Grad an Detaillierung und Konkretisierung und damit ein Höchstmaß an Konsistenz zu erreichen.

Das logische Datenmodell dient als Basis des physischen Modells (ein Datenmodell unter der Berücksichtigung von Implementations- und Performanceerfordernissen) und schließlich der konkreten Datenbank.

Im Datenmodell werden die systemrelevanten Entitäten und ihre Beziehungen untereinander dokumentiert. Die Beziehungen zwischen den Entitäten können als systemspezifische Zugriffswege definiert werden. Als solche decken diese aus Datensicht den funktionalen Aspekt ab und müssen sich im Funktionsmodell wiederfinden. Umgekehrt sind die notwendigen Zugriffswege auf die benötigten Daten im Datenmodell abzubilden.

Folgende Anlässe zur Bildung von Entitätenbeziehungen können aus der Logik des ereignisorientierten Ansatzes systematisch definiert werden:

Der Zusammenhang zwischen den Datenentitäten bei der Modellbildung.

Die Zusammenführung bzw. Trennung von Ereignissen, soweit sie Datencharakter besitzen.

Die Verallgemeinerung bzw. Differenzierung von Partnern und Objekten führen zu Spezialisierungs-Beziehungen

Bei der Attributanalyse werden durch die Anwendung der Normalisierungsregeln neue Entitätsbeziehungen aufgebaut.

Aus der ereignisorientierten Analyse auftauchende M : N - Beziehungen werden mit assoziativen Entitäten aufgelöst.

### 3.5 Begriffskategorisierung

Die wesentlichsten Bedingungen für die Möglichkeit der Wiederwendbarkeit von Systembauteilen ist die semantische Disziplin und die Kategorisierbarkeit von relevanten Sachverhalten.

Unter semantischer Disziplin wird die einheitliche, bewusste Verwendung von Begriffen zur Benennung relevanter Sachverhalte verstanden. Dabei können gewohnte Begriffe über Synonymbildung verkettet werden, deren Inhalt sich mit dem des Schlüsselbegriffs decken müssen.

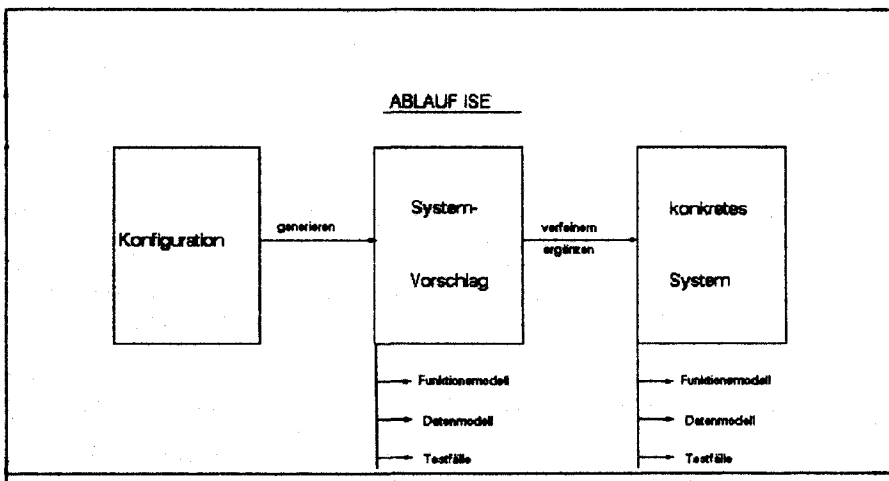
Wichtig für die Begriffsbildung in der methodisch unterstützten Systemverwaltung ist die **Zweckbestimmung des einzelnen Begriffs** für die Systementwicklung. Die funktionale zielorientierte Relevanz steht auf der Basis fachlicher bzw. technisch-organisatorischer Sachverhalte im Vordergrund. Funktionale oder attributive Gleichheit bzw. Ähnlichkeit müssen sich in der begrifflichen Fassung des Sachverhalts wiederfinden.

Diese Kategorisierung faßt begrifflich ähnliche bzw. gleiche Strukturen begrifflich zusammen. Dabei kann eine Kategorisierung über mehrere Ebenen **hierarchisch** geführt werden, um spezifische Funktionalitäten abbilden zu können.

Alle Sachverhalte, die auf einem beliebigen abstrakten Niveau Gleichartigkeiten besitzen, sind Kandidaten für die Kategorienbildung. Diese Kategorien sind der **Schlüssel zur Verkettung** schon erarbeiteter Systembausteine in das zu entwickelnde System und damit ein wesentliches Element der Wiederverwendbarkeit. Dabei sind auch zusammengesetzte Begriffe als **Schlüssel für Funktions- und Datenableitungen** einsetzbar.

#### 4. Beschreibung des Tools ISE

Mit dem Tool wird versucht, die zuvor skizzierte Logik in eine maschinelle Logik umzusetzen und die Regeln und Verfahren so zu gestalten, daß **automatisch die genannten Systemergebnisse** erzeugt werden (siehe folgende Zeichnung).



Das Werkzeug umfaßt drei wesentliche Leistungsmerkmale:

Die Pflege der Konfigurationskonstrukte von Systembausteinen und die Generierung der Systemdatenbank aus diesen Konfigurationsbausteinen  
(Funktionsbereich: Konfigurationsmanagement).

Die Pflege der Systemdatenbank, um spezifische Systemdaten, deren Wiederverwendbarkeit nicht offensichtlich ist und eine Konfigurierung zu aufwendig wäre (Funktionsbereich: **Systempflege**).

Die Erzeugung der wesentlichen Informationen bzw. Datenschnittstellen aus der Systemdatenbank (Funktionsbereich: **Auswertungen**).  
Folgende Informationen werden automatisch erzeugt:

#### **Funktionsmodell**

Diese Informationsmenge entspricht dem decomposition diagram von IEW und kann in diese überführt werden.

#### **Daten-/Info-Modell**

Das Datenmodell zeigt die Beziehungen zwischen Entitäten nach ER-Konstruktionsmerkmalen auf. Es entspricht dem entity diagram im IEW und kann in dieses überführt werden.

#### **Liste der Systemtestfälle**

Aufgrund der Daten der Systemdatenbank werden aufgrund eines Algorithmus die vollständige Liste der Systemtestfälle erzeugt. Mit diesen Testfällen können alle möglichen Zustände des Systems getestet werden, wenn sie in einer geeigneten technischen Form erfaßt werden. Die Qualität dieser Auswertung hängt naturgemäß von der Qualität der Systemdaten ab.

Grundsätzlich können auch alle dem Werkzeug IEW übertragbare Daten indirekt in Datenmanagement-Systeme übertragen werden.

Im Funktionsbereich Konfigurationsmanagement werden solche Bausteine angeboten, die wiederverwendbare Sachverhalte adressiert über Schlüsselbegriffe bzw. über Kategorien abspeichern können. Diese Datenmuster werden projekt- bzw. bereichsunabhängig gespeichert und bei jeder Generierung spezieller Systeme auf ihre Verwendbarkeit geprüft. Die wichtigsten Bausteine sind folgende:

#### **Modell**

Im Modell wird entweder geschlüsselt über das Ereignis (Initialvorgang) oder über die Verarbeitungsfunktion alle nach / von außen gerichteten Datenströme (Ereignisse, Reaktionen, Daten) in Beziehung zum Verarbeitungsprozess gesetzt. Das Modell stellt einen abgeschlossenen, separaten Funktionsbereich dar. Ein Modell wird im Generierungsprozess zur Erzeugung eines konkreten Systems einbezogen, wenn die Reaktion eines aktiven Modells kategorisch mit dem Ereignis (Initialvorgang) des Modellkandidaten übereinstimmt.

#### **Funktionsableitung**

Mit der Funktionsableitungs-Relation werden Muster von Funktionsdifferenzierungen abgebildet.

#### **Informations-Muster**

Mit diesem Baustein wird der Datenbedarf der Informationen gespeichert, soweit sie in mehreren Kontexten auftauen oder eine abstrakte Basis abbilden

### **Attributmuster**

Wiederkehrende Attributfolgen in Nachrichten, Entitäten oder Datengruppen können damit abgebildet werden. Die Regeln der Normalisierung werden auf dieser Datenbasis bei der Generierung des konkreten Systems angewandt.

Der Prozess der Generierung der Systemdatenbank kann beliebig oft durchgeführt werden. Damit wird auch der **dynamische Aspekt der Systementwicklung** unterstützt, daß aufgrund von Zwischenergebnissen der Generierung weitere Informationen anfallen, die wiederum Datenbasis der Generierung werden können.

Bei Erzeugung des konkreten System geschieht unter Beachtung der spezifischen Begriffe des analysierten Sachverhalts. Diese Synonym-Begriffe ersetzen die Kategorie-Kennzeichen bzw. deren Schlüsselbegriffe.

Im **Funktionsbereich Systempflege** wird das konkrete System entwickelt, auf dessen Basis die Auswertungsfunktionen die Systemdokumente erstellen. Über eine Dialogschnittstelle können die generierten Ergebnisse verfeinert und angepaßt werden. Die folgenden Bausteine, eine Auswahl der wichtigsten, werden benutzt, um relevante Systeminformationen so abzubilden, daß sie maschinell weiterverarbeitbar bzw. auswertbar sind (siehe oben):

#### **Funktions-Beziehung**

Diese Relation stellt die hierarchische Beziehung zwischen verschiedenen Funktionen dar.

#### **Dat-Relation**

Die Entitäten und ihre Attribute werden hiermit gespeichert, wobei die Gruppenstrukturen festgehalten werden.

#### **Funktionsdatum**

Diese Relation gibt die Daten-Beziehungen einer Funktion mit ihrer Umwelt wieder.

#### **Regel**

In der Regel-Relation werden die Bedingungen abgelegt, die das Systemverhalten steuern. Sie sind immer einer Funktionsbeziehung oder einer Funktion zugeordnet.

#### **Informationsdaten**

Diese Daten sind das Pendant der Informationsmuster in den Konfigurationsdaten. Sie geben die Informations-Daten-Beziehungen im konkreten System wieder.

Die Auswertungen, durch den **Funktionsbereich Auswertungen** erzeugt, erfolgen immer aufgrund der konkret erarbeiteten Systemdaten, so daß eine Änderung der Basis sich simultan in den entsprechenden Systemergebnissen niederschlägt.

### **5. Fazit**

Das Tool kann ab einem gewissen Reifegrad eine wertvolle Unterstützung des konzeptionell arbeitenden Entwicklers darstellen. Die Arbeitserleichterung, die Transparenz und Konsistenz der Ergebnisse sind eine wertvolle Hilfe durch den

Einsatz dieses Tool-Typs. Aber auch ohne den Einsatz der maschinellen Unterstützung kann die beschriebene dezidierte Verfahrensweise, die eine Schritt-für-Schritt-Entwicklung unterstützt zur Verbesserung der konkreten Projektarbeit führen.

Dipl. Soziologe  
Walter Kulawiak  
Monatabaurer Str. 13b  
6230 Frankfurt 80

#### Anmerkungen:

1 Es werden keine speziellen Literaturhinweise gegeben. Die Begriffswelt wird aus den nachstehenden Ausführungen entnommen:

James Martin: Information Engineering  
Savant Research Studies, 1986

Jörg Raasch: Systementwicklung mit Strukturierten Methoden  
Hanser Verlag, 1991.

2 Der Sachverhalt kann aufgrund des geringen zur Verfügung stehenden Raums nur in groben Zügen (weitgehend ohne Beispiele) dargestellt werden. Eine genaue und anschauliche Erörterung wird zweckmäßigerweise im Rahmen der Erstellung bzw. der Überarbeitung eines Vorgehensmodells einer konkreten Projektorganisation geschehen.

3 In dieser Aussage sind die Begriffe Daten und Information impliziert differenziert.

#### Definition:

Eine Information ist eine sinnvolle Aggregation von Attributen, die eine sinnhafte (zweckrationale) Handlung beim Empfänger auslöst (auslösen kann). Informationen sind aufbereitete Daten, die immer auf einen Empfänger gerichtet sind und ihre Relevanz aufgrund dessen Anforderungen erhalten.

Daten sind gespeicherte Sachverhalte, die unabhängig von spezifischen Informationsbedürfnissen festgehalten werden, ihre Bedeutung aber dadurch erlangen, daß mit ihnen alle spezifischen Informationsbedürfnisse der Benutzer eines Systems befriedigt werden können.

Deutlich wird aus diesen Definitionen der enge Zusammenhang bzw. Verzahnung von Daten, Funktion und Information, der auch in der methodischen Grundlage dieser Vorgehensweise beachtet wird.

4 Die Begriffe Funktion und Prozess werden als Synonym benutzt.

## Publikationsstrukturen im Ländervergleich<sup>+</sup>

Manfred Bonitz, Eberhard Bruckner, Andrea Scharnhorst

### Zusammenfassung

Es wird ein neuer Wissenschaftsindikator, Science Strategy Index, vorgestellt. Er beruht darauf, daß die wissenschaftliche Aktivität jedes Landes über die wissenschaftlichen Hauptdisziplinen gestreut ist. Diese Struktur wird zunächst mit der Weltstruktur verglichen. Der Strukturvergleich ist dabei nicht abhängig von der Größe eines Landes. Es muß eine Datenbasis über das Publikationsaufkommen der Länder zur Verfügung stehen, in der sich die Länderstrukturen widerspiegeln.

Der Indikator gestattet aber auch, die Wissenschaftsstruktur eines Landes mit der jedes anderen Landes zu vergleichen. Danach ist es möglich, in einem Clusterprozeß die Länder mit ähnlicher Wissenschaftsstruktur herauszufinden. Die auf solche Weise entstehenden Länderkarten bedürfen eingehender Diskussion über die unterschiedlichen Wissenschaftsstrategien der Länder und über die geographischen, politischen, kommunikativen und sozio-kulturellen Wurzeln dieses Phänomens.

### Abstract

A new indicator, Science Strategy Index, is proposed, which is based on the scattering of a country's science activity over all science fields. This structure can, firstly, be related to the world distribution of science fields. The indicator allows to compare the publication output of countries as reflected by the database, irrespective of the size of the countries.

The science structure of each country can, secondly, be related for comparison to that one of each other country. Thus it becomes possible to cluster countries according to their structural similarity. The cluster map of countries achieved in this way deserves intense discussion upon the different science strategies of countries and the geographic, political, communicative, and socio-cultural background of the phenomenon.

<sup>+</sup> *Beitrag für das 17. Internationale Kolloquium für Information und Dokumentation in Verbindung mit dem 2. Internationalen Symposium für Informationswissenschaft, 4.-7. November 1991, Oberhof*



## Einleitung

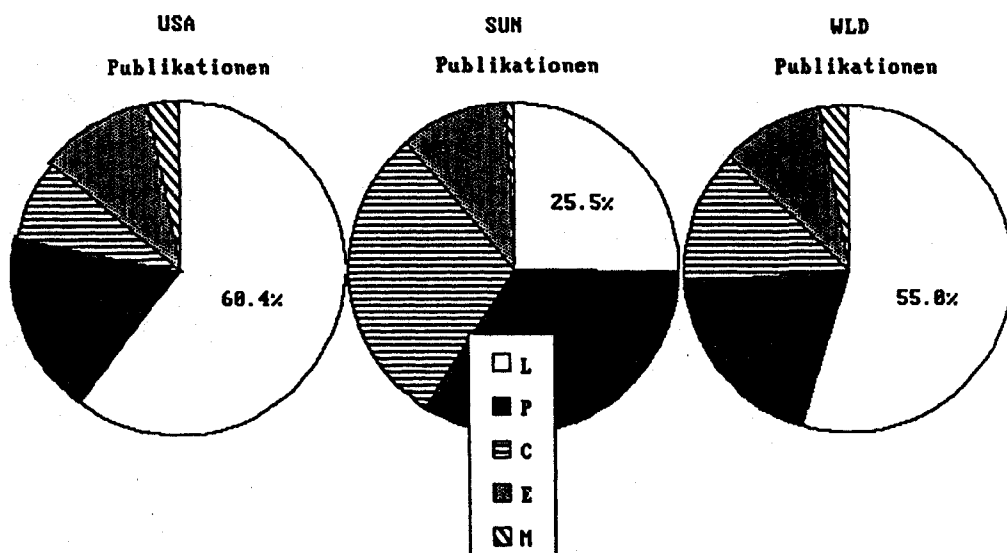
Informationswissenschaft und Wissenschaftswissenschaft sind eng miteinander verbunden. Das System der wissenschaftlichen Kommunikation ist Forschungsobjekt beider Disziplinen. Konzepte für die Beschreibung des Phänomens wissenschaftliche Information berühren sich mit Konzepten für die Beschreibung des Phänomens Wissenschaft. Ein Beispiel dafür ist die Untersuchung von Publikations- und Zitationsstrukturen und die Deutung der Ergebnisse durch die Hypothese von der Wissenschaft als einem selbstorganisierenden System. Diese Forschungen stehen am Anfang.

Zu den schwierigen Problemen bei der Prüfung der Selbstorganisationshypothese für die Wissenschaft gehört es, Strukturen im Wissenschaftssystem aufzufinden, die einfach genug sind, um noch quantitativ behandelt zu werden, aber nicht so einfach, daß wesentliche Aspekte der Wissenschaft verlorengehen.

Solch eine Struktur kann z.B. die Verteilung der Anteile wissenschaftlicher Hauptgebiete (Biowissenschaften, Physik, Chemie, Ingenieurwissenschaften, Mathematik) für ein bestimmtes Land sein (vgl. Abbildung 1), wobei diese Anteile aus dem Publikationsaufkommen dieses Landes hergeleitet werden. Auch das "Land" 'Welt' zeigt eine Struktur, die sich aus dem Weltpublikationsaufkommen ergibt.

Beim Vergleich absoluter Werte, etwa für das Publikationsaufkommen, liegen selbstverständlich die großen Wissenschaftsländer immer "vorn", die kleinen "hinten", nicht selten wird dann gefragt, welchen Beitrag kleine Wissenschaftsländer überhaupt zur Weltwissenschaft leisten können.

Ganz anders sieht es aus, wenn Strukturen, "Wissenschaftsmuster", miteinander verglichen werden. Hier sind alle Länder "gleichberechtigt", und geeignete Maße für den Strukturvergleich zu finden, wird vielleicht mit tieferen Einsichten in den Mechanismus der Wissenschaft und sinnvollen Anwendungsmöglichkeiten für die praktische Wissenschaftspolitik belohnt.



**Abbildung 1.** Die Wissenschaftsstrukturen der USA, UdSSR (SUN) und der Welt (WLD). L - Biowissenschaften, P - Physik, C - Chemie, E - Ingenieurwissenschaften, M - Mathematik

### Der Science Strategy Index

Der Science Strategy Index für ein Land  $i$ ,  $SSI_i$ , ist ein Maß für die Übereinstimmung der Landesstruktur mit der Weltstruktur, und wird durch einfachen arithmetischen Vergleich der Anteile der wissenschaftliche Hauptgebiete nach der Formel (1) bestimmt.

$$SSI_i = 100 - 1/2 \sum_f ( |P_{if} - P_{wf}| ) \quad [\%], \quad (1)$$

$P_{if}$  und  $P_{wf}$  sind dabei die Anteile von Land bzw. Welt für das Hauptgebiet  $f$ . Der Höchstwert von  $SSI_i$  ist 100%, der niedrigste Wert ist 0%. Tabelle 1 und die Rangordnung der Abbildung 2 zeigen, daß Länder mit sehr hohem Publikationsaufkommen nicht notwendig "vorn" liegen müssen beim Vergleich ihrer Wissenschaftsstruktur mit der Weltstruktur. Besonders deutlich ist das an den großen Wissenschaftsländern Indien, UdSSR und China zu erkennen.

Tabelle 1 SSI-Werte und Publikationszahlen von 45 Ländern

Land	Dreibuchstabencode	SSI	Publikationen
'Welt'	WLD	100.0%	1937470
Schweiz	CHE	97.1%	23672
Österreich	AUT	95.5%	10670
BRD	BRD	95.4%	116138
Argentinien	ARG	95.3%	5495
Canada	CAN	95.0%	82567
Frankreich	FRA	94.7%	88201
Italien	ITA	94.5%	45191
Niederlande	NDL	94.4%	33717
Belgien	BEL	94.3%	16943
Venezuela	VEN	94.2%	1421
Israel	ISR	93.9%	21130
USA	USA	93.6%	718334
Mexiko	MEX	92.0%	3413
Hongkong	HKG	91.6%	1788
Brasilien	BRA	90.1%	6885
Australien	AUS	90.1%	43589
Großbritannien	UKD	89.6%	172924
Japan	JPN	89.2%	138722
Singapur	SGP	89.0%	1057
DDR	DDR	88.3%	17143
Portugal	PRT	88.2%	1210
Irland	IRE	86.2%	3548
Saudiarabien	SAR	86.0%	1632
Türkei	TUR	85.9%	1489
Neuseeland	NZL	84.9%	9510
Griechenland	GRE	84.5%	4573
Südafrika	ZAF	84.4%	10253
Ungarn	HUN	84.1%	9406
Tschechoslowakei	CSK	83.9%	15159
Jugoslawien	YUG	83.0%	4541
Spanien	ESP	82.8%	16551
Nigeria	NIG	82.5%	3871
Norwegen	NOR	82.3%	10189
Finnland	FIN	81.9%	12443
Dänemark	DNK	80.3%	16219
Taiwan	TWN	80.1%	2594
Schweden	SWE	79.6%	32313
Indien	IND	78.8%	50015
Polen	POL	73.0%	17927
Bulgarien	BGR	72.4%	3026
Ägypten	EGY	70.8%	5088
China	PRC	70.4%	7596
UdSSR	SUN	68.8%	126613
Rumänien	ROM	64.4%	2969
Korea	KOR	60.9%	1590

Die Publikationsstrukturen wurden einer umfangreichen Arbeit von Schubert, Glänzel, Braun /1/ entnommen, welche das Weltpublikations- und Zitationsaufkommen der Jahre 1981-1985 enthält, wie es sich im *Science Citation Index* widerspiegelt.

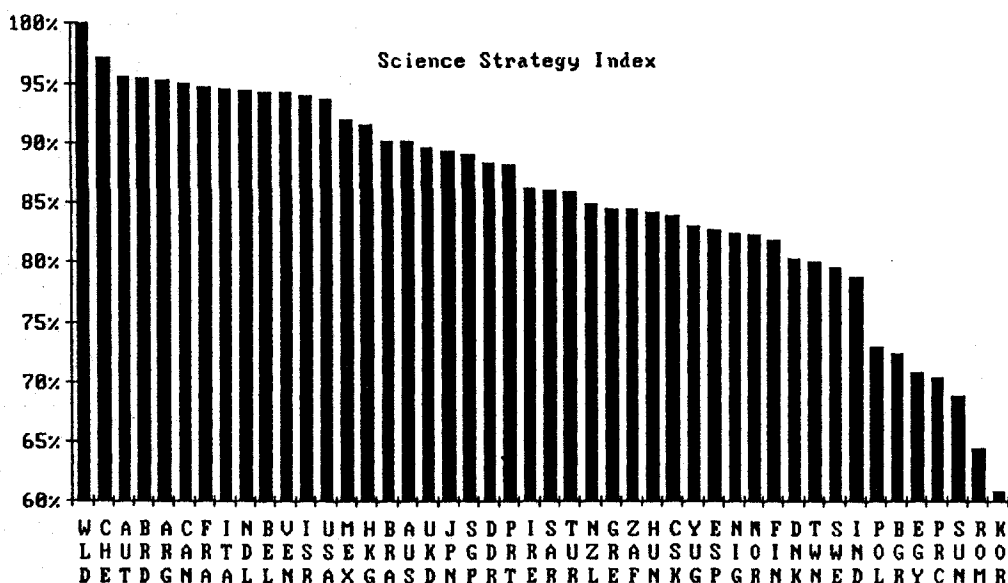


Abbildung 2. Rangordnung von 45 Ländern nach dem Wert ihres Science Strategy Index, SSI

Der Bezug einer Länderstruktur auf die Weltstruktur ist allerdings nicht einleuchtend. Warum sollte diese eine hervorgehobene Bedeutung besitzen? So haben beispielsweise Ungarn und Südafrika annähernd gleiche SSI-Werte (und sogar ein vergleichbares Publikationsaufkommen), aber völlig verschiedene Wissenschaftsstrukturen! Logischer ist doch, daß die Wissenschaftsstruktur eines beliebigen Landes mit der jedes beliebigen anderen Landes verglichen und daß für jedes Länderpaar das Ausmaß der gegenseitigen strukturellen Ähnlichkeit bestimmt wird. Dann lassen sich vielleicht Ländergruppen (Cluster) ausfindig machen, in denen die Länder eine ähnliche Wissenschaftsstruktur aufweisen.

## Der Co-Structure-Clustering-Prozeß

Für den Strukturvergleich zweier Länder braucht die Formel (1) nur dahingehend erweitert zu werden, daß anstelle der WLD-Werte für die Disziplinanteile die Werte für jedes beliebige Land eingesetzt werden können:

$$SSI_{ij} = 100 - 1/2 \sum_f ( | p_{if} - p_{jf} | ) \quad [\%], \quad (2)$$

wobei  $p_{if}$  und  $p_{jf}$  jetzt die Anteile des wissenschaftlichen Hauptgebietes  $f$  in den Ländern  $i$  und  $j$  sind.

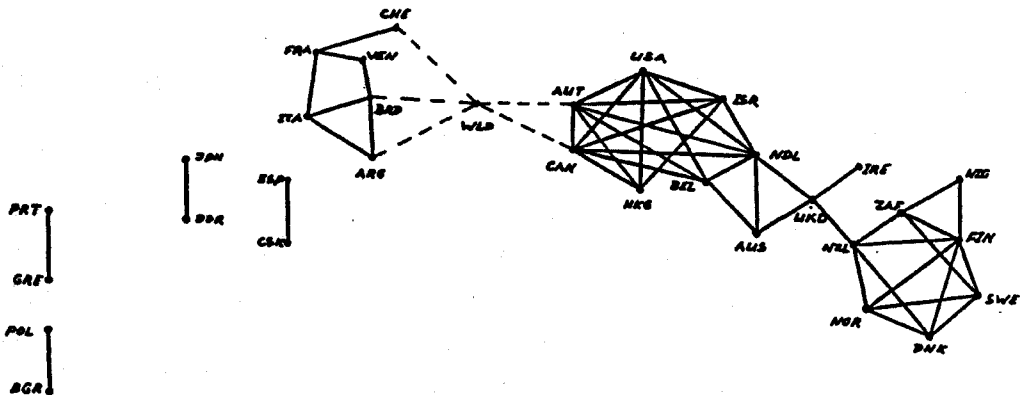
Wird SSI für alle Länderpaarungen berechnet, - was  $(46*45)/2 = 1035$  Werte ergibt -, so können diese Werte als Matrixelemente in einer Länder-Länder-Matrix Grundlage für das "co-structure clustering" bilden: Länder mit struktureller Ähnlichkeit liegen dann im Clusterbild nahe beieinander.

### Länderkarte der Wissenschaftsstrukturen

Wir haben die in Tabelle 1 angeführten 45 Länder einschließlich WLD untersucht, von denen unter der Bedingung sehr hoher Ähnlichkeit ( $SSI_{ij} > 95\%$ ) nur 31 im ersten Clusterbild (Abbildung 3) auftauchen. Innerhalb einer Gruppe weisen die Länder eine hohe strukturelle Ähnlichkeit auf, diese kann erheblich absinken, wenn die Länder unterschiedlichen Gruppen angehören, wie das genannte Beispiel Ungarn-Südafrika zeigt.

In der Abbildung 3 lassen sich deutlich unterschiedliche Gruppen von Ländern ausmachen. Auf den ersten Blick ist zu sehen, daß sich alle skandinavischen Länder, Schweden, Dänemark, Norwegen und Finnland in einem Cluster befinden. Eng beieinander liegen Südafrika und Nigeria. Die Commonwealthländer England, Australien, Neuseeland und Irland haben sehr ähnliche Publikationsstrukturen. Belgien und Holland bilden eine Art Ähnlichkeitsbrücke zu dem größten Cluster, in dem sich außer den bedeutenden

Wissenschaftsländern USA und Canada auch Österreich, Hongkong und Israel befinden. Hohe Ähnlichkeit mit der Weltstruktur weisen die



**Abbildung 3.** Clusterbild für Länderpaare mit  $SSI_{ij}$ -Werten von mindestens 95%.

fünf Länder Österreich, Canada, BRD, Argentinien und Schweiz auf. Wie selbstverständlich ähneln sich Frankreich, Italien und die Schweiz, aber auch Argentinien und Venezuela. Das Bild zeigt aber auch vier isolierte Länderpaare: Spanien-Tschechoslowakei, Japan-DDR, Portugal-Griechenland, Polen-Bulgarien. Die großen Wissenschaftsländer UdSSR, Indien und China fehlen in diesem Bild.

Ist diese Konstellation zufällig? Vielleicht durch das verwendete Ähnlichkeitsmaß  $SSI_{ij}$ , oder durch Meßfehler, oder durch die benutzte Datenbasis bedingt? Diese Fragen sind noch nicht endgültig beantwortet, sie bedürfen eingehender Untersuchungen.

An dieser Stelle wird lediglich auf einen Sachverhalt hingewiesen, der den unterschiedlichen Charakter der verschiedenen

Ländergruppen sehr grob erklären kann. Diese Unterschiedlichkeit ist durch das Verhältnis bestimmt, das sich in einem Land zwischen den Biowissenschaften und den übrigen wissenschaftlichen Hauptgebieten herausgebildet hat. Beispielsweise besitzen alle skandinavischen Länder (Schweden, Norwegen, Finnland, Dänemark) einen dominierenden Anteil von Biowissenschaften, während das entgegengesetzte Bild - dominierender Anteil der "klassischen" Wissenschaften - von Polen und Bulgarien, und bei weiterem Absenken der Ähnlichkeitsschwelle, von der UdSSR, China, Indien, Korea u.a. geboten wird. Ein stark zu den Biowissenschaften neigendes Übergewicht weisen die Commonwealthländer England, Australien, Irland, Neuseeland und Südafrika auf, weniger ausgeprägt ist dies bei Belgien und Holland, immer noch deutlich bei der Gruppe USA, Canada, Österreich, Hongkong und Israel, eine ungefähre Balance zeigen die Länder BRD, Frankreich, Italien sowie Brasilien, Argentinien und Venezuela.

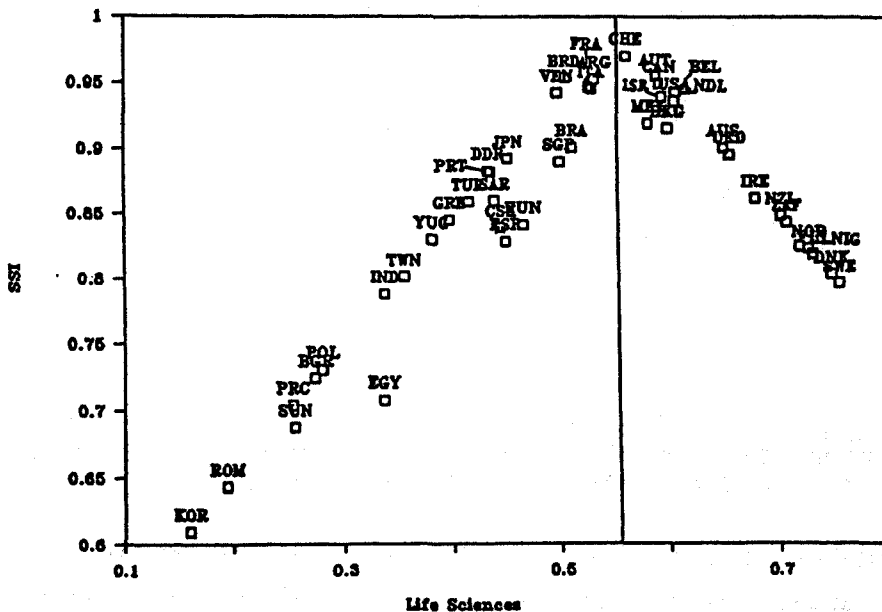


Abbildung 3. Für jedes Land ist der SSI-Wert, bezogen auf WLD, über dem Landesanteil an den Biowissenschaften (Life Sciences) aufgetragen. Die senkrechte Linie entspricht dem Weltanteil der Biowissenschaften.

In Abbildung 4 wiederholt sich dieser Sachverhalt gleichsam. Als "Länderpunkte" zeigt sie die Wissenschaftsstrukturen der Länder: weiß erscheint der Anteil der Biowissenschaften, schwarz der Anteil aller übrigen wissenschaftlichen Hauptdisziplinen. Außerdem sind (bewirkt durch das Absenken der Ähnlichkeitsschwelle) alle in Abbildung 2 noch fehlenden Länder vertreten. Die Schlingelinien kennzeichnen bei diesen Ländern die Anbindung an das "nächstähnliche" Land.

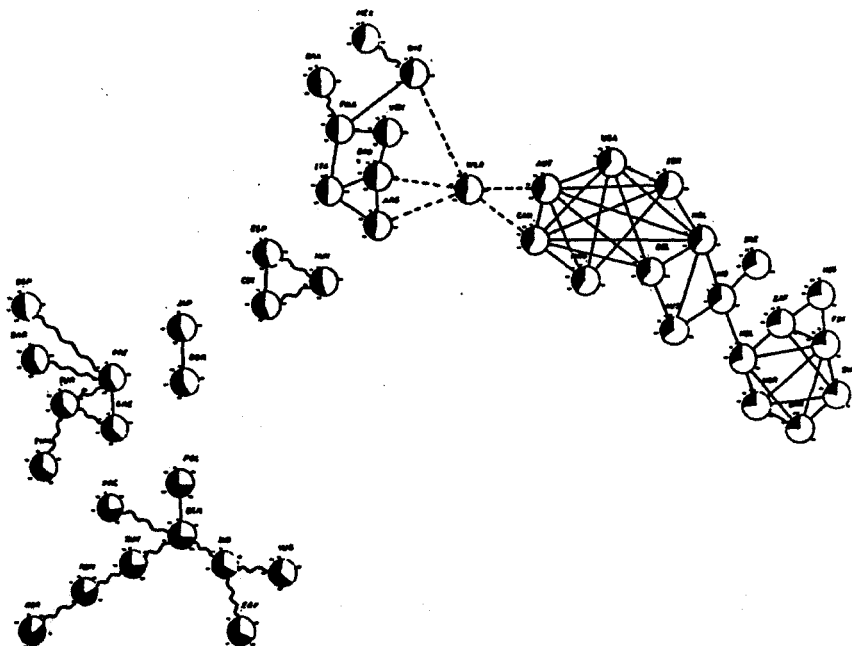


Abbildung 4. Die *Länderatome* im *Weltwissenschafts-Molekül* unterstreichen die dominierende Rolle des Verhältnisse von Biowissenschaften zu den "klassischen" Wissenschaften.

Weitere Bemühungen bei diesen Forschungen werden darauf gerichtet sein, mit unterschiedlichen Methoden und Datensätzen, aber auch durch Betrachtung verschiedener Zeitabschnitte sicherzustellen, daß es sich bei den Ergebnissen nicht um Artefakte handelt. Erst dann können die sich aufdrängenden Erklärungsvarianten sowie die Auswirkungen auf Wissenschaftswissenschaft und Wissenschaftspolitik ernsthaft erörtert werden.



Als wichtigste Schlußfolgerung für die Wissenschaftspolitik auch kleiner Länder würde uns dann erscheinen, daß nicht der Ressourceneinsatz allein, sondern vor allem die "richtige" Strukturierung der wissenschaftlichen Disziplinen Kriterium für die richtige Wissenschaftsstrategie eines Landes bilden müssen.

Die Autoren, die erste Untersuchungen zum Science Strategy Index kürzlich an anderer Stelle /2/ vorgetragen haben, bedanken sich bei R. Fisch, W. Glänzel, A. Schubert, L. Leydesdorff, U. Schoepflin und T. Braun für interessante Hinweise, die in diesen Beitrag mit eingeflossen sind.

### Literatur

- /1/ A. Schubert, W. Glänzel, T. Braun, *Scientometric Datafiles. A Comprehensive Set of Indicators on 2649 Journals and 96 Countries in all major Science Fields and Subfields, 1981-1985. Scientometrics*, 16(1989)3-478.
- /2/ M. Bonitz, E. Bruckner, A. Scharnhorst, *Science Strategy Index. Workshop on Scientometric Methods of Research Evaluation in the Sciences, Social Sciences and Technology*, April 13-17, Potsdam 1991 (im Druck).

### Autoren

Dr. rer. nat. habil. Manfred Bonitz  
Zentralinstitut für Kernforschung  
O-8051 Dresden, PSF 19

Dr. phil. Eberhard Bruckner und Dr. phil. Andrea Scharnhorst  
Institut für Theorie, Geschichte  
und Organisation der Wissenschaft  
Prenzlauer Promenade 149-152  
O-1100 Berlin

# Das Konzept der informationswissenschaftlichen Ausbildung in Ilmenau

Harald Killenberg, Ilmenau

## Zusammenfassung:

Der Beitrag beschreibt die Einordnung der informationswissenschaftlichen Ausbildung in den Studiengang Wirtschaftsinformatik an den TH Ilmenau. Ziele, Konzept und Inhalt dieses Studiums werden dargestellt.

## Abstract:

The paper describes the connection of studies in information sciences with the course of studies in business informatics at the Technische Hochschule Ilmenau (F.R.G.). Goals, concept and contents are shown.

An der Technischen Hochschule Ilmenau wurde mit dem Wintersemester 1990/91 eine neue Form der informationswissenschaftlichen Ausbildung als Studienrichtung im Studiengang Wirtschaftsinformatik eingeführt (Vorbehaltlich der endgültigen Bestätigung durch die thüringische Landesregierung).

Bisherige Formen der Aus- und Fortbildung waren

- das postgraduale Studium der Informations- und Dokumentationswissenschaft, Dauer zwei Jahre, Erwerb der Berufsbezeichnung Fachinformatiker und
- die Vertiefungsrichtung "wissenschaftlich-technische Information" im Direktstudium der Fachrichtung Informationstechnik, Dauer viereinhalb Jahre, Abschluß Diplomingenieur,

die für die immatrikulierten Jahrgänge zu Ende geführt werden /1/.

Der Studiengang Wirtschaftsinformatik seinerseits existiert in Ilmenau seit 1988 mit folgenden Immatrikulationszahlen:

Immatrikulationsjahr	Anzahl StudentInnen
1988	24
1989	23
1990	49
1991	vorauss. 80

Er wird getragen von der sich gegenwärtig in Gründung befindenden Fakultät für Wirtschafts- und Rechtswissenschaften. Bestimmte Lehrgebiete des Grundstudiums wie z.B. Mathematik und Teile der Informatik werden von anderen Fakultäten erbracht.

Die bereits laufende Ausbildung ist gegenwärtig durch Umstellungsprozesse gekennzeichnet, darauf soll hier jedoch nicht eingegangen werden.

Im Weiteren wird die ab Immatrikulationsjahrgang 1990 gültige Studienordnung /2/ zugrunde gelegt.

Wirtschaftsinformatik wird als Wissenschaft von Entwurf und Anwendung computergestützter Informationssysteme in Unternehmen verstanden. Sie ist als interdisziplinäre Wissenschaft zwischen den Wirtschaftswissenschaften, im engeren Sinne der Betriebswirtschaftslehre, und der Informatik angesiedelt.

Sie befaßt sich mit der Analyse, Planung, Bewertung, Entwicklung und dem Betrieb von Informationssystemen, die überwiegend in Wirtschaftsunternehmen, aber auch in Banken, Versicherungen, Organisationen eingesetzt werden. Sie betrachtet dabei auch die sozialen, ökonomischen und organisatorischen Voraussetzungen und Konsequenzen.

Der Studiengang Wirtschaftsinformatik ist ein wirtschaftswissenschaftlich orientierter Studiengang mit umfassender durchgängiger Ausbildung in Informatik und Wirtschaftsinformatik und führt zum Abschluß als Diplom-Wirtschaftsinformatiker/in.

Ziel der Ausbildung ist ein Absolvent, der selbständig und in Zusammenarbeit mit Informatik-Ingenieuren, Betriebswirten und Ingenieuren anderer Fachrichtungen Informations- und Kommunikationssysteme nutzerorientiert gestalten und ihre Funktionsfähigkeit gewährleisten kann.

Nach /2/ sind spezielle Ziele der wissenschaftlichen Hochschulbildung von Absolventen der Wirtschaftsinformatik Kenntnisse, Fähigkeiten und Fertigkeiten auf den Tätigkeitsfeldern

- Analyse, Beurteilung und Projektierung computergestützter betriebswirtschaftlicher Informations-, Organisations- und Steuerungssysteme
- effektive Entwicklung von Software, insbesondere in betriebswirtschaftlichen Anwendungsgebieten
- sachkundige Beschaffung, Anpassung und Wartung von Software für den Anwenderbereich
- sachkundige Präsentation und Verhandlung beim Absatz von Software
- theoretische und angewandte Forschung zur disziplinären und interdisziplinären Anwendung moderner Informationstechnologien
- konzeptionelle Vorbereitung und Durchführung von Schulungen für die Arbeit mit Informationssystemen in Unternehmen sowie zur Aus- und Fortbildung von Wirtschaftsinformatikern
- koordinierende und leitende Tätigkeit in Rechenzentren, in Fachbereichen mit computergestützter Arbeitsweise, in DV-Unternehmen, in Beratungsfirmen u.a.

Mit dieser Profilierung wird einerseits die hohe Informatik-Orientierung des Studienganges Wirtschaftsinformatik erkennbar, die ihn deutlich von anderen wirtschaftswissenschaftlichen Studiengängen mit vertiefter Wirtschaftsinformatik-Ausbildung im Hauptstudium abgrenzt. Andererseits unterscheidet sich der Studiengang Wirtschaftsinformatik durch seine Anwendungsorientiert-

heit auf einen wirtschaftswissenschaftlich einzuordnenden Gegenstandsbereich von ingenieurtechnisch ausgerichteten Informatik-Studiengängen.

Die Haupteinsatzbereiche von Absolventen des Studienganges Wirtschaftsinformatik liegen auf Grund ihres Ausbildungsprofils nach der von der Bundesanstalt für Arbeit vorgenommenen Gruppierung von mit Informationsverarbeitung befaßten Tätigkeiten in den sogenannten Mischberufen, für die eine Gleichwertigkeit von Fach- und Informationsverarbeitungsaufgaben kennzeichnend ist. Ihr Einsatz ist auch in Berufen mit dominierender Rolle der Informationsverarbeitung möglich.

In diesen "Kontext" des Studienganges Wirtschaftsinformatik fügt sich die neu konzipierte informationswissenschaftliche Ausbildung organisch ein. Sie nutzt die besonders durch die Gebiete Informatik, Wirtschaftsinformatik und Betriebswirtschaft gegebenen Voraussetzungen und vertritt ihrerseits den Aspekt der Gestaltung von Informationsprozessen in Wirtschaft, Verwaltung und Gesellschaft.

Sie vermittelt Studieninhalte zur Analyse des Informationsbedarfs von Unternehmen bzw. Einrichtungen, zur Analyse des relevanten Informationsaufkommens, zu Systemen des Informationsmarktes, zur Auswahl und Verteilung von Informationen bzw. Wissen an den Nutzer.

Sie vertieft die Inhalte zur Gestaltung und Anwendung integrierter Informationssysteme einschließlich von Methoden und Systemen der Wissensdarstellung und -verarbeitung sowie zur Nutzung von Informations- und Kommunikationstechnologien.

Das Tätigkeitsfeld von Absolventen dieser Studienrichtung umfaßt insbesondere auch die

- bedarfs- und nutzergerechte Konzeption, die Entwicklung und den Betrieb von Informations- und Kommunikationssystemen,
- die Herstellung von fachspezifischen Daten- und Wissensbasen sowie
- die Informationsvermittlung.

Die Regelstudienzeit im Studiengang beträgt einschließlich der

berufspraktischen Ausbildung und der Diplomprüfung 9 Semester. Das Studium ist gegliedert in das Grundstudium, das vier Semester umfaßt, sowie das Hauptstudium, das einschließlich der berufspraktischen Ausbildung 5 Semester umfaßt und mit der Diplomprüfung abschließt. Der zeitliche Gesamtumfang der Lehrveranstaltungen im Pflicht- und Wahlpflichtbereich beträgt 186 Semesterwochenstunden.

Die berufspraktische Ausbildung hat einen Umfang von insgesamt 26 Wochen. Sie besteht aus dem Grundpraktikum, das vor bzw. während des Grundstudiums zu absolvieren ist, und dem Fachpraktikum, das nach der Vordiplomprüfung liegt.

Die Praktika finden in Wirtschaftsunternehmen und Institutionen auf einem dem Studienziel entsprechenden Aufgabengebiet statt und umfassen jeweils 13 Wochen. Das Grundpraktikum kann seinerseits in zwei Abschnitte von 6 und 7 Wochen Dauer gesplittet werden.

Die Diplomarbeit hat eine Bearbeitungsdauer von 6 Monaten.

Das Grundstudium enthält Lehrveranstaltungen im Umfang von 100 Semesterwochenstunden (SWS).

Sie verteilen sich wie folgt auf die Lehrgebiete:

- Wirtschaftsinformatik/Informatik	27 %
- Mathematik einschl. Statistik	20 %
- Betriebswirtschaftslehre	14 %
- Volkswirtschaftslehre	12 %
- Sprachen	12 %
- Grundlagen Elektrotechnik/Elektronik	6 %
- Studium generale	6 %
- sonstiges	3 %

Im Hauptstudium kann gegenwärtig zwischen den beiden Studienrichtungen

- computerintegrierte Systeme in der Industrie und
- Wirtschafts- und Fachinformation gewählt werden. (Voraus-

sichtlich erfolgt eine Bezeichnung der informationswissenschaftlichen Studienrichtung mit "Informationsmanagement"). Sein Umfang beträgt 86 Semesterwochenstunden, die sich in folgender Weise auf vier Fächergruppen aufteilen lassen:

- Wirtschaftsinformatik/ Informationswissenschaft 40 %  
entsprechend den Studienrichtungen
- Wirtschafts- und Rechtswissenschaft 26 %
- Informatik 23 %
- Operations Research 11 %

Die Lehrgebiete der informationswissenschaftlichen Studienrichtung sind im Bild 1 ausgewiesen.

Lehrgebiet	Ges	5.	6.	7.	8.
Informationsmethodik	5	2	3		
Integrierte Informations- und Kommunikationssysteme	3	2	1		
Wissensverarbeitung	4		4		
Projektierung von Informationssystemen	3				3
Informationsmanagement 2	3			3	
Fachspezifische Informationssysteme	2			2	
Retrievalsysteme/-sprachen	5			3	2
Projektarbeit	5			2	3
Informations- und Computerrecht	2			2	
Summe	32	4	8	12	8

Bild 1 : Lehrgebiete der informationswissenschaftlichen Studienrichtung im Hauptstudium, Angaben in Semesterwochenstunden

Träger dieser Ausbildungsrichtung ist das Institut für Informationswissenschaft, das aus zwei Fachgebieten

- Informationswissenschaft/Wissensverarbeitung und
- Informationswissenschaft/Methodik

besteht. Die genannten Lehrveranstaltungen werden auch anderen Studiengängen angeboten (vgl. Bild 2).

Für die Studiengänge Wirtschaftsinformatik und Wirtschaftsingenieurwesen ist im Grundstudium eine in das Gebiet des Informationsmanagements einführende Lehrveranstaltung enthalten. In den Lehrgebieten des Hauptstudiums ist ein gemeinsames informationswissenschaftliches Laborpraktikum vorgesehen, in dem Kenntnisse anwendungsbereit vertieft und miteinander integriert werden sollen. Dieses Praktikum umfaßt ca. 8 - 10 Versuche und behandelt Themen wie Literatur-Informationssysteme, Fakteninformationssysteme, CD-ROM-Datenbanken, Online-Datenbanken, Datenmodellierung und Datenbanksysteme, Hypertextsysteme, Expertensysteme, Elektronische Publikationssysteme. Der Stundenfonds dafür ist in dem angegebenen Umfang der Lehrgebiete enthalten.

Im Rahmen der Projektarbeit werden komplexe Aufgabenstellungen ebenfalls lehrgebietsintegrierend über einen Zeitraum von einem Jahr (4. Studienjahr) bearbeitet.

Themen sind hier z.B. Abschnitte aus der Projektierung von Informationssystemen, die Erstellung von Datenmodellen, der Aufbau und die Pflege von Daten- bzw. Wissensbasen zu bestimmten Fachgebieten.

Ergänzt werden die in der Übersicht (Bild 1) angegebenen Lehrgebiete im 9. Semester mit einem Hauptseminar Informationswissenschaft im Umfang von zwei Semesterwochenstunden.



# Ausbildung in Wirtschaftsinformatik

## Institut für Informationswissenschaft

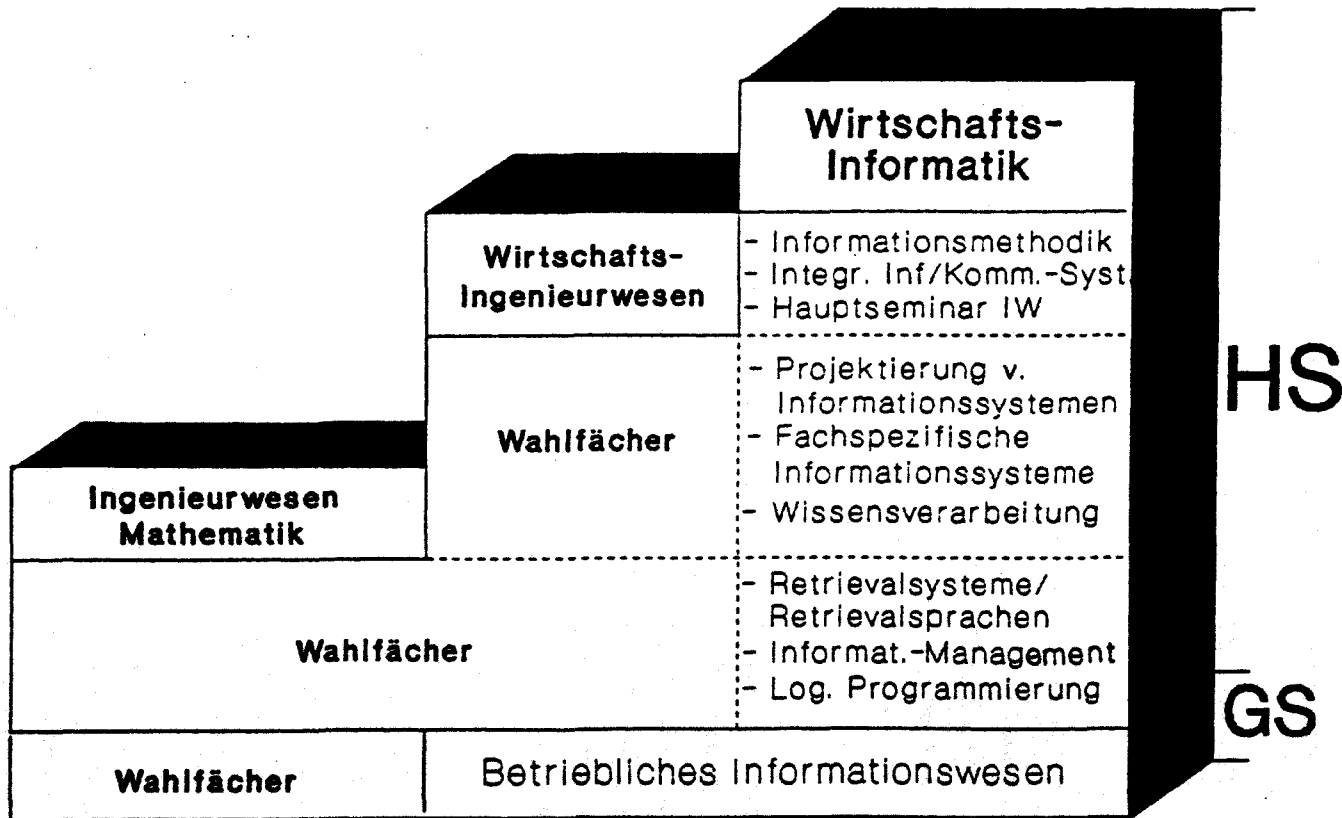


Bild 2: Lehrangebot des Instituts für Informationswissenschaft

## Literatur

- /1/ Manecke, H.-J.: Die Technische Hochschule Ilmenau/Thüringen  
Eine alte(neue) Ausbildungsstätte für Informationsfachkräfte  
Nachr.Dok.42 (1991), S.52-53
- /2/ Studienordnung Diplom-Studiengang Wirtschaftsinformatik  
TH Ilmenau, Fakultät für Wirtschafts- und Rechtswissen-  
wissenschaften (i.G.), 1990
- /3/ Schmitz, P.; Seibt, D.; Strunz, H.: Informatiker, Wirt-  
schaftsinformatiker und Ingenieure der Fachrichtung  
Informatik aus der Sicht der DV-Anwender-Erhebung zu  
Aufgaben, Ausbildungskonzeption und Bedarf  
Wirtschaftsinformatik 32 (1990) 1, S. 20-32

Prof. Dr.-Ing.habil. Harald Killenberg

TH Ilmenau

Fakultät Wirtschafts- und Rechtswissenschaft

Institut für Informationswissenschaft

O - 6300 Ilmenau

PSF 327

## **Informationswissenschaft in Saarbrücken**

### **- Stand und Perspektiven**

Ilse Harms, Angelika Schulz

Einleitend wird die Fachrichtung Informationswissenschaft in Saarbrücken und ihre Inhalte in Forschung und Lehre seit 1984 skizziert. Neue Fragestellungen und ein sich veränderndes universitäres Umfeld verlangen Überlegungen zu einer modifizierten informationswissenschaftlichen Ausbildung. In Saarbrücken wird darauf mit einer neuen Schwerpunktsetzung reagiert, die im Bereich Informationsarbeit und -präsentation liegt.

As an introduction, the Information Science Department (Fachrichtung Informationswissenschaft) in Saarbrücken and the contents of its research and teaching activities since 1984 are described. New problem fields and a changing universitarian background call for reflections on a modified education in Information Science. The Saarbrücker reaction to this development is a shift emphasis towards information work and information presentation.

An der Philosophischen Fakultät der Universität des Saarlandes wird seit 1984 Informationswissenschaft als Haupt- und Nebenfach im Magister- bzw. Promotionsstudiengang angeboten. (Bereits vor 1984 konnte das Studium aufgenommen werden, das Fach war jedoch noch nicht durch eine Studienordnung abgesichert.) In der Folge wurde es dann möglich, Informationswissenschaft auch in Kombination mit Fächern außerhalb der Philosophischen Fakultät zu studieren. Informationswissenschaft kann heute mit Informatik, Wirtschaftswissenschaften und Rechtswissenschaft kombiniert werden. Ein universitätsinterner NC begrenzt die Zulassung von Hauptfachstudierenden auf 25 pro Jahr, die Zahl der Nebenfachstudierenden ist unbegrenzt. Hauptfachstudierende

werden jeweils zum Wintersemester aufgenommen. Abbildung 1 zeigt die progressive Entwicklung der Studentenzahlen, die sich nach dem Wintersemester 1990/91 auf ca. 400 belaufen. Für das Studienjahr 1991 werden voraussichtlich ähnliche Zahlen gelten wie für 1990:

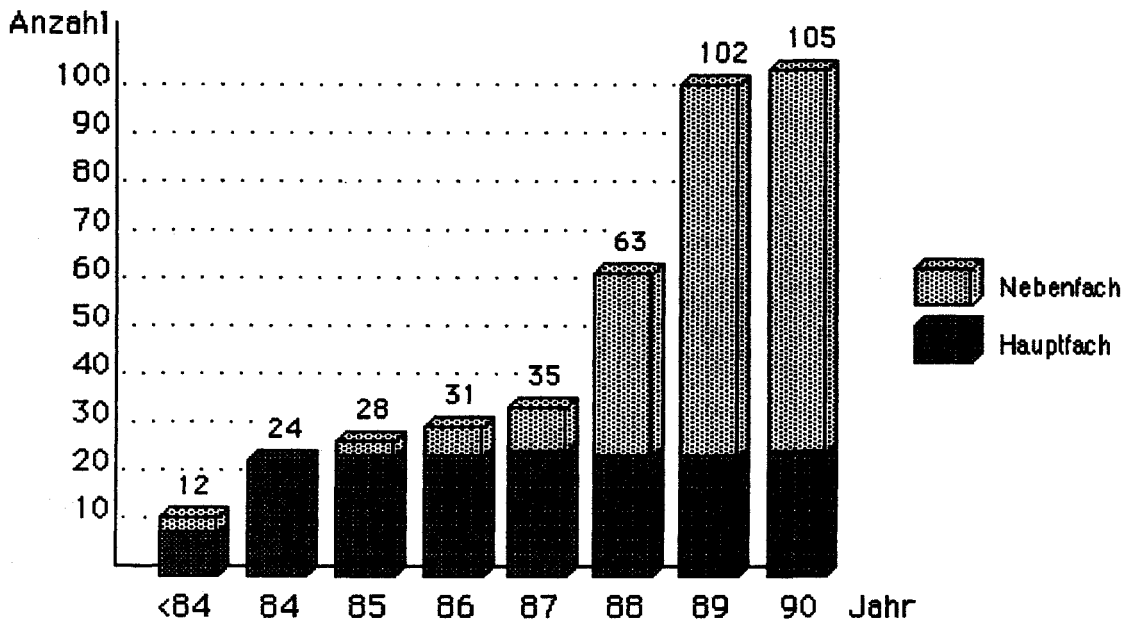


Abb. 1: Haupt- und Nebenfachstudierende Ende 1990  
nach Jahr des Studienbeginns (Quelle: Statusbericht)

Auffällig ist die starke Zunahme an Nebenfachstudierenden ab 1988; sie deutet nicht nur auf ein wachsendes Interesse an Informationswissenschaft als Nebenfach hin, sondern zeigt vielmehr das allgemeine Interesse an einem Studium der Informationswissenschaft auf. (Da für das Hauptfach ein NC gilt, fangen viele Studierende als "Nebenfächler" an in der Hoffnung, später umsteigen zu können.) Konträr zu der Entwicklung der Studentenzahlen steht die personelle Ausstattung der Fachrichtung, die sich seit 1984 nicht verändert hat: Der aktuellen Zahl von 400 Studierenden steht eine Lehrkapazität von 1 Professur und 1,5 wissenschaftlichen Mitarbeitern gegenüber. Über ein geringes Kontingent an

Lehraufträgen und durch studentische Tutoren wird das Studienangebot gewährleistet, eine umfassende Betreuung der Studierenden jedoch ist schwierig geworden.

Für die Attraktivität der Informationswissenschaft als Studienfach lassen sich eine Reihe von Gründen ausmachen. Der Einzug der Informationstechnik auch in vormals computerferne Bereiche außerhalb und innerhalb der Universität erfordert zunehmend Kenntnisse im Umgang mit dem Computer. Durch die informationstechnische Grundbildung in den Schulen wird den nachfolgenden Schülergenerationen einerseits die Relevanz des Mediums vermittelt, andererseits werden Schwellen abgebaut. Die gute bis sehr gute Arbeitsmarktsituation für die Absolventen muß nicht weiter betont werden.

Das Curriculum für den informationswissenschaftlichen Studiengang in Saarbrücken entstand zu Beginn der 80er Jahre. Es ist inhaltlich partiell geprägt von den Aspekten, die das IuD-Programm vorgab (Stichwörter Fachinformation, Informationsindustrie, betriebliche Information und Kommunikation), wurde jedoch erweitert durch den Schwerpunkt Publikumsinformation. Durch das Zusammenwachsen der Informations- und Kommunikationstechnologie, d.h. von maschineller Informationsverarbeitung und Vermittlungstechnik, werden die Grenzen von Individual- und Massenkommunikation aufgehoben (vgl. Zimmermann 1984a). Der Wissenstransfer an ein breites Publikum wurde deshalb auch Gegenstand der Informationswissenschaft in Saarbrücken. Zimmermann (1990) beschreibt ausführlich die Inhalte der einzelnen tätigkeitsfeldbezogenen Schwerpunkte. Vorbereitet wird deren vertiefte Bearbeitung durch vier sogenannte "Querschnittsbereiche", die, neben praxisorientierten Übungen, den ersten Studienabschnitt bis zur Zwischenprüfung prägen, aber während der gesamten Studiendauer Einfluß nehmen auf die thematische Ausgestaltung der Schwerpunkte.

Es ergibt sich daher folgender Studienaufbau:

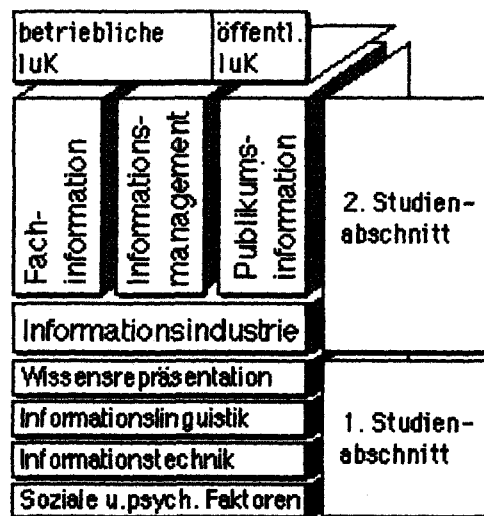


Abb. 2: Aufbau des Studiums der Informationswissenschaft in Saarbrücken

Dieses Modell hat sich bisher aus mehreren Gründen bewährt:

- es ist thematisch breit angelegt
- es gewährleistet vielfältige Orientierungsmöglichkeiten für Studierende (was sich auch an der Themenvielfalt der bisher eingereichten Masterarbeiten zeigt)
- die Schwerpunktsetzung kann so gewählt werden, daß sie die im Masterstudiengang erforderlichen weiteren zwei Fächer inhaltlich ergänzt.

Unter den oben skizzierten Bedingungen aber führt die thematische Vielfalt zu Schwierigkeiten:

- Bei der verfügbaren Lehrkapazität erfordert das pro Semester wechselnde Lehrangebot eine fortwährende Umorientierung der Dozenten
- Forschungsschwerpunkte können nur selten unmittelbarer Gegenstand der Lehre sein
- die Lehrangebote, die über Lehraufträge abgedeckt werden müssen, wechseln ständig

Die informationswissenschaftliche Forschung an der Universität des Saarlandes konzentrierte sich bislang vor allem auf "Fragen der Überwindung von Informations- und Sprachbarrieren in der technischen Information und Kommunikation mit dem Ziel, grundsätzlich bereits vorhandenes Wissen besser zu distribuieren." (Zimmermann 1984b) Die Tradition der Forschungsarbeiten aus dem Bereich der maschinellen Übersetzung und Indexierung in Saarbrücken wird auch heute in einer Reihe von Projekten weitergeführt.

Die Vermittlung theoretischer Grundlagen in der Lehre wird ergänzt durch anwendungsbezogene Forschung. Seit 1984 steht hierfür das IAI, das "Institut der Gesellschaft zur Förderung der Angewandten Informationsforschung e.V. an der Universität des Saarlandes" als sogenanntes "An-Institut" zur Verfügung. Aufgabe des IAI ist es, "informations- und kommunikationswissenschaftliche Fragestellungen wissenschaftlich zu bearbeiten und durch praxisnahe Arbeiten die Anwendung informations- und kommunikationswissenschaftlicher Erkenntnisse in Wirtschaft, Wissenschaft und Behörden zu fördern (Wissens- und Technologietransfer)" (IAI-Info). Projekte am IAI sind, neben der Sprachdatenverarbeitung, auf den Gebieten Expertensysteme, Online-Schulung und -Beratung sowie Büroorganisation angesiedelt. Studierenden der Informationswissenschaft wird die Möglichkeit gegeben, an einzelnen Projekten mitzuarbeiten. Exemplarisch sollen einige Forschungsprojekte kurz vorgestellt werden.

Das derzeit größte Projekt ist EUROTRA, ein EG-Projekt zur Überwindung der Sprachbarrieren in der Europäischen Gemeinschaft. Ziel des Projekts, das von 17 Forschungsgruppen in allen 12 Mitgliedsstaaten der EG mit ca. 170 Mitarbeitern durchgeführt wird, ist hierbei die Entwicklung eines maschinellen Übersetzungssystems modernster Konzeption. Der deutsche Teil des Projekts wird in Saarbrücken durchgeführt, der Etat betrug seit dem Start im Jahre 1985 jährlich bis zu zwei Millionen DM, ca. 17 Mitarbeiter (im

wesentlichen Wissenschaftler) sind an dem Projekt beschäftigt.

Um die Generierung einer modernen Lern-Infrastruktur auf Hypertextbasis für EDV-orientierte Betriebswirtschaftslehre unter Berücksichtigung verschiedener Benutzertypen, geht es in dem Projekt WITEC (Wissens- und Technologietransfer innerhalb universitärer Ausbildungspartnerschaften - Anwendungsbereich "Wirtschaftsinformatik"). Dieses Vorhaben wird gemeinsam mit dem Institut für Wirtschaftsinformatik an der Universität des Saarlandes und der Fakultät für Wirtschaftswissenschaften der Universität Leipzig, Lehrstuhl für Wirtschaftsinformatik, durchgeführt.

Nicht die Wissenschaft, sondern die Wirtschaft ist Kooperationspartner im Projekt STEELEXPERT (System für Bearbeitung und Verwaltung auftragsbezogener technischer Parameter bei der Edelstahlproduktion). Ziel ist die Erstellung eines Expertensystems zur Auftragsbearbeitung in der Edelstahlherstellung. Ein dreiköpfiges WissenschaftlerInnen-Team hat im Sommer 1991 - nach 1 1/2-jähriger Vorarbeit - bei dem Auftraggeber, einem südkoreanischen Stahlhersteller, Installation und Schulung vorgenommen.

In die Richtung "angewandte Informationswissenschaft" unter dem Aspekt der Optimierung von betrieblicher Informationsorganisation geht das Projekt IF\*P (Informationspraktikanten). Der Wissenstransfer zwischen Wissenschaft und Wirtschaft wird von fortgeschrittenen Studierenden der Informationswissenschaft im Rahmen eines Zwei-Phasen-Praktikums geleistet. Nach der Einarbeitung in die betriebspezifischen Informationsprobleme zeigen sie die Möglichkeiten neuer IuK-Technologien auf und entwickeln "vor Ort", also in den Betrieben, Nutzungskonzepte. Das Projekt gewährleistet einerseits den kleinen und mittleren Unternehmen im Saarland eine Vertiefung des informationstechnischen Know-Hows und andererseits den Studierenden den



für die Berufsfeldorientierung notwendigen Praxisbezug. Das Projekt wird aus Fördermitteln des Bundes für die Montanregion Saarland und aus Mitteln des Saarlandes gefördert.

## Perspektiven

Die Existenz eines wechselseitigen Beziehungsgeflechts zwischen Gesellschafts- und Kommunikationsform ist unbestritten. Eine gründliche Analyse dieses Verhältnisses findet sich u.a. bei Habermas (1962). Wersig (1985, S. 95) formuliert dazu: "In mehrfacher Hinsicht ist die Geschichte der Menschheit auch die Geschichte der Entwicklung ihrer Kommunikationsmittel." Unterschiedliche Aussagen gibt es jedoch hinsichtlich der prophezeitlichen Implikationen bei der Veränderung auch nur eines Elements des Gesamtsystems. Ohne auf die unterschiedlichen Positionen in der Folgediskussion einzugehen kann festgestellt werden, daß die Wissenschaft sowohl Motor als auch Folger von Innovationen ist. Diesem Faktum trägt die Saarländische Landesregierung im Rahmen ihrer Wissenschaftspolitik Rechnung durch eine neue Schwerpunktsetzung im universitären Bereich. Der forcierte Ausbau der technischen und mathematisch-naturwissenschaftlichen Fakultäten (hier vor allem Informatik), die Einrichtung neuer Studiengänge wie Computerlinguistik, Kognitionspsychologie, Wirtschafts- und Rechtsinformatik sowie die Ansiedlung innovativer Forschungseinrichtungen wie das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI), das Max-Planck-Institut für Informatik und das Internationale Begegnungs- und Forschungszentrum für Informatik (Schloß Dagstuhl) unterstreichen das Programm. Auch die Umwandlung der Fachhochschule für Design in eine Hochschule der Bildenden Künste ist in diesem Zusammenhang von Bedeutung.

Die Informationswissenschaft findet sich deshalb zehn Jahre nach ihrer Etablierung in Saarbrücken nicht nur in einem

anderen gesellschaftlichen Kontext wieder, sondern auch in einem neuen universitären Umfeld, auf das sie reagieren muß. Das bisher realisierte "Saarbrücker Modell" (vgl. Zimmermann 1990), das sich in den letzten Jahren als erfolgreich erwiesen hat, wird in seiner curricularen Ausrichtung der extrem gestiegenen Anzahl von Studierenden, den Anforderungen an ein optimiertes Studienangebot, aber vor allem auch der erwähnten veränderten Informationsumwelt mit anderen Bedürfnissen und neuen Verfahren nicht mehr in allen Aspekten gerecht.

Damit keine Mißverständnisse entstehen: Die oben genannten universitären Rahmenbedingungen "diktieren" der Saarbrücker Informationswissenschaft nicht die Gegenstände von Forschung und Lehre, sie wirken vielmehr unterstützend bei der Bewältigung der Aufgabenfelder, die die Gesellschaft stellt. Damit wurden also für die Informationswissenschaft nur die Voraussetzungen geschaffen, die es ihr ermöglichen, die ihr inhärente "Brückenfunktion" (Zimmermann 1990) einzubringen. Dafür bieten sich thematisch eine Reihe von Möglichkeiten, die teilweise auch schon realisiert wurden.

Informationswissenschaft wird aufgefaßt als eine Wissenschaft, deren zentrales Ziel als "facilitating the effective communication of desired information between human generator and human user" formuliert wird (Belkin 1978). Der Mensch mit seinen kognitiven Fähigkeiten steht damit im Mittelpunkt. Wenn man sich diesen Blickwinkel zueigen macht, muß man den Informationsprozeß von der Wissensakquisition bis hin zur Rezeption unter einem "cognitive viewpoint" sehen. Bisher hat man sich eher auf die formalen Erfolge des Information Retrieval (der Wissensbereitstellung an sich) konzentriert (vgl. Grundlagenpapier 1991). Der (Be)Nutzungsaspekt aber rückt erst seit einiger Zeit in das Blickfeld wissenschaftlicher Forschung, wie zahlreiche Veröffentlichungen und eine zunehmende Anzahl von Kongressen in den Bereichen Software-Ergonomie, Design von Benutzungsober-

flächen, Mensch-Maschine-Interaktion und Benutzermodellierung der letzten Jahre belegen.

Wissensveränderung findet nur da statt, wo Information in vorhandene Wissensstrukturen sinnvoll integriert werden kann (vgl. u.a. Kuhlen 1990). Hierzu aber bedarf es Informationen, die sich an den individuellen rezeptiven und kognitiven Kompetenzen orientieren bzw. an diese adaptierbar sind. Diese Form der benutzerorientierten Aufbereitung von Information wird durch Informationsarbeit geleistet. Informationsarbeit schließt unter dem "cognitive viewpoint" die Wissensakquisition, die mehrwerterzeugende Transformation und die Präsentation ein. Zusätzliches Gewicht erhält der letztgenannte Aspekt durch die zunehmende Verfügbarkeit neuer elektronischer Medien mit ihren Möglichkeiten der nutzeradäquaten Darstellung von Wissen.

Die sach-, medien- und benutzergerechte Aufbereitung von Information sowie deren ergonomische und ästhetische Gestaltung wird aus allen genannten Gründen der Schwerpunkt der Informationswissenschaft in Saarbrücken sein. Die kognitionswissenschaftliche Ausrichtung wird durch das in Saarbrücken vorhandene wissenschaftliche Umfeld gefördert.

#### Literatur:

Belkin, Nicholas J. (1978): Information Concepts for Information Science, in: Journal of Documentation, 1978/34, S. 55-85.

Grundlagenpapier: Informationswissenschaft 2000 im Rahmen einer Europa-Universität Saar-Lor-Lux-Rheinland/Pfalz, 6/91, unveröffentlichtes Manuskript.

Habermas, Jürgen (1962): Strukturwandel der Öffentlichkeit. Darmstadt.

IAI-Info: IAI-Info-Mappe 2/91. Institut der Gesellschaft zur Förderung der Angewandten Informationsforschung e.V. an der Universität des Saarlandes.

Kuhlen, Rainer (1990): Zum Stand pragmatischer Forschung in der Informationswissenschaft, in: J. Herget, R. Kuhlen (Hrsg.): Pragmatische Aspekte beim Entwurf und Betrieb von Informationssystemen. Proceedings des 1. Internationalen Symposiums für Informationswissenschaft. Konstanz, S. 13-18.

Statusbericht: Statistischer Statusbericht der Fachrichtung Informationswissenschaft, 5/91.

Wersig, Gernot (1985): Die kommunikative Revolution. Strategien zur Bewältigung der Krise der Moderne. Opladen.

Zimmermann, Harald H. (1984a): Studienführer zum Fach Informationswissenschaft. 1. Aufl., Saarbrücken.

Zimmermann, Harald H. (1984b): Informationswissenschaftliche Ausbildung und Forschung an der Universität des Saarlandes. Saarbrücken.

Zimmermann, Harald H. (1990): Informationswissenschaft an der Universität des Saarlandes ("Saarbrücker Modell"). In: M. Buder, W. Rehfeld, T. Seeger (Hrsg.): Grundlagen der praktischen Information und Dokumentation, 3. Aufl., München etc., S. 1100-1107.

Dr. Ilse Harms  
FR 5.5 - Informationswissenschaft  
Universität des Saarlandes  
Im Stadtwald, Gebäude 4  
6600 Saarbrücken

Angelika Schulz, M.A.  
FR 5.5 - Informationswissenschaft  
Universität des Saarlandes  
Im Stadtwald, Gebäude 4  
6600 Saarbrücken

# Die Informationsgesellschaft: eine Utopie im Spiegel postmoderner Theorie ?

## Das Beispiel Amerika

Steffen Wawra

Der Begriff einer kommenden Informationsgesellschaft ist Gegenstand zahlreicher Diskussionen. Für alle im Prozeß von Produktion und Distribution von Information Stehende ist die Prüfung des Realitätsgehaltes des Begriffes Informationsgesellschaft von Bedeutung, führt sie doch zu einer Einordnung der eigenen Tätigkeit in soziale, ethische und politische Entwicklungen. Für diese Realitätsprüfung, die uns zeigen soll, ob dem Begriff der Informationsgesellschaft Wirklichkeit entspricht, ist die Methode des postmodernen Denkens geeignet, da sie von einer Pluralität sozialer Prozesse ausgeht.

Wir beginnen erst zu begreifen, mit welcher Macht die Information unsere Welt verändern wird. Die Digitalisierung von Information wird zu einer Neuordnung des traditionellen Gefüges wissenschaftlicher Kommunikation führen. Traditionelle Institutionen müssen ihr Selbstverständnis ändern, um auf einer neuen Ebene fortbestehen zu können.

Risiken eines Autonomieverlustes des politischen Handelns gegenüber einer multi-medialen Präsenz von Wirklichkeit werden beschrieben.

We talk a great deal about the coming of the information age. For all professionals whose common concern is information it is important to examine the reality of information society. We need a clear perspective of "who we are" and "what we have to do". Our view of the reality may or may not coincide with the true reality.

We can use the post-modern thinking for examine the reality of information society, which is involving a plurality of social development.

We are beginning to recognize the power that information brings to our live and our society. The digitalization of information will create a post-traditional system of scientific communication. Traditional institutions such as university libraries must change their individuality for survival in a higher level. Risks of a loss of autonomy in the political interactions in opposition to "a multi-media vision of reality" created by the mass media and computer industry are described.

### 1. Die Informationsgesellschaft - eine neue Qualität?

Die Entwicklung der modernen Gesellschaft zu einer Informations- und Kommunikationsgesellschaft, die aufgrund der neuen Stellung von Wissen und Information zu einer zunehmenden Informatisierung und Digitalisierung sämtlicher gesellschaftlicher Lebensprozesse führt, scheint evident. Zu groß sind die sichtbaren ökonomischen, sozialen und politischen Konsequenzen aus dieser Entwicklung, als die Tendenz negiert werden könnte. Die gegenwärtige und zukünftige Entwicklung sozialer Systeme ist von diesen Prozessen in einem Maße abhängig geworden, das den Schluß nahelegt: die industrielle (oder "postindustrielle") ist eine Informationsgesellschaft. Nach SHULTZ /14, S.9/ verändert die "information revolution" die Routine unseres Planeten in einer Art, wie es die industrielle Revolution des letzten Jahrhunderts tat. Die Basis unserer Ökonomie bewege sich rapide von einer industriellen Produktion zu "information-based goods and services".

Das bekannte Beispiel, daß in den USA etwa 46 % des Bruttosozialproduktes

und mehr als 50 % des Arbeitsaufkommens aus Waren und Dienstleistungen erzielt werden, die aus Informationsverarbeitung resultieren und jeder zweite US-Dollar, der in der Wirtschaft der USA erarbeitet oder verbraucht wird, durch die Produktion und Distribution von Informationen geschaffen, bzw. durch den Kauf von Informationen verbraucht wird /10, S. 380/, wird nachvollziehbar, wenn bedacht wird, daß der Weltfinanzmarkt einen elektronischen Markt geschaffen hat, der pro Jahr 3 Billionen USD Umsatz macht und vom Geschäft des "handling money" sich tendenziell in den Bereich des "handling informations" verlagert. Ökonomische, politische und soziale Stärke von Systemen hängt zunehmend von der Fähigkeit ab, Informationen zu produzieren, zu verteilen, zu managen und zu nutzen /1, S. 20/.

Zwar ist die Informationstechnologie nur eine von verschiedenen Basistechnologien, stellt aber offensichtlich einen Katalysator dar, der andere Entwicklungen ermöglicht und in ihrer Durchsetzung forciert. Am "Schwarzen Montag" von 1987 wurden in den Folgetagen des Börsencrashes Kursverluste von 1000 Mrd USD verzeichnet. Bereits in der ersten halben Stunde des Crashes waren in der Wall Street mehr als 50 Mio Aktien verkauft worden. Dabei trugen maßgeblich zu einer Konfusion des Marktes computergesteuerte Verkäufe bei, da die auf bestimmte Mindestkurse programmierten "sell orders" eine Eigendynamik entwickelten, die gewissermaßen in Echtzeit und weltweit wirkte /vergl. 12/. Einen für die Prägung des Begriffes Informationsgesellschaft wesentlichen Ansatz lieferte BELL /4/. Er klassifiziert die gesellschaftliche Entwicklung in eine "vorindustrielle" Gesellschaft, in eine "Industriegesellschaft" und in eine "nachindustrielle" Gesellschaft. Vollzog sich das Leben in den vorindustriellen Gesellschaften als ein Kampf gegen die Natur, erschien es in der Industriegesellschaft als ein Spiel gegen die technisierte Natur. Nicht mehr bloßes Überleben, sondern die Rationalität der Gesellschaft besteht in Wirtschaftlichkeit. Die nachindustrielle Gesellschaft schließlich ist definiert als ein Spiel zwischen Personen. Sie ist eine durch fünf Merkmale zu beschreibende Dienstleistungsgesellschaft: durch den Wirtschaftssektor (der Übergang von einer güterproduzierenden zu einer Dienstleistungsgesellschaft ist vollzogen), durch die Berufsstruktur (der Vorrang einer Klasse professionalisierter und technisch qualifizierter Berufe), durch das axiale Prinzip (die Zentralität des theoretischen Wissens als Quelle von Innovation und Ausgangspunkt der gesellschaftlichen und politischen Problematik), durch die Zukunftsorientierung (die Steuerung des technischen Fortschrittes und Bewer-

tung der Technologie) und die Entscheidungsbildung (Schaffung einer neuen Entscheidungstechnologie) /4/. Der österreichische Wirtschaftswissenschaftler SCHMORANZ stellt darauf aufbauend die These auf, daß die Gesamtentwicklung des Dienstleistungssektors durch die Dynamik der Informationsaktivität bestimmt wird und zieht daraus den Schluß: die postindustrielle Gesellschaft ist eine Informationsgesellschaft /13a, S. 45/.

## 2. Die Kritik: der postmoderne Ansatz

Nachdem die Möglichkeit der Betrachtung der modernen Gesellschaft als Informationsgesellschaft dargestellt wurde, soll dieser Ansatz unter dem pluralen Denkmuster der postmodernen Theorie von Individuum und Gesellschaft hinterfragt werden.

Der hier verwendete Sinninhalt des Begriffes Postmoderne zielt in Anlehnung an LYOTARD und WELSCH auf eine veränderte Aneignung der Welt: nicht mehr die Singularität der Konstruktion eines Systems ist der Anspruch des Denkens an sich selbst, sondern die Bereitschaft zur Aufnahme der Welt als ein pluraler und komplexer Organismus. Postmoderne meint also nicht einen Zustand, sei es individuell oder zeitgeschichtlich, nach einer bestimmten Epoche, sondern sieht sich gleichsam als Klammer der Pluralität - entsprungen allerdings einer veränderten Geisteshaltung, die als Haupttopos das "Ende der großen Entwürfe" /9, S. 112/ für sich setzt. Das Ende der großen Entwürfe als das Ende der "Metaerzählungen" zeigt sich nach WELSCH im Scheitern des Totalitätswahnes von Leitideen, so der Emanzipation durch Wissenschaft in der Aufklärung, der Teleologie des Geistes im Idealismus, der Befreiung der Menschheit durch die Revolution des Proletariats im Marxismus, der Beglückung aller durch Reichtum im Kapitalismus und, auf unseren Gegenstand bezogen: der "endgültigen Problemlösung durch Informationsmaximierung im technologischen Zeitalter" /18/. Postmoderne sieht sich so zwar als vom geistigen Ansatz der Moderne verschieden an (als deren Idealtypus in der Regel HEGEL aufgefaßt wird, der eine Versöhnung der Welt im objektiven Geist

seines Idealismus erreichen wollte, dabei aber bereits postmoderne Pluralität zu erfassen suchte: "Das Prinzip der neueren Welt überhaupt ist die Freiheit der Subjektivität, daß alle wesentlichen Seiten, die in der geistigen Totalität vorhanden sind, zu ihrem Recht kommen, sich zu entwickeln." /8, S. 439/). Sie feiert aber keinen Abschied von den Metaerzählungen in einer Weise, die auf einen Zustand nach diesen Systemen abhebt, sieht sich also nicht als Diskontinuität, wie es HABERMAS kritisch wertet ("Mit diesem "nach" wollen sich die Protagonisten von einer Vergangenheit absetzen; der Gegenwart können sie einen neuen Namen noch nicht geben" / 6, S. 11/), wohl unter zu deutlicher Bezugnahme in Literatur und Architektur auftretenden Tendenzen, in denen es sich in der Tat in den 70er Jahren (Charles Jencks) um eine provokative Absage an moralische Grundsätze der modernen Architektur handelte.

Die postmoderne ist eine Philosophie des Abschieds von der Moderne nur in dem Sinne, daß sie deren "letzte Gründe" der großen Metaerzählungen nicht mehr anerkennt. Eine philosophiegeschichtliche Parallele zur Moderne ist hierbei aufschlußreich: In der Geburtsstunde der griechischen Philosophie wurde die Frage nach dem Verhältnis von Einheit und Vielheit gestellt, die "moderne" Frage nach dem Prinzip, "woraus alles Seiende besteht und woraus es ursprünglich entsteht, und in das es letztlich vergeht, in dem die Substanz zwar beharrt, aber in ihren Zuständen wechselt"/2, S. 983b /. In dieser Substanz ist dann zwar EINES auch VIELES und umgekehrt, aber die Suche nach dem letzten Prinzip konnte noch nicht aufgegeben werden. Postmodernes Denken versucht, die Akzeptanz der Vielheit, der Pluralität als Kriterium einer reifen Menschheit zu definieren. In dem Sinne, daß die großen Entwürfe natürlich weiter bestehen, reproduziert in Denken und Handeln der Individuen, aber eingestehen müssen, ihren Anspruch auf eine totalitäre Umfassung von Wirklichkeit verloren, aber durchaus in einer Pluralität Bestand zu haben. Einer Pluralität, die im Gegensatz zu einem "einfachen Pluralismus", der durch Kommensurabilität gekennzeichnet ist, einer Inkommensurabilität verpflichtet ist. Dieser "postmoderne Pluralismus" ist die Denkweise einer erwachsen gewordenen Menschlichkeit, die begriffen hat, daß eine Vergleichbarkeit über das Anlegen eines gleichen Maßes nicht gegeben ist. Besonders interessant für unsere Fragestel-



lung scheint die logische Konsequenz aus diesen Sätzen darzustellen, daß nach WELSCH nämlich eine externe Pluralisierung eine interne hervorruft, in der die Dekonstruktion der Wirklichkeit keine Destruktion darstellt / 17 /. Und LYOTARD: "Die narrative Funktion verliert ihre Funktoren, den großen Heroen, die großen Gefahren, die großen Irrfahrten und das große Ziel. Sie zerstreut sich in Wolken, die aus sprachlich-narrativen, aber auch denotativen, präskriptiven, deskriptiven usw. Elementen bestehen, von denen jedes pragmatische Valenzen sui generis mit sich führt. Jeder von uns lebt an Punkten, wo viele von ihnen einander kreuzen. Wir bilden keine sprachlich notwendigerweise stabile Kombination, und die Eigenschaften derer, die wir formen, sind nicht notwendigerweise mitteilbar" / 9, S. 15/. Kriterium des neuen Denkens ist es also, die Inkommensurabilität zu akzeptieren und zu ertragen und sich gegen den unhaltbaren Ansatz von MACHT zu wehren, der da lautet: "Wirkt mit, seid kommensurabel, oder verschwindet" / 9, S. 15/.

Ist die Informationsgesellschaft nun eine solche Metaerzählung, sich vielleicht ergebend aus einer "subjektiven Philosophie", einer bloßen Erfolgs-story multimedialer, globaler Unternehmen, die diese Philosophie verkaufen und somit einen expansiven Markt vorbereiten / 8a / ? Oder sind die Begriffe "Informationsgesellschaft" und "postmoderne Gesellschaft" zwei verschiedene begriffliche Belegungen ein und derselben Sache, nämlich auf die qualitativ veränderte Stellung von Wissen in der modernen Gesellschaft zielend ? Oder setzt sich Postmoderne über die Informationsgesellschaft gleichsam durch, ohne sie begrifflich anzuerkennen ? Hat die Empfindlichkeit, mit der die "Metaerzählung Marxismus" auf den Begriff der Informationsgesellschaft reagierte und ihn aus einer "surrogativen" Informationsideologie kommend beschrieb, die über die wirklichen Triebkräfte der Gesellschaft nichts aussagt, den logischen Grund, daß hier das eine Ziel einem anderen begegnet ist ?

### 3. Die Auflösung

Die Logik dargestellten postmodernen Denkens muß die Belegung der modernen Gesellschaft mit dem Begriff Informationsgesellschaft ablehnen. Postmoderne Theorie differenziert jedoch das Herangehen an die Problematik: Bei der Betrachtung des komplizierten Wechselverhältnisses von Wissen - Information - Gesellschaft\* kann die dargestellte neue Rolle des Wissens in der Gesellschaft - als wesentliche Voraussetzung für das Greifen der Produktivkraft Information und als Faktor zur Ausübung von Kontrolle und Macht - nicht übersehen werden. So konstatiert LYOTARD, das Wissen ändere mit dem Eintreten der Gesellschaften in das postindustrielle und das Eintreten der Kulturen in das postmoderne Zeitalter sein Statut /9, S.30/. Und unter dieser Betrachtung ist es bis zur Betrachtung des Szenarios der Informatisierung der höchstentwickelten Länder nur noch ein kleiner Schritt. Damit kann die Transformation des Wissens erhellt und seine Auswirkungen auf die Gesellschaft untersucht werden und zudem habe dieses Szenario gute Chancen, sich durchzusetzen, da nicht ersichtlich ist, welche andere Orientierung die Technologien der Gegenwart einnehmen sollen, die sich als Alternative zur Informatisierung der Gesellschaft herausstellen könnte.

Allerdings wird diesem Ansatz kein prognostischer Wert gegenüber der Wirklichkeit eingeräumt, sondern ein strategischer in bezug auf die Untersuchung der Veränderung der Natur des Wissens.

Die Ambivalenz des Denkens postmoderner Kultur in bezug auf die Informationsgesellschaft wird auch im folgenden sichtbar: Der postmoderne Ansatz setzt sich über eine zu Ende gedachte Informatisierung der Gesellschaft durch, da ein abstrakt vorstellbarer allgemeiner freier und ungehinderter Zugang zur Information logisch dazu führen muß, daß die verschiedenen Sprachspiele (also die voneinander verschiedenen Ansätze der Betrachtung von Wirklichkeit) in diesem Moment Spiele mit vollständiger Information sind. Mit dieser Konstruktion von Sprachspielen mit vollständiger Information schafft LYOTARD jedoch eine neue Metaerzählung: Betrachtet man etwa die Idologie von totalitären Systemen, so lebt deren Sprachspiel doch gerade vom bewußten und unbewußten Zurückhalten von

---

\* Auf das komplizierte Verhältnis von Wissen - Nichtwissen - Information - Desinformation soll hier nicht näher eingegangen werden.

"Spieleinsätzen" - also Information über das eigene und fremde Systeme - glaubend, damit seinen Wirklichkeitsanspruch besser durchsetzen zu können.

Aber: festzuhalten bleibt, daß selbst das modernste gegenwärtige Denken als postmodernes die Informationsgesellschaft als ein Sprachspiel von übergeordneter Bedeutung gelten läßt, die sie einzunehmen durch die qualitative neue Rolle von Wissen und Information in der modernen Gesellschaft in der Lage ist. In der Konsequenz postmoderner Theorie muß hier gleichzeitig Wert auf die Feststellung gelegt werden, daß die Informationsgesellschaft als Parameter die Qualität hochentwickelter Gesellschaften zwar wesentlich, aber nicht ausschließlich bestimmt. Und die durch demokratische Mechanismen gelenkte Informationsgesellschaft kann mit dem Anspruch einer plural verstandenen Freiheit der Information dem postmodernen Wunsch nach Gerechtigkeit und nach Unbekanntem in der Tendenz entsprechen und einen "Polylog" ermöglichen, "in dem Denk- und Handlungsmöglichkeiten geprüft und begründet werden, die dann im Bewußtsein ... einsichtig werden" /11, S. 210/.

### 3. Die Realität der Informationsgesellschaft

Es wurde gezeigt, daß die Möglichkeiten elektronischer Information und Kommunikation für die gegenwärtige und zukünftige Entwicklung sozialer Systeme eine wesentliche Funktion - sowohl industriell als auch kulturell - einnehmen werden.

An dieser Stelle sollen nun einige Probleme angesprochen werden, die sich an der Schnittstelle postmoderne Kultur/Informationsgesellschaft zeigen. Hatte zwar jede Stufe der gesellschaftlichen Entwicklung ihre eigene, spezifische Informationsstruktur, mit der in jener Zeit umgegangen werden mußte, zeigen sich durch die neue Qualität der Informationsgesellschaft Probleme, mit denen der moderne Mensch erst umzugehen lernen muß. Wobei sicher nicht linear gedacht werden kann: Der Urmensch, für den die Information, wo sich denn nun das Mammut aufhält, überlebensnotwendig war, hatte sicher eine große Menge von Informationen auf Relevanz zu prüfen. Für den Menschen der Gegenwart gestaltet sich die Frage des Überlebens in einer

multimedialen Gesellschaft komplizierter, da er nicht nur mehr Informationen erhält, sondern durch die zunehmende Komplexität der Produktion, Verarbeitung und Distribution von Informationen die Auswahl relevanter schwieriger wird.

Um beim Beispiel zu bleiben: Der moderne Mensch erhält nicht nur Informationen, die ihn das Mammut wohl ewig suchen lassen würden, sondern auch solche über Beschaffenheit des Fells, Fleischverarbeitungsmöglichkeiten, etc.

Die differenzierte moderne Gesellschaft weist eben sehr differenzierte Interessenlagen auf, deren unaufhörlichem Einfluß sich das Individuum ausgesetzt sieht. Und führt dann der Einfluß der Massenmedien, die die "Enthemmung des Informationswesens" /vergl. 15/ professionell betreiben, über die Überflutung von Bewußtseinskapazitäten in anthropologisch bedrohlicher Weise zu einer "Totalsynthese" durch "Totaladdition" (SLOTERDIJK), werden Rahmenbedingungen deutlich, in welchen sich das Individuum als soziales Wesen zu bewegen hat.

Diese empiristische Totalsynthese ist nicht das Ziel einer aufgeklärten postmodernen Kultur, deren "vollständige Information" meint: "Gebt mir die für mich notwendigen Informationen und dann laßt uns reden."

In den USA zeichnen sich die Informations- und Kommunikationsstrukturen der Zukunft bereits ab. In wohl keinem anderen Land der Welt wird deutlich, was man gemeinhin mit "Gefährdungen" und "Möglichkeiten", resultierend aus einer digitalisierten Welt, umschreibt.

Untersuchungen zur politischen Kultur der Informationsgesellschaft zeigen in den USA eine Verschiebung sozialen Handelns in Richtung einer "telekratischen" Teilhabe ("with the public watching without partizipation" /5, S. 63/) an sozialen Prozessen. Diese Tendenz tragen ca. 20 % der amerikanischen Bevölkerung.

Aber auf der anderen Seite sind gerade die USA ein Paradebeispiel dafür, wie die Möglichkeiten des elektronischen Zugriffs auf "public information", die z. B. sämtliche Forschungsdaten und -ergebnisse von lokal- oder bundesstaatlich unterhaltenen Forschungseinrichtungen enthält, im Sinne einer demokratischen Kultur diskutiert und durchgesetzt werden.

Institutionen, die traditionell der Freiheit von Forschung und Informa-

tion verpflichtet sind, geraten durch die neuen Möglichkeiten elektronischer Kommunikation in Handlungszwang. Vielleicht gibt es in der Zukunft Universitäten, aber deren Bibliotheken werden nur noch zu historischen Zwecken genutzt. Der Student besitzt seine eigene digitale Bibliothek auf dem PC, der interaktive Medien bereitstellt. Simulation von chemischen Prozessen, Musik-Synthese, originale Reden von herausragenden Wissenschaftlern - alles ist möglich und im "NeXT-PC" von STEVE JOBS bereits möglich. Und sollten doch einmal Originalquellen benötigt werden - die "Bank of America" übernimmt die Beschaffung ...

#### Literatur

- / 1 / Allan, K. B.: A National Strategy for the Information Age  
- In: Information Management Review 4 (1989) 4, S. 17-27
- / 2 / Aristoteles: Metaphysik, A 3, 983b, 6
- / 3 / Bartels, Klaus: Schöne neue Medien-Welt ? . - In: Cogito (1991) 4, S. 19-23
- / 4 / Bell, Daniel: Die nachindustrielle Gesellschaft. - Frankfurt, 1985
- / 5 / Demokratierituale: zur polit. Kultur d. Informationsgesellschaft / hrsg. von Fritz Plasser. - Wien, 1985
- / 6 / Habermas, Jürgen: Die neue Unübersichtlichkeit. - Frankfurt, 1985
- / 7 / Habermas, Jürgen: Der philosophische Diskurs der Moderne. - Frankfurt, 1989
- / 8 / Hegel, G. W. F.: Suhrkamp-Werkausgabe. - Band 7, S. 439
- / 8a/ Krollner, Franz: Die Bibliotheken und die Informationsgesellschaft. - In: ABI-Technik 5 (1985) 4, S. 267-272
- / 9 / Lyotard, Jean-Francois: Das postmoderne Wissen. - Graz, 1986
- / 10/ Ploman, E. W.: The Communication Revolution. - In: ASLIB-Proceedings 33 (1987) 10, S. 373-382
- / 11/ Richter, Frank: Philosophie in der Krise. - Berlin, 1991
- / 12/ Rohde, Erwin: Schwarzer Montag: Börsenkrach `87. - Berlin, 1988
- / 13/ Rorty, Richard: Bildung ohne Dogma. - In: Dialog (1990) 1, S. 44-47
- /13a/ Schmoranz, Ingo: Makroökonomische Analyse des Informationsmarktes. - Wien, 1980
- / 14/ Shultz, George: The Role of Information Scientists. - In: The Scientist, Oktober 5, 1987, S. 9
- / 15/ Sloterdijk, Peter: Kritik der zynischen Vernunft. - Frankfurt, 1983
- / 16/ Max Weber: Soziologie - Universalgeschichtliche Analysen - Politik. - Stuttgart, 1973
- / 17/ Welsch, Wolfgang: Postmoderne. - Köln, 1988
- / 18/ Welsch, Wolfgang: Topoi der Postmoderne. - Vortrag

---

Dr. Steffen Wawra  
Akademie-Bibliothek Berlin  
0 - 1086 Berlin, PSF 1313  
Unter den Linden 6

0 - 1144 Berlin  
Ingolstädter Str. 55

## ÜBERLEGUNGEN ZUM PAPERWORK REDUCTION ACT 1980

Thomas Weitzendorf, Graz

als verheißungsvoller Beginn eines erfolgreichen, öffentlichen Information-Managements in den USA gesehen. Die vorliegende Publikation durchleuchtet die nunmehr elfjährige Geschichte des PRA inklusive seiner gesetzlichen Vorgänger, um diese Sichtweise kritisch zu überprüfen. Weiters wird auf theoretischer Ebene untersucht, ob die Methodik der Kosten-Nutzen-Analyse verbunden mit einem Preismechanismus auf von der US-amerikanischen Bundesregierung zu erhebende Informationen anwendbar ist.

Eine der Zielsetzungen des PRA war und ist die bürokratische Entlastung der US-amerikanischen Öffentlichkeit. Die bürokratische Belastung der Öffentlichkeit stammt von Formularen, die ausgefüllt werden müssen, um den Informationsbedarf der Bundesverwaltung zu decken. Während von Bundesseite Formulare als Informationsquelle gesehen werden, sieht die Öffentlichkeit hauptsächlich die bürokratische Belastung und den unbequemen Papierkrieg in diesen Formularen. Während also die Bundesverwaltung mehr den Nutzenaspekt betrachtet, konzentriert sich die Öffentlichkeit auf die hohen Kosten der Informationserhebung.

Diese beiden konträren Sichtweisen haben sich schon seit längerem in der Diskussion innerhalb der US-amerikanischen Legislative - im Senat und im Abgeordnetenhaus - widerge-

spiegelt. Der PRA ist nur eine der letzten, nach außen hin sichtbaren Resultate dieser Debatten.

Der PRA war bei weitem nicht der erste Anlauf - wie in Europa vielfach angenommen - der US-amerikanischen Legislative, Information und ihre Kosten auf Bundesebene besser zu kontrollieren. Es war jedoch der erste Versuch, Informationstechnologie zu nutzen, um die Effizienz der Informationserhebung in der Bundesverwaltung zu erhöhen.

Vorgänger des PRA wurden aufgrund der damals verwendeten und andersartigen Technologie auch anders benannt. Am Anfang des Jahrhunderts wurde in den USA dem Vervielfältigungsproblem am meisten Aufmerksamkeit geschenkt. Die heutige Vervielfältigungstechnologie hat die Kosten von Kopien und daher dieses Problem auf ein Minimum reduziert.<sup>1</sup> Der Federal Reports Act von 1942 war der erste Gesetzesakt, der bürokratische Entlastung ("burden reduction") wörtlich erwähnte. Die Schwäche dieses und der folgenden Gesetzesakte war es, daß bloß ungefähr 25 % der gesamten Formulare von Bundesbehörden von den entsprechenden Bestimmungen betroffen waren. So wurden vor allem die Formulare des US-amerikanischen Bundesfinanzamtes (Internal Revenue Service) zur Gänze unberührt gelassen. Der Brooks Act von 1965 war der erste Vorgänger des PRA, was den Einsatz von Informationstechnologie in der US-amerikanischen Bundesverwaltung betrifft. Er beinhaltete als erster Gesetzesakt Richtlinien zum Kauf, zum Leasen, zur Instandhaltung, zum Betrieb und Gebrauch von Informationstechnologie in der Bundesverwaltung.<sup>2</sup>

Trotzdem waren Ausnahmen von Gesetzgebung bis zum PRA 1980 die Regel. Vor dem PRA waren auch die Entbürokratisierungsbestre-

---

<sup>1</sup> Federal Paper Requirements: U.S. Senate, Select Committee on Small Business, Subcommittee on Government Regulation and Paperwork, Washington D.D., 27 June 1979, p. 68.

<sup>2</sup> Caudle Sharon L.: "Federal Information Resources Management After the Paperwork Reduction Act", Public Administration Review, vol. 48 (Juli/August 1988), p. 791.

bungen unter den Regierungen Ford und Carter auf ein Minimum der bürokratischen Last beschränkt.

Der zweite Teil des Beitrags beschäftigt sich mit der Frage, ob die aufgezeigten Interessensgegensätze durch eine Kosten-Nutzen Analyse und durch ein darauf basierendes Preissystem objektiviert werden können.

Die Kosten-Nutzen Analyse ist ursprünglich ein volkswirtschaftliches Werkzeug, wobei der Nutzen sowohl allgemein als auch im speziellen Fall größere Bewertungsprobleme mit sich bringt als die Kosten.

Um die Kosten und Nutzen des Informationsbedarfs der US-amerikanischen Bundesverwaltung zu bewerten, wird im vorliegenden Beitrag folgendermaßen vorgegangen:

Zunächst wird die Verwendung der Kosten-Nutzen Analyse zur Objektivierung des Bewertungsvorgangs des Informationsbedarfs begründet und deren Vor- und Nachteile erörtert. Weiters wird der momentan von den Behörden verwendete Ansatz des Information Collection Budgets (ICB) diskutiert. Als Kern des Beitrags wird erklärt, wie die notwendigen Informationen für eine Kosten-Nutzen Analyse erhoben werden könnten. Schließlich wird darauf eingegangen, welche Schritte in Zukunft notwendig sein werden, um die Kosten-Nutzen Analyse für den öffentlichen Informationsbedarf anwendbar zu machen.

Dr. Thomas WEITZENDORF  
Institut für Informations-  
wissenschaft  
Karl-Franzens-Universität Graz  
Strassoldogasse 10  
8010 GRAZ



## CONCEPTION OF INFORMATIC TECHNOLOGIES IN MANAGEMENT

ANANYEVA T.N., cand. of techn. scien..

SVIRIDENKO J. P., doct., prof., Moscow Technologie Institute

With the growth and complication industrial enterprises structure and the increase of the workers employed, it has become more complicated to manage the purposeful activities of people. This led to a fast growth of employees who carry out administrative and management functions. And so there arose such a situation all over the world the most numerous part of employed people, that is the management body, turned out to have the lowest productivity.

Introduction of automated management systems into practice universally little raised the productivity of the final user, that is the manager, for the last decades. This can be explained by a number of situations. They are:

- system orientation at the information from only one group of documents (for example, scientific technological, economic statistical, practical ordering), all this means the absence of complex conception of different character information contained in systems;
- limited structure of solved problems dealing mainly with information organisational and summary-grouping functions;
- far distance of systems from the final user, existence of great many links between him and the system, those links that deal with information processing, and this considerably reduces management operativeness;
- absence of simple and convenient methods and means for interaction between managers and information systems.

Introduction of personal technology into management practice and elaboration on this base of personal-orienting information systems direct at every working place and the subsequent joining them by means of distributed computer networks all this creates new possibilities for elaborating complex systems for automating administrative management activity.

The problems of creating new information technologies covering all the stages of information processing from receiving and transferring to representing, storing and using are comprehensi-

vely described in the papers [1,2,3].

In this address we shall consider two problems:

- the problem of elaboration special means for representing management information in information bank;
- the problem of elaborating the user's interface as an autonomous program-apparatus module providing effective dialogue between the user and the computer.

#### THEME FIELD MODEL ELABORATING

Experience in using data bases showed a definite feature of information storage arranging consisting in the fact that besides storing factografing information which continuous operational processing text files containing a structural text (quotations, recommendations, instructions, assignments ect.) are extensively used.

Hence, we offer a conception of constructing a linguistic provision (LP) of personal knowledge base. The conception is based on combining the means for representing theme field model built on the frame base and the functional and processoriented treasurus system.

In this case theme field model (TFM) is given by a construction called a frame network, i.e. the network having empty slots which then are filled by a concrete content characterising the given situation.

Frame network consist of following frame totality: factografic material and its use. The top of the frame network is the user's frame (Fig. 1).

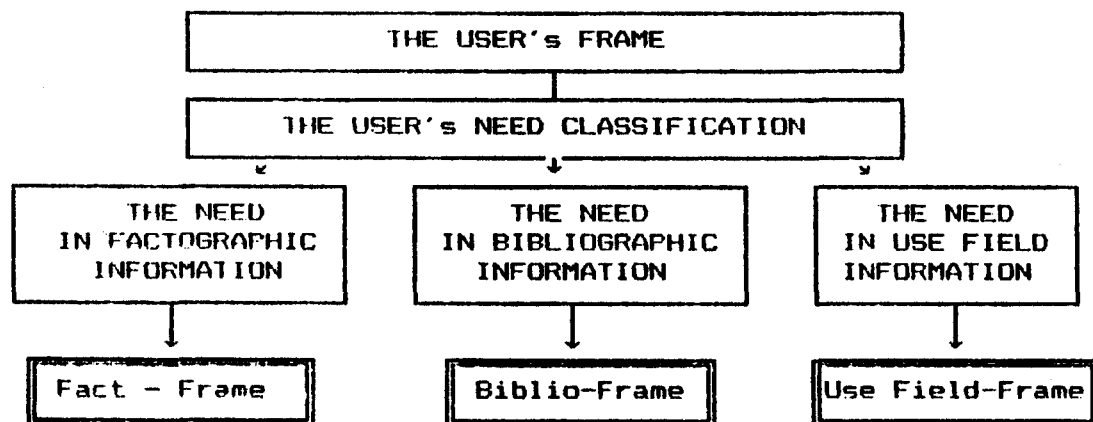


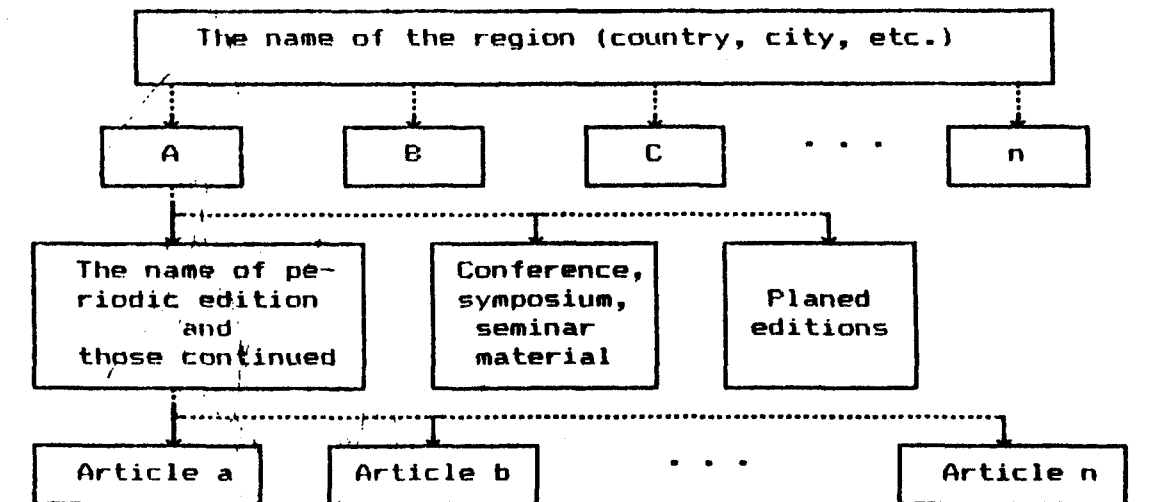
Fig.1. A fragment of user's frame model.

The USER'S FRAME serves for revealing, systematizing and forming hierarchical trees, each of them representing subframe. The frame constructing occurs in several stages. In the first stage, formal data about information system user analysed (his education, speciality, post, foreign language knowledge ect.). Then aims and tasks of activities and the range the user is interested in are defined. With the help of statistic analysing the user's inquiries, polling and interviewing ect. revealed user's interests are systematized. The user's information need classification and data analyses received previously define the structure of the user's frame.

To describe the factographic information we use a fact frame defining encyclopaedic data structure and destined for knowledge representing about THEME FIELD (TF). The Fact-Frame possible to describe both dynamic and static knowledge. The static knowledge is represented in the kind of a standart frame set. Dynamic knowledge in the kind of slots indicate concrete research results, each slot acting as an independent frame.

Representation of documental information for TF model is carried out by a bibliografic frame, "Biblio-Frame". It includes information about the source of knowledge and about the author, the name of the source, the place of publication etc..

A geographical sign may be assumed as a Biblio-Frame classification basis (Fig.2).



Bibliography (Article, Author, Name of Publication, Year, Page.

Fig.2. A fragment of a Biblio-Frame.

One of the most complicated frame network building problem is elaborating of Application Field Frame (Appli-Frame) since application of some knowledge, in broad sens, exceeds application field described, i.e. the same knowledge may be used differently. For this strive to define not how some information is used but where that or other property, fact, TF object is applied. Today's development of ceramics, for example, led to an extensive use of it, from catalists for environmental protection to high temperature reactors, the detail extent of using the same information in different fields of knowledge being different. It is obvious that for frame elaborating, for using ceramics in the mineralogy field the stress is made on material description and all the frame network branches are developed in the case of an analogous frame in the silicate material chemistry field.

In view of this we offer to use a so called Scale Change Mechanism. It has been offered in "SINTOL" System for building a paradigmic apparatus. Thus, the Use-Field-Frame makes it possible to join different information into a united network. The frame interaction model in chemistry is shown in Fig.3.

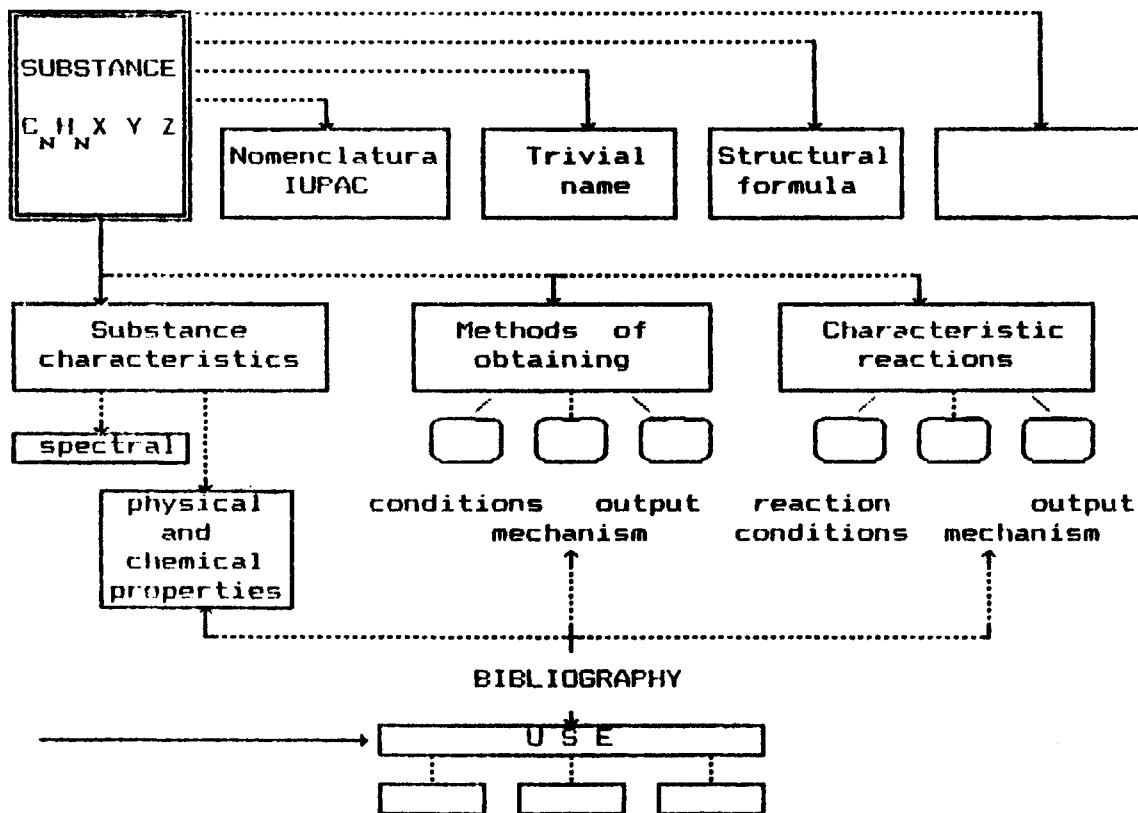


Fig.3. Interaction model of bibliographic, factographic and use field frame in problem-oriented base in chemistry.

The personal base thesaurus should not be considered as a contracted variant of the big thesaurus. Out of memory economy considerations the personal base thesaurus is laconic one concerning the word-stock and the same time it has a high level of representing information profile accuracy. Furthermore, the logical connections existing between TF ideas should be represented with maximum completeness. Thesaurus has to possess apparatus of easy accessibility for control and keeping. Being a pilot means of connections with a computer, thesaurus should be maximum handy for screen representing. It should have an optimum structure and more complete system of reference to the sources and massives of information storing it should possess a mechanism of building representations from separate TF terms.

In our case thesaurus system consist of partial thesaurus (PT) totality of different kinds that are in coordination with one another in discipline and task-independent semantic criteria. Proceeding from the concept of idea classifications according their functional signs it should be noted that partial thesaurus acquire autonomy in the rights of an owner of the names of definite category ideas (quality, materials, etc.) and they cannot be compared with a thesaurus in a concrete speciality in the branch thesaurus system.

The following categories are recommended as a base for classifying thesaurus system according to the partial thesaurus: Objects, Relations, Qualities, Properties. This categories are universal and independent from specific character of their use. Each of the big partial thesaurus undergo to futher subdivision into functional classes and integration stages (that is vertical and horisontal idea grouping). So obtained totally of partial frames that are more particular in their functions represents multitude ideas which intersect each other. By this intersection between functional classes of only one category is assumed. In case of need, hierarchical associative connections are defined only of the same category.

At the stage of defining a thesaurus form it is necessary for elaborator to choose the most characteristic and actual types of paradigmic connections and to give the thesaurus filling specialists a chance to use them.

At the first stage of compiling a thesaurus it is offered to single out an objective field of ideas and their names which are to be systematized within a partial thesaurus or within its part. For creating an objective field it is most effective to use ideas grouping according to their permanent signs.

Modern means of programming make it possible to work out an automatic register in which not only words but whole expressions can be arranged in alphabetical order. In particular, it is possible to make a list of words which are supposed to be introduced to a thesaurus of expressions having identical elements, i.e. keys, can be chosen with the help of the keys and then grouped and arranged in alphabetical order. This makes the process of idea organization on objective field much easier. So formed objective fields may be conjunctively connected with one another forming great partial thesaurus which have great word stock and many paradigmic connections but retain the signs of all their field components up to the level of functional classes. All such organisation finally results in a thesaurus. The same methods of idea grouping can be applied within any objective field.

The next stage of thesaurus structuring is the task of defining connections between classes in to represent paradigmic connections.

As a normal case of defining hierarchical connections polyhierarchy is chosen. Polyhierarchy implies such an idea organisation in which every idea has more than one direct connection owing to taking into account of greater number of differentiation signs.

To make an automatic polyhierarchical thesaurus easier it is recommended to mark connections within it so that it could be possible to single out monohierarchical thesaurus. This means for each idea which may be classified polyhierarchically one is chosen out of possible variants which mostly agrees with the chosen classification strategy. This connection will be automatized in a kind of automatic substitution of a descriptor by a disjunction which is above it.

All the rest distributions and connections will be characterised as secondary ones. Not being automatized they will be marked in the thesaurus for the case of traditional using.

The necessity of polyhierarchy introduction for elaborating a personal knowledge base thesaurus is obvious compiling of which is given into the hands of a professional indexator. It will make possible for him to express his ideas about the structure of the knowledge base used, i.e. it will make his orientation in thesaurus system easier and will provide the due comfort in his work with it in view of the user's idea classification according to hierarchical levels.

All the connections between ideas represented by order of the group organization should be indicated distinctly.

With due regard for technology of representing connections the following system of mutual references is offered:

D-title descriptor;

E-nondesoriptor that refers, as an equivalent idea, to D and is substituted by this descriptor in the indexing process;

K-nondesoriptor that refers to the descriptor conjunction by which it will substituted;

O-above-standing descriptor of the monohierarchical system;

P-above-standing descriptor of the polyhierarchical system;

U-below-standing descriptor of the monohierarchical system;

A-associative descriptor: every D can belong only to one A;

B-associative descriptor: every D can belong to several A;

C-associative descriptor: not too important for D.

There exists a number of methods for compiling out of separate thesaurus lexical units which are integral according to their sense of expression. The most traditional of them are postcoordinated term combinations entering the system in the kind of set expressions. This methods cannot be used in the presented thesaurus variant in which all the descriptors are ctriectly classified in three (non-interested) categories and cannot be represented in a kind of element combination belonging to different categories. In our case it is possible to use as a means supplementing thesaurus a regulator apparatus (i.e. typing, aspect, phase, model, (indicator), etc.). The regulators may be represented in a separate table and used for building up descriptor chains out of different category ideas. It is sufficient for the user to choose the most suitable regulators offered and to introduce them to the chain of descriptor thesaurus according to definite rules. Introducing a regulator into descriptor chain

makes possible for the user to exactly form needs, to single out idea aspects and operation stages, to introduce time and space qualifiers, to coordinate different category ideas by conjunctive relators, etc.

As mentioned above, the characteristic feature of personal knowledge base is accumulation of the great amount of both text- and factographic components in it.

Factographic information in a knowledge base is organized by a frame networks, i.e. network whoos modules are mutually structured and can be processed by means offered by operational system by the base itself. Therefore it is expedient to use inverse memory not simply for storing facts recognized on address of there source but to supply them with addresses of working files in which a given descriptor plays the role of a data element. Then the search on one or several criteria will give the follow result: the adress of information source out of the document files and file indices in which these criteria give a name to a data field structuring nuder one name different meaning of one and the same element. By this the following algorithm of action may be realized:

1. The search for descriptors on thesaurus which are in the base (in a dialogue regime);
  2. The search for working file indices on inversive memory (automatic regime);
  3. Discovery of the named files in order to operate with them.
- Thus, the approach offered applies a representation of linguistic and problem knowledge and concrete facts concerning the object fields on a single formally-logical base.



## MEANS OF SUPPLING "UNTRAINED USER-COMPUTER" INTERFACE

In connection with the necessity to take account of different factors in the process of the user's communication with the system it seems to be expedient to use Petry network theory apparatus which makes possible taking account influence of both planned and arising conditions have to carry out, to most completely solve the problems of information supplying the users both administrator and the executants of separate stages of the research program or another solved problem.

A simple representing of the system by the Petry network is based on two main ideas:

1. An event is an action taking place in a system; the system state controls arising of events and it should be described by a great number of conditions,
2. A condition is the predicate and logical description of the system. The condition can assume a value either 1 (one) or 0 (zero). A network is an aggregate  $(P, T, F)$  where  $P$  is a non-empty great number of places and  $T$  is a non-empty great number of transitions:  $F \subseteq P \times T \cup T \times P$ .

We have offered constructing principles of the software having elements of examination (expert) systems that make it possible for subscriber to formulate his inquiry for information system and to develop his practical skills of communication with the system in regime of dialog search. The concept of a semantic graph is assumed as a basis of building-up this software. The semantic graph expresses the DB structure and some concepts about the Petry network theory.

The semantic graph serves for choosing inquiry lexical elements and at the same time it is a model of a surplus (typical) inquiry. As a result of the semantic graph interpretation arises inquiry (system examination) which expresses an initial semantic graph fragment.

The dictionary of associative expansions is complied by experts as a semantic graph. The aggregation means form inquiry from chosen semantic elements with regard to requirements of the language of concrete data base management system (DBMS).

This software supplied by forecasting elements can prepare subscriber's inquiry more operatively for account of constructing

model in the kind of set expression totally of inquiry semantic elements. In the process of interaction with the subscriber the latter may be offered ready semantic constructions formed automatically on the basis of statistical information collected on the previous stages.

Object model evolves, i.e. it is renovated after every regular seance of subscriber interaction with this software (later on inquiry processor). Inquiry processor (IP) realization on personal computer with the use of interaction language of the "menu" type permits it to have a "friendly" adaptive character with regard to a subscriber.

To construct problem-oriented data base (PODB) model we shall define some concept used in relational DB and semantic model theory.

Under the object we shall understand some material substance, processor a concept information of which is stored in DB. Object identification is carried out with the help of names. Every object is characterised by a number of signs. The object sign is described by data elements. Before defining this concept more accurately we should return to the base semantic concepts.

Thus we shall consider the network uniting thematically similiar DB with different DBMS. The user of such a network usually may handle any DB in a certain unified inquiry language based on knowledge of semantic model (SM) of data common for the network and the common complex of the semantic models.

Inquiry and answer languages based on data description language (DDL) accepted in SM.

The main requirements to SM and DDL suitable for above formulated aims consist in simplicity and clearness for user-nonprogrammers, in answers for their transmissions over communication canal.

Basic semantic concepts of the offered models are already mentioned "object", "sign" (property, attribut), "event" and "index".

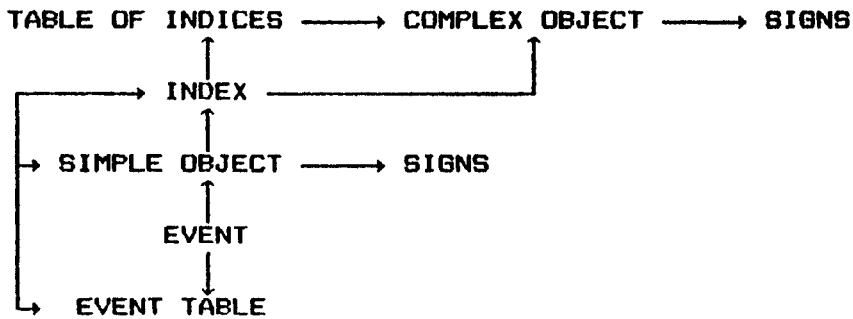
Concepts of the type of a content data aggregates are used as apilot ones. We examine such aggregates as "object-sign", "listy of object and their signs" and "table of indices".

Object are united into classes (great numbers) of signs

homogenous by their composition. Object identification in a given class is carried out through its name (code).

Indices may be united into tables. Common data parts are put into title, into column and line names.

Connections existing between different content data aggregates are shown in the following Figure:



Primary information is an information arisen as a result of events. This information may be transformed into new signs of objects or into indices. Indices too are "object-sign" data processing. Complex objects have signs in the kind of indices or index table.

Let's translate said above into the language of semantic models. Thus, object signs are described by data elements (DE).

DE is the smallest indivisible aggregate of symbols in this model.

Meaning totality of one DE forms a domain:

Domain name (meaning 1, ... meaning - ).

One domain may unite DE with definite semantic aggregate . Below we shall operate with domains only with text DE. Such models are set in a SM by a list of this data. In the range of one domain all DE are different.

We shall call elementary data forming a domain as a domain positions. Between domain positions, relations of the type of synonymic hierarchical ones may be set. Between different domains, relations may be set which are usual for great number.

A construction consisting of a name and meaning we shall call a semantic unit (SU). In a simplest case domain's name is used as SU, domain's position is used as SU meaning.

Semantic category (SC) is a concept expressing most common and substantial properties, signs and connections between semantic language components in a given object field (object, pro-

cess, different types of relations).

A semantic field (SF) is a structure expressing lexic semantic classification of a given object field which unites term on the basis of their semantic totality.

Analyzing inquiries to such DB in a natural language one may notice that their semantic base touches a number of semantic categories. With this it is possible to quite simply construct a due semantic field for every revealed SC. A semantic field is constructed so that the term-candidates to the semantic field are selected intuitively and offered by experts who are specialists in this object field. So constructed semantic field will be both representative and quite compact.

Analysis of a whole semantic field formed by uniting semantic fields for every semantic category shows that it is possible to single out  $H^k$  of aspects depending on the point of view at the problem.

For constructing a generalized model we single out a group of domains  $\{Dom_i^k\{P_{ij}\}\}$  which expresses data connected with one of aspects.

Every domain generally contains  $i \times j$  of positions of  $P_{ij}$  elements,  $i=1, I, j=1, J$ .

Appearing in the role of experts we shall introduce connections between positions of neighbour domains of a given group. This will lead to appearing a certain graph on many domain positions. By analogy with this it is possible to construct a graph for the other aspects. We shall introduce additional top-source in every graph that will be connected with all domain positions and will be the first one the  $k$  group.

It should be noticed that so constructed graphs cross one another through the common top source but do not through the ribs since they express different aspects of one and the same field.

In a word, graphs  $\{H^k\}$  are parts of a certain graph  $\{H^0\}$  expressing semantic connections between the terms in a problem-oriented data base

$$H^0 = \bigcup_{k=1}^K H^k \quad (2)$$

Taking into account that the inquiry to the DB represents an additional-multiplication function (arguments-lexical inquiry are connected with each other by the logic "and", "or", "not"), and also the fact that a graph may be presented in the kind of

a successive construction of the type "each one with the other" between positions of the first domain of the k group and the first domain of the k+1 group. If we assume that

$$(\forall H^k, \forall \text{Dom}_i^k) P_{i_i} = \text{Dom}_r^j(P_{r,s}), \quad r = \overline{1,R}, \quad s = \overline{1,S} \quad (3)$$

then we can say that the data base model will be G graph, including  $H^0$  and complement  $\bar{H}$  which in its turn is a straight sum (on the tops and the ribs) of non intersected parts of  $\{H^k\}$  constructed on domain positions.

$$G = H^0 + \bar{H}, \quad \text{where } \bar{H} = \bigcup_{i \in I} H^i \quad (4)$$

Now we introduce rib orientation from the top-source into graph G. As semantic unit selection is carried out by transact's passing along one of the routes it will be logical to go over from a semantic graphs to an algorithm graph-diagram (AGD) of semantic unit selection. It should however be noticed that the graph in common case must be complete.

Attention should be paid to the combination of the two apparatuses: domain position out of the semantic model theory and relation data bases that is synonymously expressed in the position of Petry network apparatus. In this case the presentation of the semantic model in the kind of non-marked Petry network simplifies the process of constructing inquiry preprocessor AGD.

Any terms, their bases, standart abbreviations may be domain position meanings. In the domain position meanings representing which transform word combination into common terms. It is only important that terms tied to the domain positions be coordinated with lexis of terminological dictionary of problem-oriented data base.

By practical realization inquiry preprocessor, switch functions may be exercised by indicator light pen or similar device with the help of which choice of one of the alternative is carried out. The list of alternatives is given on a screen whenever a transition from one domain to the another takes place.

Out of marked semantic units aggregation module forms a primary inquiry variant which at wish of the user may be specified.

In the common case a semantic graph for inquiry preprocessor may be constructed as follows:

1. To limit a number of aspects  $\{H^k\}$  characterizing objects from

certain most important points of view.

2. For every aspect, a group of domains  $\{Dom_i^k\}$  with the aggregate names of similar object signs.
3. For every domain group a corresponding graph  $G$  is formed as follows:
  - a) on the base of expert data a great number of the domain position text meanings in the group is defined;
  - b) a great number  $\{P_{ij}\}$  of tops of the graph  $G$  corresponding to position meaning is formed;
  - c) on the base of expert information connections of the kind of arcs between top are introduced, these connections express relations between domain positions.
4. Out of parts  $\{H^k\}$  so formed a straight line sum on ribs is drawn according to expressions (2).
5. In case of need expression (3) is taken into account, after that the graph  $G$  is drawn.

#### CONCLUSION

The present stage society development is characterized not only by technical inventions but also by introducing new technologies in management are most closely tied up with manager practical activists, they make information sources which are in disposal of administrative bodies easily accessible.

The methods and means of data and knowledge introducing and also methods and means of data manipulating offered in this paper are realized within the range informational systems for scanning fulfilment of purposeful complex program on material-technical supply.

This system was tested by carrying out experiments on applying the means offered in the limits of an informational system in chemistry and chemical technology.

The experience of practical application of this system shows that the set and instrument content are practically invariant by transition from one branch to another. This gives the grounds to think that the elaborated means may be successively applied by constructing flexible systems for a certain class of complex objects (e.g. state power bodies, ministries, corporations, firms, etc.). Their application is even possible when according to changing conditions there arises a need to reconstruct management model components by a given process.

#### REFERENCES

1. T. N. Ananyeva, Avtomatizirovannye sistemy obrabotki, poiska i hraneniya upravlenchesroi informatsii. M., MGIAI, 1984.

# Bereitstellung von Agrardatenbanken auf CD-ROM - ein Projekt für die neuen Bundesländer

Münch, E., Thulke, M., König, E.

## Zusammenfassung

Auf der Grundlage eines Förderprojektes des Bundesministeriums für Forschung und Technologie erhielten 15 wissenschaftliche Einrichtungen auf dem Gebiet der ehemaligen DDR CD-ROM-Stationen und wichtige agrarwissenschaftliche Literaturdatenbanken (AGRIS, AGRICOLA, CAB ABSTRACTS, FSTA und ASFA).

Es werden erste Ergebnisse zur Nutzung der CD-ROM-Technologie vorgestellt. Ausgehend von statistischen Erhebungen werden Aussagen zur Gesamtzahl der Recherchen, zu Nutzergruppen, zu recherchierten thematischen Gebieten und zur Nutzungshäufigkeit der CD-ROM-Literaturdatenbanken getroffen.

## Abstract

Within the scope of a development project of the Federal Ministry of Research and Technology, fifteen scientific institutions in the former German Democratic Republic received CD-ROM stations and CD-ROM holding major documentary data bases relating to the agricultural sciences (AGRIS, AGRICOLA, CAB ABSTRACTS, FSTA and ASFA). The paper describes the first results of the CD-ROM technology used. On the basis of statistical surveys, statements are made regarding the total of searches, regarding user groups, searched subjects, and on the frequency of CD-ROM databases used.



# Bereitstellung von Agrardatenbanken auf CD-ROM - ein Projekt für die neuen Bundesländer

Münch, E., Thulke, M., König, E.

## 1 Einleitung

Bis zum 3.10.1990 existierte auf dem Gebiet der neuen Bundesländer das Informationssystem Wissenschaft und Technik der Land-, Forst- und Nahrungsgüterwirtschaft (IWT der LFN). Unmittelbare Aufgaben sind die Umstrukturierung sowie Neuorientierung des IWT der LFN und die Überführung dieses Systems in eine nach föderalistischen Prinzipien konzipierte Informationslandschaft (Fachinformationssystem Ernährung, Land- und Forstwirtschaft - FIS/ELF)/1,2,3,4,5/. Die Probleme der Informationsversorgung unter Einsatz der elektronischen Datenverarbeitung spielen dabei eine besondere Rolle. Entsprechend den Möglichkeiten in der ehemaligen DDR wurden im IWT der LFN auf diesem Gebiet vergleichsweise gute Erfolge erreicht /6,7,8,9,10/.

Der Einsatz von CD-ROM spielte in der ehemaligen DDR keine herausragende Rolle, obwohl man sich seit 1989 intensiver mit dieser Technik befaßte /11,12,13,14,15/. Aus finanziellen und technischen Gründen waren einer praktischen Anwendung enge Grenzen gesetzt. Diese Situation war auch typisch für das IWT der LFN, wobei das Institut für Landwirtschaftliche Information und Dokumentation (ILID) als damaliges nationales AGRIS-Inputzentrum die Datenbank AGRIS auf CD-ROM bezog und ein technisch-technologischer und methodischer Vorlauf existierte. Dieser Erkenntnisgewinn schlug sich unmittelbar in dem mit Unterstützung der Zentralstelle für Agrardokumentation und -information (ZADI) vom ILID beim BMFT beantragten Förderprojekt "Informationsversorgung im Bereich Ernährung, Land- und Forstwirtschaft in den neuen

Bundesländern durch Einsatz von CD-ROM-Technologie" nieder. Dabei flossen die Erfahrungen der ZADI beim Umgang mit der CD-ROM-Technologie in die Aufgabenstellung des Projektes ein /16,17/

2 Das Förderprojekt "Informationsversorgung für den Bereich Ernährung, Land- und Forstwirtschaft (ELF) in den neuen Bundesländern durch Einsatz der CD-ROM-Technologie"

Das Projekt beruht auf der Tatsache, daß die Infrastruktur auf dem Gebiet des Post- und Fernmeldewesens in den fünf neuen Bundesländern zum gegenwärtigen Zeitpunkt und auch in absehbarer Zeit sich auf einem technischen Niveau befindet, daß der on-line-Zugriff zu Datenbanken nicht möglich ist. Dabei ist es völlig klar, daß die CD-ROM-Technologie keine direkte Konkurrenz für den on-line-Zugriff darstellt, sondern daß beide Formen spezifische Aufgaben bei der Informationsbereitstellung zu erfüllen haben.

Das Projekt hat zwei grundlegende Ziele. Erstens soll durch die Bereitstellung von CD-ROM-Stationen und Literaturdatenbanken auf CD-ROM ein unmittelbarer Innovationsschub bei einer effektiven Informationsversorgung erreicht werden, um den Zeitraum bis zum Aufbau einer modernen Kommunikationsinfrastruktur wirksam zu überbrücken /18/. Zweitens soll ein Beitrag zur Erweiterung des Wirkungsbereiches des FIS-ELF in Deutschland in organisatorischer, struktureller, inhaltlicher und technisch-technologischer Hinsicht durch Einbeziehen oder Nutzung von Elementen des ehemaligen IWT der LFN auf dem Gebiet der bisherigen DDR geleistet werden.

Bei der Auswahl der Datenbanken wurden vor allem die berücksichtigt, die ein möglichst breites Profil aufweisen und für die Versorgung eines großen Nutzerkreises eingesetzt werden können. Dafür boten sich folgende Datenbanken an:

- AGRIS (Fachgebiet: Agrarwissenschaften; Hersteller: FAO-Mitgliedsländer)
- AGRICOLA (Fachgebiet: Agrarwissenschaften; Hersteller: National agricultural Library, USA)

- CAB ABSTRACTS (Fachgebiet: Agrarwissenschaften; Hersteller: CAB International, GB)
- FSTA (Fachgebiet: Ernährungswissenschaften; Hersteller: IFIS, GB und D)
- ASFA (Fachgebiet: Meereswissenschaften und Fischerei; Hersteller: FAO-Mitgliedsländer)

Auf eine genauere Beschreibung der Datenbanken wird verzichtet. Es sei auf weiterführende Publikationen verwiesen /16,19/.

In Tabelle 1 ist die Verteilung der CD-ROM-Stationen und der Datenbanken auf CD-ROM auf die neuen Bundesländer und Berlin (Ost) ausgewiesen. Bei der Berücksichtigung der Länderstruktur ergaben sich Schwierigkeiten dahingehend, daß die Forschungseinrichtungen im Bereich ELF für die neuen Länder eine ungleichmäßige Verteilung aufweisen, was z.B. im Raum Brandenburg/Berlin zu einer Häufung von Stationen und CD-ROM führte.

Tabelle 1: Verteilung der CD-ROM-Stationen und CD-ROM-Datenbanken nach Ländern

Land	Stationen			CD-ROM-Datenbank		
	(Anzahl)	AGRIS	AGRICOLA	CAB	FSTA	ASFA
Mecklenburg/Vorpommern	2	2	2	1	-	1
Brandenburg	5	4	4	3	1	-
Sachsen-Anhalt	3	3	3	2	2	-
Sachsen	2	2	2	2	-	-
Thüringen	1	1	1	1	-	-
Berlin(Ost)	2	2	2	1	2	1
insgesamt	15	14	14	10	5	2

Neben der territorialen Verteilung spielte die Vergabe an Einrichtungen der universitären und außeruniversitären Forschung eine Rolle (siehe Tabelle 2). In das Projekt wurden 5 Universitäten, 5 Institute der ehemaligen Akademie der Landwirtschaftswissenschaften und 5 sonstige wissenschaftliche Einrichtungen einbezogen.

Tabelle 2: Verteilung der CD-ROM-Stationen und CD-ROM-Banken nach universitärer und außeruniversitärer Unterstellung

Einrichtungen	CD-ROM-Datenbanken
1. Humboldt-Universität zu Berlin	AGRIS, AGRICOLA, FSTA
2. Universität Rostock	AGRIS, AGRICOLA, ASFA
3. Universität Leipzig	AGRIS, AGRICOLA, CAB
4. Friedrich-Schiller-Universität Jena	AGRIS, AGRICOLA, CAB
5. Martin-Luther-Universität Halle-Wittenberg	AGRIS, AGRICOLA, CAB, FSTA
6. Forschungszentrum für Tierproduktion Dummerstorf	AGRIS, AGRICOLA, CAB
7. Forschungszentrum für Bodenfruchtbarkeit Müncheberg	AGRIS, AGRICOLA, CAB
8. Institut für Landw. Information u. Dokumentation Berlin	AGRIS, AGRICOLA, CAB, FSTA, ASFA
9. Agrarwissenschaftl. Zentrum Chemnitz, Sitz Lichtenwalde	AGRIS, AGRICOLA, CAB
10. Forschungszentrum für Mechanisierung und Energieanwendung in der Landwirtschaft, Schlieben	AGRIS, AGRICOLA, CAB
11. Institut für angewandte Tierhygiene Eberswalde-Finow	AGRIS, AGRICOLA
12. Forschungsanstalt für Forst- und Holzwirtschaft Eberswalde	AGRIS, AGRICOLA, CAB
13. Institut für Milchwirtschaft Oranienburg	FSTA
14. Hochschule "Thomas Müntzer" Bernburg	AGRIS, AGRICOLA, CAB, FSTA
15. Institut für Rinderproduktion Iden-Rohrbeck	AGRIS, AGRICOLA

### 3. Ergebnisse des CD-ROM-Einsatzes

Insgesamt kann festgestellt werden, daß die Mehrzahl der Einrichtungen in kürzester Zeit den Betrieb aufnahm und in zunehmendem Maße eine erfolgreiche Nutzerversorgung absichern konnte. Begleitend dazu wurde die neue Möglichkeit der Informationsversorgung im Einzugsbereich der Informationseinrichtungen propagiert. Die im Antrag für das Förderprojekt getroffene Aussage, "daß die vorgeschlagene Möglichkeit der Bereitstellung umfassender Datenbanken unter Nutzung der CD-ROM-Technologie problemlos von den Informationseinrichtungen übernommen und genutzt wird", hat sich vollauf bestätigt.

Auswertungen von statistischen Erhebungen zur Nutzung der CD-ROM-Technologie können wertvolle Hinweise auf die Effektivität und Nutzerakzeptanz des Verfahrens vermitteln und ermöglichen Schlußfolgerungen für die weitere methodische, inhaltliche und technisch-technologische Gestaltung der Informationsarbeit /20,21,22,23,24,25/.

Auf Grund der geringen Laufzeit des Projektes (Projektbewilligung Dezember 1990, Übergabe der Stationen und CD-ROM-Datenbanken ab Februar 1991) sind natürlich abschließende Analysen nicht möglich. Angaben zur Zahl der durchgeführten Recherchen, thematischen Gebiete, Nutzergruppen und Nutzungshäufigkeit der Datenbanken stellen deshalb nur den unmittelbaren aktuellen Stand dar und erlauben keine umfassende Beurteilung. Dies bleibt späteren Publikationen vorbehalten.

Aus Rückinformationen wird deutlich, daß diese Form der Informationsversorgung von den Nutzern akzeptiert wird. Von Februar bis August 1991 wurden in den beteiligten Einrichtungen 2.672 Recherchen in den CD-ROM-Datenbanken durchgeführt (siehe Tabelle 3).

Tabelle 3: Anzahl der Recherchen in CD-ROM-Datenbanken von Februar bis August 1991

Einrichtung	Febr	März	April	Mai	Juni	Juli	Aug	Gesamt
1. Uni Berlin	-	-	-	19	100	46	29	194
2. Uni Rostock	-	-	10	13	58	95	39	215
3. Uni Leipzig	-	18	48	55	39	40	40	240
4. Uni Jena	-	-	28	36	31	48	27	170
5. Uni Halle	62	104	136	36	161	37	13	549
6. FZT Dummerstorf	12	15	-	10	16	37	5	95
7. FZB Müncheberg	-	-	14	8	7	4	6	39
8. ILID Berlin	-	-	11	23	3	8	12	57
9. AWZ Lichtenwalde	Stat.		Mai übergeben		23	4	3	30
10. FZM Schlieben	-	10	8	8	-	3	6	35
11. IaT Eberswalde	5	55	65	65	67	37	45	339
12. FA Forst/Holz Eberswalde	-	10	4	5	4	+) )	+) )	23
13. Inst. Milchwirt. Oranienburg	-	-	35	15	13	4	-	67
14. HS Bernburg	-	-	40	51	138	127	64	420
15. Inst. Rinderprod. Iden-Rohrbeck	-	-	15	43	39	41	61	199
Gesamt	79	212	414	387	699	531	350	2672

+) noch nicht gemeldet

In Tabelle 4 sind die Zahlen der mit Recherchen versorgten Nutzer ausgewiesen. Von Februar bis August wurden Recherchen für 1.272 Nutzer durchgeführt. Nach Auskunft der Informationseinrichtungen war die überwiegende Mehrzahl von Nutzern Wissenschaftler bzw. Studenten. Aus dem Bereich der Verwaltung kamen nur einzelne Nutzer. Potentiellen Nutzern aus Industrie und Landwirtschaft wurden vereinzelt Angebotsrecherchen geliefert. Im erfaßten Zeitraum lag der Schwerpunkt der Informationsversorgung im unmittelbaren Einzugsbereich der jeweiligen Einrichtung (interne Versorgung).

Der relative Anteil der Anfragen an wichtigen Gebieten stellt sich wie folgt dar:

- Tierproduktion/Veterinärmedizin	47 %
- sonstige Gebiete	18 %
- Pflanzenproduktion	12 %
- allgemeine Landwirtschaft	10 %
- Bodenkunde/Pflanzenernährung	7 %
- Wirtschafts-/Sozialwissenschaften	6 %

Tabelle 4: Anzahl der mit Recherchen versorgten Nutzer von Februar bis August 1991

Einrichtung	Feb	März	April	Mai	Juni	Juli	Aug	Ges.
1. Uni Berlin	-	-	-	19	88	29	16	152
2. Uni Rostock	-	-	8	12	4	7	4	35
3. Uni Leipzig	-	17	48	54	39	40	40	238
4. Uni Jena	-	-	28	36	31	16	9	120
5. Uni Halle	11	30	34	13	29	9	6	132
6. FZT Dummerstorf	7	10	-	9	10	10	3	49
7. FZB Müncheberg	-	-	14	8	6	3	5	36
8. ILID Berlin	-	-	11	22	3	5	12	53
9. AWZ Lichtenwalde	Stat.	Mai übergeben			23	4	3	30
10. FZM Schlieben	-	10	8	8	-	3	3	32
11. IaT Eberswalde	2	11	12	18	8	12	17	80
12. FA Forst/Holz Eberswalde	-	10	4	5	4	+) )	+) )	23
13. Inst. Milchwirt. Oranienburg	-	-	14	4	7	2	-	27
14. HS Bernburg	-	-	8	24	31	24	8	95
15. Inst. Rinderprod. Iden-Rohrbeck	-	-	24	48	27	35	36	170
<b>Gesamt</b>	<b>20</b>	<b>88</b>	<b>213</b>	<b>280</b>	<b>310</b>	<b>199</b>	<b>162</b>	<b>1272</b>

Tabelle 5 weist den Anteil der einzelnen CD-ROM-Datenbanken bei den Rechercheanfragen aus. Es wird deutlich, daß die drei thematisch breiten agrarwissenschaftlichen Datenbanken AGRIS, AGRICOLA und CAB am meisten genutzt wurden. Die geringere Nutzung von CAB rührt von der Tatsache her, daß diese CD-ROM nur 10 Einrichtungen zur Verfügung steht. Die geringe Nutzung der Datenbanken FSTA und ASFA ist bedingt durch die thematische Spezifität.

Tabelle 5: Anteil der CD-ROM-Datenbanken bei Rechercheanfragen von Februar bis August 1991

Datenbank	Febr	März	April	Mai	Juni	Juli	August	Gesamt
AGRIS	12	113	218	251	320	243	214	1371
AGRICOLA	28	105	217	268	382	269	176	1443
CAB	37	76	175	152	256	140	94	930
FSTA	-	13	72	44	67	43	4	243
ASFA	-	-	1	-	2	3	1	7

Die zum gegenwärtigen Zeitpunkt verfügbaren Daten erlauben nur eine erste grobe Schätzung aus der Anlaufphase des Förderprojektes. Es wurde jedoch deutlich, daß mit dem breiten Einsatz der CD-ROM-Technologie eine wesentliche Verbesserung der Informationsversorgung erreicht werden kann.

#### 4 Ausblick

Die CD-ROM-Technologie hat für den landwirtschaftlichen Bereich in den neuen Bundesländern ihren Platz gefunden. Die Realisierung des Förderprojektes des BMFT kann durchaus als Initialwirkung bezeichnet werden. Am Ausbau der Leistungsfähigkeit und an der Verbreiterung des thematischen Angebotes wird in den beteiligten Einrichtungen intensiv gearbeitet /26,27,28,29/.

Um einen möglichst intensiven und vielseitigen Einsatz der CD-ROM-Datenbanken zu erreichen, wurde durch das ILID ein Programm zum downloading der Rechercheergebnisse entwickelt und zur kostenlosen Nutzung an die Partnereinrichtungen übergeben. Die Grundlage dafür bildete das Programmsystem MIDOS-16, das in der Vergangenheit auf dem Gebiet der ehemaligen DDR eine große Verbreitung gefunden hatte, so daß in den meisten Informationseinrichtungen z.T. langjährige Erfahrungen bei der Arbeit mit diesem Programmsystem vorlagen und daraus eine unmittelbare Praxiswirksamkeit resultierte.

Ein weiteres Anwendungsfeld ist die Nutzung von CD-ROM-Rechercheergebnissen aus mehreren CD-ROM-Datenbanken zur Erstellung von gedruckten Profildiensten zu thematischen Gebieten, ebenfalls unter Nutzung der Möglichkeiten des Programmsystems MIDOS-16 bei der Manuskripterstellung.

Bei weiterführenden Überlegungen zum Einsatz der CD-ROM-Technologie ist zu berücksichtigen, daß diese Formen der Informationsversorgung nur eine, wenn auch sehr moderne Möglichkeit aus der breiten Palette der Informationsleistungen darstellt und immer im Zusammenhang mit dem Online-Zugriff zu Literaturdatenbanken auf Hosts zu sehen ist.



Auch in Zukunft werden die Nutzung von Online-Datenbanken und der Einsatz von CD-ROM nebeneinander Bestand haben. Der Nutzer wird letztendlich die Entscheidung fällen, welches Medium er aus Gründen der Zweckmäßigkeit anwendet. Es wird jeweils die Technologie eingesetzt, die bei langfristig geringeren Kosten den größten praktischen Nutzen bewirkt. Eine wichtige Fragestellung ist dabei der Vergleich der Vor- und Nachteile von Online- und CD-ROM-Datenbanken. In der zu diesem Problemkreis recht umfangreichen Literatur werden folgende Kriterien in den Mittelpunkt gestellt: Speicher- menge, Verfügbarkeit, Bedienungskomfort, Aktualität, Nutzungshäufigkeit, Preis und Kombination von Online- und CD-ROM-Recherchen /18,30,31,32,33,34,35,36/. Mit zunehmender Entwicklung der Kommunikationsinfrastruktur in den neuen Bundesländern erhalten diese Fragestellungen einen bestimmten Stellenwert und beeinflussen Arbeitsweisen in Informationseinrichtungen im Wechselspiel mit den Nutzern.

Eine wichtige Voraussetzung zur effektiven Nutzung von verschiedenen Online- oder CD-ROM-Datenbanken, auch in Kombination, ist, den Nutzern und Informationsvermittlern Werkzeuge in die Hand zu geben, die es ihnen ermöglichen, Recherchen durchzuführen, ohne sich auf die Unterschiede von Datenbanken, Retrievalsprachen und spezifischen Anforderungen der Hosts einzustellen zu müssen. Diese Überlegungen führten zu der Idee, ein Expertensystem zur Unterstützung des Retrievals in landwirtschaftlichen Datenbanken zu schaffen /37/.

In einer Reihe von Informationseinrichtungen des IWT der LFN wurden spezifische Datenbanken betrieben, die sich in thematischer Hinsicht teilweise mit der im ILID Datenbank AGROSELEKT überschneiden, aber auch Themenbereiche erfaßten (z.B. Nahrungsgüterwirtschaft, Ökonomie oder Forstwirtschaft), die nicht für die Datenbank AGROSELEKT erschlossen wurden. Diese Datenbanken werden in der Mehrzahl heute noch betrieben und haben, unter Berücksichtigung der notwendigen Veränderungen bei der Vereinheitlichung der Informationslandschaft, z.T. weiterhin ihre Daseinsberechtigung. An dieser Stelle sei darauf hingewiesen, daß es Überlegungen dahin gehend gibt, wertvolle Datenbestände aus dem ehemaligen IWT der LFN durch die Übernahme auf CD-ROM für eine

weitere Nutzung zu sichern und sie auf diese Weise einem breiten Nutzerkreis anzubieten. Zu diesem Zweck wird durch das ILID in Zusammenarbeit mit der ZADI ein Projektvorschlag erarbeitet. Es bleibt zu hoffen, daß die Möglichkeit, zum gegebenen Zeitpunkt über die Realisierung dieses Vorhabens berichten zu können, keine Utopie bleibt.

#### Literatur

- /1/ Informationstechnik im Agrarbereich. Entwicklungsstand und Handlungsbedarf. - BML, Bonn, 1990. - 127 S.
- /2/ Fachinformationssystem Ernährung, Land- und Forstwirtschaft. - ZADI, Bonn, 1991. - 12 S.
- /3/ Selle, K.O.v.; Vogels, J.C.: Das Datenbankangebot im Fachinformationssystem Ernährung, Land- und Forstwirtschaft (FIS-ELF). - In: Agrarinformatik. - Stuttgart 19(1990). - S.109-116
- /4/ König, E.: Die Datenbasis Ernährungs-, Land- und Forstwirtschaftliches Informationssystem (ELFIS). - In: Mitt. GBDL. - Karlsruhe (1988)42. - S.43-50
- /5/ Koehler, C.U.; Winkelmann, U.: Nutzung von Datenbanken im Agrarbereich heute und in der Zukunft. - In: Forschungsreport Ernährung, Landwirtschaft, Forsten. - Bonn(1990)5. - S. 14-17
- /6/ Münch, E., Kutschke, G.: Rechnergestützte Informationsverarbeitung in der LFN. - In: Tag.-Ber., Akad. Landwirtsch.-Wiss. DDR. - Berlin (1987)255. - S. 201-208
- /7/ Münch, E.: EDV-Einsatz im IWT der LFN. - In: Mitt. ILID. - Berlin (1987)3. - S. 3-8
- /8/ Thulke, M.: Zu den Leistungen des Zentralen Fonds. - In: Mitt. ILID. - Berlin (1989)1. - S. 40-44
- /9/ Pätzsch, M.: Leistungen des ZF AGROSELEKT mit seinen Dateien A, B und C. - In: Mitt. ILID. - Berlin (1988)2. - S. 38-42
- /10/ Krüger, H.-D.: Möglichkeiten zur Arbeit mit dem Datenbanksystem DIALOG-2. - In: Mitt. ILID. - Berlin (1989)2. - S.32-34
- /11/ Rindfleisch, B.: Optische Speicherplatten - ein neues Informationsmedium - In: Berichte für Wissenschaftsinformation und -kommunikation. - AdW, Berlin (1989)3. - S 371 S.
- /12/ Sonntag, H.: CD-ROM-Einsatz in der Wissenschaftsinformation. - In: Informatik. - Berlin 36(1989)3. - S. 114-116
- /13/ Reddman, H.-D.: CD-ROM- und PC-Technologie für Speicherung und Recherche von Sachinformationen. - In: Informatik. - Berlin 37(1990)3. - S.109
- /14/ Kneitzschel, F.: CD-ROM - ein neues Medium in der Informationsversorgung. - In: Informatik. - 37(1990)3. - S. 91-93
- /15/ Kneitzschel, F.: CD-ROM: Immer breitere Anwendung, - In: Informatik. - Berlin 38(1991)3. - S. 115-118
- /16/ Winkelmann, U.; König, E.: Agrardatenbanken auf CD-ROM. - In: Agrarinformatik. - Stuttgart 19(1990). - S. 227-234
- /17/ Böckly, R.; Michels, H.: Ersetzen Datenbanken auf neuen PC-Speichermedien die klassischen Datenbanken am Großrechner? - In: Agrarinformatik. - Stuttgart 19(1990). - S. 235-240
- /18/ Klaes, G.: Online oder CD-ROM. - In: Cogito. - (1990)1. - S. 4-12
- /19/ Deseleers, H.; König, E.: AGRICOLA, AGRIS und CAB im Vergleich. - In: Mitt. GBDL. - Karlsruhe (1985)37. - S. 29-44

- /20/ Lobeck, M.A.: Erfahrungen mit CD-ROM in einer Industrie-Informationsabteilung. - In: Nachr. Dok. - 40(1989)3. - S. 137-150
- /21/ Gloecker-Rist, A.; Lehmler, W.; Wettler, M.: Endnutzerrecherchen in Bibliographischen CD-ROM-Datenbanken. - Nachr. Dok. - 40(1989)3. - S. 151-159
- /22/ Bowman, F.: Selbstrecherchieren - Pro und Contra. Erfahrungen mit wissenschaftlichen und mit Literaturdatenbanken. - In: Nachr. Dok. - 40(1989). - S. 77-82
- /23/ Kaltenborn, K.-F.: Endnutzerrecherchen in der CD-ROM-Datenbank Medline. Teil 1. - In: Nachr. Dok. - 42(1991)2. - S. 107-114
- /24/ Kaltenborn, K.-F.: Endnutzerrecherchen in der CD-ROM-Datenbank Medline. Teil 2. - In: Nachr. Dok. - 42(1991)3. - S. 177-190
- /25/ Wißmann, Ch.: Online- und CD-ROM-Literaturrecherchen an der Universitätsbibliothek und der Zentralen Medizinischen Bibliothek Marburg. - In: Bibliotheksdienst. - 25(1991)2. - S. 196-201
- /26/ Material zu den Fragen an die außeruniversitären Forschungseinrichtungen. - ILID, Berlin 1991. - 91 Seiten
- /27/ Werner, M.: Konzeption zur Ausgestaltung der fachlichen Rahmentätigkeit. - ILID, AG Osteuropainformation, Berlin 1991. - 11 S.
- /28/ Dornick, H.: 25 Jahre Informationsarbeit am Institut für Forstwissenschaften und wie soll und kann es weitergehen? - In: IFE-Berichte zur Forschung und Entwicklung. - Eberswalde-Finow 1990, 21. - 19 S.
- /29/ Wolf, J.; Große, R.; Müller, U.: Leichter Griff zur Datenbank. - In: Deutsches Landblatt Nr. 125 v. 02.06.1991. - S. 7
- /30/ Thuss, J.: CD-ROM: Wie schnell geht die Reise? - In: Nachr. Dok. - 37(1986)6. - S. 286-287
- /31/ Meyer, R.: From Online to Ondisc. - In: CD-ROM - the new papyrus. The current and future state of the art. - Redwood, 1986. - S. 467-481
- /32/ Funk, R.: Wieweit ersetzt die CD-ROM die Online-Datenbank? - In: 10. Frühjahrstagung der Online-Benutzergruppe der DGD. - Frankfurt/Main, 1988. - S. 345-359
- /33/ Hatvany, B.: Comparison of CD-ROM and Online. - In: Online Information 87. Proceedings, 11th international online information meeting. - Oxford, 1988, - S. 285-290
- /34/ Zick, W.: Zehn Legenden über die CD-ROM. - In: Nachr. Dok. - 39(1988). - S. 116-119
- /35/ CD-Technik und Anwendungen. - In: Cogito. - (1988)1. - S. 32-37
- /36/ Manak, A.: CD-ROM: Das neue Medium für Fachinformation. - In: Fakten, Daten, Zitate. - Wien 8(1988)1. - S. 6-8
- /37/ König, E.; Koehler, C.U.: Unterstützung der Retrieval in landwirtschaftlichen Datenbanken durch ein Expertensystem. - In: Agrarinformatik. - Stuttgart 20(1991). - (im Druck)

Dr. Eberhard Münch  
 Dr. Michael Thulke  
 Institut für Landwirtschaftliche  
 Information und Dokumentation  
 Krausenstraße 38/39  
 O - 1086 Berlin

Dr. Eckehard König  
 Zentralstelle für  
 Agrardokumentation und -information  
 Villichgasse 17  
 5300 Bonn 2

# **Einsatz eines optischen Speichersystems in einer Wirtschaftspresseudokumentation**

Martin Michelson

## **Zusammenfassung**

Die Wirtschaftspresseudokumentation in der zentralen IuD-Abteilung einer Großbank ist und bleibt ein wichtiger Bestandteil einer effizienten Informationsversorgung von Haus und Kunden. Herkömmliche Erschließungs-, Ablage- und Nutzungsverfahren bilden allerdings mehr und mehr einen Schwachpunkt der Infrastruktur. Zur Verbesserung der Informationsversorgung, zur Arbeitserleichterung und zur Raumeinsparung hat die IuD-Abteilung der Dresdner Bank ihre Wirtschaftspresseudokumentation auf ein System auf Basis der optischen Speichertechnologie umgestellt.

Voraussetzungen, konzeptionelle Überlegungen, Planung und Durchführung werden erläutert, die gewählte Systemkonfiguration, die software- und hardwaretechnische Lösung dargestellt.

Die bisher gemachten Erfahrungen lassen sich an den gesetzten Zielen messen. Ein erstes Resümee ist, daß entsprechende Anwendungen auch für andere Banken und Unternehmen mit einer Wirtschaftspresseudokumentation erfolversprechend erscheinen.

## **Die Presseudokumentation im bankbetrieblichen Informationswesen**

In den bankbetrieblichen Informations- und Dokumentationsabteilungen spielt die herkömmliche Presseudokumentation nach wie vor eine wichtige Rolle. Banken und ihre Kunden benötigen Wirtschaftsinformationen im breiten Umfang. Die Hintergrundberichte der Wirtschaftspresse sind als Informationsträger kaum zu ersetzen. Daher führen zumindest die größeren Institute mehr oder weniger umfangreiche Archive, in denen über Jahre zurück Materialien vorgehalten werden. Zwar ist die Presseudokumentation mit physischer Ablage der Dokumente im Zeitalter der Online-Recherche nur mehr ein Teilbereich der möglichen Informationsmittel, aber dieser Teilbereich wird weiterhin gepflegt.

Die zentrale Informations- und Dokumentationseinrichtung der Dresdner Bank betreibt zur Bewältigung der an sie gestellten Anforderungen unter anderem eine umfassende Wirtschaftspresseudokumentation. In arbeitsteiliger Zusammenarbeit mit anderen Banken werden täglich die wichtigsten überregionalen Zeitungen und Zeitschriften mit wirtschaftlichem Schwerpunkt des In- und teilweise des Auslandes ausgewertet, die relevanten Artikel erschlossen, auf DIN-A4 umbrochen und in die Ablagesysteme jedes Hauses geordnet und damit zur Verfügung gehalten.

Zu unterscheiden sind die Bereiche Sachdokumentation, Firmendokumentation und Personendokumentation, sie waren bisher in der Dresdner Bank auch räumlich und arbeitsorganisatorisch getrennt.

Die Sachdokumentation enthält Artikel zu einem weitgefächerten Themenbereich mit den Hauptgruppen Politik, Volkswirtschaft, Kreditwesen, Wirtschaftszweige und Betriebswirtschaft. Die Erschließung und Zuordnung der Artikel erfolgt nach einer mit den Partnern gemeinsam entwickelten und genutzten Klassifikation.

Ablage und Zugriff bestand bisher in der Dresdner Bank in gekennzeichneten Schubfachboxen, die sich wiederum in offenen Schränken befanden. Inland und Ausland waren in unterschiedlichen Räumen untergebracht, jeweils mit Mitarbeiterschreibtisch, Lesetisch und Kopierer. In den je 15 Schränken mit 1100 Schubfächern hatte Material von bis zu drei Jahren zurück Platz, älteres Material wurde regelmäßig ausgelagert.

Die Firmendokumentation war nach Ländern und alphabetisch, innerhalb der einzelnen Firma chronologisch geordnet. Ablagemittel war hier der Ringordner. Die Firmendokumentation beanspruchte drei Räume. Auch hier wurde älteres Material in einen Archivkeller ausgelagert.

Die Personendokumentation war in acht Schränken in der Bibliothek untergebracht. Ablage- und Suchkriterium in den Ringordnern bildete allein der Name. Wegen des geringeren Mengenanfalls und der geringeren Aktualitätsbezogenheit der Dokumente gab es hier keine Auslagerung.

Der Zuwachs aller drei Dokumentationsbereiche beläuft sich auf etwa 500 Dokumente pro Arbeitstag. Jeden Tag werden insgesamt rund 100 Recherchen durchgeführt.

## **Grenzen des herkömmlichen Verfahrens**

Mehrere Faktoren führten zu intensiven Überlegungen, wie die Pressedokumentation neuen Anforderungen anzupassen ist. Der Service der IuD-Einrichtung wird nicht nur von den Mitarbeitern der Dresdner Bank in immer stärkerem Maße genutzt, vor allem auch von den nicht in der Zentrale arbeitenden, sondern steht auch den Kunden der Bank zur Verfügung. Der Wert von geschäftsbegleitenden Informationen, die - bisher noch - unentgeltlich zur Verfügung stehen, wird von immer mehr Externen geschätzt und genutzt. Die Pressedokumentation bleibt einer der wichtigsten Quellen für den steigenden Informationsbedarf. Die bisherige Struktur konnte, wenn sie auch im Rahmen des immanent Möglichen angepaßt wurde, den steigenden Anforderungen auf die Dauer nicht mehr gerecht werden.

Besonders die Mitarbeiter sollten in der Lage sein, möglichst viele Leistungen der IuD-Einrichtung über moderne Kommunikationsnetze von ihrem Arbeitsplatz in Anspruch zu nehmen, und das in "Echtzeit", also ohne Zwischenschaltung von IuD-Mitarbeitern, die mehr dem Kundenbedarf zur Verfügung stehen sollten.

Zudem erforderte eine räumliche Neuordnung der Zentrale in Frankfurt rasches Handeln. Für den Sommer 1991 stand der Umzug der Geschäftsleitung samt wichtiger Stabsabteilungen in ein neues Gebäude an. Als organisatorisch zum Bereich (Hauptabteilung) Volkswirtschaft und Marktforschung gehörende und stark von den zentralen Abteilungen genutzte Einrichtung sollte die Information und Dokumentation mit in das neue, leistungsfähig verkabelte Gebäude ziehen - aber aus Platzgründen ohne große Archivbestände.

Schließlich bestehen schon längerfristig Überlegungen, die in der gesamten Bank und den Tochtergesellschaften zahlreich vorhandenen papiergebundenen Bestände auf ein platzsparendes und für ein Bürokommunikationssystem geeignetes Speichermedium zu bringen. Die IuD-Abteilung wurde ausgewählt, um das Pilotprojekt durchzuführen.

Vergegenwärtigen wir uns noch einmal die Grundzüge und die Grenzen des bisherigen Systems:

Die Pressedokumentation der IuD-Abteilung enthielt im direkten Zugriffsbereich ca. 200 laufende Meter Presseschnitte in der Sachdokumentation, 350 laufende Meter in der Firmendokumentation und 50 laufende Meter in der Personendokumentation. Zu diesen Beständen, die zwischen drei und fünf Jahren umfassen (in der Personendokumentation das komplette Material seit Beginn der Sammlung) kommen die in verschiedene Archivkeller ausgelagerten Materialien von ca.

weiteren fünf Jahren. Alles in allem bestand ein nicht unerheblicher Raumbedarf innerhalb einer zentralen und daher teuren Gegend in Frankfurt.

Der Zugriff auf die ausgelagerten Bestände war (und ist) umständlich. Das Ordnen und Aussortieren in den Kellern kann nur sporadisch geschehen, es herrscht notgedrungen nur eine grobe, im Prinzip zeitlich orientierte Ordnung. Ältere als in der Regel 10 Jahre zurückdatierende Dokumente werden weggeworfen.

Die in der Pressedokumentation vorgehaltenen Dokumente waren z.T. durch häufiges Rein- und Rauslegen recht abgenutzt. Zudem führte die Möglichkeit der Selbstbedienung durch den Nutzer zu häufigen Fehlablagen. Dokumente, die sich am falschen Ort oder nicht in der richtigen chronologischen Folge befanden, entzogen sich weitgehend dem Zugriff.

In der großen Mehrheit der Fälle werden Anfragen allerdings von den IuD-Mitarbeitern bearbeitet. Abgesehen von der Ablage, die in jedem Speichersystem in irgendeiner Weise erfolgen muß, haben die rein manuellen Tätigkeiten bei einer Recherche in der Pressedokumentation viel Zeit gekostet: Aufsuchen der entsprechenden Ablageorte, Herausnehmen der Schubfächer, Durchsuchen der Stöße nach relevanten Dokumenten, Gang zum Kopierer, Anfertigen von Kopien, Aufsuchen der Ablageorte und Zurücksortieren der Originalbelege.

Eine Mehrfachablage entsprechend der verschiedenen Aspekte eines Artikels war nur begrenzt möglich. Bei der Recherche ließ das eindimensionale Ablagesystem einen verknüpften Zugang nicht zu. So war es z.B. in der Firmendokumentation nicht möglich, ohne zeitraubende Suche Artikel zu selektieren, die sich mit der Fusion zweier Firmen befassten.

## **Planung der Umstellung**

Als neue Speicher- und Organisationsform kam nur ein System auf Basis der optoelektronischen Speichertechnologie in Frage. Um einen Eindruck zu gewinnen und Erfahrungen zu nutzen besuchten verantwortliche Mitarbeiter der IuD-Abteilung Organisationen mit bereits bestehenden Anwendungen (u.a. Gruner + Jahr, AEG, Staatskanzlei NRW). Parallel studierte man Prospekte von Anbietern und besuchte Vorführungen.

In Zusammenarbeit mit dem Münchner Softwarehaus iXOS wurde von der IuD-Abteilung und der Organisation der Dresdner Bank ein Anforderungsprofil für ein den künftigen Aufgaben der Pressedokumentation entsprechendes System erar-

beitet. Die wesentlichen Aussagen der Studie fanden Berücksichtigung in einem Pflichtenheft für Hard- und Software. Es seien hier einige Punkte genannt:

#### Hardware:

Wegen der zur erwarteten Standardisierung sind Jukeboxen mit 5 1/4"-WORM-Platten zu verwenden.

Die Auflösung beim Scannen, bei der Anzeige und im Ausdruck wird auf 300 dpi festgesetzt.

Die Bildschirme sollen wenigstens eine Diagonale von 19 Zoll besitzen.

Die Druckerleistung muß wenigstens zehn Seiten DIN A4 pro Minute betragen.

#### Software:

Als Betriebssystem geeignet erscheint UNIX V. Es entspricht den Anforderungen von X/OPEN und ermöglicht damit durch feste und bekannte Schnittstellen die Erweiterbarkeit der Anlage.

Die Dokumente sind in komprimierter Form zu speichern und im Netz zu übermitteln.

Für die Übertragung im lokalen Netzwerk gilt TCP/IP als Protokoll.

Für das Retrieval soll eine Datenbank mit SQL-Schnittstelle eingesetzt werden.

Die Benutzeroberfläche ist so zu gestalten, daß sie auch für ungeübte Benutzer leicht zu handhaben ist. Sie ist durch Tastatur und durch Maus zu bedienen. Geeignet erscheint OSF/Motif, das sich im UNIX-Bereich als Standard etabliert hat. OSF/Motif basiert auf X-Windows und bietet damit die Sicherheit eines etablierten Produktes.

Es muß die Möglichkeit bestehen, Dokumente direkt aus dem System heraus per Fax zu übertragen.

Der Rechercheur soll in der Lage sein, einen Drucker im Netz auszuwählen und den Druckauftrag an diesen Drucker zu geben.

Es wurden mehrere Angebote über die Hardwarekonfiguration und Software eingeholt, die Entscheidung fiel zugunsten des Angebots der Firma SNI. Das Softwareunternehmen iXOS entwickelte auf der Basis des relationalen Datenbanksystems Informix eine speziell auf die Bedürfnisse der Pressedokumentation zugeschnittene Anwendungssoftware. Die in den Anwendungsmenues enthaltenen Vorgaben wurden dabei in Anlehnung an bewährte bisherige Ordnungsmerkmale und Verfahren gewählt.

Bis es soweit war, bedurfte es aber bei der Entwicklung intensiven Austausches zwischen der betroffenen Fachabteilung der Dresdner Bank, der Organisation der Bank, die das Projekt in ihre langfristige und bankweite Systemplanung



einbezieht und den Experten des Softwarehauses. Im Winter und im Frühjahr 1991 erfolgte das "Prototyping", aus dem schließlich die Anwendungsfunktionen und die Benutzeroberfläche, wie sie sich heute darbieten, hervorging.

Die Programmfunktionen sind vom Systemadministrator zu beeinflussen. Es stehen eine Reihe von Prozeduren zur Verfügung, die es erlauben, die Anwendungsprogramme zu überwachen, zu gestalten und den sich ändernden Bedürfnissen anzupassen. So können z.B. Listen geändert, Felder neu definiert oder den verschiedenen Arbeitsstationen variierende Funktionen zugeordnet werden. Zur Systemadministration gehört auch das Rauf- und Runterfahren der Anlage und das Importieren und Exportieren von Platten.

### **Durchführung der Umstellung**

Um in der Einführungsphase das neue System gründlich testen und Modifikationen vornehmen zu können, war es notwendig, einen genügend großen Bestand aus allen drei Dokumentationsbereichen zu erfassen und für Recherchezwecke zu speichern. Da die Dokumentationsbestände in ihrem bisherigen Umfang nicht in die neuen Räumlichkeiten mitgenommen werden konnten, mußte für einen reibungslosen Übergang wenigstens ein Bestand von einigen Monaten gespeichert sein, und das rechtzeitig vor dem Umzug im August 1991. Der Betrieb des bisherigen Verfahrens, das Erfassen der täglichen Neueingänge, das notwendige Rückerfassen älterer Bestände bei Verfügbarkeit der Anlage im Juni und schließlich die organisatorischen Maßnahmen der Umstellung waren gleichzeitig nicht zu bewältigen. Daher wurde die Rückerfassung bis zum Stichtag 1. Juli, an dem die laufende Erfassung begann, extern in Auftrag gegeben. Bei Inbetriebnahme der Anlage im neuen Gebäude stand die Personendokumentation komplett und die Sachdokumentation seit Anfang 1990 auf optischem Speicher zur Verfügung. In der Firmendokumentation erfolgte keine Zurückerfassung von Dokumenten.

Das Anlernen der Mitarbeiter erwies sich als relativ problemlos. Die Mehrheit hieß die Umstellung willkommen und zeigte sich an den Funktionen und Möglichkeiten interessiert. Die Übernahme einzelner Aufgaben im Rahmen der neuen Organisationsformen geschah weitgehend freiwillig und unter gegenseitiger Absprache.

Weniger Akzeptanz hingegen fand das System und die neue Organisation hingegen bei vielen der Direktnutzer, zumindest in der Einführungsphase. Mittlerweile scheinen die Vorzüge des neuen Verfahrens die meisten Kollegen überzeugt zu haben.

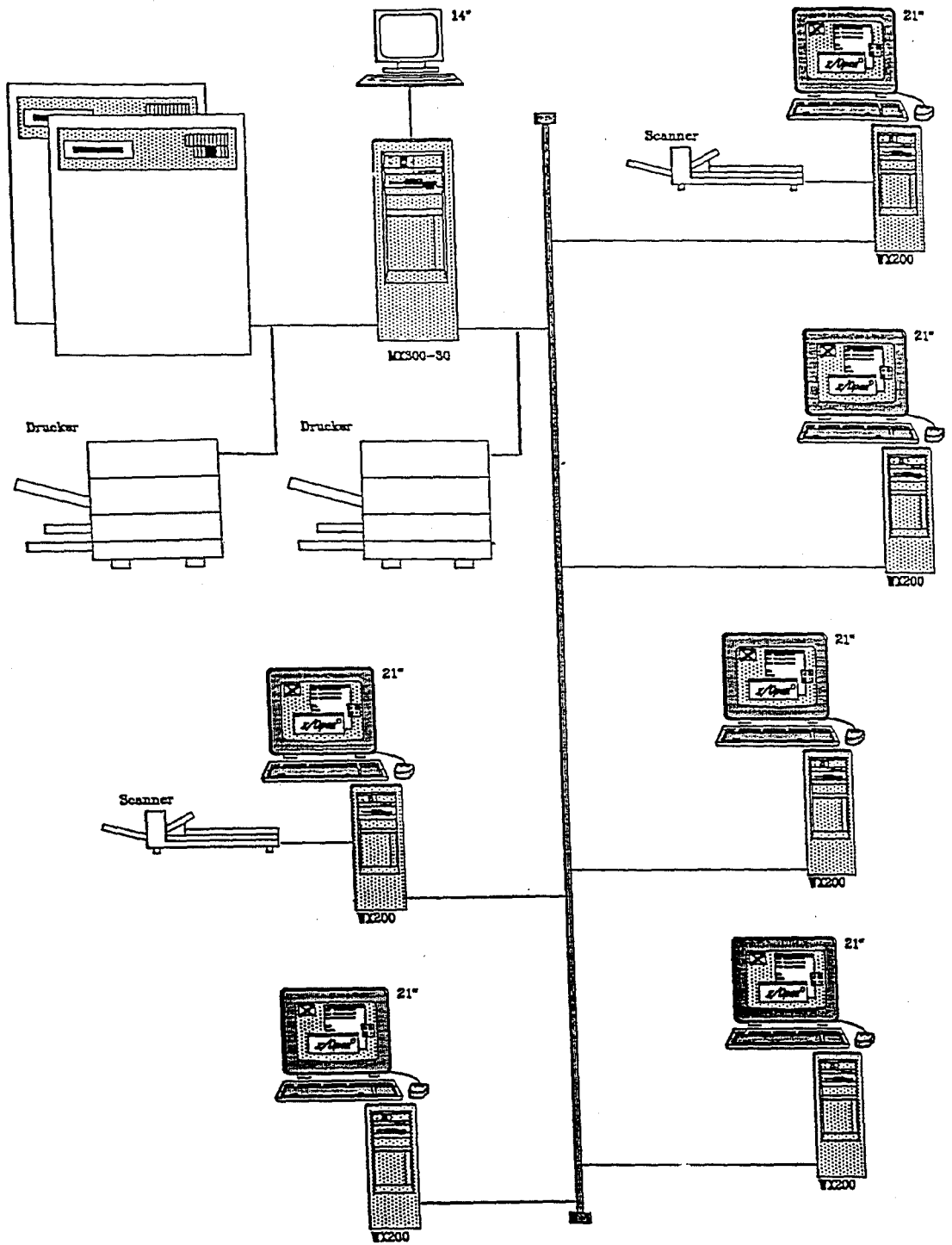


Abb. 1: Gesamtkonfiguration

Zwei für die Anlage verantwortliche Mitarbeiter erhielten eine besondere Schulung in der Systemadministration und sind in der Lage, komplexere Eingriffe ins System vorzunehmen und auch einfachere Störungen zu beheben.

## **Organisatorische Aspekte**

Zur Zeit ist folgende Hardwarekonfiguration in Betrieb:

Zwei Jukeboxen mit einer Kapazität von jeweils 24 5 1/4"-WORM-Platten, zwei Scanner, zwei Laserdrucker, ein MX 300 Server, sechs WX 200 Arbeitsstationen mit hochauflösenden DIN A3-Bildschirmen, davon je zwei zum Recherchieren, Recherchieren/Indexieren, Scannen.

Die Geräte verteilen sich auf zwei Arbeitsräume. Die reinen Recherchierplätze sind in erster Linie für die nicht zur IuD-Abteilung gehörenden Direktnutzer gedacht. Die Netzinstallation ist dabei so ausgelegt, daß jederzeit zumindest im Neubau weitere Rechercestationen an den Arbeitsplätzen von Mitarbeitern zugeschaltet werden können.

Softwareseitig sind alle Arbeitsvorgänge am System, ob Scannen, Indexieren, Recherchieren oder Ausgeben weitgehend menuegeführt. Es stehen Listen zur Verfügung, die nach jedem Dokumentationsbereich getrennt Vorgaben enthalten zur einfachen Übernahme in das jeweilige Bearbeitungsfeld. Zu ihnen gehören z.B. in der Sachdokumentation die Klassifikation, das Länderverzeichnis, die möglichen Quellen und zusätzliche beschreibende Schlagworte. Zum Ansteuern von Menüefeldern und Listenwerten ist die Maus bei einiger Übung das schnellste Hilfsmittel.

Bei der Recherche erhält man zunächst eine Trefferliste mit den wichtigsten inhaltlichen und formalen Angaben zum Dokument. Der Rechercheur trifft eine Auswahl und läßt sich die Dokumente am Bildschirm in Orginalgröße, ggf. auch mit Ausschnittsvergrößerung anzeigen. Am Ende steht der Ausdruck oder - und das ist sehr hilfreich - die direkte Faxübermittlung.

Zur Arbeitsorganisation:

Zwei Lektoren (je für die Sachdokumentation und die Firmen-/Personendokumentation) werten die Publikationen, die in der Bankenzusammenarbeit von der Dresdner Bank übernommen sind, aus und geben sie zur technischen Bearbeitung weiter. Die auf DIN-A4 umbrochenen Artikel werden dann von den Lektoren indexiert und schließlich mehrfach zur Liefe-

**Recherche**

<input checked="" type="checkbox"/> <b>Klassifizierung</b> <input style="width: 150px;" type="text" value="K"/> Langtext: <input style="width: 150px;" type="text" value="Monetärer Bereich, Kredit"/> <input type="checkbox"/> <b>Länderkennung</b> <input style="width: 150px;" type="text"/> Langtext: <input style="width: 150px;" type="text"/> <input checked="" type="checkbox"/> <b>Quelle</b> <input style="width: 150px;" type="text" value="IFO-Institut für Wirtscha"/> <input checked="" type="checkbox"/> <b>Datum</b> von: <input style="width: 50px;" type="text" value="1.1.1990"/> bis: <input style="width: 50px;" type="text" value="31.3.1990"/> <input type="checkbox"/> <b>Formales</b> <input style="width: 30px;" type="text" value="---"/> <input style="width: 30px;" type="text"/>	<input checked="" type="checkbox"/> <b>Schlagworte</b> <input style="width: 150px; height: 50px;" type="text"/> <input style="width: 150px; height: 20px;" type="text"/> <input style="width: 150px; height: 20px;" type="text"/> <input type="button" value="Löschen"/> <input style="width: 150px; height: 20px;" type="text"/> <input type="button" value="Suchen"/> <input type="checkbox"/> <b>Institution</b> <input style="width: 150px; height: 50px;" type="text"/> <input style="width: 150px; height: 20px;" type="text"/> <input style="width: 150px; height: 20px;" type="text"/> <input type="button" value="Löschen"/> <input style="width: 150px; height: 20px;" type="text"/> <input type="button" value="Suchen"/>
--	---

Abb. 2: Recherchemaske

**Trefferliste**

Funktionen	Bearbeiten		Datum	Form
<b>Klassifizierung</b>	<b>Länderkennung</b>	<b>Quelle</b>		
B	1250	Arbeit und Soziales	15.05.1991	Zch
K	2101	Das Wertpapier	21.05.1991	
K 0	1120	Der Volks- und Betriebswirt	22.05.1991	
	1000	Frankfurter Neue Presse	14.04.1991	
	1000		22.05.1991	
V 12	1000	Frankfurter Allgemeine Zeitung		
V 7.0	5000	Männer der Wirtschaft		
B	2000	Arbeit und Soziales	14.05.1991	
K 5.0	6500	Der Spiegel		Zch
K	1250	British Business	15.05.1991	
V	1250	Arbeit und Soziales	22.05.1991	

96 Dokumente gefunden

Abb. 3: Trefferliste

rung an die anderen Banken kopiert. Wenn die externen Lieferungen eingehen, sichten die Lektoren den gesamten Tagessatz auf Redundanz und geben die verbliebenen Dokumente - im Tagesdurchschnitt jeweils 100 - 150 in der Sach- und der Firmendokumentation, 15 - 20 in der Personendokumentation - mit teilweise zusätzlichen handschriftlichen Merkmalen versehen zur weiteren Bearbeitung. Ein Sachbearbeiter legt die Dokumente stapelweise in den Scanner und überprüft die Wiedergabe in Originalgröße am Bildschirm. Am separaten Indexierplatz werden sie entweder zeitgleich oder später inhaltlich gekennzeichnet. Vor der endgültigen Abspeicherung des Dokumentes auf der optischen Platte überprüft der Bearbeiter noch einmal die korrekte Zuordnung.

In jeder Phase der Bearbeitung ist eine Korrektur von Imagequalität oder Indexierung möglich. Auch eine spätere Korrektur der Einträge läßt sich bewerkstelligen. Die Überschreibung der Einträge erfolgt allerdings nur in der Datenbank auf der Festplatte. Die parallelen Einträge auf der optischen Platte bleiben erhalten, werden aber systemseitig nicht mehr angesteuert.

Unmittelbar nach der Eingabe stehen die erfaßten Dokumente für alle Rechercestationen zur Verfügung.

## **Erfahrungen**

Beginnen wir mit den Vorteilen der Neuorganisation der Wirtschaftspresseudokumentation gegenüber dem alten Verfahren. Es sei hier bewußt der Begriff "Neuorganisation" gebraucht, da die Veränderungen weit über die Einführung eines neuen technischen Verfahrens hinausgehen.

Der augenscheinlichste Vorteil ist die Platzeinsparung und die damit verbundene Gelegenheit, die verschiedenen Dokumentationsbereiche und ihre Mitarbeiter in den neuen Räumlichkeiten der IuD-Abteilung zusammenzuführen. Ein Teil des eingesparten Platzes wird übrigens für großzügig bemessenere Arbeitsplätze eingesetzt, mit dem durchaus zu beobachtenden Effekt einer höheren Arbeitszufriedenheit gerade der "technischen" Kräfte.

Wegen des geringen Raumbedarfs ist es jetzt möglich, gespeicherte Dokumente praktisch unbegrenzt lange am Ort des Geschehens zur Verfügung zu halten. Sind alle Platten auf den Jukeboxen voll, so besteht entweder die Möglichkeit, eine weitere Jukebox zu installieren (bzw. bei der raschen Entwicklung der optischen Speichertechnologie die Jukeboxen gegen speicherintensivere auszutauschen) oder noch einfacher die ältesten Platten zu exportieren, ins Regal zu stellen und damit Platz für neue zu schaffen. Auf der Datenbank sind auch die Inhalte

exportierter Platten nachgewiesen. Bei Bedarf kann dann die entsprechende Platte ohne größeren Aufwand wieder ins System importiert werden.

Ein zweiter offensichtlicher Vorteil ist die Arbeits- und damit Zeitersparnis bei den manuellen Tätigkeiten. Während der Erfassungs- und manuelle Indexiervorgang etwas länger dauert als die bisherige Ablage, entfällt sämtliches Raus- und Wiederrücksortieren sowie das Kopieren bei einer Recherche. Nicht zu unterschätzen ist auch die Einsparung der mit dem Betrieb eines großen Papierarchives einhergehenden laufenden Ordnungsmaßnahmen.

Die Pressedokumentation ist auf Grund der vorgenannten Einsparungen aktueller. Je nach personeller Situation und nach Arbeitsanfall gab es bisher zuweilen erhebliche zeitliche Diskrepanz zwischen Erscheinen und Verfügbarkeit eines Artikels.

Die datenbankgestützte Dokumentbeschreibung erlaubt eine spezifischere Indexierung und damit eine komfortablere und genauere Recherche. Dies macht die bisher übliche Durchsicht größerer Stapel von Papier auf der Suche nach bestimmten Informationen weitgehend überflüssig. Recherchen mit verknüpften Suchmerkmalen sind jetzt möglich und üblich.

Es gibt natürlich auch Nachteile gegenüber dem bisherigen Betrieb.

Die Pressedokumentation, und hier in erster Linie der Bereich Sachdokumentation hat traditionell viele Selbstnutzer, vor allem aus der Volkswirtschaftlichen Abteilung. Der Fachreferent kann naturgemäß bei seinem jeweiligen Informationsbedarf am besten selbst beurteilen, was er benötigt. Einfach und schnell ging das Durchblättern der relevanten Stapel. Hier traf der selbst recherchierende Nutzer zuweilen auf Randgebiete, die durchaus interessant für die Problemstellung waren, in einem Rechercheauftrag aber nicht formuliert worden wären. Das Blättern zwischen den angezeigten Dokumenten ist am Bildschirm wenig sinnvoll, da es technisch bedingt zu viel Zeit benötigt. Das System ist ja auch so konzipiert, daß man sich die Artikel, die man ansehen bzw. ausdrucken will, nach der aus der Datenbank generierten Trefferliste auswählt. Die inhaltlichen Angaben auf der Trefferliste können naturgemäß nur eine begrenzte Aussage über die Relevanz des Dokumentes geben, da eine tiefgehende Indexierung wie bei umfangreichen Fachaufsätzen bei Presseartikeln zu aufwendig ist. Dieses Manko führt dazu, daß die in der konventionellen Selbstrecherche geübten Nutzer mit dem Rechercheergebnis nicht immer zufrieden sind und sich deshalb gern alle nachgewiesenen Dokumente zu einem Themenbereich über einige Monate zurück einfach aus-

drucken lassen, um dann nach gewohnter Manier vom Stapel nur einen Bruchteil zu verwenden.

Ein weiterer Nachteil ist, daß manche bisherige Selbstrechercheure nicht an das neue System gehen. Das mag an ihrer Abneigung gegenüber technischen Neuerungen liegen oder auch, daß sie im Rechercheraum untätig rumsitzen, während das System eine Anfrage verarbeitet. (Letzteres wird dann nicht mehr bestehen, wenn möglichst viele regelmäßige Nutzer direkt von ihrem Platz aus recherchieren können und während der Wartezeiten sich anderen Arbeiten widmen können und auch telefonisch erreichbar sind). Erwünscht ist im Hause aber gerade die Unabhängigkeit der Informationssuchenden von der Unterstützung des IuD-Personals.

Welche Weiterentwicklungen sind geplant?

Schritt für Schritt sollen in der Zentrale in Frankfurt regelmäßige Nutzer mit Recherchetationen ausgestattet werden und per Netz an den zentralen Speicher angeschlossen werden. Die Mitarbeiter der IuD-Abteilung werden das System betreiben und Recherchen hauptsächlich für Bankkunden durchführen.

Auf der Inputseite ist eine weitere Automatisierung von Arbeitsvorgängen wünschenswert. So könnte eine elektronische Schere und ein programmgesteuerter Umbruch das aufwendige manuelle Bearbeiten der Artikel ersparen. (Das DIN-A4-Format bleibt vorläufig wegen der Fax-Kompatibilität notwendig).

Auch bei der Indexierung und damit bei der Recherche läßt sich noch einiges verbessern. Mit Hilfe einer OCR-Lösung können die gescannten Texte oder zumindest Textbereiche im ASCII-Format abgelegt werden und wären mittels eines Volltext-Retrieval-Systems mit integriertem Synonym- und orthographischem Abgleich zugänglich. Dies kann nur als Ergänzung der intellektuellen Indexierung verstanden sein, würde deren Wert aber erheblich verbessern.

Alles in allem lassen die in der Dresdner Bank gemachten Erfahrungen den Schluß zu, daß der Einsatz eines optischen Speichersystems in der Wirtschaftspresdokumentation auch anderer Häuser von Nutzen sein kann.

Martin Michelson  
Dresdner Bank  
Jürgen-Ponto-Platz 1  
W-6000 Frankfurt am Main

# Der Einsatz neuer Informationstechniken am Richterarbeitsplatz auf der Ebene der Amtsgerichte

Friedrich Scheuermann

## Inhalt

1. Einleitung
2. Die Bedeutung des Einsatzes
3. Gegenwärtiger Zustand
4. Anforderungen für den Einsatzes
- 4.1. Allgemeine Anforderungen
- 4.2. Anforderungen an die Ausstattung
5. Einsatzgebiete
- 5.1. Individuelle Anwendungen
- 5.2. Informationsbeschaffung und Kommunikation
- 5.2.1. Gerichtseigene Datenbanken
- 5.2.2. Externe Datenbanken
- 5.2.3. CD-ROM-Angebot
- 5.2.4. Bildschirmtext
6. Schluß

## Referat

"Neue" Informationstechniken sind an Amtsgerichten noch wenig verbreitet. Trotz Überlast und zu langer Verfahrenslaufzeiten wird noch immer davon abgesehen elektronische Werkzeuge am Arbeitsplatz einzuführen. Nur experimentell werden Erfahrungen mit dem Einsatz von PC zur Textverarbeitung, Nutzung von Datenbanken und Mitteilungsversand gesammelt. Im vorliegenden Beitrag wird ein Spektrum einsetzbarer Techniken am Richterarbeitsplatz vorgestellt.

## Abstract

New forms of processing and transmitting information have not yet been fully accepted by German judges. Advanced electronic equipment have yet to be completely incorporated for such work, despite ever increasing workloads which result in costly overruns. At the moment PC's for word processing purposes databases and electronic mail are being provided to and used by the judges on an experimental bases only. The following paper will provide an overview of how advanced office automation is presently being used and will suggest new applications to further benefit the judge.



### 1. Einführung

Bei der Untersuchung des Einsatzes neuer Informationstechniken ist wegen der Unterschiedlichkeit der Aufgabenbereiche eine Eingrenzung erforderlich. "Den" Richterarbeitsplatz gibt es nicht, ebenso wie es nicht "die" Anforderungen an die Informationstechnologien gibt. Einerseits unterscheiden sich die Gerichte und die Tätigkeit der Richter durch verschiedene Zuständigkeiten, Organisationsstrukturen und somit unterschiedlichem Bedarf, andererseits werden auch innerhalb der Gerichte nicht die gleichen Aufgaben wahrgenommen. Zumindest in vertikaler Hinsicht, von Gerichten der ersten Instanz bis zum Bundesverfassungsgericht erscheinen getrennte Analysen notwendig.

### 2. Die Bedeutung der Einsatzes von Informationstechniken

Im Themenkomplex "Richterarbeitsplatz" wird schon von "neuen Techniken" gesprochen, wenn damit der Einsatz von PC's unterschiedlichster Leistungsfähigkeit und Nutzungsmöglichkeiten gemeint ist. Die Notwendigkeit des Einsatzes ergibt sich aus einer allgemein nicht zufriedenstellenden Situation an den Amtsgerichten. So sind viel zu lange Verfahrenslaufzeiten und eine nur eingeschränkte Servicebereitschaft in Amtsgerichten festzustellen. Aus interner Sicht der Gerichte werden als Hauptprobleme die Ineffizienz der Abläufe bei extremer Arbeitsteilung und hohem Transportaufwand, unzureichend vorhandene sachliche Hilfsmittel, die zudem, falls vorhanden, unzureichend genutzt werden sowie die Demotivation vieler Mitarbeiter verantwortlich gemacht [Kienbaum,1990,MS-1]. Durch Umstrukturierung und Modernisierung auf der gesamten Amtsgerichtsebene erwartet die Kienbaum Unternehmensberatung GmbH (in ihrem Abschlußbericht zur Strukturanalyse der Rechtspflege (Organisation der Amtsgerichte) vom Dezember 1990) eine erhebliche Verbesserung der Situation. Die neuen Informations- und Kommunikationstechniken können dabei einen beträchtlichen Beitrag leisten: Richter arbeiten zum Beispiel bereits mit einer umfangreichen Sammlung von Formularen mit verschiedenen Textbausteinen, die auf einen Großteil der Verfahren anwendbar sind, eine Tätigkeit, für die sich der Einsatz von EDV anbietet. Neben der Unterstützung von Büroarbeiten, kann dies zu einer Vereinfachung fachjuristischer Routinetätigkeit führen. Bestehende Überlast kann abgebaut und die Qualität richterlichen Arbeitens verbessert werden [Berkemann 1989, 34].

### 3. Gegenwärtiger Zustand

Richter waren bislang bei der Unterstützung durch Datenverarbeitung ausgeklammert und bis heute konnten EDV-Lösungen am Arbeitsplatz der Rechtsprechenden ein experimentelles Stadium nicht überwinden [Hoffmann,1990,8]. Man konzentrierte sich zu Beginn allein auf den Bereich der Geschäftsstellenverwaltung. Dieser isolierte Einsatz von Informationstechniken greift zu kurz, da der Richter innerhalb eines Systems mit Geschäftsstelle, Rechtspfleger, Kanzlei, Kostenbeamte etc. zusammenarbeitet.

Der Untersuchung der Kienbaum Unternehmensberatung GmbH zufolge sind gegenwärtig sind fast alle Arbeitsplätze der Richter mit Diktiergerät und Handliteratur ausgestattet. Die meisten Richter verfügen über ein von ihnen alleine genutztes Büro mit Telefon, wobei nur ungefähr zwei Drittel der Richter auf einen eigenen Amtsanschluß zurückgreifen können. Selten dagegen sind sie mit PC's oder Terminal ausgestattet, häufiger noch mit Rechen- und Schreibmaschinen.

Die Richter kritisieren den unvollständigen und unzureichenden Buchbestand, vor allem aber die Auswahl an aktuellen Fachzeitschriften. Immer lauter wird die Forderung nach dem Einsatz elektronischer Datenverarbeitungsanlagen zur Unterstützung des Geschäftsablaufes. Van Raden dokumentiert die Meinungen der Richter ausführlich in der Auswertung seiner Akzeptanzuntersuchungen [Raden, 1989].

Die individuellen Erfahrungen bei erstmaliger Nutzung schwanken dabei oftmals zwischen Euphorie und Resignation [RÜHLE,1989,23]. Hohe Akzeptanz korreliert stark mit privatem Engagement. Dies zeigt sich zum Beispiel an der Vielfalt eigenerstellter Programme. Für die Ablehnung des Computereinsatzes kann die als unzureichend bemängelte Schulung und Betreuung verantwortlich gemacht werden. Auf der anderen Seite werden Befürchtungen offenbar, Kontroll- und Steuerungsmöglichkeiten des Arbeitsprozesses zu verlieren. Häufig beobachtet wurden Schwierigkeiten, die persönlichen Bedürfnisse exakt zu definieren.

Im Zeitalter digitaler Kommunikationsverfahren klagen viele Richter über unzulängliche Verfügbarkeit der herkömmlichen Fernmeldetechnik. Man beanstandet fehlende Erreichbarkeit nach Dienstschluß der Zentrale, häufige Überlastung der Zentrale und keine Wahlwiederholung" [Kienbaum,1990,89].

#### 4. Anforderungen für den Einsatz

##### 4.1. Allgemeine Anforderungen

Die richterliche Tätigkeit stellt besondere Grundanforderungen an den Einsatz neuer Informationstechniken.

a) Richterliche Unabhängigkeit: Die richterliche Unabhängigkeit muß in jedem Fall gewährleistet werden: einerseits soll die Informationstechnik die volle Beherrschung des Arbeitsprozesses und der Falldaten (Kontrolle und Steuerung ermöglichen, andererseits muß dem Richter die Auswahl geeigneter Produkte und Informationsdienstleistungen selbst überlassen werden. Nicht selten errechnen verschiedene Programme durch Anwendung differenter Algorithmen und Modellierungen voneinander abweichende Ergebnisse.

b) Datenschutz: Der Zugang auf das System und der Zugriff auf bestimmte Daten müssen durch Identifikationskontrollen und Passwortschutz gewährleistet sein.

##### 4.2. Anforderungen an die Ausstattung

Für die Ausstattung der Richterarbeitsplätze wird ein leistungsstarker PC aufgrund seiner Flexibilität und Offenheit favorisiert. Mit Vernetzungsfähigkeit innerhalb eines lokalen Netzwerkes (LAN) wird dezentrale Datenverarbeitung mit dem Funktionsspektrum eines Multi-User-Systems gekoppelt [Endrös,1987,3]. So kann

- den individuellen Bedürfnissen an Optionen (in Form sogenannter Toolboxes) und Software Rechnung getragen werden,
- die Aktualisierung der Software flexibler durchgeführt werden,
- der Datenaustausch mit den bereits vorhandenen, jedoch auf andersartiger Betriebssystembasis operierenden Geschäftsstellenrechnern, stattfinden kann und
- die Tätigkeit am heimischen PC-Arbeitsplatz vereinfacht wird.

Den besonderen richterlichen Anforderungen soll durch eine einheitliche, an Tätigkeit und Anspruch orientierten Benutzeroberfläche nachgegangen werden, wie sie zum Beispiel von der Firma SIEMENS AG mit "Comfodesk" grafikorientiert und mit Mausunterstützung (Ikonen, Pull-Down-Menüsteuerung) angeboten wird.

Technisch realisierbar wäre die Verwirklichung eines papierlosen Szenarios der Kienbaum Unternehmensberatung GmbH:

- "- Die Schriftsätze der Parteien/Anwälte werden eingescannt oder gleich als Electronic Mail über Modem zugesandt.
- Der Richter liest am Bildschirm den Text und markiert die entsprechenden Begriffe, z.B. "Schadensersatz" mit Maus und Lichtgriffel.
- Der Computer ordnet den Text entsprechend den Vorgaben des Richters, sucht die entsprechenden Passagen heraus und stellt die gegensätzlichen Parteivorträge geordnet gegenüber.

Eine solche Aufbereitung der Texte ist auch in automatisierter Form möglich: Eine intelligente Software nimmt den Zuordnungsprozeß selbst anhand bestimmter, vorher "gelernter" Schlüsselbegriffe vor." [Kienbaum, 1990, 149].

Die grundsätzlichen Probleme des papierlosen Büros werden in diesem Zusammenhang durch juristische Bedenken (z.B. gegenüber der Vereinfachung von Dokumentmanipulation) weiter verschärft

## 5. Einsatzgebiete der Informationstechniken

Verschiedene Verwendungsmöglichkeiten der Techniken finden sich in den Bereichen der aufgabenbezogenen Gestaltung des Arbeitsprozesses, der Informationsbeschaffung aus elektronischen Datensammlungen und einer ökonomischeren Organisation in der Sachbearbeitung und Entscheidungsfindung.

### 5.1. Individuelle Anwendungen

Jedes Dezernat benötigt dabei eigene Lösungen zu Detailproblemen, die zum Beispiel Kostenentscheidungen, Blutalkoholbestimmung etc. betreffen. Dazu existiert bereits eine Fülle von Programmen, meist von Richtern für ihren jeweiligen Aufgabenbereiche selbst entwickelt. Die daraus resultierende unübersichtliche Vielfalt nicht-standartisierter Programme wirft aber Probleme auf, die bundeseinheitliche Verfügbarkeit normierter Software erforderlich werden läßt.

### 5.2. Informationsbeschaffung und Kommunikation

Auch auf Amtsgerichtsebene wird der Rechnereinsatz bei der Beschaffung und Organisation von Information, sowie der individuelle Kommunikation unterstützend wirken können. Im Einzelnen betrifft dies

- die interne Datenbank des Gerichts,
- externe Datenbanken,
- Publikationen auf Massenspeichermedien wie CD-ROM,
- Mailboxen oder electronic mail-Systeme zum Nachrichten- und Informationsaustausch.

Die daraus gewonnenen Informationen könnten in nach persönlichen Präferenzen ausgerichteten Datenbanken gespeichert werden. Kriterien ergonomischer Arbeitsplatzgestaltung sollte dabei Rechnung getragen werden, statt dem Benutzer unflexiblen Eingaben abzunötigen.

#### 5.2.1. Gerichtseigene Datenbanken

Durch den Einsatz gerichtseigener Datenbanken kann das Know-How einzelner Richter dem gesamten Gericht zugänglich gemacht werden. Die Sachbearbeitung bestimmter Probleme durch den Vergleich mit bereits abgeschlossenen, vergleichbaren Fällen kann dadurch erheblich erleichtert werden.

#### 5.2.2. Externe Datenbanken

Für den richterlichen Gebrauch steht die Online-Datenbank juris zur Verfügung. Die meisten Bundesländer haben Pauschal-Nutzungsverträge mit der juris GmbH abgeschlossen. Diese schließen bestimmte Nutzungskontingente mit festgesetzter maximaler Stundenzahl ein. Angeboten werden Datenbanken und Spezialbibliotheken zu allen Rechtsgebieten. Mangelnde bzw. schlechte Erfahrungen führten allerdings bei den wenigen bislang nutzungsberechtigten Gerichten zu Akzeptanzproblemen von Seiten der Richter.

#### 5.2.3. Entscheidungssammlungen/Kommentare auf CD-ROM

Ein expandierender Markt entsteht für CD-ROM-Publikationen. Vorreiter auf diesem Markt war die juris GmbH, die als preisgünstige Online-Alternative (teilweise auch im CD-ROM-Antiquariat) eine Reihe von Entwicklungen anbietet (siehe Kasten).

Momentan sind folgende CD-ROM-Produkte verfügbar:

- juris GmbH:
  - Rechtsprechung der Finanzgerichtsbarkeit
  - Rechtsprechung des Bundesgerichtshofes
  - Rechtsprechung der Sozialgerichtsbarkeit
  - Rechtsprechung der Arbeitsgerichtsbarkeit
  - Rechtsprechung der Verwaltungsgerichtsbarkeit
  - Bundesrecht
- Verlag C.H.Beck
  - Vollständige Sammlung der arbeitsgerichtlichen Praxis (AP)
  - Volltexte NJW (Neue Juristische Wochenzeitschrift) 1981 - 1985
  - Volltexte NJW 86-90, Kurztex te NJW 81-85, Volltexte NJW-RR 86-90
- Carl Heymann Verlag:
  - Entscheidungssammlungen des Bundesgerichtshofes BGHE (BGH-DAT und BGHR)
- MediConsult GmbH:
  - Einigungsvertrag
  - Einigungsvertrag mit Erläuterung der Bundesregierung
  - Einigungsvertrag und weitergeltendes Recht der früheren DDR

- Luchterhand Verlag:
  - EZA Entscheidungssammlung zum Arbeitsrecht
- Werner-Verlag:
  - Baurecht auf CD-ROM.

Eine Reihe weiterer Veröffentlichungen ist in Vorbereitung (z.B. Mietrecht, Familienrecht, Wettbewerbsrecht), ebenso wie klassische Standardkommentare in Zukunft auf CD-ROM angeboten werden sollen. Dadurch ist ein beträchtliches Informationsvolumen am eigenen Arbeitsplatz zugänglich. Für Richter bietet die Nutzung von CD-ROM ferner den Vorteil, zeitungsbunden und in Ruhe recherchieren zu können, ohne die kostenabhängige Nutzungszeit berücksichtigen zu müssen.

### 5.2.3. Bildschirmtext

Wenig Erwähnung in der einschlägigen Literatur findet Bildschirmtext in Bezug auf die Informationsgewinnung. Dieser Dienst der Deutschen Bundespost Telekom bietet einerseits den Vorteil (gegenüber Briefpost und Telefon) eines elektronischen Kommunikationsmittels, jederzeit und schnell Mitteilungen an andere Teilnehmer zu übermitteln. Einige Anbieter stellen darüberhinaus juristische Information und Datenbanken zur Verfügung. Im Btx-Angebot des Boorberg-Verlags sind beispielsweise wöchentlich bzw. zweiwöchentlich die neusten Leitsätze des BGH und des BAG abrufbar. Nahezu problemlos läßt sich auch ein automatisiertes Abspeichern in die eigene Datenbank realisieren.

## 6. Schlußbemerkung

Trotz der Vielfalt möglicher Einsatzformen wird eine flexible Integration neuer Techniken aus dem experimentellen Stadium heraus in den juristischen Alltag schwerfallen. Haushaltsbedingte Schranken und organisatorisch Handlungsschwierigkeiten der Gerichte verhindern, daß anwendbares Know-How auf absehbare Zeit in die Praxis umgesetzt wird. Dem qualitativen Nutzen der Informationstechnik in der Urteilsfindung steht quantitativ das klassische Problem gegenüber, mehr Zeit für Suche, Selektion und Verarbeitung der anfallenden Daten investieren zu müssen. Der Einsatz neuer Informationstechniken ist aber im Gesamtzusammenhang zu betrachten. Die Effektivität der Zusammenarbeit zwischen dem Richter und seinen Mitarbeitern könnte gestärkt und Verfahren beschleunigt werden unter der Voraussetzung einer organisatorischen Neuordnung.

## 7. Literatur

BERKEMANN, JÖRG: ANFORDERUNG AN INFORMATIUNSTECHNIK AM ARBEITSPLATZ VON RICHTERN. In: van Raden/Weihermüller (Hrsg.), Informatiunstechnik am Arbeitsplatz von Juristen. Berlin / Bonn / München, 1989

ENDRÖS, ALFRED: Rechtsprechung und Computer in den neunziger Jahren - Am Beispiel von Begriff und Typologie der Körperschaft des öffentlichen Rechts. Berlin/ Heidelberg/ New York/ London /Paris/ Tokio 1988

FIEDLER, HERBERT: Die Bedeutung der Informationstechnik für die Rechtspflege - Chancen und Gefahren - in: van Raden/Weihermüller: Informationstechnik am Arbeitsplatz von Juristen, Köln / Berlin / Bonn / München, 1989

GESELLSCHAFT FÜR MATHEMATIK UND DATENVERARBEITUNG MBH/INSTITUT FÜR DATENVERARBEITUNG IM RECHTSWESEN: Bericht im Rahmen der Hauptuntersuchung über die Büroautomation im Teilbereich der Geschäftsstellenverwaltung in der Justiz. Bonn, 1980

HOFFMANN, HELMUT: Informationstechnik am Arbeitsplatz des Richters. Unter besonderer Berücksichtigung des Zivilrichters. Dissertation, Saarbrücken, 1990

KIENBAUM UNTERNEHMENSBERATUNG GmbH: Abschlußbericht: Strukturanalyse der Rechtspflege: Organisation der Amtsgerichte. Düsseldorf, 1990

RADEN, LUTZ VAN: Rechner, Richter, Realitäten - Computer in der Justiz. Köln, Berlin, Bonn, München, 1989.

RÜHLE, KLAUS: Einsatz von Informationstechnik am LG Hamburg. In: van Raden/Weihermüller: Informationstechnik am Arbeitsplatz von Juristen, Berlin, Bonn, München, 1989

SCHNEIDER, JOCHEN: Information und Entscheidung des Richters - Zu einer juristischen Entscheidungs- und Kommunikationstheorie automatischer Informationssysteme im Recht. Ebelsbach, 1980

### Anschrift:

Stud.cand.phil. Friedrich Scheuermann  
Universität des Saarlandes  
FR 5.5 Informationswissenschaft  
Im Stadtwald  
W-6600 Saarbrücken 11  
E-Mail: scheuermann@rz.uni-sb.de

**INFORMATIONSSYSTEME ÜBER AUSBILDUNGSSOFTWARE:  
STAND UND ENTWICKLUNGSPERSPEKTIVEN IN BULGARIEN**

R. Tsankova, R. Koleva

Das automatisierte Informationssystem über Ausbildungssoftware kann vielseitig genutzt werden. Es sammelt Informationen über Ausbildungssoftware, speichert und verarbeitet sie und stellt sie bei Bedarf zur Verfügung.

Dadurch wird folgendes erreicht:

- Unterstützung von Lehrern und Lektoren bei der Auswahl von Ausbildungssoftware für die Lösung spezifisch didaktischer Aufgaben;
- Verbreitung moderner pädagogischer Methoden zur Verbesserung der Methodologie und der Qualität der Ausbildung;
- Bereitstellung von Informationen über Angebote und Bezugsbedingungen von bzw. für Ausbildungsprogrammprodukte;
- Schaffung von Kriterien für Entscheidungen über Ausbildungsstrategien sowie für ihre Analyse und Prognose;
- Verbesserung der Qualifikation von Lehrern und Lektoren bei der Anwendung von Computern im Ausbildungsprozeß.

Damit diese Ziele erreicht werden, mußten mit dem Informationssystem folgende Aufgaben gelöst werden:

- Bereitstellung bestimmter statistischer Informationen sowie von Möglichkeiten zur Beantwortung von Nutzeranfragen;
- Veröffentlichung eines Katalogs über die Ausbildungsprogrammprodukte in Bulgarien;
- Veröffentlichung eines Bulletins, das methodische Materialien und Ergebnisse der Anwendung von Ausbildungssoftware enthält.

Die Lösung dieser Aufgaben erforderte ein gemischtes System, das sowohl informationssuchende als auch "informationsleitende" Funktionen hat.

Die Nutzer benötigen Informationen über Inhalt und Qualität des Unterrichtsmaterials, über Unterrichtsverfahren und über die Methodologie ihrer Anwendung. Diese Art von Informationsbedürfnissen kann leicht gedeckt werden durch die Verwendung geeigneter Schlüsselwörter. Sie bestimmt den informationssuchenden Teil des Systems. Das System ähnelt somit den Systemen für wissenschaftliche und pädagogische Information der UNESCO. Dadurch wird auch ein Informationsaustausch zwischen beiden möglich.

Die Gewährleistung der "informationsleitenden" Funktionen des Systems ist eine Frage des genauen Beschreibens, Gruppierens, Klassifizierens und Identifizierens der einzelnen Programmprodukte. Dazu werden speziell für das System geschaffene Klassifikatoren benutzt.

Die Benutzer des Systems können anhand ihrer Informationsbedürfnisse folgendermaßen gruppiert werden:

- Autoren von Programmprodukten für Ausbildungszwecke.

Die ihnen vom System gegebenen Informationen sind eine gute Vergleichsbasis zwischen ihren Programmprodukten und den schon existierenden in diesem Bereich. Die gesammelten Informationen über die Programmprodukte für Ausbildungszwecke ermöglichen die Bestimmung von Entwicklungstendenzen bei ähnlichen Programmprodukten;

- Lehrer und Lektoren. Das System schafft gute Bedingungen für ein näheres Kennenlernen fertiger Ausbildungssoftware. Das ermöglicht die Auswahl von passenden Programmen für konkrete didaktische Aufgaben sowie für das Erlernen der notwendigen methodischen Schritte;

- Studenten und Personen, die sich weiterbilden. Die im System gespeicherten Informationen enthalten auch Angaben über die Art der Verbreitung der Ausbildungssoftware.



# WIDAB - Konzeption und Pilotrealisierung eines experimentellen Information Retrieval-Systems

Kurt Englmeier

## Kurzfassung

Die Konzeption des Informationssystems WIDAB orientiert sich an dem Ziel, aus einer wahrhaftig unübersehbaren Menge an Wirtschaftsdaten, die bei den statistischen Ämtern und Wirtschaftsforschungsinstituten gehalten werden, Informationen werden zu lassen, die so einfach zugänglich sind wie der Inhalt einer Zeitung.

Dies ist vor dem Hintergrund zu sehen, daß einerseits für viele Unternehmen die Bedeutung von Wirtschaftsinformationen aus Datenbanken stetig wächst, die derzeitigen Informationsdienste die Ansprüche der Nutzer aber nur insoweit erfüllen können, als sie ihr Informationsangebot erweitern und damit die Informationsversorgung in quantitativer Hinsicht verbessern. Daraus resultiert in zunehmendem Maße das Problem der Informationsüberlastung, dem durch eine qualitative Verbesserung des Informationsangebots begegnet werden muß. Dieser Aspekt zielt auf die Gestaltung des Retrieval-Systems.

Im Information Retrieval (IR) - System WIDAB soll anstelle einer Retrieval-Sprache ein Menüsystem die Nutzer in die Lage versetzen, ihre Suchanfrage präzise zu formulieren. Zu diesem Zweck wird Sachwissen in Form von Begriffsrelationen in den Retrievalprozeß integriert. Diese zusätzlichen Daten werden zentral von Informationsmittlern für alle Benutzer verwaltet und diesen zur Formulierung ihrer Suchanfragen zur Verfügung gestellt. Die Nutzer recherchieren nicht direkt in der Datenbank, von der sie vermuten, ihr Inhalt könnte zur Reduzierung ihres Informationsproblems beitragen, sondern sie bedienen sich eines automatisierten Informationsmittlers, der die Recherche für sie durchführt.

## Abstract

The basic concept of the WIDAB information system is centered on the goal of presenting information from the vast quantities of economic data compiled by the statistical offices and research institutes in a form that is as easily accessible as the daily newspaper.

This is to be seen against the background of the fact that despite an increasing importance of economic information from databases for many companies, the existing information services are only able to satisfy the requirements of the users by enlarging their information offer, thereby improving the information supply as far as its quantity is concerned. This increasingly results in the problem of information overload which has to be countered by improving the quality of the information offered. This aspect is of prime importance for the design of the retrieval system.

Instead of a language, in the information retrieval (IR) system WIDAB the users are to be enabled by a menu-driven system to precisely formulate their queries. For this purpose expert knowledge in the form of concept relations is intergrated into the retrieval process. The additional data are for all users centrally managed by information mediators who make them available to users for formulating their queries. Instead of, therefore, directly searching the contents of a data base which might contribute to solving their information problem, users let an automated information mediator do the searching for them.

Seit der jüngsten Vergangenheit kann beobachtet werden, daß der Markt für Informationsdienste auf dem Sachgebiet Wirtschaftsinformationen zunehmend an Attraktivität gewinnt. Das Angebot erstreckt sich von Detailinformationen für die Konkurrenzbeobachtung bis zu den Daten für Branchen- oder gesamtwirtschaftliche Betrachtungen. Dabei werden neben Text- immer mehr numerische Dokumente in das Angebot integriert.

Die Nachfrage nach derartigen Informationen wird von den sich stärker verändernden Marktsituationen belebt. Informationen über neue und veränderte Märkte spielen in den Planungsprozessen der Unternehmen eine bedeutende Rolle: Die Informationen, die in den Entscheidungsprozeß eingehen, werden mehr denn je zu entscheidenden Parametern der Konkurrenzfähigkeit.

Die in der Zwischenzeit stark gestiegene Menge an Informationen wird in der Fachwelt zwar als ausreichend erachtet, um eine zufriedenstellende Informationsversorgung in vielen Planungssituationen sicherzustellen. /1/ Dieser Quantitätsaspekt ändert aber offensichtlich nichts an der geringen Attraktivität der Informationsdienste. So wird bedauert, daß sie nicht so verbreitet sind wie es im Interesse der Wettbewerbsfähigkeit der Unternehmen notwendig wäre. Oftmals werden als Gründe mangelndes Wissen um die Informationsdienste angeführt und - vor allem bezogen auf kleine und mittlere Unternehmen - mangelnde Bereitschaft und Fähigkeit der Entscheidungsträger, ihr Informationsverhalten zu ändern, d.h. sich auf neue Informationsmedien einzustellen. /2/

Es zeigt sich aber, daß neben der Preisgestaltung, bei der viele Informationsdienste nicht gerade zimperlich sind, die Unterstützung bei der Informationsrecherche und die Aufbereitung der Recherche-Ergebnisse wichtige und in manchen Fällen sogar entscheidende Faktoren für viele potentielle Nutzer sind. Der Umgang mit dem Computer ist den meisten vertraut und die Bereitschaft, neue Informationsmedien zu nutzen, vorhanden. Die Ausweitung des Informationsangebots ist zwar eine Voraussetzung für eine verbesserte Informationsversorgung, gleichzeitig forciert sie aber auch die Informationsüberlastung.

D.h. der Aufwand, den Benutzer betreiben müssen, um aus dem gesamten Informationsangebot die für sie relevanten Informationen zu selektieren, steht oftmals in keinem Verhältnis zum erreichten Nutzen. /3/

Dieser Aspekt zielt direkt auf die Gestaltung des Retrievalprozesses. Die heute angebotenen Retrieval-Systeme stützen sich größtenteils auf Retrieval-Verfahren aus den sechziger Jahren. Boolesche Frageformulierungen auf der Basis einer ungewichteten Indexierung mit einem vorgegebenen Deskriptorensystem und/oder als Freitextsuche in Textdokumenten, dazu wenig Unterstützung des Rechercheprozesses durch das System machen den Informationssuchenden das Leben alles andere als leicht /4/.

Mit WIDAB (Wirtschaftsinformationen- und Datenbank) soll die Schaffung eines Systems zur Informationsorganisation erreicht werden, das eine effiziente Nutzung der wahrhaftig unübersehbaren Mengen an Wirtschaftsinformationen erlaubt, die bei den statistischen Ämtern und Wirtschaftsforschungsinstituten gehalten werden. Die Informationen sollen dabei so aktuell sein wie eine Tageszeitung, gezielt und problemorientiert der augenblicklichen Entscheidungssituation der Nutzer entsprechen und dabei noch so einfach zugänglich sein wie der Inhalt einer Zeitung. Da WIDAB aufgrund dieses Anspruchs grundsätzliche Bedeutung für die Wirtschaft gewinnt, wird das Projekt vom Bundesministerium für Wirtschaft gefördert.

Aus dem genannten Anspruch können die Ziele abgeleitet werden, auf deren Grundlage das Information Retrieval (IR)-System WIDAB entworfen wird:

Der Wunsch nach Aktualität verlangt, daß neue Informationen unmittelbar nach ihrer Entstehung in die Informationssammlung aufgenommen werden. Um ein Höchstmaß an Vollständigkeit zu erreichen, muß das IR-System den Zugang zu einem großen Anteil potentiell relevanter Informationen gewährleisten /5/.

Weiter ist es Aufgabe des Systems, durch ein hohes Maß an Problemorientiertheit eine Reduktion der Informationsflut zu erreichen. D.h. es soll einerseits den Nutzern die Möglichkeit eröffnen, in der Suchanfrage ihr Informationsproblem möglichst präzise zu umreißen, und andererseits in der Lage sein, aufgrund der Suchanfrage aus dem Gesamtangebot der Informationsautoren die Nachrichten zu selektieren, die sich genau auf den Informationsbedarf beziehen. Der von den Nutzern artikulierte Informationsbedarf resultiert dabei aus dem Wissensdefizit in der jeweiligen Entscheidungssituation. Durch den Informationsgehalt der übermittelten Dokumente wird ihr Wissensdefizit reduziert /6/.

Den Nutzern muß die Möglichkeit geboten werden, die erhaltenen Informationen weiterzuverarbeiten, um sie möglichst optimal in die jeweilige Entscheidungssituation integrieren zu können (Integrationsfähigkeit). Weitere Informationen können aus den übermittelten Dokumenten gewonnen werden, sobald sie mit den entsprechenden Werkzeugen bearbeitet werden.

Damit das System bei einem großen Kreis von Nutzern Anklang findet, soll die Zugangsschwelle so niedrig wie möglich gehalten, d.h. ein leichter und kostengünstiger Zugang gewährt werden. Diese Bedingung beeinflusst in erster Linie die Gestaltung der Benutzeroberfläche und der Preise.

## **1. Das Informationsmittlerkonzept im Überblick**

Um ein Höchstmaß an Aktualität und Vollständigkeit zu erreichen, sieht das Konzept von WIDAB vor, alle Funktionen, die mit der Sammlung und Pflege aber auch der Indexie-

zung der Dokumente (numerische Daten und Texte) in Zusammenhang stehen, am Ort der Informationsentstehung - bei den Informationsautoren - zu belassen.

Um den Nutzern einen umfassenden, aber einheitlichen Zugang zu den dezentralen und heterogenen Datenbanken der Informationsautoren bieten zu können, sieht WIDAB eine Vermittlerstelle vor, die eine für die Nutzer virtuell zentrale Datenbasis realisieren soll. Diese recherchieren folglich nicht direkt in der jeweiligen Datenbank, von der sie vermuten, ihre Informationssammlung könne ihren Bedarf decken. Sie bedienen sich eines Informationsmittlers, dem sie ihr Informationsproblem anvertrauen und der aufgrund seines Fachwissens in der Lage ist, daraus eine geeignete Recherche in den relevanten Datenbanken abzuleiten. Der Informationsmittler weiß, in welchem Themengebiet das Informationsproblem anzusiedeln ist und er weiß auch, welche Informationsautoren welche Informationen zu diesen Themen auf ihren Datenbanken vorhalten. Darüberhinaus kennt er auch die Arbeitsweise für den Abruf der nötigen Informationen. WIDAB will versuchen, den Teil der Informationsmittlerfunktionen in das IR-System zu integrieren, der vollständig automatisierbar ist. Beim Design des Systems wird deshalb besonderes Augenmerk auf drei Bereiche gerichtet:

- Die Umsetzung des Informationsproblems in eine genaue Suchanfrage
- die Organisation eines geeigneten Rechnernetzes zur Generierung der virtuellen Gesamtdatenbank und
- die präzise Verknüpfung der Suchanfrage mit den relevanten Informationen.

## **2. Mensch-Computer-Interaktion**

### **2.1. Das Sach- und Verständigungsproblem**

Beim Rechercheprozeß lassen sich die Nutzer von mehr oder weniger wagen Partialmodellen über das Wirkungsgeflecht ökonomischer Größen leiten. Sie treffen Entscheidungen über die Ausprägungen bestimmter Modellgrößen (als Entscheidungsparameter) unter Berücksichtigung der durch die Entscheidungen ausgelösten Rückwirkungen auf die Elemente der Partialmodelle.

Das Sachproblem orientiert sich an dieser Entscheidungssituation und zielt darauf, bei den Nutzern das Wissen über die Modellstruktur und die Ausprägungen der Modellparameter zu verbessern. In diesem Verbesserungspotential spiegelt sich das Wissensdefizit wieder. Die Benutzer wollen folglich durch das Retrieval-System Informationen finden, die zur Reduzierung ihres Wissensdefizits beitragen. Das Sachproblem liegt in der Artikulation des Informationsdefizits, d.h. in der Formulierung der Suchanfrage.

Jeder ökonomischen Größe ordnen die Nutzer charakterisierende Begriffe zu. Ihr Informationsdefizit kann folglich anhand der Begriffe kurz umrissen werden, die mit den Modellgrößen verbunden sind, über die sie Informationen benötigen. Das Sachproblem kann in dem hier beschriebenen Anwendungsgebiet als das Problem verstanden werden, Begriffscluster zu bilden, die das Informationsdefizit beschreiben.

Bei der Recherche mit dem IR-System aber werden die Benutzer nicht nur mit dem Sachproblem sondern auch mit dem Verständigungsproblem konfrontiert, das der Funktionsumfang des Systems aufwirft. Krause /4/ verweist darauf, daß Benutzer einen nicht unerheblichen Teil ihrer Energien darauf verwenden, zu lernen und darüber nachzudenken, wie sie ihr Sachproblem so formulieren, daß sie die Funktionalität des Systems nutzen können. Deshalb ist die Mensch-Computer-Schnittstelle so zu gestalten, daß sie zur Verbesserung dieses Verständigungsproblems zwischen Mensch und Computer beiträgt. Damit die Benutzer die für sie relevanten Informationen aus der Datenbank abfragen können, müssen sie sich mit Verständigungsproblem und dem Sachproblem auseinandersetzen. Bullinger /7/ weist darauf hin, daß sich in derartigen Situationen Wissen auf unterschiedliche Aspekte und Ebenen der Mensch-Computer-Interaktion bezieht. Zum einen kann es Wissen über die Interaktionsmöglichkeiten selbst sein. Zum zweiten kann Wissen über das Anwendungsgebiet verwendet werden, um Erklärungen zu liefern. Der Bereich des aufgabenbezogenen Wissens bietet ein erhebliches Potential für eine verbesserte Unterstützung der Benutzer.

Das eigentliche Retrievalproblem zielt damit bei WIDAB nicht nur auf den Zugang zur eigentlichen Zielinformation, sondern auch auf die Integration von zusätzlichem Sachwissen. Zentrales Moment bei der Gestaltung der Benutzeroberfläche, der Interaktionskomponente, ist die Integration des Sachwissens der Informationsmittler. WIDAB bietet den Nutzern eine Übersicht der Themen, die auf deren Fachkompetenz gründen und einen Ausschnitt seines Sachwissens abbilden. Diese Übersicht enthält allerdings nicht eine schlichte Auflistung aller Themen sondern vielmehr eine Begriffshierarchie, die auch den nötigen Kontext verdeutlichen soll. D.h. die Nutzer können so erkennen, aus welchen übergeordneten Punkten ein betrachteter Begriff abgeleitet wurde und welche Punkte darunter zu subsumieren sind. Diese Relationen vermitteln theoretische Zusammenhänge der in der virtuellen Gesamtdatenbank gespeicherten Sachgebiete.

Hierarchische Begriffsrelationen setzen semantisches Wissen über die Strukturierung des Gegenstandsbereichs ein. "Durch die schrittweise Weitergabe dieses Wissens an die Benutzer reduziert sich die Komplexität der Suchanfrage" /4/. Sie erleichtern den Nutzern den Einblick in die thematischen Zusammenhänge der Sachgebiete, die ihr Informationsdefizit tangieren.

Die Benutzeroberfläche präsentiert die Begriffshierarchie. Die Darstellung der Begriffshierarchie und die Gestaltung der Navigationsmöglichkeiten wird deshalb zum zentralen Baustein für ein effizientes und benutzerfreundliches IR-System. Die Benutzer sollen die

Schlüsselbegriffe, die sie intuitiv ihrem Informationsdefizit zuordnen, mit den Begriffen der entsprechenden Kontextpositionen in der Begriffshierarchie assoziieren können. Ausgangspunkt ist dabei der benutzerseitig formulierte Begriffscluster.

Während der Navigation durch die Begriffshierarchie kann der ursprüngliche Cluster in der Auseinandersetzung mit den Begriffsrelationen durchaus modifiziert werden. In jedem Fall aber entsteht in der Markierung relevanter Begriffe ein neuer Cluster, der Ergebniscluster, der dem des Benutzers gegenübergestellt wird. Die Suchanfrage gilt dann als formuliert, wenn die Benutzer eine ausreichende Übereinstimmung der gegenübergestellten Cluster feststellen. Der Ergebniscluster repräsentiert dann die Suchanfrage.

## **2.2. Der Interaktionsmodus**

Die WIDAB-spezifische Wahl und Gestaltung der Interaktionsform orientiert sich an dem dargestellten Verständigungsproblem. Nachdem sich zeigte, daß die von den meisten Informationsdiensten bevorzugten formale Technik (Kommandosprache) bei vielen Benutzern - selbst wenn sie hoch geübt sind - starke Belastungen verursachen, wird diese Form zunächst außer Betracht gelassen.

Bei WIDAB stehen die Darstellung von Begriffsrelationen und die Navigationsmöglichkeiten in diesen im Vordergrund. Deshalb fand die grafische Darstellung zusammen mit der direkten Manipulation Eingang in die Realisierung des Prototypen. Das heißt nicht, daß eine natürlichsprachliche Schnittstelle und die erwähnte formale Technik gänzlich unberücksichtigt bleiben. Experimente mit dem Prototypen sollen Ergebnisse liefern, aus denen neue Gestaltungsrichtlinien abgeleitet werden können. Die können durchaus zur Folge haben, daß in die bestehende Interaktionsform Elemente der beiden anderen Modi integriert werden, so daß sich ein Mix aus diesen "drei basalen (generischen) Interaktionsformen" /7/ ergibt.

Den Benutzern stellt WIDAB insgesamt drei Begriffshierarchien entsprechend den drei Hauptbezugsebenen der Wirtschaftsinformationen zur Verfügung. Jedes Dokument weist im Zusammenhang mit Wirtschaftsdaten einen regionalen, Branchen- und Variablenbezug auf. Die Navigationsmöglichkeiten wurden folglich dahingehend gestaltet. Durch die Markierung werden Begriffstripel dem Ergebniscluster zugeordnet. Zusätzlich können die Benutzer noch entscheiden, ob die Begriffe des Tripels isoliert, d.h. ausschließlich, oder unter Hinzunahme ihrer nachgeordneten Begriffshierarchie übernommen werden sollen.

Selbstverständlich sind die Ansprüche an die Benutzeroberfläche auch bei der Form der Ergebnisrepräsentation zu berücksichtigen, der Repräsentation der Dokumente, wie sie

das System als Antwort auf die Suchanfrage der Nutzer ausgibt. Der Aufwand für das Sichten der Ergebnisse aus der Informationssuche wird zunächst dadurch reduziert, indem die auf die Anfrage zugeschnittene Dokumentensammlung in die Begriffshierarchie eingeordnet und damit gegliedert wird. Die Bezeichnungen in den Begriffshierarchien werden dann zu Gliederungspunkten der Dokumentensammlung.

Für die numerischen Daten werden unterschiedliche Repräsentationen angeboten, da bspw. die gemeinsame Darstellung von Zeitreihen in einem Diagramm die Information über deren Entwicklung eher verständlich als die parallele Darstellung in Tabellen.

### **3. Das distribuierte System**

#### **3.1. Die logisch-technische Struktur**

Während die Benutzerschnittstelle den intellektuellen oder physischen Aufwand beeinflusst, den die Nutzer treiben müssen, um eine Suchanfrage zu formulieren und das Suchergebnis einzusehen, beeinflusst die logisch-technische Struktur den Aufwand für die Durchführung der Informationsermittlung. Dieser Aufwand dient den Nutzern als Parameter, inwieweit ihrem Wunsch nach leichtem und kostengünstigen Zugang zu den relevanten Datenbanken entsprochen wird.

Nachdem die Kosten einen nicht zu unterschätzenden Faktor bei der Akzeptanz des Systems darstellen, ist es wirtschaftlich sinnvoll, den gesamten Teil der Datenhaltung am Entstehungsort der Informationen zu belassen und über einen Verbund dieser Datenbanken, der Informationsmittler-Stationen und der Benutzerrechner nur diejenigen Daten zwischen ihnen auszutauschen, die zum jeweiligen Zeitpunkt für die Nutzer von Bedeutung sind.

Um den gesamten Prozeß der Informationserstellung durchführen zu können, wird ein temporäres Rechnernetz realisiert. D.h. die Kommunikationsverbindungen zwischen den jeweiligen Datenstationen werden nur für den Zeitraum der Nachrichtenübermittlung aufrecht erhalten.

Unter Nachrichten sind in diesem Zusammenhang

- a) die Suchanfragen der Benutzer,
- b) die Begriffshierarchien der Informationsmittler (Sachwissen),
- c) die aus diesen beiden Teilen abgeleitete Liste der Schlüssel, die die für die Suchanfrage relevanten Dokumente repräsentieren,
- d) die mit Hilfe der Liste an den Datenbanken der Informationsautoren produzierte Dokumentensammlung selbst und
- e) das Informationspaket (als Suchergebnis) mit der Dokumentensammlung und modifizierten Begriffshierarchien.

Die Datenstationen der Informationsautoren stellen dabei im Gesamtsystem diejenigen Systemelemente dar, an denen die Informationen (d) entstehen, die dem Netz zur Übermittlung an die Datenstationen der Informationsmittler übergeben werden. Von dort erhalten sie über das Netz deren Anforderungen (c). Die Datenstation des Informationsmittlers, die aufgrund ihrer Funktionszuordnung auch mit dem Fachbegriff "Terminal Interface Processor" (TIP) bezeichnet werden kann, wickelt im Gegenzug den Nachrichtenverkehr (bzgl. a,b und e) mit den Nutzern ab. /8/ Sie empfängt die Suchanfrage der Nutzer, übersetzt sie in Algorithmen der jeweiligen Datenbank-Recherche und führt diese an den Datenbanken durch. Nachdem sie alle notwendigen Dokumente gesammelt und aufbereitet hat, übermittelt sie das fertige Informationspaket an die Nutzer. Die Datenkommunikation basiert ausschließlich auf Filetransfer und ist automatisiert, so daß für die Nutzer hier kein Aufwand entsteht.

Das System WIDAB realisiert aufgrund dieser Darstellung ein temporäres Rechnernetz, das aus einer Menge von Funktionskomponenten besteht, die einerseits in räumlicher Distanz zueinander stehen und andererseits Benutzeraufträge in kooperativer Autonomie bearbeiten. Der einzelne Nutzer braucht die Arbeitsgliederung und den Ort der jeweiligen Funktionskomponenten nicht zu kennen. Ein System mit derartigen Eigenschaften nennt Jessen /9/ "Distribuiertes System".

### **3.2. Die funktionale Struktur**

Die Gestaltung der funktionalen Struktur, d.h. der Struktur der Daten und Prozesse sowie deren systeminterne Wechselwirkungen orientiert sich an dem Ziel, im Rahmen des definierten distribuierten Systems die Datenbestände mehrerer Datenstationen (die Dokumente der Informationsautoren und das Sachwissen der Informationsmittler) logisch zusammenzufassen. /8/ /9/

Die Benutzer definieren mit Hilfe der Menünavigation, die auf den Begriffshierarchien der Informationsmittler basieren, ihre Suchanfrage. Dies geschieht lokal ohne Verbindung zu irgendeiner Datenstation des Netzes. Erst wenn die Benutzer diesen Prozesses abgeschlossen und den Auftrag zur Übermittlung der Suchanfrage an den Informationsmittler erteilt haben, wird die Kommunikationsverbindung zur TIP-Datenstation aufgebaut, die entsprechende Datei dorthin geschickt und die Verbindung wieder abgebaut. Nach dem Eingang der Suchanfrage (d.h. der sie charakterisierenden Liste mit Begriffsstripeln) bei der TIP-Datenstation führt das System einen Abgleich mit den Begriffsstrukturen durch. Dabei werden die Teile extrahiert und neu zusammengefügt, die notwendig sind, um die markierten Begriffe der Suchanfrage anzusteuern. Diese neu entstandenen Strukturen sollen später die Nutzer bei der Einsicht der Dokumentensammlung führen. Irrelevante Verzweigungsmöglichkeiten wären dabei nur hinderlich.



Die Dokumentensammlung bildet zusammen mit den an ihrem Inhalt orientierten Begriffsstrukturen das Informationspaket (Suchergebnis).

Die Verwaltung der Verweise aus den Begriffsstrukturen auf die Schlüssel der Dokumente fällt ausschließlich in den Tätigkeitsbereich der Informationsmittler. Jeder Kombination aus jeweils einem Knoten der drei Begriffshierarchien (entsprechend den Bezugsebenen Branchen, Regionen und Variablen) ist eine (durchaus auch leere) Liste der Schlüssel für die entsprechenden Dokumente zugeordnet. Auf diese Weise erzeugt das System eine für die Suchanfrage spezifische Referenzliste. D.h. es erarbeitet eine Anforderungsliste für die Datenbanken, in der alle Schlüssel für die Dokumente enthalten sind, die in Zusammenhang mit der Suchanfrage stehen. Diese neu entstandenen Dateien werden den zuständigen Informationsquellen zugeordnet und anschließend per Filetransfer an die jeweiligen Datenbanken transferiert.

Dort werden vorgefertigte Queries angestoßen, die die genannten Anforderungslisten abarbeiten und das Query-Ergebnis, die Dokumentensammlung in weiteren Dateien in einem zwischen Informationsmittler und -autor vereinbartem Format abgelegt. Nach dem die automatische Recherche durchgeführt wurde, senden die Datenstationen der Informationsautoren die Dateien zurück an die TIP-Stationen.

Nach Erstellung dieser Dokumentensammlung bei den Datenanbietern und ihrer Übermittlung an den Informationsmittler wird die Generierung des Informationspaketes (d.h. die Verknüpfung der neu erstellten Begriffshierarchien mit der Dokumentensammlung und ihre Abspeicherung in einer Datei) an den TIP-Datenstationen vorgenommen. Das letzte Glied der Transaktionskette aktiviert den relevanten Transportweg zu den Nutzern und realisiert den Filetransfer des gewünschten Informationspaketes. Mit Hilfe der Menünavigation kann nun die Dokumentensammlung eingesehen werden.

#### **4. Wissensrepräsentation des automatisierten Informationsmittlers**

WIDAB stellt keine direkte Verbindung zu den unterschiedlichen, für die Fragestellungen der Nutzer relevanten Datenbanken her, sondern bedient sich eines automatisierten Informationsmittlers. Diese zentrale Vermittlungsstelle generiert aus der Suchanfrage der Nutzer unter Zuhilfenahme ihres Sachwissens (Begriffshierarchie mit Verweisen auf relevante Dokumente) eine Menge an Antwortobjekten (Dokumenten) als Ausschnitt aus der virtuellen Gesamtdatenbank. Die realisierte Datenbank-Recherche liefert dann Informationen, die stark auf die Begriffe des Informationsproblems unter Berücksichtigung ihres jeweiligen Kontexts zugeschnitten sind, wodurch der notwendigen Problemorientiertheit Rechnung getragen wird.

#### **4.1. Das Sachwissen abgebildet in Begriffshierarchien**

Während die Benutzeroberfläche von WIDAB dafür verantwortlich ist, daß die Benutzer des IR-Systems ihr Informationsproblem in der Suchanfrage soweit präzisieren bzw. modifizieren können, um es möglichst scharf abzugrenzen, muß der automatisierte Informationsmittler für eine genaue Übersetzung der Suchanfrage in geeignete Datenbank-Recherchen sorgen. Beide Prozesse aber stützen sich auf weitere Daten, die - in das System integriert - das Sachwissen des Analytikers, d.h. des Informationsmittlers, in strukturierter Form abbilden.

Durch den rekursiven Prozeß der thematischen Untergliederung entsteht für jedes Sachgebiet (z.B. eine Branche) eine hierarchische Begriffsstruktur. Diese bildet dann nicht nur Begriffe ab, sondern auch die Beziehungen zwischen ihnen und erzeugt so die Repräsentation der Zusammenhänge. Jeder der darin enthaltenen Begriffe wird deshalb nicht mehr isoliert betrachtet, sondern in Wechselbeziehung mit den ihn umgebenden Begriffen. Dadurch wird deutlich, in welchem Kontext er zu sehen ist. Mittels dieser Abstraktion wird dem Nutzer eine schnelle Information über das gesamte Wissensterrain geliefert, mit dem sein Wissensdefizit, d.h. sein Informationsproblem, abgedeckt werden kann /10/. Die Nutzer ordnen intuitiv ihrem Informationsproblem bestimmte Schlüsselbegriffe zu und grenzen es dabei anhand charakterisierender Begriffe ab (siehe Abschnitt 2). Diese oder deren Synonyma finden sie im durch die Begriffsstruktur dargestellten Kontext wieder. Diese Identifikation der Begriffe ihres Interesses, dieses "Anklicken" ihrer relevanten Knoten, verkörpert die Artikulation ihres Informationsproblems.

Die Begriffshierarchien unterstützen im gesamten Prozeß der Informationserstellung folgende Funktionen:

Sie steuern das Menüsystem, mit dessen Hilfe die Nutzer in ihrem Arbeitsbereich ihre Suchanfrage einfach, aber präzise definieren können. Durch die gleiche Herangehensweise können sie die Dokumentensammlung einsehen. Das System bietet dann an den Begriffsknoten Funktionen zur Darstellung der Dokumente (Texte und Zeitreihen).

Sie dient der automatischen Generierung der Informationspakete beim Informationsmittler (siehe Abschnitt 3.2).

Sie unterstützt die automatische Zuordnung der Dokumente - genauer gesagt ihrer Schlüssel - anhand ihrer charakterisierenden Begriffe.

#### **4.2. Die Repräsentation des Sachwissens in einem semantischen Netz**

Für WIDAB wird derzeit ein Modul geschaffen, das die komplexen Begriffshierarchien, sowie ihre vielfältigen Beziehungsarten verwaltet. Diese Entwicklung orientiert sich an

der strukturierten Wissensrepräsentationsmethode, d.h. an der Darstellung von Objekten und Beziehungen in einem semantischen Netz /10/.

Die Syntax eines semantischen Netzes, d.h. die Symbole und ihre Verwendungsmöglichkeiten, sind einfach: Es gibt Objekte, und es gibt Beziehungen zwischen Objektpaaren. In grafischer Form sind Objekte durch benannte Kreise und die Beziehungen zwischen ihnen durch benannte Pfeile gekennzeichnet. In der Standardterminologie werden die benannten Kreise Knoten und die benannten Pfeile Kanten genannt. Die Beschreibungen der Objekte (Begriffshierarchien) liegen in natürlicher Sprache vor, wodurch das semantische Netz, das in Zusammenhang mit WIDAB Anwendung findet, auf der deskriptiven Semantik basiert /10/. Weiter wird vereinbart, "daß die Slots eines Knoten den unterschiedlich benannten Kanten entsprechen" /10/.

In Anlehnung an die Graphentheorie /11/ kann eine Begriffshierarchie als ein Wurzelbaum mit seiner begrifflichen Hauptbezugsebene als Wurzel ("Einstiegspunkt") betrachtet werden. In dieser Betrachtung werden die Begriffe durch Knoten und die Beziehungen zwischen ihnen durch Pfeile repräsentiert. Die Slots eines Knotens entsprechen den unterschiedlich benannten Pfeilen. Die unterschiedlichen Funktionen, die das Netz unterstützen soll, spiegeln sich auch in den unterschiedlichen Slot-Klassen wieder.

Die KONTEXT-Slots verkörpern die Kontextabhängigkeit der Knoten und erzeugen damit die thematische Gliederung eines Sachgebiets. Im IR-System WIDAB unterstützen sie den Aufbau der Substrukturen für die Sachgebiete Branchen, Variablen und Regionen.

Die Kanten der KONTEXT-Slots legen die Pfade der Menünavigation fest. Im Nutzerbereich wird so die Möglichkeit geschaffen, die Suchanfrage zu definieren bzw. die Dokumentensammlung einzusehen. Im Informationsmittler-Bereich kommt ihnen eine wichtige Funktion im Rahmen der Verwaltung der Strukturdaten zu: Sie ermöglichen die automatische Zuordnung neuer Dokumente (genauer, ihrer Schlüssel) durch Abgleich ihrer charakterisierenden Begriffe mit den Knotenbegriffen.

Die SUBSUMTION-Slots stellen Beziehungen zu Begriffen her, die unter dem Begriff des Knotens zu subsumieren sind, von dem sie ausgehen. Sie unterstützen die Funktion der KONTEXT-Slots, indem sie diese weiter spezifizieren. Deren Aussagekraft wird erhöht, da sie um weitere charakterisierende Elemente ergänzt werden. Sie steigern damit die Fähigkeit des Systems, die korrekte Zuordnung der Dokumente automatisch vorzunehmen. Für den Benutzerbereich sind sie unbedeutend, da sie in der Menüauswahl nicht dargestellt werden.

Die DETAIL-Slots hingegen realisieren die Verbindung zwischen Knoten in den Begriffshierarchien der drei thematischen Hauptbezugsebenen (Begriffstripel) und den Schließeln der Dokumente. Sie stellen das zentrale Bindeglied zwischen den Begriffshierarchien und den Datenbankinhalten der Informationsautoren dar.

Im Mittelpunkt des besagten Moduls für das Retrieval-System WIDAB steht nicht die automatische Generierung der Begriffshierarchien. Sie werden ebenso wie die Zuordnung der Dokumente im derzeitigen Prototyp auf manuellem bzw. halbautomatischem Wege erzeugt. Es wird aber das Ziel verfolgt, diesen Zuordnungsprozeß so weit als möglich zu automatisieren.

## Literatur

- /1/ Häußler, E.: Perspektiven und Strukturen elektronischer Informationssysteme. In: Häußler, E.: Auf dem Weg zum europäischen Informationsverbund: Von CD-ROM bis Informationsdatenbanken. Velbert 1990. Pickels, Ch.: Historical and Realtime Information: Increased Importance of Information Services to the Single Market 1992. In: Häußler, E.: Auf dem Weg zum europäischen Informationsverbund: Von CD-ROM bis Informationsdatenbanken. Velbert 1990. Reyes, G.: Die Qualität von Datenbanken mit Unternehmensdaten: Ein internationaler Vergleich. In: Häußler, E.: Europäischer Congress für Mehrwertdienste, Informations-Datenbanken und CD-ROM. Velbert 1991. Steven, G.: Gemeinschafts-Tätigkeiten zur Entwicklung des Europäischen Informationsmarktes. In: Häußler, E.: Auf dem Weg zum europäischen Informationsverbund: Von CD-ROM bis Informationsdatenbanken. Velbert 1990.
- /2/ Ahrend, W.: Die Nutzung von externen Informationsdatenbanken durch kleine und mittlere Unternehmen: Ergebnisse einer empirischen Untersuchung. In: Herget J.; Kuhlen, R.: Pragmatische Aspekte beim Entwurf und Betrieb von Informationssystemen. Proceeding des 1. Internationalen Symposiums für Informationswissenschaft. Konstanz 1990. S.553-566. Commission of the European Communities. Directorate General for Telecommunications, Information Industries and Innovation (DGXIII/B): Electronic Information Users in Europe. Sixth Survey. Luxemburg 1990. Koch, A.: Informationsbeschaffung im Zeichen des EG-Binnenmarktes. In: Häußler, E.: Auf dem Weg zum europäischen Informationsverbund: Von CD-ROM bis Informationsdatenbanken. Velbert 1990. Wietersheim, B. von: Markt für Online-Datenbanken in der Bundesrepublik Deutschland, in Europa und weltweit: Ergebnisse einer Studie. In: Häußler, E.: Auf dem Weg zum europäischen Informationsverbund: Von CD-ROM bis Informationsdatenbanken. Velbert 1990.
- /3/ Mang, E.; Schweiger, J.: Der Akzeptanztest bei den Unternehmen. In: Englmeier, K. u.a.: WIDAB-Projekt. Ergebnisse der Praxisphase. Deutsches Institut für Wirtschaftsforschung, Berlin, und Ifo-Institut für Wirtschaftsforschung, München 1991. S.47-68.
- /4/ Krause, J.: Zur Architektur von Wing: Modellaufbau, Grundtypen der Informationssuche und Integration der Komponenten eines intelligenten Information Retrieval. WING-IIR-Arbeitsbericht 7. LIR Regensburg 1990. S.8-48.
- /5/ Salton, G.: Information Retrieval - Grundlegendes für Informationswissenschaftler. Hamburg u.a. 1987. S.1-17.

- /6/ Bössmann,E.: Information. In: Albers,W. u.a.: Handwörterbuch der Wirtschaftswissenschaften. Stuttgart 1978. S.185/186.
- /7/ Bullinger,H.-J.: Software-Ergonomie: Stand und Entwicklungstendenzen. In: Schönplug W.: Software-Ergonomie '87. Nützen Informationssysteme dem Benutzer. Stuttgart 1987. S.17-30.
- /8/ Schnupp,P.: Rechnernetze. 2.Auflage. Berlin 1982. S.19-27.
- /9/ Jessen,E.: Rechnernetze. Vorlesungsskript, Wintersemester 1988/89.
- /10/ Winston,P.H.: Künstliche Intelligenz. Bonn u.a. 1987. S.265-278.
- /11/ Domschke,W.: Logistik: Transport. 3.Auflage. München 1989. S.3-15. Schmidt,G.; Ströhlein,T.: Relationen und Graphen. Berlin u.a. 1989. S.81-143.

Kurt Englmeier  
Ifo-Institut für Wirtschaftsforschung  
Poschinger Str. 5  
8000 München 86

# **AIR/PHYS - Automatische Indexierung in der Praxis**

**Michael Schwantner**

## **Zusammenfassung**

Seit 1985 setzt das Fachinformationszentrum Karlsruhe das automatische Indexierungssystem AIR/PHYS in der Inputproduktion der Datenbank PHYS ein. AIR/PHYS teilt englischsprachigen Referatetexten Deskriptoren aus einem vorgegebenen Indexierungsvokabular zu. Wichtigstes Werkzeug dabei ist ein weitgehend automatisch erstelltes Wörterbuch, das zum großen Teil aus statistischen, auf Assoziationsfaktoren beruhenden Term-Deskriptor-Beziehungen besteht. In der vorliegenden Arbeit werden der Ablauf der Indexierung mit AIR/PHYS und der Aufbau des Wörterbuches beschrieben. Ferner wird auf die Problematik der Wörterbuch-Pflege eingegangen und entsprechende, beim Fachinformationszentrum Karlsruhe geplante Arbeiten werden vorgestellt.

## **Abstract**

### **AIR/PHYS - Automatic Indexing in Practice**

Since 1985, the automatic indexing system AIR/PHYS is being applied by the Fachinformationszentrum Karlsruhe for the input production of the physics database PHYS. AIR/PHYS assigns descriptors chosen from a controlled indexing vocabulary to abstracts written in English. An essential tool for that is a mainly automatically constructed dictionary. It consists to a major part of statistical term-descriptor relations based on association factors. In this paper, the process of indexing with AIR/PHYS and the dictionary structure itself are outlined. Furthermore, problems of dictionary maintenance are discussed and corresponding developments planned for the future at the Fachinformationszentrum are presented.

## **1 Einleitung**

Das Fachinformationszentrum Karlsruhe (FIZ) ist Deutschlands führender Datenbank-Produzent für die Bereiche Energie, Physik und verwandte Gebiete. Zusammen mit der American Chemical Society (ACS) und dem Japan Information Center of Science and Technology (JICST) betreibt es das Scientific and Technical Information Network, STN International. Über dieses Datennetz werden ca. 110 wissenschaftliche und technische Datenbanken angeboten. Etwa zur Hälfte sind es bibliographische, die übrigen sind Faktendatenbanken oder numerische Datenbanken. Der Zugang erfolgt mit der Retrievalsprache Messenger über Datennetze, in Deutschland z.B. über DATEX-P.

Die vom FIZ produzierte Datenbank PHYS ist eine umfassende bibliographische Datenbasis für die Gebiete Physik, Astronomie und Astrophysik sowie relevante Randgebiete. Es wird weltweit publizierte Fachliteratur (Zeitschriftenartikel, Bücher, Konferenzbeiträge u.a.) er-

faßt. PHYS deckt den Zeitraum seit 1979 ab und enthält heute (September 1991) mehr als 1,5 Mio. Dokumentationseinheiten (DE). Bei dem vierzehntäglichen Update kommen jeweils ca. 5000 DE neu hinzu. Eine DE besteht neben bibliographischen Angaben aus dem englischsprachigen Titel und Abstract des Originalartikels sowie aus inhaltserschließenden Angaben. Letztere beinhalten u.a. Klassifikationsnummern und Deskriptoren (Schlagwörter), die auf der Basis eines z.Zt. etwa 24.000 Deskriptoren umfassenden Thesaurus vergeben werden.

Das Indexierungssystem AIR/PHYS wurde mit dem zugehörigen Indexierungswörterbuch PHYS/PILOT\* von der TH Darmstadt im Rahmen des Projekts PILOTANWENDUNG AIR/PHYS entwickelt, wo man sich seit 1977 mit den Problemen der automatischen Indexierung beschäftigt. Dem Indexierungssystem liegt der Darmstädter Indexierungsansatz (DIA) /1, 2/ zugrunde. Dieser beruht auf folgenden Voraussetzungen:

- Es erfolgt eine Zuteilung von Deskriptoren, die aus einem vorgegebenen Indexierungsvokabular stammen.
- Es wird nicht der Volltext des Dokuments indexiert, sondern nur Titel und Abstract, die in englischer Sprache vorliegen müssen.
- Die Indexierung erfolgt mit einem Wörterbuch, zu dessen Erstellung eine größere Anzahl manuell indexierter Dokumente vorhanden sein muß.

Aus der Vorgabe, daß die Indexierung mit kontrolliertem Vokabular zu erfolgen hat, ergibt sich ein hoher Anspruch: Die Indexierung muß begriffsorientiert sein und muß, im Gegensatz zu einer Freitextindexierung, von den konkreten Formulierungen des Textes losgelöst sein.

Die Leistung von AIR/PHYS wurde vom FIZ zunächst in einer mehrmonatigen Pilotanwendungsphase geprüft. Auf Grund der Ergebnisse beschloß man, AIR/PHYS für den Routinebetrieb einzusetzen. Zum gegenwärtigen Zeitpunkt wurden bereits mehr als 400.000 Dokumente durch dieses System indexiert.

AIR/PHYS ist damit das erste Indexierungssystem, das in der Praxis für eine anspruchsvolle Indexierung zur Inputproduktion einer Datenbank in diesem Umfang eingesetzt wird.

## **2 Ablauf der Indexierung mit AIR/PHYS**

In der Inputproduktion werden zunächst die für die Datenbank PHYS relevanten Dokumente bestimmt. Nach der Erfassung der Dokumente, die Korrekturen und die Bestimmung der bibliographischen Angaben einschließt, werden Titel und - sofern vorhanden - Abstract der Dokumente dem System AIR/PHYS übergeben, dessen Aufgabe nun die Zuteilung der Deskriptoren ist. Im Zuge weiterer, intellektuell durchzuführender Erschließungsarbeiten wie das Zuordnen der Sachklassifikation wird anschließend die automatische Indexierung von Fachreferenten kontrolliert und gegebenenfalls korrigiert.

---

\* In Zusammenarbeit mit der Gesellschaft für Information und Dokumentation (GID) und dem FIZ; gefördert vom BMFT als Förderungsmaßnahme 10131138.

Der Indexierungsprozeß kann in einen Beschreibungsschritt und einen Entscheidungsschritt unterteilt werden.

Im Beschreibungsschritt wird zunächst das Dokument  $d$  (d.h. dessen Titel und Abstract) in seine einzelnen Wörter zerlegt, grammatikalische Funktionswörter werden anhand einer Stoppwortliste entfernt und die verbleibenden Wörter werden mit einem Algorithmus nach Kuhlen /3/ auf ihre Grundform reduziert. Mittels eines von Muhler und Repp /4/ entwickelten Verfahrens werden die in den Texten enthaltenen chemischen und physikalischen Formeln analysiert und auf standardisierte Formelbezeichner abgebildet. Außerdem werden die in den Texten vorkommenden Mehrwortgruppen (aus mehreren Wörtern zusammengesetzte Fachausdrücke) mit einem wörterbuch-abhängigen, einfachen Matchingverfahren identifiziert. Das dafür notwendige Mehrwortgruppen-Wörterbuch wurde mit einem wörterbuch-unabhängigem Verfahren aufgebaut, das auf dem Begrenzerverfahren von Jaene und Seelbach /5/ beruht und von Kienitz-Vollmer und Reichardt /6/ weiterentwickelt wurde. Wörter, Mehrwortgruppen und Formelbezeichner werden zusammenfassend als Terme bezeichnet. Zu jedem Term wird die Form des Vorkommens festgehalten, zu der Angaben u.a. darüber gehören, ob der Term im Titel oder Abstract identifiziert wurde oder wieviele andere Wörter zwischen den einzelnen Komponenten einer Mehrwortgruppe stehen. Die Form des Vorkommens liefert dem System Informationen über die Signifikanz des Terms für das betreffende Dokument und über die Sicherheit mit der er identifiziert werden konnte. Anschließend wird für jeden Term  $t$  geprüft, ob er im Indexierungswörterbuch enthalten ist und dort mit einer (gewichteten) Relation mit einem Deskriptor  $s$  verknüpft ist. Ist dies der Fall, hat damit das System einen Hinweis  $t \rightarrow s$  auf den Deskriptor bestimmt, der intern durch die Hinweisbeschreibung dargestellt wird. Diese setzt sich zusammen aus

- Angaben über die Form des Vorkommens von  $t$ ,
- Signifikanzangaben: statistische Angaben über  $t$  und  $s$ , z.B. die Zuteilungshäufigkeit,
- Angaben über die Art der Relation und über das Gewicht, falls der Hinweis auf einer gewichteten Relation beruht.

Abschließend werden jeweils alle Hinweise auf  $s$  zur Relevanzbeschreibung  $x(d,s)$  als Vektor zusammengefasst:  $x(d,s) = (x_1, \dots, x_n)$ . Der Relevanzbeschreibungsvektor enthält zusätzlich noch Angaben, die von den Hinweisen unabhängig sind, wie etwa die Dokumentlänge.

Der Entscheidungsschritt besteht in der Anwendung der Indexierungsfunktion  $a$ , einem linearem Polynom, auf  $x(d,s)$ . Das Ergebnis ist das Indexierungsgewicht  $g(d,s)$ :

$$g(d,s) = a(x(d,s)) = a_0 + \sum_{1..n} a_i x_i .$$

Die Koeffizienten  $a_i$  werden vor dem eigentlichen Einsatz von AIR/PHYS mit einem als Adaption bezeichneten Verfahren der multiplen Regressionsanalyse bestimmt /7, 8/, das außerdem über die Signifikanz der Variablen  $x_i$  Aufschluß gibt.

In einer zweiten Indexierungsstufe /9/ wird u.a. aus  $a(x(d,s))$ , einigen Beschreibungselementen  $x_i$  von  $x(d,s)$  und den im Wörterbuch enthaltenen Deskriptor-Deskriptor-



Relationen ein zweiter Beschreibungsvektor  $x'(d,s)$  aufgebaut, auf den eine weitere, ebenfalls polynomiale Indexierungsfunktion  $a'$  angewandt wird:  $g'(d,s) = a'(x'(d,s))$ .

Mit  $g'(d,s)$  erhält man eine gewichtete Indexierung; die für die Datenbank PHYS erforderliche ungewichtete Indexierung wird erzeugt, indem all die Deskriptoren zugeteilt werden, für die  $g'(d,s)$  einen vorgegebenen Schwellenwert  $g_0$  übersteigt.

Eine über diese knappe Zusammenfassung hinausgehende Beschreibung findet sich in /7, 10, 11/.

### 3 Das Indexierungswörterbuch

Aus dem Ablauf des Beschreibungsschrittes wird deutlich, daß ein Term nur dann zum Indexierungsergebnis beitragen kann, wenn er im Indexierungswörterbuch enthalten ist und dort durch eine Relation mit einem Deskriptor verknüpft ist. Dementsprechend muß das Wörterbuch zu möglichst allen Termen des Fachgebiets möglichst viele Relationen enthalten.

Das Wörterbuch PHYS/PILOT wurde weitgehend maschinell erstellt. Dazu war eine repräsentative und möglichst große Menge  $D$  manuell indexierter Dokumente notwendig. Zu diesem Zweck wurden vom FIZ 392.000 Dokumente zur Verfügung gestellt. PHYS/PILOT enthält:

- Das Indexierungsvokabular  $V$ ,
- das Termvokabular  $T$ , d.h. alle Terme in der Grundform, die mit mindestens einem Deskriptor durch eine Relation verknüpft sind,
- Signifikanzangaben zu den Elementen aus  $T$  und  $V$ :
  - Für  $t \in T$ :  $f(t)$ , die Anzahl der Dokumente aus  $D$ , in denen der Term  $t$  im Titel oder Abstract vorkommt (Dokumenthäufigkeit). Dabei wird  $t$  in einem Dokument  $d$  dann als vorkommend angesehen, wenn  $d$  alle Komponenten von  $t$  in Grundform enthält und zwischen den Komponenten höchstens sechs andere Wörter stehen.
  - Für  $s \in V$ :  $m(s)$ , die Zuteilungshäufigkeit für  $s$  bzgl.  $D$ .
- sowie folgende Relationen:
  - Die aus dem Thesaurus übernommenen Relationen Broader Term, Narrower Term und USE sowie weitere Deskriptor-Deskriptor-Relationen.
  - Die Identitäts-Relation, eine aus technischen Gründen in das Wörterbuch aufgenommene Relation, die es dem System erlaubt, einen als Term im Titel oder Abstract des Dokuments vorkommenden Deskriptor zu identifizieren und
  - die Relation  $Z$ .

Die Relation  $Z$  ist die quantitativ und qualitativ wichtigste Relation für AIR/PHYS. Sie wird maschinell berechnet und beruht auf dem Assoziationsfaktor  $z(t,s)$ , der durch

$$z = \frac{h(t,s)}{f(t)}$$

definiert ist /12/, wobei

$h(t,s)$  die Anzahl der Dokumente ist, denen der Deskriptor  $s$  manuell zugeteilt ist und in denen  $t$  im Titel oder Abstract vorkommt (Paarhäufigkeit).

Der  $z$ -Wert  $z(t,s)$  ist damit eine Schätzung der bedingten Wahrscheinlichkeit  $P(s|t)$ , daß der Deskriptor  $s$  dem Dokument manuell zugeteilt wird, wenn der Term  $t$  darin vorkommt.

Im Indexierungswörterbuch sind diejenigen Terme und Deskriptoren miteinander durch die Relation  $Z$  verbunden, für die  $z(t,s)$  und  $h(t,s)$  vorgegebene Schwellenwerte  $c_1$  und  $c_2$  überschreiten (für PHYS/PILOT wurde  $z(t,s) \geq 0.3$  und  $h(t,s) \geq 3$  gefordert). Durch die Verwendung dieser nach heuristischen Gesichtspunkten gewählten Schwellenwerte wird zum einen der Umfang des Wörterbuches in handhabbaren Grenzen gehalten, zum anderen werden statistisch allzu unsichere  $z$ -Werte von der Aufnahme ausgeschlossen. Aus diesem Grund auch enthält das Wörterbuch die Terme in Grundform, da  $f(t)$  dann höher ist, als wenn die Vollformen betrachtet würden. In die Beschreibung eines auf der Relation  $Z$  beruhenden Hinweises  $t \rightarrow s$  wird neben  $z(t,s)$  auch  $h(t,s)$  mit aufgenommen, so daß die von  $h(t,s)$  und  $f(t)$  abhängende statistische Unsicherheit der Schätzungen  $z(t,s)$  mit berücksichtigt wird.

Die Methoden für die Erstellung von PHYS/PILOT sind an anderer Stelle /13, 14, 15/ ausführlich dokumentiert.

Tab. 1: Angaben zu PHYS/PILOT

Dokumente  D	392.000
Indexierungsvok.  V  (Stand 1985)	22.683
an Relation Z beteiligte Deskriptoren	10.002
in D enthaltene verschiedene Terme	1.554.600
an der Relation Z beteiligte Terme	≈190.000
Anzahl der berechneten $z$ -Werte	805.852
- davon in PHYS/PILOT aufgenommen	355.932
Relationen USE, Identität	72.821
Deskriptor-Deskriptor-Relationen	192.897

Aus Tabelle 1 geht hervor, daß nur etwa 44% aller Deskriptoren von der Relation  $Z$  erfaßt sind und weniger als 13% aller in  $D$  vorkommenden Terme. Die daraus resultierende Unvollständigkeit des Wörterbuches, die nur zum Teil durch die Aufnahme der Thesaurus-Relationen kompensiert werden kann, wird in Abschnitt 5 ausführlich diskutiert werden. Von den ursprünglich mehr als 800.000  $z$ -Werten (die alle die Bedingungen  $z(t,s) \geq 0.3$  und  $h(t,s) \geq 3$  erfüllten) konnte nur ein Teil in PHYS/PILOT übernommen werden, da sich sonst Probleme bei der Speicherplatzverwaltung und dem Laufzeitverhalten ergeben hätten. Es wurden daher auf kleinen Häufigkeiten beruhende  $z$ -Werte weggelassen, wobei aber darauf

geachtet wurde, daß alle bislang von der Relation Z erfaßten Deskriptoren dies auch weiterhin blieben.

#### 4 Bewertung der Indexierungsergebnisse

Zur Bewertung des Systems AIR/PHYS wurden Indexierungstests durchgeführt, bei denen die automatische mit der korrigierten Indexierung verglichen wurde. Hierzu wurden die den Retrievalmaßen Precision und Recall entlehnten Vollständigkeitsquote  $r$  und Trefferquote  $p$  verwendet, die wie folgt definiert sind:

$$r = \frac{\sum_{d \in D} |M_d \cap A_d|}{\sum_{d \in D} |M_d|} \quad p = \frac{\sum_{d \in D} |M_d \cap A_d|}{\sum_{d \in D} |A_d|}$$

wobei

$M_d$  die Menge der dem Dokument  $d$  aus  $D$  manuell zugeteilten Deskriptoren und  $A_d$  die Menge der dem Dokument  $d$  aus  $D$  automatisch zugeteilten Deskriptoren ist.

Bei einem 1987 mit ca. 20.000 Dokumenten durchgeführten Indexierungstest ergaben sich folgende Werte:

$$p = 0,655 \quad , \quad r = 0,658 \quad .$$

Da die Tiefe der automatischen Indexierung durchschnittlich 12 Deskriptoren je Dokument beträgt, bedeutet dies, daß bei der Korrektur etwa 4 Deskriptoren gestrichen werden und ebenso viele wieder hinzugefügt werden.

Bei der Interpretation der Ergebnisse sind zwei Umstände zu berücksichtigen: Erstens trifft AIR/PHYS die Indexierungsentscheidung allein auf Basis von Titel und Abstract, während den Indexierern zusätzlich der Volltext zur Verfügung steht. Ein Grund, warum sich die Realisierung des Darmstädter Indexierungsansatzes bisher auf die Indexierung von Titel und Abstract beschränkte, war, daß es erst in jüngster Zeit maschinenlesbare Volltexte gibt. Zweitens wurde AIR/PHYS so parametrisiert, daß es durchschnittlich 12 Deskriptoren je Dokument zuteilt, wogegen die Tiefe der manuellen Indexierung früher bei etwa 8,8 lag. Die Absicht dabei war, den Indexierern zusätzliche Deskriptoren als Vorschlag anzubieten. Als Folge davon gingen die Indexierer ebenfalls zu einer höheren Indexierungstiefe über, so daß sich durch die korrigierte automatische Indexierung beim Retrieval bei gleichbleibender Precision bessere Ergebnisse für den Recall ergeben /16, 17/.

#### 5 Input, Wörterbuch-Qualität, Indexierungsqualität

Bei der Planung und Erstellung eines Wörterbuches sind Fragen über die Beziehung zwischen Input, Wörterbuch-Qualität und Indexierungsqualität von Interesse. Der Input kann durch den Umfang  $|D|$  der verwendeten Dokumentmenge angegeben werden, wobei die durchschnittliche Dokumentlänge, die Indexierungstiefe, der Umfang des Indexierungsvokabulars und die Größe des Anwendungsgebietes mit zu berücksichtigen ist. Eine Möglichkeit,

die Indexierungsqualität zu messen, wurde in Abschnitt 4 vorgestellt. In /2, 18/ werden weitere Möglichkeiten diskutiert.

Die Wörterbuch-Qualität läßt sich durch die Aspekte Vollständigkeit und Genauigkeit beschreiben. Zwei einfache Maße für die Vollständigkeit eines auf der Relation Z beruhenden Wörterbuches sind durch

$V_1$  : Anteil der Deskriptoren am Indexierungsvokabular V, die durch die Relation Z erfaßt sind, und

$V_2$  : die durchschnittliche Anzahl der z-Werte für einen an der Relation Z beteiligten Deskriptor gegeben.

Bei einem idealen, d.h. vollständigen Wörterbuch gilt  $V_1=1$ , d.h. alle Deskriptoren sind von der Relation Z erfaßt und  $V_2$  ist mit dem Umfang des Fachvokabulars identisch. In der Praxis läßt sich ein solches Wörterbuch allerdings nicht realisieren, da dazu ein unendlich großer Input notwendig wäre.

Die Vollständigkeit eines Wörterbuches läßt sich auch dadurch messen, indem man angibt, wieviele zusätzliche Einträge man erhalten würde, wenn man bei der Erstellung den Input um ein Dokument vergrößern würde. Erwartungswerte hierfür wurden von Hüther in /19/ hergeleitet.

Ein weiteres Merkmal für die Qualität eines Wörterbuches ist die Genauigkeit, d.h. die statistische Sicherheit der darin enthaltenen z-Werte. Diese läßt sich mit der zu erwartenden Abweichung des Schätzwertes  $z(t,s)$  von der tatsächlichen Wahrscheinlichkeit  $P(s|t)$  wiedergeben.

Es werden folgende Abkürzungen verwendet:  $z := z(t,s)$ ,  $h := h(t,s)$ ,  $f := f(t)$ ,  $p := P(s|t)$ .  $z$  ist binomialverteilt und nach dem Gesetz der großen Zahl ist die Abweichung  $|p-z|$  höchstens von der Größenordnung

$$\sqrt{\frac{p(1-p)}{f}}$$

Für  $p=0,5$  ist  $p(1-p)$  maximal, so daß

$$\Delta(f) := \frac{1}{2\sqrt{f}}$$

die Größenordnung der Abweichung des z-Wertes  $z(t,s)=h/f$  vom tatsächlichen Wert  $P(s|t)$  widerspiegelt. Um die Genauigkeit aller z-Werte des Wörterbuches global anzugeben, wird folgender gewichteter Mittelwert definiert:

$$\underline{\Delta} = \frac{\sum_f \Delta(f) \cdot f \cdot \#(f)}{\sum_f \#(f)} = \frac{\sum_f \sqrt{f} \cdot \#(f)}{2 \sum_f \#(f)}$$

Dabei ist  $\#(f)$  die Anzahl der z-Werte im Wörterbuch, die auf dem Wert  $f$  beruhen.  $\underline{\Delta}$  ist ein einfach zu berechnendes, aber auch verhältnismäßig grobes Maß. Genauere Schätzungen für

$|p-z|$  erhält man, wenn man die Binomialverteilung durch die Normalverteilung oder F-Verteilung nähert. Auf einfachere Weise erhält man mit  $u \cdot \Delta$  eine zwar ebenfalls grobe, aber für die hier verfolgten Ziele hinreichend genaue Schätzung der maximalen Abweichung  $|p-z|$  zur Sicherheitswahrscheinlichkeit  $\alpha$ , wobei  $u$  das  $(1+\alpha)/2$ -Quantil der Normalverteilung ist /18/.

Tabelle 2 zeigt die Werte für den Input und die Wörterbuch- bzw. Indexierungsqualität für PHYS/PILOT und drei weitere Wörterbücher. Letztere decken das Gebiet der Ernährungswissenschaften ab und wurden mit Dokumenten aus den Food Science and Technology Abstracts ebenfalls von der TH Darmstadt entwickelt. Da sie jeweils unter gleichen Voraussetzungen aufgebaut wurden, lassen sich an ihnen die Beziehungen der Größen untereinander verdeutlichen.

Tab. 2 : Input, Wörterbuch-Qualität, Indexierungsqualität.

	FST1	FST2	FST4	PHYS/PILOT
Input  D	5500	11000	22000	392000
V	≈3000	≈3000	≈3000	22683
V <sub>1</sub>	0,317	0,432	0,522	0,441
V <sub>2</sub>	5,0	8,2	14,3	78,0
$\Delta$	0,135	0,125	0,120	0,074
$u \cdot \Delta$	0,265	0,245	0,235	0,145
$u = 1,96, \alpha = 0,95$				
p	0,638	0,622	0,629	0,658
r	0,557	0,623	0,655	0,655

Die Daten für die Wörterbücher FST1, FST2 und FST4 zeigen erwartungsgemäß, daß bei größerem Input die Vollständigkeit wächst. Auch die Genauigkeit vergrößert sich. Dies liegt daran, daß bei kleinen Inputmengen die statistisch ungenauen z-Werte mit kleinen Häufigkeiten  $f(t)$  stärker vertreten sind. Aus einer größeren Vollständigkeit ergibt sich ein höherer r-Wert; einzig für p läßt sich kein einheitlicher Trend feststellen. Tabelle 2 zeigt auch, daß sich funktionale Zusammenhänge zwischen den einzelnen Größen auf Wörterbücher, die unter anderen Bedingungen erstellt wurden, nicht übertragen lassen: PHYS/PILOT wurde für ein sehr viel größeres Fachgebiet erstellt, was sich u.a. an dem umfangreicheren Indexierungsvokabular zeigt. Trotz des vielfachen Inputs gegenüber FST4 sind aber nur 44,1% der Deskriptoren durch die Relation Z erfaßt.

Die Beziehung zwischen Input und Wörterbuch- bzw. Indexierungsqualität läßt sich durch Wachstumfunktionen beschreiben, mit deren Hilfe man Erwartungswerte für die einzelnen Größen angeben kann. Dazu sei auf den Beitrag von Tzeras in diesem Band und /20/ verwiesen.

## 6 Der Übergang zum dynamischen Wörterbuch

Das bisherige Wörterbuch kann als statisches Wörterbuch charakterisiert werden, da es, einmal aufgebaut, nicht mehr geändert wird. Die Nachteile sind evident:

- Die aufgebauten Wörterbücher sind, auch wenn sie mit großem Input und damit hohem Aufwand erstellt werden, notwendigerweise unvollständig und ungenau (siehe Abschnitt 5).
- Die Weiterentwicklung im Anwendungsgebiet, die sich z.B. in Änderungen des Indexierungsvokabulars niederschlägt, werden im Wörterbuch nicht nachvollzogen. So sind seit 1985 etwa 1450 Deskriptoren neu in den Physik-Thesaurus aufgenommen worden. Ebensovienig werden neu gebildete Fachausdrücke in das Wörterbuch eingetragen und bleiben damit, wie die neuen Deskriptoren, bei der Indexierung unberücksichtigt.
- Der prinzipiell mögliche Neuaufbau des Wörterbuches in regelmäßigen Zeitabständen ist nur eine Behelfslösung, da das neue Wörterbuch wiederum unvollständig und ungenau ist und der Aufbau mit hohem Aufwand verbunden ist.

Aus diesem Grund werden beim FIZ z.Zt. Arbeiten durchgeführt, um das bisherige Wörterbuch-Konzept durch ein dynamisches Wörterbuch-Pflege-System zu ersetzen. Dazu soll AIR/PHYS um die drei Module Korrekturauswertung, Schwachstellenanalyse und Wörterbuch-Update erweitert werden.

### a) Korrekturauswertung

Zweck dieses Moduls ist es, die Indexierungskorrekturen als Feedback für das System zu nutzen. Dazu werden für jeden Deskriptor  $s$  aus dem laufenden Input in PHYS folgende Daten erhoben:

COM( $s$ ): Anzahl der Dokumente, denen  $s$  automatisch und manuell zugeteilt wurde (COMmon descriptors),

ADD( $s$ ): Anzahl der Dokumente, denen  $s$  nur manuell zugeteilt wurde (ADDitions),

DEL( $s$ ): Anzahl der Dokumente, denen  $s$  nur automatisch zugeteilt wurde (DELetions).

Daraus lassen sich die Signifikanzangaben  $r(s)$  und  $p(s)$  ableiten:

$$r(s) = \frac{\text{COM}(s) + 0.5}{\text{COM}(s) + \text{ADD}(s) + 1} \quad p(s) = \frac{\text{COM}(s) + 0.5}{\text{COM}(s) + \text{DEL}(s) + 1}$$

Diese Angaben werden in das Wörterbuch aufgenommen und dort ständig aktualisiert. In einer in /18/ beschriebenen Untersuchung wurden für  $r(s)$  und  $p(s)$  20.000 Dokumente ausgewertet, eine Menge, die etwa dem Input von zwei Monaten entspricht.  $p(s)$  und  $r(s)$  wurden als zusätzliche Angaben in die Relevanzbeschreibungen  $x(d,s)$  aufgenommen. Bei der Adaption der Indexierungsfunktion  $a(x(d,s))$  stellte sich heraus, daß  $p(s)$  und  $r(s)$  signifikanter als die meisten der bisher verwendeten Beschreibungselemente  $x_i$  sind. Indexierungstests ergaben eine um 4% verbesserte Trefferquote, die Vollständigkeitsquote erhöhte sich um 2%.

## b) Schwachstellenanalyse

Die Schwachstellenanalyse lokalisiert Lücken und Ungenauigkeiten des Wörterbuches, die das Indexierungsergebnis am meisten negativ beeinflussen. Als Indikator dienen ADD(s) und DEL(s). Für Deskriptoren, deren Werte die vorgegebenen Schwellenwerte übersteigen, werden die wörterbuch-abhängigen Ursachen für die Indexierungsfehler bestimmt, die sich an der Anzahl der Werte  $z(t,s)$  ablesen lassen, an denen der Deskriptor beteiligt ist:

- Von der Relation Z nicht oder ungenügend erfaßte Deskriptoren.  
Neu in den Thesaurus aufgenommene Deskriptoren und in der zur Berechnung der z-Werte verwendeten Dokumentmenge selten zugeteilte Deskriptoren werden von der Relation Z nicht erfaßt. Für diese Deskriptoren können daher nur vergleichsweise schwache, auf der Identitäts- oder USE-Relation beruhende Hinweise gebildet werden, so daß sie von AIR/PHYS nur selten zugeteilt werden und bei der Korrektur häufig hinzugefügt werden müssen. Es muß daher versucht werden, für diese Deskriptoren mit aktuellen Dokumenten z-Werte zu berechnen. Dies gilt auch für Deskriptoren mit nur wenigen z-Werten.
- Deskriptoren mit (über-)durchschnittlich vielen z-Werten.  
Hier werden sich Indexierungsfehler in der Regel in hohen Werten für DEL(s) und einer niedrigen Trefferquote  $p(s)$  bemerkbar machen. Dies kann daran liegen, daß zu diesen Deskriptoren auch viele statistisch unsichere z-Werte im Wörterbuch eingetragen sind. Die z-Werte dieser Deskriptoren müssen daher überprüft werden.

Entsprechend der Art der Indexierungsfehler und der Ursache werden die Deskriptoren im Wörterbuch als Schwachstelle markiert. Die Berechnung oder Überprüfung der Relationen erfolgt dann durch das Wörterbuch-Update.

## c) Wörterbuch-Update

Die wesentlichen Aufgaben des Wörterbuch-Update-Moduls sind:

- Eintragen neuer Deskriptoren in das Wörterbuch.  
Damit gewährleistet wird, daß neue Deskriptoren dem Indexierungssystem sofort bekannt sind, werden sie zusammen mit der Identitäts-Relation in das Wörterbuch eingetragen. Bis zu dem Zeitpunkt, an dem sie ausreichend oft manuell zugeteilt wurden, so daß für sie z-Werte berechnet werden können, werden sie als "neu" markiert. Mittels einer Default-Indexierungsfunktion, die nur auf Relevanzbeschreibungen zu solchen Deskriptoren angewandt wird, läßt sich dafür sorgen, daß sie immer dann zugeteilt werden, wenn für sie auf der Identität beruhende Hinweise vorliegen.
- Berechnen neuer Relationen.  
Für neue Deskriptoren mit ausreichender Zuteilungshäufigkeit und Deskriptoren, die von der Schwachstellenanalyse entsprechend markiert wurden, wird versucht, z-Werte zu bestimmen, die die Bedingungen  $z(t,s) \geq 0.3$  und  $h(t,s) \geq 3$  erfüllen.
- Überprüfen vorhandener Relationen.  
Für die von der Schwachstellenanalyse hierzu bestimmten Deskriptoren müssen die im Wörterbuch vorhandenen Werte  $z(t,s)$  überprüft werden.

Wenn für Deskriptoren, die als Schwachstelle markiert waren, Relationen berechnet werden konnten, so muß die Markierung anschließend zurückgesetzt werden.

Für die Überprüfung und Berechnung der z-Werte ist es wesentlich, daß dazu die aktuellen Dokumente zur Verfügung stehen. Dazu werden invertierte Listen bereitgestellt, die die Terme in Grundform bzw. die Deskriptoren enthalten und die Häufigkeiten  $f(t)$  bzw.  $m(s)$  sowie die Nummern der Dokumente, in denen die Terme vorkommen bzw. denen die Deskriptoren zugeteilt sind. Dafür notwendige Daten fallen bereits zum Teil während der Indexierung an, wenn nach der Zerlegung der Dokumente die Wörter auf ihre Grundform reduziert werden und Mehrwortgruppen bestimmt werden. Bezüglich der invertierten Listen für Deskriptoren ist zu untersuchen, ob dafür nicht die in PHYS ohnehin vorhandenen Dateien verwendet werden können.

Die einzelnen Module müssen nicht explicit von außen gestartet werden, vielmehr wird der neue Input in die Datenbank von der Korrekturauswertung und dem Wörterbuch-Update automatisch ausgewertet. Sind von der Korrekturauswertung Daten in einem bestimmten Umfang erhoben worden, wird die Schwachstellenanalyse aktiviert. Um die im Wörterbuch neu eingetragenen Informationen stets optimal zu nutzen, wird nach umfangreichen Änderungen eine Neu-Adaption der Indexierungsfunktion durchgeführt.

## 7 Ausblick

Die in Abschnitt 6 beschriebenen, vom FIZ geplanten Erweiterungen des Indexierungssystems AIR/PHYS (vgl. /18/) werden das System in einigen wichtigen Punkten verbessern:

- In den Thesaurus neu aufgenommene Deskriptoren werden im Wörterbuch verankert und stehen zur automatischen Indexierung zur Verfügung.
- Da der neue Input in PHYS benutzt wird, um z-Werte zu überprüfen oder zu berechnen, ist der Bestand der Relation Z und des Termvokabulars stets aktuell.
- Durch die deskriptor-orientierte Pflege werden die jeweils größten Schwachstellen des Wörterbuches behandelt. Lücken im Wörterbuch werden beseitigt und unsichere z-Werte durch genauere ersetzt.

Die dadurch steigende Qualität der automatischen Indexierung hat zusätzlichen, positiven Einfluß auf die Konsistenz der korrigierten Indexierung. Nachdem AIR/PHYS schon mehr als 5 Jahre erfolgreich eingesetzt wird, soll das System nun auch für Datenbanken anderer Bereiche, z.B. der Energieforschung und -technologie eingesetzt werden. Bei dem Aufbau des dafür erforderlichen, fachspezifischen Wörterbuches sind die Eigenschaften des dynamischen Wörterbuch-Systems von Anfang an ein Vorteil: Das Wörterbuch kann zunächst mit einem geringerem Input als PHYS/PILOT aufgebaut werden, da das Wörterbuch, während das System schon einsatzfähig ist, dort, wo es am nötigsten ist, automatisch vervollständigt werden kann.



## Literatur

- /1/ Fuhr, N. (1985). *A Probabilistic Model of Dictionary Based Automatic Indexing*. In: Proceedings of the riao 85 ( Recherche d'Information Assistée par Ordinateur), Grenoble, S. 207-16.
- /2/ Fuhr, N. (1986). *Probabilistisches Indexing und Retrieval*. Dissertation, TH Darmstadt, Fachbereich Informatik.
- /3/ Kuhlen, R. (1977). *Experimentelle Morphologie in der Informationswissenschaft*. Verlag Dokumentation, München.
- /4/ Muhler, I.; Repp, J. (1986). *Die Verarbeitung von Formeln in physikalischen Texten*. In: /11/, S. 31-42.
- /5/ Jaene, H.; Seelbach, D. (1975). *Maschinelle Extraktion von zusammengesetzten Ausdrücken aus englischen Fachtexten*. (ZMD-A-29), Beuth Verlag, Berlin, Frankfurt.
- /6/ Kienitz-Vollmer, B.; Reichardt, J. (1986). *Bestimmung von Mehrwortgruppen mit Hilfe des Begrenzerverfahrens*. In: /11/, S. 18-30.
- /7/ Knorz, G. (1983). *Automatisches Indexieren als Erkennen abstrakter Objekte*. Dissertation, TH Darmstadt, Fachbereich Informatik.
- /8/ Schürmann, J. (1977). *Polynomklassifikatoren für die Zeichenerkennung. Ansatz, Adaption, Anwendung*. Oldenbourg Verlag, München, Wien.
- /9/ Knorz, G. (1986). *Die Anwendung von Polynomklassifikatoren für die automatische Indexierung*. In: /11/, S. 98-126.
- /10/ Biebricher, P.; Fuhr, N.; Knorz, G.; Lustig, G.; Schwantner, M. (1988). *Entwicklung und Anwendung des automatischen Indexierungssystems AIR/PHYS*. Nachrichten für Dokumentation 39, (1988), S. 135-43.
- /11/ Lustig, G. (Hrsg.) (1986). *Automatische Indexierung zwischen Forschung und Anwendung*. Georg Olms Verlag, Hildesheim, Zürich, New York.
- /12/ Fangmeyer, H.; Lustig, G. (1969). *The EURATOM Automatic Indexing Project*. In: International Federation for Information Processing (IFIP) (Hrsg.): IFIP Congress 68, Edinburgh. S. 1310-4. North Holland Publishing Company, Amsterdam.
- /13/ Fuhr, N.; Jäger-Beck, R.; Schwantner, M. (1986). *Die Gewinnung von statistischen Relationen zwischen Terms und Deskriptoren*. In: /11/, S. 43-51
- /14/ Biebricher, P.; Fuhr, N.; Knorz, G.; Schwantner, M. (1986). *Indexierungswörterbücher für das Fachgebiet Physik*. In: /11/, S. 52-72.
- /15/ Schuckmann, A.; Süner, W. (1987). *Dokumentation des Systems ZWERG*. Studienarbeit, TH Darmstadt, Fachbereich Informatik.
- /16/ PIL (1987). *Projekt PILOTANWENDUNG AIR/PHYS (Pilotanwendung von automatischen Indexing- und verbesserten Retrievalverfahren mit der Datenbasis PHYS)*. Abschlußbericht. TH Darmstadt, Fachbereich Informatik, Fachgebiet Datenverwaltungssysteme II.
- /17/ Lück, W. (1988). *Erfahrungen mit dem automatischen Indexierungssystem AIR/PHYS*. In: Deutsche Gesellschaft für Dokumentation e.V. (DGD) (Hrsg.): Deutscher Dokumentartag 1987, S. 340-52. VCH Verlagsgesellschaft, Weinheim.

- /18/ **Schwantner, M. (1991).** *Aufbau und Pflege eines Wörterbuches für die automatische Indexierung.* Dissertation, TH Darmstadt, Fachbereich Informatik (Erscheinen angekündigt).
- /19/ **Hüther, H. (1990).** *On the Interrelationship of Dictionary Size and Completeness.* In: Vidick, J.-L. (Hrsg.): Proceedings of the 13th International Conference on Research and Development in Information Retrieval, S. 313-26. ACM Press, New York.
- /20/ **Hüther, H. (1989).** *Wachstumsfunktionen in der automatischen Indexierung.* Dissertation, TH Darmstadt, Fachbereich Informatik.

Anschrift des Autors:

Dipl.-Inform. Michael Schwantner  
Fachinformationszentrum Karlsruhe  
Abt. STABEND  
7514 Eggenstein-Leopoldshafen 2

## **Moderne Technik und betriebliche Realität im Informationsmanagement am Beispiel eines Automobilkonzerns**

Schael, Fritz

### Kurzfassung

High-Tech-Systeme sind vielversprechende Medien für das Informationsmanagement. An zwei Beispielen aus der Praxis eines Automobilkonzerns wird jedoch gezeigt, daß im Einzelfall wesentliche Tätigkeiten bzw Funktionen durch solche modernen Techniken nicht optimal abzudecken sind. Es handelt sich um die Arbeit mit dem Patentarchiv und den Einsatz von Mikroformen in der Bibliothek.

Ein Beispiel für die Aufgeschlossenheit gegenüber modernen Systemen ist das Bibliotheksverwaltungssystem, das konzernweit, d.h. weltweit, eingesetzt wird.

### Abstract

High-tech information technology applies for advantageous application in information management. Yet two examples from an automotive company show that not in any case these technologies can serve for optimizing work and efficiency: working with the patent documentation and the copy shop of the library.

Nevertheless, Volkswagen is implementing an advanced library management system for corporate wide use, that means for use at Volkswagen companies all over the world.

Der Siegeszug der elektronischen Speichermedien im Informationsmanagement scheint unaufhaltsam - zumindest wenn man den Experten auf diesem Gebiet glauben darf, seien es die Produzenten solcher Systeme oder die Theoretiker auf diesem Gebiet.

Wie aber bewähren sich diese Systeme in der Praxis der betrieblichen Informationsvermittlung? Sind sie für den einzelnen Anwendungsfall besser und schneller als einfachere, weniger mit elek-

tronischen Komponenten ausgestattete, konventionelle Systeme? Und - letzter aber wichtigster Aspekt - erhöhen sie die Effizienz der Mitarbeiter? In betriebswirtschaftlichen Größen ausgedrückt: wie hoch ist das "Return on Investment"? In der Industrie ist der letzte Punkt der entscheidende. An ihm wird jedes neue System, d.h. jede Investition auch im Informationsmanagement gemessen.

Bei Volkswagen existieren zwei große Literatursammlungen, die für die modernen Massenspeicher prädestiniert scheinen: das Patentarchiv und die Literatursammlung zur Literaturdatenbank (LIDAS intern; "VWWW" bei FIZ-Technik und DATA-STAR). Beide unterschiedlich in der äußeren Form der Literatur und beide mit spezifischen Anforderungen an die Arbeit mit dem Archiv.

Da es naheliegend ist, die umfangreichen Sammlungen - auch aus Gründen des Platzbedarfes - auf Mikroformen zu bringen, stehen wir seit vielen Jahren in Kontakt mit Systemanbietern verschiedener Art. Dabei ergab sich sehr schnell, daß für eine Fortentwicklung unserer Speichersysteme und der Arbeitsorganisation als Speichermedium nur die optische Speicherplatte in Frage kommt. In der Abteilung Fachinformation und -bibliothek bei Volkswagen muß sie jedoch mit den eingeführten Systemen auf Basis der Mikrofilmkarte konkurrieren.

#### Optische Speicherplatte und Patentarchiv

Das Patentarchiv basiert auf den europäischen und deutschen Patentschriften in Mikrofilmform, bis heute 380.000 Schriften mit 625.000 Karten. (Die Mikrofilmkarten enthalten jeweils 8 DIN A4-Seiten). Hinzu kommen rd. 100.000 US-Patentschriften, die im Original vorliegen, mit ca. 1 Mio Seiten.

In seiner Hauptfunktion dient das Patentarchiv als Dokumentationsbasis für die Patentingenieure. Es ist daher aufgebaut nach der "Internationalen Patentklassifikation" ("IPC"), d.h. nach sachlichen Kriterien, und untergebracht in einer

Paternosterregistratur. Die Feinaufgliederung des Archivs ist je nach Sachgebiet und der Häufigkeit von Anmeldungen unterschiedlich und ausgerichtet am Interessenspektrum von Volkswagen. Das Prinzip besteht darin, zu große Kartens Stapel in einer Einheit zu vermeiden.

Die Arbeitsweise der Patentingenieure ist dabei wie folgt: Es liegt eine Patentanmeldung, eine interne Erfindungsmeldung, eine Anfrage zum Stand der Technik zu einem Bauteil o.ä. vor. Die Aufgabe besteht nun darin, mit geringstem Aufwand von Zeit und geistiger Anstrengung relevante Patentschriften herauszufinden. Die "IPC" bietet dazu das am besten geeignete Zugriffskriterium. Die Mikrofilmkarten der passenden Klasse sind schnell aus der Paternosterregistratur herausgeholt und in ein speziell für Patentschriften konstruiertes Lesegerät eingelegt. Dieses Gerät zeigt 4 Seiten auf einem Bildschirm an; mittels eines Hebels können die übrigen 4 Seiten angezeigt werden. (Es existieren andere Geräte, die alle 8 Seiten gleichzeitig zeigen, darunter leiden jedoch die Projektionsqualität und die Lesbarkeit erheblich). Der Kartenwechsel erfolgt über einen fußbetätigten Druckschalter.

Die wichtigste und erste Informationsquelle sind Zeichnungen, die zu jedem Patent gehören. Benennungen bzw. Titel sind meistens nichtssagend ("Fahrzeugsbremse", "Vorrichtung zum Halten von Gegenständen", "Schaltungsanordnung für KFZ", "Preßpapiermaschine" usw.).

Der Patentingenieur läßt nun in rascher Folge die Mikrofilmkarten durch das Gerät laufen, er erkennt anhand der Zeichnung, ob die jeweilige Patentschrift einen Bezug zur Aufgabenstellung hat. Erst danach liest er gegebenenfalls das Abstract oder - wenn erforderlich - Patentansprüche und Text. Durch die Anzeige von mindestens 4 Seiten der Patentschrift und dem einfachen Umschwenken auf weitere 4 Seiten ist ein laufender Vergleich von Zeichnung, Beschreibung

und Patentansprüchen untereinander in einfacher und rationaler Weise möglich.

Bei diesem Verfahren können in sehr kurzer Zeit größere Mengen von Patentschriften durchgesehen und bewertet werden. Das Sortierkriterium "IPC" grenzt das Sachgebiet sehr gut ab, und in vielen Fällen reicht das Durchsehen einer Rubrik aus, um alle relevanten Patentschriften zu erfassen. Nur in wenigen Fällen müssen alternative Facetten oder Nachbargebiete zusätzlich recherchiert werden (z.B. "Bremsen" und "Hydraulik"). Diese Arbeitsweise erlaubt außerdem, einzelne Schriften herauszunehmen und in einem zweiten Schritt eine genauere Analyse vorzunehmen.

Alle Analysen von Archivsystemen auf Basis von optischen Speicherplatten, rechnerunterstütztem Mikrofiche-Archiv u.ä. ließen

- a) keine Effizienzsteigerung für die Patentingenieure erkennen, im Gegenteil, im Einzelfall war eine höhere Bearbeitungszeit zu erwarten.
- b) fanden keine Akzeptanz bei den Betroffenen.
- c) waren gegenüber den Lesegeräten wenig flexibel für die Bewertung, z.B. für den Vergleich von Text mit Zeichnung, Patentansprüche mit Zeichnung u.a.m., oder für das gezielte Vor- und Rückblättern.

Der Rationalisierungseffekt bei der Erstellung von knapp 4.000 Kopien pro Jahr ergab keine Wirtschaftlichkeit für die erforderliche hohe Investitionssumme: Scanner, Lesegerät, Drucker, Plattenwechsler, Server usw.

Allein die Anforderung, alle Patentschriften zu einer Klasse, Unterklasse oder Gruppe - ggf. auch noch Untergruppe - zu selektieren, bedeutet bei einem System auf Basis der optischen Speicherplatte,

entweder alle Platten "abzufragen", das Ergebnis zwischen-

zuspeichern und zu sortieren,  
oder die Recherche entsprechend der Zahl der chronologisch  
aufeinander folgenden Platten zu wiederholen,  
oder den Gesamtbestand der Speicherplatten jährlich neu sor-  
tiert auszutauschen.

### Bildplatte und Literaturarchiv

Das Literaturarchiv enthält rd. 95.000 Dokumente mit ca. 800.000  
Seiten, jede Mikrofilmkarte trägt 4 DIN A4-Seiten, somit besteht  
dieses Archiv aus rd. 200.000 Mikrofilmkarten. Der jährliche Zu-  
wachs beträgt rd. 5.000 Dokumente mit ca. 45.000 Seiten. Das Lite-  
raturarchiv ist ebenfalls in einer Paternosterregistratur unter-  
gebracht. Die Rückvergrößerungen erfolgen auf A2-Formaten und  
werden dem Kunden gerollt oder gefaltet zugestellt.

Die optische Speicherplatte bietet sich an, weil

- die Identifizierung der Dokumente nach einer Bestellnummer er-  
folgt, die chronologisch orientiert ist.
- der Hauptteil der Bestellungen in den ersten 2 Jahren nach Auf-  
nahme ins System eingehen.
- die Ausgabe im DIN A4-Format erfolgt (die Rückvergrößerungen von  
der Mikrofilmkarte sind wie beschrieben DIN A2 und damit nicht  
sehr handlich).

Das erste Problem stellt sich hier bereits bei der Aufnahme der  
Dokumente aus Zeitungen, Zeitschriften verschiedener Länder, Ta-  
gungsbänden u.a.m.

Sie sind sehr heterogen

- a) im Format
- b) in der Art der Vorlage (Grafiken und Bilder, schwarz-weiß und  
bunt)

und erfordern daher gegenüber dem bisherigen Verfahren der Mikro-  
verfilmung teilweise eine zusätzliche und aufwendige Vor- und/  
oder Nachbehandlung:

- Anfertigen von Zwischenkopien von Hand
- Einsatz von Rastern mit Prüfung

- elektronische Korrekturen des Formates.

Die bei Bildern u.ä. erforderliche höhere Punktdichte (dpi) verringert die Speicherkapazität der Bildplatte bezogen auf die Seitenzahl erheblich. Alle Berechnungen ergaben den Bedarf von einer Platte pro Jahrgang. Keines der betrachteten Systeme lieferte Ausdrücke, die eine der Mikrofilmkarte gleichwertige Qualität der farbigen Grafiken, Bilder usw. aufwiesen. Damit war die Akzeptanz beim Kunden fraglich. Sie ist schon beim heute verwendeten Mikrofilmsystem in Einzelfällen nicht mehr gegeben.

Das zweite Problem stellt die Vervielfältigung dieses Archivs für die Verteilung im Konzern dar. Heute erhalten die Informations-Vermittlungsstellen (IVS) einiger Konzerngesellschaften (z.B. AUDI, SEAT, Autolatina) Kopien aller neuen Mikrofilmkarten, die automatisch erstellt und sortiert werden. Die vor Ort anfallenden Literaturanforderungen erledigen die Stellen direkt. Eine Umstellung auf ein Bildplattensystem erfordert an jeder Stelle eine nahezu vollständige Parallelinvestition! Der andere Weg - Abrufen der Dokumente über eine Standleitung von der Zentrale in Wolfsburg - reduziert zwar den Aufwand auf Bildschirm und Drucker, belastet aber die EDV-Verbindungen derart, daß die zuständigen Systemstellen dies ablehnen, bzw. dieser Anwendung eine sehr niedrige Priorität einräumen, was die Verbindung nahezu illusorisch macht.

Bleibt als letztes eine Variante des "Online-Ordering": die einzelne Stelle bestellt Kopien online im System der Zentrale; sie werden über Nacht ausgedruckt und über die normale Hauspost ausgeliefert. Allein der Zeitverzug ergibt eine deutliche Verschlechterung der Dienstleistung der externen IVS und zusätzlichen Aufwand in der Zentrale.

Das dritte Problem resultiert aus der Struktur und Arbeitsweise der zentralen IVS. Ein Bildplattensystem benötigt nur die Funktionen "Archivieren" und "Drucken", diese aber in perfekter Form.



Im Normalfall ist keine Anzeige der zu druckenden Seiten erforderlich. Nach Eingabe der Bestellnummer kann das Ausdrucken sofort beginnen. Recherchefunktionen sind nicht gefragt. Fachrecherchen erfolgen in der eigenen oder in externen Literaturdatenbanken. Es fallen pro Jahr rd. 50.000 Kopien mit rd. 400.000 Seiten DIN A4 an (entsprechend 100.000 Mikrofilmkarten). Dafür steht eine halbe Arbeitskraft zur Verfügung. Sie sucht heute die angeforderten Dokumente aus der Paternoster-Registrieratur heraus und gibt die Mikrofilmkarten in Stapeln in das danebenstehende Rückvergrößerungsgerät. Dies benötigt pro Blatt ca. 11 Sekunden (Stapelverarbeitung). Entsprechend müßte ein System mit DIN A4-Ausgabe unter 3 Sekunden pro Blatt liegen (ohne Pause zwischen zwei Dokumenten).

Insgesamt gesehen stehen dem Vorteil einer Qualitätssteigerung - erhöhter Komfort für den Kunden durch das Format A4 gegenüber z.Z. A2 - die hohen Kosten einer solchen Anlage gegenüber, sowie Qualitätseinbußen bei Bildern und einigen Schrifttypen. Die Wirtschaftlichkeit einer solchen Investition ist nicht nachweisbar. Die Akzeptanz bei den Mitarbeitern bereitet in diesem Fall keine Probleme, da in gewissem Rahmen eine Arbeitserleichterung eintritt, z.B. kein Falten oder Rollen der DIN A2-Blätter.

### Zusammenfassung

Moderne Technologien in Fachinformation und Bibliothek bedeuten nicht a priori eine Steigerung von Effizienz und Qualität der Dienstleistung. Eine eingehende Prüfung unter Einbeziehung der Arbeitsabläufe und der Akzeptanz durch die Mitarbeiter ist unumgänglich. Dabei kann es sich - wie gezeigt - herausstellen, daß die derzeitigen Verfahren ein Optimum darstellen, und durch die am Markt vorhandene moderne Technologie keine Verbesserung erreichbar ist. Die Analyse bekannter Anwendungen zeigt in jedem Falle, daß die gegebenen Arbeitsweisen, Dokumentarten und Randbedingungen für den jeweiligen Einsatz spezifisch und nicht übertragbar sind.

Eine Voreingenommenheit modernen Technologien gegenüber ist aber keinesfalls gegeben. Neben unserer Literaturdatenbank LIDAS, die unter STAIRS läuft und seit 1971 existiert, bauen wir z.Z. ein sehr umfangreiches und komplexes Bibliotheksverwaltungssystem auf, in das wir im Endausbau unsere derzeitige Literaturdatenbank integrieren können. Damit hätten wir dann Bibliotheksverwaltung, Erwerbung, Ausleihe und Literaturdatenbank in einem System. Das Ergebnis wird sein: Qualitätsverbesserung der Dienstleistung, Vereinfachung von Abläufen, Wirtschaftlichkeit u.a. durch Entfall eines Systems. Die Akzeptanz bei den Mitarbeitern ist bereits jetzt gegeben.

Dipl.-Ing. Fritz Schael  
Volkswagen AG  
E/Fachinformation und -bibliothek  
Postfach

W-3180 Wolfsburg 1

Hillebrand, Volker

Ausgehend vom teilmechanisierten Informationssystem wird die Entwicklung der Informationsarbeit sowie die Stellung des Fachinformationssystems Dieselmotoren im Unternehmen aufgezeigt. Wesentliche Etappen waren 1985 das automatisierte Informations-Recherche-System unter Anwendung des PS AIDOS OS/ES in Kopplung mit dem PS AIDOS FORTE für die internationale Zusammenarbeit und der Einsatz moderner Bürocomputer im Jahre 1987. Die eingesetzte Software für den Informationsverarbeitungsprozeß war und ist MIDOS.

Gegenwärtig wird das Fachinformationssystem zur Informationsvermittlungsstelle umgestaltet. Als neue Dienstleistungen werden den Nutzern Online-Recherchen sowie CD-ROM-Recherchen angeboten. Erste Erfahrungen im Umfeld dazu. Neue zukünftige Aufgaben werden durch die Einkopplung der Informationsvermittlungsstelle in das Management des Unternehmens im Aufbau eines gesamten Informations- und Organisationsprozesses gesehen.

Starting from the mechanical information system it shall be shown the development of the information service included the position of the technical information system "Diesel engine" in the enterprise.

There were main steps in the year 1985 with the automated information searching system using the PS AIDOS OS/ES in connection with the PS AIDOS FORTE for the international cooperation and in the year 1987 the application of modern personal computers.

MIDOS it was and is the applicated software of the information processing.

At present the technical information system shall be transformed into information transfer agency, offering the new services online-information searching and CD-ROM-searching for users.

Showing first experiences and new tasks at the transformation, especially with regard to the connection of the informa-

tion transfer agency with the management of the enterprise, it will be represented the rebuilding of a completed information and organisation system.

## 0. Einleitung

Daß der Fachinformation auf dem Gebiet des Dieselmotorenbaues, insbesondere in der SKL Engineering Roßlau GmbH, schon immer eine große Bedeutung beigemessen wurde, verdeutlicht die Entwicklung des Informationssystems und zeigt sich im besonderen in der frühzeitigen Verfügbarkeit der wissenschaftlich-technischen Ressourcen in Form eigener Datenbanken und einer integrierten Fachbibliothek.

Der Gesamtkomplex der wissenschaftlich-technischen Information wurde frühzeitig in seiner Funktion als Bestandteil des technologischen Prozesses der Forschung und Entwicklung und im Ansatz in seiner Planung und Leitung gesehen.

Die Leistungsfähigkeit des Fachinformationssystems einschließlich des Patentinformationssystems hatte, bedingt durch die Eigenverfügbarkeit umfassender externer und interner Forschungs- und Entwicklungserkenntnisse, sowohl für die Entscheidungen zu Grundrichtungen und Schwerpunkten der wissenschaftlich-technischen Arbeit anfallenden Leitungsebenen als auch bei jeder einzelnen F/E-Aufgabe außerordentliche Bedeutung.

Unter den mit dem wirtschaftlichen Umbruch verbundenen Veränderungen stellten sich jedoch folgende Fragen:

1. Unterstützen die bisher erbrachten Informationsleistungen den dringend notwendigen wirtschaftlichen Aufbau und den Erfolg des Unternehmens?
2. Ist, aufbauend auf der vorhandenen Basis, der reibungslose Übergang zur Anwendung moderner Informations- und Kommunikationstechnologien gegeben?
3. Ist eine Integration des Informationssystems in die bundesdeutsche Fachinformationslandschaft möglich?

Die Beantwortung dieser Fragen ist Gegenstand nachfolgender Betrachtungen.

## 1. Entwicklung des Fachinformationssystems

### 1.1 Basis

#### - Informationsbedarf

Entsprechend der damaligen Informationsverantwortlichkeit des Fachinformationssystems auf dem Gebiet des Dieselmotorenbaues setzte sich der Informationsbedarf aus den Forderungen

- . der Betriebe des Kombinates SKL,
  - . der Betriebe außerhalb des Kombinates SKL,
  - . der Hoch- und Fachschulen sowie
  - . der interessierenden Nutzer innerhalb des RGW
- zusammen.

#### - Informationsquellenbasis

Die wesentlichen Literaturkategorien des Informationsaufkommens im Fachgebiet Dieselmotoren waren:

- . Fachzeitschriften,
- . Firmenschriften,
- . Interna,
- . Tagungsmaterial,
- . F/E-Berichte,
- . Patente.

#### - Arbeits- und Ordnungsmittel

Die Arbeitsorganisation im Fachinformationssystem wird durch folgende Arbeitsmittel unterstützt:

- . Informationsthemenplan,
- . Zeitschriftenkatalog,
- . Richtlinie für die inhaltliche Erschließung und Erfassung der Dokumente.

Zur inhaltlichen Erschließung der Dokumente liegt der zweisprachige Informations-Recherche-Thesaurus vor.

## 1.2 Etappen des Aufbaues

1972 wurden mit dem Aufbau eines teilmechanisierten Informationssystems die Grundvoraussetzungen für den Auf- und Ausbau zum Fachinformationssystem geschaffen, die eine ständige Weiterentwicklung in technischer und organisatorischer Hinsicht entsprechend den Anforderungen der Forschung und Entwicklung sowie der Planung und Leitung ermöglichten.

Die Forderung zur Steigerung der Qualität und Quantität der Informationsversorgung mit wissenschaftlich-technischen Informationen, die Notwendigkeit zur Rationalisierung der Prozeßabläufe in der Informationsbe- und -verarbeitung und die Forderung nach Vertiefung der Zusammenarbeit mit den damaligen Mitgliedsländern des RGW, insbesondere mit der UdSSR, waren 1982 die maßgeblichen Gründe für den Aufbau einer Datenbank Dieselmotoren auf einem Großrechner unter Anwendung des PS AIDOS OS/ES und des Programmpaketes AIDOS FORTE.

1985 erfolgte die Zusammenfassung und Integration in den Fachabteilungen vorhandener, fachlich stark spezialisierter Datensammlungen zu

- . erzeugnispezifischen Daten der Dieselmotoren,
- . statistischen Daten der Auswertung von Störungsmeldungen für Dieselmotoren,
- . spezifischen Daten des Planungsprozesses

zu einer einheitlichen computergestützten Datenbank unter Anwendung des PS REDABAS/REDAFAKT in das Fachinformationssystem.

Die Realisierung eines langfristig tragfähigen und zugleich praxisnahen computergestützten Fachinformationssystems erfolgte 1986 unter Nutzung des vom EAW Teltow entwickelten PS MIDOS 8 und später MIDOS 16.

## 1.3 Organisation

### - Informationserschließung und -erfassung

Auf der Grundlage des Speicherprofils erfolgte die Informationserschließung und -erfassung.

Eine Schlüsselposition nahm seit 1972 bei der Erschließung und Erfassung der Informationsquellen das Indexieren der

Fachinhaltskomponenten nach dem Fachthesaurus Dieselmotoren ein, da von der Genauigkeit des Indexierens das Wiederauffinden der Informationen im Rechercheprozeß abhängt.

#### - Recherchedurchführung

Die Abarbeitung der Recherchefragen auf dem Großrechner erfolgte bei retrospektiven Recherchen in der Regel in einer Frist von 14 Tagen und bei der selektiven Informationsverbreitung im Rahmen des zweimonatlichen rechentechnischen Abarbeitungszyklus. Mit der Einführung des PS MIDOS 16 und der damit zeitlich beeinflussbaren Stapelabarbeitung der Nutzerprofile und der Dialogrecherche wurde die für den Informationsnutzer vorteilhafte, effektive und damit kostengünstige Form gefunden.

#### 1.4 Informationsleistungen

Entsprechend dem Bedürfnis der Nutzer wurden folgende Informationsleistungen erbracht:

- . selektive Informationsverbreitung,
- . retrospektive Recherche und
- . synthetische Informationsmittel.

Zur Befriedigung des Informationsbedarfs bei Problemen aus Randgebieten des Informationsprofils oder zu Querschnittsfragen bestanden Kooperationsbeziehungen zu anderen Informationsfonds bzw. Fachbibliotheken. Sämtliche nachgewiesenen Informationsquellen sind durch bibliographische Angaben, Kurzreferate, MF-Nr. und/oder Standortvermerk gekennzeichnet, die die Beschaffung der vollständigen Quelle zur persönlichen Informationsverarbeitung garantieren.

#### 1.5 Internationale Beziehungen

Für die internationalen Beziehungen des Kombinates bezüglich der Informationen aus Wissenschaft und Technik der spezifischen Belange des Dieselmotorenbaues war das Fachinformationssystem verantwortlich.

Hier waren es besonders die Informationseinrichtungen der Zweckeinrichtung Forschung - Fachgebiet Dieselmotorenforschung - der CSSR, VR Polen und der UdSSR, zu denen enge Kon-

takte bestanden.

Der Aufbau einer internationalen Datenbasis DDR/UdSSR konnte unter Nutzung einheitlicher Rechentechnik (ESER) und problemorientierter Systemunterlagen 1985 so weit vorangetrieben werden, daß der Magnetbandaustausch in der Probephase betrieben wurde.

Auf Grund der beim sowjetischen Partner auftretenden Unsicherheiten hinsichtlich der qualitativen und quantitativen Weiterführung des Informationsfonds und der Erfüllung der für den Informationsaustausch vereinbarten inhaltlichen und wirtschaftlichen Bedingungen wurde das Projekt nicht weitergeführt.

Als wichtigstes Ergebnis dieser Zusammenarbeit ist der zweisprachige IRT Dieselmotoren zu nennen.

## 2. Erfordernisse für die Weiterentwicklung

Im Ergebnis einer Analyse nach dem wirtschaftlichen und politischen Umbruch mußte festgestellt werden, daß unter den neuen Bedingungen

- der Überblick über den nationalen und internationalen edvgestützten Informations-Dienstleistungsmarkt fehlt,
- das Unternehmen nicht nur wissenschaftlich-technische Fachinformationen benötigt, sondern die Wirtschaftsdaten an Bedeutung gewinnen,
- die Bereitstellung umfassender, transparenter und ständig aktualisierter Informationen aus dem Gesamtinformations- und Organisationsprozeß des Unternehmens für die Geschäftsführung gesichert sein muß,
- durch den Aufbau der computergestützten Datenbank Dieselmotoren der Übergang zu den modernen Informationstechnologien zwar geschafft wurde, jedoch die erforderliche Informations- und Kommunikationstechnik für den Onliner-Betrieb noch aufgebaut werden muß,
- die Einführung der erforderlichen neuen Hard- und Software ohne unnötige organisatorische und technische Schnittstellen in das Gesamtkonzept einer Informationsvermitt-



lungsstelle (IVS) erfolgen kann,

- die im Fachinformationssystem existierenden Ressourcen in Form der Datenbank Dieselmotoren zur Erweiterung eines gesamtdeutschen Fonds beitragen können.

### 3. Einkopplung in das Management

#### 3.1 Organisatorische Veränderungen

Da die Information im Unternehmen als wichtiger Produktionsfaktor gesehen wurde, erfolgte die Integration des Kaderpotentials des Fachinformationssystems zur Absicherung eines erforderlichen Gesamtinformations- und Organisationsprozesses im Unternehmen ins Management.

Dabei wird davon ausgegangen, daß die Information, Kommunikation und Organisation als einheitlicher Prozeß für den Erhalt der Wettbewerbsfähigkeit des Unternehmens unabdingbar ist und daß die dadurch gegebene strategische Ausrichtung des Gesamtsystems, orientiert an den Gesamtanforderungen des Managements, schneller erreicht werden kann.

#### 3.2 Anforderungen an die künftige Arbeit

Zielvorgabe ist, Möglichkeiten und Wege zur wirtschaftlichen Erzeugung betriebsinterner und -externer Informationen zu erschließen und diese im Interesse einer Effektivitätssteigerung voll auszuschöpfen.

Das setzt voraus:

1. Ausbau des Fachinformationssystems zu einer IVS durch den Aufbau eines Online-Anschlusses zur Absicherung eines gezielten und schnellen Zugriffs auf die auf dem Informationsmarkt weltweit angebotenen Datenbanken.
2. Aufbau eines operativen und strategischen Informationsmanagements als Führungsinstrumente der Unternehmensleitung durch schrittweise Einführung durchgängig rechnergestützter Arbeitsweisen im gesamten Informations- und Kommunikationsprozeß des Unternehmens im Interesse einer differenzierteren Aufbereitung und Darbietung der Information.

Das wiederum erfordert auf die ganzheitliche Aufgabenabwicklung bezogen die Beschaffung, Erfassung, Verarbeitung und Aussteuerung wissenschaftlich-technischer, operativer und strategischer Informationen und den organisatorischen Aufbau von wettbewerbsorientierten Ablaufstrukturen und Koordination der Informationsflüsse zu einem selbständigen effizienten Kommunikationssystem.

### 3.3 Aufbau unter Einbeziehung des Förderprojektes

Der schnelle Vorstoß wurde durch das BMWI im Rahmen des Förderprojektes "Aufbau einer Informationsvermittlungsstelle" zur Versorgung des eigenen Unternehmens und der Unternehmen in der Region ermöglicht.

Die finanziellen Zuschüsse gestatteten es, die notwendige technische Ausstattung zum Betrieb einer dialogorientierten Online-Recherchestation in Form von effektiver Hardware, ausgestattet mit geeigneter Kommunikationssoftware für den Zugang zu den Hosts sowie Weiterverarbeitungssoftware, zur Aufbereitung der Online-Rechercheergebnisse, zu installieren.

Mit Einführung der Hard- und Software und ihrer effizienten Anwendung war neben einer Stärkung der Selbständigkeit der IVS auch die Möglichkeit zum Aufbau eines strategischen Führungsinstrumentes für eine erfolgreiche Tätigkeit des Unternehmens mit wirtschaftlich vertretbarem Anpassungsaufwand gegeben.

Die Nutzungsmöglichkeiten der installierten Hard- und Software lassen die Selektion und Integration großer externer und interner Informationsmengen im Gesamtinformationsprozeß des Unternehmens zu. Der Aufbau eines auf die Größe des Unternehmens bezogenen Marketing-, Planungs-, Steuerungs- und Controlling-Systems ist somit gegeben.

Der Übergang zur Nutzung der modernen Informations- und Kommunikationstechnologien wurde mit Unterstützung der Betreuer im Rahmen des Förderprojektes kurzfristig bewältigt. Schwerpunkte der Schulung waren das Ausloten der Möglichkeiten und Grenzen der Nutzung der verfügbaren Datenbanken unter Beach-

tung der organisatorischen, ökonomischen und rechtlichen Vielfalt.

### 3.4 Erste Erfahrungen beim Aufbau der Informationsvermittlungsstelle

Die umfassende Nutzung des Online-Informationsmarktes zeigt eindeutig, daß es in der Regel mit der Übergabe der Literaturnachweise in Form von Computerausdrucken nicht getan ist.

Lag es früher weitestgehend im Ermessen des Fachinformationssystems, wie und in welcher Form die Information an den Nutzer weitergegeben wurde, so wird heute durch den Nutzer zum einen die schnelle und rationelle Bereitstellung von Primärliteratur für die Vorbereitung und Durchführung einer Entwicklungs- und/oder Beratungsaufgabe gesehen und zum anderen eine meist differenziertere Bearbeitung der Anfragen und eine gezielte Verdichtung und Ergänzung bis hin zur Erarbeitung von Entscheidungsvorlagen entsprechend dem Informationsbedarf gefordert.

Die Praxis zeigt auch, daß die Integration des Nutzers in die Recherchetätigkeit durch Darlegung der gegebenen Möglichkeiten und Erläuterung der vorhandenen Ordnungsmittel bis hin zur Online-Sitzung diesen stark motiviert und sich auf die Qualität und Quantität des Rechercheergebnisses auswirkt.

Dabei wird vorausgesetzt, daß der Informationsvermittler

- . sich im Informationsangebot im Detail auskennt,
- . im Umgang mit den Informations- und Kommunikationstechnologien sowie mit Bibliotheken und Literaturquellen eine gewisse Routine besitzt,
- . Kenntnisse über das jeweilige Kosten/Nutzen-Verhältnis beim Zugriff auf die jeweilige Datenbank hat,
- . den konkreten Informationsbedarf sowohl des Managements als auch der Fachleute kennt bzw. ermittelt sowie
- . Engagement und Einfühlungsvermögen zeigt.

Zur Verkaufsförderung der Informationsleistungen außerhalb des Unternehmens wurde ein Marketingkonzept/eine Verkaufsstrategie erarbeitet. Weiterhin unterstützt das Unternehmen die

IVS, in dem sie die Informationsleistungen als Marketinginstrument einsetzt, um so den Auftraggeber von der Qualität seiner Sach- und Fachkompetenz mit zu überzeugen.

### 3.5 Erste Vorstellungen zum Aufbau eines Führungsinstrumentes

Der Aufbau des Führungsinstrumentes ist z. Z. in der Entstehung. Es bedarf noch weiterer genauerer Analysen der Aufgabeninhalte innerhalb des Unternehmens und ihrer informationellen Beziehungen zueinander.

Vorgesehen sind

- . der Aufbau eines klar strukturierten und aussagefähigen flexiblen Systems,
- . die weitgehende Vereinheitlichung der Vorgehensweisen im Interesse schneller Ablaufzyklen,
- . die Definition einheitlicher Arbeits- und Ordnungsmittel und Absicherung ihrer Anwendung.

Im Ergebnis muß der Zugriff auf verdichtete Daten und ihre grafische Darstellung in Form von Tabellen und Diagrammen zur besseren Interpretation der Daten, zur frühzeitigen Information über die Lage und Entwicklung des Unternehmens, z. B. Liquiditätsabsicherung, Umsatz/Rentabilität, garantiert werden.

## 4. Zusammenfassung

Die Leistungsfähigkeit des vorhandenen Fachinformationssystems ist die solide Basis für eine effektive Lösung der neuen Aufgaben.

Durch die Anwendung der neuen Informations- und Kommunikationstechnologien und den damit gegebenen Zugang zu den Informationsressourcen der Welt kann sowohl den neuen Anforderungen im Unternehmen als auch denen der Klein- und Mittelbetriebe entsprochen werden.

Die Einkopplung des Kaderpotentials ins Management eines Mittelbetriebes zum Aufbau und zur Wahrung einer komplexen Informationsstruktur stellt eine Herausforderung für die Mitarbeiter dar.

## Veränderung des Aufgabenprofils einer Informationseinrichtung in einem Unternehmen des Schwermaschinenbaus

Hans Kruske

Die Umwandlung des ehemaligen Schwermaschinenbau-Kombinates SKET Magdeburg in die SKET Schwermaschinenbau Magdeburg GmbH brachte auch der Informationseinrichtung grundlegende Veränderungen in der Struktur und in ihren Aufgaben. Aus ihrer unmittelbaren Verbindung zu den F/E-Abteilungen herausgelöst, ist sie einem Zentralbereich zugeordnet, der den Unternehmensbereichen, in der die Fertigung erfolgt, vorgelagert ist. Dadurch wurde der langjährige Informationsfluß, der vorgeschriebene Planungsablauf und die Rückkopplung vom Informationsnutzer verändert oder unterbrochen. Wurde z. B. der Informationsbedarf über viele Jahre weitgehend durch die F/E-Aufgaben langfristig im Informationsthemenplan erfaßt, so sind heute die Aufträge in der Regel in wenigen Stunden oder Tagen auszuführen.

Waren die Informations-Ingenieure bisher für festumrissene Fachgebiete verantwortlich, so strömt jetzt ein Potpourri von Anfragen und Aufgaben auf jeden einzelnen ein.

Erwartete man früher von den Mitarbeitern der Informationseinrichtung vor allem technische Auskünfte, so nehmen heute Informationen zum Einkauf, Absatz, Markt und Management den ersten Platz ein.

Wurden Kosten für die Informationsermittlung und Dokumentenbeschaffung, sofern sie in DDR-Mark zu zahlen waren, akzeptiert, müssen diese jetzt vor jeder Auftragsbearbeitung möglichst genau kalkuliert werden. Diese Aufgabenverschiebung ergibt sich aber nur zum Teil aus den neuen Wirtschaftsbedingungen.

Denn schon vor der Wende war die Informationsstelle Anlaufpunkt für Aufgaben und Anfragen, für die sich im Kombinat niemand kompetent fühlte. So bestand schon seit vielen Jahren die Möglichkeit, Firmenschriften aus aller Welt zu beschaffen. Gelegentlich wurden sogar Angebote von westlichen Firmen eingeholt, um so den offiziellen langwierigen Weg zu unterlaufen. Diese Unterlagen wurden nach verschiedenen Gesichtspunkten katalogisiert. Damit lagen bereits mit der Wende nutzbare Firmenverzeichnisse vor, die anfangs die wichtigsten Quellen der Marktarbeit darstellten. Auch über rationale Verfahren, neue Werkstoffe und moderne Bauelemente werden seit langem Dossiers angelegt. Früher dienten sie nur zur Orientierung, jetzt aber werden sie unmittelbar für den Konstruktionsprozeß genutzt. Diese langfristig entwickelten Beziehungen zu den Nutzern und die kurzfristige Ausstattung der Informationseinrichtung mit leistungsfähigen Rechnern sowie einem Kopiergerät ermöglichten die Übernahme neuer Aufgaben.

#### Technische Ausstattung der Informationseinrichtung

Seit mehreren Jahren werden Dokumentennachweise rechnergestützt verarbeitet. Anfangs auf dem Robotron PC 1715, danach auf dem A 7150. Wachsende Fondsgrößen ließen besonders bei Komplexrecherchen die Rechnerzeiten immer mehr anwachsen. Seit kurzem konnten nun die internen Datenbanken auf einen leistungsfähigeren Siemens PCD-2 M übernommen werden. Noch vor der Währungsunion ergab sich die Möglichkeit, ein Angebot der FIZ Technik zu nutzen und zwei Philips-Rechner zu erwerben. Die Geräte P 3120 und P 3230 ermöglichen einerseits Recherchen über CD-ROM und andererseits Online-Recherchen in Datenbanken. Mit Hilfe dieser Rechner verbesserte sich die Informationsversorgung ganz wesentlich. Vor allem durch die CD-ROM-Speicher wurden Teile einiger westdeutscher Datenbanken nutzbar, die im Online-Betrieb wegen der mangelhaften Übertragungsmöglichkeiten uns anfangs noch verschlossen waren. Besonders in der ersten Zeit nach der Währungsunion reichte die normale Arbeitszeit kaum aus, um alle Anfragen mit diesen Speichermedien beantworten zu können.

Die wichtigsten von uns genutzten Dateien auf CD-ROM sind:

Wer liefert was?  
FIZ Technik DOMA  
FIZ Technik BEFO  
PATOS

und seit ihrem Erscheinen ABC-EUROPEX mit zunehmender Tendenz. Nach dem Anschluß der Datenbanken über Online-Verbindung werden auch diese Informationsspeicher immer häufiger genutzt. Datenbanken, aus denen regelmäßig Informationen bezogen werden, sind:

FIZ BFAI Außenhandelsinformationen  
DOMA Maschinenbau und Fertigungstechnik  
SDIM Metallische Werkstoffe  
EURD Europäische Exportindustrie

Genios VD - Creditreform  
HB - Handelsblatt  
TABL- Wirtschaftstabellen  
WGZW- Wer gehört zu wem?

DATA-STAR PTSP Amerikanische Wirtschaftsdatenbanken

Ergänzt werden konnten die Informationsleistungen durch die Einrichtung eines Bildschirmarbeitsplatzes (Btx). Über diese Anlage werden vor allem hochaktuelle Wirtschafts- und Firmeninformationen ermittelt. Auch für die Werbung der eigenen Erzeugnisse wird dieses Medium genutzt. Während einer Übergangszeit beziehen wir über Btx auch Fernsprechnummern sowie Fahrpläne der Bundesbahn.

Seitdem die Informationseinrichtung auch über ein eigenes Xerographie-Kopiergerät verfügt, sind auch eilige Kopieraufträge problemlos zu realisieren. Eine Arbeitserleichterung, die wir besonders schätzen, da vorher alle zu vervielfältigenden Dokumente etwa einen Kilometer weit zur zentralen Kopierabteilung transportiert werden mußten.

## Veränderungen bei der Auskunftstätigkeit und den Informationsleistungen

Die bessere technische Ausstattung und die erweiterten Kommunikationsmöglichkeiten sind Voraussetzung, um den durch die neue Wirtschaftsform veränderten Informationsbedürfnissen weitgehend nachkommen zu können. Die Aufgaben der Informationseinrichtung bestanden früher vor allem darin:

- den F/E-Prozeß mit entsprechenden wissenschaftlichen und technischen Informationen zu unterstützen;
- die relevanten Patentnachweise aus dem Express-Informationsdienst zu ermitteln und aufzubereiten;
- von der im Betrieb vorhandenen Literatur Nachweise anzufertigen und diese in Datenbanken zu speichern;
- Informationen über Konkurrenzzeugnisse zu beschaffen und auszuwerten sowie
- die Fakten solcher Erzeugnisse in Karteien zu verdichten und in Studien zu verarbeiten.

Jetzt werden besonders Zuarbeiten mit folgendem Inhalt erwartet:

- Initial-Informationen, die zu einer wesentlichen Qualitätserhöhung beitragen oder die zur Entwicklung alternativer Produkte mit langzeitigem Absatz führen;
- Verbesserung der Markt-Situation durch Informationen über potentielle Kunden;
- Nachweis von Bezugsquellen sowie von Kaufteil- und Baugruppenlieferanten;



- Ermittlung von Mitbewerbern mit dem Ziel der Kontaktanbahnung, um gemeinsam auf Märkten aufzutreten, Lizenzen auszutauschen oder zu kooperieren;
- Nachweis von Arbeitsangeboten, die zu einer besseren und kontinuierlichen Auslastung der eigenen Werkzeugmaschinen beitragen könnten;
- Hinweise auf Kaufgesuche nach gebrauchten Maschinen und Geräten, um den Verkauf nicht mehr benötigter Maschinen zu erleichtern;
- Unterstützung der Werbeabteilung bei der Auswahl von Agenturen, die für die Werbung unserer Erzeugnisse geeignet sind sowie von Datenbanken, in denen sich das eigene Unternehmen darstellen kann;
- Ermittlung und Aufbereitung von Wirtschaftsinformation für die Geschäftsleitung nach folgenden Aspekten:
  - Einrichtungen, die für die Vergabe staatlicher Fördermittel zuständig sind, möglichst mit Hinweisen auf den formalen Prozeß der Antragstellung,
  - Umsatzdaten und Absatzmengen von Mitbewerbern gegebenenfalls als Trendaussage,
  - Regelmäßige und kurzfristige Nachweise von Ausschreibungen, für die Bewerbungen abgegeben werden sollten,
  - Länder und Firmeninformationen zur Vorbereitung auf Dienstreisen, Messen und Tagungen.
- Überwachung der Patentsituation. Diese Aufgaben werden gemeinsam mit der Patentabteilung durchgeführt. Viele Anfragen zu dieser Problematik lassen sich umgehend und vor allem kostengünstig durch Recherchen in CD-ROM "Patos" beantworten. Auch für die Vorbereitung auf Online-Recherchen wird "Patos" gern genutzt, vor allem mit dem Ziel, Recherchezeit und -kosten gering zu halten.

## **Vor- und Nachteile, die sich aus den veränderten Aufgaben ergeben**

Sowohl die neuen Aufgaben als auch die technische Ausstattung der Informationseinrichtung führten zu einem Wandel im Tätigkeitsbild der Informationsingenieure. Der Spezialist eines Fachgebietes wird abgelöst von einem vielseitig einsetzbaren und aussagefähigen Ansprechpartner. Die Verringerung der Mitarbeiter in der Informationsabteilung erfordert darüber hinaus erhöhte Bereitschaft zur Zusammenarbeit. Wechselseitige Übernahme von Arbeiten ist bei kurzfristigen Terminen unerlässlich. Diese Prozesse werden zwischen den einzelnen Mitarbeitern weitgehend selbständig abgestimmt. Geplant werden lediglich solche Aufgaben, die zeitaufwendig sind (z. B. Studien) oder erhebliche Kosten verursachen (z. B. Komplex-Recherchen). Die Sichtung und Auswertung der Dokumente dient nicht mehr ausschließlich dazu, die Datenbanken zu ergänzen, vielmehr soll auch dazu beigetragen werden, bisher nicht bekannte Apparate, Maschinen, Verfahren und Werkstoffe besser zu nutzen. Eine Aufwertung erfährt die Informationsarbeit besonders durch Anfragen von der Geschäftsleitung sowie den Vertriebs- und Einkaufsbereichen. Bei Entscheidungen über die Unternehmensstrategie sowie zur Produkt- und Marktentwicklung werden Fachinformationen permanent gefordert, ein Prozeß, der erst mit der Unternehmensausbildung einsetzte.

Die noch nicht abgeschlossene Unternehmensstruktur sowie die nicht durchgängig geplante Produktentwicklung drängen die Informationseinrichtung dazu, die Fachinformationsflüsse im Betrieb neu zu bestimmen und wiederholt zu ändern. So wurde festgestellt, daß wichtige Nutzer aus dem Informationsfluß ausgeschlossen wurden, obwohl sie im Verteiler ausdrücklich genannt waren. Dieses falsche Konkurrenzdenken und unternehmensschädigende Verhalten einiger Mitarbeiter ließ sich nur durch breit angelegte Informationsflüsse umgehen. Der weltweite Online-Zugriff auf Datenbanken verringert die Bedeutung der internen Informationsspeicher. Die zu erwartende häufige Umstellung der Produktion auf neue Erzeugnisse läßt den aufwendigen Aufbau neuer eigener Datenbanken nicht zu.

Geprüft wird gegenwärtig, für welche Fachgebiete weiterhin interne Informationsspeicher existieren müssen und in welchen Fällen auf externe Datenbanken zurückgegriffen werden sollte. Kriterien für diese Entscheidung sind neben den Kosten vor allem die notwendige Trefferquote sowie die geforderte Recherchetiefe.

Noch unklar ist Art und Höhe der künftigen Finanzierung der Informationseinrichtung. Angestrebt wird eine Umlage, die von den Unternehmensbereichen zu tragen ist. Jedoch soll auch ein Teil der Kosten durch Eigenfinanzierung abgedeckt werden. Arbeitsmöglichkeiten und Leistungsfähigkeit werden von dieser Entscheidung nachhaltig beeinflusst.

Die personellen, organisatorischen und technischen Voraussetzungen für eine wirksame Informationsversorgung wurden geschaffen. Zu keiner Zeit ihres 26jährigen Bestehens wurde die Informationseinrichtung mehr genutzt und stärker gefordert als vom Zeitpunkt der Wende an. Die früher häufig vermißte Anerkennung der Informationsleistungen hat sich fast von selbst eingestellt. Durchgesetzt hat sich die Erkenntnis, daß sich mit einer gut funktionierenden Informationseinrichtung viele Aufgaben leichter lösen lassen. Besonders vorteilhaft für die Nutzer ist, daß sie alle zur Lösung eines Problems erforderlichen Informationen mit einem Auftrag auslösen und aus einer Hand erhalten.

**17. Internationalen Kolloquium für Information und  
Dokumentation**

**in Verbindung mit ISI'91**

**Akzeptanzhürden spielerisch überwinden -**

"Informationsmanagement" in der berufsbegleitenden Weiterbildung.

Ein Vermittlungskonzept dargestellt am Beispiel  
"Kosten- und Leistungsrechnung für Informationsdienstleistungen"

Gisela Frerk, Achim Oßwald

**Abstract:**

Ziel des Vortrages ist es, konkrete Erfahrungen bei der Vermittlung des Konzeptes "Informationsmanagement" am Beispiel einer Unterrichtseinheit zur "Kosten- und Leistungsrechnung für Informationsdienstleistungen" zur Diskussion zu stellen. Nach einer kurzgefaßten Darstellung des organisatorischen und zielgruppenspezifischen Rahmens der Weiterbildung zum Wissenschaftlichen Dokumentar am Lehrinstitut für Dokumentation (LID), Frankfurt, umreißt der Vortrag die wesentlichen Inhalte und den Stellenwert der Unterrichtseinheit "Kosten- und Leistungsrechnung für Informationsdienstleistungen" innerhalb des Unterrichtsmoduls "Informationsmanagement". Im Mittelpunkt der Ausführungen steht das didaktische Konzept der Unterrichtseinheit, das getragen wird von der Wechselwirkung von Präsentation und aktiver, spielerischer Aneignung von Wissen in Arbeitsgruppen. Dabei wird die spezifische Problemstellung des Themas "Kosten- und Leistungsrechnung" in zweifacher Weise berücksichtigt: einerseits werden die ihm gegenüber aufgebauten Akzeptanzhürden überwunden, andererseits die

zentralen Aspekte des Themas herausgearbeitet und in den Fachinformationsbereich übertragen. Ein entscheidender Ansatzpunkt ist dabei die konkrete Vorerfahrung der Weiterbildungsteilnehmer.

Vor dem Hintergrund der Lernerfolge des bis zum November 1991 in insgesamt fünf Durchläufen optimierten Konzeptes wird eine Einschätzung gegeben, inwieweit das Vermittlungskonzept - unter Einbeziehung entsprechender Vorerfahrungen - auf andere Zielgruppen übertragbar ist und damit zur Überwindung von Akzeptanzhürden gegenüber diesem für die Fachinformation ungemein wichtigen Themenbereich beitragen kann.

It is the aim of this paper to discuss the experiences in teaching the concept of "Informationmanagement" at the concrete example of cost and benefit calculation. First, there is a brief description of the organizational aspects and target groups of the "Wissenschaftlichen Dokumentar" training at the Lehrinstitut für Dokumentation (LID), Frankfurt. Further, main contents and importance of the teaching unit "Cost and benefit calculation of information services" are pointed out. The focus of the interest is on the didactical concept which consists of the alternating effects of knowledge presentation and acquisition in working groups. This special concept is supposed to overcome barriers of acceptance in a more "game-like" way of teaching and learning.

## **1. Überblick**

Informationsverarbeitende Aufgaben erlangen in unterschiedlichen gesellschaftlichen Bereichen zunehmend an Bedeutung. Die Qualifizierung für solche Aufgaben durch Fort- und Weiterbildung wird daher immer wichtiger. Neben technisch-methodischen Fragen ist dabei in den letzten Jahren auch die Wirtschaftlichkeit von Informationsprodukten und -dienstleistungen wieder in den Kernbereich des Interesses

gerückt. Ihre Einbeziehung in entsprechende Curricula wirft Fragen auf, zu deren Lösung das nachfolgend vorgestellte Konzept erfolversprechende Hinweise gibt.

## **2. Der Bedarf: Berufsbegleitende Weiterbildung**

Der für das Tätigkeitsfeld IuD typische, hohe Weiterbildungsbedarf wird verschärft durch die große Anzahl von beruflichen Quereinsteigern. Diese sind zumeist erfahrene Fachleute ohne dokumentarische Qualifikation oder junge Wissenschaftler, die im Tätigkeitsgebiet Fachinformation eine zukunftssträchtige Ergänzung ihrer Fachausbildung sehen. Ein spezielles Aufbaustudium wie z.B. in Konstanz /1/ steht ihnen aus formalen (Zulassungsbedingungen) oder finanziellen und zeitlichen Gründen faktisch selten offen. Universitäre Ausbildungsgänge für Berufspraktiker, wie sie bislang an der TH Ilmenau /2/ angeboten wurden, bestanden in den alten Bundesländern der BRD bislang nicht. Dort widmet sich stattdessen seit 1967 das **Lehrinstitut für Dokumentation (LID)** in der DGD dieser Aufgabe, indem es Dokumentationspraktiker theoretisch und praktisch, produkt- und herstellerneutral für unterschiedliche Tätigkeitsbereiche im Rahmen der Fachinformation qualifiziert /3//4/. Die inhaltliche und methodische Verzahnung der Unterrichtsmodule in den einzelnen Curricula bzw. Fortbildungsveranstaltungen hat dabei zum Ziel, die Arbeitsabläufe im Fachinformationsbereich exemplarisch nachzubilden und dadurch die Arbeit der Praktiker mit theoretischem und handlungsorientiertem Wissen zu unterstützen.

Dieses Konzept wurde Ende der 80er Jahre um Lehrinhalte der "Kosten- und Leistungsrechnung" (KLR) ergänzt. Sie wurden in den Themenbereich "Informationsmanagement" eingebettet und haben dort heute einen hohen Stellenwert.

### **3. Die Zielgruppe**

Nur ein geringer Anteil der Kursteilnehmer kann auf positive Erfahrungen mit dem Thema KLR zurückblicken. Zudem bleibt im beruflichen Alltag selten Zeit, solche Fragen systematisch aufzuarbeiten. Begriffe und die damit angesprochenen Sachverhalte werden häufig unsicher und ungenau benutzt. Vermittlungskonzepte müssen daher nicht nur themenspezifische Aspekte berücksichtigen, sondern auch die mittel- und langfristige, sachlich und persönlich zu sehende Relevanz des Themas aufgreifen und vermitteln.

Darüberhinaus besteht bei den Teilnehmern von Fort- und Weiterbildungsveranstaltungen das Bedürfnis nach kompakter, im Hinblick auf andragogische (erwachsenenbildnerische) Aspekte optimierte didaktische Aufbereitung der Sachverhalte /5/ /6/. Diese müssen von den Dozenten auf ihre praktische Relevanz hin reduziert werden, ohne daß der methodisch-theoretische Kontext verloren gehen darf.

### **4. Das inhaltliche Umfeld: Lehrinhalte des Unterrichtsmoduls "Informationsmanagement"**

In der 13-wöchigen berufsbegleitenden Weiterbildung am LID, die als Fallbeispiel im weiteren zugrunde gelegt wird, umfaßt das Unterrichtsmodul "Informationsmanagement" zwei Unterrichtswochen. Bei seiner Durchführung können die theoretischen und praktischen Lehrinhalte der Lehrgebiete "Entwurf von Informationssystemen", "Inhaltserschließung" sowie "Information Retrieval" als bekannt vorausgesetzt werden.

In der begrenzt zur Verfügung stehenden Unterrichtszeit findet eine Konzentration auf die folgenden Themenbereiche statt:

Begriffliche Orientierungen; Benutzerforschung; Projektmanagement; Organisation einer Fachinformationsstelle; Informationsmarketing; Informationsdienstleistungen und ihre Erstellung; KLR für Informationseinrichtungen;

Servicekonzepte und Nutzen-Wirksamkeitsanalysen; Rechtsfragen; Bedingungsfaktoren des Informationsmarktes.

##### **5. Einbettung der KLR in das didaktische Konzept des Unterrichtsmoduls "Informationsmanagement"**

Aus andragogischen Überlegungen wird die Unterrichtsform getragen vom Wechsel zwischen theoretischem Unterricht und Projektarbeit in Gruppen, in denen die theoretischen Kenntnisse auf eine Projektfragestellung umgesetzt und konkretisiert werden müssen. Daraus entsteht zusätzlicher Informationsbedarf, der in den nachfolgenden Unterrichtssequenzen befriedigt wird.

Die Projektarbeit in Arbeitsgruppen von max. 5 Lehrgangsteilnehmern hat zum Ziel, in Fachinformationsstellen unterschiedlicher Organisationsform Neuerungen bzw. Änderungen personeller, organisatorischer, inhaltlicher oder technischer Art einzuführen. Hierfür werden Rahmenbedingungen von fiktiven, aber im Hinblick auf die Informationspraxis jedoch prototypischen Stellen vorgegeben. Die Arbeitsgruppen sind dadurch mit typischen Problemen des Informationsmanagements konfrontiert, deren "Lösung" sie am Ende der Projektphase dem Plenum vorstellen und dort auch verteidigen müssen. Die Einteilung der Lehrgangsteilnehmer in Arbeitsgruppen erfolgt u.a. im Hinblick auf ihre berufliche Vorerfahrung, so daß im Projekt aus diesen Erfahrungen geschöpft werden kann.

Ein wesentlicher Aspekt der Problemstellungen bezieht sich auf Fragen der KLR, die in der ersten der beiden Unterrichtswochen in einer eintägigen Unterrichtseinheit theoretisch und methodisch eingeführt wurde. Das hierfür entwickelte Unterrichtskonzept ist Thema der nachfolgenden Ausführungen.

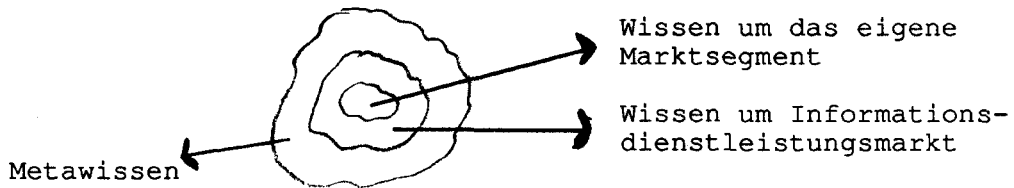


## **6. Vermittlungskonzept für den Themenbereich "Kosten- und Leistungsrechnung"**

### **6.1 Vermittlungsziele**

Die einzelnen Projekte können ohne die Reflexion und konkrete Analyse von KLR-Aspekten nicht sinnvoll durchgeführt werden. Der hohe Stellenwert der KLR im Projekt-, aber auch im Gesamtkonzept des Unterrichtsmoduls IM spiegelt seine Bedeutung in der Realität der Informationspraxis wieder. Durch diese Einbettung ist es möglich, daß die Lehrgangsteilnehmer die Bedingtheit des Handelns im Informationsbereich im Hinblick auf Wirtschaftlichkeitsfragen erkennen, Einflußfaktoren benennen können und wesentliche Verfahren für die Kostenermittlung sowie die Preisfestsetzung von Informationsdienstleistungen oder -produkten benutzen können.

Hierzu ist es notwendig, dem Kursteilnehmer ein auf verschiedenen Ebenen angesiedeltes Informationsbewußtsein zu vermitteln (vgl. information onion, /7/): Die "Informationszwiebel" in Fachinformationsbereich setzt sich aus folgenden Informationsebenen zusammen: Auf der Ebene des Metawissens (Meta information awareness) sieht sich der Informationsdienstleister als ein Subsystem, welches in die ihn umgebende Umwelt eingebettet ist. Hier muß er Wissen in Bezug auf Gesetze, Verordnungen, Techniken sowie ökonomische und psychologische Aspekte von Informationsdienstleistungen haben. Auf der Ebene der Informationsdienstleistungsmärkte (Domain specific information skills) muß ein Wissen um Bedürfnisse, Angebote, verfügbare Datenbasen, usw. vorhanden sein. Der engste Kreis umfaßt das Wissen um das eigene Marktsegment (Disciplinary core), das auch das Wissen um die eigene Wirtschaftlichkeit (Kosten- und Leistungsrechnung) einschließt.



Wirtschaftlichkeit ist jedoch ein vielschichtiger Begriff, da Kosten- und Nutzenbewertungen auf verschiedenen ökonomischen Ebenen durchgeführt (beispielsweise Stückkostenrechnungen, Gewinn-Verlust Rechnungen, Kosten-Nutzen-Analysen, Nutzwertanalysen) und weiter differenziert werden können (z.B. nach betrieblichen und gesellschaftlichen Kosten, nach individuellen und gesellschaftlichen Nutzen). Eine ausführliche Darstellung der Lehrinhalte des Themenbereichs "Kosten- und Leistungsrechnung" ist jedoch zu umfangreich und besonders am Anfang zu abstrakt. Es wird eine Auswahl des Lehrstoffes aus dem Bereich der Kosten- und Leistungsrechnung erforderlich, der keinen Leistungsdruck und Stress erzeugt, wohl aber die Informationen vermittelt, die den Ziel der Ausbildung entsprechen.

Von besonderer Bedeutung für die Auswahl des Lehrstoffes ist das Vorwissen der Kursteilnehmer. Erfahrungen im Bereich der Kosten- und Leistungsrechnung sind selten und viele Kursteilnehmer bringen eine gewisse Voreingenommenheit gegen mathematische Darstellungsweisen und ökonomische Fachbegriffe wie Rentabilität oder Effizienz mit. Effizienznachweise werden aber immer häufiger von der übergeordneten Einrichtung, dem Nutzer oder der Öffentlichkeit den Fachinformationseinrichtungen verlangt und sind darüberhinaus auch für die eigene Planung unabdingbar. Hieraus ergeben sich spezielle Probleme, die bei der Vermittlung der Lehrinhalte berücksichtigt werden müssen.

## **6.2 Spezifische Problemstellung des Themas**

In Bezug auf die beabsichtigten Ergebnisse (Lernziele) muß beachtet werden, daß das Wissen im Bereich der Kosten- und Leistungsrechnung nicht nur reproduzierbar, sondern auch in der konkreten beruflichen Situation anwendbar sein muß. Kosten- und Wirtschaftlichkeitsabwägungen sollten für die unterschiedlichen Leistungen der Fachinformationsstelle möglich sein. Besondere Anforderungen stellen sich hier angesichts der sich ständig wandelnden Anforderungen durch neue Technologien (wie Online, CD-ROM). Auf der anderen Seite macht die Verschiedenartigkeit der Fachinformationseinrichtungen und ihrer Leistungen Abstraktionen notwendig. Einfache Ordnungssysteme, grundsätzliche Zusammenhänge und Kennzahlen sollen verstanden und angeeignet werden. Beispielsweise das Denken in einfachen Betriebsabrechnungssystemen (wie der Betriebsabrechnungsbogen), Zusammenhänge zwischen den grundsätzlichen Möglichkeiten und Arten der Kostenzuweisung und -verrechnung sowie die Fähigkeit zum Umgang mit einfachen Kennzahlen wie Personalkosten im Verhältnis zu Sachkosten, Gesamtkosten im Verhältnis zu Mitarbeitern oder Personalkosten im Verhältnis zu Mitarbeitern.

Um die spezielle Problemstellung des Themas didaktisch erfolgreich zu operationalisieren, müssen die drei Prinzipien der Strukturiertheit, Einfachheit und zusätzlichen Stimulanz beachtet werden /8/:

### **a. Strukturiertheit**

Wichtig ist eine klare, logische, geordnete Strukturierung des Stoffes, die grundlegende Prinzipien betont und Nebensächlichkeiten vermeidet. Hierzu erfolgt zuerst die Einordnung der Kosten- und Leistungsrechnung in das betriebliche Rechnungswesen. Danach wird auf die Begriffsmerkmale und Ziele der Kosten- und Leistungsrechnung eingegangen. Weiterhin werden ihre Gestaltungsalternativen, die Varianten des Kostenbegriffs und wichtige Kostenkategorien erläutert.

Um Licht in das Begriffswirrwarr im Zusammenhang mit der ökonomischen Analyse von Fachinformationseinrichtungen zu bringen, erfolgt im Rahmen der Leistungsrechnung zuerst die Klärung und Abgrenzung der Grundbegriffe wie Informationsbedürfnis, Informationsbedarf, Informationsdienstleistung, Umsatz, Gewinn, Leistung, monetärer Nutzen, Nutzen. Anschließend werden die unterschiedlichen Bewertungsebenen von Fachinformationseinrichtungen bzw. für die von ihnen erstellten Dienstleistungen vorgestellt und diskutiert.

b. Einfachheit

Resultierend aus den Unterrichtserfahrungen wurden relativ komplizierte Kostenverrechnungssysteme wie die "Kostendeckungsbeitragsrechnung" nach Riebel nicht mehr in den Unterricht aufgenommen, so daß jetzt eine Konzentration auf das grundsätzlichere Vollkosten- und Teilkostenrechnungssystem möglich ist. In der Kosten- und Leistungsrechnung als Teil des internen Rechnungswesens ist die Verwendung von ökonomischen Fachausdrücken unumgänglich. Da dies jedoch häufig auf Voreingenommenheit bzw. Akzeptanzprobleme bei den Kursteilnehmer stößt, wird sie durch Beispiele aus der Praxis und dem täglichen Leben unterlegt.

Zu beachten ist ferner, daß sich zwischen Einfachheit und Strukturiertheit Zielkonflikte ergeben können: gerade am Anfang erscheinen Ziele und Vorgehen der Kosten- und Leistungsrechnung sehr theoretisch und abstrakt, so daß eine eher praxisbezogene Vorgehensweise zur Überwindung von Akzeptanzhürden beitragen könnte. Diese jedoch, (Beispiel Betriebsabrechnungsbogen) ist wiederum ohne vorangehende Erläuterung der Begrifflichkeiten nur schwer verständlich.

c. Zusätzliche Stimulierung

Lernpsychologische Erkenntnisse der Bedingungen menschlichen Lernens zeigen, daß ein erfolgreiches Lernen besonders stark von einer hohen "inneren Motivation", d.h. von einem starken Interesse am Stoff selbst abhängen /8/. Motivationsfördernd wirkt hier eine zusätzliche Stimulanz, ein spezifisches Vermittlungskonzept.

**6.3 Vermittlungskonzept: Wechselwirkungen zwischen Präsentation und erarbeitendem Lernen**

Von besonderer Bedeutung für die innere Motivation sind die sozialen und kommunikativen Bedingungen des Lernens. Wichtig für die Kursteilnehmer ist ein breiter Handlungsspielraum, indem sie die Inhalte mitbestimmen und eigenverantwortliche handeln können. Das Vermittlungskonzept für die Unterrichtseinheit Kosten- und Leistungsrechnung sieht den Abbau der einseitigen Kommunikation Dozent-Kursteilnehmers und die Wechselwirkung zwischen Präsentation und erarbeitendem Lernen vor.

**6.3.1 Rezeptive Transformation theoretischer Grundlagen anhand von ausgewählten Beispielen**

Der Unterricht besteht aus einem Wechsel von theoriebezogenem Unterricht durch eine KLR-Expertin, gestützt durch Beispielsrechnungen und konkrete Berichte aus der Praxis (durch einen LID-Dozenten), wodurch die Reflexion und der Vergleich der eigenen Berufserfahrung mit den vorgestellten Beispielen intensiviert wird. Die Beispiele werden aus der täglichen beruflichen Praxis der Kursteilnehmer gewählt und beziehen sich somit auf deren berufliches Erfahrungswissen. So werden im Rahmen der Kostenträgerrechnung die Kosten einer CD-ROM und einer Online Recherche berechnet und verglichen. Im Rahmen der Wirtschaftlichkeitsanalyse wird beispielsweise eine empirische Kosten-Nutzenanalyse im Bereich der

Endbenutzerrecherchen in Literaturdatenbanken an der Universität Konstanz vorgestellt /9/.

### **6.3.2 Aktives Verarbeiten durch spielerisches Lernen**

Untersuchungen zeigen, daß bei steigender Mitwirkung der Kursteilnehmer die Vergessensquote abnimmt. So lag sie bei einer auditiven und visuellen Darbeitung des Stoffes bei etwa 50%, bei einem gemeinsamen Erarbeiten und Tun bei etwa 10%. Diesen Untersuchungsergebnissen wird durch das "Spiel zur Kostenrechnung" Rechnung getragen. Ziel des Spiels zur Kostenrechnung ist, daß sich möglichst alle Kursteilnehmer aktiv beteiligen, Ideen einbringen und kooperieren. In einem gruppenweise aufeinanderbezogenen Arbeiten und Lernen sollen die Kursteilnehmer selbständig Entscheidungen treffen und die Stellung der Fachinformationseinrichtung im volkswirtschaftlichen Gesamtsystem erkennen (Abb.1/Anhang).

Es soll ein Bewußtsein dafür vermittelt werden, daß der Informationsdienstleister als wirtschaftliche Einheit in ein marktwirtschaftliches System und dieses wiederum in ein gesellschaftliches System eingebettet ist. Als Subsystem steht er mit seiner Umwelt (Beschaffungs-, Absatz-, Kapitalmärkte und Staat) in zahlreichen Transaktionsbeziehungen. Dabei fließen sowohl Ströme finanzieller wie auch materieller Art zwischen Informationsdienstleister und Lieferanten, Kunden, Staat und Kapitalgebern. Die Kosten- und Leistungsrechnung bildet die wirtschaftlich bedeutsamen Vorgänge ab, die sich innerhalb der Unternehmung abspielen, also den unternehmerischen Transformationsprozeß. Die Hauptaufgabe ist, den Verzehr von Produktionsfaktoren (Input) und die damit verbundene Leistungsentstehung (Output) mengen- und wertmäßig zu erfassen und die Wirtschaftlichkeit der Leistungsentstehung zu überwachen. In Fachinformationseinrichtungen entstehen Kosten beispielsweise durch den Einsatz menschlicher Arbeitskraft; die Nutzung von Grundstücken und Gebäuden; den

Verbrauch/Gebrauch materieller Güter (Büromöbel, Computer, usw.).

Gleichzeitig soll auf die bei der Kostenrechnung in einzelnen Fachinformationseinrichtungen auftretenden Schwierigkeiten im Prozeß der Informationsgewinnung aufmerksam gemacht werden /10/:

- Informationseinrichtungen sind häufig Teil unterschiedlicher Institutionen mit verschiedenen Zielen und Verfahren der Kostenrechnung. Dies führt dazu, daß bestimmte Kostenarten gar nicht oder aber in sehr unterschiedlichem Grad in den Kosten der Informationsaktivitäten berücksichtigt werden: z.B. Verwaltungs-, Gebäude- und Grundstückskosten (im Spiel wird dies in den Anlagekarten berücksichtigt).

- In kleineren Informationseinrichtungen üben einzelne Personen oft mehrere Funktionen aus. Dies erschwert die Zurechnung der Personalkosten auf einzelnen Teile des Prozesses der innerbetrieblichen Leistungserstellung (im Spiel wird dies in den Zeitreporten berücksichtigt).

- Werden aus einer bestimmten Datenbasis verschiedene Arten von Informationsdiensten angeboten, wie gedruckte Dienste, Online-Recherchen, Profildienste, so ergeben sich Probleme bei der Zurechnung der Inputkosten auf die verschiedenen Dienstleistungsarten (im Spiel wird dies in den Zurechnungssätzen in der Kostenträgerrechnung berücksichtigt).

Ablauf des Kostenrechnungsspiels (Abb. 2/ Anhang):

1. die Vorstellung der Spielidee erfolgt im Plenum
2. Arbeitsgruppen werden gebildet
3. Verteilen und Erläuterung der Arbeitsblätter
4. Problemlösen in der Gruppe
5. Problemlösungsvollzug im Plenum
6. Abstimmung der Ergebnisse (Feed back in Gruppe)
7. Diskussion der Ergebnisse (Feed back im Plenum)

## 8. Diskussion des Spiels (Feed back zur Spielidee)

Die Anforderungen, Entscheidungen zu treffen für Fragen, für die sich die Kursteilnehmer erst noch qualifizieren sollen, kann Schwierigkeiten bereiten. Deshalb werden den einzelnen Arbeitsgruppen während des Problemlösungsprozesses Betreuer zur Seite gestellt. Dies geschieht einmal zur Beobachtung und Förderung der gruppenspezifischen Prozesse. Zum anderen soll durch fachliche Hilfeleistung die ständige Rückkopplung zur Spielidee gesichert werden.

### **6.3.3 Kommunikative Verarbeitung und Ergänzung des Gelernten**

Die kommunikative Erarbeitung und Ergänzung des Gelernten findet einmal in der Spieldiskussion statt. Hier wird hinterfragt, inwieweit das Spiel geeignet ist, sich mit dem Thema vertraut zu machen. Verbesserungsvorschläge aus dem Plenum werden in den Spielaufbau- und ablauf eingebaut. Zum anderen werden die Lehrinhalte der Kosten- und Leistungsrechnung in der konkreten Projektarbeit berücksichtigt und weiter diskutiert.

### **6.4 Lernerfolge in kurz- und mittelfristiger Betrachtung**

Der Lernerfolg, der durch dieses Vermittlungskonzept erreicht wurde, wird im Rahmen der Weiterbildung auf mindestens zwei Zeitebenen überprüft. Im Rahmen des Projektunterrichtes durch KLR-spezifische Fragestellungen, die von den Arbeitsgruppen bearbeitet werden sowie - zu einem späteren Zeitpunkt - durch Klausurfragen zum Thema KLR. Die erste Phase läßt vor allem für die Kursteilnehmer selbst erkennen, ob sie die Prinzipien und Verfahren der KLR verstanden haben. In der Gruppe und unter Bezugnahme auf die spezielle Fragestellung im Projekt ist so eine Ergänzung der Kenntnisse möglich. Hierbei übernimmt der betreuende Dozent vor allem die Aufgabe, durch korrigierende und ergänzende Hinweise den Lernerfolg für alle



Kursteilnehmer auf ein vergleichbares Niveau zu bringen. Die Überprüfung des Lernerfolges in einer Klausur ermöglicht zu erkennen, ob das Gelernte individuell und auch nach einer längeren Zeitspanne angewendet werden kann. Die bisher in den Klausuren erzielten guten Ergebnisse bestätigen die mit dem Konzept erreichten Lernerfolge.

### **7. Übertragbarkeit des Vermittlungskonzeptes auf andere Themen und Zielgruppen**

Charakteristisch für die Zielgruppe in den LID-Lehrgängen ist die Heterogenität der Vorkenntnisse im Thema KLR, die z.T. geringe Affinität zu den Fragestellungen und Verfahrensprinzipien der KLR sowie eine Gemeinsamkeit: das Interesse am Anwendungsgebiet Dokumentation/Fachinformation. Diese Charakteristika sind bei Mitarbeitern in zahlreichen anderen Anwendungsgebieten gegeben, wo das Konzept bedarfsspezifisch übertragen und eingesetzt werden kann. Darüberhinaus sehen wir mit diesem Konzept auch die Chance, Berufstätige aus sozialen Zusammenhängen, in denen das Thema KLR bislang nur eine geringe Rolle gespielt hat (z.B. Bürger der ehemaligen DDR, aber auch Übersiedler aus osteuropäischen Ländern) in z.T. spielerischer Weise mit derartigen Fragen und Konzepten vertraut zu machen. Die pädagogische und didaktische Kompetenz der agierenden Dozenten für die Umsetzungen des Konzeptes bleibt hierfür jedoch eine unabdingbare Voraussetzung.

### **8. Zusammenfassung der Ergebnisse**

Die Lernerfolge weisen das am LID praktizierte spezifische Vermittlungskonzept als erfolgsversprechend aus. Ein didaktisches Konzept der Wechselwirkungen von Präsentation und spielerischer Aneignung von Wissen hilft einmal Akzeptanzhürden zu überwinden. Zum anderen erscheint es zur Übertragung auf andere Zielgruppen und Anwendungsgebiete geeignet.

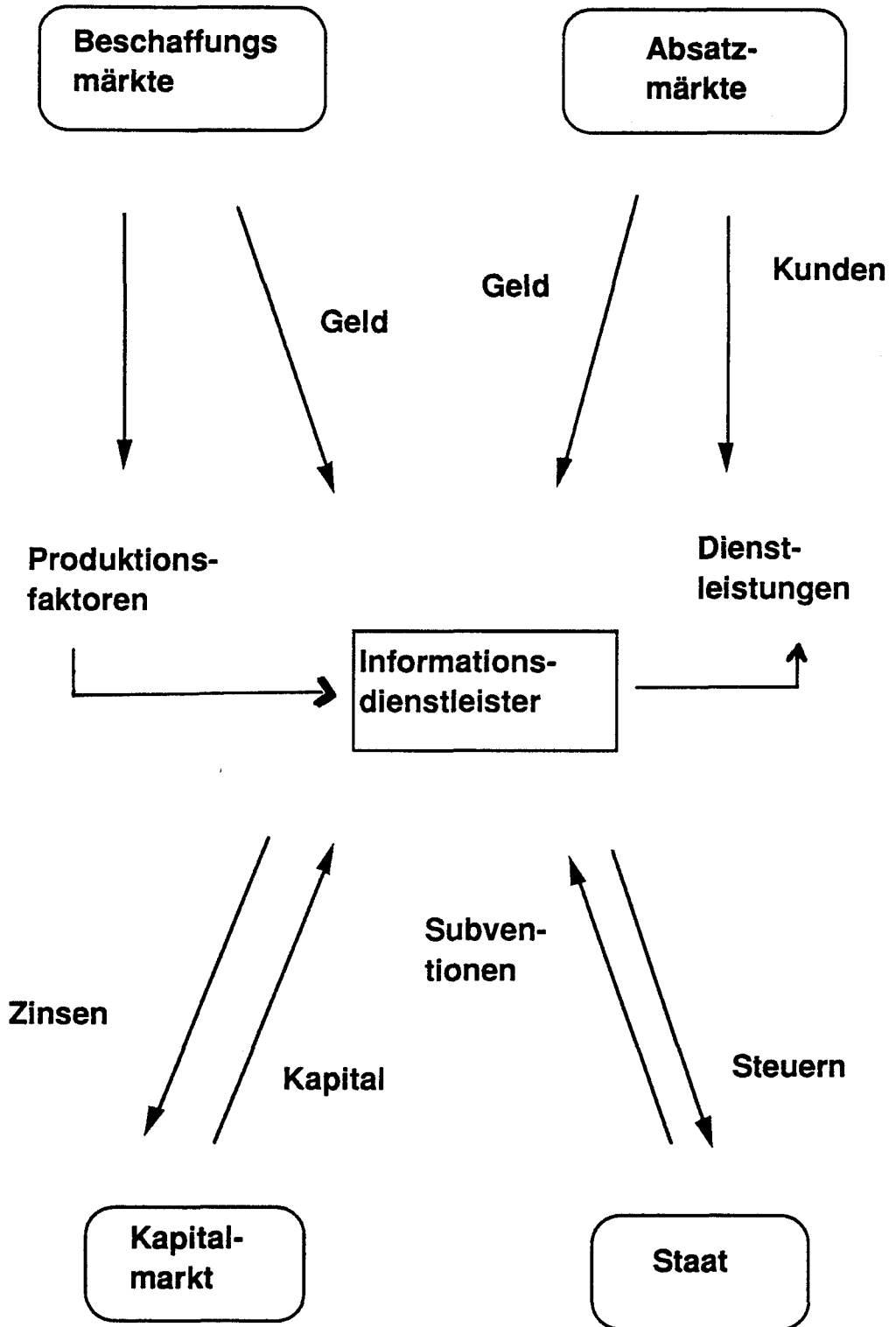


Abb. 1

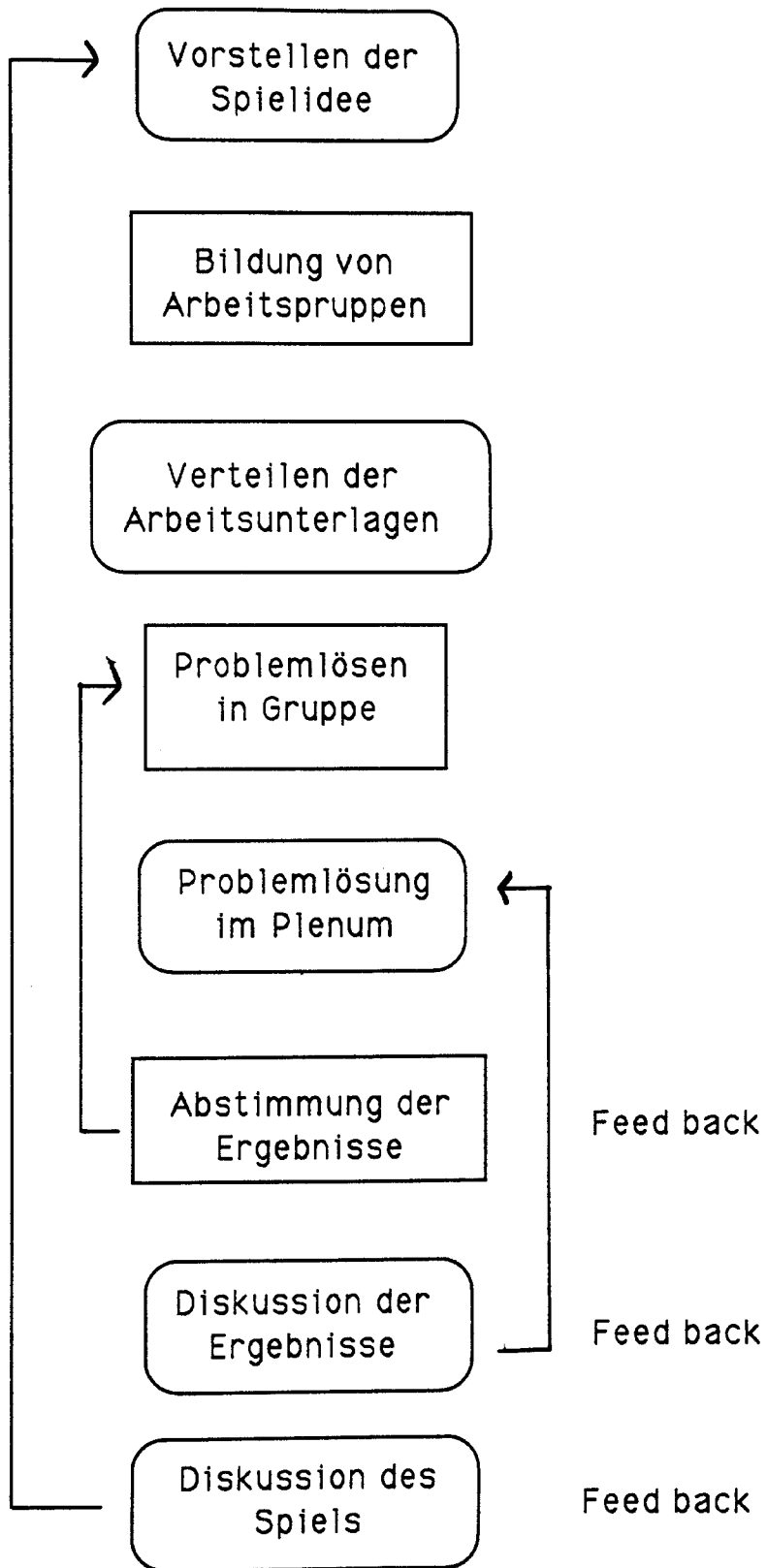


Abb. 2

### Literaturhinweise

- /1/ **Kuhlen**, Rainer: Lehre und Forschung der Informationswissenschaft an der Universität Konstanz. - In: Buder, Marianne (Hrsg.); Rehfeld, Werner (Hrsg.); Seeger, Thomas (Hrsg.): Grundlagen der praktischen Information und Dokumentation. Ein Handbuch zur Einführung in die fachliche Informationsarbeit, 3. völlig neu gefaßte Ausgabe, Band 2, München u.a. 1990, 1073-1099
- /2/ **Manecke**, Hans-Jürgen: Die Technische Hochschule Ilmenau/Thüringen. - In: Nachrichten für Dokumentation 42(1991) 52-53
- /3/ **Oßwald**, Achim: Fort- und Weiterbildung von Dokumentationspraktikern. Konzeption und Ziele des Lehrinstitut für Dokumentation. - In: Informationsspezialisten für Europa. Die Konsequenzen des Europäischen Binnenmarktes für die Ausbildung von Informationsfachleuten, Hannover, 17.-19.10.1989, Proceedings, Hannover 1990, 347-356
- /4/ **Anders**, Anni; **Buder**, Marianne; **Seeger**, Thomas: 25 Jahre Aus- und Weiterbildung in der Deutschen Gesellschaft für Dokumentation e.V. (DGD). - In: Nachrichten für Dokumentation 33 (1982) 237-245
- /5/ **Watson**, David G.; **UNESCO** (Hrsg.): Guidelines for the organization of short courses and workshops on the dissemination of data in science and technology, General Information Programme and UNISIST, Paris 1986 (PGI-86/WS/11)
- /6/ **Marloth**, Heinz: Zur organisatorischen und inhaltlichen Gestaltung von Fortbildungsmaßnahmen, Sankt Augustin 1989 (Arbeitspapiere der GMD 351)
- /7/ **Cronin**, Blaise: Future users: how can we plan, train and educate?. In: Cronin, Blaise; Davenport, Elisabeth (Hrsg.): Post-professionalism: transforming the information heartland. London 1988, 316-328
- /8/ **Vogel**, Elisabeth: Didaktische Hinweise zur Veranstaltungsplanung und -bewertung im Rahmen des Diplom-Aufbaustudiengangs Informationswissenschaft, Bericht 2/84, Konstanz, FB Informationswissenschaft, Konstanz 1984
- /9/ **Lehmler**, Wilfrid: CD-ROM: Endnutzerrecherchen in Literaturdatenbanken. Endbericht eines Projektes an der Bibliothek der Universität Konstanz, Berlin 1990 (dbi-materialien 97)

/10/ **Schwuchow**, Werner: Informationsökonomie. - In:  
Buder, Marianne (Hrsg.); Rehfeld, Werner (Hrsg.);  
Seeger, Thomas (Hrsg.): Grundlagen der  
praktischen Information und Dokumentation. Ein  
Handbuch zur Einführung in die fachliche  
Informationsarbeit, 3. völlig neu gefaßte  
Ausgabe, Band 2, München u.a. 1990, 928-1006

**Gisela Frerk** GMD Forschungsstelle für Informations-  
wirtschaft  
Schönhauser Str. 64  
5000 Köln 51  
0221 / 3767331

**Achim Oswald** Lehrinstitut für Dokumentation in der DGD  
Hanauer Landstr. 126-128  
6000 Frankfurt 1  
069 / 430313

# Das "Online-Unternehmensbild" anhand von Wirtschaftsdaten und informetrischen F&E-Indikatoren

Wolfgang G. Stock

---

*Zusammenfassung.* Das aus öffentlich zugänglichen Online-Datenbanken erstellbare Unternehmensbild setzt sich aus direkt abrufbaren Informationen (u.a. Wirtschaftsnachrichten, Unternehmensdossiers, Bilanzen) sowie aus informetrischen Forschungs- und Entwicklungsindikatoren zusammen. Letztere verdichten verstreute Einzelinformationen insbesondere aus Patentdatenbanken bzw. bibliographischen Datenbanken und lassen thematische Schwerpunkte, Informationsflüsse sowie Entwicklungstrends der F&E-Aktivitäten des Unternehmens und seiner Mitwettbewerber erkennen.

---

## **Problemstellung**

Welche Möglichkeiten bieten online zugängliche elektronische Datenbanken, ein adäquates Bild eines Unternehmens, sei es nun ein Industriebetrieb oder ein Forschungsinstitut, zu vermitteln? In den verschiedensten Datenbanken, ja auch in unterschiedlichen Datenbanktypen können vereinzelt Informationen zu einem Unternehmen verstreut vorkommen. Der Wust von Einzelnachweisen aus bibliographischen Datenbanken, Volltexten, Unternehmensdossiers, Projektdatenbanken, Patentstatistiken usw. muß zu einem Ganzen verdichtet werden.

Zum einen Teil können die Informationen aus den Online-Datenbanken direkt abgerufen werden und bedürfen allenfalls einer knappen redaktionellen Bearbeitung, so etwa Wirtschaftsnachrichten im Abstract oder Volltext oder Bilanzdaten. Zum anderen Teil müssen die relevanten Informationen über das Unternehmen erst auf der Basis von Rohdaten, die in den Online-Datenbanken vorhanden sind, erstellt werden. Zur Verdichtung der Informationen werden Methoden der Informetrie eingesetzt. Insbesondere können auf diese Weise aussagekräftige Indikatoren zur Stellung des Unternehmens in der Forschung und Entwicklung generiert werden. Erst die Vereinigung von Wirtschaftsdaten und F&E-Indikatoren vermittelt ein Gesamtbild eines Unternehmens.

Seinen praktischen Einsatz findet ein solches Online-Unternehmensbild bei der Eruierung der Stellung einer Firma im internationalen Wettbewerb (im Auftrag des Unternehmens selbst oder von der Konkurrenz), bei der Suche nach möglichen Kooperationspartnern oder bei betriebs- bzw. volkswirtschaftswissenschaftlichen Analysen.

## **Unser Beispiel: Die Krones AG in Neutraubling**

Wir wollen die Erstellung eines Online-Unternehmensbildes an einem konkreten Beispiel vorführen. Gefunden haben wir unser Beispiel als Resultat einer Online-Recherche an einer Wirtschaftsnachrichtendatenbank (FITT-Presseindex der GBI) mit einer Suche nach einem deutschen Unternehmen, das in seinem Bereich Weltmarktführer ist (bzw. von dem dies in der Literatur behauptet wird). Die Suche nach

C4WGE.CN and (Weltmarktfuehrer\$ or (weltweit neben fuehrend\$))

ergab eine Reihe von Treffern, die ausschließlich mit dem CO-Feld (Unternehmensname) ausgegeben wurden. Willkürlich wurde eine Firma selektiert: Die Krones AG Hermann Kronseder Maschinenfabrik in Neutraubling (nahe Regensburg in der Oberpfalz).

# Abb. 1: Bilanz der Krones AG

## \*\*\* KENNZAHLEN \*\*\*

AN: 509, FINN 19. 5.91

CR: 823001051

CO: Krones AG H.Kronseder Maschinenfabrik

ST: Boehmenwaldstr. 5

PL: 8402

LO: 8402 Neutraubling

CN: Bundesrepublik Deutschland D C4WGE; C4E Europa

DT: 3112

GC: G2200 Anlagenbau und Maschinenbau, G2000

CU: DM

DI: MIO

LF: AG

TY: Einzel Abschluss 03

YR: 88

CA: 71 MIO

TA: 264 MIO

SA: 382 MIO

EM: 2609

	1983	1984	1985	1986	1987	1988
Cash-Flow	13.9	15.0	17.6	20.2	27.3	38.0
Verschuldungsgrad	2.4	2.0	2.4	2.9	2.7	2.7
Selbstfinanz. quote	2.9	3.0	3.0	2.7	3.0	2.5
Return on Invest.	20.0	18.7	18.2	16.2	15.8	19.4
Eigenkapitalrend.	16.8	14.8	15.6	15.8	12.7	17.6
Gesamtkapitalrend.	5.9	5.8	5.2	5.1	4.8	6.3
Umsatzrentabilitaet	3.3	3.5	3.2	2.8	2.7	3.3
Eigenkapitalquote	28.7	33.4	29.0	25.4	27.1	27.0
Bilanzkurs in %	266.2	193.9	189.1	193.0	273.0	310.4
Anlagendeckung	86.7	97.6	82.5	80.0	77.2	76.0
Kapitalintensitaet	33.1	34.2	35.1	31.7	35.1	35.5
Arbeitsintensitaet	61.9	62.4	62.7	66.3	64.6	64.2
Umschlagskoeff.Lager	5.6	5.7	4.4	4.2	4.5	6.6
Umschlagskoeff.Ford.	4.4	4.2	4.9	4.4	3.8	3.5
Umschlagsdauer Lager	64.4	62.9	82.5	86.0	80.7	54.6
Umschlagsdauer Ford.	82.6	86.1	73.0	81.7	94.2	102.0
Materialaufwandsquo.	28.7	27.3	33.8	36.7	32.4	36.3
Personalaufwandsquo.	49.9	46.9	47.1	45.9	41.3	37.4
Rationalisierungsgr.	4.7	4.4	5.1	5.3	6.5	6.6
Umsatz/Mitarb. in TSD	75.6	76.0	77.9	100.3	115.8	141.8

BILANZ AKTIVSEITE	1983	%	1984	%	1985	%	1986	%	1987	%	1988
Anlagevermoege	39.9	14.5	45.7	15.3	52.7	5.3	55.5	46.5	81.3	15.5	93.9
Sach- u. immat. Anl	30.2	19.2	36.0	19.2	42.9	6.5	45.7	45.7	66.6	11.0	73.9
Finanzanlagen	9.7	.0	9.7	1.0	9.8	.0	9.8	50.0	14.7	36.1	20.0
Umlaufvermoege	74.6	11.8	83.4	12.8	94.1	23.2	115.9	28.9	149.4	13.9	170.1
Vorraete	31.1	6.8	33.2	45.2	48.2	22.6	59.1	12.9	66.7	-13.0	58.0
Forderungen	39.9	13.8	45.4	-6.2	42.6	31.9	56.2	38.4	77.8	39.3	108.4
Wertpapiere	.0	.0	.0	.0	.0	.0	.4	-100.0	.0	.0	.0
Fluessige Mittel	3.6	33.3	4.8	-31.2	3.3	-93.9	.2	2350.0	4.9	-24.5	3.7
Bilanz-Verlust	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
Rechnungsabgrenzung	.1	.0	.1	.0	.1	300.0	.4	75.0	.7	-14.3	.6
Sonstige Aktiva	6.0	-26.7	4.4	-29.5	3.1	.0	3.1	-100.0	.0	.0	.2
Bilanzsumme aktiv	120.6	10.8	133.6	12.3	150.0	16.6	174.9	32.3	231.4	14.4	264.8

BILANZ PASSIVSEITE	1983	%	1984	%	1985	%	1986	%	1987	%	1988
Eigenkapital	34.6	28.9	44.6	-2.5	43.5	2.1	44.4	41.4	62.8	13.7	71.4
Gezeichnetes Kapi.	13.0	76.9	23.0	.0	23.0	.0	23.0	.0	23.0	.0	23.0
Ruecklagen	11.1	21.6	13.5	51.9	20.5	-31.7	14.0	153.6	35.5	17.5	41.7
Ausgleichsposten	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
Gewinn/ Verlust	10.5	-22.9	8.1	-100.0	.0	.0	7.4	-41.9	4.3	55.8	6.7
Sonst. Eigenkapital	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
Sonderp. mit Rueckl.	1.2	-8.3	1.1	-9.1	1.0	-10.0	.9	-22.2	.7	-14.3	.6
Rueckstellungen	50.8	1.4	51.5	7.8	55.5	4.5	58.0	12.8	65.4	33.9	87.6
Verbindlichkeiten	32.4	7.4	34.8	25.9	43.8	44.1	63.1	62.4	102.5	2.5	105.1
Anleihen	.0	.0	.0	.0	.0	.0	.0	.0	20.0	.0	20.0
Bankverbindlichk.	16.0	-38.1	9.9	-61.6	3.8	286.8	14.7	-7.5	13.6	142.6	33.0
Vb. a. Lief.u. Lei.	6.2	19.4	7.4	79.7	13.3	-32.3	9.0	53.3	13.8	44.2	19.9
Uebrig Verb.	10.2	71.6	17.5	52.6	26.7	47.6	39.4	39.8	55.1	-41.6	32.2
Davon langfr. Verb.	.0	.0	.0	.0	.0	.0	.0	.0	NA	.0	.5
Rechnungsabgrenzung	.2	-100.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
Sonstige Passiva	1.4	14.3	1.6	287.5	6.2	37.1	8.5	-100.0	.0	.0	.1
Bilanzsumme passiv	120.6	10.8	133.6	12.3	150.0	16.6	174.9	32.3	231.4	14.4	264.8

GEWINN- / VERLUSTRECHNUNG	1983	%	1984	%	1985	%	1986	%	1987	%	1988
Umsatzerloese	173.8	9.3	189.9	10.7	210.2	17.7	247.5	20.2	297.4	28.6	382.6
sonst. bet. Ertraege	NA	.0	NA	.0	NA	.0	NA	.0	NA	.0	NA
Materialaufwand	49.9	4.0	51.9	36.8	71.0	28.0	90.9	6.1	96.4	44.1	138.9
Personalaufwand	86.8	2.5	89.0	11.3	99.1	14.7	113.7	8.1	122.9	16.4	143.1
Abschreibungen	8.1	3.7	8.4	28.6	10.8	22.2	13.2	46.2	19.3	31.6	25.4
sonst. betr. Aufwand	NA	.0	NA	.0	NA	.0	NA	.0	NA	.0	NA
Ertr. aus Beteilig.	.0	.0	.0	.0	.0	.0	.0	.0	2.8	25.0	3.5
Zinsertraege + Aehnli	1.1	.0	1.1	-54.5	.5	100.0	1.0	-60.0	.4	125.0	.9
Zinsaufwand + Aehnli	1.3	-7.7	1.2	-16.7	1.0	90.0	1.9	68.4	3.2	28.1	4.1
Ausserord. Ergebnis	NA	.0	NA	.0	NA	.0	NA	.0	NA	.0	NA
Steuern	10.2	-2.0	10.0	-3.0	9.7	-15.5	8.2	12.2	9.2	44.6	13.3
Saldo so. Ert./Aufw.	-12.8	-86.7	-23.9	48.5	-12.3	-10.6	-13.6	-205.9	-41.6	-19.2	-49.6
Jahresueberschuss	5.8	13.8	6.6	3.0	6.8	2.9	7.0	14.3	8.0	57.5	12.6
Beschaeftigte	2300	8.7	2500	8.0	2700	-8.6	2467	4.1	2568	5.1	2698

## Das Online-Unternehmensbild. Teil 1: Wirtschaftsdaten

Einen ersten Einblick in das Unternehmen gestatten Datenbanken mit Wirtschaftsnachrichten. Recherchen sind hierbei etwa möglich bei den Predicast-Datenbanken oder beim schon erwähnten FITT-Presseindex. Hier erfahren wir, daß Krones ein Hersteller von Maschinen für Abfüll- und Verpackungstechnik ist und daß unser Unternehmen 1990 einen Umsatz von 612 Mill. DM erwirtschaftete sowie daß dies eine Umsatzsteigerung gegenüber dem Vorjahr um 21% bedeutet.

Einige der Meldungen erscheinen uns derart interessant, daß es zweckmäßig erscheint, die Volltexte der Nachrichten einzusehen. Hierzu wechseln wir in Volltextdatenbanken der Tages- bzw. Wirtschaftspresse. Infragekommen u.a. Reuters Textline, NEXIS von Mead Data Central, die Newspapers-Sammlungen bei Dialog, die FT Profile-Volltextdatenbanken sowie die deutsche Wirtschaftspresse bei GENIOS. Ein Artikel im Handelsblatt (vom 29.1.1991) zeichnet ein recht umfassendes Bild der Unternehmenssituation 1990 mit Ausblicken auf die Entwicklung im Jahre 1991. Recherchiert wurde im File HB bei GENIOS:

AB gw MUENCHEN. Die Krones AG Hermann Kronseder Maschinenfabrik, Neutraubling, verzeichnete 1990 ein zweistelliges Wachstum und eine weiterhin gute Ertragslage. Saemtliche Konzerngesellschaften arbeiteten mit Gewinn. Firmenchef Hermann Kronseder kuendigt wieder eine Dividende von 8 DM je Stamm- und 10 DM je Vorzugsaktie an. Ausserdem soll das Eigenkapital durch Ruecklagendotierung gestaerkt werden.

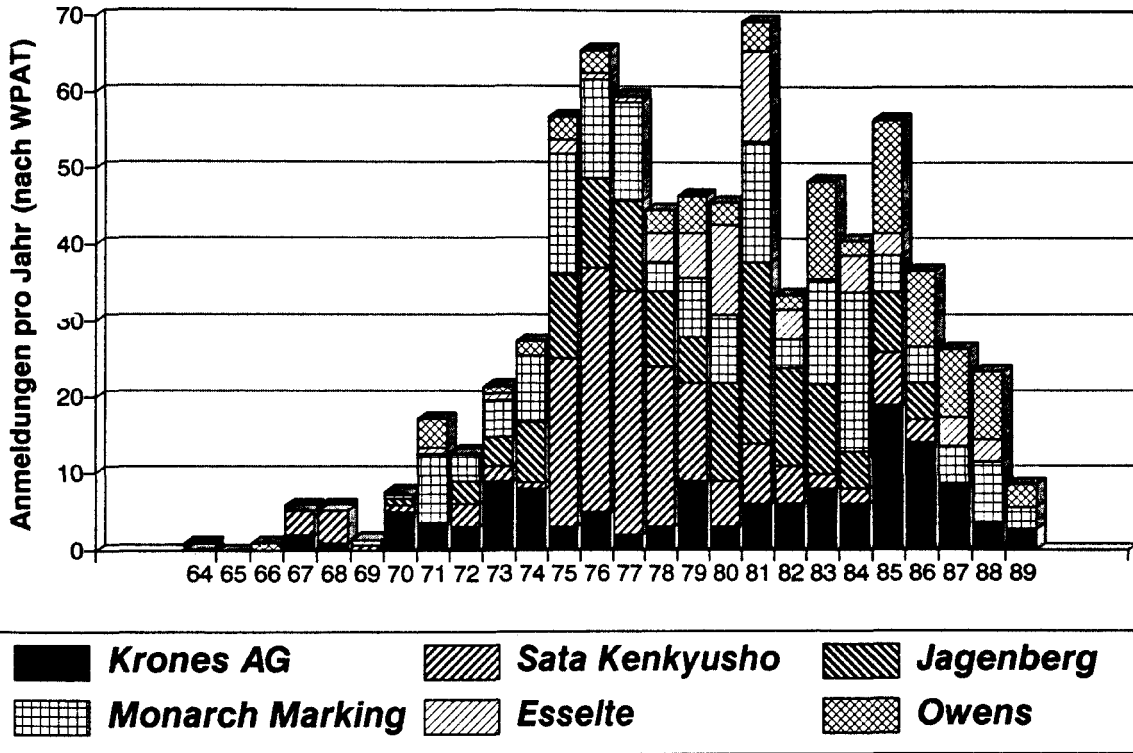
TX Der Umsatz der Krones AG stieg 1990 um 21% auf 612 (505) Mill. DM. Der Exportanteil lag wieder mit 75 (77)% extrem hoch. Der Weltumsatz kletterte um 20% auf 820 Mill. DM. Aufgrund der Wechselkursveraenderungen wuchs das Auslandsgeschaefit langsamer als der Inlandsumsatz. Hauptumsatztraeger war mit 50% der Bereich Etikettiermaschinen, gefolgt von Leerflaschen-Inspektionsmaschinen. Die groesste Expansion verzeichneten Flaschenreinigungsmaschinen mit plus 50% und Flaschen-/Dosen-Fuellmaschinen mit 42%. Aufgrund einiger beachtlicher Geschaefitabschluesse stieg der Auftragseingang um 35% auf 603 Mill. DM. Allein aus den neuen Bundeslaendern kamen Auftraege in Hoehe von ueber 50 Mill. DM. Besonders stark expandierten die Auftraege fuer Fuellmaschinen in Osteuropa (+61%). Krones will nach der Gruendung einer Tochter in Polen und der Errichtung von Vertriebsbueros in Berlin und Moskau demnaechst eine Niederlassung in der CSFR aufbauen. Ende 1990 lag der Auftragsbestand mit 430 Mill. DM um 48% ueber dem Vorjahresniveau. Die volle Kapazitaetsauslastung fuer die naechsten zwelf Monate sei damit gesichert. Die Krones-Belegschaft stieg um 15% auf 3897 Mitarbeiter im Inland. Weltweit erhoehrte sie sich um 16% auf 5036 (4335) Beschaeftigte. Die Investitionen stiegen 1990 um rund 30% auf ueber 60 Mill. DM und auf 20 Mill. DM bei den Auslandstoechtern. Bei Krones Zierk, Flensburg, hat sich die Kapazitaet innerhalb von drei Jahren verdreifacht.

Auch fuer 1991 erwartet Krones ein zweistelliges Wachstum und eine befriedigende Ertragsentwicklung. Fuer Forschung und Entwicklung werden wieder 8% ausgegeben. Der anhaltende Aufwaertstrend spiegelt sich in den fuer 1991 geplanten Investitionen von 70 Mill. DM in der AG und 85 Mill. DM im Konzern wieder.

Eine systematische, wenngleich nicht so aktuelle Ergänzung zu Wirtschaftsnachrichten bieten Unternehmensdossiers, wie sie z.B. Dun & Bradstreet, Hoppenstedt oder Creditreform offerieren. Angeboten werden Unternehmensdaten wie Anschrift und Telefon, Bankverbindungen, Management, Aufsichtsräte, Aktienbesitz, Zweigwerke, Beteiligungen, Umsatz, Beschäftigtenzahl usw. Wirtschaftsdaten eines Unternehmens werden plastisch dargestellt an seiner Bilanz. Bilanzdatenbanken, etwa EcoSaldo bei Ecodata oder die Creditreform-Bilanzdatenbanken, der wir die Krones-Bilanz (Abb. 1) entnommen haben, gestatten einen Online-Zugriff auf die gewünschten Zahlen. Für die Jahre 1983 bis 1988 verfügen wir nunmehr über eine ausreichende Basis zur Beurteilung der geschäftlichen Lage unseres Unternehmens. Rechnen wir die aus den Wirtschaftsnachrichten erhaltenen Umsatzzahlen mit ein, so liegt für den Zeitraum 1983 bis 1990 eine lückenlose Umsatzkurve vor.



**Abb. 2: Patente für Etikettiermaschinen. Die führenden Unternehmen**



Quelle: ORBIT Search Service; File WPAT

Weitere Datenbanken zu unternehmensrelevanten Wirtschaftsinformationen liegen online vor. Zu denken ist an Selbstdarstellungen von Firmen, abgefaßt etwa als Jahresberichte (Beispiel: ANNR / Company Annual Reports bei KIWINET) oder als Text von Hauptversammlungsreden von Aktiengesellschaften (Beispiel: HVR / Hauptversammlungsreden deutscher Aktiengesellschaften bei der GBI). Relevanz haben auch Produktdatenbanken (z.B. Thomas New Industrial Products bei DIALOG), Datenbanken zu Joint Ventures (z.B. Mergers and Acquisitions International bei FT Profile), Handelsstatistiken (z.B. Tradstat World Trade Statistics bei DATA-STAR) und diverse andere Datenbanken (Butcher 1988, Koenig 1988, Parkinson 1991).

## **Das Online-Unternehmensbild. Teil 2: Informetrische Indikatoren zu Forschung und Entwicklung**

Unternehmensspezifische Informationen zum Stand der F&E-Aktivitäten sowie zum internationalen Vergleich sind in der Regel nicht direkt abrufbar in Online-Datenbanken vorhanden. Sehr wohl gespeichert sind aber Patente der Firma sowie Literatur von Mitarbeitern in Fachzeitschriften und Büchern. Die so erhaltenen Einzelinformationen bedürfen einer zusammenfassenden Informationsverdichtung, sei es nun thematischer Art (Konstellation der F&E-Schwerpunkte) oder formaler Art (z.B. Stellung eines Autors in der entsprechenden scientific community, bevorzugte Publikationsorgane, Beschreibung der Informationsflüsse ins Unternehmen hinein bzw. aus diesem heraus). Die zur Informationsverdichtung einsetzbaren informetrischen Methoden umfassen einfache Rangordnungen ("Hitparaden"), Methoden der deskriptiven Statistik sowie der Clusteranalyse.

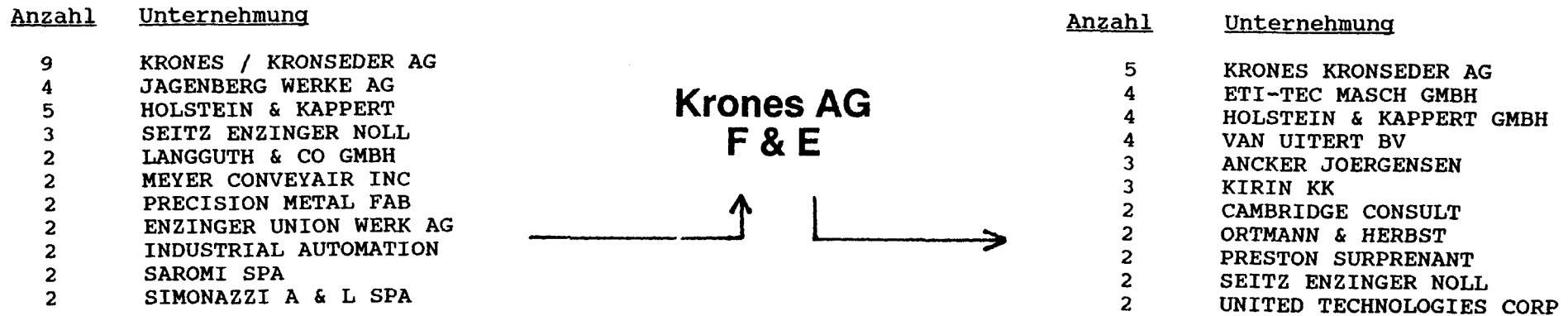
Die Informetrie zerfällt - analytisch betrachtet - in zwei große Gruppen. Zum einen geht es darum, mittels quantitativer Verfahren Regelmäßigkeiten oder Gesetzmäßigkeiten zu eruieren, die der fachlichen Information zu eigen sind. Datenbanken dienen hierbei als Experimentierfeld für systematische Untersuchungen. Ergebnisse sind Gesetze des "Verhaltens" von Informationen in Raum und Zeit (Bonitz 1986a, 1986b), z.B. das Bradfordsche Gesetz zur Konzentration von Zeitschriften zu einem Fachbereich, das Gesetz der "Halbwertszeit" wissenschaftlicher Informationen oder das Garfieldsche Gesetz zur Konzentration von Zitationen. Die erste Gruppe informetrischer Aktivitäten wollen wir "nomothetische Informetrie" benennen (Egghe/Rousseau 1990, 291 ff.). Zum andern geht es darum, informetrische Verfahren dazu einzusetzen, um konkrete Objekte zu vermessen. Diese Objekte können bestimmte Autoren, Zeitschriften, Länder, Wissenschaftsdisziplinen, aber auch Unternehmen sein. Die zweite Gruppe sei die "deskriptive Informetrie". Klar ist, daß in der konkreten informetrischen Arbeit beide Gruppen nicht zu trennen sind, sind doch Teile der deskriptiven Informetrie Rohdaten für die nomothetische Betrachtung. - Macht man seine Studien ausschließlich an öffentlich zugänglichen Online-Datenbanken, so hat sich dafür die Bezeichnung "Online-Informetrie" etabliert (Stock 1991). - Unsere Studie ist dem Bereich der deskriptiven Online-Informetrie zuzuordnen.

Bei informetrischen Recherchen zum Unternehmensbild arbeiten wir mit zwei Datenbanktypen: zum einen mit Patentdatenbanken, zum anderen mit bibliographischen Datenbanken auf dem Fachgebiet des Unternehmens. Ergänzt werden die Analysen ggf. - soweit dort Einträge vorhanden sind - durch Zitationsdatenbanken.

### **Online-Patentstatistik**

Bei unserer Online-Patentstatistik benutzen wir durchgängig die World-Patents-Index-Datenbank beim Host ORBIT (Orbit 1985). Ein erster und recht einfacher F&E-Indikator eines Unternehmens ist seine Patentanmeldeaktivität. Unter Anwendung des GET-SHOW-Befehls auf Derwent-Patentdatenbanken erhält man eine Zeitreihe der Patentanmeldungen pro Jahr. Die Kronos AG zeigt in der Datenbank des World Patents Index eine rege Anmeldeaktivität. In den 60er Jahren liegen etwa ein bis zwei Patente pro Jahr vor, in den 70er und 80er Jahren steigt der Wert auf ca. 15 Anmeldungen im Jahr mit einem Höhepunkt (1985) von knapp 30 Patenten. Insgesamt weist der World Patents Index 244 Patente der Kronos AG nach.

**Abb. 3: Der Technologiefuß von und zur Krones F&E**

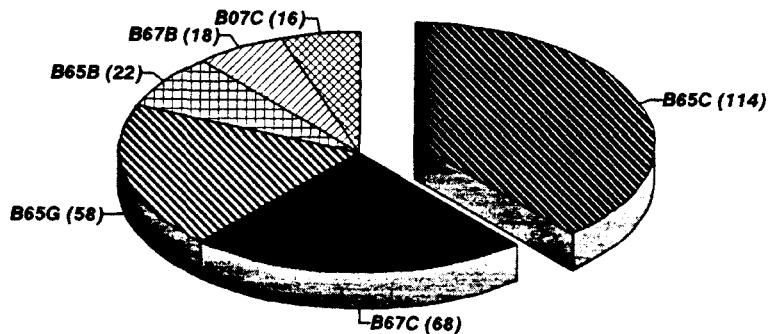


**Technologieinput:** 177 Patente, die in Patenten von Krones (mit EP- und PCT-Anmeldung) zitiert werden.

**Technologieoutput:** 66 Patente (mit EP- und PCT-Anmeldung), die Patente von Krones zitieren.

Quelle: ORBIT Search Service; File WPAT

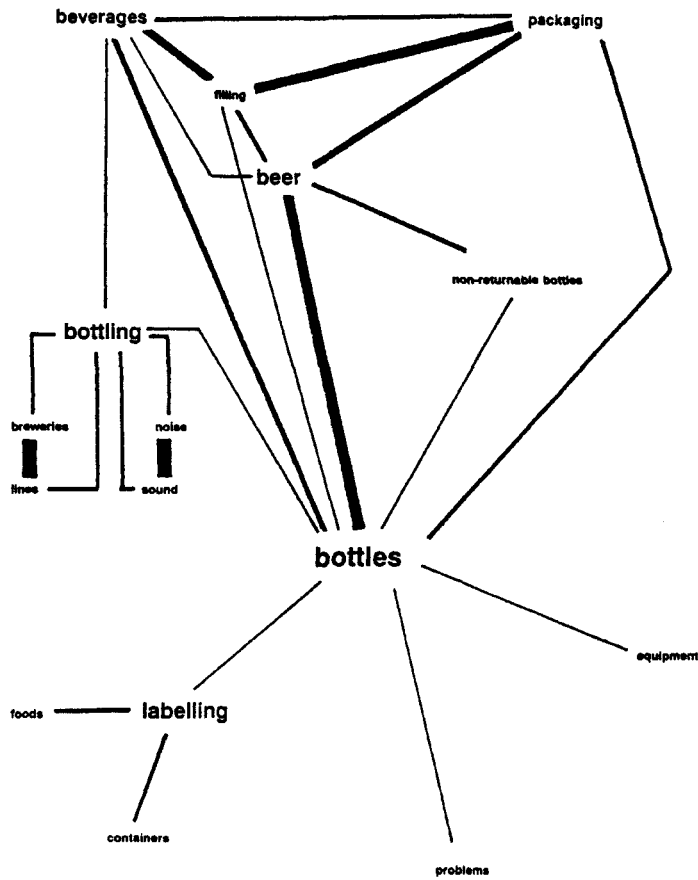
Ein thematischer Wissenschaftsindikator ist die rangierte Ausgabe der Patentinformationen nach Patentklassen (nach der Internationalen Patentklassifikation IPK). Hier wird vom GET-RANK-Befehl bei ORBIT Gebrauch gemacht, indem über die Treffer für Krones die thematische Statistik läuft. Demnach entfallen von den 244 Krones-Patenten insgesamt 114 auf IPK B65C (Etikettiermaschinen), gefolgt von B67C (Abziehen von Flüssigkeiten; 68 Patente), B65G (Transport- und Lagervorrichtungen; 58 Patente), B65B (Verpackungsmaschinen; 22 Patente), B67B (Öffnen/Verschließen von Flaschen; 18 Patente) und B07C (Sortieren; 16 Patente); im Überblick (Mehrfacheinordnungen von Patenten in verschiedenen Klassen sind möglich):



Für die eruierten F&E-Themen kann man nunmehr an den Patentdatenbanken Konkurrenzvergleiche durchführen. Eine Recherche nach IPK B65C und eine Hitparade, sortiert nach Anmeldern, gibt einen Überblick über die weltweit führenden F&E treibenden Unternehmen im Bereich der Etikettiermaschinen. Insgesamt sind im World Patents Index 3.179 Patente unter IPK B65C eingeordnet. Davon entfallen auf Sato Kenkyusho 135 Anmeldungen, auf Jagenberg 130, auf Krones 114, auf Monarch Marking 95, auf Owens-Illinois 67 und auf Esselte Pendaflex 51. Wenn man bedenkt, daß die japanischen Anmeldeaktivitäten stärker detailorientiert sind als international üblich, ist die Spitzenposition von Sato Kenkyusho durchaus anzuzweifeln. Wichtiger als die Absolutzahl der Patentanmeldungen in einem Technikfeld ist die zeitliche Entwicklung der Aktivitäten. Für die sechs patentreichsten Unternehmen in der Branche der Etikettiermaschinen wird die Entwicklung in Abb. 2 skizziert. Demnach liegt der Höhepunkt der F&E auf dem Gebiet der Etikettiertechnik in der Zeit zwischen 1975 und 1985. Danach ist ein Aktivitätsabfall zu beobachten. Das japanische Unternehmen Sato Kenkyusho spielte zwischen 1975 und 1979 eine große Rolle in diesem Technikbereich, aber - ebenso bei Jagenberg - ab 1987 sind keine Nachweise mehr vorhanden. Auch und gerade in den letzten Jahren sind Monarch Marking und Owens-Illinois stark zu beachten, während Esselte Pendaflex auf einem recht niedrigen quantitativen Niveau Patentaktivitäten vorweist. Krones hat seit dem Anmeldehöhepunkt 1985 an Schwung verloren. Möglicherweise wurden die F&E-Bemühungen von der Etikettiertechnik in andere Forschungsbereiche verlagert.

Wie fließen die wissenschaftlich-technischen Informationen in unserem Technikbereich? Zur Untersuchung von Informationsflüssen folgen wir der Methode der Zitationsanalyse. Im Rahmen der Online-Patentstatistik können wir jedoch nicht alle Patente in unsere Studie miteinbeziehen, weil in den Patentdatenbanken Zitate nur für Patente des Europäischen Patentamtes und für Weltpatente notiert werden. Der Technologiestrom von und zur Krones AG konnte an 177 von Krones zitierten und an 66 Krones-Patente zitierenden Patenten eruiert werden. Der Technologieinput wurde über ein Ranking des IPK-Feldes derjenigen Patente errechnet, die von Krones-Patenten (mit EP- bzw. PCT-Anmeldung) zitiert werden. Thematisch sieht es so aus, daß die Krones F&E bevorzugt IPK B65G (Transport- und Lagervorrichtungen) zitiert (in 17 Patenten), gefolgt von B65C (Etikettiermaschinen; in 13 Patenten) und B67C (Abziehen von Flüssigkeiten). Zum Errechnen des Technologieoutputs wurde analog vorgegangen. (Für Interessierte folgt eine detaillierte Beschreibung der Strategie: Es wurden alle Krones-Patente gesucht. Zwischengespeichert wurden nur die Patentnummern (PRT SELECT). Der Feldname PN (patent number) wurde mit dem Befehl REQUAL umdefiniert zu CT (Cited Patents). Nach allen CT=<Treffermenge der gefundenen Patentnummern> wurde mittels SELECT

**Abb. 4: Thematisches Cluster der Krones-Publikationen im Bereich Lebensmitteltechnologie**



Rangplatz	Thema	rel.Häufigkeit
1	bottles	55,00
2	bottling	40,00
3	beer	35,00
3	labelling	35,00
5	beverages	25,00
6	packaging	20,00
7	containers	15,00
7	equipment	15,00
7	noise	15,00
7	sound	15,00
11	breweries	10,00
11	filling	10,00
11	foods	10,00
11	lines	10,00
11	non-returnable bottles	10,00
11	problems	10,00

N = 20

gesucht. Bei der erhaltenen Treffermenge fand ein Ranking mit dem Befehl GET RANK IC (1-4) nach den Vierstellern gemäß IPK statt.) Krones-Patente werden vor allem im Kontext von B65G (26 zitierende Patente), B67C und B65C (jeweils 14 zitierende Patente), aber auch nicht unerheblich bei G01N (Analysieren von Stoffen; in 9 zitierenden Patente) erwähnt.

Jetzt interessiert uns der Technologiefuß relativ zu Unternehmen. Patente welcher Unternehmen zitiert die Krones F&E? und: Welche Unternehmen zitieren Krones-Patente? Die Suchstrategie ist wie oben beschrieben, lediglich das Ranking beim IC-Feld wurde mit dem Auflisten des PA-Feldes (patent assignee) gewechselt. Die Ergebnisse zeigt Abb. 3. Die Krones F&E bedient sich vor allem bei eigenen Vorarbeiten, aber auch bei der Konkurrenz (etwa Jagenberg, Holstein & Kappert oder Seitz/ENZINGER/NOLL). Im Gegenzug nutzt auch die Konkurrenz in Patenten publizierte Ergebnisse von Krones (so ETI-TEC, Holstein & Kappert, van Uiter, Ancker Joergensen oder Kirin).

### **Online-Infometrie bei bibliographischen Datenbanken**

Über ein Ranking der Namen der Patentinhaber der Krones-Patente kommen wir zu Namen von führenden Ingenieuren bzw. Wissenschaftlern beim untersuchten Unternehmen. Diese Namen sowie die teilweise mit gespeicherten Affiliationsangaben von Autoren sind für uns der Einstieg für Recherchen in bibliographischen Datenbanken. Die Namen mit den meisten Patenten sind bei der Krones AG H.Kronseder, H.Winter, E.Schneider, W.Weiss, R.Zodrow, K.Dullinger, E.Eder und G.Gau.

Es kann zunächst - insbesondere bei relativ kleinen Nachweismengen (bis ca. 30 Dokumentations-einheiten) - sinnvoll sein, die Literaturstellen und hierbei vor allem die Abstracts zu studieren, um detailliert F&E-Aktivitäten zu verfolgen. So erfahren wir beispielsweise aus einem Nachweis bei der Dokumentation Maschinenbau (DOMA) beim FIZ Technik, daß Krones schon 1986 den CIM-Einsatz vorbereitete und in diesem Zusammenhang das CAD-System MEDUSA einsetzte.

Zum Erkennen der Schwerpunkte nützen allerdings die einzelnen Nachweise wenig; gefragt ist auch hier eine infometrische Informationsverdichtung. Die themenanalytischen Methoden setzen bei bibliographischen Datenbanken an (Stock 1990). Außer ORBIT bieten auch andere Hosts infometrische Software an, z.B. STN, Questel oder DIMDI. Zur statistischen Aufbereitung von Daten anderer Hosts oder zur Durchführung von Clusteranalysen dient Spezialsoftware, z.B. LBase. Die hier vorgestellten Ergebnisse wurden im deskriptiv-statistischen Teil bei DIMDI mit dem Befehl REPORT=STAT (Kaiser 1991) sowie im clusteranalytischen Teil mit Hilfe von LBase (Werba/Stock 1989) erarbeitet.

Unser Beispiel betrifft das Online-Unternehmensbild der Krones AG im Bereich der Lebensmitteltechnologie. Große Relevanz auf diesem Wissenschafts- bzw. Technikfeld hat die Datenbank Food Science and Technology Abstracts (FSTA). Andere Blickwinkel sind ebenso möglich, so etwa hinsichtlich der thematischen Konstellation der Publikationen von Krones-Mitarbeitern im Bereich Maschinenbau, eruiert z.B. bei der DOMA. Während das DOMA-Online-Bild von Krones die Themen Etikettiermaschine, Flasche, Etikett, Etikettierstation, Klebstoff bis hin zur Optoelektronik bei Flaschenbehandlungsmaschinen in den Vordergrund rückt, zeigt FSTA zentral die Themen Flasche, Abfüllen, Bier, Etikettieren und Getränke (Abb. 4). Die Rangtabelle der Themen erstellt der DIMDI-Befehl REPORT=STAT; gelaufen ist die Statistik über das CT-Feld bei FSTA. Zur Berechnung der thematischen Nähe der Deskriptoren wurde das Download (nur CT-Feld) nach LBase importiert und mit dem PERM-Befehl weiterverarbeitet. Für alle Themen, die mindestens zweimal gemeinsam in einem Nachweis vorkommen, wird ein Maß zwischen größer 0 und 1 errechnet, wobei 1 bedeutet, daß beide Themen immer gemeinsam vorkommen. Im Cluster wird die Bindungsstärke zwischen den Themen durch die Strichstärke angedeutet (der Strich zwischen beer und bottles entspricht einer Stärke von 0,500). Die Druckgröße der Deskriptoren entspricht der Wichtigkeit (relativen Häufigkeit) des Themas in der betrachteten Dokumentenmenge. Demnach gibt es im Cluster des Krones-Bildes bei FSTA vier Subcluster. Alle sind mit dem zentralen Thema "Flaschen" verknüpft. Zum einen geht es um das Etikettieren von Lebensmitteln und Behältern (in Abb. 4 links unten), zweitens um Einwegflaschen und Bier (rechts Mitte), drittens um die Beziehungen zwischen Getränken, speziell

Bier, und dem Verpacken bzw. Füllen derselben (desselben) in Flaschen (oben), viertens um das Abfüllen, und dies in zwei Aspekten: bzgl. von Produktionslinien und Brauereien (links außen) sowie bzgl. des dabei entstehenden Lärms (links innen).

Für als wichtig erkannte Themen können weitere Rangtabellen erstellt werden. Über eine Suche nach der Gesamtliteratur zum Deskriptor BOTTLING in FSTA und einer rangierten Ausgabe nach Autorennamen können die Krones-Autoren in ihrer Stellung innerhalb der fachspezifischen Publikationsaktivitäten bewertet werden. Unter den ersten zwölf Rangplätzen in dieser Aufstellung finden sich zwei Mitarbeiter von Krones (unter insgesamt 1.009 Verfassern).

Bei Rangtabellen insbesondere zu Namen kann es immer wieder dazu kommen, daß unterschiedliche Schreibweisen die Ergebnisse verzerren. So ist unser Beispielunternehmen etwa unter Krones AG sowie unter Krones Kronseder AG im Feld PA beim World Patents Index vertreten. Solange die Datenqualität nicht völlig adäquat ist (Eisberg 1991), hilft nur die intellektuelle Nachbereitung der Ergebnisse.

Die Online-Infometrie in der Form, gewisse Retrievalergebnisse wieder als Suchargumente zu verwenden (und diesen Vorgang möglicherweise mehrfach zu wiederholen), erinnert durch das "Navigieren" durch die betreffende(n) Datenbank(en) sehr an Hypertext (Gloor 1990; Maurer 1991). Bei jedem Navigationsschritt entstehen (nach der Datenaufbereitung) Rangtabellen, Cluster, Informationsflußdiagramme usw. zunächst zum Ausgangsthema (etwa zu einem Unternehmen) und dann zum Umfeld (etwa zur Konkurrenz): Jeder gefundene Aspekt kann letztlich wieder Ausgangspunkt weiterer vertiefender (aber ggf. auch vom Thema wegführender) Analysen werden.

Wolfgang G. Stock  
ifo Institut für Wirtschaftsforschung  
Arbeitsgruppe Information / Dokumentation  
Poschingerstr. 5  
D-8000 München 86

## **Literatur**

(Bonitz 1986a)

Manfred Bonitz: Wissenschaftliche Information und wissenschaftliches Verhalten, Berlin: Zentralinstitut für Information und Dokumentation, 1986

(Bonitz 1986b)

...: Holographie- und Tempoprinzip : Verhaltensprinzipien im System der wissenschaftlichen Kommunikation, Informatik 33 (1986), 191-193

(Butcher 1988)

Helen Butcher: Business and economics : United Kingdom, in: C.J.Armstrong; J.A.Large (Hrsg.): Manual of online search strategies, Aldershot [u.a.]: Gower, 1988, 559-598

(Egghe/Rousseau 1990)

Leo Egghe; Ronald Rousseau: Introduction to informetrics : quantitative methods in library, documentation and information science, Amsterdam, New York, N.Y., Oxford, Tokyo: Elsevier, 1990

(Eisberg 1991)

Anita Eisberg: Rechnergestützte Überprüfung der Datenqualität in klinischen Studien, Praxis Medizinischer Dokumentation 11 (1991) 2, 19-22

(Gloor 1990)

Peter A. Gloor: Hypertext und Hypermedia : von theoretischen Konzepten zur praktischen Anwendung, Berlin; Heidelberg: Springer, 1990 (Informatik-Fachberichte ; 249)

(Kaiser 1991)

Dieter Kaiser: Statistische Auswertungen in Literaturdatenbanken, GRIPS-News / Deutsches Institut für Medizinische Dokumentation und Information (DIMDI) Nr. 02/91 (Juni 1991), 1-3 - 1-6

(Koenig 1988)

Mike Koenig: Business and economics : United States, in: C.J.Armstrong; J.A.Large (Hrsg.): Manual of online search strategies, Aldershot [u.a.]: Gower, 1988, 599-622

(Maurer 1991)

Hermann Maurer (Hrsg.): Hypertext Hypermedia '91 : Tagung der GI, SI und OCG ; Graz 27./28. Mai 1991, Berlin; Heidelberg: Springer, 1991 (Informatik-Fachberichte ; 276)

(Oppenländer/Faust 1990)

Karl Heinrich Oppenländer; Konrad Faust: Patentanmeldungen als Frühindikatoren, in: Hermann J. Schuster (Hrsg.): Handbuch des Wissenschaftstransfers, Berlin [u.a.]: Springer, 1990, 239-249

(Orbit 1985)

ORBIT Search Service: World Patents Index database manual (TM-5859/000/01), McLean: Orbit, 1985

(Parkinson 1991)

Helen Parkinson: Online business and company databases 1991, London: Aslib, 3.Aufl. 1991

(Persson 1986)

O. Persson: Online bibliometrics, Scientometrics 8 (1986), 69-75

(Persson 1988)

...: Measuring scientific output by online techniques, in: A.F.J. van Raan (Hrsg.): Handbook of quantitative studies of science and technology, Amsterdam; New York; Oxford; Tokyo: North-Holland, 1988, 229-252

(Stock 1990)

Wolfgang G. Stock: Themenanalytische informetrische Methoden, in: Mechtild Stock; Wolfgang G. Stock: Psychologie und Philosophie der Grazer Schule : eine Dokumentation, Amsterdam, Atlanta: Rodopi, 1990, 7-31

(Stock 1991)

...: Möglichkeiten der Online-Informetrie bei Themenanalysen, Praxis Medizinischer Dokumentation 11 (1991) 4

(Werba/Stock 1989)

Helmut Werba; Wolfgang G. Stock: LBase - Ein bibliographisches und faktographisches Informationssystem für Literaturdaten, in: Wolfgang L. Gombocz; Heiner Rutte; Werner Sauer (Hrsg.): Traditionen und Perspektiven der Analytischen Philosophie, Wien: Hölder-Pichler-Tempsky, 1989, 631-647



# BIOTECHNOLOGY INFORMATION SERVICE and its Value for Research

Eleonore Poetzsch

## SUMMARY

In the first part, the paper gives a survey of the situation in the field of biotechnology information. The results of an analysis of the **information market** regarding factual and bibliographic biotechnology information have been compiled in directories. Special emphasis is laid on **users' needs**, and on the use of databases for different purposes. Some aspects of **quality** and further development of databases including possibilities for international co-operation are discussed.

In the second part, the establishment of the **"Biotechnology Information Service"** is presented. The **"Biotechnology Information Service"** has the following tasks: 1. Online access and offline using, 2. Establishment of special databases, 3. Publications. The database **"Who-What-Where"** with information on institutions and companies working in various fields of biotechnology in the former GDR and Eastern Europe is presented. Some conclusions relating to the further development of biotechnology information, especially **information brokerage** in this field, are derived.

## 1. THE STATE-OF-THE-ART IN THE FIELD OF BIOTECHNOLOGY INFORMATION

The availability of high quality, up-to-date and comprehensive information is an important requirement in biotechnology and its applications in ever wider fields of medicine, ecology, agriculture and food industry. The last ten years have seen a considerable and at the same time necessary increase in the amount of available information on biotechnology.

We have analysed the **information market** regarding factual and bibliographic databases relevant to basic and applied fields of biotechnology. The results have been compiled in two directories [1][2]; and **the users' needs** in selected areas of biotechnology, especially in the field of genetic engineering.

Starting out from the results of these investigations we can state: The consequences of the rapid development in biotechnology with the increasing growth of data on the one hand, and the possibilities of information technology on the other hand, are **very large databases** containing different kinds of information (bibliographic, factual, referral and full-text information) and **highly sophisticated software** for processing and using of these databases.

**As a result a new discipline has been developed: BIOINFORMATICS.** The development of bioinformatics is connected with

- the organization of a comprehensive **information service** including
  - . various possibilities for the availability of databases (magnetic tape, floppy disk, CD-ROM);
  - . communication networks (online access, electronic mail, computer conferencing);
  - . special information processing services;
- the enhancement of **international co-operation** with the purpose of increase and coordination of the information exchange, and establishing of only a few, but powerful information centres for collection, processing, and exchange of data;
- high governmental and industrial **investment** to provide the pre-conditions necessary for the establishment of databases at which there are great differences between funding of bioinformatics in the USA and Europe on the one hand, and between Western Europe and Eastern Europe on the other hand [3];
- very different **information needs** of users in research and industry influenced by the multidisciplinary character of biotechnology, and the different kinds of necessary information (bibliographic information; factual information, e.g. sequences, structures, spectra, as well financial and business information; referral information, etc.).

The new discipline **"Bioinformatics"** is an essential part of biotechnology that underpins not only research and development but also commercial, legal, and regulatory activities.

The enlargement of the information market in biotechnology is also connected with specific problems. Some of them are of general character, and another are connected with the former GDR structures:

#### **1. Insufficient Appreciation of "Information" as a "Knowledge" Resource**

*Reasons:*

- an underdeveloped information awareness;
- we have not learned to use "information" as a resource comprehensive and to transform the provided information into knowledge.

#### **2. The Necessary Pre-conditions for the Setting-up and Use of Databases are Insufficiently Provided**

*Reasons:*

- the willingness to national and international co-operation does not meet the requirements;
- the effort in setting-up databases is underrated and thus there is a lack of provision of the necessary manpower, financial resources and equipment;
- the information management is underdeveloped.

### 3. The Interfaces Between Research and Information are Insufficiently Developed

*Reasons on the side of information:*

- the "classical" understanding of documentation continues to prevail;
- the necessary pre-conditions for a comprehensive information service do not exist.

*Reasons on the side of research:*

- a lack of possibilities as a consequence of inadequate pre-conditions or insufficient data quality, but also unwillingness to feed research findings into databases and thus make them more widely available;
- occasionally a dangerous ignorance is exhibited with regard to information.

### 4. The Guarantee of the Quality of the Database

concerning

- **the actuality of the database** determined by the time it takes to make new information available on a database, and by the possibility of rapid access to the database;
- **the reliability and accuracy of data** connected with the necessity
  - . to integrate evaluation algorithms,
  - . to expand the syntactic and semantic representation possibilities,
  - . to formulate requirement characteristics;
- **the complexity of data and databases** connected with the necessity
  - . to crosslink different (kinds of) information which are part of different (kinds of) databases,
  - . to combine different forms of information representation,
  - . to extend the numeric data by means of supplementary, descriptive information,
  - . to use standardized or easily translatable formats which must be interconnected.

### 5. The Necessity of a Referral Database System

with information of the kind "Who, What, Where" building one of the bases of information brokerage.

## 2. THE ESTABLISHMENT OF THE "BIOTECHNOLOGY INFORMATION SERVICE"

### 2.1. Starting Point

The GDR's situation in the field of biotechnology information was characterized by the fact that:

- the information possibilities were (are) inadequate;
- the existing information possibilities were (are) insufficiently known and the routes were (are) troublesome;
- precisely in regard of information from or about the former GDR there is a lack of knowledge.

Starting out from this situation it was determined in March 1988 to establish a *"Biotechnology Information Service"* [4].

## 2.2. Objective

The *"Biotechnology Information Service"* improves the information provision in the field of biotechnology and is a powerful information tool. It serves the scientist as an information guide which steers him through the information landscape to the desired information. This service helps especially the scientists in the five new states of Germany as well as in Eastern Europe.

## 2.3. Tasks

The activities of the *"Biotechnology Information Service"* include

### - Information brokerage

- . **online access** to the international hosts Data-Star, STN International, FIZ Technik, ECHO, ICECC, GENIUSnet, etc;
- . **searches in its own databases:** "Information Sources in Biotechnology", "Who-What-Where in Biosciences and Biotechnology", "Conferences in Biotechnology";
- . **offline use of other databases** which are obtained on CD-ROM or diskettes, and implemented by the *"Biotechnology Information Service"* such as Science Citation Index, EMBL Nucleotide Sequence Data Library, SWISS-PROT Protein Sequence Database, and others;

### - Setting-up of own databases

- . **Information Sources in Biosciences and Biotechnology**  
This database contains information on information sources (databases, special information materials, periodicals, books, software, etc.) in the fields of biotechnology.
  - . **Who-What-Where in Biosciences and Biotechnology**  
This database contains information on institutions and companies working in the various fields of biotechnology including information on research projects, products, technologies, services, culture collections, etc.
    - for the five new states
    - for East European countries.
  - . **Conferences in Biosciences and Biotechnology;**
- **Publications** on the basis of the own databases, and as a result of information analysis and synthesis (state-of-the-art reports, reviews, bibliographies);
- **Consulting** concerning bioinformatics projects, research support, as well as making research and business contacts.

## 2.4. Contents

Because of the multidisciplinary character of biotechnology and the variety of questions directed to the *"Biotechnology Information Service"*, the **basic fields of biotechnology**

(molecular biology, genetics, cell biology, biochemistry, biophysics, microbiology, and others) as well as the **fields applied biotechnology** (medicine and pharmacy, chemistry and chemical engineering, agriculture and nutrition, ecology, etc.) are included.

## 2.5. Pre-conditions

The pre-conditions necessary for the establishment of the "*Biotechnology Information Service*" include

- a) **Data Input.** The acquisition of the necessary information is one of the main difficulties. At present we employ the distribution of questionnaires, the extraction of information from literature and from other referral databases;
- b) **Controlled Vocabulary.** For indexing and retrieval of information are used the "Thesaurus Biotechnologie" (in German) and the BIOREP thesaurus [5][6];
- c) **Investment.** The establishment of databases requires considerable investment. The role of information will be determined by the fact how a society answers the economic questions concerning the production of information, the distribution of information, and especially concerning the **development of an information infrastructure**;
- d) **Price "Policy".** Everybody has to accept that an information has its price. This insight depends on the information awareness and the importance of this information;
- e) **A Qualified Staff including Information Management.** The interaction between biotechnology and information science demands a "hybrid" knowledge. It is necessary to take into account that, at an international level, considerable attention is paid to information management.

## 2.6. Availability

- a) **In the printed form** for ready availability of information as reference books:
  - Directory of factual databases in the field of biotechnology [1]
  - Directory of bibliographic databases in the field of biotechnology [2]
  - Directory of information sources in the field of biotechnology published in the GDR [7]
  - Directory of periodicals in the field of biotechnology [8]
  - Who-What-Where in biosciences, biotechnology and genetic engineering in the former GDR [9]
  - Conferences, workshops, fairs in the field of biotechnology [10]
  - BioTechnologie Das Jahr- und Adreßbuch 91/92 [13];

b) **In the form of databases** containing more information than the printed versions:

- **The database "Information Sources"** provides answers to the following questions:  
Which information sources are where available in the field of biotechnology?  
Does a particular source of information exist and where is it available?  
What software is available for processing biotechnological data (e.g. sequences, structures)?
- **The database "Who-What-Where"** provides answers to the following questions:  
Which research institutions and industrial firms in the former GDR and in Eastern Europe are working on what research projects and supply which products or provide which services relevant to biotechnology?  
The database "Who-What-Where" contains
  - . addresses
  - . research information
  - . information on R&D projects
  - . product information
  - . technology information
  - . service information
  - . information on culture collections
  - . information on basic and further training opportunities
  - . co-operative relationships.
- **The database "Conferences"** provides answers to the questions concerning time and locations of conferences, workshops, and fairs.

c) **Online** as a part of the database BIKE via the hosts Data-Star and DIMDI.

## 2.7. Further Development

The aims of the further development of the "*Biotechnology Information Service*" are to participate in the **BIOREP project**, as well as in other directory projects of relevance to facilitating coordination of national biotechnology programs and transnational collaboration; and to provide a basis for the **coordination** of research and the application of research results in the field of biotechnology. In this way, the "*Biotechnology Information Service*" could help to support the **planning** of research, and to prevent parallel work.

Another aim will be to connect the both German biotechnology information services: **BIKE** (Biotecnologie-Informations-Knoten für Europa), and **AuBit** (Auskunftsdienst Biotecnologie).

At a later point of time we will investigate the necessity and the possibility of application of artificial intelligence methods. The aim could be to develop an **expert system** which would assist both information scientists and end-users in a referral situation.

### 3. CONCLUSIONS FOR THE FUTURE OF BIOTECHNOLOGY INFORMATION

1. **Formation of a new information awareness and behaviour**  
in the conviction that information comprises an indispensable component of all fields of science and technology.
2. **Implementation of a European biotechnology information infrastructure**  
in connection with the creation of the necessary pre-conditions
  - for the use of the large international databases for which there are no national equivalents;
  - for the setting-up of specific databases which contain information on research findings and are available either as independent databases or as data input for international databases;
  - for the inclusion and improvement of the East European information market.
3. **Development of a highly sophisticated software**  
With the large databases we have a huge information potential at our disposal, but we are not able to utilize it to a high degree. A highly sophisticated software is required in order to make the knowledge contained implicitly in these large databases explicitly available. Especially artificial intelligence methods should be introduced in order to improve the processing of different kinds of information.  
In the future, various forms of representation patterns will have to be combined in order to be able to adequately and completely represent objects properties, relations and interrelations, and to transfer information into knowledge.
4. **Importance of factual information**  
Factual databases derive their importance from the following:  
The provide a research instrument, which
  - as a consequence of its indispensability for certain types of problem solving, has led to the formation of a new attitude toward information on the part of scientists;
  - arises at the interface between subject area and information science whereby the scientist increasingly assumes the role of producer and user of information;
  - bears witness to the increasing influence of information technologies on the research process [11].The factual databases have to provide information required for Computer Aided Design, e. g. protein design, and as a part of the knowledge bases of expert systems [12].
5. **Participation in the international information exchange**  
The increasing amount of information and the considerable expenditure involved cause the gathering, processing and distribution of information to be possible only with international scientific co-operation by means of modern

**information technology.** Maintenance of international databases is particularly important to provide **timely access** to information from research.

Proceeding from the fact that significant databases in the field of biotechnology like nucleotide sequence, protein sequence, hybridoma and culture collections databases, etc. can only be created through **international co-operation**, it is **imperative** that all countries participate in the international information exchange.

At present time, the basis for an international co-operation is so good than never before. **We should use this chance!**

#### **ACKNOWLEDGEMENT**

The project "*Biotechnology Information Service*" is supported by the Federal Minister for Research and Technology (BMFT), Germany.

#### **REFERENCES**

- [1] POETZSCH, E., Verzeichnis faktographischer Datensammlungen und Datenbasen zu Grundlagen- und Anwendungsgebieten der Biotechnologie, Sci. Inf. Centre, Berlin (1988).
- [2] EBNER, A., Verzeichnis bibliographischer Datensammlungen und Datenbasen zu Grundlagen- und Anwendungsgebieten der Biotechnologie, Sci. Inf. Centre, Berlin (1989).
- [3] POETZSCH, E., Databases in molecular biology - the international situation, *Studia Biophysica* 129 2-3 (1989) 109-116.
- [4] POETZSCH, E., Vorstellungen zum Aufbau des Auskunftsdienstes Informationen zur Biotechnologie, *Informatik*, 35 5 (1989) 174-176.
- [5] *Biotechnologie-Thesaurus*, Inst. of Biotechn. Res., Brunswick (1988).
- [6] *BIOREP Biotechnology Thesaurus*, Royal Netherlands Acad. of Art and Sciences, Amsterdam (1990).
- [7] LUECKE, E.-M., MUENNICH, B., POETZSCH, E., Verzeichnis von Informationsmitteln zu Grundlagen- und Anwendungsgebieten der Biotechnologie, Sci. Inf. Centre, Berlin (1989).
- [8] *Zeitschriftenbestandsverzeichnis*, Sci. Inf. Centre, Berlin (1990).
- [9] LUECKE, E.-M., POETZSCH, E., SCHILLER, G., *Wer-Was-Wo in Biowissenschaften, Biotechnologie und Gentechnologie im Gebiet der DDR*, DECHEMA, Frankfurt a. M. (1990).



- [10] PAULIGK, S., Tagungskalender zu Biowissenschaften und Biotechnologie 1991, Sci. Inf. Centre, Berlin (1990).
- [11] POETZSCH, E., Stand und Entwicklung faktographischer Informationssysteme auf dem Gebiet der Biotechnologie, Informatik, 34 2(1987) 73-76.
- [12] POETZSCH, E., "Forderung von CAD und Expertensystemen an die Entwicklung faktographischer Datenbasen auf dem Gebiet der Biotechnologie", Wissensvermittlung-Informatik (Proc. Int. Symp. Erfurt, 1987) 178 - 191.
- [13] BioTechnologie Das Jahr- und Adreßbuch 91/92, polycom Verlagsgesellschaft mbH, Braunschweig (1991).

Dr. Eleonore Poetzsch  
Auskunftsdienst Biotechnologie  
Wissenschaftliches Informationszentrum  
Schiffbauerdamm 19  
0-1040 Berlin

### **Komponenten von Informationsvermittlung und -consulting für die veränderten Informationsbedürfnisse von Forschung und Wirtschaft in den neuen Bundesländern**

Die Situation der Informationsvermittlung in den neuen Bundesländern (NBL) und den alten Bundesländern (ABL) wird als unterschiedlich charakterisiert. In der ehem. DDR war die Informationsvermittlung durch administrative und zentralistische Strukturen gekennzeichnet. Die Akzeptanz der Informationsleistungen lag generell höher als in den ABL. Die Informationsvermittlung aus Datenbanken beinhaltete hauptsächlich die Bereitstellung von Literaturhinweisen. Für die Informationsvermittlung der neu entstehenden Wissenschaftsparks wird das Konzept des Wissenschaftstransfers favorisiert. Literaturinformationen stehen nach wie vor im Vordergrund. Darüberhinaus zeigt sich ein zunehmender Bedarf an Informationen u.a. über Firmen, Wettbewerbsaktivitäten, Marktchancen. In gleichem Maße wächst die Nachfrage nach dem Aufzeigen von Kontakt-, Auftrags-, Austausch- Projektfördermöglichkeiten sowie nach Beratungen über Know-How-Transfer, Informationstechnologien und Ausbildung. Erfahrungen mit einem breiten Nutzerkreises werden mitgeteilt.

### **Components of information mediation and consulting in view of changed information needs of research and economy in the new Federal Lands.**

The situation of information intermediation is characterized as different in the new Federal Lands (NFL) in comparison with the old federal Lands (OFL). In former GDR information services were marked by administrative and centralized structures. In general, value estimation of information services by clients was higher than in OFL. Information intermediation with databases enclosed mainly hints of literature items.

For the ongoing creation of science and technology parks the concept of science transfer will be favorized. Literature information is still in front. Furthermore a growing need is indicated of information on firms, competition activities, market chances. In the same manner the demand is growing on pointing out possibilities for contacts, tenders, exchanges, project promotions as well as consulting on know-how-transfer, information technology and training. Experiences with a broad clientele are reported.

# Komponenten von Informationsvermittlung und -consulting für die veränderten Informationsbedürfnisse von Forschung und Wirtschaft in den neuen Bundesländern

Dr. Stephan Kurby, Dipl.-Ing. Joachim Schwandt

## 1. Das Problem

Die alten administrativen Wirtschafts- und Forschungsstrukturen in den neuen Bundesländern (NBL) existieren nicht mehr. Neue marktgerechte Strukturen existieren noch nicht. Die FuE-Einrichtungen und innovativen Unternehmen in den neuen Bundesländern stehen vor der Aufgabe, wettbewerbsfähig unter den Bedingungen eines offenen Forschungsmarktes zu sein. Die Umstrukturierung der Wirtschaft und Forschung initiiert z.Z. eine Umstrukturierung des Bedarfs an Wissenschafts-, Wirtschafts- und Marktinformationen. Die Informationsbedürfnisse der potentiellen Nutzer von Informationsvermittlungsstellen (IVS) verändern sich signifikant. IVS müssen nach Wegen suchen, diese zu befriedigen sowie das Konzept und die Strukturen der Informationsvermittlung neu zugestalten.

## 2. Die Ausgangssituation der Informationsvermittlung in den neuen Bundesländern

Für das Verständnis der Umstrukturierung der Informationsvermittlung in den NBL erscheint es sinnvoll, die Ausgangslage kurz (Ist-Stand der jüngsten Vergangenheit) zu betrachten.

## 2.1. Konzepte und Leistungsangebote der Informationsvermittlung in der ehem. DDR

Im Gegensatz zu den alten Bundesländern (ABL) wirkte die Informationsvermittlung in einem geschlossenen System mit entsprechenden Auswirkungen auf Konzepte und Leistungsangebote sowie Akzeptanz der Leistungen.

Die Informationstätigkeit war vorwiegend an größeren, meist überregionalen Einrichtungen zentralisiert. Sie war gekennzeichnet durch administrative Strukturen und einen z.T. enormen technologischen Rückstand /1/. Als Konzepte dominierten Informationszentren neben der traditionellen Auskunftstätigkeit in Bibliotheken.

Die Leistungsangebote von IVS beinhalteten im wesentlichen die Bereitstellung von Literaturhinweisen:

- SIV (Selektive Informationsverbreitung)
  - Profildienste mit eigenerstellten Datenfonds,
  - Profildienste mit geleasteten Datenbasen (z.B. INSPEC);
- Online Datenbankrecherchen
  - Zugriff zu bibliographischen und faktographischen eigenerstellten Datenfonds,
  - "grenzüberschreitender Fernzugriff" zu internationalen Hosts.

Die Vermittlung von Literaturhinweisen unterlag insbes. bei den internationalen Online-Datenbankrecherchen einer restriktiven Verfahrensweise. Trotz der z.T. schwierigen Arbeitsbedingungen gab es dennoch einige IVS, die mit ihren Leistungen den Vergleich von IVS mit den ABL nicht scheuen brauchten.

Um den objektiv vorhandenen Informationsbedarf wenigstens teilweise zu decken, bildete sich z.B. im Wissenschaftlichen Informationszentrum der ehem. Akademie der Wissenschaften bereits vor 10 Jahren eine kleine Gruppe, die sich der Aufgabe Informationsvermittlung aus kommerziellen Datenbanken widmete. Diese Gruppe, die heutigen Fachinformationsdienste (FID),

betreiben Informationsvermittlung für den Bereich der ehemaligen Akademie der Wissenschaften, den universitären Bereich sowie in zunehmenden Maße für die Wirtschaft.

## 2.2. Akzeptanz der Informationsleistungen

Die Akzeptanz der Informationsleistungen von IVS und IuD-Stellen lag in der DDR generell höher als in den ABL, wie. Ergebnisse einer Felduntersuchung kurz vor der Wende, 1988, belegen /2/.

Tab. 1 zeigt die von Wissenschaftlern aus der ehemaligen Akademie der Wissenschaften angegebene Rangfolge von Informationsmöglichkeiten.

<u>Informationsart</u>	<u>Rang</u>
Zeitschriften	1
Datenbanken	2
Monographien	3,5
pers. Gespräche/ Konferenzen	3,5
andere	5

Tab. 1: Rangfolge der Informationsmöglichkeiten für Wissenschaftler aus der ehem. Akademie

Bei vergleichbaren internationalen Untersuchungen liegen Datenbanken auf mittleren Rangplätzen, nach persönlichen Gesprächen, Konferenzbesuchen und Monographien /3,...,5/.

Gleiches gilt für die Akzeptanz von Informationsdiensten, d.h. die Bereitschaft von Wissenschaftlern, Datenbanken zu nutzen (s. Abb. 1). Während die Felduntersuchung eine relativ hohe Akzeptanz ausweist, ist in den ABL, aber auch international ein gewisser 'Konservatismus' zu verzeichnen /6,...,8/.

Ursache für die höhere Akzeptanz und Rangplazierung von ID-Leistungen waren die Schwierigkeiten, Kontakte mit der internationalen Wissenschaftlergemeinschaft herzustellen und zu pflegen sowie die beschränkten Möglichkeiten der Verfügbarkeit von Monographien und Tagungsberichten.

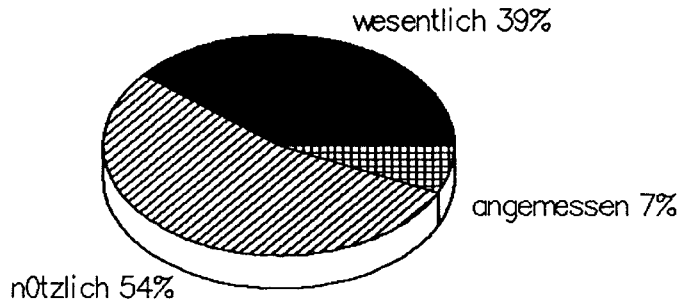


Abb. 1: Akzeptanz von ID in FuE

Die Ausgangslage der Informationsvermittlung in den ABL und den NBL war somit durchaus unterschiedlich. Während als Umfeld in den ABL eine nach den geltenden Maßstäben funktionierende Wirtschafts- und Wissenschaftslandschaft existiert, ist dies in den NBL nicht der Fall. In den ABL ist die Akzeptanz von Informationsdienst-Leistungen trotz großer Förderprogramme und guter Technikausrüstung mittelmäßig, in der ehem. DDR dagegen relativ hoch. Die Umfeldbedingungen haben sich in den NBL wesentlich verändert. Ungehinderte Kontakte zu Wissenschaftlern, Tagungsbesuche und Literaturbezug sind nun möglich, haben jedoch ihre Limitierung durch die finanziellen Möglichkeiten der FuE-Einrichtungen. Die Akzeptanz der ID-Leistungen durch Wissenschaftler in den NBL ist, entgegen mancher Erwartung, nicht gesunken. Von einem einheitlichen Informationsmarkt kann z.Z. noch nicht ausgegangen werden.

### 3. Konzepte der Informationsvermittlung in den NBL

#### 3.1. Strukturen der Informationsvermittlung

Das Umfeld der Informationsvermittlung in den NBL entwickelt sich noch schleppend. Marktgerechte Strukturen in der

Wirtschaft beginnen sich zögernd und widerspruchsvoll abzuzeichnen. Mit der Abwicklung der Forschungsakademien und der Neugliederung der Hochschulen ist die Umstrukturierung der Wissenschaftslandschaft z.Z. erst am Anfang.

Für die Informationsvermittlung bedeutet dies eine grundlegende Neuausrichtung. Im wesentlichen sind z.Z. 3 Formen von IVS im Aufbau:

- rechtlich selbständige, private IVS (Informationsbroker-Konzept);
- wirtschaftsnahe IVS (Technologievermittlungs-Konzept);
- IVS in der Forschung,
  - IVS an Hochschulen,
  - IVS an FuE-Einrichtungen,
  - IVS in Wissenschaftsparks.

Eine schematische Übernahme der Formen privater und wirtschaftsnaher IVS wird aus Gründen der unterschiedlichen Ausgangsbedingungen, sowohl des Umfeldes als auch der Informationsvermittlung selbst, als äußerst problematisch angesehen. Die interessanteste Entwicklung wird sich u.E. bei den IVS in den Wissenschaftsparks ergeben.

### 3.2. Informationsvermittlung in Wissenschaftsparks

Wissenschaftsparks in den NBL orientieren sich an den Vorbildern aus den USA und Westeuropa /9, 10/. In räumlicher Nähe auf einem Territorium arbeiten dort universitäre und außeruniversitäre Forschungseinrichtungen sowie innovative Unternehmen in engem Dialog und zum gegenseitigem Nutzen nebeneinander.

Eine wesentliche Grundlage für die Effektivität von Wissenschaftsparks ist eine gut funktionierende **gemeinsame** Infrastruktur, insbes. für den Informations- und Bibliotheksbereich. Externe IVS können dem komplexen und z.T. sehr speziellen Informationsbedarf der in den Parks angesiedelten Einrichtungen i.d.R. nicht gerecht werden. Für

Leistungsangebote weiterzuentwickeln. Die IVS bieten in den Parks nicht nur die reine Informationsvermittlung durch Datenbanken an. Durch Kombination der Datenbankinformation mit anderen konventionellen (auch informellen) Informationsquellen, verbunden mit dem Expertenwissen und dem vorhandenen Vertrauensverhältnis zu den Nutzern können sie zusätzlich zur Organisation des Wissenschaftstransfers beitragen. Der Wissenschaftstransfer als Spielart des Wissenstransfers zielt darauf ab, "die in den wissenschaftlichen und technischen Erkenntnissen enthaltenen Informationen als Handlungswissen in außerwissenschaftliche Anwendungsbereiche zu transferieren und für weitergehende Problemlösungen nutzbar zu machen" /6/. Er unterscheidet sich deutlich vom Technologietransfer, bei dem technische Lösungen aus FuE dem Praxisbereich angepaßt werden /6/.

Der Wissenschaftstransfer erfolgt notwendig in zwei Ebenen:

- in und zwischen wissenschaftlichen Einrichtungen des 'Parks',
- intersektoral zwischen wissenschaftlichen und wirtschaftsnahen Einrichtungen.

Zum Wissenstransfer zählen hauptsächlich:

- Informationstransfer,
- Know-How-Transfer,
- Kontakt-Initiierung,
- Erkenntnistransfer,
- Ideentransfer,
- Ergebnistransfer.

Wie diese Aufgaben am Beispiel der FID im Wissenschaftspark Berlin-Adlershof konzipiert und realisiert werden, versuchen die nachfolgenden Ausführungen an Teilbereichen zu verdeutlichen.

### 3.3. Veränderungen im Informationsbedarf

Die Umorientierung von Unternehmen und Forschungsgruppen ist besonders gut an den Aufträgen für Online-Datenbankrecherchen ersichtlich.



Seit 1987 nutzen wir den Host Data-Star routinemäßig. Verträge für fünf weitere Hosts mit Datenfonds zu Wissenschaft und Wirtschaft kamen im Laufe der Zeit hinzu.

Das Angebot an Wirtschaftsdaten wurde in den Jahren vor 1990 nur wenig in Anspruch genommen. Wirtschaftsdaten interessierten neben wissenschaftlicher Literatur im wesentlichen nur im Zusammenhang mit der Entwicklung neuer Produkte oder neuer Technologien. Die Inanspruchnahme von Wirtschaftsdatenfonds mit der genannten Orientierung lag in der Regel unter 30% (s. Tab.2)

Monat/Jahr	<u>7/89</u>	<u>9/89</u>	<u>3/90</u>	<u>7/91</u>
Datenfonds pro Auftrag	4,2 (1...8)	3,1 (1...8)	3,1 (1..7)	3,0 (1...7)
wissenschaftl. Datenfonds	88%	88%	72%	72%
Wirtschafts- datenfonds	12%	12%	28%	28%
Aufträge nur zur Wirtschaft	--	--	19%	18%

Tab. 2: Änderung des Informationsbedarfs nach 1989  
(Stichproben)

Seit der Jahreswende 1989/1990 nahm die Nachfrage nach Wirtschaftsdaten deutlich zu. Jetzt wurden auch Anfragen gestellt, die nur auf Wirtschaftsdaten, nicht auf wissenschaftliche Literatur, ausgerichtet waren (s. Abb. 2). Ziele dieser Anfragen sind im wesentlichen, die Marktchancen für die in der Entwicklung befindlichen Produkte und Technologien

festzustellen, sowie Konkurrenten einerseits, mögliche Kooperationspartner andererseits zu ermitteln und Auskünfte über sie einzuholen.

Ab März 1990 konnten wir uns auch für Rechercheaufträge aus Betrieben öffnen. Diese hielten besonders nach Joint-Venture-Partnern Ausschau und holten Auskünfte über die Wirtschaftskraft potentieller Partner ein. Sie suchten aber auch das bisherige Defizit an Literaturinformationen zu decken.

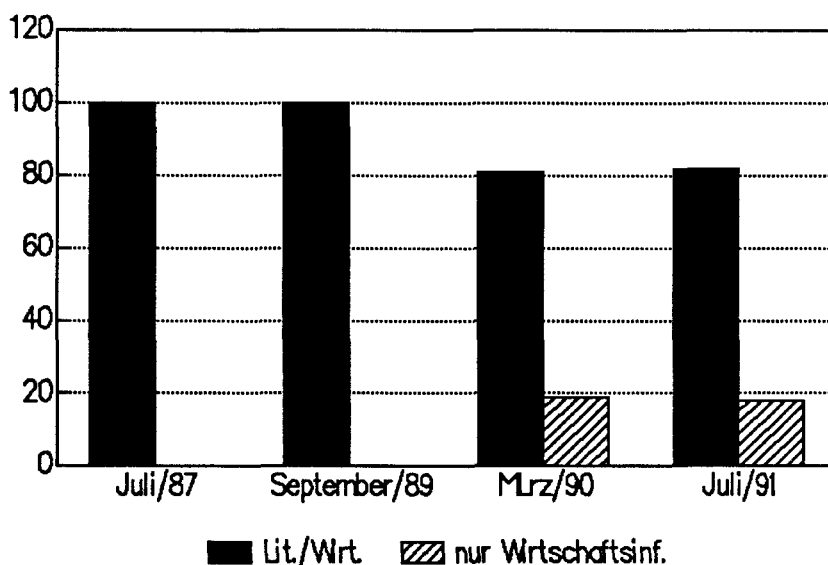


Abb. 2: Änderung des Informationsbedarfs (Stichproben)

### 3.4. Änderung der regionalen Verteilung der Auftraggeber für Online-Datenbankrecherchen

Mit unterschiedlichen Förderprogrammen sind im Verlaufe des letzten Jahres zahlreiche Online-Stationen in Betrieben und

Instituten installiert worden. Einrichtungen mit hohen Nutzungsfrequenzen können sich jetzt selbst versorgen. Von diesen Möglichkeiten machten besonders Institute Gebrauch, die ihren Sitz außerhalb der Region (Ost-)Berlin/Brandenburg haben. Außerdem wurden in den großen Städten Informationsvermittlungs-Stellen oder ähnliche Einrichtungen eröffnet, die ihre Dienste besonders den jeweils ortsansässigen Betrieben anbieten. Beides führte zum Rückgang der überregionalen Inanspruchnahme zentraler Online-Informationsvermittlungsleistungen.

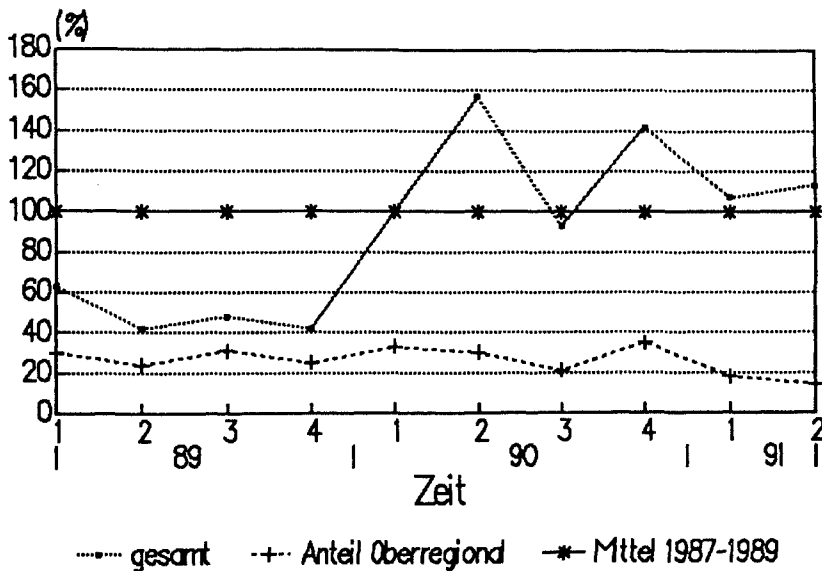


Abb. 3: Auftragsentwicklung (rel. Anz. d. Rechercheaufträge)

Veränderungen ergaben sich aber auch aus dem Wegfall von Hürden für die Inanspruchnahme von Online-Informationen, die mit solchen Dingen wie Valutamittel, Staatsplanthemen oder Geheimnisschutz im Zusammenhang standen. Mußte früher die Unterschrift des Institutsdirektors unter den Auftrag für eine

Online-Recherche eingeholt werden, so entscheidet heute der Projektleiter über die Mittel und Wege der Informationsbeschaffung. Das führte zum Anstieg der regionalen Inanspruchnahme der Online-Informationsleistungen.

Die Auswirkungen der Dezentralisierung in der Online-Informationsvermittlung wurden für uns erst 1991 sichtbar. Während die Zahl der Aufträge aus den Instituten und Einrichtungen der ehemaligen Akademie der Wissenschaften in der Summe anstieg, ging die Zahl der externen Kunden wieder zurück; während die Zahl der Aufträge aus überregionalen Standorten zurückging, nahm die Zahl der Aufträge aus der Region Berlin/Brandenburg zu. Die Änderung der regionalen Verteilung ist in Abb. 3 dargestellt.

### 3.5. Die Komponenten der Informationstätigkeit zur Deckung der neuen Anforderungen

Online-Datenbankrecherchen sind ein wichtiges Gebiet aber eben nur ein Teilgebiet der Informationsversorgung. Die von den Mitarbeitern von Forschungsinstituten und innovativen Unternehmen sowie von Informationsstellen und Bibliotheken an uns herangetragenen Anforderungen beziehen wesentlich weiterreichende Tätigkeitsfelder einer IVS, im wesentlichen die Elemente des Wissenschaftstransfers, ein.

Zur inhaltlichen Begründung von Projekten sowie für laufende Arbeiten in Forschung und Entwicklung stehen nach wie vor Literaturinformationen im Vordergrund. Darüber hinaus besteht aber ein zunehmender Bedarf an Informationen besonders über Produkte, deren Absatzmöglichkeiten, Wettbewerbsbeobachtung und Marktchancen. Ebenso wächst die Nachfrage nach Beratung über Informationsmöglichkeiten und -technologien incl. Ausbildung, Schulung und Betreuung sowie nach dem Aufzeigen individueller und institutioneller Kontaktmöglichkeiten sowie nach Auftragsgelegenheiten und Projekten.

Den erweiterten Anforderungen begegnen wir mit folgender Angebotspalette:

- Profildienste mit den Datenfonds
  - = INSPEC und
  - = Current Contentsauf dem institutseigenen Rechner;
  
- Online-Datenbankrecherchen dem Informationsbedarf unserer Hauptnutzer entsprechend vorrangig zu Wissenschaft und Wirtschaft mit Zugang zu den Hosts
  - = Data-Star (Wissenschaft, Wirtschaft),
  - = DIMDI (Biowissenschaften, Medizin, Landwirtschaft),
  - = ECHO (EG-Wirtschaftsförderung),
  - = FIZ-Technik (Wissenschaft/Technik, Wirtschaft),
  - = GENIOS (Wirtschaft),
  - = STN (Wissenschaft (bes. Chemie), Patente);
  
- Beratung / Consulting
  - = individuelle Beratung zu Informationsmöglichkeiten,
  - = Bearbeitung und problemorientierte Aufbereitung von Einzelinformationen,
  - = Bereitstellung von institutionellen und informellen Wissens- und Know-How-Quellen;
  
- Kontakte
  - = Hinweise auf individuelle und institutionelle Kontakt-, Auftrags-, Austausch- und Projektfördermöglichkeiten in West- und Ost-Europa,
  - Unterstützung des Transfers von wissenschaftlichen Ergebnissen und Know-How zwischen Wissenschaft und Wirtschaft;
  
- Ausbildungsservice
  - = Durchführung von Seminaren und Trainingskursen zu allen Fragen der Informationsvermittlung und -beschaffung.

#### 4. Ausblick

Die Umstrukturierung von Forschung und Wirtschaft verändert signifikant die Informationsbedürfnisse der potentiellen Nutzer. Informationsstellen müssen die Struktur der Informationsvermittlung neu gestalten. Den Erfahrungen aus den alten Bundesländern folgend wird die Nachfrage nach Einzelinformationen zurückgehen und die Forderung nach Komplexleistungen zunehmen /11/. So wie sich zur Anhebung des allgemeinen technologischen Niveaus die Methode des Technologietransfers durchgesetzt hat, gilt es zur Beschleunigung der Innovation wissenschaftlicher Ergebnisse die Methode des Wissenschaftstransfers weiterzuentwickeln.

Die Vorstellung, mit einem Brokerbüro ein privates Unternehmen zuführen, hat in den NBL jedoch geringe Erfolgsaussichten. Die in der Vergangenheit durchgeführten Förderprogramme zur Durchsetzung der Information aus Datenbanken haben dazu geführt, daß von öffentlichen Stellen, Hochschulen, Universitäten sowie von Kammern und Verbänden Informationsleistungen zu nicht gewinnorientierten Preisen angeboten werden. Die Preise, die ein privatwirtschaftlich geführtes Unternehmen erheben müßte, liegen nach unseren Erfahrungen oberhalb der Grenze, die potentiellen Kunden derzeit bereit oder gar in der Lage sind zu bezahlen.

#### Literatur

- /1/ Manecke, H.-J.  
Strukturen im Wandel. Stand und Perspektiven der  
Industrieinformation in der DDR  
in: Neubauer, W. (Hrsg.); Schneider-Briehn, U. (Hrsg.) :  
Deutscher Dokumentartag 1990, 25. - 27.9.1990;  
Proceedings, Frankfurt a. Main. - 1991. - S. 71 - 89
  
- /2/ Kurby, St.  
Nutzung automatisierter Informationsdienste - Darstellung  
der Wechselbeziehung Informationsnutzer-Informationsdienst  
und Anforderungen an die Gestaltung von Informations-  
diensten  
Humboldt-Univ., Diss. A. - Berlin 1989. - 212 S.

- /3/ Aloni, M.  
Patterns of information transfer among engineers and applied scientists in complex organizations  
Scientometrics. - Amsterdam, Budapest 8(1986)5. - S. 279 - 300
- /4/ Dervin, B.; Nilan, M.  
Information needs and users  
Ann. Rev. Inf. Sci. Technol. - New York (1986)21. - S. 3 - 33
- /5/ Takayama, M.  
Analysis of technological information transfer among japanese computer scientists at a research front  
Inf. Serv. & Use. - Amsterdam 6(1988)1. - S. 9 - 25
- /6/ Schmidt, R.  
Informationssysteme und Datenbanken  
in: Handbuch des Wissenschaftstransfers, Schuster, H.-J. (Hrsg.), Springer Verlag. - Berlin 1990
- /7/ Deutscher Dokumentartag 1987: Von der Information zum Wissen - vom Wissen zur Information  
Nachr. Dok. - Weinheim 39(1988)1. - S. 30 - 37
- /8/ McKinney, G.; Mosby, A. P.  
Online in academia: a survey of online searching in US colleges and universities  
Online Rev. - Oxford 10(1986)2. - S. 107 - 124
- /9/ Everett, M.; Larsen, J.K.  
Silicon Valley Fieber  
Diogenes Verlag. - Zürich 1986
- /10/ Fiedler, H.; Wodtke, K.-H.  
Innovationszentren in Deutschland, Österreich und der Schweiz 1990/91  
Weidler Buchverlag Berlin. - Berlin 1990
- /11/ Schmidt, R.; Wellems, C.  
Der Modellversuch Informationsvermittlung - eine Bilanz  
in: Fraunhofer Institut für Systemtechnik und Innovationsforschung, Arbeitspapier Nr. 101. - Karlsruhe 1990. 26 S.

**Autoren:**

Dr. Stephan Kurby und Dipl.-Ing. Joachim Schwandt  
Rechenzentrum (RZA)  
Fachinformationsdienste FID  
Mohrenstr. 39  
O-1086 Berlin

## Managementinformationen auf der Basis von statistischen Untersuchungen in Patentinformationsdiensten

Babst, P.

Zur erfolgreichen Orientierung auf dem Markt sind Managementinformationen über Trends von Erzeugnissen, Verfahren, Märkten und über Firmen unerlässlich. In der Praxis gibt es daher zahlreiche Methoden, um Trends der wirtschaftlichen Entwicklung und des Marktes zu erkennen. Wesentliche Parameter für die Marktforschung sind z.B.:

- Produktionskennziffern
- die Auslastung bestehender Produktionskapazitäten
- der Bau neuer Produktionslinien
- Export-/Importzahlen
- Preisentwicklungen.

Obwohl im Laufe der Zeit die zugrundegelegten Beobachtungsparameter ständig erweitert und die Beobachtungsmethoden sehr verfeinert worden sind, gibt es immer wieder unvorhergesehene Turbulenzen auf dem Markt, die durch technische, technologische oder gesamtwirtschaftliche Änderungen hervorgerufen werden.

Die Ursache für das nicht rechtzeitige Erkennen solcher bevorstehenden Ereignisse liegt unter anderem auch im Charakter der zugrundegelegten Beobachtungsparameter, die ihrem Wesen nach Änderungen meistens erst dann erkennen lassen, nachdem sie eingetreten sind. Im Prinzip läßt sich mit den herkömmlichen Methoden der Marktbeobachtung zwar der gegenwärtige Zustand erkennen, aber kaum oder nur im begrenzten Maße das, was vorbereitet wird.



Unvorhergesehene Änderungen auf dem Markt werden häufig durch die Einführung neuer Technologien oder Erzeugnisse ausgelöst, die die im Handel befindlichen Waren durch neue ablösen, aber auch den Bedarf an herkömmlichen Erzeugnissen durch das Entstehen neuer Einsatzgebiete ansteigen lassen. Um Voraussagen zur künftigen Marktentwicklung sicherer zu gestalten, müssen technische und technologische Wandlungen stärker als bisher noch vor ihrer Produktionswirksamkeit und Auswirkung auf den Markt wahrgenommen, objektiviert und bewertet werden.

Eine geeignete Methode zur Erkennung dynamischer Technologiefelder der Zukunft ist die statistische Analyse wissenschaftlich-technischer Veröffentlichungen, deren Häufigkeit zu einem Objekt in Korrelation mit der erwarteten zukünftigen wirtschaftlichen Bedeutung steht.

Qualitativ besonders hochwertige Informationen liefert die statistische Auswertung von Patentinformationsdiensten, weil diese bekanntermaßen die umfassendste wissenschaftlich-technische Informationsquelle sind und in komprimierter Form das gesamte verfügbare technische Wissen beinhalten. Bei statistischen Analysen in Patentinformationsdiensten stehen nicht die patentjuristischen Aspekte einer Anmeldung im Vordergrund, sondern die aus den Patentschriften entnehmbaren Informationen wie z.B.:

- der technische Inhalt der Anmeldung
  - die Anmeldehäufigkeit zu einem Sachgebiet, die ein Ausdruck der internationalen Forschungsintensität ist
  - der Anmelder, der zugleich Wettbewerber und Konkurrent ist
  - die Anmeldeländer, die als zukünftiges vorgesehenes Marktgebiet des Konkurrenten anzusehen sind
  - der Länderumfang einer Anmeldung, der ein Anzeichen für die zukünftige wirtschaftliche Bedeutung der Erfindung ist, die ihr der Anmelder beimißt
- sowie die zahlenmäßigen Veränderungen der genannten Parameter im Verlauf der Zeit.

Von der Anmeldung einer Erfindung zum Patent bis zur Realisierung der Erfindung in der Produktion vergeht im allgemeinen ein Zeitraum von mehreren Jahren. Die maschinelle statistische Auswertung der in den Patentschriften enthaltenen bibliografischen und technischen Angaben kann daher frühzeitig Auskunft über mögliche Wandlungen in Technologiefeldern, der Forschungsaktivitäten und Marktabsichten von Firmen sowie über den Trend für Erzeugnisse und Verfahren geben. Im Prinzip können Analysen auf der Basis statistischer Untersuchungen der Patentanmeldetätigkeit für folgende Gebiete angefertigt werden:

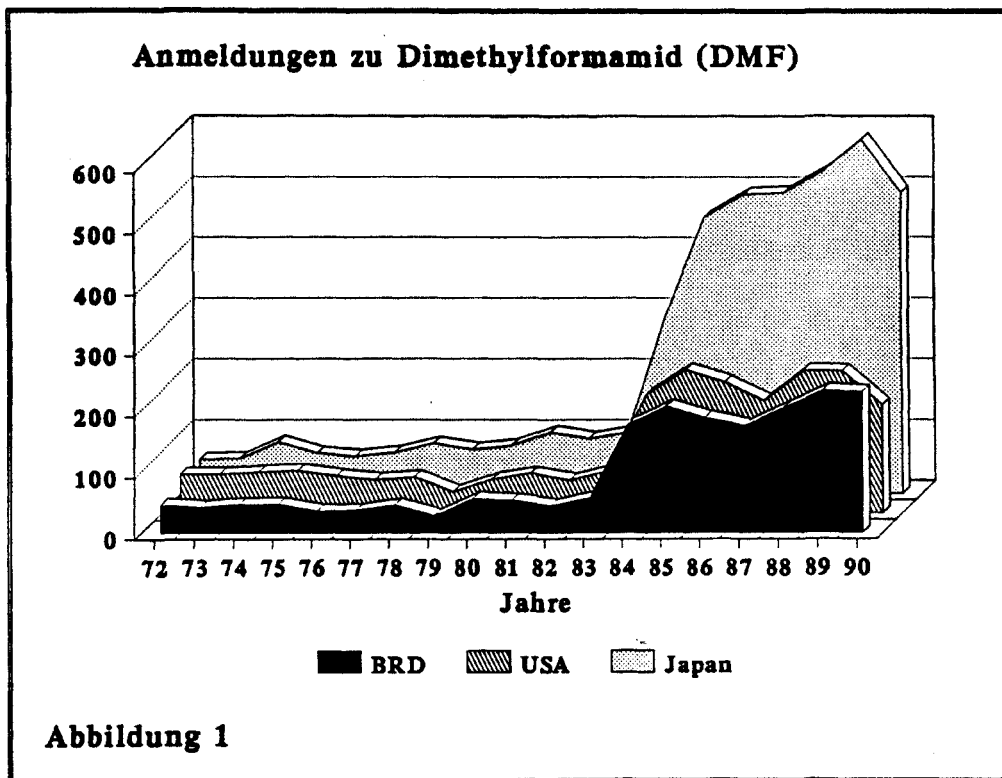
- Trendanalysen zu Erzeugnissen und deren Herstellungsverfahren
- Firmenanalysen über die Forschungstätigkeit, die Veränderung der Forschungsgebiete und die Marktabsichten von Wettbewerbern
- Trendanalysen zum Erkennen von Schwerpunkten der Forschungstätigkeit innerhalb von Erzeugnisgruppen oder größeren Technologiefeldern
- Länderanalysen

Die Ausgangsdaten für die Analysen werden durch Online-Recherchen in internationalen Datenbanken und durch Mitspeichern der Rechercheergebnisse gewonnen. Die eigentliche statistische Analyse und die Verdichtung der erhaltenen Ergebnisse zu Trend- und Firmenanalysen erfolgt auf einem Personalcomputer unter Verwendung geeigneter Software. Für einen Teil der interessierenden Fragen, die mit der bekannten Software nicht oder nur ungenau bearbeitet werden können, wurde von uns spezielle Software entwickelt, die seit geraumer Zeit erfolgreich angewandt wird.

Nachfolgend sollen einige Ergebnisse statistischer Analysen, die unter Nutzung des Patentinformationsdienstes der Firma Derwent gewonnen wurden, vorgestellt werden. Im ersten Komplex sind es Ergebnisse aus Trendanalysen zu Erzeugnissen und Verfahren, die die prinzipiellen Möglichkeiten der Patentin-

formationsdienste für die Gewinnung marktwirtschaftlicher Informationen andeuten sollen.

Abb.1 zeigt die Anmeldetätigkeit zu Dimethylformamid (DMF) in der BRD, den USA und in Japan von 1972 - 1990. Obwohl dieses Erzeugnis seit vielen Jahren großtechnisch hergestellt wird und nicht neu ist, steigen die Anmeldezahlen von Erfindungen, in denen DMF von Bedeutung ist, seit 1984 in allen drei Ländern extrem an. Die Zahlen für 1990 sind noch nicht komplett, aber es ist mit hoher Wahrscheinlichkeit zu erwarten, daß die Zahlen der Anmeldungen 1990 über denen von 1989 liegen.

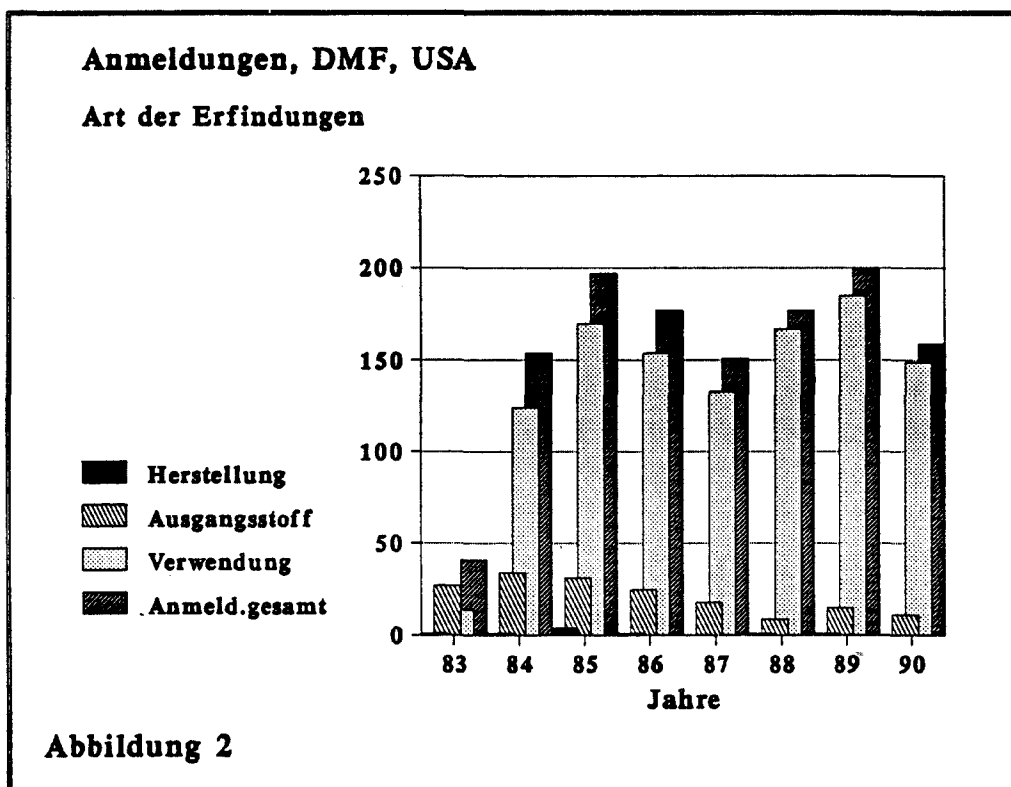


Die Derwent-Kodierung ermöglicht es, die Anmeldungen für DMF nach drei weiteren Kriterien zu analysieren, und zwar nach Anmeldungen,

- die die Herstellung von DMF zum Gegenstand haben

- in denen DMF als Ausgangsstoff eingesetzt wird, d.h. chemisch umgesetzt wird
- in denen DMF als solches verwendet wird, d.h. z.B. als Lösungsmittel oder in Gemischen

Abb. 2 spiegelt das Ergebnis der Recherche nach diesen Kriterien für die USA im Zeitraum von 1983 bis 1990 wider und zeigt, daß für DMF augenscheinlich neue Einsatzgebiete gefunden worden sind. Der Wert dieser Erkenntnis für einen Produzenten von DMF besteht darin, daß er sich zielgerichtet auf die zu erwartende Situation eines erhöhten Bedarfes an DMF rechtzeitig einstellen kann.



Darüber hinaus ist es möglich, ihm die zukünftigen potentiellen Käufer zu nennen, indem nach den Anmeldern recherchiert wird. Tabelle 1 zeigt einen Ausschnitt aus der Liste der aktivsten Patentanmelder, die in den USA Anmeldungen vorgenommen haben, in denen DMF verwendet wird.

Tabelle 1

DMF-Verwendung, USA, Aktivste Patentanmelder

Patentee	Number	%	
DOWC	DOW CHEMICAL CO	34	2.7
FARB	BAYER AG	34	2.7
DUPO	DU PONT DE NEMOURS CO	30	2.4
FARH	HOECHST AG	28	2.2
RAYC	RAYCHEM CORP	23	1.8
ALLC	ALLIED CORP	21	1.7
MINN	MINNESOTA MINING MFG CO	20	1.6
TORA	TORAY IND INC	20	1.6
BERG/	BERG L	19	1.5
CANO	CANON KK	17	1.4
CIBA	CIBA GEIGY AG	16	1.3
ESSO	EXXON RES & ENG CO	16	1.3
GENE	GENERAL ELECTRIC CO	15	1.2
FUJF	FUJI PHOTO FILM KK	14	1.1
BADI	BASF AG	13	1.0
CELA	CELANESE CORP	12	
ICIL	ICI AMERICAS INC	12	
USAS	NAT AERO & SPACE ADMIN	12	
NITL	NITTO ELECTRIC IND KK	11	
STAH	STANDARD OIL CO (OHIO)	11	
EAST	EASTMAN KODAK CO	10	
TOYJ	TOSOH CORP	10	
	Others =	860	
	Total =	1258	
	CUTOFF =	10	

Eine Firma, die eine Erfindung als wirtschaftlich bedeutsam einstuft, wird sie in vielen Ländern anmelden, d.h., die Anmeldung hat einen großen Länderumfang. Dagegen werden Anmeldungen von geringerer wirtschaftlicher Bedeutung vom Anmelder meist nur in einem Land oder in wenigen Ländern hinterlegt. Tabelle 2 gibt die Verteilung der DMF-Anmeldungen (Verwendung) in den USA wieder.

Interessant ist auch die Entwicklung der Anmeldetätigkeit bei bedeutenden Anmeldungen. Abbildung 3 offenbart, daß in den USA seit 1983 zunehmend bedeutende Anmeldungen, die die Verwendung von DMF zum Gegenstand haben, angemeldet werden.

Tabelle 2

DMF-Verwendung, USA,  
Verteilung der Anmeldungen nach  
dem Länderumfang

Länder- umfang	Anzahl der Anmeldungen	%
1	284	25.9
2 - 5	183	16.7
6 - 10	284	25.9
11 - 15	228	20.8
16 - 20	89	8.1
21 - 25	16	1.5
26 - 30	4	0.4
31 - 35	4	0.4
36 - 40	4	0.4

insges.: 1096

Anmeldungen, DMF-Verwendung, USA

Länderumfang 16-40

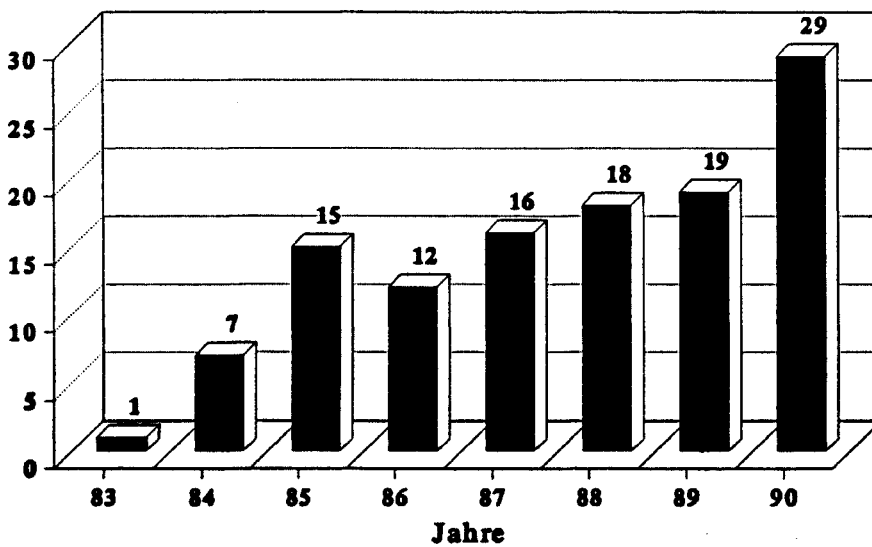


Abbildung 3

Auf Basis dieser Analysen ist es möglich, nach denjenigen Firmen zu recherchieren, die bedeutende Anmeldungen vorgenommen haben. Tabelle 3 ist ein Auszug aus einer Aufstellung derjenigen Firmen von Anmeldungen mit einem Länderumfang von 16-40.

Tabelle 3

DMF-Verwendung, USA, Firmen mit bedeutenden Anmeldungen (Länderumfang 16-40)

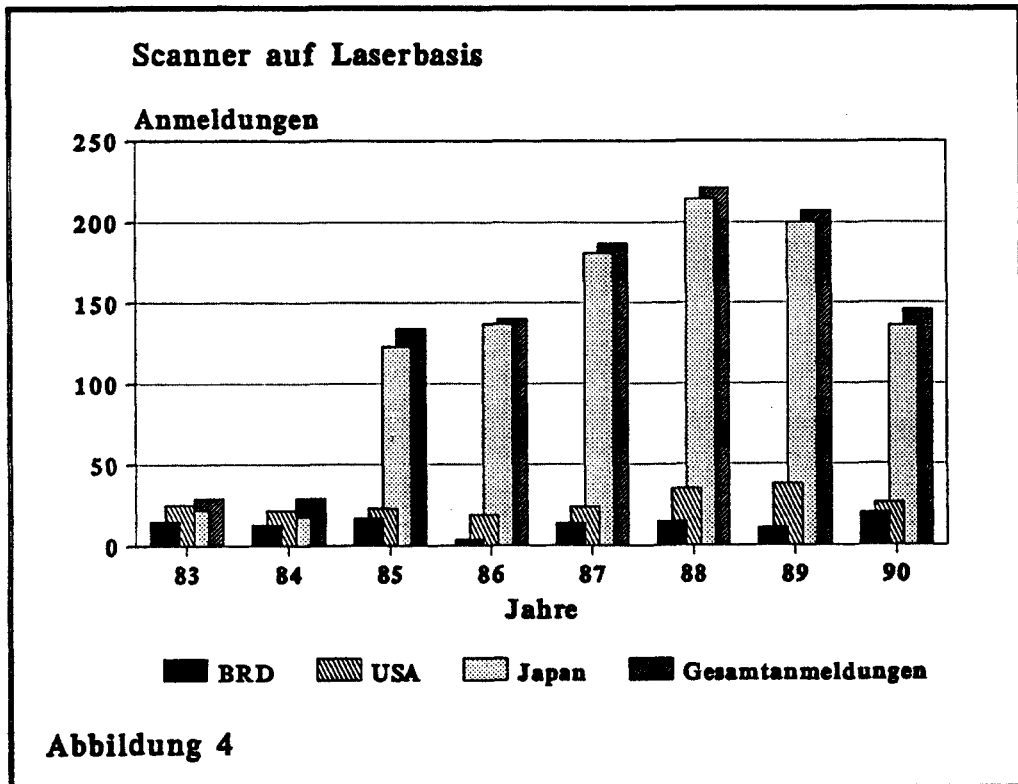
Patentee		Number	%
DOWC	DOW CHEMICAL CO	6	3.9
FARH	HOECHST AG	5	3.3
ALLC	ALLIED CORP	4	2.6
FARB	BAYER AG	4	2.6
RHON	RHONE-POULENC CHIMI	4	2.6
DUPO	DU PONT DE NEMOURS CO	3	2.0
ICIL	ICI AMERICAS INC	3	2.0
RAYC	RAYCHEM CORP	3	2.0
ROHM	ROHM & HAAS CO	3	2.0
BEND-	BEND RESEARCH INC	2	1.3
COLG	COLGATE PALMOLIVE CO	2	1.3
EAST	EASTMAN KODAK CO	2	1.3
ENQU-	ENQUAY PHARM	2	1.3
KORO/	KOROL B	2	1.3
MARK-	MARKWELL MED INST	2	1.3
STAD	AMOCO CORP	2	1.3
THOR-	THORATEC LAB CORP	2	1.3
UNIC	UNION CARBIDE CHEM	2	1.3
	Others =	99	
	Total =	152	
	CUTOFF =	2	

Marketing und Forschung interessieren sich darüber hinaus aber auch dafür, in welchen technischen Gebieten das DMF in verstärktem Maße verwendet wird. Auch diese Frage kann mit Hilfe der statistischen Analyse beantwortet werden. Aus Tabelle 4 geht zusammengefaßt hervor, daß ein hoher Anteil der Erfindungen, in denen DMF Verwendung findet, die Herstellung von Spezialpolymeren, Spezialanwendungen und die Produktion

Tabelle 4

DMF-Verwendung, USA, Anmeldungen nach Sachgebieten

Manual Codes	Number	Sachgebiet
A05-J	262	Spezialpolymere
A12-W	260	Spezialanwendungen
A12-S	209	Spezielle Halbfertig-Produkte
A10-E	202	Modifizierung von Polymeren
A11-B	165	Verformungsverfahren
A12-E	165	Elektrotechnik
A12-L	120	(Elektro) fotografie





von Halbfertigprodukten zum Gegenstand haben. Eine tiefere-  
 hende statistische Analyse weist aus, daß es sich z.B. bei den  
 Spezialpolymeren insbesondere um die Herstellung und Verwen-  
 dung von Polyimiden und bei den Spezialanwendungen in hohem  
 Maße um umkehrbare Osmoseprozesse und halbdurchlässige Mem-  
 branen handelt.

Als weitere Beispiele für den Trend bei bestimmten Erzeugnis-  
 sen sollen nachfolgend noch Scanner sowie Ausrüstungen zum  
 Drucken und Färben erwähnt werden.

Abbildung 4 zeigt die Entwicklung der Anmeldetätigkeit bei  
 Scannern auf Laserbasis in der BRD, den USA und in Japan in  
 der Zeit von 1983-1990. Die japanische Dominanz ist ein-

Tabelle 5

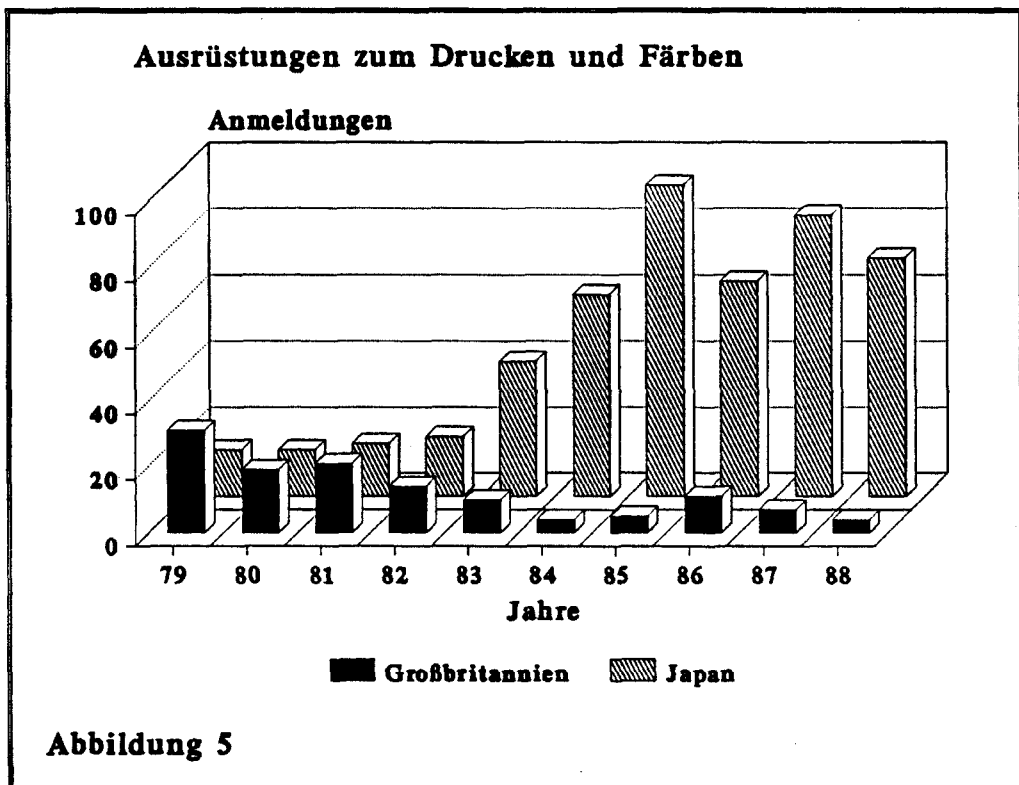
Anmeldungen zu Scannern auf Laserbasis in den  
 Ländern BRD und USA

Recherche nach den Herkunftsländern der Prio-  
 ritätsanmeldung

PR Country	BRD	%	USA	%
JP	68	56,7	143	65,9
US	32	26,7	63	29,0
DE	6	5,0	4	1,8
GB	4	3,3	2	
FR	2	1,7	0	
IL	2	1,7	2	
CA	1		0	
CH	1		1	
EP	1		0	
FI	1		1	
LU	1		0	
NL	1		1	
Others =	0		0	
Total =	120		217	
Cutoff = 1				

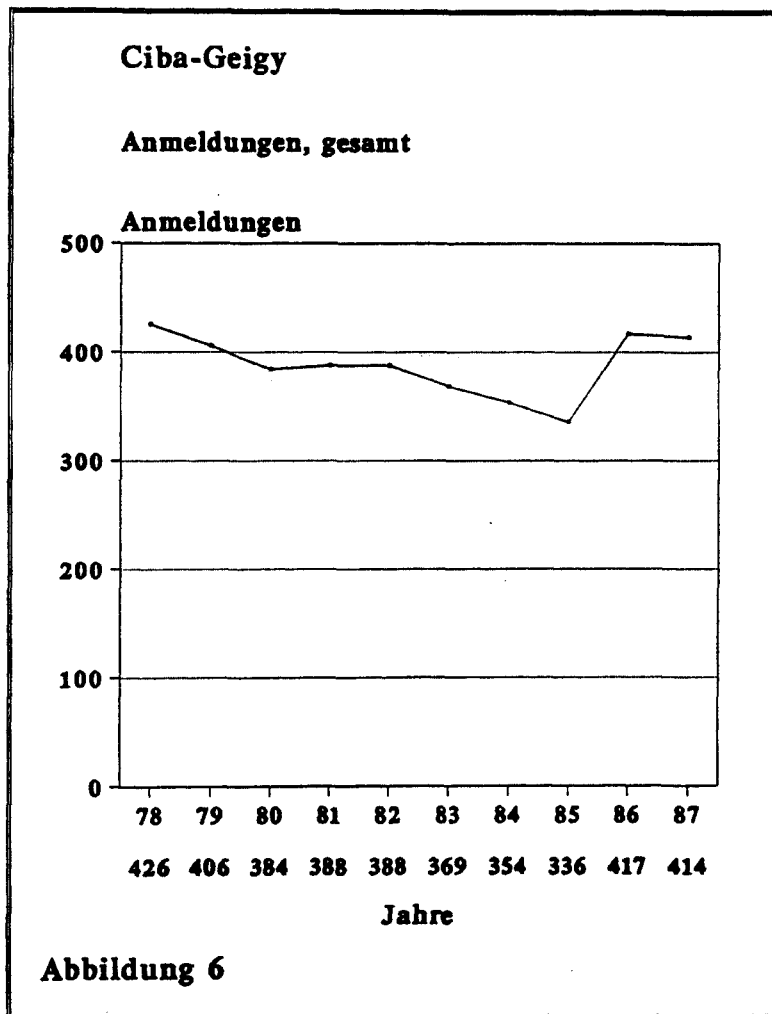
drucksvoll und wird durch die Aussage der Tabelle 5 bestätigt. Nach Tabelle 5 liegen den 120 Anmeldungen in der BRD mehr als 55% und den 217 Anmeldungen in den USA mehr als 65% Anmeldungen mit japanischer Priorität zugrunde.

Eine statistische Analyse nach Ausrüstungen zum Drucken und Färben in Großbritannien und Japan (Abbildung 5) zeigt ein völlig entgegengesetztes Anmeldeverhalten, das für Marketingfachleute auf diesem Gebiet zumindest Anlaß zum Nachdenken sein sollte.



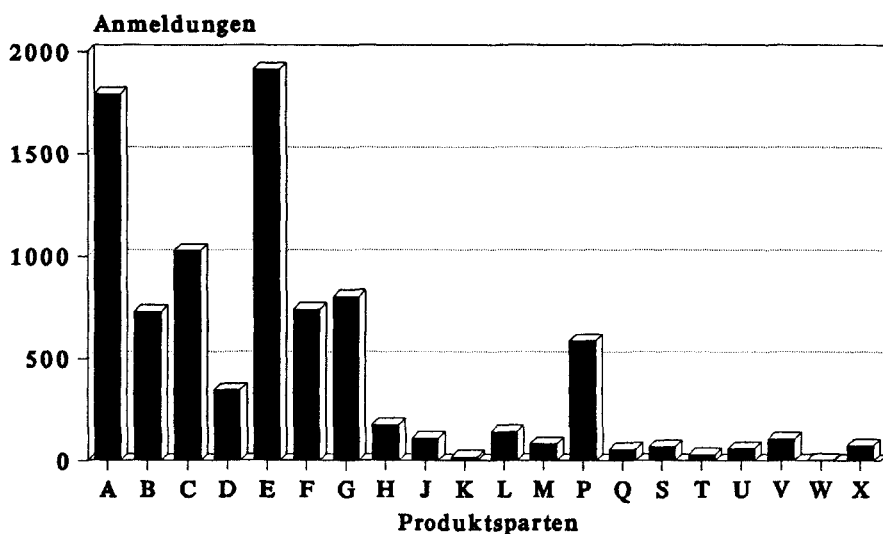
Die prinzipiellen Möglichkeiten einer Firmen- oder Wettbewerberanalyse sollen nachfolgend am Beispiel einer statistischen Analyse der Patentanmeldungen des Chemiekonzerns Ciba-Geigy gezeigt werden.

In Abbildung 6 wird die gefundene Entwicklung der Anmeldezahlen Ciba-Geigy von 1978-1987 und in Abbildung 7 die Verteilung der Anmeldungen nach einzelnen Produktparten nach der Derwent-Klassifizierung dargestellt. Abbildung 6 läßt erkennen, daß die jährlichen Anmeldungen Ciba-Geigy von 1978-1985 von 426 auf 336 sanken und erst ab 1986 wieder das Niveau von 1978 erreichten. Die für die Forschung wichtigsten Produktparten sind chemische Erzeugnisse, Polymere, Agrochemikalien, Druck/Beschichtungen/Fotografie sowie Pharmaka.



# Ciba-Geigy

## Anmeldungen nach Produktparten

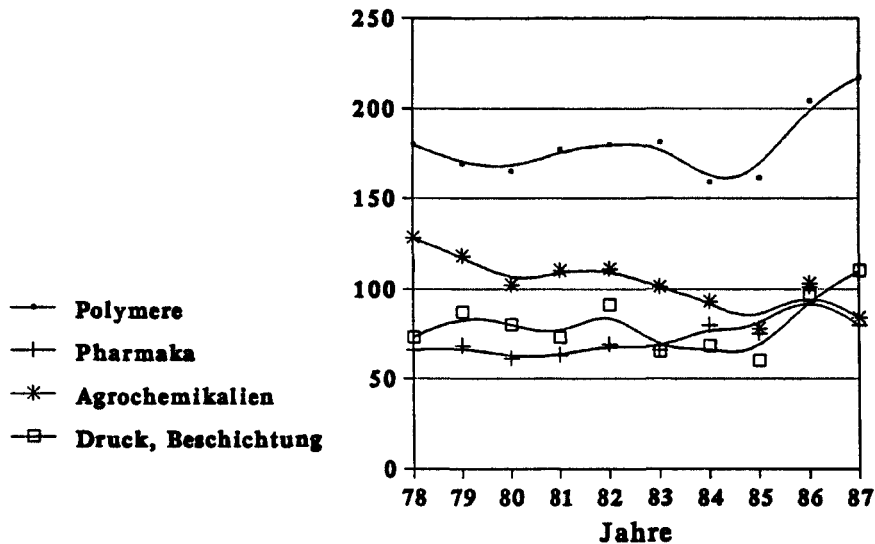


- |                                     |  |
|-------------------------------------|--|
| A - Polymere                        | L - Keramik, Glas, Elektroorganika         |
| B - Pharmaka                        | M - Metallurgie                            |
| C - Agrochemikalien                 | P - Tägl. Lebensbedarf, Arbeitsprozesse    |
| D - Nahrung, Kosmetik, Detergentien | Q - Maschinenbau, Fahrzeuge, Transport     |
| E - Allgemeine Chemie               | S - Elektrische Instrumente                |
| F - Textil, Papier, Cellulose       | T - Datenverarbeitung, Kontrolle           |
| G - Druck, Beschichtung, Fotografie | U - Halbleiter, elektronische Schaltkreise |
| H - Erdöl, Brennstoffe              | V - Elektronik                             |
| J - Chemische Verfahrenstechnik     | W - Kommunikation                          |
| K - Kerntechnik, Sprengstoffe       | X - Elektrotechnik                         |

Abbildung 7

Die Untersuchung der Anmeldetätigkeit über den Zeitraum von 1978-1987 für die Produktparten Polymere, Agrochemikalien, Pharmaka und Druck/Beschichtungen/Fotografie weist steigende Anmeldezahlen bei Polymeren, Pharmaka sowie Druck/Beschichtungen/Fotografie und sinkende bei Agrochemikalien aus (Abbildung 8). Offensichtlich haben sich die Forschungsschwer-

**Ciba-Geigy,  
Entwicklung der Anmeldungen in  
ausgewählten Produktparten**



**Abbildung 8**

punkte in diesem Sinne innerhalb von Ciba-Geigy im Lauf der Zeit verändert.

Am Beispiel der Produktparte Druck/Beschichtungen/Fotografie soll nun untersucht werden, worauf die Steigerung der Anmeldezahlen in den letzten Jahren zurückzuführen ist (Abbildung 9). Die statistische Analyse dieser Produktparte zeigt, daß von insgesamt 6 Teilgebieten insbesondere in den beiden Bereichen

- Beschichtungen/Farben/Tinten (Abbildung 10)
- Klebstoffe (Abbildung 11)

steigende Anmeldezahlen zu beobachten sind.

Innerhalb des Teilgebietes Beschichtungen/Farben/Tinten (Abbildung 10) kann nach den Davonpositionen

- Farben und Lacke (anorganisch)
- Farben und Lacke (organisch)

### Ciba-Geigy

Druck, Beschichtung, Fotografie

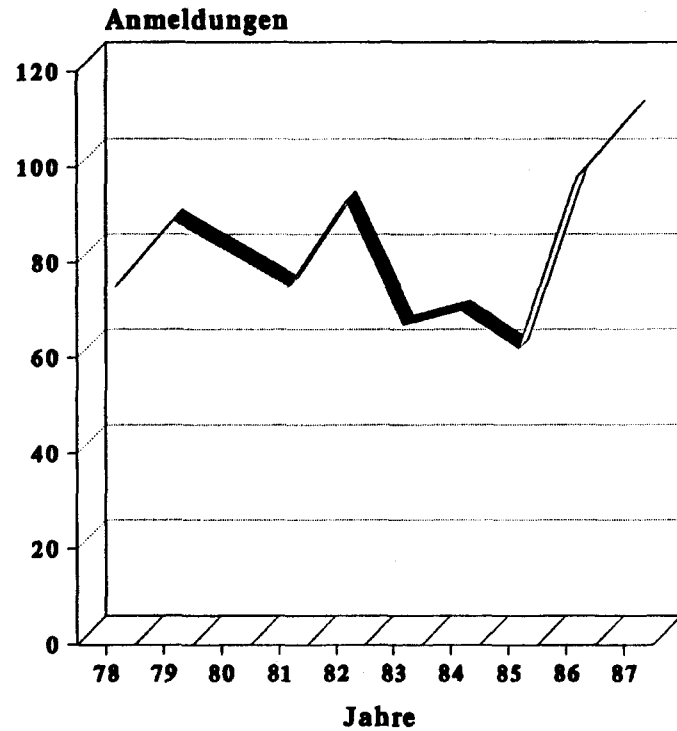


Abbildung 9

### Ciba-Geigy

Beschichtungen, Farben, Tinte

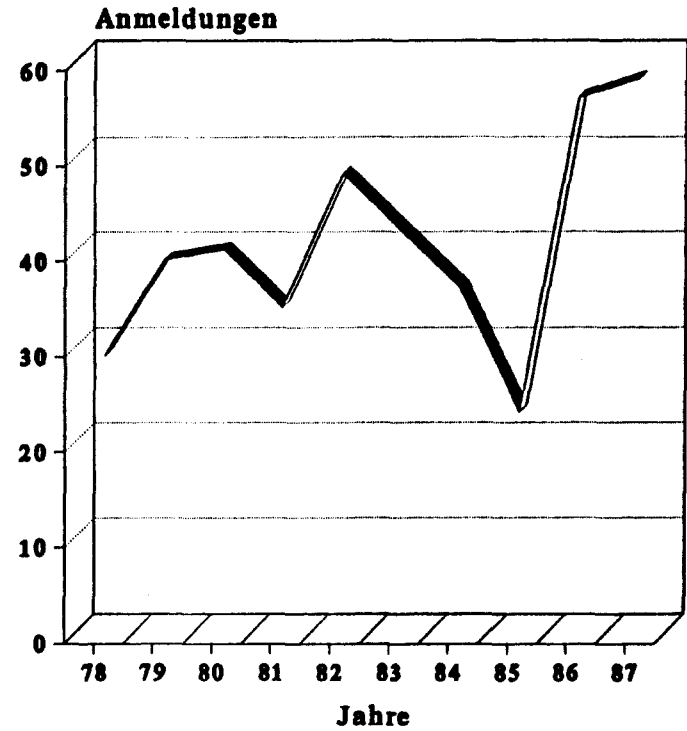


Abbildung 10

### Ciba-Geigy

#### Klebstoffe

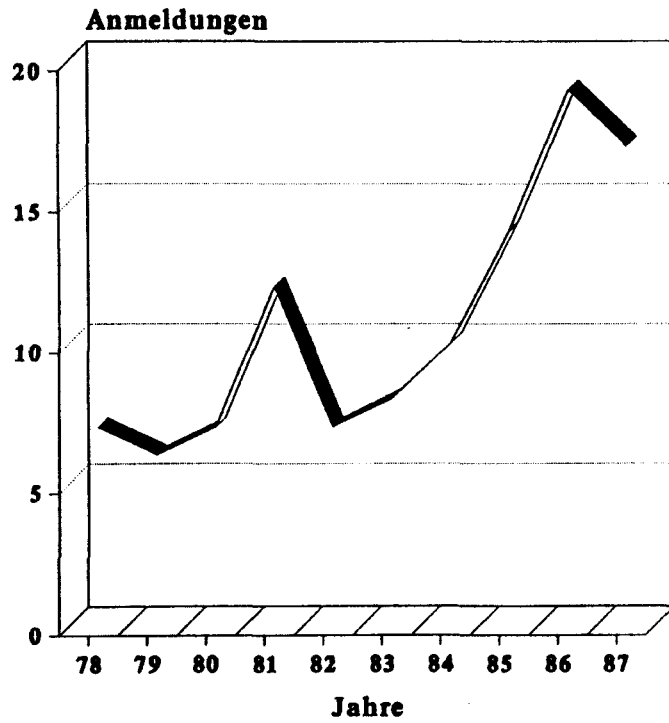


Abbildung 11

### Ciba-Geigy

#### Farben und Lacke (organisch)

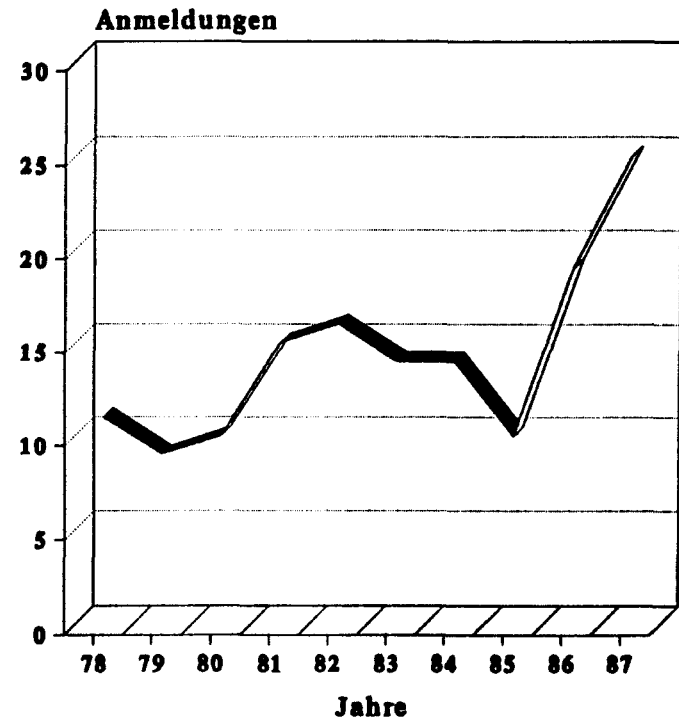


Abbildung 12

- Beimengungen für Farben und Lacke
- Druck- und Schreibtinten
- sonstige Beschichtungszusammensetzungen

recherchiert werden. Steigende Anmeldezahlen sind in den letzten Jahren in der Davonposition Farben und Lacke (organisch) angesiedelt (Abbildung 12). Innerhalb der Davonposition Farben und Lacke kann in 8 Unterpositionen nach der Art der Filmbildner recherchiert werden. Die Recherche ergab, daß die Steigerung der Patentanmeldungen insbesondere auf Anmeldungen zurückzuführen ist, in denen Epoxidharze als Filmbildner verwendet werden (Abbildung 13).

Innerhalb des Teilgebietes Klebstoffe (Abbildung 11) konnte in analoger Vorgehensweise durch statistische Analysen der hierarchisch untergeordneten Bereiche in drei Arbeitsschritten herausgefunden werden, daß die steigenden Anmeldezahlen insbesondere auf Erfindungen beruhen, die die Verwendung von Epoxidharzen für Klebstoffe beinhalten (Abbildung 14).

Zusammengefaßt sind also die steigenden Anmeldezahlen innerhalb der Produktparte Druck/Beschichtungen/Fotografie auf eine erhöhte Forschungsaktivität Ciba-Geigys auf dem Gebiet des Einsatzes von Epoxidharzen als Grundstoff für Farben und Lacke sowie für Klebstoffe zurückzuführen.



### Ciba-Geigy

#### Farben/Lacke auf Epoxidharzbasis

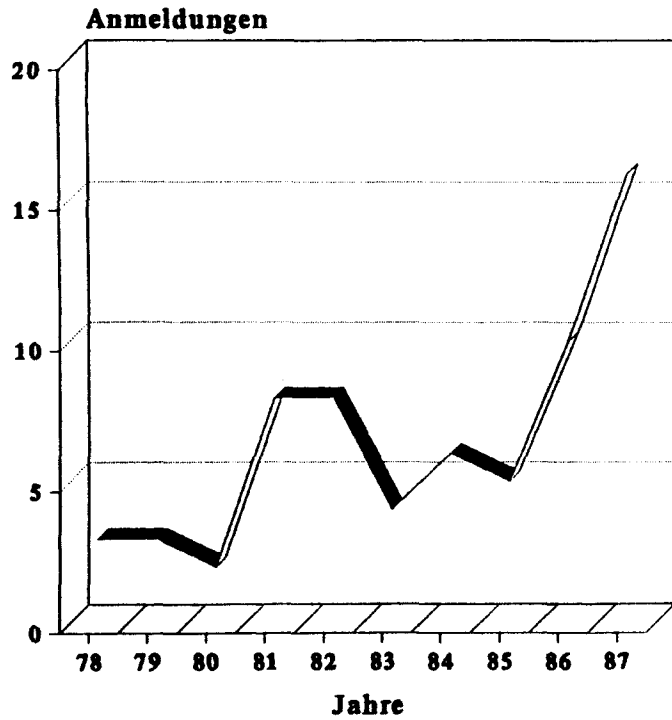


Abbildung 13

### Ciba-Geigy

#### Klebstoffe auf Epoxidharzbasis

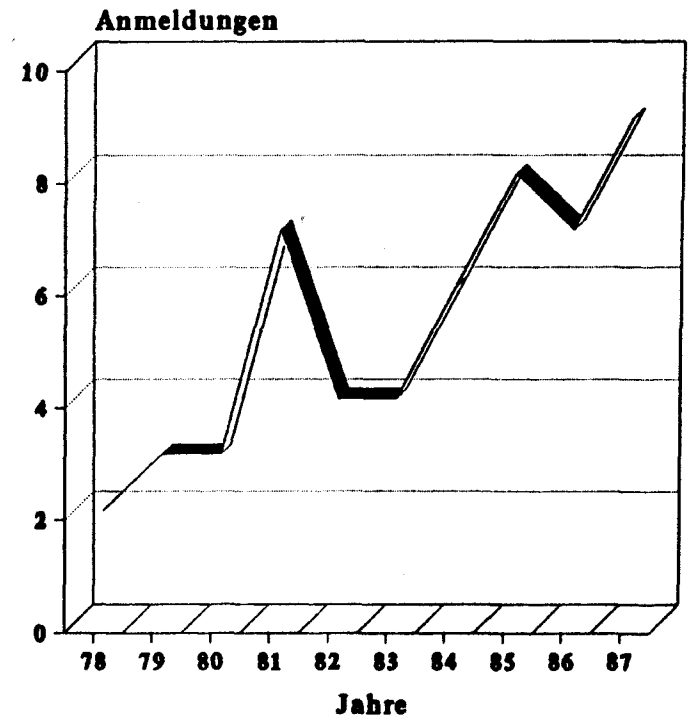
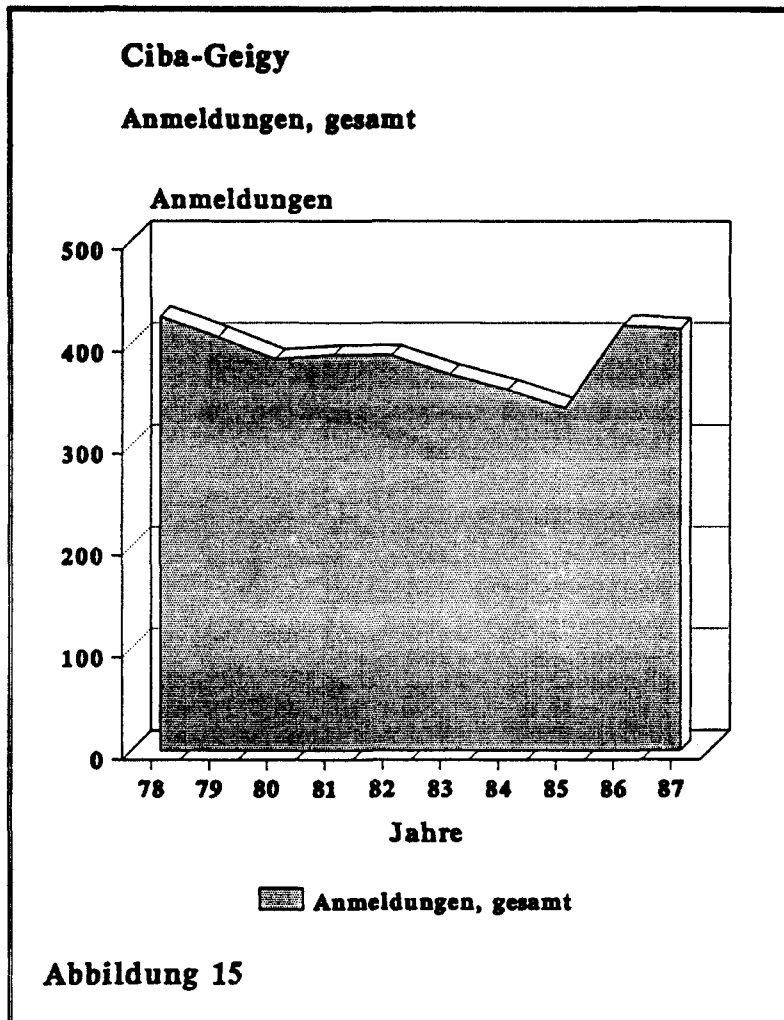
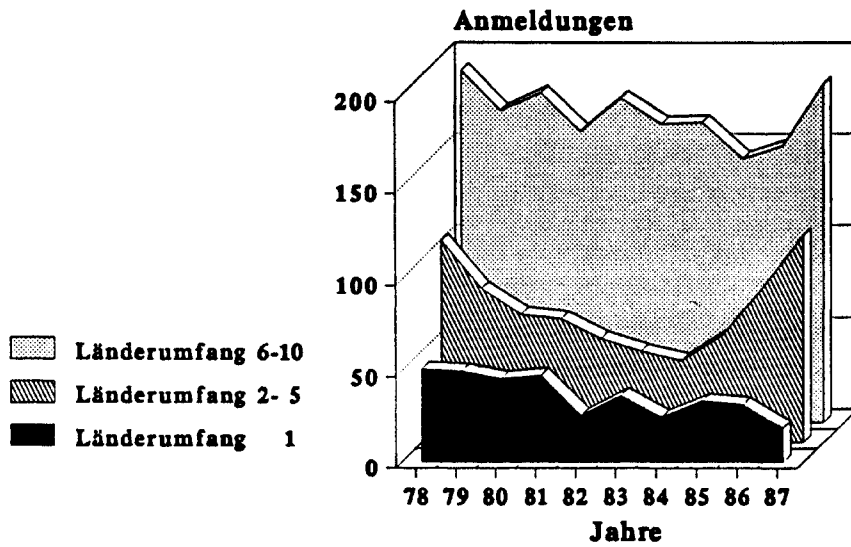


Abbildung 14



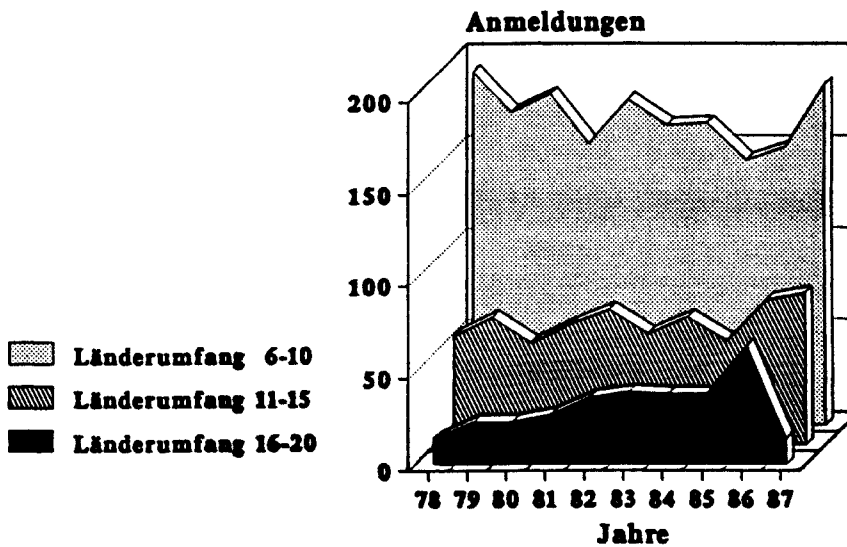
Eine Möglichkeit, die Verwertungskonzeption einer Firma für ihre Forschungsergebnisse zu erkennen, ist die Analyse der Anmeldungen nach ihrem Länderumfang, d.h. ihrer beigemessenen voraussichtlichen wirtschaftlichen Bedeutung. Abbildung 15 zeigt die Anmeldungen Ciba-Geigys von 1978-1987 insgesamt. In den Abbildungen 16 und 17 sind die Ergebnisse dargestellt, die durch die statistische Analyse der Anmeldungen nach ihrem Länderumfang und ihrer zeitlichen Veränderung erhalten wurden. Aus den Grafiken ist zu entnehmen, daß die Anmeldungen, die nur in einem Land vorgenommen werden, abnehmen und bei Ciba-Geigy in den letzten Jahren generell der Trend besteht, die Forschungsergebnisse in immer mehr Ländern schutzrecht-

**Ciba-Geigy**  
 Entwicklung der Anmeldezahlen  
 nach dem Länderumfang



**Abbildung 16**

**Ciba-Geigy**  
 Entwicklung der Anmeldezahlen  
 nach dem Länderumfang



**Abbildung 17**

Tabelle 6

## Ciba-Geigy, Länderumfang 16-20

Sachgebiet	Anzahl der Anmeldungen	%
Therapeutika, Pestizide, Herbizide	53	67,1
Antibiotika, Impfstoffe	12	15,2
Tierische Extrakte, mikrobiologische und allgemeine Extrakte	10	12,7
Prozesse, Apparaturen	4	5,0
insgesamt:	79	

Tabelle 7

## Ciba-Geigy, Impfstoffe, 1986

Von 12 Anmeldungen mit einem Länderumfang von 16-20 wurden angemeldet in:

Anmeldeländer	Anzahl der Anmeldungen	%
AT Österreich	12	100
AU Australien	12	100
BE Belgien	12	100
CH Schweiz	12	100
DE BRD	12	100
ES Spanien	12	100
FR Frankreich	12	100
GB Großbritannien	12	100
IT Italien	12	100
LI Liechtenstein	12	100
LU Luxemburg	12	100
NL Niederlande	12	100
GR Griechenland	11	92
JP Japan	11	92
ZA Südafrika	11	92
DK Dänemark	10	83
PT Portugal	10	83
FI Finnland	4	33
HU Ungarn	4	33
NO Norwegen	4	33
US USA	2	17

lich zu sichern. Besonders deutlich läßt sich dieser Trend bei Anmeldungen mit einem Länderumfang von 16-20 erkennen, wobei zu berücksichtigen ist, daß die Zahlen für 1987 aus objektiven Gründen nicht komplett sind. Auffallend ist der starke Anstieg von Anmeldungen mit einem Länderumfang von 16-20 von 38 Anmeldungen im Jahr 1985 auf 64 im Jahr 1986 (Abbildung 17). Die statistische Untersuchung nach den Sachgebieten gibt Tabelle 6 wieder, nach der es insbesondere Therapeutika, Pestizide und Herbizide sowie Antibiotika und Impfstoffe sind.

Die vorgesehenen Marktgebiete z.B. für die Antibiotika und Impfstoffe können durch die statistische Analyse nach den Anmeldeländern gefunden werden. Tabelle 7 zeigt, daß für die Impfstoffe und Antibiotika die EG-Länder, Australien, Japan und Südafrika als Hauptabsatzgebiete vorgesehen sind.

Diese Beispiele belegen deutlich den Wert der Patentinformation und der maschinellen statistischen Auswertung von Patentinformationsdiensten für die Gewinnung von Marktinformationen für die Unternehmensführung sowie für die Marketing- und Forschungsbereiche.

Dipl.-Chemiker Babst, Peter  
IPO-GmbH  
Informations-, Patent- und Online-Service  
Leninallee 70  
O-4020 Halle/Saale

## Zusammenfassung

Auf Grund des hohen Informationsgehaltes von Patentanmeldungen sind maschinenlesbare Patentinformationsdienste für statistische Analysen hervorragend geeignet. Das Prinzip der Ausarbeitung von Trendanalysen über Erzeugnisse und von Firmennalysen durch statistische Analysen von Patentanmeldungen wird an praktischen Beispielen erläutert.

By reasons of the high information content of patent specifications, patent information services on machinereadable data carriers are excellently suitable to serve as a basis for statistical analyses. The principle of elaboration of trend analyses about products and analyses of companies by means of the statistical analysis of patent applications is explained by practical examples

# **Elektronische Märkte - Ihre Entwicklung, Bedeutung und der Einsatz von Informations- und Kommunikationstechnologien\***

Andreas Himberger, Stefan Zbornik

---

## **Inhalt**

1. Motivation
  - 1.1. Veränderungen im Handel von Gütern und Dienstleistungen
  - 1.2. Entwicklung und Nutzungspotential der IKT
2. Die Idee des Elektronischen Marktes
  - 2.1. Der Elektronische Markt als neue Entwicklung
  - 2.2. 'Elektronischer Markt' als neuer Begriff
3. Entstehungsmuster Elektronischer Märkte
4. Die Bedeutung Elektronischer Märkte für die Unternehmung
  - 4.1. Die Teilnahme an Elektronischen Märkten: ein strategischer Vorteil?
  - 4.2. Wer profitiert von Elektronischen Märkten?
  - 4.3. Alternative Formen des überbetrieblichen IKT-Einsatzes
5. Die Realisierung Elektronischer Märkte
  - 5.1. Kommunikationsmodell Elektronischer Märkte
  - 5.2. Technische Grundelemente Elektronischer Märkte
  - 5.3. Planungsansätze für Kommunikationsarchitekturen Elektronischer Märkte
6. Resümee

## **Referat**

Der vermehrte Einsatz von Informations- und Kommunikationstechnologien im Rahmen des überbetrieblichen Güter- und Dienstleistungsaustausches bewirkt zunehmend die Entstehung Elektronischer Märkte. Elektronische Märkte können konventionelle Handelsbeziehungen tiefgreifend verändern. Eine umfassende, interdisziplinäre Betrachtung dieser Entwicklung drängt sich auf, um sämtliche Potentiale Elektronischer Märkte nutzen zu können.

## **Abstract**

The increasing use of information and communication technologies for the interorganizational exchange of goods and services leads to the development of electronic markets. These markets are able to change the conventional links between the market participants in a tremendous way. A comprehensive and interdisciplinary focus is necessary to use all the potentials of electronic markets.

---

\* Dieses Referat basiert auf der Forschungsarbeit aller Mitglieder des 'Kompetenzzentrums Elektronische Märkte' an der Hochschule St.Gallen.

## **1. Motivation**

Während sich über lange Zeit nur eine kleine Schar von Praktikern und Wissenschaftlern mit technologieinduzierten Veränderungen von Handelsstrukturen befaßte, hat der Begriff 'Elektronischer Markt' in jüngster Zeit Einzug in die öffentliche Diskussion gehalten. Dabei haben empirische Erhebungen [Himberger/Schmid 1991; Zbornik 1991] gezeigt, daß heute der Begriff häufig nur als Schlagwort verwendet wird. Die Vorstellungen über das Einsatzpotential von Informations- und Kommunikationstechnologien (IKT) im Handelsbereich sowie über wirtschaftliche Konsequenzen realisierbarer Veränderungen sind sehr oft verschwommen. Dabei ist bereits heute in verschiedenen Bereichen zu erkennen, in welcher grundlegenden Weise sich Handelsbeziehungen in den kommenden Jahren verändern können.

### **1.1. Veränderungen im Handel von Gütern und Dienstleistungen**

Die Unterstützung des Handels durch Informations- und Kommunikationstechnologien wird meist im Zusammenhang mit dem Trend zu globaleren und dynamischeren Marktstrukturen genannt. Der Handel über weltweite Netzwerke relativiert die Bedeutung geographischer Standorte und ermöglicht die unmittelbare und rationelle Verarbeitung von Informationen. Effizientere Marktstrukturen sowie Märkte mit höherer Transparenz und Liquidität sind in der Regel Folge des computerunterstützten Handels.

Allerdings sind für den Trend zu globaleren Marktstrukturen nicht allein technologische Faktoren ursächlich. Verschärfungen im internationalen Wettbewerb zwingen Unternehmen zu einer globalen Präsenz. Kulturelle Angleichungen ermöglichen den weltweiten Handel variantenarmer und damit elektronisch vergleichsweise leicht handelbarer Produkte. Daraus ist ersichtlich, daß nicht nur die verfügbare Technologie die weltweiten Wirtschaftsstrukturen beeinflussen kann, sondern daß auch umgekehrt Veränderungen der wirtschaftlichen Lage die Automatisierung von Wirtschaftsabläufen fördern können. Insbesondere Änderungen des handelsrechtlichen Rahmens (z.B. GATT, EWR) nehmen hier einen maßgeblichen Einfluß.

### **1.2. Entwicklung und Nutzungspotential der IKT**

Neben der Nutzbarmachung von Neuerungen im Bereich der Basistechnologien wird vor allem die Weiterentwicklung von Systemarchitekturen sowie Programmier- und Datenbankkonzepten auf absehbare Zeit die Gestaltungsmöglichkeiten computergestützter Handelssysteme vergrößern.

Zur Zeit werden außerdem beträchtliche Anstrengungen zur Etablierung weltweit gültiger Standards zum Austausch von Handelsdaten unternommen. Ziel ist es, die bisherigen spezifischen Lösungen durch branchen- und herstellerunabhängige Standards - im speziellen UN/EDIFACT - zu ersetzen.



Zwischen verschiedenen Branchen bestehen beträchtliche Unterschiede, was den Entwicklungsstand des computerunterstützten Handels betrifft. Am weitesten fortgeschritten sind elektronische Handelssysteme im Finanzbereich, in der Reise- bzw. der Tourismusindustrie sowie in der Logistik. In diesen Anwendungsfeldern existieren integrierte Systeme, welche mehrere Phasen einer Markttransaktion (Anbahnung, Verhandlung und Abwicklung) unterstützen oder automatisieren. Die Gründe für die unterschiedlichen Branchenentwicklungen sind vielfältig und häufig 'politischer' Natur. Zudem weisen die Branchen ein unterschiedliches Eignungspotential für den elektronischen Handel auf.

Eines der Hauptprobleme bestehender elektronischer Handelssysteme ist, daß sie oft ohne eine übergeordnete Gesamtkonzeption erstellt wurden. So entstanden unabhängige Teilsysteme mit eingeschränkter Funktionalität, die zudem auf spezifische Märkte beschränkt sind. Gravierend ist auch der Nachteil, daß in derartigen Systemen in der Regel bestehende Handelsabläufe eins-zu-eins auf das neue Kommunikationsmedium übertragen wurden, ohne eine mögliche Umgestaltung in Betracht zu ziehen.

## **2. Die Idee des Elektronischen Marktes**

### **2.1. Der Elektronische Markt als neue Entwicklung**

Im Gegensatz zu bisherigen Ansätzen steht bei Elektronischen Märkten nicht der Rationalisierungs-, sondern vermehrt der Integrations- bzw. Flexibilitätsgedanke [Adam u.a. 1989] im Zentrum. Angestrebt wird die Neugestaltung wirtschaftlicher Koordinationsmechanismen durch Informations- und Kommunikationstechnologien. Diese Aspekte äußern sich wie folgt:

**Integration:** Durchgehende Computer-Unterstützung der Interaktionsprozesse in meist sämtlichen Phasen einer Transaktion, insbesondere des Vertragsabschlusses.

**Flexibilität:** Zunehmende Ersetzung starrer Lieferanten-Kunden-Beziehungen durch den fallweisen Abgleich von Angebot und Nachfrage (Viele-zu-viele-Beziehung statt Eine-zu-viele-Beziehung).

Diese Neugestaltung von Austauschbeziehungen bringt zum einen die Möglichkeit einer (Teil-)Automatisierung von Transaktionen bzw. einzelner Phasen einer Transaktion mit sich. Zum anderen entstehen neue Möglichkeiten der transaktionsbezogenen Verknüpfung von Branchen durch die Bildung von Güter- und Dienstleistungsbündeln.

### **2.2. 'Elektronischer Markt' als neuer Begriff**

Der aktuelle Stand der Begriffsbildung in der Literatur weist ein sehr uneinheitliches Bild aus. Frühe Beiträge, welche die nachfolgende Begriffsbildung auch im deutschsprachigen Raum nachhaltig beeinflusst haben, stammen von einer Forschergruppe am Massachusetts

Institute of Technology [Malone u.a. 1987]. Integrierender Bestandteil der Mehrzahl der Definitionen ist das Vorhandensein von mindestens zwei Marktteilnehmern pro Marktseite. Weitaus weniger eindeutig präsentiert sich die Situation in bezug auf die zu unterstützenden Phasen einer Transaktion [vgl. z.B. Hanker 1990, 349].

Vor dem Hintergrund der in 2.1 vorgestellten Zielsetzungen der Integration und Flexibilität soll der Begriff 'Elektronischer Markt' hier wie folgt beschrieben werden [Krähenmann 1990, 15f.].

Ein **Elektronischer Markt** umfaßt die Gesamtheit der informations- und kommunikationstechnologisch organisierten Austauschbeziehungen innerhalb einer Gruppe von hierarchisch gleichberechtigten Marktpartnern, die ohne Zwang institutionalisierte Transaktionen abschließen.

Elektronische Märkte sind demnach durch Gleichberechtigung und zwanglose Teilnahme gekennzeichnet. In diesem Punkt unterscheiden sie sich wesentlich von den meisten bisherigen überbetrieblichen Systemen, die als elektronische Hierarchien bezeichnet werden. Der Güter- und Dienstleistungsaustausch erfolgt in Elektronischen Märkten durch den Abgleich von Angebot und Nachfrage auf der Grundlage wohldefinierter Regeln (und in diesem Sinne 'institutionalisiert'). Diese Regeln betreffen im Falle Elektronischer Märkte vor allem die elektronisch unterstützte Kommunikation. Durch Hervorkehrung dieses Marktgedankens lassen sich Elektronische Märkte von computerunterstützten Unternehmensverbänden, sogenannten elektronischen Allianzen, unterscheiden.

Die hier vorgestellte Definition erweist sich aus einer (bestimmten) volkswirtschaftlichen Sichtweise als sinnvoll. Es wird dabei deutlich, daß nicht alle Aspekte des interdisziplinären Forschungsfeldes 'Elektronischer Markt' in dieser Kürze wiedergegeben werden können.

### **3. Entstehungsmuster Elektronischer Märkte**

Elektronische Märkte sind - zumindest bislang - in der Regel nicht von Grund auf neu geplant worden, sondern aus bestehenden Computersystemen hervorgegangen. In bezug auf die Entstehung Elektronischer Märkte sind aus heutiger Perspektive vier Muster beobachtbar [Ritz 1991].

#### **Öffnung elektronischer Hierarchien**

Von einzelnen Anbietern realisierte, auf proprietären Kommunikationsstandards basierende Bestell- bzw. Reservationssysteme wurden häufig als wettbewerbsbeschränkende 'competitive weapons' konzipiert. Ihr Ziel war die Erhöhung der Kundenbindung [Petre 1985]. Aufgrund veränderter Rahmenbedingungen ist gegenwärtig eine vermehrte Öffnung solcher Hierarchien durch die Aufnahme konkurrierender Anbieter zu beobachten.

### **Schaffung eines gemeinsamen Absatzkanals**

Ein Elektronischer Markt kann durch mehrere Marktteilnehmer (z.B. einen Branchenverband) gemeinsam geschaffen werden. Hauptsächliches Motiv eines solchen horizontalen Zusammenschlusses ist die gemeinsame Bewältigung der Anforderungen an Know-how und Finanzen. Nicht selten geht ein solcher Elektronischer Markt aus einem gemeinsamen Markt-Informationssystem hervor, indem über die Bereitstellung von Marktinformation hinaus zusätzlich der Vertragsabschluß computerunterstützt wird.

### **Geschlossene elektronische Handelssysteme**

Hauptsächliches Motiv einer geschlossenen Gruppe zur Realisierung eines gemeinsamen elektronischen Handelssystems besteht in der Schaffung effizienter, transaktionskostenminimierender Handelsbeziehungen [vgl. z.B. Picot 1982]. Geschlossene Handelssysteme können sowohl der horizontalen (z.B. bei Wertpapierbörsen [Krähenmann 1990, 46ff.]) als auch der vertikalen Koordination (z.B. bei Frachtbörsen [Schmid/Hohagen 1991, 60]) von Austauschbeziehungen dienen. Solche elektronischen Koordinationssysteme sind - zumindest von der Idee her - auch innerhalb von Unternehmen denkbar.

### **Erweiterung der Tätigkeit von Mehrwertdienste-Anbietern**

Anbieter von Mehrwertdiensten (VANS) bieten ihren Kunden z.B. den Zugriff auf Transaktionssysteme und Online-Datenbanken an. Deren Kommunikationsnetze können künftig über die bloße Informationsvermittlung hinaus auch zur Unterstützung des Abschlusses von Markttransaktionen genutzt werden. VANS-Anbieter besitzen zwei wesentliche Vorteile: die Verfügbarkeit von Infrastrukturen und Know-how sowie eine neutrale Position in bezug auf die abzuwickelnden Transaktionen. Allerdings werden Anbieter und Vermittler kaum widerstandslos zulassen, daß Branchenfremde in ihre Absatzkanäle eindringen [Butler Cox 1990, 28ff.].

## **4. Die Bedeutung Elektronischer Märkte für die Unternehmung**

### **4.1. Die Teilnahme an Elektronischen Märkten: ein strategischer Vorteil?**

Bei der Einführung neuer Technologien stehen in der Regel zunächst Rationalisierungsüberlegungen im Vordergrund. Das Streben nach Wettbewerbsvorteilen schließt sich an. Dies gilt vor allem dann, wenn Beziehungen zu Kunden oder Lieferanten betrachtet werden. Der Einsatz neuer Technologien verändert diese Außenbeziehungen des Unternehmens und damit den Wettbewerb in grundlegender Weise [vgl. Porter/Millar 1986, 31ff.]. Es ist allerdings zu erkennen, daß die zwischenbetriebliche Nutzung neuer Technologien in einigen Bereichen bereits vom Wettbewerbsvorteil zur -notwendigkeit geworden ist. So ist die Einführung von EDI heute oft als Reaktion auf die Aktivitäten von Marktpartnern zu sehen und nicht auf eigenen Entschluß zurückzuführen [Benjamin u.a. 1990, 35f.].

Wie unsere Praxisbefragung zeigt, dominieren bei Elektronischen Märkten vielfach noch operative Überlegungen [Himberger/Schmid 1991, 20ff.]. 'Pionier'-Vorteile gegenüber Mitbewerbern, die sich in vielen Bereichen (noch) bieten, werden meist nicht in entsprechendem Maße gesehen. Allerdings müssen innerbetriebliche Abläufe der technologiemotivierten Um-

gestaltung im zwischenbetrieblichen Bereich folgen [Sedran 1991, 18ff.], damit die erreichten Vorteile nicht der Imitation durch Konkurrenten preisgegeben und damit von kurzer Dauer sind. Zur Wettbewerbsnotwendigkeit ist das Engagement in Elektronische Märkte bislang nur in wenigen Branchen, vor allem im Finanzsektor und in weiten Teilen des Touristikbereichs, geworden. Eine Änderung dieser Situation ist angesichts derzeitiger Entwicklungen jedoch absehbar (vgl. Abschnitte 1.1 und 1.2).

#### **4.2. Wer profitiert von Elektronischen Märkten?**

Die Entwicklung Elektronischer Märkte läßt sich nur dann konkret beurteilen, wenn branchenbezogene Überlegungen berücksichtigt werden. Unterschiedliche Entwicklungen in der Vergangenheit zeigen dies (vgl. Abschnitt 1.2). Neue Technologien spielen vor allem dann eine bedeutende Rolle, wenn die Informationsintensität der Wertschöpfungskette oder des Produkts in der jeweiligen Branche potentiell hoch ist [Porter/Millar 1986, 34].

Einen weiteren Einblick vermittelt die Differenzierung nach Anbietern und Nachfragern. So scheint der Einsatz neuer Technologien Nachfragern tendenziell mehr Vorteile zu bieten als Anbietern [Sedran 1991, 21] und dies nicht (nur), weil die meisten überbetrieblichen Systeme käufer- oder vermittlerorientiert sind. Der Grund wird hauptsächlich in den unterschiedlichen Marktpositionen von Anbietern und Nachfragern gesehen [Benjamin u.a. 1990, 35 ff.].

Die größten Auswirkungen Elektronischer Märkte dürften sich für solche Unternehmen ergeben, die beide Positionen auf diesen Märkten einnehmen [Malone u.a. 1989, 168]. Für das produzierende Gewerbe werden die Vorteile aus der Optimierung des Informationsflusses entlang der Wertschöpfungskette überwiegen. Vermittler, die das Produkt selbst nicht verändern (Broker, Händler), sind hingegen gefordert, ihre Existenzberechtigung unter Beweis zu stellen. Allerdings werden durch Elektronische Märkte auch neue Tätigkeitsfelder geschaffen, die in erster Linie diesen Vermittlern nahestehen (z.B. der Betrieb eines Elektronischen Marktes bzw. die Erbringung spezieller Dienstleistungen, vgl. Abschnitte 3 und 5.2).

#### **4.3. Alternative Formen des überbetrieblichen IKT-Einsatzes**

Da die (künftige) Nutzung von Informations- und Kommunikationstechnologien im überbetrieblichen Bereich weitgehend unbestritten scheint, wird die Fragestellung häufig auf Umfang und Form dieses Einsatzes aus Sicht der Unternehmung konzentriert [vgl. z.B. Malone u.a. 1987].

Der Elektronische Markt wird künftig eine stärkere Bedeutung erlangen [ebd.]. Allerdings stellt er 'nur' eine von verschiedenen Ausprägungen überbetrieblicher Systeme zur Unterstützung des Güter- und Dienstleistungsaustauschs dar (vgl. Abschnitt 2.2 und [Benjamin u.a. 1990, 31]). Der Einsatz bzw. die Auswahl eines konkreten Systems hängt neben zahlreichen anderen Kriterien vor allem vom Zweck des Systems ab.

So werden insbesondere elektronische Allianzen, deren Aufgabe in der Bereitstellung eines stabilen, häufig bilateralen Verbindungskanals zwischen Unternehmen besteht, nur selten durch Elektronische Märkte ersetzt werden. Als Beispiel sei hier die CAD/CAM-Verbindung zweier Unternehmen zur gemeinsamen Produktentwicklung genannt [Malone u.a. 1987, 494]. Wichtig ist hierbei, daß alle Seiten freiwillig an dieser 'Wertschöpfungspartnerschaft' [Hautz 1991] teilnehmen und gemeinsam davon profitieren.

Elektronische Hierarchien, die auf die 'Anbindung' von Kunden oder Lieferanten durch den Aufbau hoher Wechselkosten abzielen, werden sich dagegen angesichts derzeitiger Standardisierungsentwicklungen meist nicht auf Dauer erhalten lassen (vgl. Abschnitt 3). Unternehmen, die solche Systeme etabliert haben oder etablieren wollen, müssen sich auf deren künftige Öffnung einstellen [Malone u.a. 1987, 496 und 1989, 168].

## **5. Die Realisierung Elektronischer Märkte**

Auf der Grundlage vorwiegend ökonomisch motivierter Überlegungen zu elektronischen Märkten soll in diesem Kapitel aufgezeigt werden, welche informations- und kommunikationstechnologischen Aspekte für die Realisierung Elektronischer Märkte von besonderer Relevanz sind.

### **5.1. Kommunikationsmodell Elektronischer Märkte**

Der Austausch von Gütern und Dienstleistungen in Märkten erfordert eine intensive Kommunikation zwischen den Marktteilnehmern [Szyperski/Nathusius 1975]. Eines der wesentlichen Merkmale Elektronischer Märkte ist, daß die bislang konventionellen Kommunikationsformen durch elektronische beziehungsweise elektronisch unterstützte Formen abgelöst werden. Vor diesem Hintergrund lassen sich Systeme zur Realisierung Elektronischer Märkte im wesentlichen als Kommunikationssysteme auffassen [Schmid 1991, 81].

Als Grundlage für die Planung entsprechender Systemarchitekturen müssen spezifische Marktkommunikationsmodelle dienen, welche sich einerseits aus Grundmodellen der Kommunikationstheorie und andererseits aus Modellen im Bereich der Marketing-Kommunikation [z.B. Kotler 1982, 491-493; Tietz 1978, 62-73] ableiten lassen. Ausgehend von der Beobachtung konventioneller Märkte, realisierter Elektronischer Märkte und kommunikationstheoretischer Überlegungen, lassen sich alternative Modelle ableiten, welche die Kommunikationsmechanismen in (Elektronischen) Märkten erklären [Schmid/Zbornik 1991].

Elektronische Märkte haben in den Phasen der Anbahnung, des Abschlusses und der Abwicklung einer Markttransaktion unterschiedliche Informations- und Kommunikationsdienstleistungen für die Marktteilnehmer zu erbringen. Ausgehend von diesen ökonomisch motivierten Phasen einer Markttransaktion (vgl. Abschnitt 1.2), lassen sich unter dem Aspekt der Kommunikationsabläufe und -beziehungen fünf spezifischere Phasen identifizieren (s.u.). Die

Bedeutung dieser Kommunikationsphasen ist je nach Marktform und Marktorganisation, aber auch nach der Art der Handelsobjekte und der beteiligten Marktteilnehmer unterschiedlich. In zahlreichen Fällen werden nicht alle fünf Phasen durchlaufen oder es treten Zyklen auf. In jeder der fünf Phasen lassen sich bestimmte Kommunikationsbeziehungen unter den Marktteilnehmern und den jeweiligen Marktsystemen identifizieren [Schmid/Zbornik 1991, 42-49].

### **Marktinformationsbeschaffung**

In der ersten Phase beschaffen sich die Teilnehmer an einem bestimmten Markt Informationen über das Marktgeschehen, vor allem über die Preissituation. Diese Phase hat in allen Märkten eine große Bedeutung. Insbesondere bei elektronischen Börsen ist diese Funktionalität realisiert. Die Kommunikationsketten zwischen den Marktteilnehmern und den Informationsquellen können einseitig oder zweiseitig sein. Im ersten Fall werden die Marktinformationen zu bestimmten Zeitpunkten oder laufend an die Marktteilnehmer gesendet. Im zweiten Fall können die Marktteilnehmer nach Bedarf Informationen abfragen.

### **Marktpartnersuche**

In der zweiten Phase sind die Kontakte zwischen den Marktteilnehmern mit komplementären Interessen herzustellen. Dies erfolgt im allgemeinen durch einen zentralen Vermittlungsdienst. Das Vermittlungskriterium besteht dabei aus den Angeboten und Nachfragen der Marktteilnehmer. Ein zentraler Bestandteil dieser Angebote und Nachfragen ist die Beschreibung des Handelsobjektes [vgl. Schmid 1990, 11f.]. Es kann sich dabei um Referenzen auf bekannte Handelsobjekte oder um ausgedehnte Spezifikationen dieser Objekte handeln.

### **Partnerinformationsbeschaffung**

Häufig hat der Marktteilnehmer ein Interesse an der Person bzw. Institution, mit der er eine Markttransaktion abschließen will. Es ist somit sinnvoll, von einer 'externen' und möglicherweise unabhängigen Stelle Informationen über den Marktpartner erhalten zu können. Es stellen sich in diesem Zusammenhang Fragen des Datenschutzes sowie der Qualität (Zuverlässigkeit beziehungsweise Richtigkeit) der bereitgestellten Informationen.

### **Vertragsaushandlung**

In vielen Fällen genügt die Vermittlung zwischen Anbietern und Nachfragern aufgrund der Angebote und Nachfragen nicht zum Abschluß einer Markttransaktion. Eine zusätzliche Aushandlung zwischen den beiden Marktpartnern ist dann notwendig. Der Aushandlungsdialog kann sowohl über synchrone als auch über asynchrone Kommunikation zwischen den Marktpartnern erfolgen. Das Spektrum der auszutauschenden Informationen reicht von einer einfachen Bestätigung der Markttransaktion bis hin zu ausführlichen Spezifikationen der Handelsobjekte und komplexen Abschlußbedingungen und -modalitäten.

### **Transaktionsabwicklung**

Nach dem Abschluß einer Markttransaktion wird in der Abwicklungsphase häufig eine Kommunikation mit weiteren Wirtschaftssubjekten erforderlich. Die Problembereiche der informationstechnischen Logistik und der Finanzlogistik [Szyperski/Kronen 1991, 5f.] sind hier von Bedeutung. Das Muster der Kommunikationsbeziehungen und -netze in der Abwicklungsphase kann unterschiedlich aussehen. Es hängt vom jeweiligen Markt, der angestrebten

beziehungsweise realisierten Integration verschiedener elektronischer Marktsysteme untereinander und der Integration dieser in betriebliche Applikationen ab.

Die gemeinsame Betrachtung der fünf diskutierten Phasen führt zu einem Phasenmodell der Marktkommunikation. Es handelt sich dabei um die Darstellung eines möglichst allgemeingültigen Kommunikationsschemas für Elektronische Märkte (s. Abb. 1).

Kommunikationsphasen	Marktinformationsbeschaffung	Partnersuche	Partnerinformationsbeschaffung	Vertragsaushandlung	Transaktionsabwicklung
Transaktionsphasen	Anbahnung		Abschluß		Abwicklung
Funktionen	Bereitstellung von Informationen über das Marktgeschehen, um den Marktteilnehmern das Absetzen von Angeboten bzw. Nachfragen zu ermöglichen.	Zusammenführung von Marktteilnehmern aufgrund komplementärer Angebote und Nachfragen.	Bereitstellung detaillierter Informationen über konkrete Marktpartner.	Unterstützung bei der Aushandlung eines Abschlusses.	Unterstützung bei der Abwicklung einer Markttransaktion und Zuschaltung weiterer Marktdienstleister bzw. Elektronischer Märkte.
Auszu-tauschende Informationen	<ul style="list-style-type: none"> <li>- Marktinformationen</li> <li>- allgemeine Produktinformationen</li> <li>- Marktteilnehmerinformationen</li> <li>- Brancheninformationen</li> <li>- gesamtwirtschaftliche Rahmeninformationen</li> </ul>	<ul style="list-style-type: none"> <li>- Informationen über potentielle Handelspartner</li> <li>- gezielte Produktinformationen</li> </ul>	<ul style="list-style-type: none"> <li>- Unternehmensinformationen</li> <li>- Bonitätsinformationen</li> <li>- Referenzinformationen</li> </ul>	<ul style="list-style-type: none"> <li>- Informationen über Abschlusskonditionen</li> <li>- vertragsrechtliche Informationen</li> </ul>	<ul style="list-style-type: none"> <li>- Transportinformationen</li> <li>- Versicherungsinformationen</li> <li>- Finanzinformationen</li> <li>- Informationen über weitere zugängliche Elektronische Märkte</li> </ul>

Abb. 1: Kommunikationsphasen Elektronischer Märkte

## 5.2. Technische Grundelemente Elektronischer Märkte

Elektronische Märkte, verstanden als verteilte (Hard- und Software-)Systeme, bedürfen zu ihrer logischen Funktionsfähigkeit mindestens vier wesentlicher Elemente, auf die sich die Marktteilnehmer einigen müssen. Dies sind in der Regel standardisierte Kommunikationskanäle, standardisierte und möglichst formale Marktsprachen, elektronische Marktdienste und Marktapplikationen bei den Teilnehmern [Schmid/Zbornik 1991, 54-64].

### Gemeinsame Kommunikationskanäle

Unter Kommunikationskanälen werden hier die Kommunikationsverbindungen unter den Marktteilnehmern und zu Marktdiensten verstanden. In bezug auf die Realisierung Elektronischer Märkte wird es notwendig sein, sich auf (international) standardisierte Kommunikationseinrichtungen zu einigen. Die gemeinsamen Kommunikationskanäle werden mit Vorteil auf öffentlichen Kommunikationsnetzen und -diensten (z.B. X.25, ISDN, X.400 usw.) zu realisieren sein.

## **Gemeinsame Marktsprachen**

Für die direkte Kommunikation zwischen den Marktteilnehmern kann es ausreichend sein, sich lediglich auf eine gemeinsame natürliche Sprache zu einigen, in der nach bestimmten Regeln Marktinformationen ausgetauscht werden. Sollen diese Informationen aber von elektronischen Marktdiensten und Marktapplikationen verarbeitet werden, so ist eine Formalisierung der Marktsprache erforderlich. Je weitergehend diese Formalisierung gelingt, desto umfangreicher kann die Funktionalität des Elektronischen Marktes insbesondere in bezug auf die Vermittlungs- und Entscheidungsfunktion ausgestaltet werden. Als Marktsprachen beziehungsweise Bestandteile von Marktsprachen kommen zum Beispiel UN/EDIFACT, STEP [Nowacki 1988, 351ff.] und EAN-Codes in Frage.

## **Elektronische Marktdienste**

Marktdienste sind informations- und kommunikationstechnische Einrichtungen in Märkten, welche den Marktteilnehmern elektronische Dienstleistungen anbieten, die bis anhin auf konventionelle Weise erbracht wurden. Unter diesen Marktdiensten sind zum Beispiel Kommunikationsdienste, direkt handelsunterstützende Dienste (Marktvermittlungsdienste, Program-Trading-Dienste, Clearing-Dienste), Assistenzdienste (Informationsdienste, Übersetzungsdienste, Verzeichnisdienste, Gatewaydienste) und Sicherheitsdienste (Zertifikationsdienste, Protokollierungsdienste, Notariatsdienste) zu verstehen.

## **Marktapplikationen bei den Teilnehmern**

Jeder Teilnehmer an Elektronischen Märkten benötigt möglichst auf seine individuellen Bedürfnisse zugeschnittene Marktapplikationen, welche den Zugang zum Handel über elektronische Kanäle erst ermöglichen. An die Funktionalität von Marktapplikationen sind unterschiedliche Anforderungen zu stellen, welche vom einfachen Hinterlegen eines Angebots oder einer Nachfrage bis hin zur Erstellung umfassender Güter- und Dienstleistungsbündel reichen.

## **5.3. Planungsansätze für Kommunikationsarchitekturen Elektronischer Märkte**

Im Rahmen der Planung von Kommunikationsarchitekturen für Elektronische Märkte existieren zwei grundsätzliche Vorgehensweisen [Schmid/Zbornik 1991, 74-78]. Der eine Ansatz zielt von vornherein auf die Nutzung bzw. Integration der bereits vorhandenen öffentlichen informations- und kommunikationstechnischen Infrastrukturen bzw. Ressourcen. Dies sind einerseits Kommunikationsnetze und -dienste, andererseits aber auch öffentliche Online-Datenbanken, VANS usw. Der andere Ansatz geht zunächst davon aus, daß die für Elektronische Märkte erforderlichen Ressourcen neu zu konzipieren und bereitzustellen sind. Hierdurch wird der Gestaltungsspielraum für die Realisierung Elektronischer Märkte grundsätzlich erweitert. Erst in einem nächsten Schritt werden bestehende kommunikationstechnische Basisinfrastrukturen explizit in die Überlegungen einbezogen. Abhängig von den jeweiligen Märkten bzw. Branchen und ökonomischen Rahmenbedingungen können beide Ansätze ihre Berechtigung haben.



Bezüglich der künftigen Entwicklungen im Bereich der Kommunikationstechnologien für Elektronische Märkte lassen sich zwei Schwerpunkte identifizieren. Die Standardisierung von Kommunikationsdiensten, insbesondere im Rahmen des ISO/OSI-Referenzmodells, schlägt sich allmählich auf breiter Basis in Form von Produkten und verfügbaren Infrastrukturen nieder. Neben der Standardisierung der anwendungsorientierten Dienste (MHS, FTAM, VTS, ...) ist auf die Entwicklungen im Bereich der Transportdienste hinzuweisen. Mit Schmalband- und Breitband-ISDN ist eine flächendeckende Netzinfrastruktur am Entstehen [Knüsel 1991], deren Übertragungskapazität diejenige der heute verfügbaren Fernsprechnetze sowie der Paketvermittlungsnetze bei weitem übersteigt und den Austausch von Grafiken, Bildern, bewegten Bildern und Ton erlaubt. Diese Entwicklungen werden in naher Zukunft die effiziente Übertragung von komplexen und multimedialen Produktbeschreibungen in Elektronischen Märkten erlauben.

## **6. Resümee**

Wie in diesem Referat deutlich wird, sind 'Elektronische Märkte' ein neues und interdisziplinäres Forschungsgebiet. Ein wissenschaftliches Fundament, auf dem die Forschung aufbauen kann, ist nur in Ansätzen vorhanden. Überlegungen verschiedener wissenschaftlicher Disziplinen haben hier einzufließen, um das Gebiet inhaltlich fassen zu können. Mit Ökonomie sowie Informations- und Kommunikationstechnologie werden in diesem Referat zwei Richtungen ausgewählt, die neben methodischen und softwaretechnologischen Überlegungen Schwerpunkte unserer Arbeit darstellen.

Als neues und interdisziplinäres Betätigungsfeld haben es 'Elektronische Märkte' schwer, von Forschung und Praxis akzeptiert zu werden. Unsere Arbeit ist in dieser Hinsicht als Anstoß für weitere Diskussionen zu verstehen.

## Literaturverzeichnis

[Adam u.a. 1989]

Adam, D.; Backhaus, K.; Meffert, H.; Wagner, H.: Integration und Flexibilität, Eine Herausforderung für die Allgemeine Betriebswirtschaftslehre, 51. Wissenschaftliche Jahrestagung des Verbandes der Hochschullehrer für Betriebswirtschaftslehre e.V., Wiesbaden: Gabler Verlag, 1989.

[Benjamin u.a. 1990]

Benjamin, R.I.; de Long, D.W.; Scott Morton, M.S.: Electronic Data Interchange: How Much Competitive Advantage? In: Long Range Planning, 1/90, S. 29-40.

[Butler Cox 1990]

Butler Cox Foundation: Electronic Marketplaces. Research Report No. 77, London, 1990.

[Cash/Konsynski 1985]

Cash, J.I.Jr.; Konsynski, B.R.: IS Redraws Competitive Boundaries. In: Harvard Business Review, 2/85, S. 134-142.

[CEEM 1990, 31ff.]

Projektgruppe CC EM: Positionierung und Darstellung des Problembereichs 'Elektronische Märkte'. Arbeitsbericht IM2000/CEEM/6, Version 2.0, St.Gallen, 31.8.90.

[Hanker 1990]

Hanker, J.: Die strategische Bedeutung der Informatik für Organisationen, industrieökonomische Grundlagen des strategischen Informatik-Managements. Stuttgart: Teubner, 1990.

[Hautz 1991]

Hautz, E.: Elektronische Beschaffung und Logistik. In: Unterlagen zur 2. Jahrestagung des Forschungsprogramms IM 2000: 'Die elektronische Revolution der Märkte: Chancen und Gefahren neuer Koordinationsformen in der Wirtschaft'. Thun, 13.9.91.

[Himberger/Schmid 1991]

Himberger, A.; Schmid, M.: Empirische Studie zum Themengebiet 'Elektronische Märkte'. Arbeitsbericht IM2000/CEEM/11, St.Gallen, 26.5.91.

[Hubmann 1989]

Hubmann, H.E.: Elektronisierung von Beschaffungsmärkten und Beschaffungshierarchien. München: VVF, 1989.

[Knüsel 1991]

Knüsel, G.: SWISSNET - ISDN in der Schweiz. In: Online, 5/91, S.20-29.

[Kotler 1982]

Kotler, P.: Marketing-Management. 4.Auflage, Stuttgart: Poeschel, 1982.

[Krähenmann 1990]

Krähenmann, N.: Die Veränderung institutioneller Abläufe durch Elektronische Märkte -

dargestellt an einem Beispiel aus dem Finanzbereich. Unveröff. Diplomarbeit, Hochschule St.Gallen, 1990.

[Malone u.a. 1987]

Malone, T.W.; Yates, J.; Benjamin, R.I.: Electronic Markets and Electronic Hierarchies. In: Communications of the ACM, 6/87, S. 484-497.

[Malone u.a. 1989]

Malone, T.W.; Yates, J.; Benjamin, R.I.: The Logic of Electronic Markets. In: Harvard Business Review, 3/89, S. 166-172.

[Nowacki 1998]

Nowacki, H.: Übertragung von Produktbeschreibungen über Rechnernetze. In: Informati- onstechnik, 5/88, S. 349-356.

[Petre 1985]

Petre, P.: How to Keep Customers Happy Captives. In: Fortune, 2.9.85, S. 42-46.

[Picot 1982]

Picot, A.: Transaktionskostenansatz in der Organisationstheorie; Stand der Diskussion und Aussagewert. In: Die Betriebswirtschaft, 2/82, S. 267-284.

[Picot u.a. 1991]

Picot, A.; Neuburger, R.; Niggli, J.: Ökonomische Perspektiven eines 'Electronic Data In- terchange'. In: Information Management, 2/91, S. 22-29.

[Porter/Millar 1986]

Porter, M.E.; Millar, V.E.: Wettbewerbsvorteile durch Information. In: Harvard Manager, 1/86, S. 26-35.

[Ritz 1991]

Ritz, D.: Entstehungsmuster Elektronischer Märkte. Arbeitsbericht IM2000/CCEM/10 Version 2.0, St.Gallen, 20.8.91.

[Schmid 1991]

Schmid, B.: Neue Dimension dank Informationsverbund - Entstehung orts- und zeitloser Märkte. In: Neue Zürcher Zeitung, 28.5.91, S. 81.

[Schmid 1990]

Schmid, M.: Ein objektorientiertes Modell Elektronischer Märkte. Arbeitsbericht IM2000/ CCEM/5, 28.5.90.

[Schmid/Hohagen 1991]

Schmid, M.; Hohagen, U.: Stand und Entwicklungstendenzen elektronischer Koordinati- onssysteme in der Logistik. In: Thexis, 3/91, S. 57-61.

[Schmid/Zbornik 1991]

Schmid, M.; Zbornik, S.: Kommunikationsmodelle für Elektronische Märkte und Infra- strukturen zu deren Realisierung. Arbeitsbericht IM2000/CCEM/12, St.Gallen, 9.7.91.

[Sedran 1991]

Sedran, T.: Wettbewerbsvorteile durch EDI? In: Information Management, 2/91, S. 16-21.

[Sloman/Kramer 1988]

Sloman, M.; Kramer, J.: Verteilte Systeme und Rechnernetze. Wien: Hanser, 1988.

[Szyperski/Nathusius 1975]

Szyperski, N.; Nathusius, K.: Information und Wirtschaft. Frankfurt: Campus Verlag, 1975.

[Szyperski/Kronen 1991]

Szyperski, N.; Kronen, J.: Informationstechnik- und Unternehmensstrategien im Wechselspiel - Outsourcing und strategische Allianzen als wichtige Alternativen. In: Schwichtenberg, G. (Hrsg): Organisation und Betrieb von Informationssystemen, Berlin u.a.: Springer, 1991.

[Tietz 1978]

Tietz, B.: Marketing, Tübingen: Mohr 1978.

[Zbornik 1991]

Zbornik, S.: Studie zur elektronischen Unterstützung des Gebrauchcomputerhandels. Arbeitsbericht IM2000/CCEM/10, St.Gallen, 8.4.91.

---

## Kontaktadressen

Dipl.-Wirt.-Ing. Andreas Himberger

X.400: s=himberger;ou=csghsg5a;o=bitnet;p=switch;a=arcom;c=ch

Bitnet: himberger@csghsg5a

Dipl.-Inf.-Wiss. Stefan Zbornik

X.400: s=zbornik;ou=csghsg5a;o=bitnet;p=switch;a=arcom;c=ch

Bitnet: zbornik@csghsg5a

Hochschule St.Gallen  
Institut für Wirtschaftsinformatik  
Kompetenzzentrum Elektronische Märkte  
Dufourstr. 50  
CH-9000 St.Gallen

Tel.: ++41 71 302 297

Fax: ++41 71 302 771

## **Europäisches Informationsrecht: Neuere Entwicklungen**

Herbert Burkert, Jürgen W. Goebel

Informationen und ihre Nutzung lassen sich nicht nationalisieren. Der grenzüberschreitende Datenverkehr ist zu einer Selbstverständlichkeit geworden und man spricht bereits von einem europäischen Informationsmarkt. Diese Entwicklung bringt auch rechtliche Fragestellungen mit sich, deren Lösung allein auf nationaler Ebene notwendigerweise unvollständig und unbefriedigend bleiben muß. Sowohl der Europarat als auch die Kommission der Europäischen Gemeinschaften haben sich daher schon frühzeitig den damit verbundenen Fragestellungen zugewandt und auf einigen Gebieten Lösungsvorschläge unterbreitet. Dieser Beitrag will schwerpunktmäßig einige besonders wichtige Themenbereiche vorstellen.

### **(1) Informationszugangsrecht**

Immer häufiger wird auf die Wichtigkeit rechtlicher Vorschriften hingewiesen, die einen allgemeinen Zugang zu Informationen im öffentlichen Sektor gewährleisten sollen. Für den Bereich der Umweltinformationen hat die Europäische Gemeinschaft (EG) bereits eine entsprechende Richtlinie erlassen. Vergleichbare Empfehlungen zum Informationszugang generell hat der Europarat bereits im Jahre 1981 abgegeben und einige europäische und überseeische Staaten haben nationale Zugangsregelungen erlassen. Wenig Beachtung hat diese Thematik bisher in der Bundesrepublik Deutschland gefunden.

### **(2) Datenschutz**

Anders sieht es im Bereich des Datenschutzrechts aus; dort existieren bereits viele nationale Regelungen, die allerdings in den einzelnen Staaten zu einem unterschiedlichen Schutzniveau

veau führen. Auf internationaler Ebene erwähnenswert zur Beseitigung dieses Datenschutzgefälles sind die Guidelines der OECD sowie die Datenschutzkonvention des Europarats aus dem Jahre 1985, die erste völkerrechtliche Regelung dieser Thematik. Die EG ist erst relativ spät in diesem Bereich tätig geworden und hatte ihre Mitglieder zunächst nur aufgefordert, die Europarats-Konvention zu ratifizieren. Ende 1990 hat sie dann selbst einen Entwurf für eine Richtlinie zum Schutze personenbezogener Daten vorgelegt.

### **(3) Urheberschutz für Software und Datenbanken**

Trotz der Berner Übereinkunft ist im Hinblick auf den Urheberschutz von Computerprogrammen und Datenbanken nach wie vor eine uneinheitliche Rechtslage festzustellen. Die EG hat daher nach langwierigen Verhandlungen im Mai 1991 eine Richtlinie über den Rechtsschutz von Computerprogrammen verabschiedet, die bis Ende 1992 in nationales (Urheber-)Recht umgesetzt werden muß. Insbesondere in der Bundesrepublik ergeben sich wegen der bisherigen Rechtsprechung des Bundesgerichtshofs zum Urheberschutz von Software erhebliche Probleme, die bis dahin gelöst sein müssen. Derzeit werden in der EG-Kommission (insbesondere in der Generaldirektion XIII) Überlegungen angestellt, eine ähnliche Richtlinie für den Rechtsschutz von Datenbanken zu erarbeiten.

### **(4) Haftung für fehlerhafte Informationsprodukte**

Noch relativ ungeklärt ist die Frage, welche Rechte dem Nutzer von fehlerhaften Informationsprodukten zustehen, wenn ihm aus diesen ein weitergehender Schaden entstanden ist. Die EG-Kommission befürchtet wegen der uneinheitlichen Rechtslage in den einzelnen EG-Staaten Wettbewerbsverzerrungen im Informationsmarkt. Eine vereinheitlichende Regelung, über die sowohl die EG als auch der Europarat schon seit einiger Zeit nachdenken, stößt aber auf ganz erhebliche Bedenken, da dadurch massiv in das nationale Gewährleistungs- und Haftungsrecht eingegriffen würde.

## **Referenten**

Herbert Burkert

Wissenschaftlicher Mitarbeiter bei der Gesellschaft für Mathematik und Datenverarbeitung mbH (GMD) - Forschungsstelle für Informationswirtschaft

Schönhauser Str. 64

D-5000 Köln 51

Dr. Jürgen W. Goebel

Rechtsanwalt und Lehrbeauftragter für Informationsrecht und Rechtsinformatik

Melibocusstr. 52a

D-6000 Frankfurt am Main 71

# **Der Europäische Markt für elektronische Informationsdienste: Entwicklungstrends und Perspektiven**

Werner Schwuchow, Karl A. Stroetmann

## **Inhalt**

1. Einführung: Von der Informationswirtschaft zur intellektuellen Dienstleistungsgesellschaft
2. Initiativen der EG zur Förderung des Informationsmarktes
3. Entwicklung elektronischer Informationsdienste
  - 3.1 Begriffliche Abgrenzung
  - 3.2 Online-Informationendienste
  - 3.3 Videotex-Dienste
  - 3.4 Audiotex-Dienste
  - 3.5 CD-ROM-Markt
4. Trends in anderen Märkten
5. Ausblick: Informationspolitische Erfordernisse

## **1. Einführung: Von der Informationswirtschaft zur intellektuellen Dienstleistungsgesellschaft**

Der Begriff "**Informationswirtschaft**" (Engl. information economy) wurde Mitte der 70er Jahre von den Amerikanern Edwin B. Parker und Marc U. Porat geprägt /1/, /2/. Er basiert auf der Beobachtung der starken Zunahme des Dienstleistungssektors in den hochindustrialisierten Ländern und dem damit eng verbundenen Vordringen elektronischer Informationstechnologien und Informationsdienstleistungen. Der industrialisierte Teil der Welt erlebt heute ein



Zusammenwachsen von Computer- und Telekommunikationstechnologien ("**Telematik**" oder "**Compunication**"), dessen gesellschaftliche Folgewirkungen denen der Ersetzung von Muskelkraft durch Maschinen gleichkommen.

Der Trend zur sogenannten "**Informationsgesellschaft**" wurde bereits in mehreren Beiträgen zu dieser Tagung angesprochen und diskutiert. Dieser Terminus spielt darauf an, daß "Information" zunehmend als der entscheidende "Produktionsfaktor" in Unternehmungen und Organisationen sowie als Schlüsselgröße für das ökonomische Wachstum und die gesellschaftliche Entwicklung der industrialisierten Länder angesehen wird. Der "**Informationssektor**" ist heute der am schnellsten wachsende Sektor entwickelter Volkswirtschaften. Er umfaßt alle Aktivitäten, die mit der Produktion, Beschaffung, Speicherung, Verteilung und Verwendung von Informationsgütern und -dienstleistungen zu tun haben. Seine Abgrenzung - soweit es überhaupt Versuche gibt, ihn statistisch präzise zu erfassen - wird unterschiedlich gehandhabt, aber er sollte unserer Meinung nach mindestens folgende Bereiche umfassen: Die Computer-Industrie (Hardware und Software), den gesamten Telekommunikationsbereich (sowohl Basisdienste als auch "Value-Added-Network-Services (VANS)") das Verlagswesen, die Bibliotheken, das Erziehungswesen, Rundfunk und Fernsehen, elektronische Datenbanken, E-Mail-Dienste, "Electronic Data Interchange" (EDI), "Electronic Funds Transfer at Points of Sale" (EFT-POS), Electronic Banking, Electronic Shopping u.ä. Das rasante Vordringen **elektronischer** Informationsgüter und -dienstleistungen besonders in Teilbereichen des Dienstleistungssektors (Banken, Versicherungen, Handel, Tourismus usw.) führt dazu, daß man diese Bereiche z.T. bereits ganz dem "Informationssektor" zurechnet.

Die **Informationswirtschaft** ist von Natur aus grenzüberschreitend. Der kanadische Kommunikationswissenschaftler Herbert Marshall Mc Luhan hat für die elektronisch vernetzte internationale Gemeinschaft den sehr anschaulichen Begriff "**globales Dorf**" geprägt. Die bereits heute weltweite Verbreitung des Fernsehens oder die globale Telekommunikation und VANS sind Beispiel für diese Entwicklung.

Die ökonomischen Wurzeln für diese Entwicklung sind im Zusammenwachsen von Computer-, Telekommunikations- und Informationsverarbeitungstechnologien zu sehen, was wiederum angeregt wird von einem ständig wachsenden

Dienstleistungssektor - und zugleich auf diesen zurückwirkt /3/. In den hochindustrialisierten Ländern arbeiten inzwischen bis nahezu drei Viertel aller Beschäftigten im Dienstleistungsbereich - gegenüber weniger als der Hälfte vor 50 Jahren. In den USA sind z.B. rund 85 % des gesamten Bestandes an Informationstechnologieprodukten im Dienstleistungssektor eingesetzt. Während die Industrie dort nur ca. 10 % des eingesetzten Kapitals in moderne Informationstechnologien investiert, haben sich die entsprechenden Aufwendungen des Dienstleistungssektors im letzten Jahrzehnt auf fast 20 % verdoppelt.

Der Dienstleistungssektor selbst läßt sich in zwei ganz unterschiedliche Bereiche aufteilen: persönliche Dienstleistungen und Informationsdienstleistungen. Während der Anteil der Beschäftigten, die persönliche Dienstleistungen erbringen (z.B. Hoteliers, Friseure, Schneider) in den USA seit der Jahrhundertwende bei etwa 15 bis 20 % aller Beschäftigten stagniert /2/, ist der Anteil der auf dem Informationssektor Tätigen (Bankangestellte, Versicherungsvertreter, Anwälte, Wissenschaftsjournalisten usw.) auf heute bereits ca. 55 % aller in den USA Beschäftigten gestiegen /3/. Dies unterstreicht den Trend zur **intellektuellen Dienstleistungsgesellschaft**, d.h. zu höherwertigen, individuellen und informationsintensiven Diensten.

Nach neueren Schätzungen trägt in den hochindustrialisierten Ländern der Informationssektor bereits 30 bis 40 % zum Bruttosozialprodukt bei /3/. Besonders hohe Wachstumsraten haben innerhalb dieses Sektors die Märkte für **elektronische** Informationsgüter und -dienstleistungen. Mit diesen Märkten wollen wir uns deshalb im folgenden besonders beschäftigen.

## 2. Initiativen der EG-Kommission zur Förderung des Informationsmarktes

Im Wettbewerb der drei großen Wirtschaftsräume "Nordamerika", "Ostasien" und "Europa" kommt der Entwicklung der Märkte für **elektronische** Informationsgüter und -dienstleistungen (im folgenden kurz: elektronische Informationsdienste) eine entscheidende Bedeutung zu. Dem trägt die Politik der Europäischen Gemeinschaften Rechnung, indem sie die Entwicklung eines **europäischen Marktes** für elektronische Informationsdienste fördert (vgl. im einzelnen dazu den Vortrag von Wolfgang Huber im Rahmen dieser Tagung): zunächst durch den

Aktionsplan IMPACT 1 für die Jahre 1989 und 1990 und dann durch den erweiterten Plan IMPACT 2, der im laufenden Jahr verabschiedet wurde und in den Jahren 1992 bis 1995 zur Ausführung gelangt. Zu berücksichtigen sind hier auch die im Jahre 1989 erlassenen "Leitlinien für ein besseres Zusammenwirken zwischen dem öffentlichen und privaten Sektor auf dem Informationsmarkt" /4/ sowie eine ganze Reihe von rechtlichen und regulativen Aspekten bei der Entwicklung des europäischen Informationsmarktes - vgl. hierzu im Einzelnen die Ausführungen im vorangegangenen Vortrag von Jürgen Goebel und Herbert Burkert.

Diese informationspolitischen Initiativen der EG-Kommission gehen vom Ansatz her weit über entsprechende Aktivitäten in einzelnen Mitgliedsländern hinaus. Zu wünschen wäre allerdings eine bisher nur wenig zu erkennende - aber nach dem bisher Gesagten unbedingt notwendige - engere Abstimmung zwischen der EG-weiten Förderung von elektronischen Informationsdiensten (in Luxemburg) und von elektronischen Informationstechnologien (in Brüssel).

Im Rahmen der obengenannten IMPACT-Programme wurde bereits im Frühjahr 1989 das europäische "**Informationsmarkt-Observatorium**" (IMO) eingerichtet. Ziel des IMO ist die Vernetzung aller Institutionen in den EG-Ländern, die sich mit der Analyse der Märkte für elektronische Informationsdienste beschäftigen. Es sollen alle verfügbaren Daten zusammengeführt, kritisch ausgewertet und die Ergebnisse über elektronische Datenbanken auf dem EG-Host in Luxemburg für alle IMO-Mitglieder (und später auch gegen Entgelt für Dritte) bereit gehalten werden. Darüber hinaus werden aus EG-Mitteln empirische Erhebungen mitfinanziert, die zum Ziel haben, einheitliche Definitionen und Erhebungsmethoden für diese Märkte EG-weit durchzusetzen. Das IMO gibt in Form von "Arbeitspapieren" in unregelmäßigen Abständen Übersichten über wichtige Marktentwicklungen heraus. Außerdem sind bisher ein Jahresbericht für 1988 /5/ und ein Zweijahresbericht für 1989-90 /6/ erschienen. Beide sind eine wichtige Basis für die späteren Ausführungen.

### **3. Entwicklung elektronischer Informationsdienste**

Das folgende Kapitel wird die Marktentwicklung in vier Hauptbereichen elektronischer Informationsdienste behandeln. Zunächst folgt eine begriffliche Erläuterung, um dann auf die einzelnen Märkte detaillierter einzugehen.

### 3.1 Begriffliche Abgrenzung

#### **Online-Informationendienste**

Hierzu gehören **Online-Datenbankdienste** (Nutzer haben die Möglichkeit, in verschiedenen elektronischen Datenbanken rückwirkend zu recherchieren. z.B. in bibliographischen Datenbanken, Volltext-Datenbanken, numerischen Datenbanken) und **Online-Real-Time-Dienste** (den Nutzern werden laufend sehr aktuelle Informationen online angeboten: z.B. Börsenkurse, Pressenachrichten, Unternehmensinformationen, Finanzinformationen, Wechselkurse). Zwischen diesen beiden Arten von Diensten besteht eine zunehmende Konvergenz, indem z.B. die Real-Time-Dienste durch "historische" Datenbankdienste unterstützt werden (etwa im Finanzbereich, wo neben der Kursentwicklung auch weitere Wirtschaftsinformationen bereitgestellt werden).

Die Online-Informationendienste werden häufig als "1. Generation elektronischer Informationendienste" bezeichnet, da ihre Verbreitung bereits Ende der 60er Jahre (zunächst vor allem in den USA) begann und mit der Entwicklung von modernen Datenübertragungsnetzen vor allem in den 70er und in der ersten Hälfte der 80er Jahre stark voranschritt.

#### **Videotex-Dienste**

Hierbei handelt es sich um Dienste, die neben Text auch einfache grafische Informationen in Farbe übermitteln können, was wiederum spezifische, in der Regel nur national gültige Standards bedingt. In Deutschland ist Bildschirmtext ein Videotex-Dienst. Als Kommunikationsband genügt eine Telefonverbindung, als Terminal können auch spezielle Fernsehgeräte eingesetzt werden. Entwickelt wurden diese Dienste in den 70er Jahren; erste Feldversuche begannen in Europa jedoch erst Anfang der 80er Jahre.

## - Audiotex-Dienste

Hierbei handelt es sich um Weiterentwicklungen der bekannten Telefonansagen (Zeitansage, Wettermeldungen, Sportergebnisse, Wechselkurse usw.), indem diese auch für den Dialogverkehr und damit auch für (einfache) Recherchen ausgebaut wurden. Erste Entwicklungen gab es bereits in den 70er Jahren, interaktive Audiotex-Dienste wurden jedoch in größerem Umfang erst in den 80er Jahren mit der Verbesserung der Telefontechnologie möglich.

## - Optical Disc-Dienste

Hier begann die Entwicklung Anfang der 80er Jahre mit der Videodisc-Technologie. Ein gewisser Durchbruch wurde in einigen Bereichen Ende der 80er Jahre mit der Entwicklung der **optischen Speichermedien** erzielt, insbesondere mit der sog. **CD-ROM (Compact Disc-Read Only Memory)**. Sie kann große Datenmengen relativ kostengünstig speichern und eignet sich für "Recherchen vor Ort". Außerdem wird durch diese Technologie der Weg für **Multimedia-Produkte** (Kombinationen von Text, Bild und Ton) eröffnet.

Diese vier Arten von Informationsdiensten können auch als verschiedene "Medien" oder "Kanäle" angesehen werden, die Zugang zu einer Vielfalt von Informationsinhalten, Präzisions- und Suchmöglichkeiten bieten. Wir wollen im folgenden versuchen, die Verschiedenartigkeit der "Märkte" kenntlich zu machen, die jedes dieser "Medien" bedienen kann: von Börsenmaklern über Banken, Unternehmen und Forschungsinstituten bis hin zum einzelnen Konsumenten. Es liegt auf der Hand, daß die heute in den EG-Ländern installierten insgesamt ca. 120 Millionen Telefone, ca. 6 Millionen Videotex-Endgeräte und ca. 3,2 Millionen Fax-Geräte bereits eine geeignete Infrastruktur für "**Massen-Märkte**" bei verschiedenen der genannten Informationsdienste bilden. Besonders die sprunghafte Zunahme der Fax-Geräte bietet Chancen z.B. für ein starkes Wachstum der Audiotex-Dienste (s. dazu weiter unten).

### 3.2 Online-Informationendienste

In allen EG-Ländern zusammen wurden 1989 etwa 1,8 Milliarden ECU (1 ECU = 2,08 DM in 1989) für Online-Informationendienste ausgegeben. Die weltweiten **Umsätze** mit Online-Informationendiensten lagen vergleichsweise dazu 1989 bei annähernd 8 Milliarden US-Dollar (davon allein über 4,5 Mrd. US-Dollar durch US-amerikanische Anbieter) /7/. Die Nutzung dieser Dienste in den EG-Ländern lag schwerpunktmäßig im professionellen Bereich (Banken, Börsen, Unternehmen, Wissenschaft, Bibliotheken, öffentliche Verwaltung), wenig oder kaum im Bereich privater Haushalte oder einzelner Bürger. Konsumenteninformationendienste werden eher durch Videotex und Audiotex abgedeckt. In den USA werden hingegen Online-Informationendienste auch in großem Umfang im Bereich der Unterhaltung genutzt; daneben setzen sich auch sog. "Transaktionsdienste" (Elektronic Banking, Electronic Shopping, elektronische Platzreservierung und Reisebuchungen usw.) mehr und mehr durch.

1989 nutzten lediglich ca. 6 % aller EG-Unternehmen und -Organisationen Online-Informationendienste (inklusive Zugang über Videotex-Terminals). Von diesen ca. 700 000 Institutionen hatten 54 % ihren Standort in Frankreich, 16 % in Großbritannien, 16 % in Deutschland, 10 % in Italien.

Die Gesamtausgaben für Online-Informationendienste in der EG (1989) verteilten sich nach Ländern allerdings folgendermaßen:

- 33 % auf Frankreich
- 28 % auf Großbritannien
- 11 % auf Deutschland
- 11 % auf Italien
- 17 % auf die anderen EG-Länder.

Die durchschnittlichen Online-Ausgaben pro Institution waren in Großbritannien dreimal höher als in Frankreich oder Deutschland. Dies erklärt sich durch die intensive Nutzung dieser Dienste durch den Finanzsektor in Großbritannien.

Im Gegensatz zu Videotex-Diensten, deren grenzüberschreitende Ausbreitung durch unterschiedliche Normen behindert wird, sind Online-Informationendienste

potentiell von jedem Punkt der Welt aus zugänglich, d.h. dies ist von Natur aus ein **internationaler Markt**.

Die erste in den EG-Mitgliedsländern flächendeckend durchgeführte Erhebung der Umsätze von Hosts (= Rechenzentren, die Online-Informationendienste für externe Nutzer im Fernzugriff anbieten), die ihren Hauptsitz in den EG-Ländern haben /8/, kam bei einer Befragung von ca. 350 solcher Hosts zu Gesamtumsätzen (weltweit) von annähernd 2,5 Milliarden ECU im Jahre 1988. Davon wurden ca. 56 % (ca. 1,4 Milliarden ECU) in der EG umgesetzt. Diese Zahlen beziehen sich auf sog. "professionelle" Dienste der Hosts (d.h. Dienste für die Unterhaltung und für Konsumenten wurden nicht in diese Untersuchung einbezogen). 57 % dieser Umsätze entfielen auf **Real-Time-Dienste** und ca. 15 % auf den Zugriff über Videotex-Dienste (vor allem in Frankreich), der Rest auf den unmittelbaren Zugriff auf retrospective Datenbank-Dienste. Mehr als 50 % der Gesamtumsätze entfielen auf nur zwei Anbieter (die beide in Großbritannien ihren Hauptsitz haben): Reuters und die EG-basierten Hosts des US-amerikanischen (und weltweit führenden) Anbieters von Firmen- und Kreditinformationen Dun & Bradstreet.

Natürlich steht damit Großbritannien auch im Ländervergleich weit an der Spitze (ca. Angaben in Millionen ECU 1988):

Großbritannien	2094
Frankreich	185
Italien	99
Niederlande	55
Deutschland	39
Dänemark	14
Spanien	4
Irland	2
Belgien/Luxemburg	1

(In Portugal und Griechenland konnten 1988 keine Umsätze ermittelt werden).

In einigen Ländern sind offenbar in gewissem Umfang auch interne Umsätze (sog. "geschlossenen Benutzergruppen") enthalten, was die Aussagefähigkeit dieser Zahlen mindert. Generell kann man aber zu den Ergebnissen dieser ersten gemeinsamen Erhebung sagen, daß noch nie so genaue EG-Zahlen vorlagen.

Eine zweite gemeinsame Erhebung (für 1989 und 1990), die auch Umsätze von Datenbasisproduzenten umfaßt, steht kurz vor dem Abschluß. Hier sind weiter verbesserte Ergebnisse über diesen "Markt" zu erwarten.

### 3.3 Videotex-Dienste

Im Gegensatz zu den Online-Informationendiensten handelt es sich hier aufgrund inkompatibler nationaler Standards um bisher relativ abgeschlossene nationale Märkte, die sich auch sehr unterschiedlich entwickelt haben. Im Jahre 1989 wurden in allen EG-Ländern zusammen etwa 700 bis 800 Millionen ECU für Videotex-Dienste ausgegeben. Hier sind neben Verbraucherinformationen, Unterhaltung sowie professionellen Informationsdiensten auch reine "Transaktionsdienste" (z.B. Geldüberweisungen, Hotelbuchungen) und Mischformen zwischen diesen Dienstearnten eingeschlossen. Auf **professionelle** Informationsdienste entfielen im Jahre 1988 in verschiedenen EG-Ländern folgende Umsätze (alle ca. Angaben in Millionen ECU) /8/:

Frankreich	168
Großbritannien	157
Niederlande	45
Italien	3
Deutschland	3
Irland	2

Der Videotex-Markt verzeichnet inzwischen höhere Wachstumsraten als der "Online-Markt der 1. Generation". Ursprünglich war Videotex für private Haushalte (besonders in Frankreich) konzipiert und sollte - durch Verbindung von vorhandenen Telefon- und Fernsehgeräten - einen relativ leichten Zugang zu (einfachen) Informations- (Telefonauskunft) und auch Transaktionsdiensten ermöglichen. France Telecom gelang es dann, durch die unentgeltliche Lieferung einfacher und billiger Terminals (den sog. Minitels) relativ schnell einen Massenmarkt für diesen Dienst zu schaffen (Der Hauptgrund für diese unentgeltliche Abgabe war die Vermeidung von Druckkosten für Telefonbücher). Hinzu kam dann die Einführung eines vereinfachten Abrechnungsverfahrens (das sog. "Kiosk"-Prinzip), das die Entwicklung dieses Marktes in Frankreich erleichterte. Die Umsätze im französischen Videotex-System (Teletel) werden heute noch durch allgemeine Konsumentendienste ("Adult Services", elektronisches Einkaufen,



Zahlungsverkehr, Reisebuchungen usw.) dominiert, doch verzeichnen bereits (einfach zu handhabende) professionelle Dienste (Finanzinformationen, Wirtschaftsnachrichten usw.) höhere Wachstumsraten. Für diese sog. "high-level"-Kiosk-Dienste wuchs die Inanspruchnahme zwischen 1989 und 1990 um 180 % (Zahl der Anrufe, Anschaltstunden). Die Nutzung dieser Dienste erfolgt fast ausschließlich durch Wirtschaftsunternehmen. Postverwaltungen in anderen EG-Ländern versuchen heute (bisher mit relativ wenig Erfolg), dieses durch hohe Anfangssubventionen, einfachste Technik und dezentrale, daher flexible Datenbereitstellung gekennzeichnete Marktentwicklungskonzept von France Telecom in Teilen zu kopieren.

In anderen EG-Mitgliedstaaten dominierte von vornherein die geschäftliche Nutzung der Videotex-Dienste. In Belgien z.B. sind 85 - 90 % aller Videotex-Terminals in Wirtschaftsunternehmen installiert, in den Niederlanden 75 %, in Portugal und Deutschland 60 % und in Großbritannien 55 %. In letzterem kommt man der französischen Entwicklung noch am nächsten. Besonders Verkehrs- und Börsendienste sind in Großbritannien über Videotex verbreitet.

Ende 1990 gab es in Frankreich ca. 5,6 Millionen Minitels (das waren 90 % aller in den EG-Ländern installierten Videotex-Terminals). In Deutschland waren es zu diesem Zeitpunkt 128 000, in Großbritannien 100 000, in Italien 145 000, in Spanien 60 000 und in den Niederlanden 25 000. In Frankreich wurden 1989 allerdings "nur" ca. 70 % der gesamten EG-Umsätze mit Videotex-Diensten erzielt (France Telecom allein erzielte 1989 Gesamteinnahmen von 520 Millionen ECU aus dem Videotex-System). Dies hängt mit der vornehmlich geschäftlichen (professionellen) Nutzung in den anderen Ländern zusammen.

Der Durchbruch ist jedoch nur auf den Massenmärkten zu erzielen, wie auch das Beispiel USA mit seinen äußerst erfolgreichen Konsumenteninformationsdiensten Prodigy und CompuServe (Ende 1990: 750 000 bzw. 700 000 Nutzer) zeigt. In Frankreich waren Ende 1990 bereits 19 % aller Haushalte und 80 % aller Unternehmen mit einem Minitel ausgerüstet.

Allerdings sind die öffentlichen Anbieter von Telekommunikations-Diensten angesichts der Liberalisierungs- und Privatisierungstendenzen mit abnehmenden Gewinnaussichten konfrontiert. Die Verzinsung des in Videotex-Dienste eingesetzten Kapitals war niedriger (in der Regel negativ) als vorher angenommen

und führte sowohl in Frankreich als auch in Deutschland zu öffentlichen Debatten. Außerdem war öffentliches Kapital für andere Dienste bzw. Technologien (Kabelfernsehen, ISDN, Mobilfunk) erforderlich. Diese zwei Gründe (verminderte Gewinnaussichten, Kapitalknappheit) führten in verschiedenen EG-Ländern (Niederlande, Belgien, Irland, Spanien, Italien) zu Reprivatisierungen oder zu "joint ventures" mit Banken, Industriefirmen und lokalen öffentlichen Verwaltungen.

Ein zweiter Trend ist die Entwicklung von "Kiosk-Systemen" (seit 1986 in Frankreich) ab 1991 auch in anderen EG-Ländern: Italien, Belgien, Luxemburg, Portugal, Irland. Man verspricht sich hierdurch einen leichteren Zugang für kleine und mittlere Unternehmen und Einzelnutzer und eine Stimulation in der Nutzung sowohl von Transaktions- als auch von Informationsdiensten.

Ein dritter Trend auf diesem Markt ist die Vernetzung der bisher getrennten nationalen Videotex-Systeme, welche jedoch nur sehr schleppend vorankommt (bisher sind nur bilaterale Übereinkünfte in Sicht).

### **Audiotex-Dienste**

Dies ist sicher ein Markt mit großen Entwicklungsmöglichkeiten, der aber fast überall noch in den Anfängen steckt. Ähnlich wie Videotex ist Audiotex relativ einfach zu nutzen. Es erfordert keine Extra-Anschaffung von teuren Geräten. Die Grundlage bilden die ca. 120 Millionen Telefonanschlüsse in den EG-Ländern. Dies bietet Möglichkeiten für einen Massenmarkt, der sowohl Einzelnutzer (private Haushalte) als in gewissem Umfang auch die Geschäftswelt einschließt. Audiotex-Dienste reichen von der traditionellen Form von auf Band gesprochenen und per Telefon abrufbaren Mitteilungen bis zur verbalen Suche (voice recognition) in relativ komplexen elektronischen Datenbasen. Letzteres ist noch im Versuchsstadium, wird jedoch mit der Weiterentwicklung der Telefontechnologie und den Übertragungsleitungen technisch möglich.

Audiotex-Dienste können **passiv** (der Nutzer wählt eine Telefonnummer und hört eine vorher gespeicherte Mitteilung ab) oder **interaktiv** sein (der Nutzer wählt nach bestimmten Kriterien Informationen aus einem Speicher aus). Diese Interaktion kann auf einfache Auswahlvorgänge beschränkt sein (sog. "Menüs") oder eine komplexere Auswahl durch Spracheingabe ermöglichen; die erste Möglichkeit wird bereits durch die herkömmliche Telefontastatur geboten, während

für die zweite Alternative eine höher entwickelte Fernsprechtechnologie erforderlich wird.

Interessante Kombinationsmöglichkeiten dieser Dienste bieten sich an mit dem inzwischen sehr verbreiteten Telefax-Dienst und ebenso mit dem Videotex-Dienst. Beim **Audio-Fax** recherchiert der Nutzer per Audiotex und erhält das Ergebnis (Dokumente, Abbildungen, Tabellen usw.) per Fax. Analog beim **Audio-Videotex**. Auch der Audiotex-Dienst ist von Natur aus grenzüberschreitend. Man kann sich heute schon ein großes Potential für paneuropäische Audiotex-Dienste vorstellen: z.B. im Bereich von Unternehmensinformationen, Aktienkursen, Wechselkursen, Reisebuchungen. In fünf EG-Ländern (Frankreich, Großbritannien, Irland, Niederlande, Belgien) gibt es bereits "Kiosk-Einrichtungen" für anspruchsvolle Audiotex-Dienste: Der Telefonnetzbetreiber zieht die (höheren) Gebühren, die bei der Anwahl spezieller Nummern entstehen, ein; er behält den Teil, der die reinen Telefonkosten deckt und überweist den Rest an die Anbieter der Audiotex-Dienste.

In der EG hat Großbritannien den am weitesten entwickelten Audiotex-Markt mit dem größten Spektrum verschiedenartiger Dienste. 83 % an diesem Markt hält British Telecom, den Rest zwei hinzugekommene private Anbieter von Telefonnetzen: Racal (15 %) und Mercury Communications (2 %).

Der Markt für interaktive **professionelle** Audiotex-Dienste in der EG (d.h. unter Ausschluß des Bereichs Unterhaltung und "voice messaging") hatte 1989 ein Volumen von etwa 300 Millionen ECU.

1990 hatten nach dem Kiosk-Prinzip angebotene Dienste ("premium rate services") ein Volumen von 499 Millionen ECU in nur drei Ländern:

Großbritannien	48 % davon,
Frankreich	32 % davon,
Niederlande	20 % davon.

Etwa 50 % der in diesen Ländern hiermit erzielten Einnahmen (vor allem: Unterhaltung, Kontaktanzeigen, Wettermeldungen, Verkehrsmeldungen, Sportnachrichten usw.) wurden von den Telefonnetzbetreibern an die Anbieter der Audiotex-Dienste weitergegeben. Die Umsätze in den USA lagen im gleichen

Zeitraum mit ähnlichen Diensten nicht wesentlich höher. Jedoch nehmen in Großbritannien jetzt die Umsätze mit höherwertigen sog. "seriösen Diensten" über Audiotex signifikant zu: z.B. Finanzinformationen, Bankdienstleistungen, Bilanzinformationen, Bestellungen. Erschwert wird die weitere Entwicklung der Audiotex-Dienste durch komplexe rechtliche und regulatorische Probleme (Schutz der Privatsphäre, vertragsrechtliche Fragen, grenzüberschreitender Datenschutz) und zum Teil durch technische Barrieren, die die grenzüberschreitende Nutzung von Audiotex-Diensten behindern (unterschiedliche Telefongebührensyste me usw.).

Abgesehen von diesen Barrieren bestehen große Chancen für die Entwicklung eines **internationalen Audiotex-Massenmarktes** vor allem in Bereichen wie: Finanzen, Tourismus, Verkehr, Spedition, Sport, Telemarketing.

### **Der Europäische CD-ROM-Markt**

Im Gegensatz zu Videotex und zum Teil auch zu Audiotex war die CD-ROM-Produktentwicklung (das gilt auch für andere optische Medien) in erster Linie auf **professionelle** Nutzer ausgerichtet. Die Orientierung auf die **Massenmärkte** für Konsumenten erfolgt erst jetzt (besonders in Japan im Bereich der Unterhaltung, Weiterbildung, der Spiele usw.). Die von der Technologie gebotenen Möglichkeiten (Verbindung von Bild, Text und Ton; Hypertext-Software usw.) müssen erst noch durch geeignete Produktentwicklungen von Seiten der Anbieter (besonders der Verlage) ausgeschöpft werden. Diese Entwicklung ist in den USA und Japan schon weiter vorangeschritten als in den EG-Ländern. In der EG sind beträchtliche Zunahmen in der Anzahl der veröffentlichten "CD-ROM-Titel" zu verzeichnen, aber vergleichsweise geringe Umsatzsteigerungen.

Die Anzahl der weltweit auf dem Markt angebotenen kommerziellen **CD-ROM-Titel** stieg von 52 im Jahre 1986 auf 1430 im Jahre 1990 (ohne die nur "inhouse" genutzten Titel), mit extrem hohen Wachstumsraten von über 100 % pro Jahr.

Die USA war führend mit 440 Titeln in 1989, gefolgt von Japan mit 161 Titeln und der EG mit 118. Die in der EG angebotenen Titel verteilten sich auf einzelne Länder wie folgt /6/:

Italien	28
Großbritannien	28
Deutschland	21
Frankreich	17
Niederlande	9
Spanien	5
Belgien	3
Irland	1
Griechenland	1
EG-weit	5

Bereits 1990 schob sich Deutschland mit 93 Titeln an die Spitze, gefolgt von Großbritannien (90), Frankreich (81) und Italien (72).

Die weltweiten **CD-ROM-Umsätze** (einschließlich der Umsätze mit Abspielgeräten, kommerziellen Titeln und dem monetären Wert der inhouse-Titel) betragen 1989 ca. 773 Millionen ECU, davon entfielen 437 Millionen ECU auf kommerzielle Titel, die sich wiederum folgendermaßen aufteilten /10/:

289	in den USA	(66%)
90	in Japan	(20 %)
58	in der EG	(13 %)

Die 58 Millionen ECU in der EG verteilten sich auf folgende Länder:

41	in Italien	(71 %)
7	in Großbritannien	(12 %)
3,5	in Deutschland	(6 %)
3	in Frankreich	(5 %)
3,5	auf alle anderen EG-Länder	(6 %)

Die Anbieter von CD-ROM-Titeln in der EG konzentrieren sich auf andere Bereiche (Recht, Wirtschaft) als diejenigen in den USA (Geographie, Medizin, Software). 19 % der in der EG 1989 angebotenen Titel z.B. können dem Bereich "Recht" zugerechnet werden, 16 % dem Bereich "Wirtschaft", 12 % sind irgendwelche Nachschlagewerke, Verzeichnisse, Handbücher und Produkt-Kataloge

("directories") und 11 % sind dem Bereich "Bibliotheksautomatisierung" zuzurechnen. 20 % der in den USA angebotenen Titel sind dem Bereich "Geographie", 12 % dem Bereich "Software", 11 % dem Bereich "Medizin" und 9 % dem Bereich "directories" zuzurechnen.

In der Vergangenheit hat man CD-ROM überwiegend als ein alternatives Medium für die Verbreitung bestimmter Datenbasen angesehen, die auch online angeboten werden (für häufige Benutzung; bei Inhalten, die relativ selten aktualisiert werden müssen usw.). Heute zeigt sich, daß CD-ROM's zunehmend für die Verbreitung von Informationsinhalten verwendet werden, die wenig (oder kaum) Bedeutung für den traditionellen Online-Markt haben: z.B. **Multimedia-Anwendungen** in Bereichen wie Erziehung, Ausbildung und Unterhaltung (vgl. die Entwicklung in Japan). In der Tat findet der prognostizierte Transfer von "online" zu "CD-ROM" kaum statt: Nur 4 % der weltweit online angebotenen Datenbasen sind auch auf CD-ROM verfügbar. Weniger als 30 % der heute angebotenen kommerziellen CD-ROM-Titel sind aus online-Datenbasen abgeleitet (z.B. durch Entnahme von Teilinhalten oder Inhaltskombinationen oder -synthesen) /6/.

#### **4. Trends in anderen Märkten**

Neben diesen Entwicklungen in ausgewählten Informationsdienstleistungsmärkten lassen sich einige übergreifende, hier subjektiv ausgewählte Trends festmachen, auf die im folgenden kurz eingegangen wird:

- Konzentrationstendenzen und Aktivitäten der Medienkonzerne
- Electronic Data Interchange (EDI)
- Auswirkungen des Gemeinsamen Marktes 1993
- Telematikinfrastruktur.

##### **4.1 Konzentrationstendenzen und Aktivitäten der Medienkonzerne**

Die zunehmende ökonomische Bedeutung von Informationsdienstleistungen - und damit auch der wachsende Wettbewerbsdruck - lassen sich an dem weltweit seit ein paar Jahren zu beobachtenden Konzentrationstrend ablesen; er dokumentiert

gleichzeitig das große Potential, das elektronischen Informationsdiensten in den Führungsetagen großer Medienkonzerne zugerechnet wird. So drängen große Firmen, die bisher eher traditionelle Informationsdienste anboten (z.B. die britische Nachrichtenagentur Reuters; die weltweit führende Anbieter von Kreditinformationsdiensten Dun & Bradstreet und insbesondere große Medienkonzerne und Verlagshäuser wie Bertelsmann, McGraw Hill, Maxwell, Reed, Knight Ridder, Pearson, Elsevier oder Dow Jones) - meist durch Aufkäufe von Firmen - in diese Märkte ein.

Daneben bildeten sich "Joint ventures": z.B. zwischen den Wirtschaftszeitungsverlagen Dow Jones (USA) und Holtzbrink (Deutschland) bei der Lieferung von Finanzinformationsdiensten (auf einer 50/50-Basis); zwischen den Verlagen Bonnier (Schweden) und Hoppenstedt (Deutschland) bei der Lieferung von Firmeninformationen (ebenfalls auf einer 50/50-Basis); zwischen der Maxwell Communication Corporation (Großbritannien) und VINITI (Sowjetunion) bei der Lieferung von wissenschaftlichen und technischen Informationen; zwischen der Royal Society of Chemistry (Großbritannien) und der Dechema (Deutschland) beim Angebot der weltweit größten Datenbasis im chemischen Apparatewesen und der industriellen Biotechnologie.

Zu nennen wäre hier auch die geplante Zusammenarbeit zwischen Amadeus (Lufthansa, Air France, Iberia) und Sabre (American Airlines) beim weiteren Aufbau ihrer elektronischen Flugreservierungssysteme. Die Investitionskosten (bei Start/Amadeus bisher ca. 2 Mrd. DM) und das damit verbundene Risiko werden auch für große Konzerne unüberschaubar.

Die traditionellen Verlage in den EG-Ländern scheinen hingegen nicht in dem Maße an einem online-Angebot ihrer Volltextdatenbasen interessiert zu sein, wie ihre amerikanischen Kollegen. Allerdings erkunden die europäischen Verlage zunehmend ihre Möglichkeiten auf den Videotex- und Audiotex-Märkten.

So sind z.B. in den USA mehrere Hundert der großen Zeitungen heute im Volltext bei verschiedenen Hosts aufgelegt, in der EG dagegen nur wenige. Ein Grund dafür mag der hohe Investitionsaufwand für das online-Angebot sehr umfangreicher Volltextdatenbanken sein, der bei den national und sprachlich aufgesplitterten Märkten in Europa kaum ausreichende Kapitalerträge erlaubt. Stattdessen nehmen Verlage in der EG an den Videotex-Versuchen teil, indem z.B.

Tageszeitungen (oder Teile davon) per Bildschirm publiziert werden, aber unter Berücksichtigung der technischen Einschränkungen dieses Mediums. Die "elektronische Zeitung" hat bisher noch kein großes Publikumsinteresse gefunden. Die einzig nennenswerte Ausnahme kommt auch hier aus Frankreich, wo Tageszeitungen und Magazine von der Dynamik des Teletel-Systems profitieren konnten. Neuerdings konzentrieren sich Presseverlage eher auf Videotex-Angebote in den Bereichen Spiele, Kontaktanzeigen, Konsumenteninformationen. Daneben bieten anspruchsvollere Zeitungen (wie Le Monde) aber auch höherwertige Dienste über Videotex wie Expertensysteme oder Portfolio Management für Börseninformationen.

Das Interesse europäischer Presseverlage an fax-basierten Informationsdiensten ist bisher auf Newsletters beschränkt. Von den 24 führenden europäischen Presse- und Medienkonzernen haben bisher 10 signifikante Host-Aktivitäten entwickelt oder erworben, 13 produzieren mindestens eine Online-Datenbasis, 14 produzieren mindestens einen CD-ROM-Titel und 8 sind als "Service-Büro" im CD-ROM-Publishing involviert.

## **4.2 Electronic Data Interchange (EDI)**

Von nicht zu unterschätzender Bedeutung sind die Zukunftsaussichten des Electronic Data Interchange (EDI). Nach neueren Vorhersagen soll sich der Markt für EDI-Produkte und -Dienstleistungen in Europa von 86 Millionen US-Dollar in 1990 auf 400 Millionen US-Dollar in 1991 steigern /13/. Die Leitvorstellung bei EDI ist eine (im Prinzip weltweite) branchenübergreifende, Wirtschaft und öffentliche Verwaltung einschließende standardisierte Geschäftskommunikation zwischen Computern (bzw. deren Programmen) /9/. Bestimmte Nachrichtentypen wie z.B. Bestellungen oder Rechnungen werden von den Anwenderprogrammen des Kunden statt auf Papier als Datei ausgegeben, auf elektronischem Weg übertragen und durch das Anwenderprogramm des Lieferanten maschinell interpretiert. Die meisten der in Deutschland täglich produzierten Bestellungen werden zwar heute schon häufig durch ein Computersystem erzeugt, aber sodann auf Papier gedruckt und verschickt oder mündlich per Telefon übermittelt /6/. Die Rationalisierungsmöglichkeiten dieser "elektronischen Geschäftskommunikation" in sehr vielen Bereichen der Wirtschaft und der öffentlichen Verwaltung liegen auf der Hand. EDI-Anwendungen gibt es heute bereits als Reise-Buchungssysteme (Hotel, Autovermietung, Flug), integrierte Hafenmanagement-Systeme (die Spediteure,



Makler, Reeder, Banken, Versicherungen, Zoll- und Hafenamts einbinden) und im Banken- und Handelsbereich ("electronic funds transfer"). Voraussetzung für die Automatisierung unternehmensübergreifender Geschäftstransaktionen sind standardisierte Datenaustauschformate. Erst diese gewährleisten offene Kommunikation zwischen Programmen, die in unterschiedlichen Unternehmen oder Verwaltungen auf unterschiedlicher Hardware laufen /6/. Ein solcher Standard ist die UN-Norm EDIFACT (Electronic Data Interchange for Administration, Commerce and Transport), deren EG-weite Anwendung von der EG-Kommission gefördert wird. Die weitere Ausbreitung von EDI wird sich auch auf die Entwicklung der anderen Märkte für elektronische Informationsgüter und -dienstleistungen auswirken: An immer mehr Arbeitsplätzen werden PC's stehen, die - auch über EDI-Netze - den Zugang zu VANS, E-Mail, Bulletin Boards usw. ermöglichen und damit ein ganz neues Marktsegment erschließen.

### 4.3 Auswirkungen des gemeinsamen Marktes 1993

Wie bereits eingangs angedeutet, können von der weiteren Entwicklung der Märkte für elektronische Informationsdienste entscheidende Impulse für den angestrebten europäischen Binnenmarkt (für alle Arten von Gütern und Dienstleistungen) ausgehen. Dazu müssen diese elektronischen Dienste jedoch auch von den europäischen Firmen genutzt werden. 69 % der gesamten Aufwendungen von Firmen in der EG für elektronische Informationsdienste kamen 1988 aus dem Finanzbereich (Banken, Versicherungen), 9 % aus dem Handel, 9 % aus dem Tourismus und nur 4 % aus dem industriellen Sektor /11/. Man kann sagen, daß im Dienstleistungssektor der EG die Inanspruchnahme von elektronischen Informationsdiensten heute schon in etwa auf dem Niveau der USA liegt - im industriellen Sektor dagegen noch lange nicht!

Um voll die Vorteile des europäischen Binnenmarktes nutzen zu können, müssen die EG-Firmen über neuartige Dienste leichten Zugang zu mehr Informationen über andere nationale Märkte in der EG erlangen. D.h. es sind **EG-weite** Marktdaten-, Kreditinformations-, Produktinformations- und Herstellerinformationsdienste, Listen potentieller Kunden (mailing lists) usw. erforderlich. Um solche Dienste erstellen zu können, sind schwierige Harmonisierungen der heute noch unterschiedlichen Abgrenzungen, Nomenklaturen und Klassifikationen in den verschiedenen EG-Ländern notwendig. Da viele dieser EG-weiten Dienste auf Basisdaten zurückgreifen müssen, die von Regierungen und Behörden (z.B. statistischen

Ämtern) produziert werden, müßten auch diese Datenbasen harmonisiert werden /4/. Auf die hiermit zusammenhängenden vielfältigen rechtlichen und regulatorischen Aspekte wurde schon hingewiesen. Auch eine Verbesserung der schon diskutierten Volltext-Presse-Datenbanken könnte zur Befriedigung der EG-weiten Informationsbedürfnisse beitragen. Hier wird vermutlich der zu beobachtende Trend der Konzentration und Kooperation in der europäischen Wirtschaftspresse die Bedingungen für die Entwicklung europaweiter Wirtschaftsnachrichtendienste schaffen.

#### **4.4 Telematik-Infrastruktur**

Auch im technologischen Bereich scheint sich vieles in Richtung einer **EG-weiten Telematik-Infrastruktur** zu entwickeln:

- Videotex und Audiotex können in hybriden Anwendungen kombiniert werden: z.B. Audio-Fax, Audio-Videotex, Video-Fax; Konvergenz zwischen Videotex- und Online-Diensten der 1. Generation; Vernetzung nationaler Videotex-Systeme. All dies wirft komplizierte Fragen der Rechnungsstellung und Preisstrukturen auf (Pan-Europäisches Kiosk-System).
  
- Der Fax-Dienst übernimmt zunehmend eine Schlüsselstellung. 1990 waren bereits weltweit 10 Millionen Faxgeräte installiert (1985 erst 2 Millionen), davon in der EG 3,2 Millionen (513 000 in Deutschland, 432 000 in Frankreich, 944 000 in Großbritannien, 773 000 in Italien, 234 000 in Spanien). In den USA dringt Fax bereits in die privaten Haushalte vor (1990 hatten 1 % aller amerikanischen Haushalte einen Fax-Anschluß). Die Entwicklung sog. "fax boards", die jeden PC in ein Fax-Gerät verwandeln können, und die Verbindung von "fax hosts" mit Local Area Networks (LANs) in Organisationen wird diese Entwicklung weiter vorantreiben.
  
- Mobile Telefondienste werden die EG-weite Ausdehnung von elektronischen Informationsdiensten (besonders von Audiotex) stimulieren.

- Europaweite Normen für Multimedia-Dienste und Breitbandnetze werden die EG-weite Ausdehnung der optischen Speichermedien fördern.

## 5. Ausblick: Informationspolitische Erfordernisse

All diese skizzierten Entwicklungen und Trends zeigen eins sehr deutlich: Auch der Europäische Markt für elektronische Informationsdienste bewegt sich in Richtung Integration, Globalisierung, Vernetzung, gleichzeitig aber auch Differenzierung und Individualisierung.

Will europäische Informationspolitik zukunftsweisende "policies", eine mittelfristig tragfähige Programmatik entwickeln, so wird sie sich diesen Entwicklungen nicht nur stellen müssen, vielmehr ist es unabdingbar, der Vermaschung von Technologiepolitik, Medienpolitik, Telekommunikations- und Informationspolitik (letztere bisher eher im konventionellen IuD- oder Fachinformations-Sinne verstanden) durch **neue konzeptionelle Ansätze** Rechnung zu tragen.

Stand bisher bei der öffentlichen Informationspolitik eher eine statisch-funktionale, institutionen- und inhaltsbezogene Betrachtungsweise im Vordergrund, so wird in Zukunft immer mehr ein **prozeßorientierter, informationsflußbezogener, integrierender Ansatz** im Vordergrund stehen müssen.

Die Informationspolitik wird, gerade auch auf europäisch-multinationaler Ebene, sich auf diese außerordentlich dynamischen Marktentwicklungen einstellen und einen sog. "Paradigma-Wechsel" in Richtung "**Informationsflußansatz**" vornehmen müssen, d.h. auf eine integrierende Unterstützung und Optimierung von Informationsverarbeitungsprozessen setzen müssen, will sie nicht jedwede relevante Steuerungsfähigkeit verlieren. Das Setzen flexibel wirksamer Anreize, die Bereitstellung längerfristig verlässlicher Rahmenbedingungen, eine moderate, insbesondere auch soziale, ökologische und kulturelle Aspekte berücksichtigende Regulations- und Interventionspolitik zur Förderung des europäischen Informationsmarktes sollte hier unseres Erachtens im Vordergrund stehen, nicht das projekt- und einzelfallbezogene Setzen u.U. völlig falscher, den Marktentwicklungen nicht Rechnung tragender Signale.

## Literatur

1. Parker, E.B.: Social implications of computer / telecoms systems, in: Telecommunications policy, December 1976, p. 3 - 20.
2. Porat, M.U.: The information economy: definition and measurement, vol. 1., Washington, D.C.: US Department of Commerce 1977
3. Wright, K.: The road to the global village, in: Scientific American, March 1990, p. 57 - 66.
4. Kommission der Europäischen Gemeinschaften. Generaldirektion Telekommunikation, Informationsindustrie und Innovation: Leitlinien für ein besseres Zusammenwirken zwischen dem öffentlichen und dem privaten Sektor auf dem Informationsmarkt. Luxemburg 1989
5. Kommission der Europäischen Gemeinschaften. Generaldirektion Telekommunikation, Informationsindustrie und Innovation: Report on main events and developments on the information services market (annex III to SEC (90) 1778 final) Luxemburg, September 1990.
6. Kommission der Europäischen Gemeinschaften. Generaldirektion Telekommunikation, Informationsindustrie und Innovation: Report on main events and developments on the electronic informations services markets in 1989 - 1990 (annex III to) Luxemburg, September 1991.
7. SCIENTIFIC CONSULTING. Der internationale Markt für elektronische Informationsdienste / Datenbanken 1986 - 1990. Untersuchung des Weltmarktes auf der Anbieterseite und Analyse der Online-Nutzung in der Bundesrepublik Deutschland. Bd. I u.Bd. II (Abbildungen). Köln, Februar 1991.
8. European Information Industry Association: CEC/EIIA Coordinated Survey of the EC Host Organisations engaged in the Provision of Electronic Information Services for Professional Users 1988. Project Report. Luxemburg, December 1990
9. Klein, S.: Der Stellenwert von EDI in der Informationslogistik; in: Stroetmann, Karl A. (Hrsg.): Informationslogistik - Managementkonzepte, Fallbeispiele und Anwendererfahrungen bei Informationsprozessen. Proceedings der 6. Internationalen Fachkonferenz der Deutschen Gesellschaft für Dokumentation (DGD) in Zusammenarbeit mit der Gesellschaft für Informatik (GI) vom 27. - 29. Mai 1991 in Garmisch-Partenkirchen, Frankfurt/Main: DGD 1992 (in Vorbereitung).
10. IMO working paper 91/2 "Overview of the CD-ROM market", Luxemburg, June 1991

11. SCICON: The market for value-added services in Europe, vol I (Main Report), December 1989.
12. Schwuchow, W.: The Situation of the Online Information Services Industry in the European Community (with special consideration of the FRG), in: International Forum on Information and Documentation, vol. 16, no. 1, January 1991, p. 6 - 9.
13. Diebold Management Report 1/91, S. 21.

Dipl.-Volkswirt Werner Schwuchow  
Dr. Karl A. Stroetmann

Gesellschaft für Mathematik und  
Datenverarbeitung mbH (GMD)  
Forschungsstelle für Informationswirtschaft  
Schönhauser Str. 64

5000 Köln 51

Tel.: +221/3 76 73 43 und +221/3 76 73 14

Fax.: +221/3 76 73 83

E-Mail: [kas@svinfow.koeln.gmd.de](mailto:kas@svinfow.koeln.gmd.de)

## **Graphische Benutzeroberflächen: Methoden und Werkzeuge für Prototyping und Entwicklung**

Barbara Jochum

Das Thema wurde im Rahmen einer Masterarbeit behandelt, die im Dezember 1990 an der Fachrichtung Informationswissenschaft der Universität des Saarlandes eingereicht wurde. Die Anregung, dieses Thema im Rahmen einer Masterarbeit aufzugreifen, erhielt ich von der Unternehmensberatung Collogia GmbH (Köln), die sich vornehmlich auf den Bereich der Software-Erstellung und -Portierung spezialisiert hat und durch deren Unterstützung eine praxisnahe Arbeit ermöglicht wurde. Der Arbeit angegliedert ist ein Fallbeispiel, das den Entwurf einer Graphischen Benutzeroberfläche mit einem modernen Entwicklungswerkzeug erläutert.

Graphische Benutzeroberflächen erlauben das direkte Agieren mit und das direkte Verändern von Objekten am Bildschirm mit unmittelbarer Ergebnisrückmeldung. Das Interaktionsprinzip, dem diese Benutzeroberflächenform folgt, wird deshalb auch als Direkte Manipulation bezeichnet.

Obwohl man angesichts der unterschiedlichen Ausprägungsformen Graphischer Benutzeroberflächen nicht von dem graphischen Interaktionsstil sprechen kann, hat sich diese Technik der Oberflächengestaltung in den letzten Jahren zu einem "Quasistandard" entwickelt. Der Erfolg dieser Dialogform ist vor allem unter dem Gesichtspunkt der Bedienfreundlichkeit zu sehen. Systeme mit Graphischen Benutzeroberflächen ermöglichen vor allem ungeübten und gelegentlichen Benutzern den leichten, spielerischen Einstieg und sind relativ leicht erlernbar. Der Umgang mit der speziellen Syntax des Computers entfällt weitgehend, und das Handeln mit Repräsentationsformen real existierender Objekte kommt den angeborenen Fähigkeiten des Menschen entgegen.

Die zunehmende Verbreitung von Benutzeroberflächen mit Direkter Manipulation verstärkt die Forderung nach geeigneten Werkzeugen und Methoden für die Entwicklung.

Bei der Entwicklung Graphischer Benutzeroberflächen ergeben sich eine Reihe von Problemen. Zum einen ist der Markt gekennzeichnet durch eine starke Heterogenität bezüglich der zur Verfügung stehenden Endgeräte, Konfigurationen, Hard- und Softwarestandards und der Werkzeuge für ihre Erstellung. Zum anderen ist auch der Entwurfs- bzw. Entwicklungsprozeß an sich nicht unproblematisch. Graphische Benutzeroberflächen bieten dem Benutzer simultan mehrere alternative Vorgehensweisen an und erfordern ein schnelles unmittelbares Feedback. Da es nur sehr allgemein gehaltene

Entwurfsleitlinien gibt, ist die Prototypenerstellung in Zusammenarbeit mit dem späteren Benutzer die oftmals einzige Möglichkeit qualitativ gute Produkte zu erstellen. Für die Entwicklung dieser Benutzeroberflächen werden also spezielle Werkzeuge benötigt, die dazu dienen, die Komplexität und den Aufwand bei der Entwicklung besser beherrschen zu können.

Ziel des Vortrages ist es, Werkzeuge und Methoden zur Entwicklung Graphischer Benutzeroberflächen aufzuzeigen sowie ihre Einsatzmöglichkeiten zu diskutieren. Der Begriff Werkzeug bezeichnet in diesem Zusammenhang Computer-Softwareprodukte, die die Erstellung von Benutzeroberflächen erleichtern helfen. Unter Methoden werden in erster Linie Gestaltungsmethoden verstanden, d.h. systematische Vorgehensweisen, die von den Gestaltern in den Phasen des Entwicklungsprozesses eingesetzt werden können.

Der erste Teil des Vortrages beschäftigt sich vornehmlich mit den Grundlagen für die Gestaltung der Mensch-Computer-Schnittstelle und mit der Konzeption Graphischer Benutzeroberflächen. Neben einem Überblick über bestehende Normen und softwareergonomische Kriterien für die benutzergerechte Oberflächengestaltung wird u.a. der Frage nachgegangen, inwieweit die graphische Gestaltung von Benutzeroberflächen aus der Sicht der Kognitionspsychologie die menschliche Informationsverarbeitung unterstützt. In Form eines Exkurses werden Entwicklungstendenzen und Forschungsansätze im Bereich der Mensch-Maschine-Schnittstelle diskutiert. Es wird auf die noch bestehenden Probleme und Einschränkungen dieser Technologien eingegangen sowie mögliche Einsatzbereiche besprochen.

Im zweiten Teil des Vortrages stehen die Werkzeuge für die Entwicklung Graphischer Benutzeroberflächen sowie der Entwurfsprozeß der Benutzeroberfläche selbst im Vordergrund.

Die Dialogform Direkte Manipulation stellt neue und höhere Anforderungen an die Applikations-Entwicklung und -Architektur in vielerlei Hinsicht. Neben der, unter Software-Engineering Gesichtspunkten seit längerer Zeit erkannten, Notwendigkeit die Benutzerschnittstelle von der Applikation zu trennen, ergeben sich auch veränderte Anforderungen an die Programmierung. Die Entwicklung direkt manipulierbarer Benutzeroberflächen revidiert alle Grundsätze der bisherigen Programmierung. Der Endbenutzer, nicht die Applikation ist der Auslöser sämtlicher Aktionen, d.h. der Kontrollfluß wird von der Applikation weggenommen. In Zukunft werden die Applikationen als eine Sammlung von Unterrouتين zu gestalten sein, die von der Benutzeroberfläche bei Bedarf aufgerufen werden (externe Kontrollstruktur). Dieser Tatsache müssen auch Entwicklungswerkzeuge Rechnung tragen.

Der Markt für Entwicklungs-Werkzeuge ist indes gekennzeichnet durch eine starke Heterogenität, auch in bezug auf hersteller- und anwenderspezifische Systemmodelle. Dennoch haben sich in letzter Zeit einige unübersehbare Trends herausgebildet, die zum Teil bereits zu (zumindest Industrie-) Standards geführt haben.

Hinsichtlich der Entwicklungs-Werkzeuge werden grundsätzlich zwei Gruppen

unterschieden, die bezüglich Konzeption, Leistungsumfang und Systemarchitektur näher erläutert werden: Toolkits/ Fenstersysteme und User Interface Management Systeme (UIMS). Toolkits bauen i.d.R. auf Fenstersystemen auf. Da sich herkömmliche Fenstersysteme jedoch die Fensterverwaltung mit der Anwendung teilen (das Fenstersystem übernimmt die Verwaltung des Ein-/Ausgabebereiches), ergeben sich eine Reihe von Problemen, besonders unter dem Gesichtspunkt der Isolierung der Benutzeroberfläche.

Grundgedanke von UIMS ist die Trennung der Applikation in einen interaktionsspezifischen und einen rein funktionalen Teil. Sie verfügen im Gegensatz zu Toolkits über eine eigenständige aktive Komponente, die das gesamte dynamische Verhalten der Oberfläche kontrolliert. Dieser Grundaufbau, d.h. die Abtrennung der Benutzeroberfläche vom rein funktionalen Teil, macht es möglich, die Oberfläche gesondert von einer Person mit softwareergonomischen Kenntnissen entwerfen zu lassen. Informationswissenschaftler verfügen hier über wichtige Voraussetzungen.

Es wird neben der Schilderung der technischen und theoretischen Gegebenheiten immer wieder auf Beispiele aus der Praxis, d.h. auf konkrete Softwareprodukte eingegangen. Im Anschluß werden UIMS unter folgenden vier Gesichtspunkten betrachtet und den Toolkits und Fenstersystemen gegenübergestellt: Handhabung, Funktionalität, Portabilität und Evaluationsunterstützung.

Neben den neuen Anforderungen an die Applikations-Entwicklung werden Entwurfswerkzeuge für Graphische Benutzeroberflächen vor allem auch danach zu beurteilen sein, ob sie sich für einen iterativen Softwareentwicklungsprozeß in Zusammenarbeit mit dem späteren Benutzer eignen.

Es wird auf drei grundlegende Prinzipien eingegangen, die als Teile in den Gestaltungsprozeß einzubeziehen sind: iteratives Design, Benutzerbeteiligung und empirische Evaluation. Eine Hauptvoraussetzung für eine Gestaltung nach diesen drei Prinzipien ist die Erstellung von Prototypen. Beim Prototyping der Benutzeroberfläche zeigt sich, daß die Trennung von der Anwendung besonders wichtig ist.

Obwohl gerade im Bereich der UIMS noch viele Probleme bewältigt werden müssen (Portabilität, Handhabung), stellt der Einsatz von offenen Applikationsumgebungen und von UIMS eine äußerst funktionale Entwicklungsumgebung zur Verfügung, die auch den Einsatz von Rapid-Prototyping Verfahren ermöglicht.

Ausführliche Informationen über Magisterarbeit und Vortrag können eingeholt werden unter folgender Adresse:

Barbara Jochum  
Heinrichstr. 4a  
6683 Elversberg  
Tel. : 06821/73524



# Ein Konzept für die computergestützte Auftragsabwicklung in der Informationsvermittlung

Astrid Sprenger

Die Idee für dieses Thema entstand während eines Praktikums aus der Notwendigkeit heraus, daß jener Informationsvermittlungsdienst auf eine andere Software umsteigen wollte. Die dort bisher eingesetzte Software stößt zum Teil an ihre Grenzen und die Funktionsvielfalt des Datenbankmoduls und der Tabellenkalkulation entsprechen nicht dem Standard an Funktionen die heute professionelle Datenbank-Systeme bieten.

Damit für die individuelle Umsetzung dieser Auftragsabwicklung genügend Freiraum vorhanden ist, wurden zunächst alle Etappen der Ablauforganisation, angefangen vom Eingang des Auftrages bis zum Zusenden der Rechercheergebnisse an den Kunden, unabhängig von dem Werkzeug, das später eingesetzt werden soll, analysiert.

Das Ergebnis dieser Analyse stellt die einzelnen notwendigen Schritte für diese Auftragsabwicklung dar, sie zeigt, die in der bestehenden Ablauforganisation vorhandenen Schwachstellen auf und schafft so die Voraussetzungen für die Entwicklung eines neuen Systems. Im ersten Teil des Vortrages wird die Ist-Analyse vorgestellt und anhand dieser die bisher durchgeführten Schritte der Auftragsabwicklung nachvollzogen. Bei der Bewertung der Ist-Analyse muß zwischen quantifizierbaren und nicht quantifizierbaren Mängeln unterschieden werden. Quantifizierbar sind:

- Überstunden
- entgangene Aufträge
- Umsatzverluste
- Kundenbeanstandungen
- ...

Schwer oder nicht quantifizierbare Mängel sind:

- Unvollständigkeit des Datenbestände
- unzureichende Aussagefähigkeit
- ...

Die Ist-Analyse bezieht sich nur auf die Vermittlung der Informationen, nicht auf die Quellen und die Bereitstellung der Informationen in den Datenbanken.

Im nächsten Schritt erfolgte eine Untersuchung der auf dem Markt vorhandenen Informationstechnologien und Techniken. Dabei ging es um die Möglichkeit solche Werkzeuge wie Multimedia, Hypertext oder neue Speichermedien gezielt in der Ablauforganisation für die Informationsvermittlung einzusetzen.

Danach erfolgte eine Analyse der bisher eingesetzten Software hinsichtlich :

- Benutzerführung (Menü, Kommandos)
- Datenbankbeschreibungen
- Eingabe-Masken
- Abfrage- und Auswertungsmöglichkeiten
- Ausgabe-Maske

Die Ergebnisse der einzelnen Untersuchungen wurden zu einem Konzept zusammengefaßt. Das Anliegen dieses Konzeptes besteht darin, zu zeigen, wie sich bestehende Mängel in der Ablauforganisation beseitigen lassen. Weitere Aufgaben dieses Konzeptes sind die Festlegung des Leistungsumfanges und die Vorgaben für die Arbeitsabläufe. Diese Vorgaben enthalten:

- Beschreibungen der Datenstrukturen
- Aufführung der einzurichtenden Datenbanken
- Beschreibungen der Dialogabläufe
- Muster für Bildschirm-Masken

Anhand dieser Darstellungen kann der Auftraggeber entscheiden, ob er für die computergestützte Umsetzung seiner Ablauforganisation ein auf dem Markt vorhandenes Produkt nutzen oder ob die Programmierung eines eigenständigen Systems sinnvoller ist.

Das Konzept ist so allgemein gehalten, daß es sich ohne großen Aufwand an andere Auftragsabwicklungen anpassen läßt.

# Die vernetzte Weltgemeinschaft: Potentiale und Risiken

## - Entwicklungsperspektiven der Informationswirtschaft und des Informationsmanagements -

Karl A. Stroetmann

### **Kurzfassung:**

Ausgehend vom Paradigma der vernetzten Weltgesellschaft ist es Ziel des Papiers, zentrale Entwicklungsperspektiven der Informationswirtschaft und Konsequenzen für das Informationsmanagement aufzuzeigen. Die Diffusionsaspekte Individualisierung bei gleichzeitiger Globalisierung sowie die Integrationsaspekte Dezentralisation bei gleichzeitiger Re-Zentralisation dienen hierbei als Analysekategorien. Bei der Informationsprozeßsteuerung immer wesentlicher werdende Sicherheits- und Verfügbarkeitsgesichtspunkte werden kurz gestreift.

### **Abstract:**

Based on the paradigm of the globally networked society, it is the objective of the paper to outline fundamental trends of the information economy and their consequences for information management. Diffusion categories of individualization with concurrent globalization, and integration categories of decentralisation with concurrent re-centralization will serve to structure the topic. Security and availability aspects, becoming priority fields for information process controll, will be briefly touched.

### **Inhalt**

- 1. Einführung und Überblick**
- 2. Der Informationsbegriff**
- 3. Strukturierung des Informationskontextes**
- 4. Entwicklungsperspektiven**
  - 4.1 Das Paradigma der vernetzten Weltgesellschaft**
  - 4.2 Diffusionsaspekte**
    - 4.2.1 Individualisierung
    - 4.2.2 Globalisierung
  - 4.3 Integrationsaspekte**
    - 4.3.1 Dezentralisation
    - 4.3.2 Zentralisation
  - 4.4 Prozeßsteuerung**
    - 4.4.1 Informationsmanagement
    - 4.4.2 Sicherheit/Verfügbarkeit
- 5. Ausblick**

Anmerkungen und Referenzen

# 1. Einführung und Überblick

Information wird immer mehr zum alles durchdringenden Wettbewerbsfaktor. Sie ist ein zentrales Mittel für die Erzeugung von Mehrwert, sei es in der Industrie, im Dienstleistungssektor, in der Bürokratie oder in der Forschung. Der strategische Umgang mit ihr, die Fähigkeit, Informationsinhalte und -strukturen zu managen, entscheidet zunehmend über den langfristigen Erfolg von einzelnen Unternehmungen wie auch von ganzen Regionen und Nationen.

Ausgehend von dieser Grundthese werde ich im folgenden zunächst auf verschiedene Begriffsinhalte von „Information“ eingehen, um anschließend fünf Akteursebenen, einen statischen Gesamtbezugsrahmen und ein dynamisch orientiertes Analyseraster zu skizzieren. Ausgehend vom Paradigma der vernetzten Weltgemeinschaft - welches m.E. nicht nur für die Informationspraxis sondern auch für die Informationswissenschaft in diesem Jahrzehnt und darüberhinaus von grundlegender Bedeutung sein wird - werde ich im zentralen Abschnitt sodann auf einige mir besonders wichtig erscheinende Entwicklungsperspektiven eingehen, die für die Informationspraxis in all ihren Facetten prägend sein werden. Unter dem Oberbegriff Diffusion habe ich die Tendenzen der Individualisierung bei gleichzeitiger Globalisierung und unter dem Begriff Integration die der Dezentralisation bei gleichzeitiger Zentralisation zusammengefaßt. Aus der Sicht der Steuerung der in vernetzten Volkswirtschaften/Gesellschaften ablaufenden Informationsprozesse werde ich die Aspekte Informationsmanagement und Sicherheit/Verfügbarkeit kurz aufgreifen.

## 2. Der Informationsbegriff

Zum Informationsbegriff gibt es eine endlose Diskussion, die sicherlich auch in Zukunft noch viele Seiten Papier und zunehmend elektronische Bulletin-Boards u.ä. füllen wird.<sup>1</sup> Gitt erklärt Information sogar „zur dritten universellen Grundgröße“ - neben Materie und Energie.<sup>2</sup> Eine allgemeinverbindliche Definition wird es nicht geben, sie ist vielmehr immer von Fragestellung und Kontext, Erkenntnisziel und Untersuchungsobjekt abhängig. Mir scheint zunächst schon die einfache Unterscheidung zwischen transaktionsorientierter oder auf Englisch „administrative“ Information (z.B. Bestellungen, Geldüberweisungen) und dispositions-(entscheidungs-)orientierter Information hilfreich,<sup>3</sup> wobei als dritte Kategorie wissensorientierte Information (z.B. Methodeninformation, Archive)

---

<sup>1</sup>Vgl. z.B. Buckland, Michael K.(1991): Information as Thing. In: Journal of the American Society for Information Science 42 (5), 351-360; Finnemann, Niels O. (1990): Computerization as a Means of Cultural Change: On the Relations Between Information Theories and the Idea of an Information Society. In: AI & Society 4, 314-328; Folberth, Otto G. u.a. (Hrsg.): Der Informationsbegriff in Technik und Wissenschaft. München, Wien: Oldenbourg 1986; Schrader, Alvin M. (1986): The domain of the information science: problems in conceptualization and in consensus-building. In: Information Services & Use 6, 169-205.

<sup>2</sup>Gitt, Werner (1989): Den Zweck der Information klärt nur die fünfte Ebene. In: PC Woche 39, 17.

<sup>3</sup>Augustin unterscheidet zwischen Bestandsinformationen (Informationen über Potentiale) - die Entscheidungen bewirken, welche als Flüsse wirksam werden - und Flußinformationen. Augustin, Siegfried (1990): Information als Wettbewerbsfaktor: Informationsklogistik - Herausforderung an das Management. Köln: TÜV Rheinland, 15 ff.

hinzukommen sollte. *Eaton* und *Bawden* weisen darauf hin, daß das „information resource management“ (IRM) Modell, welches in seinen Ursprüngen von der Idee ausgeht, daß Information ein Produkt (commodity) ist, „is arguably most suitable for administrative information, of the sort generated within organizational bureaucracies. This [information] will probably have a clear use, and hence value, a well-defined life-cycle, and other parameters in accordance with IRM ideas. But it is the last kind of information which will be a resource in the other sense: providing value to the organization through encouraging innovation and change.“<sup>4</sup> Viele, u.U. gerade besonders relevante Informationen, liegen oft nicht in formalisierter Form vor: „Information may take the form of stories, even myths with little basis in provable fact, which nonetheless link policy ideas to community experience and values, make sense of complex realities, and motivate collective action.“<sup>5</sup> Dies mag eine teilweise Erklärung dafür sein, warum die Nutzung konventioneller elektronischer Informationsdienste z.B. im Unternehmensbereich noch so dürftig ist. Andererseits weist *Bellin* darauf hin, daß zunehmend ‚privates Wissen‘, das in den Köpfen von Menschen verfügbar ist, z.B. über Expertensysteme und damit in formalisierter Form verfügbar gemacht wird.<sup>6</sup>

Ich plädiere in erster Annäherung für einen umfassenden, ganzheitlichen Informationsbegriff. Die verschiedenen Informationsarten lassen sich in der Praxis kaum klar voneinander trennen, es gibt vielmehr situativ bedingt graduelle Abstufungen und Überlappungen. In der vernetzten Gesellschaft wird der Nutzer auf viele Arten von Informationen nebeneinander und integrativ, ja nach Situation, zugreifen, sie weiterverarbeiten, kommunikativ austauschen, verändern. - Hinweisen möchte ich auch noch auf den Service-Charakter von Information. Sie ist kein Selbstzweck - außer wenn sie zur Unterhaltung „konsumiert“ wird, in welchem Falle die Kennzeichnung „Information“ fraglich wird - sondern dient immer anderen Zielen. Oder, wie *Quinn* es ausdrückte: „Information is interesting only insofar as it affects other things.“<sup>7</sup>

### 3. Strukturierung des Informationskontextes

**Schaubild 1** gibt einen Überblick über zentrale Akteursebenen im Bereich der Informationswirtschaft. Standen bisher Institutionen/Unternehmungen als Produzenten und Nutzer von Informationen und nationale Regierungen als Hauptakteure bei der Förderung und Regulierung des Informationsmarktes im Mittel-

---

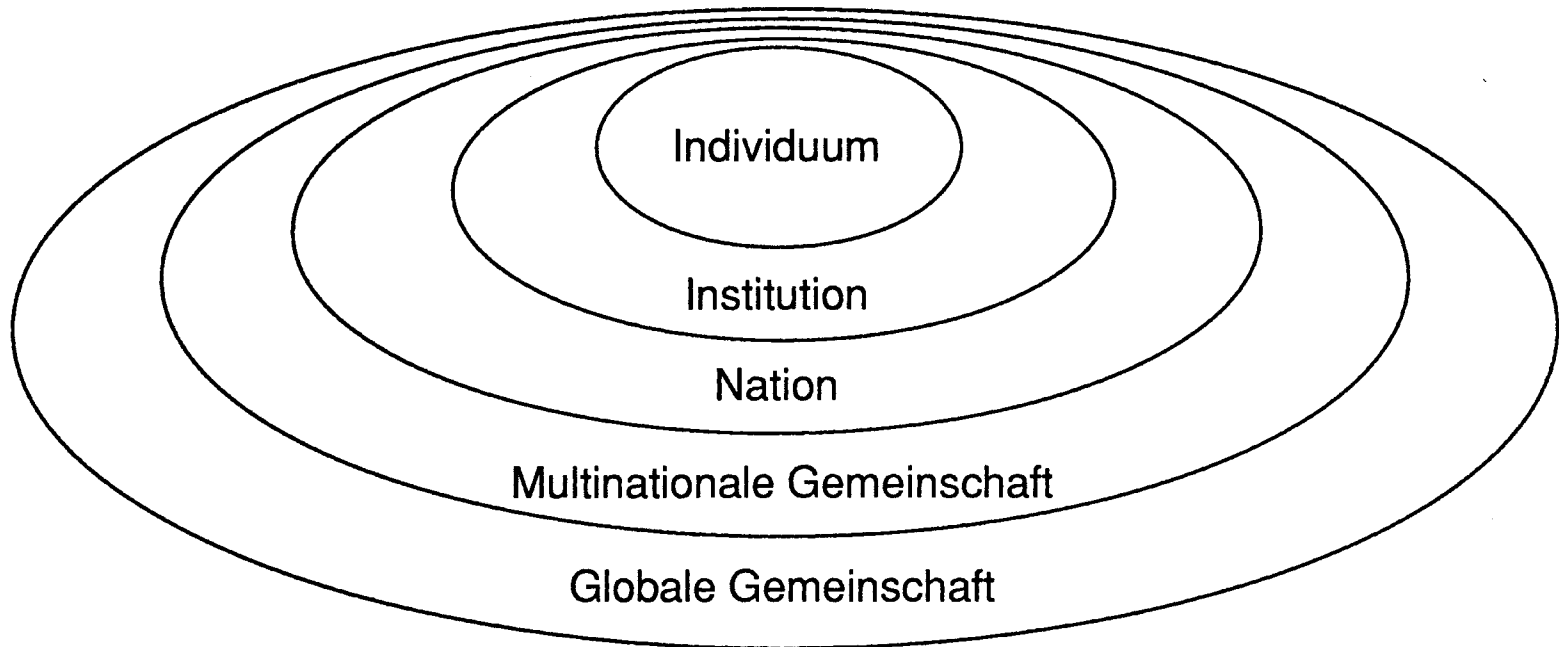
<sup>4</sup>Eaton, J.J.und Bawden, D. (1991): What Kind of Resource is Information? In: International Journal of Information Management 11, 164. - Finnemann drückt eine ähnliche Idee so aus: „The concept of information has acquired a double sense, both including human knowledge in general and those knowledge forms which we are able to formalize and hence implement in machines.“ A.a.O. (s. Fußnote 1), 321.

<sup>5</sup>Innes, Judith (1989) : The Construction of Information in Planning. In: Knowledge in Society 2 (2), 6.

<sup>6</sup>Bellin, David (1991): The Commoditization of Information - Political and Economic Realities. In: Learned Information: Proceedings of the Twelfth National Online Meeting. Medford, NJ. - Paul van der Vet, Universität Twente, Enschede/NL formulierte in der Ankündigung seines Vortrages „The Concept of Concept“ in der GMD am 16.10.1991 eine ähnliche Überlegung: „It is evident that even scientists themselves will have to get used to formalized versions of their concepts, for the massive amount of information a scientist has to take into account can no longer be handled without automated technique.“

<sup>7</sup> Quinn, James Brian, zitiert in: Wright, Karen (1990): The Global Village. In: Scientific American (March), 59.

Schaubild 1: **Ebenen der Akteure / Teilnehmer**



punkt, so dehnt sich das Aktionsfeld sowohl in Richtung Individuum/ Haushalt<sup>8</sup> als auch multinationale Gemeinschaften wie z.B. die EG und hin zu globalen Institutionen<sup>9</sup> aus. **Schaubild 2** vermittelt eine grobe Strukturierung des (statischen) Gesamtkontextes von Information, wie er für die je unterschiedlichen Akteursebenen - mit sicherlich unterschiedlich verteilten Gewichten - von Relevanz ist. Wichtig sind hier insbesondere zwei Aspekte: Ganz gleich auf welcher Akteursebene und in welchem größeren Kontext, das Management der Ressource ‚Information‘, d.h. die zielgerichtete Disposition, deren Erfolg über (ökonomische) Kriterien meßbar und kontrollierbar sein muß, spielt eine zentrale Rolle; der Informationstechnik kommt hingegen ‚nur‘ Mittelcharakter zu.

Wie schon erwähnt, ist Information kein Selbstzweck, sondern selbst als Mittel für andere Zwecke in dynamische Leistungsprozesse (Wertschöpfung) eingeordnet. **Schaubild 3** stellt ein dreidimensionales Analyseraster vor, um solche Prozesse zu analysieren.<sup>10</sup> Dabei wurde die Darstellung der ‚Wertschöpfungskette‘ bewußt sehr allgemein gehalten; sowohl auf der Ebene eines einzelnen Arbeitsplatzes als auch auf der von Organisationen (Industrie, Dienstleistungen, öffentliche Verwaltung, Forschung und Wissenschaft) und Nationen<sup>11</sup> lassen sich solche Prozesse unter Informationsgesichtspunkten analysieren. Dabei handelt sich jeweils nur um einen Ausschnitt aus verbundenen, interdependenten und auch rückgekoppelten Prozessen. Im folgenden kann nicht jeweils auf alle denkbaren Ebenen von Akteuren, Strukturaspekte oder Prozeßkomponenten eingegangen werden. Diese Grafiken sollen vielmehr nur eine erste Handreichung zur Reduzierung von Komplexität und zur Strukturierung von multidimensionalen Zusammenhängen geben.

#### 4. Entwicklungsperspektiven

Zur Gliederung dieses Abschnittes dient mir **Schaubild 4**. Damit wird der Versuch unternommen, die folgernden Ausführungen in einem konzeptionellen Rahmen zusammenzufassen und in ihrer Interdependenz zu verdeutlichen.

---

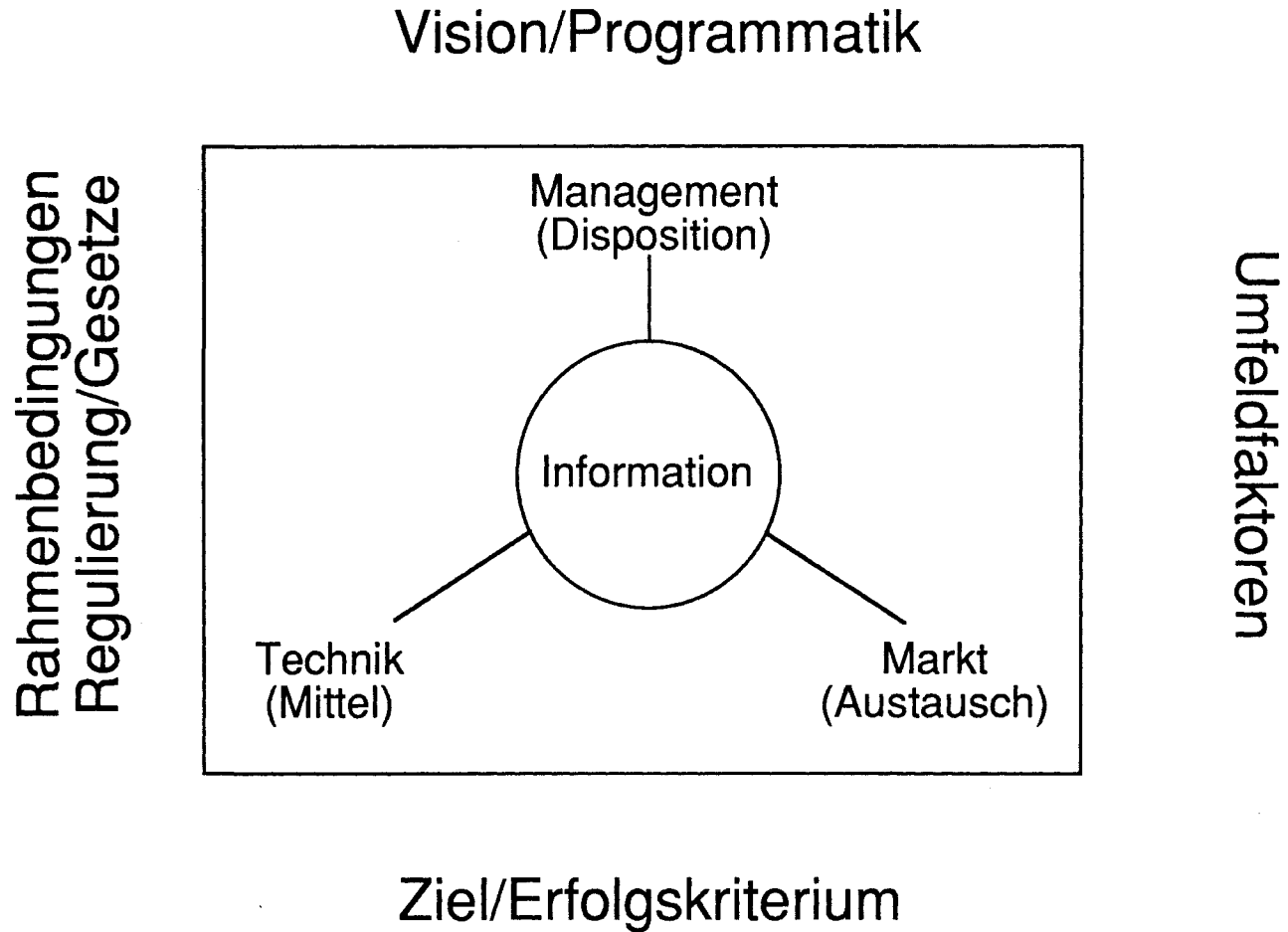
<sup>8</sup>Vgl. dazu Frerk, Giesela (1990): Economic Effects of the Use of New Information Technologies in Private Households. In: International Journal of Information and Library Research 2 (1), 11-21.

<sup>9</sup>Vgl. z.B. die Electronic Data Interchange (EDI) Aktivitäten der UN und der EG: Klein, Stefan (1992): Der Stellenwert von EDI in der Informationslogistik. In: Stroetmann, Karl A. (Hrsg.): Informationslogistik. Managementkonzepte, Fallbeispiele und Anwendererfahrungen bei Informationsprozessen. Proceedings der 6. Internationalen Fachkonferenz der Deutschen Gesellschaft für Dokumentation (DGD) in Zusammenarbeit mit der Gesellschaft für Informatik (GI) in Garmisch-Partenkirchen, 27.-29. Mai 1991. Frankfurt: DGD (in Vorbereitung). - In Kanada wurde gerade (Sept. 1991) das EDI-Weltinstitut gegründet.

<sup>10</sup>Siehe hierzu den grundlegenden Aufsatz von Porter, Michael E. and Millar, Victor E. (1985): How information gives you competitive advantage. In: Harvard Business Review (4), 149-160. Außerdem: Tulowitzki, Ulrich (1990): Pragmatic approaches to strategic information systems planning. In: Zeitschrift für Planung (2), 1-16 und Wollnik, Michael (1988): Ein Referenzmodell des Informationsmanagement. In: Information Management 3 (3), 34-43.

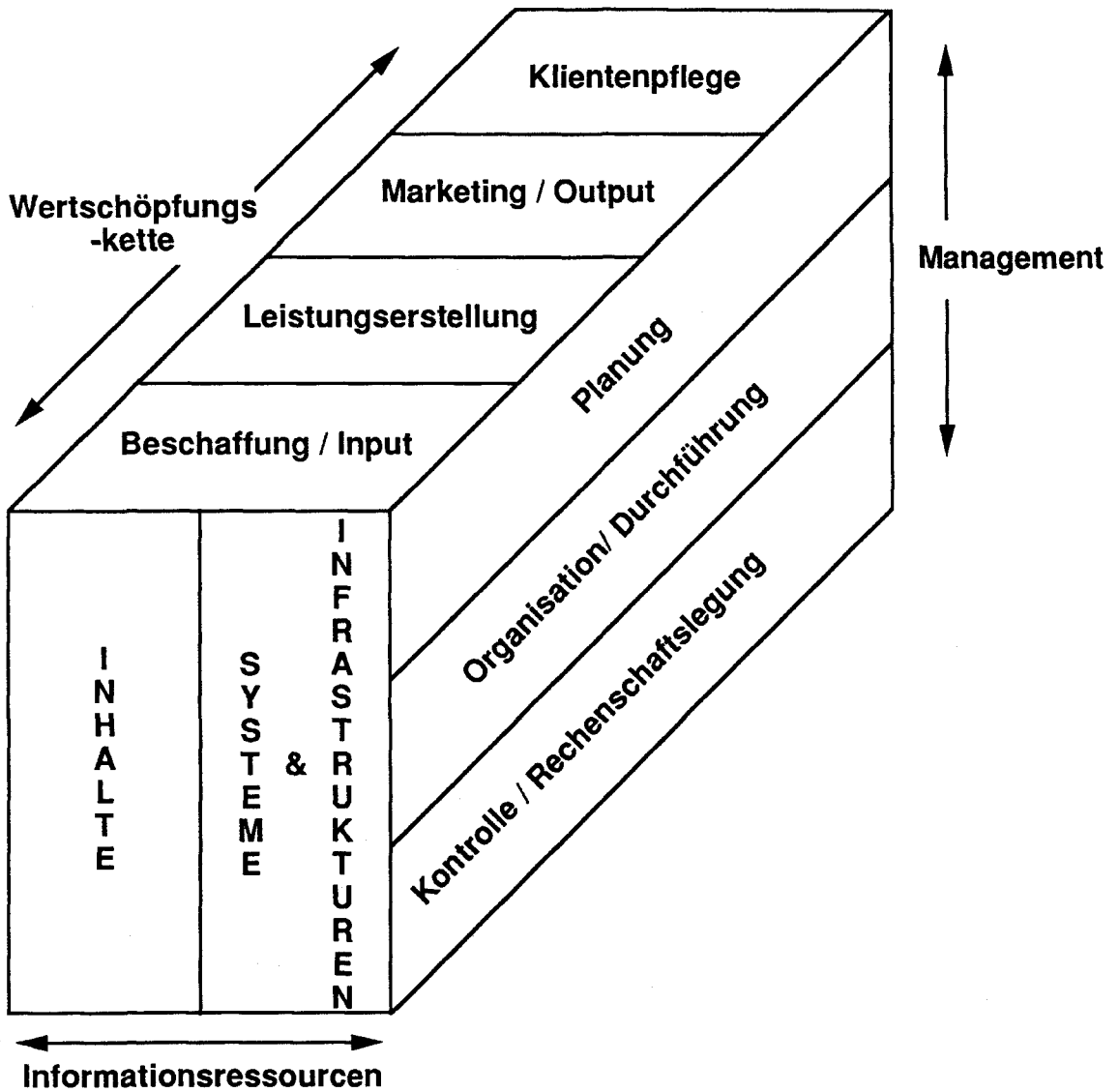
<sup>11</sup>„Economic power ... will come increasingly from the maximisation of the human intellect. ... Today, comparative advantage is something that a nation or an industry creates for itself.“ Eijsvoogel, Juurd (1990) in der Introduction zu The Discipline of Curiosity. Amsterdam: Elsevier (zitiert nach „Guest Editorial“. In: Journal of Information Science 17 (1), 78).

## Schaubild 2: Der Informationskontext





**Schaubild 3: EIN ANALYSERAHMEN**



## 4.1 Das Paradigma der vernetzten Weltgesellschaft

War noch vor kurzem die dramatische Verbesserung der Informationsverarbeitungsmöglichkeiten am einzelnen Arbeitsplatz durch PC und Workstation die dominante Entwicklung, so ist schon jetzt absehbar, daß das letzte Jahrzehnt dieses Jahrtausends vom Trend zur vernetzten Weltgemeinschaft geprägt sein wird. Was ist die Ursache? Standen Kostenreduzierung, Rationalisierung und Qualitätsverbesserung lange Zeit als Wettbewerbsfaktor im Zentrum des Interesses, so ist dies mehr und mehr der Faktor ‚Zeit‘ als strategischer Wettbewerbsvorteil.<sup>12</sup> Neue Organisationsformen zur Verbesserung der Kundennähe, institutionenübergreifende Prozeßketten, globales Risikomanagement, Outsourcing, Strategische Allianzen, elektronische Märkte verlangen nach einer entsprechenden Informationsversorgung und den dazu erforderlichen Kommunikationsnetzen. Neben die Materiallogistik tritt die globale Informationslogistik. Dieser Teil des Computermarktes zeigt die größten Wachstumsraten.<sup>13</sup> Value Added Network Services (VANS) in ihren vielfältigen Formen, von E-Mail und Electronic Data Interchange (EDI) bis zu sog. Community Systemen („gemeinschaftlich genutzte Informationsdienste für globale Aufgabenstellungen“<sup>14</sup>) sind ein schnell wachsender Markt.<sup>15</sup> Breitband-ISDN bzw. das geplante Integrierte Breitband Fernmeldenetz (IBFN) eröffnen ganz neue Möglichkeiten für Informationsdienste in unterschiedlichster Form und Kombination.<sup>16</sup> Die sich aus der globalen Vernetzung ergebenden Möglichkeiten und Risiken werden ganz entscheidend die Informationspraxis der Zukunft determinieren, nicht nur im Produktions- und Dienstleistungssektor - sei es als die Leistungsprozesse begleitende Informationsprozesse, sei es als originäre Informationsprozesse, die zu primären Informationsprodukten und -leistungen führen -, sondern z.B. auch im Wissenschaftsbereich.<sup>17</sup>

## 4.2 Diffusionsaspekte

Unter dem Begriff der Diffusion sollen die beiden zunächst marktmäßig gedachten gegensätzlichen Begriffe der Individualisierung und Globalisierung subsumiert werden.

---

<sup>12</sup>Schertler, Walter (1992): Informationslogistik und Strategische Unternehmensplanung. In: Stroetmann, Karl A. (Hrsg.): Informationslogistik. A.a.O. (s. Fußnote 9).

<sup>13</sup>O.V.(1991): Open Systems On-Line Transaction Processing. In: Computer Magazin 20 (3), 53-55.

<sup>14</sup>Beyenburg, R. und Emde, P. (1990): Interaktion im globalen Netz. In: net 44 (1-2), 18.

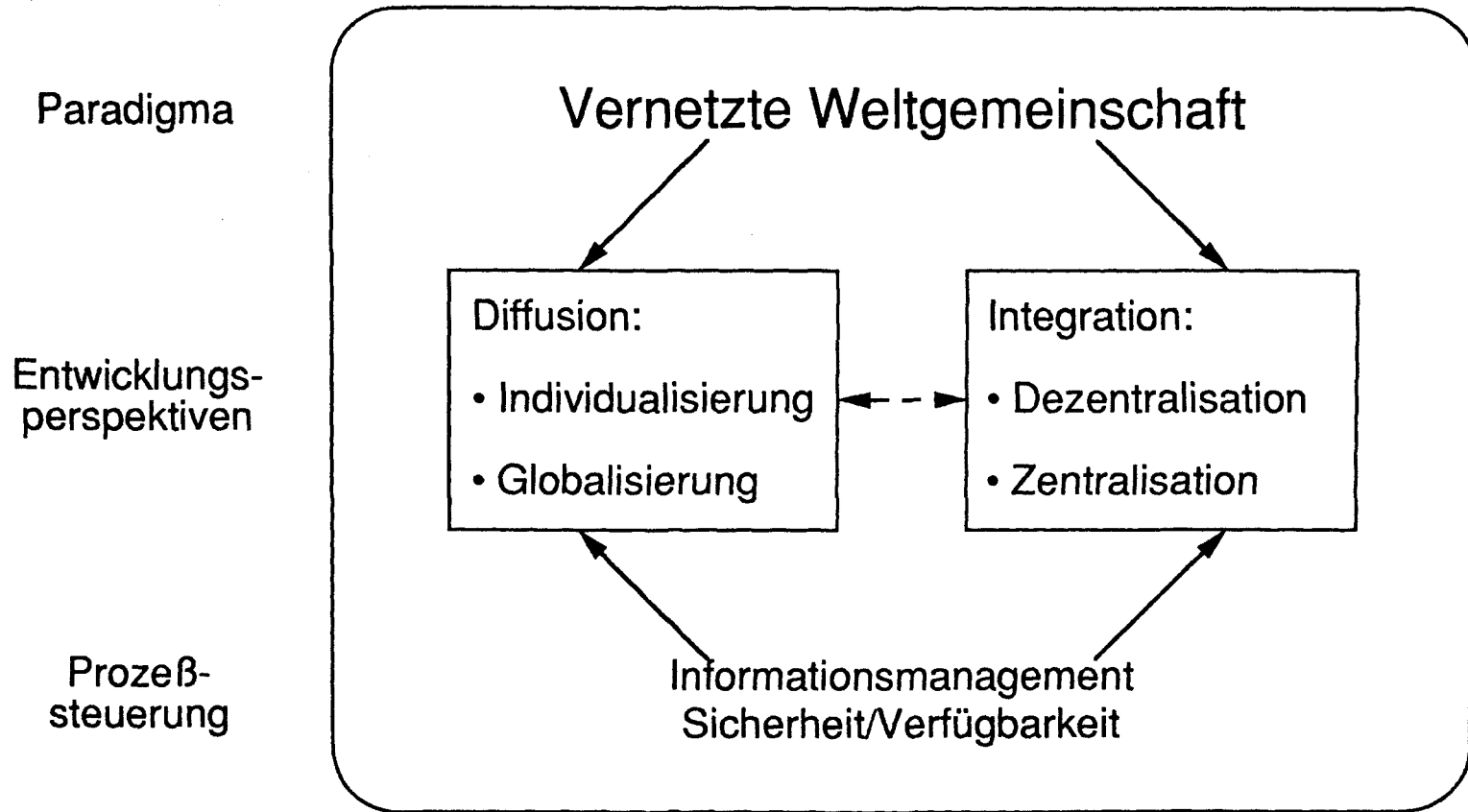
<sup>15</sup>Techke, H. (1991): Value Added Network Services im Rahmen der europäischen Integration: Stand und Entwicklung. In: Proceedings der 14. Europäischen Congressmesse für Technische Kommunikation in Hamburg, 04.-08.02.1991, Symposium V-1. Velbert: Online GmbH, V.05.01-V.05.26.

<sup>16</sup>Kanzow, Jürgen (1991): In Zusammenhängen denken. Das Projekt Berkomp - Stand und bisherige Entwicklung. In: net 45 (3), 76-81.

<sup>17</sup>Siehe Stichweh, Rudolf (1989): Computer, Kommunikation und Wissenschaft: Telekommunikative Medien und Strukturen der Kommunikation im Wissenschaftssystem. Max-Planck-Institut für Gesellschaftsforschung Discussion Paper 89/11, Köln. - An der Aston University in Birmingham wird mit einem Aufwand von ca. 12 Mio DM ein Campus-weites LAN mit 2600 „Service Points“ aufgebaut, welches über „LAN/WAN relays“ global vernetzt wird. Auf konventionelle Bibliotheksdienste soll weitgehend verzichtet werden. Press Release (ohne Datum), Project ACCENT, Aston University, Birmingham/GB.

Schaubild 4:

# Die vernetzte Weltgemeinschaft: Entwicklungs- und Steuerungsaspekte



#### 4.2.1 Individualisierung

Mit dem Begriff Individualisierung soll die Tendenz gekennzeichnet werden, daß zunehmend Informationsdienstleistungen nicht mehr nur auf ‚Massenbedienung‘ ausgerichtet werden, sondern daß dank Vernetzung und ‚intelligenter‘ Software die Bedienung spezialisierter Nischenmärkte und von Individuen mit ihren u.U. sehr persönlichen Informationsbedürfnissen einerseits möglich wird, andererseits dies ganz neue Marketing- und Servicemöglichkeiten auch für andere Produkte und Dienstleistungen eröffnet. Grundvoraussetzung werden jedoch vielfach weiterhin große, umfassende Datenbanken sein, aus denen je nach Informationsbedarf und Kontext z.T. auf globaler Ebene die jeweils relevanten Daten gefiltert werden.

Einige Beispiele sollen diese Individualisierungstendenz illustrieren: *Individual Inc.* in Cambridge, Mass./USA, produziert einen täglichen Informationsdienst, der mittels neu entwickelter Filter-Software aus Volltextdatenbanken mit einer Relevanzrate von 75-80% (nach einer gewissen Lernphase) nur wenige, auf den individuellen Informationsbedarf eines jeden Abonnenten zugeschnittene Informationen morgens zu Arbeitsbeginn per E-Mail oder Fax bereitstellt.<sup>18</sup> Andere Übertragungstechniken wie ‚point-to-multipoint‘ Satelliten- oder Rundfunkübertragung werden auch eingesetzt; Audio-Dienste sind eine weitere nutzerfreundliche Möglichkeit. Zentraler Ansatzpunkt ist, durch Reduzierung der Informationsflut und einfachste Zugangsmöglichkeiten den großen Endnutzermarkt im Wirtschaftsbereich zu erschließen. - Die Schweiz-Direkt-Versicherungsgesellschaft bietet gegen eine Schutzgebühr eine Diskette an, die es erlaubt, nach Feierabend auf dem heimischen PC den persönlichen Versicherungsbedarf zu ermitteln, den gegenwärtigen Versicherungsschutz zu optimieren, die gesamte Finanzsituation des Haushaltes zu analysieren und schließlich sich selbst ein Versicherungsangebot mit den jeweiligen Bedingungen zusammenzustellen. - In Japan haben Hitachi und die Mitsui Lebensversicherung ein Vermögensverwaltungssystem auf der Basis der fuzzy Logik entwickelt, daß selbst dem völlig Ungeübten die Optimierung seines Vermögensbestandes auf Anhieb ermöglichen soll.<sup>19</sup> - Beim sog. „Mikro-Marketing“ werden die Einkaufsinformationen für den Kunden z.B. im Supermarkt durch Individualisierung verbessert. Am Einkaufswagen befindet sich ein Boardcomputer, der dem Kunden den Weg zum gesuchten Artikel per Lageplan weist; dort angekommen, kann er ergänzende Produktinformationen abrufen, und auf dem Weg dorthin wird er auf Sonderangebote, an denen er gerade vorbeifährt, u.a. aufmerksam gemacht.

Für die Informationspraxis eröffnet sich hier ein riesiges neues Marktpotential. Eine Grundvoraussetzung ist, daß die Endnutzer mit der Informationstechnik so vertraut wie mit dem Autofahren werden (können) und daß sich diese flexibel den jeweiligen Bedürfnissen anpassen läßt. Eine andere Bedingung ist, daß den psychischen und sozio-ökonomischen Bedürfnissen der Nutzer und dem jeweiligen Umfeld stärker als bisher Rechnung getragen wird. - Individualisierte Märkte erfordern individuelle Marketinginformationen über potentielle Kunden, wiederum ein zukunftssträchtiges Feld für Informationsdienste, das gleichzeitig

---

<sup>18</sup>Shillingford, Joia (1990): Keeping the watching brief for the individual customer. In: Financial Times (13.Dec.).

<sup>19</sup>Vgl. auch Schuster, Ralf (1992): Euro-FIS - Das Software-Produkt zur Integration von internationalen Finanzinformationen. In: Stroetmann, Karl A. (Hrsg.): Informationslogistik. A.a.O. (Fußnote 9).

aber auch große Datenschutzprobleme aufwirft und ohne staatliche Regulierung kaum auskommen dürfte.<sup>20</sup>

Als Randbemerkung sei noch auf eine völlig andere Art der Individualisierung hingewiesen: Dank weiträumiger Vernetzung von z.B. ganzen Industrieanlagen oder Krankenhäusern wird es möglich, jeden Schritt von Mitarbeitern zu überwachen - oder eingehende Telefonanrufe an den jeweils nächsten Apparat weiterzuleiten. Infrarotsensoren geben die Signale von mit Transistoren bestückten Plaketten an den nächsten vernetzten PC weiter. Bei diesen „Follow-me“ genannten Systemen hakt es noch im Sicherheitsbereich: Damit sind jedoch nicht Datenschutz- oder Persönlichkeitsrechte gemeint, sondern die Tatsache, daß das System wohl die Ortveränderung der Plakette, aber nicht notwendigerweise die der dazugehörigen Person aufzeichnet.<sup>21</sup>

#### 4.2.2 Globalisierung

Mit Globalisierung soll die Tendenz bezeichnet werden, daß dank Vernetzung und VANS nationale (Informations-)Märkte zu multinationalen und globalen werden, die für Anbieter, Nutzer und Regulierungsinstanzen ganz neue Gegebenheiten schaffen. „Advanced information technology will be utilized for the advancement, the multilateralization, and the merging of industries on a global scale, transcending the boundaries among industries and nations. In order to survive in this ‚borderless economy‘ of the 21st century, each enterprise must recognize the environmental changes constantly and quickly, make adequate decisions to create a strategy to cope with the changes, and to maintain its advantages in competitions.“<sup>22</sup>

Echte Weltmärkte für Finanz-Transaktionen, Autos oder Computer wurden erst durch die Globalisierung von Informationsleistungen möglich, und immer mehr Industrie- und Dienstleistungsbereiche werden von diesem Trend erfaßt. Umgekehrt werden für viele Informationsdienstleistungen neben globaler Leistungserbringung und Kundenservice auch globaler Vertrieb, Marketing und Entwicklung unumgänglich. Auch der Arbeitsmarkt wird zu einem globalen Markt.<sup>23</sup> Electronic Immigration, d.h. der globale elektronische Transfer von Arbeitsleistungen wie Dateneingabe, Informationsaufbereitung oder Programmierung sind große Wachstumsmärkte - ohne daß die Konsequenzen für nationale Arbeitsmärkte, Arbeitnehmerrechte oder (internationale) Marktregulation auch nur angedacht wären.

Die strategischen Implikationen für Informationsdienstleistungs- (und andere) Unternehmen sind vielfältiger Art. „Network-based interactions can be highly competitive and, yet, they are dependent on agreements regarding rules, thereby

---

<sup>20</sup>Zu solchen Fragestellungen vgl. Burkert, H (1989): The European Information Services Market: Challenges for Data Protection. Vortrag auf der 11. Internationalen Konferenz der Datenschutzbeauftragten, Berlin (30.08).

<sup>21</sup>Klein, Michael (1991): Der Große Bruder mit besonders scharfem Blick - Computernetze überwachen jeden Schritt am Arbeitsplatz. In: Süddeutsche Zeitung (05.09.).

<sup>22</sup>Japan Information Processing Development Center (JIPDEC) (1991): Informatization White Paper. Tokio, 2.

<sup>23</sup>Johnston, William B. (1991): Global Workforce 2000: The new World Labor Market. In: Harvard Business Review (2), 115-127.

requiring a level of cooperation unforeseen in traditional economic analysis: *actors need to cooperate to compete*. ... Multilateral rules will increasingly need to be seen as the result of corporate strategies as well as the framework in which they are supposed to take place."<sup>24</sup> - Auch für das Setzen von Rahmenbedingungen und regulativen Regeln durch nationale und multinationale Administrationen ergibt sich eine ganz neue Qualität. Der freie Zugang zu Netzwerken und ‚öffentlichen‘ Informationen, Wettbewerbsregeln für globale Märkte, Datenschutz und Persönlichkeitsrechte, die Gefahr einer „screen-based democracy without rules“,<sup>25</sup> der Widerspruch zwischen globaler technischer und ökonomischer Standardisierung und den politischen, kulturellen und regulativen Interessen einzelner Nationen, die Verletzung nationaler Ethikkonventionen und von Gesetzen durch Informationsdienste jenseits der nationalen Grenzen sind nur einige der drängenden Fragestellungen, die z.T. bereits die Aktionsmöglichkeiten multinationaler Gemeinschaften wie die der EG überfordern - ganz zu schweigen davon, daß sie z.T. auf nationaler Ebene noch gar nicht als (informations-)politische Probleme erkannt werden.

### 4.3 Integrationsaspekte

Unter Integration wird die Bündelung und Zusammenfassung ganz unterschiedlicher Informationen und Informationsverarbeitungsprozesse verstanden. Dies ermöglicht einerseits eine zunehmende Dezentralisation der Informationsverarbeitung beim Individuum, andererseits die verstärkte Zentralisation wesentlicher Dispositions- und Kontrollaufgaben in Unternehmensspitzen, Konzernen oder nationalen Behörden.

#### 4.3.1 Dezentralisation

Die zunehmende Vernetzung von Arbeitsplätzen, anspruchsvolle ‚Client/Server‘-Konzepte, mächtige Arbeitsplatzrechner, die Einbindung von CD-Rom-Laufwerken, die Anbindung an externe Online-Dienste werden es immer mehr erlauben, dem einzelnen situationsgerecht und entsprechend seinem individuellen Kommunikationsverhalten am Arbeitsplatz die für die jeweilige Aktivität relevanten Informationen - und nur diese - in der jeweils angemessenen Form zur Verfügung zu stellen. Gleichzeitig wird die zur Verfügung stehende Informationsverarbeitungssoftware immer mächtiger, d.h. der Lernaufwand wird geringer, sie entlastet zunehmend von Routinetätigkeiten, die Nebenkosten der Informationsverarbeitung sinken, bei abnehmender Arbeitsteilung (job enrichment) kann der eigentliche Informationsgenerierungsaufwand relativ erhöht werden.<sup>26</sup> Stichworte wie ganzheitliche, kunden- oder marktbezogene Aufgabebearbeitung, systematische Ausschöpfung des Kreativitätspotentials der Mitarbeiter, Erhöhung von Flexibilität und Adaptionfähigkeit u.ä. fallen in diesem Kontext. Konzepte wie das des „advanced information management (AIM)“ des

---

<sup>24</sup>Bressand, Albert (1991): Seven Global Integration Dilemmas. In: Conflict, Competition and ... Cooperation, Vol. II. Project Prometheus Perspectives no. 16 (March). Paris: Prometheus, 12-13.

<sup>25</sup>ibidem.

<sup>26</sup>Vgl. dazu auch Reichwald, Ralf, und Bellmann, Kurt (1991): Optimale Arbeitsteilung in Büroorganisationen. Der Einfluß neuer Informations- und Kommunikationstechniken - eine kosten-theoretische Betrachtung. In: Zeitschrift für Betriebswirtschaft 61 (5/6), 621-639.

Boeing Advanced Technology Center<sup>27</sup> oder das des „Assistenz-Computer“-Projektes der Gesellschaft für Mathematik und Datenverarbeitung<sup>28</sup> sind hier zukunftsweisend. In einem Kommentar zu diesen technisch gegebenen Möglichkeiten wurde kürzlich darauf hingewiesen, daß „no one yet appears to have considered what data to feed into this magnificent beast. ... The technology is great; the information is a big weak link. A bit like buying a Rolls Royce but economising by buying the model without the fuel tank.“<sup>29</sup> Und genau in dieser Feststellung liegt eine große Herausforderung für die Informationspraxis. - Auf die Entwicklung zum „Büro in der Aktentasche“, das externe Mitarbeitern per Satelliten vernetzt oder, bei zunehmender Raumknappheit, interne Mitarbeiter an dem je nach Abwesenheit von Kollegen anderen, gerade freien Arbeitsplätzen in die interne Kommunikationsstrukturen einbindet, sei nur am Rande hingewiesen.

In diesem Kontext sei auch kurz auf einige andere Tendenzen aufmerksam gemacht. Computergestützte Gruppenarbeit (Computer supported cooperative work - CSCW) kennzeichnet einen neuen Trend, die Kommunikation in Arbeitsteams zu verbessern, determinierte Abläufe an isolierten Arbeitsplätzen zu reduzieren und flexibel organisierte Gruppenarbeit zu forcieren.<sup>30</sup> - Heimarbeit oder ‚Telework‘ als Dezentralisierungsstrategie hat sich trotz entsprechender Prognosen bis heute nicht durchgesetzt.<sup>31</sup> Ob der Strategie der „elektronischen Arbeit in Satellitenbüros“ ein größerer Erfolg beschieden sein wird, muß sich noch herausstellen; hier besteht zumindest die Chance, daß durch soziale Kontakte der entscheidende Nachteil der Heimarbeit, die häusliche Isolation, vermieden wird.<sup>32</sup> - Auf der Ebene regionaler, nationaler und multinationaler Akteure ist von Bedeutung, daß die ursprünglich erhofften regionalen Dezentralisierungstendenzen der ökonomischen Tätigkeiten dank moderner Kommunikationsnetze sich nicht realisiert haben. Vielmehr haben hier offensichtlich Agglomerationen mit einer guten Dienstleistungsinfrastruktur erhebliche Wettbewerbsvorteile. Informationsinfrastrukturen werden auch hier zu einem entscheidenden regionalen Wettbewerbsfaktor.<sup>33</sup>

#### 4.3.2 Zentralisation

Dieser Dezentralisierung steht eine daraus unmittelbar folgende, entgegengesetzte Tendenz gegenüber: Eine auf Vernetzung gestützte Re-Zentralisierung, um immer verteilte, individualisiertere Aktivitäten und Leistungsprozesse besser koordinieren und kontrollieren zu können; Unternehmenstaktiken sind

---

<sup>27</sup>„An AIM system accesses and integrates diverse information sources across heterogeneous computing systems, performs implied numerical and symbolic calculations for its users through a multi-modal interface, and presents the results in a way that is readily understood“. Hoard, James E. (1989): Natural Language Processing and Advanced Information Management. In: Telematics and Informatics 6 (3/4), 313-329.

<sup>28</sup>Hoschka, Peter und Wißkirchen, Peter (1990): Assistenz-Computer - Eine neue Generation von Unterstützungssystemen. In: Der GMD-Spiegel 20 (1), 20-25.

<sup>29</sup>O.V. (1991): Update your client / server concepts. In: Monitor 123 (May), 9.

<sup>30</sup>Vgl. dazu auch Mockler, R.J. und Dologite, D. G. (1991): Using Computer Software to Improve Group Decision-making. In: Long Range Planning 24 (4), 44-57.

<sup>31</sup>Huws, Ursula (1991): Telework: Projections. In: Futures (Jan/Febr), 19-31.

<sup>32</sup>Steinert, Harald (1989): Statt isolierter Heimarbeit lieber elektronische Arbeit in Satellitenbüros. In: Blick durch die Wirtschaft (31.03.).

<sup>33</sup>Siehe hierzu Hepworth, Mark (1991): Information cities in Europe 1992. In: Telecommunications Policy (June), 175-181; ders. (1990): Geography of the Information Economy. In: Intermedia 18 (1), 29-34.

schnellstmöglich, flexibel und mit hoher Adaption der jeweiligen Situation anzupassen. Für eine enge Vernetzung (sowohl im IT-Sinne als auch im übertragene Sinne) von Unternehmensstrategie, Unternehmensorganisation und Informationsmanagement stehen Konzepte wie „Computer Integrated Enterprise (CIE)“<sup>34</sup> oder „Computer-Integrated Business (CIB)“<sup>35</sup>. Von stand-alone Informationssystemen über bereichsintegrierte und bereichsübergreifende Informationssysteme geht die Entwicklung hin zu unternehmensweiten Informationssystemen, die in überbetriebliche Kommunikationsnetze eingebunden sind. „Plantwide Systems Integration“ (d.h. auf der Produktionsebene, z.B. flexible Fertigungssysteme) ist in den USA bereits ein Milliardengeschäft, und die konkrete Entwicklung hin zu höheren Integrationsstufen ist angelaufen.<sup>36</sup> Ott geht davon aus, daß geläufige Client/Server-Konzepte nur ein erster Schritt in diese Richtung sind und daß erst der Übergang zu objektorientierten Datenbanken es erlauben wird, „die Gesamtheit der im Unternehmen vorhandenen Informationen neu aufzubereiten und in bislang nicht gekannter Tiefe für die Entscheidungsvorbereitung bereitzustellen“; darauf aufbauend wird dann die Entwicklung von entscheidungsunterstützenden Systemen den Schritt frei machen „zu einer neuen Qualität der Entscheidungsprozesse.“<sup>37</sup>

Was sind die Konsequenzen dieser Entwicklung? Eine enge Kopplung von Unternehmensstrategie und Informationsstrategie wird überlebenswichtig, unternehmensübergreifende Kommunikationssysteme ermöglichen - und erzwingen z.T. - eine stärkere horizontale und vertikale Kooperation, kleinere Unternehmen verlieren tendenziell den größenbedingten Vorteil der schnellen, auf unterschiedliche Situationen flexibel reagierenden Entscheidungsfindung, sie werden durch die Einbindung in die Informationssysteme von Großkunden auch in ihrer Selbständigkeit bedroht. Auf Kooperation fußende Gegenstrategien des Mittelstandes sind unumgänglich, Informationsmanagement wird noch bedeutsamer.

## 4.4 Prozeßsteuerung

### 4.4.1 Informationsmanagement

„There is no relationship between expenses for computers and business profitability. ... Make management more productive, by electronic means, if you know where, when and how. Automate success, not failure.“<sup>38</sup> „Strategic IT investment

---

<sup>34</sup>Frank, Ulrich und Klein, Stefan (1991): Computer Integrated Enterprise oder: Zur Bedeutung anwendungsnaher Referenzmodelle. GMD-Arbeitspapier, St. Augustin.

<sup>35</sup>Bullinger, Hans-Jörg (1989): Computer Integrated Business: Informationsmanagement sichert Wettbewerbsvorteile für das Unternehmen der Zukunft. In: Proceedings der 12. Europäischen Congressmesse für Technische Kommunikation in Hamburg, 31.01.-03.02.1989, Plenarveranstaltung. Velbert: Online GmbH, P-4-1-P-4-27.

<sup>36</sup>O.V. (1990): Wachstum bei Systemintegrationsdiensten in den Vereinigten Staaten. In: Blick durch die Wirtschaft (19.06.).

<sup>37</sup>Ott, Werner (1989): Der Schritt zu einer neuen Qualität der unternehmerischen Entscheidungsprozesse - Warum es bei elektronischen Informationssystemen nicht bei Insellösungen bleiben kann. In: Blick durch die Wirtschaft (09.11.).

<sup>38</sup>Strassmann, Paul A. (1990): The Business Value of Computers. New Canaan, Con.: The Information Economics Press, xvii.



[„defined as IT investment made to gain competitive advantage through increased sales“] is generally made in ‚blind faith‘ that real returns will occur.“<sup>39</sup> „Die Einkommen der Informations-Systeme-Manager sind inzwischen sogar rückläufig. ... Die Entlassungen werden im wesentlichen mit der unzureichenden Mitwirkung bei der Einbringung der Informationstechnik in die strategische Zielgestaltung begründet.“<sup>40</sup> Diese wenigen, auf empirischen Erhebungen beruhenden Feststellungen unterstreichen nachdrücklich die dringende Notwendigkeit eines wohlverstandenen Informationsmanagements.

Wurde lange Zeit Informationsmanagement im wesentlichen als das Management von Informationstechniken verstanden, so setzt sich - aufgrund schmerzlicher und teurer Erfahrungen - mehr und mehr die Erkenntnis durch, daß Informationstechnik und Informationssysteme nur Hilfsmittel für die Gestaltung und Abwicklung von Informationsprozessen sind. Informationen als ‚Inhalte‘, als Ressourcen in den verschiedenen, oben in Abschnitt 2 skizzierten Ausprägungen, müssen im Mittelpunkt des Informationsmanagements stehen. Die im lokalen wie auch im globalen Wettbewerb sich hochschaukelnden Informationsanforderungen und -bedürfnisse verlangen immer dringender nach einem ganzheitlichen, zielgerichteten Informationsmanagement in diesem Sinne; dies betrifft sowohl die Informationsbedürfnisse des einzelnen - sei es als Mitarbeiter an der Verkaufsfrent oder als mündiger Bürger zur Wahrnehmung seiner demokratischen Rechte (und Pflichten) - als auch die von Institutionen und Administrationen - sei es für Vorstandsmitglieder für strategische Entscheidungen oder für Minister für konsensfähige ‚Policies‘. Entwicklung von Visionen und Leitlinien für die informationsstrategische Vorgehensweise, Planung, Organisation/Durchführung und Kontrolle sowohl auf der strategischen als auch auf der administrativen und operationalen Ebene sind wesentliche Elemente eines solchen Informationsmanagements. Ein prozeßorientierter, logistischer Ansatz wird hierbei den dynamischen, vernetzten Aspekten aller Informationsvorgänge am ehesten gerecht. Wie die obige Darstellung gezeigt hat, besteht auf absehbarer Zeit kein Mangel an konkreten Einsatzfeldern. Es ist vielmehr davon auszugehen, daß sich diese dramatisch ausdehnen und alle Bereiche moderner Gesellschaften auf allen Akteursebenen durchdringen werden.

#### 4.4.2 Sicherheit/Verfügbarkeit

Eine elektronisch vernetzte Gemeinschaft ist eine hochgradig gefährdete Gemeinschaft. Ihre Funktions- und tendenziell sogar ihre Überlebensfähigkeit wird immer enger mit der Sicherheit, unbedingten Verfügbarkeit und nationalen und globalen Regulierung ihrer Vernetzungsstrukturen verknüpft. Die Beherrschung dieser Prozeßrisiken wird zu einem zentralen Objektbereich des Informationsmanagements. Einige Beispiele sollen dies kurz erläutern: Relativ trivial ist noch, daß kürzlich die harsche Kritik des IBM-Vorsitzenden an seinen Top-Managern und der Belegschaft über das interne E-Mail-Netz nicht nur bestimmten Abteilungsleitern, sondern aus ungeklärten Gründen der gesamten Belegschaft und letztendlich der Öffentlichkeit zugänglich wurde. Computerviren

---

<sup>39</sup>Weill, Peter (1990): Strategic investment in information technology: an empirical study. In: Information age 12(3), 141, 142.

<sup>40</sup>Pechanel, Frank D. (1991): Konfliktherd Datenverarbeitung. In: Blick durch die Wirtschaft 34 (158), 1.

haben sich inzwischen zu einer ernstzunehmenden ökonomischen Gefahr entwickelt; im internationalen Kreditkartengeschäft gibt es praktisch keinen Datenschutz, der z.B. verhindert, daß Persönlichkeitsprofile erstellt und vermarktet werden; drei fehlerhafte Binärcodes in einer geringfügigen Veränderung der Vermittlungssoftware führten am 26. Juni dieses Jahres zum Zusammenbruch des Telefonsystems in Washington, D.C., und umliegenden Staaten, später auch in Pittsburgh und an der Westküste.<sup>41</sup> Gesamtwirtschaftliche Relevanz erhalten diese Risiken, wenn eine weitere Verbreitung des nur dank moderner Informationssysteme operablen Prinzips des ‚Just in Time‘ dazu führt, daß Autobahnen zu einem immobilien Warenhaus werden oder wenn Störungen des elektronischen Geldumlaufs - der Austausch von Informationen *über* Geld - zum Zusammenbruch von Banken und Finanzmärkten führen würden.

Wichtige Kriterien in diesem Kontext sind u.a. (1) Vertraulichkeit, Abhörsicherheit und Anonymität (z.B. Persönlichkeitsschutz, Anonymität von Nutzungs- und Vermittlungsdaten), (2) Kommunikationsintegrität (Partnergewißheit, Zugriffskontrolle, Sende- und Empfangsbeweis u.a.), (3) Verfügbarkeit (Funktionalität, Betriebssicherheit und -kontinuität) sowie (4) Verbindlichkeit (Aussteller-Authentizität, Rechtssicherheit etc.).<sup>42</sup> Zu ergänzen ist, zumindest im nationalen und globalen Kontext, (5) Verletzlichkeit der Gesellschaft als Ganzes.

Hier sind nicht nur individuelles und organisationsbezogenes Risikobewußtsein<sup>43</sup> und Informationsmanagement gefordert. Vielmehr besteht die Gefahr, daß ohne globale Regulierungs- und Steuerungsmechanismen sich die offene, vernetzte Gesellschaft durch den notwendigen Schutz gegen Mißbrauch und nicht mehr beherrschbare Schadenspotentiale aufgrund entsprechender Überwachungs- und Kontrollmaßnahmen selbst *ad absurdum* führt.

## 5. Ausblick

Die skizzierten Entwicklungspotentiale und Risiken der vernetzten Welt stellen eine profunde Herausforderung an die Informationspraxis dar. Hierbei sind alle Akteursebenen angesprochen. Information ist nicht nur eine grundlegende Ressource zur Rationalisierung von Leistungserstellungs- und Managemententscheidungsprozessen; effektive Informationsinfrastrukturen und -verarbeitungsprozesse stellen nicht nur einen entscheidenden Wettbewerbsfaktor im globalen Wettbewerb zwischen Unternehmen, Regionen und Nationen dar; mit der sogenannten Informatisierung gehen auch gesellschaftliche und kulturelle Veränderungen einher, die langfristig vielleicht noch viel bedeutsamer in ihren

---

<sup>41</sup>1984 war aufgrund von Systemfehlern zeitweilig das ‚long distance‘-Netz der USA zusammengebrochen.

<sup>42</sup>Nach Klein, Stefan (1991): Informationssicherheit bei der Kommunikation von Versicherungen mit Dritten. In: Lippold, Heiko u.a. (Hrsg.): Sicherheit in Informationssystemen. Braunschweig: Vieweg, 175.

<sup>43</sup>„Wer mit Informatik-Professoren, Vertretern von Hard- und Software-Häusern, Nachrichtentechnikern und Netzspezialisten zu tun hat, ... macht immer wieder dieselbe Erfahrung: Alle preisen die Vorzüge ihrer Technik wärmstens an; über die konkreten Risiken schweigen sie meist; selbst auf Befragen ist ihnen dazu nicht viel zu entlocken; das Problembewußtsein fehlt weitgehend.“ Leuze, Ruth (1988) : Das Labyrinth der Vorschriften kennt nur noch der Experte. In: Heussermann, Titus und Krautter, Horst (Hrsg.): Der Griff nach dem Menschen. Stuttgart, Radius; zitiert nach Süddeutsche Zeitung (12./13.11.), 10. - Siehe auch Roßnagel, A. u.a. (1991): Die Verletzlichkeit der Informationsgesellschaft (2. Auflage). Opladen: Westdeutscher Verlag.

Auswirkungen sind. Die global vernetzte Weltgemeinschaft hat mit der Ressource Information ein potentes Hilfsmittel zur Hand, um globale Ungleichgewichte zwischen Ost und West, Nord und Süd besser in den Griff zu bekommen, sei es im Bereich der Wirtschaft, des Umweltschutzes, der Menschenrechte, des Analphabetismus, der kulturellen Vermittlung. Diesen Möglichkeiten, und auch den damit verbundenen globalen Risiken, müssen sich Informationspraxis und Informationspolitik, unterstützt von einer breit verstandenen, fundierten Informationswissenschaft, stellen.

©

Dr. Karl A. Stroetmann  
GMD-Forschungsstelle für  
Informationswirtschaft  
Schönhauserstr. 64  
D - W 5000 Köln 51  
Fax: x49 - (0)221 - 3 76 73 - 83  
Tel.: x49 - (0)221 - 3 76 73 (0) 14  
e-mail: kas@svinfow.koeln.gmd.de

## VERZEICHNIS DER AUTOREN

Anan'eva, Tatjana A.;	Moskau	259
Babst, Peter;	Halle	410
Bertha, Eva;	Graz	157
Bonitz, Manfred ;	Dresden	218
Bruckner, Eberhard;	Berlin	218
Burkert, Herbert;	Köln	447
Dosch, Walter;	Augsburg	71
Englmeier, Kurt;	München	307
Frerk, Gisela;	Köln	358
Fuhr, Norbert;	Dortmund	94
Goebel, Jürgen W.;	Köln	447
Harms, Ilse;	Saarbrücken	237
Herget, Josef;	Konstanz	52
Hillebrand, Volker;	Roßlau	341
Himberger, Andreas;	St.Gallen	433
Hoffmann, Thorsten;	Darmstadt	94
Holbein, Ralph;	Ulm	81
Huber, Wolfgang;	Luxembourg	11
Jochum, Barbara;	Saarbrücke	472
Killenberg, Harald;	Ilmenau	228
König, Eckehard;	Bonn	274
Koleva, R.;	Sofia	305
Krüger, Frank;	Saarbrücken	104
Kruske, Hans;	Magdeburg	351
Kuhlen, Rainer;	Konstanz	26
Kulawiak, Walter;	Frankfurt	207
Kurby, Stefan;	Berlin	396
Mater, Erich;	Berlin	70
Michel, Engelbert;	Konstanz	182
Michelson, Martin;	Frankfurt/Main	286
Münch, Eberhard;	Berlin	274
Oßwald, Achim;	Köln	358
Poetzsch, Eleonore;	Berlin	387

Reiss, Peter;	Darmstadt	167
Roppel, Stephan;	Regensburg	115
Schael, Fritz;	Wolfsburg	333
Scharnhorst, Andrea;	Berlin	218
Scheuermann, Fritz ;	Saarbrücken	298
Schmidt, Sabine,	Graz	157
Schröder, Thomas A.;	Düsseldorf	197
Schulz, Angelika;	Saarbrücken	237
Schwandt, Joachim;	Berlin	396
Schwantner, Michael;	Karlsruhe	320
Schwarz, Erich J.;	Graz	157
Schwuchow, Werner;	Köln	450
Sprenger, Astrid;	Ilmenau	475
Stock, Wolfgang G.;	München	376
Stroetmann, Karl A.;	Köln	450
Teufel, Bernd;	Ulm	81
Thulke, Michael;	Berlin	274
Tsankova, R.,	Sofia	305
Tzeras, Kostas;	Darmstadt	167
Wagner-Döbler, Roland;	München	40
Wawra, Steffen;	Berlin	247
Weitzendorf, Thomas;	Graz	256
Wolff, Christian;	Regensburg	115
Yetim, Fahri;	Konstanz	145
Zbornik, Stefan;	St.Gallen	433
Zimmermann, Harald;	Saarbrücken	125

## Stichwortverzeichnis

<b>Agrardatenbanken</b> neue Bundesländer CD-ROM	274
<b>AIR/PHYS</b> Automatische Indexierung Praxis	320
<b>Akzeptanzhürden</b> Informationsmanagement	358
Applikationen, Hypertext; Feldversuche Design-Vorschläge	157
<b>Aufgabenprofil</b> Informationseinrichtung Unternehmen	
Schwermaschinenbau	351
Auftragsabwicklung Informationsvermittlung	475
<b>Ausbildung</b> , informationswissenschaftliche Ilmenau	228
<b>automatische</b> Indexierung Praxis AIR/PHYS	320
automatische Indexierung; Wörterbuch; Entwicklungsaufwand	167
automatische Trunkierung; textbezogene Informationsbanken	125
automatischer Wissenserwerb; traditionelles Indexieren	70
<b>Benutzeroberflächen</b> , graphische; Methoden Werkzeuge	
Prototyping Entwicklung	472
<b>Biotechnology</b> Information Service	387
<b>Bürodokument</b> ; Beschreibung; kontextsensitive Merkmale	81
<b>CD-ROM</b> Agrardatenbanken neue Bundesländer	274
<b>Design-Vorschläge</b> Hypertext-Applikationen; Feldversuche	157
<b>Dialogkomponente</b> , hypermedia-gestützt; objektorientierte	
Sicht; Expertensystem- Wissensbasen	145
<b>elektronische</b> Informationsdienste Entwicklungstrends	
Perspektiven europäischer Markt	450
elektronische Märkte Entwicklung Bedeutung Einsatz	
Informations-, Kommunikationstechnologie	433
Entwicklungstrends, Perspektiven; europäischer Markt	
elektronische Informationsdienste	450
Entwicklungsaufwand Leistungsfähigkeit Wörterbuch	167
<b>epidemische</b> Prozesse Informationswissenschaft; Theorie	40
<b>Erfolgsfaktorenforschung</b> Informationsvermittlung	52
erzeugnisorientierter Ansatz; Wissensrepräsentation	
Software	207
<b>europäisches</b> Informationsrecht Entwicklungen	447
europäischer Markt elektronische Informationsdienste	
Entwicklungstrends Perspektiven	450

<b>experimentelles</b> Information-Retrieval-System; WIDAB- Konzeption Realisierung	307
<b>Expertensystem</b> -Wissensbasen, objektorientierte Sicht; hypermedia-gestützte Dialogkomponente	145
<b>Fachinformationssystem</b> Informationsmanagement	341
<b>Fact Text Retrieval</b> , Probabilistic Prototyp	94
<b>graphische</b> Benutzeroberflächen Methoden Werkzeuge Proto- typing Entwicklung	472
graphische Recherche intelligente Hilfe multimodales Werkstoffinformationssystem WING	115
<b>Grundlagen</b> , wissenschaftliche; Informationstechnologie	259
<b>Hyper-Advokat</b> - Konzeption Realisierung juristisches Hypertextsystem	182
<b>hypermedia-gestützte</b> Dialogkomponente; Expertensystem- Wissensbasen; objektorientierte Sicht	145
Hypermedia-Informationssystem	197
<b>Hypertext</b> ; Information Retrieval; Navigationsstrategien	104
Hypertext-Applikationen Feldversuche; Design-Vorschläge	157
Hypertextsystem, juristisches; Konzeption Realisierung Hyper-Advokat	182
<b>Indexieren</b> , traditionell automatischer Wissenserwerb	70
Indexierung, automatische; Praxis; AIR/PHYS	320
Indexierung, automatische; Wörterbuch Entwicklungsaufwand	167
<b>Information</b> Retrieval Hypertext; Navigationsstrategien	104
Information-Retrieval-System, experimentelles; WIDAB- Konzeption Realisierung	307
informationelle Mehrwerte Theorie	26
Informationsdienste, elektronische; Entwicklungstrends Perspektiven europäischer Markt	450
Information Service, Biotechnologie	387
Informations-, Kommunikationstechnologien Bedeutung Einsatz elektronische Märkte	433
<b>Informationsbanken</b> , textbezogene; automatische Trunkierung	125
Informationsbedürfnisse Forschung Wirtschaft Informations- vermittlung -consulting	396

Informationsconsulting -vermittlung Informationsbedürfnisse	
Forschung Wirtschaft	396
Informationseinrichtung Aufgabenprofil Unternehmen	
Schwermaschinenbau	351
Informationsgesellschaft postmoderne Theorie	247
<b>Informationsmanagement</b> , Akzeptanzhürden	358
Informationsmanagement Fachinformationssystem	341
Informationsmanagement, moderne Technik Realität	333
Informationsmarkt, EG-Politik	11
<b>Informationsrecht</b> , europäisches; Entwicklungen	447
Informationssystem Ausbildungs-Software	305
Informationssystem, Hypermedia-	197
Informationstechniken Richterarbeitsplatz	298
Informationstechnologie wissenschaftliche Grundlagen	259
Informationsvermittlung, Auftragsabwicklung	475
Informationsvermittlung; Erfolgsfaktorenforschung	52
Informationsvermittlung Informationsconsulting;Komponenten	396
<b>Informationswissenschaft</b> ; Theorien epidemischer Prozesse	40
Informationswissenschaft Saarbrücken	237
informationswissenschaftliche Ausbildung Ilmenau	228
<b>informetrische</b> F&E-Indikatoren; Wirtschaftsdaten; Online-	
Unternehmensbild	376
<b>intelligente</b> Hilfe,graphische Recherche; multimodales	
Werkstoffinformationssystem WING	115
<b>juristisches</b> Hypertextsystem Konzeption Realisierung	
Hyper-Advokat	182
<b>Kommunikationstechnologien</b> , Informations-; Entwicklung	
Bedeutung Einsatz elektronische Märkte	433
kontextsensitive Merkmale; Beschreibung Bürodokumente	81
<b>Leistungsfähigkeit</b> Entwicklungsaufwand Wörterbuch; auto-	
matische Indexierung	167
<b>Management</b> , Informations-; moderne Technik Realität	333
Managementinformationen statistische Untersuchungen	
Patentinformationsdienste	410
<b>Merkmale</b> ,kontextsensitive; Beschreibung Bürodokumente	81
Methoden Werkzeuge Prototyping Entwicklung; graphische	
Benutzeroberfläche	472



<b>moderne Technik Realität Informationsmanagement</b>	333
<b>multimodales Werkstoffinformationssystem WING</b>	
graphische Recherche intelligente Hilfe	115
<b>Navigationsstrategien Information Retrieval Hypertext</b>	104
<b>objektorientierte Sicht Expertensystem-Wissensbasen</b>	
hypermedia-gestützte Dialogkomponente	145
<b>Online-Unternehmensbild Wirtschaftsdaten informetrische</b>	
F&E-Indikatoren	376
<b>optisches Speichersystem Wirtschaftspresseudokumentation</b>	286
<b>Paperwork Reduction Act</b>	256
<b>Patentinformationsdienst Managementinformationen</b>	
statistische Untersuchungen	410
postmoderne Theorie Informationsgesellschaft	247
<b>Praxis automatische Indexierung ; AIR/PHYS</b>	320
<b>Probabilistic Fact Text Retrieval Prototype</b>	94
Programmierwissen Repräsentation	71
<b>Prototype Probabilistic Fact Retrieval Text Retrieval</b>	94
Prototyping Entwicklung, Methoden Werkzeuge; graphi-	
sche Benutzeroberflächen	472
Publikationsstrukturen Ländervergleich	218
<b>Realisierung Information-Retrieval-System, experimentel-</b>	
les; WIDAB-Konzeption	307
<b>Recherche, graphische; intelligente Hilfe multimodales</b>	
Werkstoffinformatuionssystem WING	115
<b>Recht-, europäisches Informations-; Entwicklungen</b>	447
Repräsentation Programmierwissen	71
<b>Retrieval Text Retrieval Fact Prototype Probabilistic</b>	94
Retrieval-System, experimentelles Informations-; WIDAB-	
Konzeption, Realisierung	307
<b>Software, Wissensrepräsentation; erzeugnisorientierter</b>	
Ansatz	207
Speichersystem, optisches; Wirtschaftspresseudokumentaion	286
statistische Untersuchungen Patentinformationsdienste	
Managementinformationen	410
<b>textbezogene Informationsbanken automatische Trunkierung</b>	125
<b>Theorie, postmodern; Informationsgesellschaft</b>	247

Theorien epidemischer Prozesse Informationswissenschaft	40
<b>traditionelles</b> Indexieren automatischer Wissenserwerb	70
<b>Trunkierung</b> , automatische;textbezogene Informationsbanken	125
<b>Unternehmen</b> Schwermaschinenbau Informationseinrichtung Aufgabenprofil	351
<b>Werkstoffinformationssystem</b> WING, multimodales; graphische Recherche intelligente Hilfe	115
<b>Werkzeuge</b> , Methoden; Prototyping Entwicklung; graphische Benutzeroberflächen	472
<b>WIDAB</b> -Konzeption Realisierung; experimentelles Information-Retrieval-System	307
<b>WING</b> , multimodales Werkstoffinformationssystem graphische Recherche intelligente Hilfe	115
<b>Wirtschaft</b> , Informationsbedürfnisse; Informationsvermittlung	396
Wirtschaftsdaten informetrische F&E Indikatoren Online- Unternehmensbild	376
Wirtschaftspresseudokumentation; optisches Speichersystem	286
<b>Wissensbasen</b> , Expertensystem-;objektorientierte Sicht; hypermedia-gestützte Dialogkomponente	145
<b>wissenschaftliche</b> Grundlagen Informationstechnologie Management	259
<b>Wissenserwerb</b> , automatischer; traditionelles Indexieren	70
Wissensrepräsentation Software ergebnisorientierter Ansatz	207
<b>Wörterbuch</b> Entwicklungsaufwand Leistungsfähigkeit; automatisches Indexieren	167