

Gerhard Knorz / Rainer Kuhlen (Hg.)

# Informationskompetenz – Basiskompetenz in der Informationsgesellschaft

Proceedings des  
7. Internationalen Symposiums für  
Informationswissenschaft (ISI 2000)

UVK Universitätsverlag Konstanz GmbH

# **Inhalt**

ISI 2000	VII
Vorwort	IX
 <i>Thomas Mandl, Christa Womser-Hacker</i>	
Ein adaptives Information-Retrieval-Modell für Digitale Bibliotheken	1
 <i>PD Dr. Ilse Harms, Werner Schweibenz</i>	
Usability Engineering Methods for the Web	
Results From a Usability Study	17
 <i>Christian Wolff</i>	
Effektivität von Recherchen im WWW	
Vergleichende Evaluierung von Such- und Metasuchmaschinen	31
 <i>Rainer Hammwöhner</i>	
TransRouter revisited - Decision support in the routing of translation projects	49
 <i>Gerhard Rahmstorf</i>	
Wortmodell und Begriffssprache als Basis des semantischen Retrievals	71
 <i>Christian Schlögl</i>	
Informationskompetenz am Beispiel einer szientometrischen	
Untersuchung zum Informationsmanagement	89
 <i>Otto Krickl, Elisabeth Milchrahm</i>	
Integrativer Ansatz zur Wissensbewertung	113
 <i>Gabriela Mußler, Harald Reiterer, Thomas M. Mann</i>	
INSYDER - Information Retrieval Aspects of a Business Intelligence System	127
 <i>C. Goller, J. Löning, T. Will, W. Wolff</i>	
Automatic Document Classification	
A thorough Evaluation of various Methods	145
 <i>Gerhard Heyer, Uwe Quasthoff, Christian Wolff</i>	
Aiding Web Searches by Statistical Classification Tools	163

<i>Matthias N.O. Müller</i> Die virtuelle Fachbibliothek Sozialwissenschaften	179
<i>Benno Homann</i> Das Dynamische Modell der Informationskompetenz (DYMIK) als Grundlage für bibliothekarische Schulungen	195
<i>Gerhard Reichmann</i> Leistungsvergleiche zwischen wissenschaftlichen Bibliotheken	207
<i>Willi Bredemeier, Wolfgang G. Stock</i> Informationskompetenz europäischer Volkswirtschaften	227
<i>Hermann Rösch</i> Internetportal, Unternehmensportal, Wissenschaftsportal Typologie und Funktionalität der wichtigsten Portalkonzeptionen	245
<i>Marc Rittberger, Wolfgang Semar</i> Regionale Elektronische Zeitungen: Qualitätskriterien und Evaluierung	265
<i>Stephan Werner</i> Der Autor im digitalen Medium - ein notwendiges Konstrukt?	285
<i>Dr. Jaroslav Šušol</i> Access to information in electronic age - situation in Slovakia	303



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. IX – V

## Vorwort

ISI 2000 als die Fachkonferenz der Informationswissenschaft im deutschsprachigen Bereich, veranstaltet vom Hochschulverband für Informationswissenschaft (HI) und organisiert von der Fachhochschule Darmstadt, konzentriert sich auf den professionellen wissenschaftlichen Bereich. Das 7. Internationale Symposium für Informationswissenschaft - ISI 2000 - hat zum Leitthema "Informationskompetenz - Basiskompetenz in der Informationsgesellschaft". Auch oder gerade auf den Endnutzermärkten des Internet ist Informationskompetenz notwendig, um wirklichen Nutzen aus den globalen Informationsdiensten ziehen zu können. Informationskompetenz ist zum einen Können der Informationsprofessionellen - das sind die klassischen Informationsspezialisten, wie sie auf akademischem Niveau von der Informationswissenschaft und angrenzenden Disziplinen ausgebildet werden, aber auch die sogenannten Postprofessionellen, die aus allen Wirtschaftsbereichen in die Informationsmärkte dringen, um Informationsdienste zu erstellen und anzubieten. Informationskompetenz sollte zum andern aber auch Basiskompetenz für jedermann in der Informationsgesellschaft sein, um bewusst und erfolgreich sich in den elektronischen Informationsräumen bewegen zu können. Aus den Artikeln des Bandes wird deutlich, welche Bandbreite Informationskompetenz heute ausmacht. Die Herausforderung ist interdisziplinär. Jenseits einer bloß technischen Ausrichtung ist z.B. sozio-ökonomisches, kognitiv-psychologisches, linguistisches, designerisch-ästhetisches Wissen verlangt, um Information erfolgreich erarbeiten zu können. Das Internet und das WWW sind Ausgangs- und Bezugspunkt vieler Arbeiten zu Informations- und Wissensmanagement, Informationswirtschaft, Verlags- und Bibliothekswesen, Wissensrepräsentation, Information Retrieval, Data/Text Mining sowie Hypertext/Multimedia.

Die Herausgeber danken den Gutachtern, die die Bewertungsleistung schnell und weitgehend vollständig erbracht haben, so dass alle Artikel von mindestens drei, die meisten sogar von vier Fachleuten begutachtet wurden. Der Prozess wurde dadurch erleichtert, dass das gesamte Reviewing elektronisch abgewickelt werden konnte. Dr. Bernard Bekavac, Konstanz, sei für den technischen Support des Reviewing gedankt. Herr Wolfgang Semar, Konstanz, hat sich um die Koordination zwischen dem Verlag und der Projektgruppe an der FH Darmstadt gekümmert. Hier gilt der Dank besonders Herrn Alex Diegelmann, der kurzfristig die Arbeit und die Verantwortung dafür übernommen hatte, dass auch diesmal wieder die Proceedings zur Tagung vorliegen. Der besondere Dank geht natürlich an die Autoren dieses Bandes, die es geschafft haben, die verschiedenen Hürden, von der ersten Idee bis zur Drucklegung in diesem Band, zu überspringen. Möge es ihre wissenschaftliche oder professionelle Laufbahn befördern. Den Teilnehmern an der Konferenz wünschen wir viele Anregungen, zunächst durch die Vorträge und Diskussionen während der Tagung, dann aber auch durch das spätere intensive Studium der Beiträge.

Die Herausgeber, Prof. Gerhard Knorz, Darmstadt, und Prof. Rainer Kuhlen, Konstanz/Berlin



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 1 – 16

## **Ein adaptives Information-Retrieval-Modell für Digitale Bibliotheken**

**Thomas Mandl — Christa Womser-Hacker**

Informationswissenschaft — Universität Hildesheim

Marienburger Platz 22 — 31141 Hildesheim

mandl@rz.uni-hildesheim.de

womser@cl.uni-hildesheim.de

### **Zusammenfassung**

In diesem Beitrag wird ein adaptives Information-Retrieval-Modell im Kontext Digitaler Bibliotheken vorgestellt, das während des Informationsprozesses lernt, Beziehungen zwischen Repräsentationsformen und Informations-Objekten herzustellen. Die aus großen Evaluierungsstudien abgeleiteten Erkenntnisse, dass Relevance Feedback und Polyrepräsentation echte Mehrwerte darstellen und zu besserer Retrievalqualität führen, dienen dabei als Grundlage. Um eine stärkere Adaption zu leisten, wird das Modell um Strategien erweitert, die Benutzer-, Dokument- und Anfrageeigenschaften einbeziehen. Im Rahmen Digitaler Bibliotheken spielen insbesondere Transformationen unterschiedlich erschlossener Objekte eine wichtige Rolle zur Kompensation semantischer Heterogenität. Auch in diesem Rahmen kann das adaptive Information-Retrieval-Modell eingesetzt werden.

### **Abstract**

This paper presents an adaptive information retrieval system for digital libraries which learns to model the relationship between representation mechanisms and information objects during an information process. The evaluation results which form the foundation of the model show that relevance feedback and poly-representation provide value added services and lead to better retrieval quality. Features of users, documents, and queries are included to enhance the model and to guarantee a higher level of adaptation. In the context of digital libraries, transformations are an important strategy in order to resolve semantic heterogeneity. The adaptive information retrieval model can be applied to these transformations.



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

## **1. Herausforderungen für Digitale Bibliotheken**

Das Information Retrieval (IR) befindet sich derzeit in einer Umbruchsituation, da aufgrund innovativer Anwendungsbereiche (Digitale Bibliotheken, Electronic Commerce, Suche im WWW etc.) modifizierte Anforderungen entstanden sind. Nicht nur die Art der informationellen Objekte ist heterogen, sondern auch die Domänen, die Benutzer, die Infrastrukturen etc. Neue Technologien wie Multimedia und das Internet beeinflussen den Informationsprozess in starkem Maße. Aufgrund dieser Komplexität und Heterogenität der Kontexte ist die Anwendung einer generellen Lösung nicht mehr möglich. In der adäquaten Selektion und Adaption von IR-Verfahren und ihrer Kombinationen sowie in den hinzukommenden Mehrwerten liegt der Schlüssel zum Erfolg.

Der Bereich der Digitalen Bibliotheken (cf. Fuhr 1999) bezieht neben einer Vielfalt unterschiedlicher Objekttypen (traditionelle Publikationen, multimediale Objekte) und neuen Möglichkeiten des Elektronischen Publizierens eine sehr heterogene Klientel ein, deren Know-how sehr unterschiedlich geprägt sein kann. Oft handelt es sich nicht um spezialisierte, mit informationeller Kompetenz ausgestattete Benutzer. Die Systeme müssen mit entsprechender Intelligenz ausgestattet werden, um diese Situation bewältigen zu können. Die Erkenntnisse aus langjähriger Erfahrung im Bereich der Evaluierung von IR-Systemen haben gezeigt, dass Benutzerverhalten und Kontextfaktoren den Informationsprozess weitaus stärker beeinflussen als dies so manche technische Parameter tun (cf. Tague-Sutcliffe & Blustein 1995). Dennoch werden diese Einflüsse nur sehr selten berücksichtigt, was in erster Linie daran liegt, dass sich derartige Kriterien nur schwierig fassen und in operationalisiertes Systemverhalten umsetzen lassen. Damit steht man vor der Problematik der Generalisierbarkeit von Aussagen und des Treffens der richtigen Auswahl.

Ein weiterer wichtiger Punkt kommt hinzu. Die technischen Möglichkeiten der Vernetzung lassen Digitale Bibliotheken auf die Informationsprodukte verschiedener Anbieter zugreifen, da die Erweiterung der Datengrundlage die Suche für den Benutzer erleichtert. Die in einer virtuellen Bibliothek enthaltenen Objekte sind heterogen und auch ihre inhaltliche Erschließung und Repräsentation ist in der Regel unterschiedlich (cf. Krause 1998). Zum einen liegen eventuell automatisch und manuell indexierte Dokumente gemeinsam in einer Digitalen Bibliothek. Auch mit verschiedenen Thesauri manuell indexierte Dokumente führen zu Problemen bei der semantischen Konsistenz des Vokabulars. So ist z. B. das Vorkommen eines speziellen Fachbegriffs in einem Korpus von Zeitungstexten anders zu bewerten als in einer Sammlung von Dokumenten zu dem Fachgebiet. Da aus Kostengründen diese semantischen Probleme größtenteils nicht durch intellektuelle Nachbearbeitung gelöst werden, sind automatische Verfahren für die Transformation zwischen verschiedenen Repräsentationen erforderlich.

Die Anforderungen Digitaler Bibliotheken bilden den Ausgangspunkt. In Kapitel 2 wird das MIMOR-Modell beschrieben und formalisiert. Das dritte Kapitel ist den

Erweiterungen gewidmet, Kapitel 4 und 5 führen Transformationen ein, zeigen deren Rolle in Digitalen Bibliotheken und stellen ein Integrationskonzept in das MIMOR-Modell vor. Im Fazit werden die Ergebnisse zusammengeführt.

## **2. Das MIMOR-Modell**

Das im vorliegenden Beitrag dargestellte MIMOR-Modell (Mehrfachindexierung zur dynamischen Methoden-Objekt-Relationierung; cf. Womser-Hacker 1997) verfolgt das Ziel, unterschiedliche Objekteigenschaften mit speziellen Benutzerkontexten zu kombinieren.

### **2.1 Ausgangslage**

Innerhalb von Evaluierungsstudien ist selten Raum, um auf der Basis realer Kollektionen des Kontexts das adäquateste Verfahren zu ermitteln. Meist besteht also nicht die Möglichkeit, die erzielten Ergebnisse mit isolierten Funktionen (z. B. mit der Wirksamkeit linguistischer Analysen) innerhalb eines Systems zu verknüpfen. Die Maßzahlen haben Gültigkeit für den gesamten Systemansatz, wobei Wechselwirkungen zwischen den Funktionen schwer zu kontrollieren sind. Dies gilt auch für die TREC<sup>1</sup>-Studie, die derzeit größte und aktuellste Evaluierungsinitiative. Hier werden in mehreren Runden Entwicklungsprototypen aber auch kommerzielle IR-Systeme auf ihre Effektivität hin untersucht. Die teilnehmenden Institutionen durchlaufen mit ihren Systemen zeitlich kontrolliert einen Test mit vorgegebenen Aufgaben (sog. Topics) und Testdokumenten. Am NIST erfolgt die Bewertung, deren Ergebnisse zusammen mit den Systempräsentationen auf einer jährlichen Konferenz vorgestellt und anschließend publiziert werden (cf. Harman 1995, 1996; Voorhees & Harman 1998).

Auch nach mehrjähriger Laufzeit zeigt sich, dass statistische Maßzahlen wie Recall und Precision aufgrund ihrer vieldimensionalen Abhängigkeiten ohne weiterführende qualitative Analysen oder Experimentwiederholungen nur pauschale Erkenntnisse wiedergeben. So zeigte sich z. B., dass die Unterschiedlichkeit der statistischen Retrievalmodelle nur geringen Einfluss auf die Effektivität der Ergebnisse hat. Die Konsequenz daraus ist die Überprüfung der Güte der zusätzlichen Komponenten und Verfahrensweisen, die beim Informationsprozess eine Rolle spielen.

Das MIMOR-Modell (cf. Womser-Hacker 1997) baut v. a. auf einigen Ergebnissen der Serie der TREC-Konferenzen auf (cf. Voorhees & Harman 1998). Zwei der wichtigsten Ergebnisse von TREC haben MIMOR besonders beeinflusst:

Relevance Feedback durch den Benutzer erwies sich als eine sehr erfolgreiche Methode, um die Qualität von Retrieval-Ergebnissen zu verbessern. Die inhärente Interaktivität des Retrieval-Prozesses führt zur Erarbeitung von Wissen über den

---

<sup>1</sup> TREC (Text Retrieval Conference) ist beim National Institute of Standards and Technology (NIST) angesiedelt und hat nunmehr eine fast zehnjährige Laufzeit (cf. Harman & Voorhees 1998).

Benutzer, das das System z. B. für die Optimierung der Anfrage einsetzt (cf. Belkin 1993).

Viele Systeme erzielen Ergebnisse ähnlicher Qualität, d. h. vergleichbare Precision-Werte. Jedoch sind die Ergebnisse sehr verschieden, d. h. die Systeme führen zu unterschiedlichen Ergebnisdokumenten. Die Schnittmenge aus den Treffern verschiedener Systeme ist meist relativ klein. Einige Forscher setzen daher Fusionsverfahren ein. Sie kombinieren dabei die Ergebnisse verschiedener Ansätze mit dem Ziel, in dem Gesamtergebnis mehr Treffer und damit eine höhere Qualität zu erreichen (cf. Fox & Shaw 1994). Die Schwierigkeit bei Fusionsverfahren besteht in der Kombination der Ergebnisse.

MIMOR greift auf beide Erkenntnisse zurück und integriert sie in einem adaptiven Modell. Anhand von Relevance Feedback lernt das Modell, die Einzelergebnisse zu fusionieren. Dabei liegt folgendes Basiskonzept zugrunde. Während einer anfänglichen Einsatzphase wird das System einem adaptiven Selbsttest unterzogen. Die vorhandenen Deskribierungswerkzeuge werden parallel angewandt (Polyrepräsentation), wobei der Benutzer dies nach außen hin nicht wahrnimmt. Das einzige, was dem Benutzer abverlangt wird, ist eine Stellungnahme in Form von Relevance Feedback über die Relevanz der Ergebnisse. Dieses Feedback ermöglicht eine zweifache Nutzung: einerseits in seiner traditionellen Form zur Optimierung der Anfrage im Sinne eines interaktiven Retrieval, andererseits um eine kontextabhängige Qualitätseinschätzung der Deskribierungsverfahren zu erreichen. Ein systeminternes Bewertungsschema sammelt, verwaltet und analysiert die Urteile. Eine nachgeordnete Klassifikation versucht, eine Korrelation zwischen den Verfahren und bestimmten Eigenschaften der Objekte herzustellen. Ziel ist, dass sich während des Einsatzes des Information-Retrieval-Systems die geeignetsten Verfahren durchsetzen. Ein adaptierbares Lernkonzept kann zu einem ausgewählten Zeitpunkt eine Neuorganisation der Systemkomponenten vornehmen.

## **2.2 Modellbildung**

MIMOR kann als eine zusätzliche Schicht in einem Information-Retrieval-System betrachtet werden, welche die Kombination verschiedener Ergebnisse verwaltet. Zu Beginn erhalten alle Verfahren das gleiche Gewicht. Damit beeinflussen alle Verfahren das Ergebnis gleich stark. Im Lauf des Lernprozesses adaptiert MIMOR die Gewichte so, dass Verfahren, die das Ergebnis positiv beeinflussen, gestärkt werden. Die Entscheidung, welche Dokumente im Ergebnis besonders relevant sind, trifft der Benutzer durch das Relevance Feedback. Die Verfahren, die den positiv bewerteten Dokumenten hohe System-Relevanz zuweisen, erhalten höheres Gewicht.

MIMOR ist nicht auf Text-Objekte beschränkt, sondern eignet sich auch für Multimedia-Retrieval oder strukturierte Objekte.



## 2.3 Formalisierung

Jedes Information-Retrieval-System berechnet für alle Dokumente eine System-Relevanz oder Retrieval Status Value (RSV) in bezug auf eine Anfrage. MIMOR kombiniert nun die Ergebnisse verschiedener Verfahren zu einem Gesamtergebnis. Tabelle 1 skizziert die Ausgangslage für MIMOR und andere Fusionsverfahren.

	Dok. 1	Dok. 2	Dok. 3	Dok. 4	...
System A	$RSV_{A1}$	$RSV_{A2}$	$RSV_{A3}$	$RSV_{A4}$	...
System B	$RSV_{B1}$	$RSV_{B2}$	$RSV_{B3}$	$RSV_{B4}$	...
System C	$RSV_{C1}$	$RSV_{C2}$	$RSV_{C3}$	$RSV_{C4}$	...
...	...	...	...	...	...

↓

MIMOR	$RSV_1$	$RSV_2$	$RSV_3$	$RSV_4$	...
-------	---------	---------	---------	---------	-----

**Tabelle 1:** Ausgangssituation für MIMOR

Jedes System trägt mit einem Gewicht  $\omega_i$  zu der RSV jedes Dokuments im MIMOR-Ergebnis bei. Folgende Formel liefert dieses Ergebnis für ein Dokument  $j$  bei einer gegebenen Anfrage:

$$(1) \quad RSV_j = \frac{\sum_{i=1}^N (\omega_i RSV_{ij})}{N}$$

$RSV_{ij}$  Retrieval Status Value des Verfahrens  $i$  für Dokument  $j$

Zu Beginn sind alle Gewichte gleich. Im Einsatz adaptiert MIMOR diese Gewichte abhängig von den Relevanz-Bewertungen der Benutzer. In der Regel kann der Benutzer Dokumente als relevant oder nicht relevant markieren. Im formalen Modell ordnet er jedem Dokument einen Relevance-Feedback-Wert  $R$  zu. Für Dokumente, die ohne explizites Urteil bleiben, nimmt  $R$  den Wert Null an. Die Gewichte der einzelnen Verfahren ändern sich nach folgender Formel:

$$(2) \quad \omega_i = \omega_i + (\varepsilon RSV_{ij} R_j)$$

$\varepsilon$  Lernrate

$R_j$  Relevanz – Bewertung für Dokumente  $j$

Das Gewicht eines Retrieval-Systems wächst, wenn es einem positiv bewerteten Dokument eine hohe RSV zugewiesen hat. Nach der Anwendung von Formel 2 auf alle Gewichte werden diese normalisiert, so dass ihre Summe stets konstant bleibt.

Das MIMOR-Modell besteht aus dem Vektor von Gewichten, den ein Administrator des Retrievalsystems einsehen und modifizieren kann. Dieser Administrator kann

die Gewichte der einzelnen Verfahren überwachen und z.B. sehr schlechte Verfahren aus Performanzgründen aus dem Retrievalsystem nehmen.

### **3. Erweiterungen des MIMOR-Modells**

Zahlreiche Untersuchungen im Information Retrieval haben gezeigt, dass die Qualität von Information-Retrieval-Verfahren vom Korpus abhängig ist. Die eingesetzten Algorithmen eignen sich dann mehr oder weniger gut für die Eigenschaften der Dokumente im Korpus. Ebenso beeinflussen die Eigenschaften der Anfragen die Wahl eines Information-Retrieval-Systems. Spezifische Ähnlichkeitsfunktionen entstanden z. B. für kurze Anfragen (cf. Kwok & Chan 1998, Wilkinson et al. 1996).

#### **3.1 Benutzer-Eigenschaften**

Viele Information-Retrieval-Tests haben gezeigt, dass die Relevanz-Bewertungen unterschiedlicher Benutzer oft nicht übereinstimmen (cf. Voorhees 1998). Um eine optimale Ausnutzung zu gewährleisten, müsste demnach jeder Benutzer sein eigenes MIMOR-Modell mit seinen Relevanz-Bewertungen trainieren. Dazu sind aber viele Bewertungen erforderlich und nicht jeder Benutzer wird diese einbringen. Benutzer geben aufgrund des erforderlichen Zeitaufwands oft kein Relevance Feedback, obwohl dies eine der effektivsten Methoden zur Verbesserung von Information-Retrieval-Ergebnissen ist (cf. Over 1998). Auch Benutzer, die dazu bereit sind, verfügen in der Trainingsphase noch über ein suboptimales Modell, das bei erst wenigen Beispielen im Lernprozess eventuell zu extrem in eine Richtung tendiert. Das andere Extrem besteht in einem MIMOR-Modell, das alle Benutzer integriert und somit zwar für eine hohe Anzahl von Trainingsbeispielen sorgt, das aber auch keine Möglichkeit zur Individualisierung einräumt.

Eine Lösungsmöglichkeit besteht in der gleichzeitigen Pflege eines privaten und eines öffentlichen Modells, die beide zum Gesamtergebnis beitragen (cf. Womser-Hacker & Mandl 1999). Beide Modelle bestehen aus einem Vektor von Gewichten, die den Einfluss der einzelnen Verfahren in dem jeweiligen Modell beschreiben. Der Einfluss jedes Modells wird ebenfalls durch ein Gewicht bestimmt. Das private Modell trägt mit dem Gewicht  $p$  zum Gesamtergebnis bei und das öffentliche Modell mit dem Gewicht  $(1 - p)$ . Im Zeitraum der Benutzung des Systems kann sich  $p$  verändern. Zu Beginn der Arbeit mit dem System steht das private Modell auf der Basis weniger Relevanz-Bewertungen und ist damit noch unzuverlässig. Daher ist es sinnvoll,  $p$  und damit den Einfluss des privaten Modells niedrig zu halten und sich mehr auf das bereits ausgewogene öffentliche Modell zu verlassen. Je mehr Bewertungen ein Benutzer im Lauf der Arbeit mit dem System vornimmt, desto größer kann der Einfluss des privaten Modells werden. Das Gewicht  $p$  könnte auch mit steigender Sättigung des privaten Modells wachsen. Ändert sich das private Modell nur noch wenig, so kann es als weitgehend vollständig und stabil gelten. Das Gewicht könnte für fortgeschrittene Benutzer auch manipulierbar sein.

Das Gesamtergebnis berücksichtigt nun zwei MIMOR-Modelle und damit zwei Vektoren von Gewichten:

- Das private Modell:  $(\omega_{private,A}; \omega_{private,B}; \omega_{private,C}; \dots; \omega_{private,N})$
- Das öffentliche Modell:  $(\omega_{public,A}; \omega_{public,B}; \omega_{public,C}; \dots; \omega_{public,N})$

Das Gesamtergebnis integriert die RSV des privaten und des öffentlichen Modells:

$$(3) \quad RSV_j = \frac{\sum_{i=1}^N ((p\omega_{private,i} + (1-p)\omega_{public,i})RSV_{ij})}{N}$$

### 3.2 Dokument-Eigenschaften

Das MIMOR-Modell wird durch die Berücksichtigung von Dokument-Eigenschaften weiter verbessert. Als Kriterien kommen hier z. B. Länge, Dokument-Typ (Zeitungstext, Fachtext, Online-Diskussionsbeitrag, ...), Anzahl von Termen oder Schwierigkeit in Frage. In einer Digitalen Bibliothek müssen diese Eigenschaften automatisch erkannt werden<sup>2</sup>. Für die Schwierigkeit eines Textes kommen als formal erkennbare Größen z. B. syntaktische Komplexität oder Anzahl von Wörtern in Frage (cf. z. B. Ballod 2000). Diese Eigenschaften können im System als Cluster von Dokumenten, die diese Eigenschaft besitzen, modelliert werden<sup>2</sup>. Die Güte der Cluster bildet die Grundlage für den Erfolg dieses Ansatzes. Die Dokumente innerhalb eines Clusters sollten möglichst homogen und ähnlich sein, während Dokumente aus verschiedenen Clustern unterschiedlich sein sollten.

	Cluster 1			Cluster 2		
	Doc. 1.1	Dok. 1.2	...	Dok. 2.1	Dok. 2.2	...
System A	$\omega_{x,1,A}$			$\omega_{x,2,A}$		
System B	$\omega_{x,1,B}$			$\omega_{x,2,B}$		
System C	$\omega_{x,1,C}$			$\omega_{x,2,C}$		
...	...			...		

**Tabelle 2:** Dokument-Cluster mit verschiedenen Gewichten, x I {private, public}

Jedes Cluster verfügt über ein eigenes, veränderbares Gewicht für jedes Information-Retrieval-Verfahren. Damit kann ein Verfahren in einem Cluster ein

<sup>2</sup> Momentan sind die Cluster überschneidungsfrei modelliert. Da in der Praxis sicher häufig Mehrfachzuordnungen auftreten, ist eine Erweiterung auf sich überlappende Fuzzy-Cluster geplant.

hohes Gewicht und in einem anderen ein niedriges Gewicht haben. Dieses Verfahren eignet sich dann gut für Dokumente mit der Eigenschaft des ersten Clusters und weniger gut für Dokumente mit der Eigenschaft des zweiten Clusters. Tabelle 2 skizziert die Verteilung unterschiedlicher Gewichte auf die Dokument-Eigenschaften.

Je nach Zugehörigkeit des in Frage stehenden Dokuments wird für die Berechnung der RSV ein anderes Gewicht verwendet. Gehört das Dokument j zum Cluster C, dann ergibt sich in Abwandlung von Formel 1:

$$(4) \quad RSV_{j \in C} = \frac{\sum_{i=1}^N (\omega_{X,C,j} RSV_{ij})}{N}$$

Re;      Relevanz-Bewertung für Dokument j in Cluster C  
RSV<sub>ij</sub>   Retrieval Status Value von Verfahren i für Dokument j

Die Lernformel für die Veränderung der Gewichte ist fast identisch mit der Lernformel (2) für MIMOR. Die Adaption wirkt nun aber nur für die Gewichte der Verfahren für das Cluster, in dem das bewertete Dokument liegt.

$$(5) \quad \omega_{X,C,i} = \omega_{X,C,i} + (\varepsilon R_{cj} RSV_{ij})$$

$\varepsilon$       Lernrate  
R<sub>cj</sub>      Relevanz-Bewertung für Dokument j in Cluster c  
RSV<sub>ij</sub>   Retrieval Status Value von Verfahren i für Dokument j

Für diese Erweiterung des Modells benötigt das Lernverfahren mehr Trainingsbeispiele, da die Anzahl der zu trainierenden Parameter mit der Anzahl der Cluster wächst. Die Eignung bestimmter Cluster bzw. Eigenschaften kann nur empirisch überprüft werden. Es ist auch unwahrscheinlich, dass sich allgemeingültige Regeln finden, vielmehr dürften gute Kandidaten für Dokument-Eigenschaften stark von der jeweiligen Kollektion abhängen.

### 3.2 Anfrage-Eigenschaften

Da auch die Eigenschaften der Anfragen bei der Bewertung der Qualität von Information-Retrieval-Systemen eine Rolle spielen (cf. Voorhees & Harman 1998), sollte ein Fusions-Verfahren darauf Rücksicht nehmen. MIMOR berücksichtigt daher analog zu den Dokument-Eigenschaften auch die Eigenschaften von Anfragen, die wiederum durch die Zusammenfassung in Cluster modelliert werden. Dabei sind weniger Cluster zu erwarten als bei den Dokumenten, da bei den Anfragen eine geringere Varietät bezüglich ihrer formal bestimmbaren Eigenschaften vorliegt. In der Regel gibt es wesentlich weniger Anfragen als

Dokumente, die Anfragen sind meist erheblich kürzer und verschiedene Texttypen wie bei den Dokumenten sind bei den Anfragen kaum denkbar.

Die Berechnungsformel für das Endergebnis und die Lernformel für die Adaption der Gewichte der Cluster ergeben sich analog zu den Formeln für die Dokument-Eigenschaften.

#### **4. Transformationen in Digitalen Bibliotheken**

Die semantische Heterogenität in umfassenden Digitalen Bibliotheken ist eine Folge der Integration verschiedener Objekte (Volltext, Abstracts, Fachliteratur, multimediale Dokumente, ...) und deren heterogener Erschließung (automatische Indexierung, Thesauri, Klassifikationen, Verschlagwortung durch Autoren oder Dokumentare, etc.).

Ein Beispiel für ein derartiges System ist die geplante Virtuelle Fachbibliothek Sozialwissenschaften, in der Bibliotheken, Verlage, die Friedrich-Ebert-Stiftung und das Informationszentrum Sozialwissenschaften eine gemeinsame Plattform für die Integration ihrer bisher getrennten Datenbestände schaffen und so die Versorgung der Fachwissenschaftler mit Literatur verbessern (cf. Kluck et al. 2000).

Um mit der semantischen Heterogenität der unterschiedlich erschlossenen Dokumente umzugehen, sind entsprechende Komponenten erforderlich. Da die meisten Retrieval-Verfahren auf einer homogenen Repräsentation aufsetzen, müssen automatische Verfahren gefunden werden, die Dokumente zwischen verschiedenen Repräsentationen transformieren, um so eine homogene Grundlage für das Retrieval zu schaffen. Kuhlen etwa spricht von einer „postkoordinierenden Ordnung durch transformierende Anpassung“ (Kuhlen 1999:138).

Der große Umfang der Daten in Digitalen Bibliotheken spricht gegen eine ausschließlich intellektuelle Lösung. Basis von automatischen Verfahren für Transformationen ist meist ein Korpus, dessen Objekte in zwei Eigenschaftsräumen repräsentiert sind. In der Regel bestehen diese Doppelkorpora aus Texten, die nach zwei Indexierungsverfahren erschlossen wurden.

Bisher leisten meist statistische Verfahren auf der Basis einer Kookkurrenz-Analyse die Transformation (cf. z. B. Ferber 1997, Lam & Yang 1998). Mandl 1998 schlägt auch neuronale Netze vor, wobei das häufig eingesetzte Backpropagation-Netzwerk als besonders erfolgversprechend gilt. Der Backpropagation-Algorithmus ist ein Lernverfahren, das anhand von Trainingsbeispielen eine Abbildung von einem n-dimensionalen in einen m-dimensionalen Raum lernt. Eine versteckte Schicht von Neuronen zwischen Input- und Output-Schicht erhöht die Mächtigkeit und erlaubt auch das Erlernen komplexer Funktionen (cf. Scherer 1997).

Ein Vergleich der Qualität von Transformationsverfahren in Mandl 2000 führte zu interessanten Ergebnissen. Dabei wurden ein statistisches Verfahren, das auf der Basis von Kookkurrenzen Assoziationswerte zwischen den Term-Räumen berechnet, und ein neuronales Backpropagation-Netzwerk verglichen, das auf mit Latent Semantic Indexing (cf. Berry et al. 1995) reduzierte Repräsentationen zugreift.

Testgrundlage bilden Daten des Informationszentrum Sozialwissenschaften (IZ) und der Universitäts- und Stadtbibliothek Köln (USB). Das IZ indexiert sozialwissenschaftliche Dokumente mit einem Thesaurus von Schlagwörtern und zusätzlich mit einer Klassifikation wissenschaftlicher Fachgebiete. Die USB verschlagwortet teilweise die gleichen Dokumente nach ihrer Bibliotheksklassifikation. In einem Experiment wurde eine Transformation vom IZ-Thesaurus auf die IZ-Klassifikation mit 12.000 Dokumenten trainiert und getestet und in einem zweiten Experiment wurde eine Abbildung von der USB-Klassifikation zu einem Teil des IZ-Thesaurus mit 15.000 Dokumenten trainiert und getestet. Im ersten Fall war die Qualität der Transformation beim statistischen Verfahren und beim Backpropagation-Netzwerk sehr ähnlich und im zweiten Experiment erwies sich das Backpropagation-Netzwerk als besser (cf. Mandl 2000).

Interessanterweise ist in beiden Fällen die Schnittmenge der Ergebnisse sehr klein. Im ersten Experiment führen beide Verfahren zu vergleichbarer Qualität, die Treffer sind aber sehr unterschiedlich. Damit ergibt sich in Mandl 2000 für Transformationen ein ähnliches Bild wie bei TREC für Retrieval. Als Konsequenz daraus sollte auch bei Transformationen mit Fusionsansätzen experimentiert werden, um die Resultate mehrerer Verfahren zu kombinieren.

## **5. Anwendung von MIMOR auf Transformationen**

Das MIMOR-Modell lässt sich für Transformationen adaptieren. Liegen Relevance-Feedback-Informationen zu den Zuordnungen vor, dann lernt MIMOR, den Einfluss einzelner Transformations-Verfahren auf das Gesamtergebnis zu optimieren.

Transformationen bilden Dokumente aus einem Term-Raum in einen anderen Term-Raum ab. Dazu weisen sie einem Dokument für jeden Term im Ziel-Term-Raum ein Gewicht zu. Eine Transformations-Funktion ordnet also jedem Paar aus Dokument und Zielklasse ein Gewicht zu. Bei der Kombination mehrerer Verfahren muss der Fusionsalgorithmus aus den Gewichten aller Verfahren ein Gesamtgewicht bilden. Wie Tabelle 3 zeigt, ist damit die Ausgangssituation ähnlich zu der von MIMOR.

	Term 1	Term 2	Term 3	Term 4	...
Verfahren A	$G_{A1}$	$G_{A2}$	$G_{A3}$	$G_{A4}$	...
Verfahren B	$G_{B1}$	$G_{B2}$	$G_{B3}$	$G_{B4}$	...
Verfahren C	$G_{C1}$	$G_{C2}$	$G_{C3}$	$G_{C4}$	...
...	...	...	...	...	...

↓

MIMOR	$G_1$	$G_2$	$G_3$	$G_4$	...
-------	-------	-------	-------	-------	-----

**Tabelle 3:** Ausgangssituation für MIMOR für die Transformation eines Dokuments

Jedes Verfahren erhält zu Beginn des Prozesses den gleichen Einfluss auf das Gesamtergebnis, das sich aus folgender Formel ergibt:

$$(6) \quad G_j = \frac{\sum_{i=1}^N \omega_i G_{ij}}{N}$$

$G_j$       Gewicht von Term j für das Dokument im Gesamtergebnis

$G_{ij}$       Gewicht von Term j für das Dokument von Verfahren i

$\omega_i$       Gewichtung von Verfahren i

Im Laufe der Interaktion adaptiert MIMOR die Gewichte so, dass die Verfahren, die positiv auf das Gesamtergebnis wirken, ein höheres Gewicht erhalten. Um die Wirkung zu bestimmen, beurteilt der Benutzer Dokumente als relevant oder nicht relevant für die Zielklasse. Das Relevance Feedback bezieht sich jetzt also nicht mehr auf die Relevanz zu einer Anfrage, sondern auf die Adäquatheit einer Zuordnung zu einem Term. Die Gewichte der einzelnen Verfahren ändern sich ähnlich wie in der Lernformel (2):

$$(7) \quad \omega_i = \omega_i + (\varepsilon G_{ij} R_j)$$

$R_j$       Relevanz-Bewertung für Dokument j bezüglich der Zielklasse

Ein Problem besteht in der Gewinnung von Relevance-Feedback-Informationen für die Transformation zwischen Term-Räumen, da Benutzer sinnvollerweise nur das Endergebnis eines Suchprozesses und damit Dokumente bewerten können. Eine direkte Bewertung der Transformation kommt evtl. für Fachleute in Frage, jedoch ist die Abschätzung des Retrievalerfolgs aufgrund einer Liste von Termen sehr schwierig. Die Relevance-Feedback- Bewertung von Dokumenten muss also in eine Bewertung der Terme umgesetzt werden. Dabei werden die Transformations-Verfahren belohnt, die Terme hoch gewichtet haben, die wiederum zu positiv bewerteten Dokumenten führten. Dazu wird aus den Relevanz-Feedback-Informationen ein optimaler Term- Vektor gewonnen. Die Term- Vektoren der bewerteten Dokumente werden mit dem jeweiligen, numerisch umgesetzten Relevanz-Urteil multipliziert und anschließend summiert:

$$(8) \quad \vec{p}_j = \sum_{i=1}^B (R_i \vec{d}_i)$$

B Anzahl der bewerteten Dokumente

Die Lernformel berechnet den Erfolg eines Verfahrens als Mittel des Erfolgs bei allen Termen im Zielvokabular:

$$(9) \quad \omega_i = \omega_i + \varepsilon \sum_{j=1}^{Dim} G_{ij} (o_j - Erg_j)$$

Dim	Anzahl Vektoren im Zielvokabular
$O_j$	optimaler Wert für Term j laut Relevanz Feedback
$Erg_j$	berechneter Wert für Term j

Völlig analog lassen sich auch die Ansätze zur Erweiterung von MIMOR auf die Transformationen übertragen. Auch die Transformation kann in ein privates und öffentliches Modell gespalten werden, die jeweils mit einem eigenen Gewicht ins Ergebnis eingehen. Weiterhin können die Dokumente und Zielklassen je nach ihren Eigenschaften in Cluster unterteilt werden, von denen jedes ein eigenes Gewicht für alle Verfahren erhält. Damit erhält z. B. ein Verfahren, das sich für bestimmte Dokument-Eigenschaften bewährt, für diese Dokumente ein höheres Gewicht.

## 6. Fazit

Digitale Bibliotheken verändern die Anforderungen an Information-Retrieval-Systeme und erfordern zusätzliche Komponenten. Im vorliegenden



Zusammenhang wurde besonderes Augenmerk auf die unterschiedlichen Repräsentationsformen informationeller Objekte gelegt. Deren Integration in einem Fusionsverfahren erlaubt es, dass die Stärken und Schwächen einzelner Repräsentationsformen sehr differenziert zur Geltung kommen. Eine Implementierung und anschließende Evaluierung von MIMOR in JAVA ist geplant. Dabei sollen die Komponenten so realisiert werden, dass sie soweit möglich auch für Transformationen einsetzbar sind.

## **Literaturverzeichnis**

Ballod, Matthias (2000): Comprehensible Science — Computer-Supported Text Analysis of Publications on Gene Technology. In: Global Dialogue. Science and Technology — Thinking the Future at EXPO 2000 Hannover. 11. — 13. 7. 2000. <[http://www.shaping-the-future.de/pdf\\_www/207\\_paper.pdf](http://www.shaping-the-future.de/pdf_www/207_paper.pdf)>.

Belkin, N. (1993): Interaction with Texts: Information Retrieval as Information Seeking Behavior. In: Knorz, G.; Krause, J.; Womser-Hacker, C. (Eds.): Information Retrieval '93. Von der Modellierung zur Anwendung. Proc. d. 1. Tagung Information Retrieval. Konstanz. S. 55 — 66.

Berry, M.; Dumais, S.; Letsche, T. (1995): Computational Methods for Intelligent Information Access. In: Proc. of ACM Supercomputing '95. San Diego, CA. S. 1 — 38.

Croft, B.; Moffat, A.; Rijsbergen, K. van; Wilkinson, R.; Zobel, J. (1998) (Eds.): Proceedings of the 21 st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '98). Melbourne 24. — 28 .8. 1998.

Ferber, R. (1997): Automated Indexing with Thesaurus Descriptors: A Cooccurrence Base Approach to Multilingual Retrieval. In: Peters, C.; Thanos, C. (Eds.): Research and Advanced Technology for Digital Libraries. Ist European Conf. ECDL '97. Pisa, 1. — 3. 9. 1997. Berlin et al.. S. 233 — 252.

Fox, E.; Shaw, J. (1994): Combination of Multiple Searches. In: Harman, D. (Ed.): The Second Text Retrieval Conference (TREC-2). NIST Publ. 500-215. S. 243 — 252.

Fuhr, N. (1999): Information Retrieval in Digitalen Bibliotheken. In: Schmidt, Ralph (Hrsg.): Proc. 21. Online Tagung der DGI. Aufbruch ins Wissensmanagement. Deutsche Gesellschaft für Informationswissenschaft und Informationspraxis. Frankfurt, 18. — 20. 5. 1999. S. 93-102.

Harman, D. (1995): The TREC conferences. In: Kuhlen, Rainer; Rittberger, Marc (Eds.) (1995): HIM '95. Hypertext, Information Retrieval, Multimedia. Synergieeffekte elektronischer Informationssysteme. Konstanz. 5. — 7. 4. 1995. Konstanz. S. 9 — 28.

Harman, D. (Ed.) (1996): The Fourth Text Retrieval Conference (TREC-4). NIST Special Publication 500-236. National Institute of Standards and Technology. Gaithersburg, Maryland, 1. – 3. 11. 1995. <[http://trec.nist.gov/pubs/trec4/t4\\_proceedings.html](http://trec.nist.gov/pubs/trec4/t4_proceedings.html)>

Kluck, M.; Krause, J.; Müller, M.; in Kooperation mit: Schmiede, R.; Wenzel, H.; Winkler, S.; Meier, W. (2000): Virtuelle Fachbibliothek Sozialwissenschaften: IZ-Arbeitsbericht 19, IZ Sozialwissenschaften, Bonn. <<http://www.bonn.iz-soz.de/publications/series/working-papers/index.htm#Virtuelle>>.

Krause, Jürgen (1998): Innovative Current Research Information Systems in the Information Society. In: CRIS '98 Current Research Information Systems. Luxemburg, 12-14.3.1998. <<ftp://ftp.cordis.lu/pub/cybercafe/docs/krause.zip>>

Kuhlen, Rainer (1999): Die Konsequenzen von Informationsassistenten: Was bedeutet informationelle Autonomie oder wie kann Vertrauen in elektronische Dienste in offenen Informationsmärkten gesichert werden? Frankfurt a. M.

Kwok, K. L.; Chan, M. (1998): Improving Two-Stage Ad-Hoc Retrieval for Short Queries. In: Croft et al. (1998). S. 250 — 256.

Lam, Wai; Ho, Chao Yang (1998): Using a Generalized Instance Set for Automatic Text Categorization. In: Croft et al. (1998).

Mandl, T. (1998): Vague Transformations in Information Retrieval. In: Zimmermann, H.; Schramm, V. (Eds.): Knowledge Management und Kommunikationssysteme: Workflow Management, Multimedia, Knowledge Transfer. Proc. 6. Int. Symposium für Informationswissenschaft. (ISI '98). 3. – 7. 11. 1998, Prag, Konstanz. S. 312 – 325.

Mandl, T. (2000): Einsatz neuronaler Netze im Information Retrieval. Arbeitsbericht. Informationswissenschaft, Universität Hildesheim.  
Over, P. (1998): TREC-6 Interactive Track Report. In: Voorhees/Harman (1998). S. 73–81.

Scherer, A. (1997): Neuronale Netze: Grundlagen und Anwendungen.  
Tague-Sutcliffe, J.; Blustein, J. (1995): A Statistical Analysis of the TREC-3 Data. In: Harman, D. (Hrsg.): The Third Text Retrieval Conference (TREC-3). NIST Special Publication 500-225. National Institute of Standards and Technology. Gaithersburg, Maryland, 2.-4.11.1994. S. 385.

Voorhees, e. (1998): Variations in relevance judgments and the measurement of retrieval effectiveness. In: Croft et al. 1998. S. 315-323

Voorhees, e.; Harman, D. (1998): The Sixth Text Retrieval Conference (TREC-6). NIST Special Publication 500-240. August 1998.

<[http://trec.nist.gov/pubs/trec6/t6\\_proceedings.html](http://trec.nist.gov/pubs/trec6/t6_proceedings.html)>.

Wilkinson, R.; Zobel, J.; Sacks-Davis, R. (1996): Similarity Measures for Short Queries. In: Harman (1996).

Womser-Hacker, C. (1997): Das MIMOR-Modell. Mehrfachindexierung zur dynamischen Methoden-Objekt-Relationierung im Information Retrieval. Habilitationsschrift. Universität Regensburg.

Womser-Hacker C.; Mandl, T. (1999): Adapting Meta Information Retrieval to User Preferences and Document Features. In: Bullinger, H.-J.; Ziegler, J. (eds.): Human-Computer Interaction: Communication, Cooperation and Application Design. Proceedings of the HCI International '99 (8th International Conference on Human-Computer Interaction), Munich, Germany, 22 – 27. 8. 1999. Lawrence Erlbaum Associates: Mahwah, NJ; London. Vol. 2. S. 604 – 608.



# Aiding Web Searches by Statistical Classification Tools

Gerhard Heyer, Uwe Quasthoff, Christian Wolff

Leipzig University  
Computer Science Institute, NLP Dept.  
Augustusplatz 10/11  
04109 Leipzig, Germany  
{heyer, quasthoff, wolff} @informatik.uni-leipzig.de

## Abstract

We describe an infrastructure for the collection and management of large amounts of text, and discuss the possibility of information extraction and visualisation from text corpora with statistical methods. The paper gives an overview of processing steps, the contents of our text databases as well as different query facilities. Our focus is on the extraction and visualisation of *collocations* and their usage for aiding web searches.

## Introduction

We describe an infrastructure for managing large monolingual language resources. Since 1995, we have accumulated a German text corpus of more than 300 Million words with approx. 6 Million different word forms in approx. 13 Million sentences. The Project - originally called "Deutscher Wortschatz" (*German Vocabulary*) - has recently been extended to include corpora of other European languages (Dutch, English) as well, with more languages to follow in the near future (see table 1).

	<i>German</i>	<i>English</i>	<i>Dutch</i>	<i>French</i>
<i>word tokens</i>	300 Mill.	250 Mill.	22 Mill.	15 Mill.
<i>sentences</i>	13,4 Mill.	13 Mill.	1,5 Mill.	860.000
<i>word types</i>	6 Mill.	1,2 Mill.	600.000	230.000

**Table 1:** Basic Characteristics of the Corpora

The intent of the project is to collect large amounts of textual data for experimenting with corpus based semantic processing. The approach is based on the extraction of sentences from various types of texts. The sentence is chosen as the



basic structuring unit due to copyright restriction on the one hand, as a feasible level of linguistic representation adequate for giving examples for word tokens on the other. As we aim at developing an infrastructure for corpus processing rather than a single corpus of text, there (almost) no restriction on the type of text to be included in the corpus.

Starting off from a rather simple data model tailored for large data volumes and efficient processing using a relational data base system as storage we employ a simple yet powerful technical infrastructure for processing large amounts of texts to be included in the corpus. Beside basic procedures for text integration into the corpus we have developed tools for post-processing our linguistic data. The corpus is available on the WWW (<http://www.wortschatz.uni-leipzig.de>) and may be used as a large online dictionary.

## Methodological Approach

Our collection is comprehensive rather than error-free. In the long run we aim at representing a large portion of current-day word usage available from various sources. While this does not prevent inclusion of errors (like typos in newspaper text), we are able to eliminate typical sources of erroneous information by statistical as well as intellectual optimisation routines (see Quasthoff 1998a for details).

In addition, only a high data volume of the corpus allows for the extraction of information like sentence-based word collocations and information about low frequency terms. At the same time, the infrastructure should be open for the Integration of various knowledge sources and tools: We strongly believe that there is no single linguistic or statistical approach for all operational needs (optimisation tasks, information extraction etc.). Hence, we provide data for very different purposes.

The backbone of our project is a relational database system. We have chosen *MySQL* (cf. <http://www.mysql.com>) as a low cost DBMS with high Performance and availability on several major platforms. Currently, the corpus databases are stored on UNIX/Linux database servers, while Web access to the corpus-related information services is provided by an Apache Web Server running on a Windows NT machine. Using SQL as data definition and manipulation language, we can make sure that standardised APIs for access and extraction tools are available for all major programming languages.

In contrast to work by Rieger at the University of Trier, or Paprotte at the University of Münster, the primary concern of the project is the collection of a very large amount of textual and lexical data that we intend to make publicly available. In this respect, the project follows similar intentions as CISLEX (Guenther 1996), but in comparison covers a much larger set of data. Also, CISLEX seems to focus on morphological analyses, while for us there textual and lexical data are used as raw material for further semantic processing such as analysing definitions, calculating collocations, disambiguating terms, deriving co-hyponyms, etc..

## **Processing Model**

Our corpus-processing infrastructure is based on two major design objectives:

- acceptance of a wide variety of digital text and
- definition of a chain of processes for the automatic setup of data corpora, extraction mechanisms, and access services.

## **Data Sources**

Data acquisition for our corpora is based on the analysis of available electronic text from various sources. These include

- General newspaper text (major German newspapers, English newspaper text from the TREC and TIPSTER collection, cf. Voorhees & Harman 1999).
- Electronic dictionaries (general knowledge dictionaries as well as technical and domain-related like medical dictionaries).
- Electronic books and journals, mostly CD-ROM-based collections.
- Web resources with a minimum level of language quality.

In the starting phase of corpus setup, text was primarily extracted from CDROMs provided by various publishers. With more and more high quality text coded in declarative markup formats like HTML being made available via the world wide web the collection strategy in our approach has changed: We employ configurable search agents for collecting texts which also do basic feature extraction like coding main subject areas in newspaper texts on the WWW.

## **Text Processing**

The processing of input data is done in several steps which may roughly be divided into the necessary routines for the extension of the corpus by including new data, and postprocessing of information for the whole database. The pre-processing steps include format conversion, i.e. extraction of raw text from various formats like PDF, MS-WinWord or HTML, the partitioning of documents into sentences, lexical analysis (word and phrase recognition as well as identification of special phrase types like multi-word proper names) and indexing of the whole text corpus.

We maintain a complete full-text index for the whole corpus, making analysis of typical word usage a simple task. The underlying data model stores single words as well as concepts and phrases automatically extracted from the corpus. Beyond the raw data level, our data model provides for the integration of additional information of various categories:

- syntactic and morphological information at word level
- semantic information like subject areas or classification codes at word and sentence levels
- information about related words, either from knowledge sources like synonym dictionaries or thesauri, or as the result of automatic extraction (word collocations, sentence classification).

This information is collected not only from various sources (dictionaries with classification codes or subject areas), but also by applying linguistic analysis tools, some of which are used in co-operation with other NLP groups (e. g. the TNT tool for part-of-speech tagging, cf. Brants 2000).

## Information Categories in the Database

The basic structure of entries in the corpus database includes information on the absolute word frequency for each entry (i. e. each inflected word form or each identified phrase like the proper name *Helmut Kohl*). Additional frequency class is calculated based on a logarithmic scale relative to the most frequent word in the corpus. For the English corpus, the most frequent word, *the*, has frequency class 0, while an entry like *Acropolis* with an absolute frequency of 20 belongs to frequency class 18, as *the* occurs approx.  $2^{18}$  times more often. In addition to this basic statistical information, example sentences extracted from the texts most recently included in the corpus are given for each word. Table 2 gives an overview of the most important information categories in our corpus and their relative amount (German corpus database):

<i>Information Category</i>	<i>Number of Entries</i>
word list	ca. 6 Mio. word forms
example sentences	ca. 13 Mio.
grammatical information	ca. 3 Mio.
morphological information	ca. 3 Mio.
descriptions	ca. 150.000
subject categories	ca. 1,5 Mio.
semantic relations	ca. 500.000
pragmatics (e. g. usage)	ca. 35.000
collocations (at sentence level)	ca. 3,5 Mio.
collocations (immediate left and right)	ca. 1,5 Mio.
full text index	ca. 30 Mio.

**Table 2:** Information Categories in the German corpus database

If available, morphological and semantic information are presented. Fig. 1 shows an example for the (partially translated) entry *Weltanschauung* from the German corpus.

<p><b>Word</b> (<i>word number</i>: 95400): Weltanschauung <b>Frequency class</b>: 14 (Absolute count: 387) <b>Subject Area</b>: General, Chemistry, Natural Science, Science, Culture, Education, Learning, Chemie -&gt; Naturwissenschaft -&gt; Wissenschaft -&gt; Kultur Erziehung Bildung Wissenschaft) <b>Morphology</b>: welt   an   schau   ung (=welt+an=schau%ung) <b>Grammatical Information</b>: <b>Part of Speech</b>: Noun <b>Gender</b>: Feminine <b>Inflection</b>: die Weltanschauung, der Weltanschauung, der Weltanschauung, die Weltanschauung, die Weltanschauungen, der Weltanschauungen, den Weltanschauungen, die Weltanschauungen (<i>inflection dass fb</i>) <b>Relations to other Entries</b>: <b>Synonyms</b>: Anschauungsweise, Betrachtungsweise, Denkweise - <b>Compare To</b>: Fatalismus, Idealismus, Ideologie, Kommunismus, Nihilismus, Optimismus, Pazifismus, Realismus <b>Synonym of</b> Anschauungsweise, Denkart, Denkungsweise, Denkweise, Einstellung, Ideologie, Lebensanschauung, Meinung, Mentalität, Philosophie, Sinnesart, Standpunkt, Urteil, Weltbild <b>Examples</b>: Auch die Schulmedizin beinhaltet schließlich eine <i>Weltanschauung</i> - eben die rein naturwissenschaftliche. (Source: TAZ 1997) Behindert die anthroposophische <i>Weltanschauung</i> nicht zugleich die Verbreitung solcher Heilmethoden? (Source: TAZ 1997) Wenn man die Medizin zur <i>Weltanschauung</i> macht, ja. (Source: TAZ 1997)</p>
---

**Figure 1:** Sample Entry for *Weltanschauung* (German corpus)

## Types of Queries

Besides querying for single word entries, the SQL-based approach allows for a broad range of query types. Among them are searches in database fields like word descriptions (subject areas), searches for grammatical information and querying the full-text index of the sentence database as well as special pur-pose queries like retrieving all words with a given length or selecting all words attributed with a given subject area. Additionally, administrative query types allow for the management of currently active database processes and the evaluation of access statistics.

## Collocations

Beyond simple text processing we have developed a number of *information extraction tools* which are based on statistical methods. Among them the automatic calculation of sentence-based word collocations stands out as an especially valuable tool for corpus-based language technology applications.

The occurrence of two or more words within a well- defined unit of information (sentence, document) is called a collocation. For the selection of meaningful and



significant collocations, an adequate collocation measure has to be defined. In the literature, quite a number of different collocation measures can be found (for an in-depth discussion of various collocation measures and their application cf. RUGE 1994 and LEMNITZER 1998). Given two words A, B, each occurring  $a$ ,  $b$  times in  $n$  sentences, and  $k$  times together, the following table shows different measures for collocation significance:

<b>Tanimoto</b> (Percentage of double in relation to single hits)	$sig_T(A,B) = k / (a + b - k)$
<b>Mutual Information Index</b> (Digression from statistical independence)	$sig_r(A, B) = \log(kn / (ab)) [= \log(p_{AB} / (p_A p_B))]$
<b>G-Test</b> (Test for Poisson distributions)	$sig(A, B) = x - k \log x + \log k!$ with $n$ = number of sentences, $x = \frac{ab}{n}$

**Table 3:** Different Significance Measures for Collocations

Based on an evaluation of these measures, we have chosen the G-Test related measure for the calculation of collocation significance, as this measure guarantees a good scalability of results in relation to the absolute occurrence frequency of the collocation terms. Two different types of collocations are generated: Collocation based on occurrence *within the same sentence* as well as *immediate left and right neighbours* of each word. Fig. 2 shows an example listing of the top 50 collocations for the term *retrieval* taken from the English corpus, number in brackets indicate the relative strength of the collocation measure. (As the basis for calculating the collocations are inflected word forms, individual word forms, such as *text* and *Text* for instance, are case sensitive. Difference in spelling is indicative of proper names in English, in German it even indicates difference in syntactic category.)

**Top 50 significant sentence-based collocations for *retrieval*:**

storage (625), text (406), data (390), information (349), search (259), document (211), full-text (204), database (149), Topic (136), indexing (129), software (123), systems (106), documents (103), image (97), CD-ROM (87), optical (82), management (78), Text (75), Verity (71), content-based (71), an-line (62), file (61), capabilities (60), query (60), access (58), processing (58), engine (51), databases (50), electronic (47), Provides (46), archival (46), files (42), hypertext (42), stored (39), archiving (38), users (38), searching (37), Boolean (35), records (35), Gescan (33), applications (33), functions (33), user (33), images (31), queries (30), relational (30), fast (29), searches (29), Information (28), disk (28), Fulcrum (27),

**Significant left neighbours of *retrieval*:**

text (401), information (293), data (190), full-text (161), document (96), contentbased (77), image (50), Text (44), file (31), Topic (25), concept (20), fast (18), later (18), rapid (18), an-line (16), Concept (15), database (15), quick (15), Full-text (14), easy (13), Information (12), free-text (12), interactive (12), storage (12), message (11), Data (10), computer-assisted (9), subsequent (9), Boolean-based (7), faster (7), news (7), record (7), remote (7), semantic (7), DiscPassage (6), associative (6), index-only (6), allows (5), efficient (5), instant (5), knowledge (5), literary-quote (5), quote (5), sequential (5), Content-based (4), archival (4), conceptbased (4), legal-information (4), search (4)

**Significant left neighbours of *retrieval*:**

software (175), systems (119), engine (83), capabilities (58), program (31), functions (25), package (23), service (19), services (17), capability (15), tool (13), methods (10), packages (10), process (10), engines (9), programs (9), technology (9), utility (9), stations (8), times (8), clients (7), mechanism (7), operations (7), facilities (6), mechanisms (6), product (6), subsystems (6), techniques (6), time (6), method (5), performance (5), products (5), speed (5), strategies (5), tools (5), client (4), purposes (4), speeds (4)

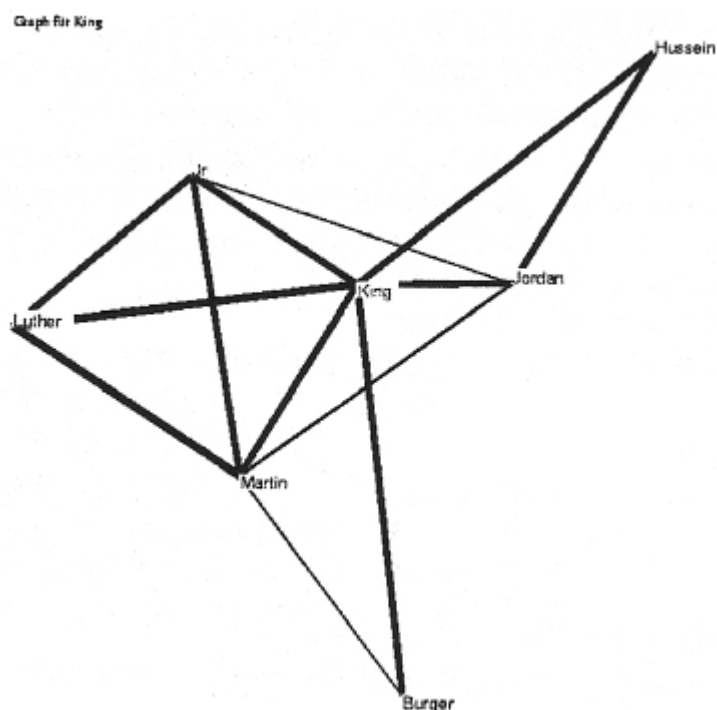
**Figure 2:** Collocation Sets for *retrieval* (English corpus)

Although the calculation of collocations for a large set of terms is a computationally expensive procedure, we have developed efficient trie-based algorithms which allow for a collocation analysis of the complete corpus. Beyond retrieving the different collocation sets for a given word, the infrastructure provides for what may be called "second order queries" on collocations: For example, the *intersection of collocation sets* for two words will contain words that have a strong relationship to both query terms. Intersecting the terms *amerikanische* (*American*) and *Präsident* (*president*) in the German corpus, yields a result set, that - among other entries - contains the names of American presidents *Bill Clinton* and *George Bush* with *Bill Clinton* carrying the highest significance measure for that query. The introduction of part-of-speech information additionally allows a more precise selection of collocation sets: Using the sets of immediate left and right neighbour collocations, it is possible to retrieve typical adjectives that appear to the left of a given noun or, verbs that appear to the right of a given noun.

## Visualisation

Based on the set of collocations for any given word with a minimum number of significant sentence-based collocations we have implemented a real-time visualisation algorithm using simulated annealing (cf. Davidson & Harel 1996). The intention is to display selected relationships from the set of collocations in the resulting graph. In effect, the graphs can be used for representing different meanings of homonyms: In fig. 3 different meanings of *King* as a proper name

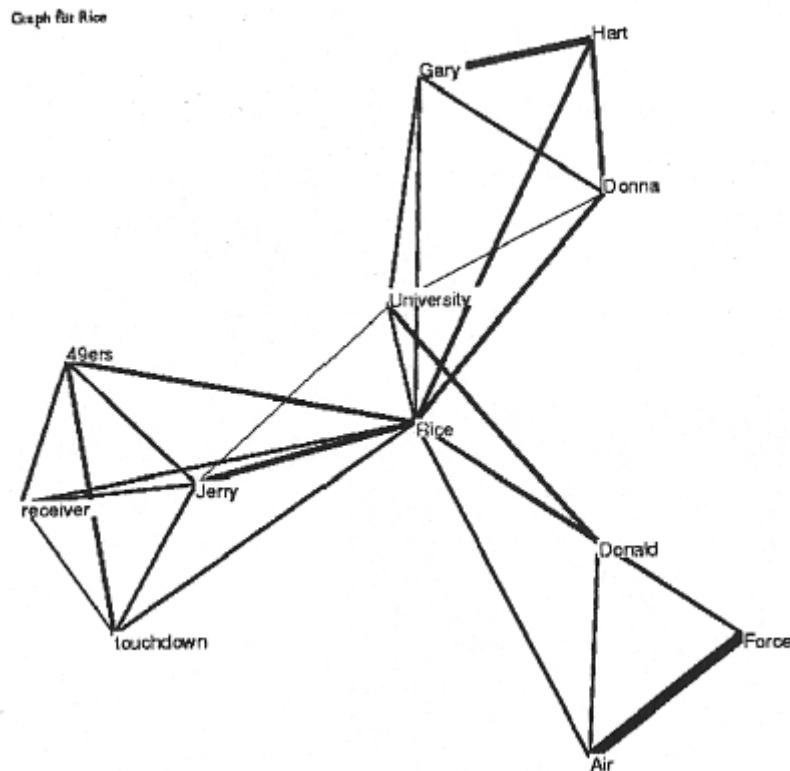
(*Martin Luther King, Jr, Burger King*) and as a title (head of a monarchy, *King Hussein of Jordan*) become apparent:



**Figure 3:** Collocation graph for King (English Corpus)

A second example for the word Rice taken from the English corpus, shows a nice division in the collocation set for different persons and institutions with the proper name Rice:

- an American Secretary of Defense (*Donald Rice*),
- a famous baseball player (*Jerry Rice*),
- *Rice University* and
- *Donna Rice* (ex-lover of presidential candidate Gary Hart).



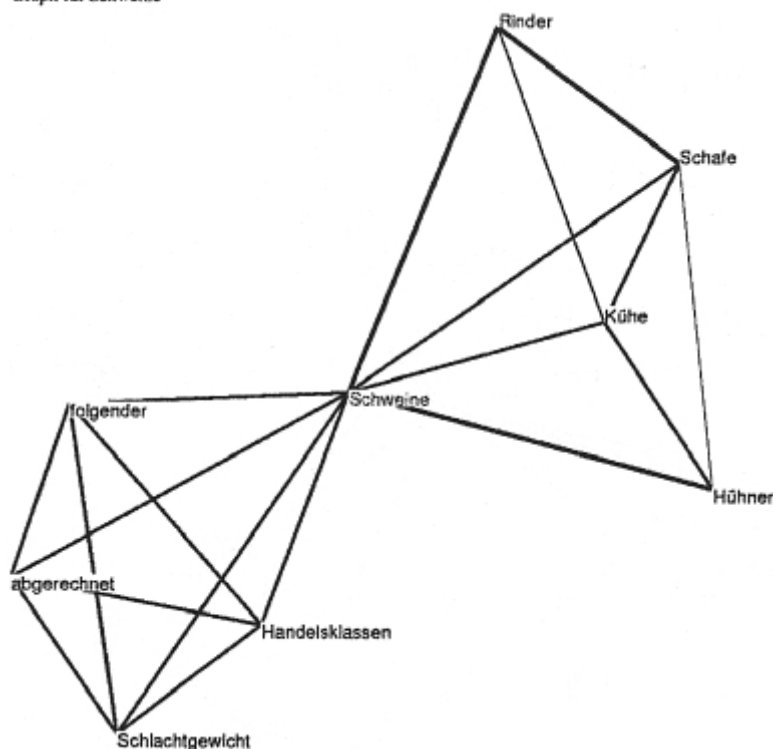
**Figure 4:** Visualisation of the collocation set for Rice

## Separation of Semantic Relations

The calculation of collocations is based on simple statistical measures and does not explicitly *name* the type of semantic relation which holds for a given collocator and its collocates. A further processing of collocation sets is highly desirable, though.

Collocation sets and their visualisation may be employed displaying and partitioning of multiple meanings for single entries. As different meanings of a given word tend to include different subsets of collocation terms which are more closely related to each other, the resulting collocation graph can show a distinct separation of the entire collocation set. The following example shows this for the set of collocations for *Schweine* (pigs): On the right side of the image typical co-hyponyms like *Rinder* (cattle), *Kühe* (cows), *Schafe* (sheep) and *Hühner* (chicken) are displayed, while the left side includes collocations which illustrate aspects of pigs as a food product: *Handelsklassen* (grade of goods), *folgender* (following), *Schlachtgewicht* (weight at slaughtering time), *abgerechnet* (discounted).

Graph für Schweine



**Figure 5:** Collocation Graph for Schweine (pigs)

While this type of visualization does not result in an explicit partitioning of collocation sets, extraction of meaningful subsets may be done by identifying typical representatives for the relations involved: If a typical term for a certain type of relation is known (*extractor*), the intersection of the sets of collocations for the original term and for the extractor results in a subset which contains collocation for the selected relationship. This works fine for sets of collocations in which words are polysemic or refer to concepts as well as to proper names. An example taken from the English corpus shall illustrate this method. The set of collocations for *board* (approx. 90.000 tokens in the corpus, thus frequency class 7) contains terms related to the meaning of board as a component of a Computer as well as for *board* in the sense of a set of people serving a special purpose within an institution or company:

directors (3710)	video (754)	circuit (552)
16-bit (345)	seats (319)	chip (268)
bulletin (1692)	VGA (737)	elected (529)
appointed (338)	trustees (298)	president (263)
chairman (1165)	executive (643)	RAM (462)
expansion (338)	add-in (296)	8-bit (250)
members (944)	shareholders (608)	director (462)
named (335)	slots (289)	proposal (242)
memory (910)	offer (571)	slot (443)
school (330)	shareholder (275)	accelerator (221)
member (890)	company's (570)	approved (436)
across (329)	chief (273)	authorized (220)

meeting (881)	fax (557)
Graphics (324)	Coprocessor (272)

**Table 4:** Top 40 of several hundred significant collocations for board

Given an adequate term representative of one of the typical meanings of board, the set of collocations can be semantically partitioned. The following table shows the collocation subsets for the intersection of collocation sets of *board* and *memory* and *board* and *members* (ordered in decreasing significance).

<i>common collocations for board and memory</i>	<i>common collocation for board and members</i>
upgrade	directors
includes	elected
drive	eight
expansion	meeting
card	proposal
bus	committee
monitor	join
boards	representatives
video	appointed
graphics	vote
processor	seats
cache	voted
chips	elect
PC	membership
controller	
chip	
serial	
slot	
slots	
Intel	
CPU	
sockets	
PS	
RAM	
adapter	
add-in	
VGA	
coprocessor	
motherboard	
SIMMs	

**Table 5:** Selection for collocation subset related to different word meanings

## **Applications**

One major advantage of the infrastructure developed for this project is its immediate portability for different languages, text domains, and applications: The basic structure consisting of text processing tools, data model, and information extraction algorithms may be applied to any given corpus of textual data. This makes this approach applicable to a wide variety of basic language technology problems like

- text classification,
- document management, or
- information retrieval.

Beside the project's WWW interface and its usage as a general purpose dictionary (basic statistical, syntactic and semantic information, typical usage examples), current applications include collocation-based query expansion in Web search engines. The latter shall be discussed in more detail.

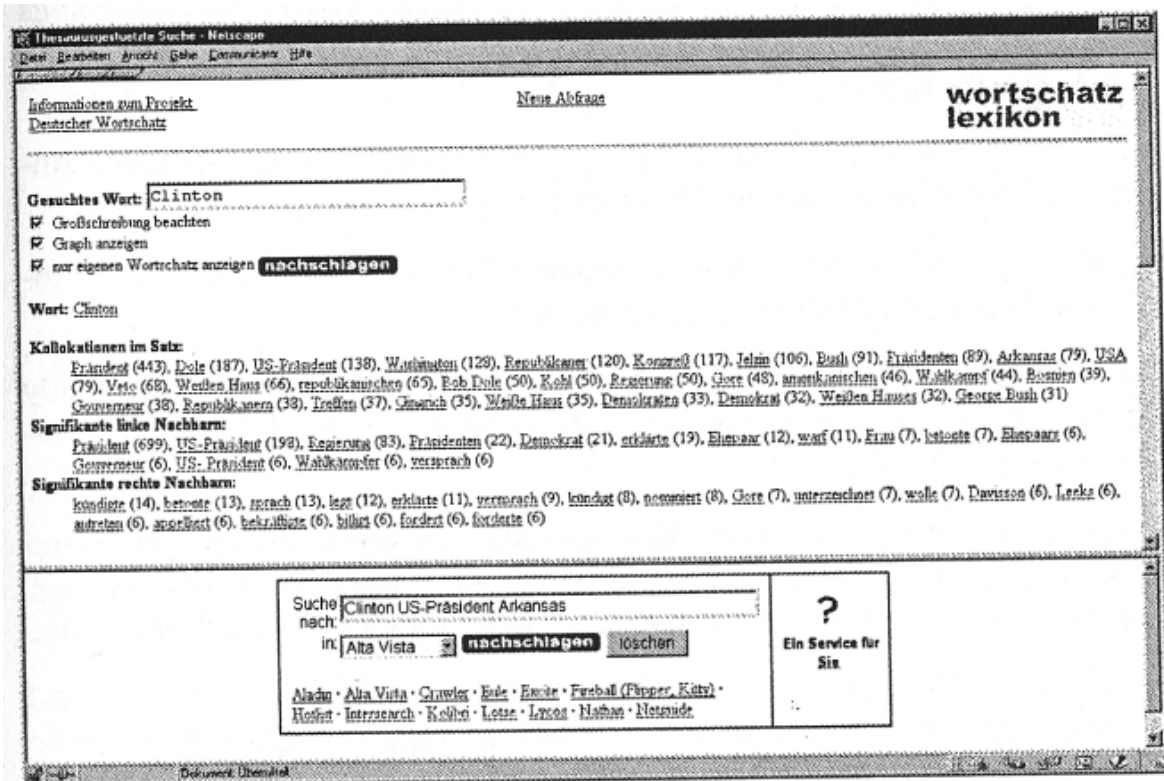
Recent studies (cf. Silverstein et al. 1999, Jansen et al. 2000) have shown that information retrieval on the Web is remarkable different from the interaction with more traditional types of information retrieval engines like bibliographic databases or full text archives. The following list of properties characterises the problem of using web search engines:

- the Web contains mass data (approx. 1 billion documents by the beginning of 2000) with little or no coherent structure
- users are not knowledgeable with respect to information retrieval systems and their interfaces
- queries tend to be very short (less than three terms on average), usage of search operators is an exception and they are often used in a wrong way
- searches tend to result in large document sets which are evaluated only partially by the searchers.

While there are a number of approaches for an optimisation of this situation, like improving retrieval models, the interfaces to search engines or using clustering and filtering techniques, we concentrate an user-driven query expansion as a technique for which the results of our information extraction tools may be employed:

For every search term the user may review collocations and synonyms and add them to his query definition before submitting it to a search engine. We have developed a search engine interface (see Fig. 6) which allows the user to start from a single search term and select additional query terms from information available

in the corpus (collocations, synonyms etc.). A simple mechanism for expanding the query is implemented using JavaScript and dynamically generated hyperlinks.



**Figure 6:** Search Interface for Web Search Enhancement Using Collocations

## Outlook

While the web search application makes use of our "standard" data corpus, the infrastructure can be applied to new and different data sets or text collections without modification. Thus, further applications like comparing special purpose document collections with the general language corpus are possible. The difference in the statistical data can help identifying important concepts and their relations. Applications of this analysis are, amongst others,

- domain specific terminology extraction and
- support of object oriented modeling of business processes.

In the latter example, business reengineering according to the methods proposed by Ortnner (cf. Ortnner 1997) is supported by generating significant semantic relations from software documentation for further use in modeling object-oriented software models.



## References

- Brants, T. (2000). TnT - A Statistical Part-of-Speech Tagger. In Proceedings of the Sixth Applied Natural Language Processing Conference ANLP-2000, Seattle, WA [to appear].
- Davidson, R., Harel, D., 1996. Drawing Graphs Nicely Using Simulated Annealing, *ACM Transactions on Graphics* 15(4), 301-331.
- Guenther, F. (1996), Electronic Lexica and Corpora Research at CIS. In International Journal of Corpus Linguistics 1(2).
- Jansen, B. J. et al. (2000), Real Life, Real Users, and Real Needs: A Study and Analysis of User Queries on the Web. In *Information Processing & Management* 36(2), 207-227.
- Läuter, M., Quasthoff, U. (1999), Kollokationen und semantisches Clustering. In Gippert, J. (ed.) 1999. *Multilinguale Corpora. Codierung, Strukturierung, Analyse. Proc. 11. GLDV-Jahrestagung*. Prague: Enigma Corporation, 34-41.
- Lemnitzer, Lothar (1998). "Komplexe lexikalische Einheiten in Text und Lexikon." In: Heyer, Gerhard; Wolff, Christian (edd.). *Linguistik und neue Medien*. Wiesbaden: Dt. Universitätsverlag, 85-91.
- Ortner, Erich (1997). *Methodenneutraler Fachentwurf*. Stuttgart & Leipzig: Teubner.
- Quasthoff, Uwe. 1998A. Tools for Automatic Lexicon Maintenance: Acquisition, Error Correction, and the Generation of Missing Values." In: Proc. First International Conference on Language Resources & Evaluation [LREC], Granada, May 1998, Vol. II, 853-856.
- Quasthoff, Uwe. 1998B. Projekt der deutsche Wortschatz. In Heyer, G., Wolff, Ch. (eds.). *Linguistik und neue Medien*. Wiesbaden: Dt. Universitätsverlag, 93-99.
- Ruge, Gerda (1994). *Wortbedeutung und Termassoziation. Methoden zur automatischen semantischen Klassifikation*. Hildesheim & New York: Olms.
- Silverstein, C. et al. (1999), Analysis of a Very Large Web Search Engine Query Log. In *SIGIR Forum* 33(1), 6-12.
- Voorhees, E.; Harman, D. (eds.) 1999. Overview of the Seventh Text RE-trieval Conference (TREC-7). In Voorhees, E.; Harman, D. (eds.), Proc. TREC-7. The Seventh Text Retrieval Conference. Gaithersburg/MD: NIST [- NIST Special Publication 500-242].



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 179 – 194

## **Die virtuelle Fachbibliothek Sozialwissenschaften**

**Matthias N.O. Müller**  
IZ Sozialwissenschaften  
Lennstraße 30  
D-53113 Bonn  
mr@bonn.iz-soz.de

**Wolfgang Meier**  
**Stefan Winkler**  
TU Darmstadt  
Institut für Soziologie  
Residenzschloss  
D-64283 Darmstadt  
meierlwinkler@ifs.tu-darmstadt.de

### **Zusammenfassung**

Das DFG-Projekt "Virtuelle Fachbibliothek Sozialwissenschaften" (ViBSoz) zielt auf die integrierte Bereitstellung sozialwissenschaftlicher Literaturinformationen aus verteilten, verschieden strukturierten und inhaltlich unterschiedlich erschlossenen Datenbeständen, die sich in miteinander nicht verbundenen, heterogenen Organisationsstrukturen und Zugänglichkeitskontexten befinden. Schwerpunkte bei den zu lösenden Problemen sind der Aufbau einer geeigneten Systemarchitektur als Voraussetzung für höhere Dienste, eine Lösung der inhaltlichen Heterogenitätsproblematik mit Hilfe von statistischen und intellektuellen Transferkomponenten sowie die Schaffung einer benutzerorientierten Oberfläche, die der verteilten Struktur Rechnung trägt.

### **Abstract**

The Social Science Virtual Library project aims at presenting an integrated view to distributed, heterogeneous data of social science literature. The main emphasis has been put on solving problems of access to such diverse document sets. As a prerequisite for higher services like this, an adequate system architecture has to be implemented. The heterogeneity in content will be solved by transfer components, which will realize a switching of vocabulary. Last but not least a user oriented interface will be implemented, which takes the diversity of the given data into consideration.



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

## 1. Einleitung

Das DFG-Projekt "Virtuelle Fachbibliothek Sozialwissenschaften" (ViBSoz) zielt auf die integrierte Bereitstellung sozialwissenschaftlicher Literaturinformationen aus verteilten, verschieden strukturierten und inhaltlich unterschiedlich erschlossenen Datenbeständen, die sich in miteinander nicht verbundenen, heterogenen Organisationsstrukturen und Zugänglichkeitskontexten befinden (Institutsbibliotheken, Sondersammelgebiete der Universitätsbibliotheken (SSG), wissenschaftliche Spezialbibliotheken, Referenzdatenbanken, digitale Volltexte).

Mit Hilfe der Virtuellen Fachbibliothek soll dem Benutzer somit die Möglichkeit gegeben werden in nur *einem* Suchvorgang in unterschiedlichen, verteilten Datenbeständen zu recherchieren. Dazu soll er die ihm vertraute Sacherschließung nutzen können, d.h. das System wird in der Lage sein, die vom Benutzer gestellte Anfrage adäquat in andere Sacherschließungen umzusetzen. Eine geeignete flexible Architektur vorausgesetzt, sind vorrangig zwei Problembereiche zu bearbeiten:

- Lösung der Heterogenitätsproblematik mit Hilfe von statistischen, intellektuellen und neuronalen Transferkomponenten
- eine benutzerorientierte Oberfläche, die der verteilten Struktur Rechnung trägt und die Rechercheintelligenz für den Benutzer transparent macht.

ViBSoz ist ein Gemeinschaftsprojekt des Informationszentrums Sozialwissenschaften, Bonn (Prof. Dr. J. Krause) und des Instituts für Soziologie der TU Darmstadt (Prof. Dr. R. Schmiede). Weitere Kooperationspartner sind das Sondersammelgebiet Sozialwissenschaften der Universitäts- und Stadtbibliothek Köln (USB Köln), die Bibliothek der Friedrich-Ebert-Stiftung (FES), Bonn, das Zentrum für Interdisziplinäre Technikforschung (ZIT) der TU Darmstadt, sowie der Westdeutsche Verlag, Wiesbaden und der Leske + Budrich Verlag, Leverkusen-Opladen. Im Oktober 1999 hat das Wissenschaftszentrum Berlin (WZB) Interesse geäußert, sich ebenfalls an dem Projekt zu beteiligen und ist mit seinen Datenbeständen einbezogen worden.

## 2. Heterogene Datenbestände

Zwar bemühen sich die Projektpartner, eine weitmögliche Vereinheitlichung bei den verwendeten Datenbanksystemen bzw. Datenmodellen zu erreichen, realistischerweise wird jedoch prinzipiell von einem heterogenen Umfeld in diesem Bereich ausgegangen, um größtmögliche Offenheit und Erweiterbarkeit des Systems zu gewährleisten.

Die Datenbestände der Projektpartner verbleiben vor Ort, d.h. es gibt keinen zentralen Datenserver, Zudem soll die Möglichkeit bestehen, die lokal bereits existierenden Retrievalsysteme zu nutzen. Gegenwärtig sind im Projekt folgende Datenquellen verfügbar:

- Die Datenbestände des Darmstädter Virtuellen Gesamtkatalogs (Monografien und Graue Literatur) waren zunächst aus dem Allegro-Bibliothekssystem in eine relationale Datenbank (Oracle) zu überführen. Es handelt sich dabei um einen OPAC mit den Beständen von über 60 Bibliotheken (490.000 Datensätze), darunter qualitativ hoch erschlossene Fachbibliotheken, wie die des ZIT.
- Die Oracle-Datenbanken des IZ Sozialwissenschaften, Bonn (SOLIS und FORTS) enthalten ca. 300.000 Monografien, mit Grauer Literatur, Zeitschriftenaufsätzen und Sammelwerksbeiträgen.
- Der SISIS OPAC der USB Köln (Gesamtbestand ca. 1,5 Mio) verweist mit ca. 700.000 Einträgen auf Dokumente des Sondersammelgebiets Sozialwissenschaften.
- Der OPAC der FES-Bibliothek verzeichnet 500.000 Dokumente (Monografien, Graue Literatur und Zeitschriftenaufsätze in mehreren Sprachen), die über ein Allegro-System zugänglich sind. Hinzu kommen 600 eigene Publikationen als elektronische Volltexte.
- Im OPAC des WZB werden insgesamt 278.700 Dokumente nachgewiesen. Als Information-Retrieval-System wird BRS/Search der Firma IHS verwendet, für das auch eine Z39.50 Schnittstelle zur Verfügung steht.

Hinsichtlich der im Projekt vertretenen unterschiedlichen Sacherschließungen siehe Kapitel 5.1

### **3. Systemarchitektur**

Die technische Implementierung eines verteilten Retrieval-Systems muss die für das Projekt zentralen Überlegungen zur Heterogenitätsbehandlung berücksichtigen. Die Programmlogik muss in der Lage sein, zwischen der Benutzeranfrage und den jeweiligen Informationsquellen zu vermitteln, wobei deren formale und inhaltliche Unterschiede zu berücksichtigen sind.

Zwei grundlegende Ziele sind daher:

- dass der Benutzer seine Anfrage in einer allgemeinen und adäquat zu bedienenden graphischen Benutzungsoberfläche formulieren kann, ohne im Einzelnen über die formalen und inhaltlichen Besonderheiten der angesprochenen Quellen informiert sein zu müssen.
- dass weitere Informationsquellen jederzeit in die Architektur aufgenommen werden können. Die Schnittstellen zwischen ViBSoz und den Systemen der teilnehmenden Projektpartner müssen deshalb allgemein anerkannten Standards genügen.

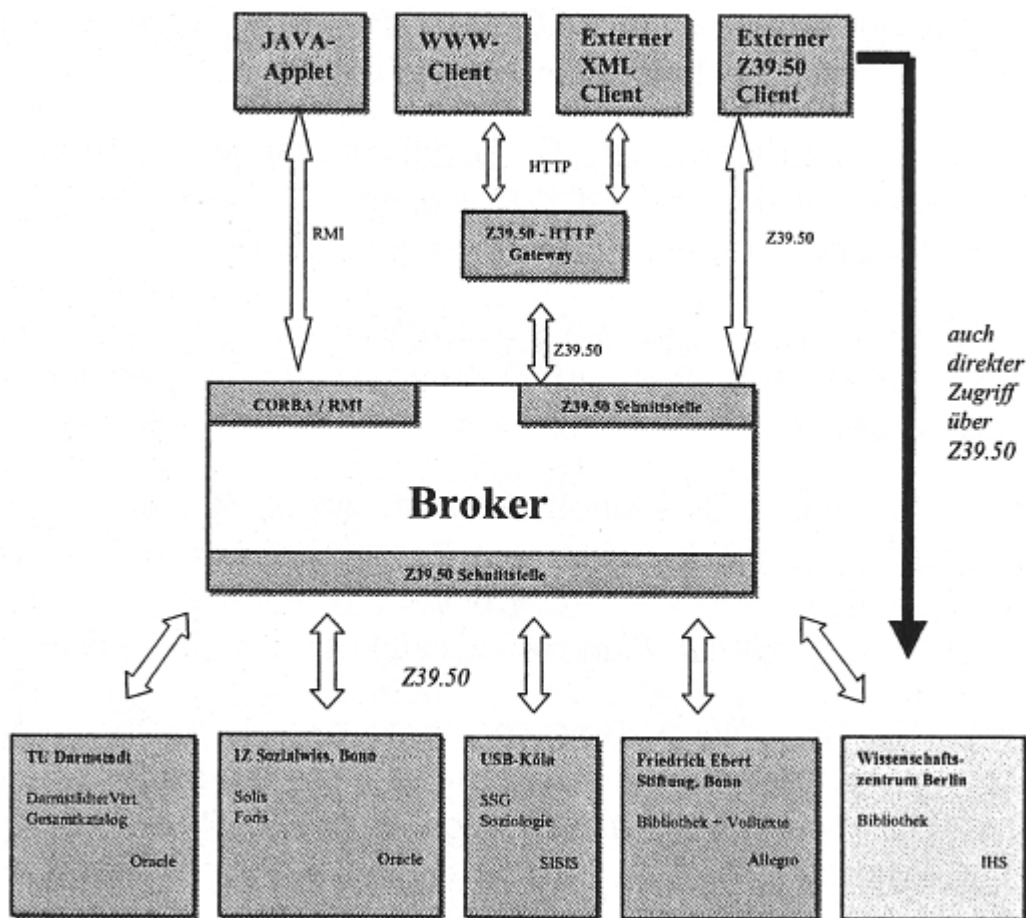


Abbildung 1: Systemarchitektur

Die Projektpartner halten ihre Datenbestände jeweils auf unterschiedlichen Informationssystemen. Vertreten sind relationale Datenbanken, verschiedene Bibliothekssysteme und Volltextdatenbanken (siehe Kapitel 2). Um erweiterbar zu sein, muss die Architektur deshalb von den speziellen Eigenschaften und Möglichkeiten des betreffenden Systems abstrahieren. Damit auf diese Systeme über eine einheitliche Schnittstelle zugegriffen werden kann, wird das im Bibliotheksbereich verbreitete Z39.50 Protokoll als Kommunikationsprotokoll verwendet.

IZ Bonn und TU Darmstadt halten ihre Daten auf Oracle Datenbanken. Um auf diese Daten über Z39.50 zugreifen zu können, wurde von uns ein eigenes Z39.50-Server-Modul implementiert. Dieses dient zunächst dazu, ankommende Z39.50-Anfragen auf die entsprechenden SQL-Kommandos umzusetzen, bzw. Datensätze in verschiedenen Formaten (UNIMarc, MAB, XML) aus der Datenbank auszugeben. Der Z39.50-Server wurde so implementiert, dass er auch außerhalb des Projekts von Nutzen sein kann, sowohl, um andere Datenbanksysteme mit einer Z39.50-Schnittstelle auszustatten, wie auch, um eine Übersetzung zwischen Z39.50 und anderen Protokollen zu gewährleisten.

An der TU Darmstadt wurde ein erster Prototyp der Broker-Komponente fertiggestellt (siehe Abb. 1). Realisiert sind bereits die parallele Abfrage der Targets und die Zusammenführung der Ergebnisse.

Da die teilnehmenden Server unterschiedliche Datenformate (Marc, MAB, XML) unterstützen, werden alle Titelsätze zunächst in ein internes Format transformiert. Dieses basiert auf XML, da somit die Verarbeitung möglichst vieler unterschiedlicher Datensatzstrukturen gewährleistet ist. Der Broker muss lediglich wissen, wie er die für Sortierung und Dublettenkontrolle notwendigen Informationen (Titel, Autor, ISBN etc.) aus dem jeweiligen Datensatz gewinnen kann. Die Konvertierung von XML nach Marc, MAB etc. und umgekehrt wird durch editierbare Stylesheets abgearbeitet, wodurch Anpassungen ohne Änderungen am Programmcode möglich sind.

Aus Sicht der Clients ist der Broker ein „normaler“ Z39.50-Server und kann von jeder Z39.50-fähigen Software angesprochen werden. Durch die offene Architektur können die Benutzer zwischen diversen Möglichkeiten des Zugriffs wählen, z.B. über eine Java-Oberfläche, ein Web-Interface, spezielle Z39.50-Clients oder über Bibliothekssoftware mit entsprechender Schnittstelle.

## 4. Benutzungsoberflächen

### 4.1 Java Oberfläche

**Abbildung 2: Formulareingabe**

**Abbildung 3: Eingabe-Grid**

Der primäre Zugang zu ViBSoz wird über eine Java-Applikation erfolgen. Die Verwendung von Java eröffnet gegenüber einem einfachen HTML Zugang alle Möglichkeiten einer eigenständigen Applikation, wie z.B. eine detailreiche, dynamische Oberfläche oder das Speichern und Laden einer Anfrage und deren Ergebnis. Die Umsetzung basiert auf dem WOB-Modell (Krause 1997), welches grundlegende softwareergonomische Prinzipien umsetzt. Dadurch wird eine dem

Problem angemessene und intuitive Benutzung gewährleistet. Der vorliegende Entwurf dient als Arbeitsgrundlage um zu prüfen, in wie weit die Anforderungen des WOB-Modells und die Bedürfnisse der Domäne 'Literatursuche in verteilten Datenbanken' mit Java umgesetzt werden können. Er spiegelt also nur das Grundgerüst der entgültigen Oberfläche wider.

#### **4.1.1 Suchfenster**

Das Suchfenster gliedert sich in die drei Bereiche:

- Filter
- graphisch orientierten Eingaben
- formalsprachliche Eingaben

Die im Suchfenster oben liegenden Filter dienen der vorherigen Festlegung, in welchen Datenbeständen gesucht werden soll. Hierzu zählen die zu durchsuchenden Datenbanken, die für den Benutzer relevanten Dokumenttypen (Monographien, Sammelwerke, usw.) und die Sprache der einzelnen Werke. Hinzu kommt die Angabe, welche Sacherschließung für die Anfrageformulierung benutzt werden wird<sup>1</sup>. Der Benutzer kann somit einerseits vorab den Suchraum auf die Dokumente einschränken, die für ihn von Interesse sind, andererseits kann er auch während des Suchprozesses den Suchraum erweitern oder weiter einschränken, ohne die Suchanfrage selbst verändern zu müssen. Die Formulierung der eigentlichen Suchanfrage geschieht wahlweise auf graphisch, formularorientierte Weise oder formalsprachlich.

Zu den graphisch orientierten Eingaben zählen die Formulareingabe (Abbildung 2, Mitte) und das Eingabe-Grid (Abbildung 3, Mitte). Die aus der Formalerschließung bekannten Felder wie Autor/Herausgeber, Titelstichwort, Erscheinungsjahr oder Verlag können über das Formular gesucht werden.

Für die Formulierung einer Schlagwort- und/oder einer Klassifikationssuche können zwei Grids genutzt werden. Durch vorgegebene syntaktische Strukturen erlauben sie die einfache Formulierung einer Boole'schen Anfrage, wodurch der Großteil der von Benutzern formulierten Anfragen abgedeckt wird.

Syntaktisch komplexere Anfragen können in der formalsprachlichen Anfrage durch verschiedene formale Sprachen formuliert werden. Vorgesehen sind die formale Sprache des SISIS Opacs, welche die Benutzer der USB Köln und zum Teil die Benutzer aus Darmstadt gewöhnt sind, und die Sprache MESSENGER des STN<sup>2</sup>, die Kunden des IZ -Sozialwissenschaften vertraut ist.

---

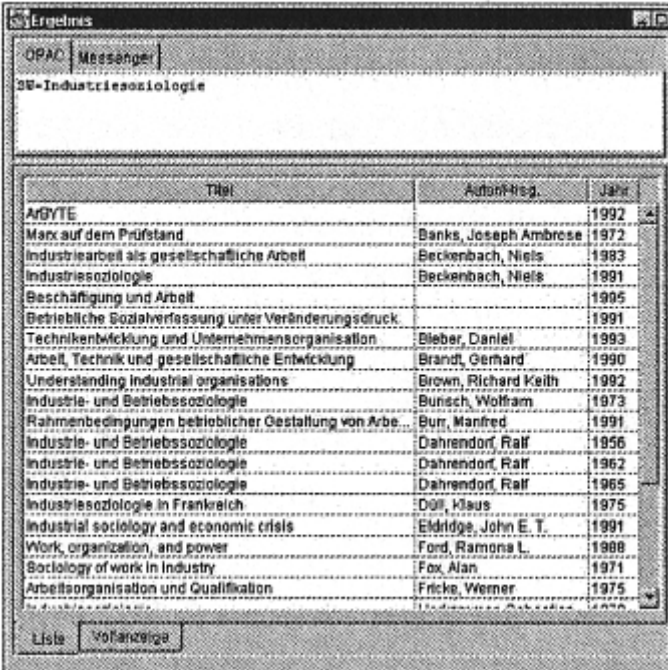
<sup>1</sup> Mehr zur Sacherschließung siehe Abschnitt 5.1.

<sup>2</sup> The Scientific & Technical Information Network. URL: <http://www.fizkarlsruhe.de/stn.html>

Alle drei Eingabeebenen sind durch das Prinzip der kontextsensitiven Durchlässigkeit<sup>3</sup> miteinander verbunden, d.h. Eingaben in eines der Elemente werden systemgesteuert in die anderen Elemente übertragen und gegebenenfalls transferiert. Eingaben in das Formular oder eines der Grids z.B. werden automatisch in eine formalsprachliche (Teil-)Anfrage übersetzt und im formalsprachlichen Eingabefenster angezeigt und vice versa. Der Benutzer kann also z.B. den Grundstock seiner Anfrage bequem über die graphisch orientierten Eingaben aufbauen, und diesen dann formalsprachlich präzisieren.

#### 4.1.2 Ergebnisdarstellung

Die Ergebnisdarstellung gliedert sich in zwei Bereiche, den formalsprachlichen Dialog und die eigentliche Ergebnispräsentation.



The screenshot shows a window titled 'Ergebnis' with a sub-header 'OPAD Message' and '3E-Industrie- und Betriebssoziologie'. Below this is a table with three columns: 'Teil', 'Autor-Frag.', and 'Jahr'. The table lists various works related to industrial sociology and labor sociology, including titles like 'ARGYTE', 'Max auf dem Prüfstand', 'Industriearbeit als gesellschaftliche Arbeit', etc., along with authors and years.

Teil	Autor-Frag.	Jahr
ARGYTE		1992
Max auf dem Prüfstand	Banks, Joseph Ambrose	1972
Industriearbeit als gesellschaftliche Arbeit	Beckenbach, Niels	1993
Industrie- und Betriebssoziologie	Beckenbach, Niels	1991
Beschäftigung und Arbeit		1995
Betriebliche Sozialverfassung unter Veränderungsdruck		1991
Technikentwicklung und Unternehmensorganisation	Bleber, Daniel	1993
Arbeit, Technik und gesellschaftliche Entwicklung	Brandt, Gerhard	1990
Understanding industrial organisations	Brown, Richard Keith	1992
Industrie- und Betriebssoziologie	Bunisch, Wolfram	1973
Rahmenbedingungen betrieblicher Gestaltung von Arbe...	Burr, Manfred	1991
Industrie- und Betriebssoziologie	Dahrendorf, Ralf	1956
Industrie- und Betriebssoziologie	Dahrendorf, Ralf	1962
Industrie- und Betriebssoziologie	Dahrendorf, Ralf	1965
Industrie- und Betriebssoziologie in Frankreich	Doll, Klaus	1975
Industrial sociology and economic crisis	Eldridge, John E. T.	1991
Work, organization, and power	Ford, Ramona L.	1988
Sociology of work in industry	Fox, Alan	1971
Arbeitsorganisation und Qualifikation	Fricke, Werner	1975

Abbildung 4: Ergebnisfenster

Der formalsprachliche Dialog gleicht exakt dem aus der Suchanfrage, d.h. er dient hier einerseits als Zustandsanzeige über die bislang erstellte Anfrage<sup>4</sup>, andererseits kann der Experte hiermit die Anfrage umformulieren und das Ergebnis der Umformulierung gleich bewerten. Eine solche editierbare Zustandsanzeige erspart also im iterativen Retrieval den zusätzlichen Interaktionsschritt der Rückkehr zur Anfrageformulierung und den erneuten Wechsel zur Ergebnisanzeige.

Gerade der Bereich der eigentlichen Ergebnisdarstellung ist noch nicht im Detail entworfen, da das Vorgehen und die Einflussmöglichkeiten auf das Ranking noch nicht endgültig geklärt sind, sie jedoch Auswirkungen auf die Darstellung des Ergebnisses haben.

<sup>3</sup> Mehr zum Thema kontextsensitive Durchlässigkeit siehe zum Beispiel (Müller 1999).

<sup>4</sup> Zur Funktion der Zustandsanzeige im WOB-Modell siehe zum Beispiel (Marx 1996).



Sicher festgelegt werden kann aber schon die Aufteilung in eine Listendarstellung und eine Detailansicht. In der Listendarstellung werden alle Treffer tabellarisch zusammengestellt. Der Benutzer kann sich somit einen Überblick über die Treffermenge und deren Inhalt machen. Für einzelne Dokumente wird in die detailliertere Vollansicht gewechselt. Dort wird alle verfügbare Information über ein einzelnes Dokument angezeigt. Dazu zählen neben der Formalerschließung z.B. die unterschiedlichen Sacherschließungen und - falls vorhanden - ein Verweis auf den Volltext.

## **4.2 Z39.50-Web-Gateway**

Um gelegentlichen Benutzern, die die Installation und Einarbeitung in die funktionsreichere Java Oberfläche scheuen, einen einfachen Zugang zu ermöglichen, wurde ein rein XML basierter Zugang geschaffen.

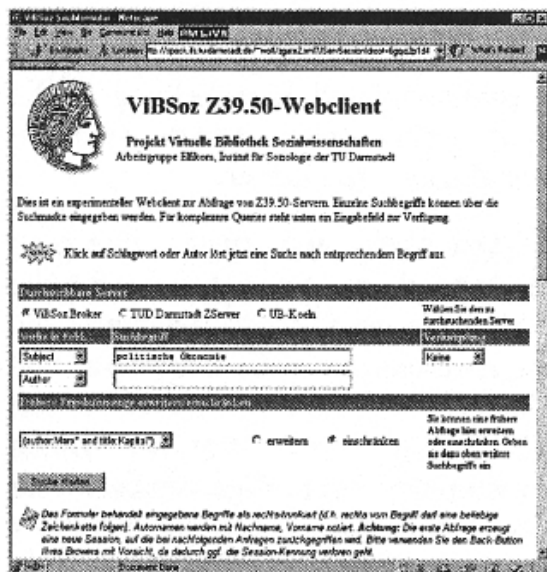
Da der in ViBSoz verwendete Broker die Datensätze wie erwähnt intern im XML-Format verarbeitet (siehe Kap. 3), kann dieses Format auch für die externen Schnittstellen genutzt. Diese Variante ermöglicht einerseits die Kommunikation mit in naher Zukunft zu erwartenden XML-Browsern, und andererseits die Schaffung eines einfachen HTML Zugangs mit Hilfe von Standard Technologien. Für den HTML-Zugang werden die XML-Daten serverseitig per XSLT nach HTML konvertiert. Zukünftige Webbrowser werden diesen Schritt nicht mehr benötigen.

Generell haben Systeme, die verschiedene Varianten von Benutzungsoberflächen anbieten, das Problem, dass Benutzer potentiell zwischen den beiden Zugängen wechseln können. Deshalb müssen die Grundprinzipien der Bedienung miteinander kompatibel sein. Für ViBSoz bedeutet dies, dass das Design des Z39.50-Webclients und der Java Applikation aufeinander abgestimmt sein müssen. Diese Abstimmung erfolgt auf der Basis der ersten Benutzertests im zweiten Projektabschnitt.

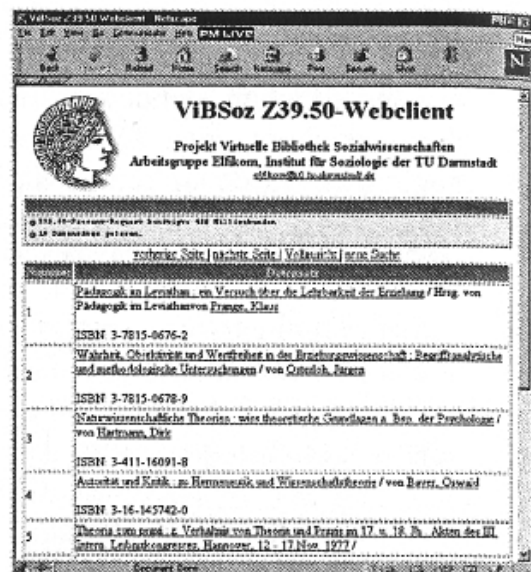
Das Z39.50-Web-Gateway implementiert einen standardkonformen Z39.50-Client, der als XML-/Web-Schnittstelle zu jedem beliebigen Z39.50-Server dienen kann. Die Verwendung von XML ermöglicht eine größtmögliche Trennung von Form, Inhalt und Funktion. Das Layout der Suchseite, sowie der verschiedenen Datenansichten ist vollständig veränderbar, da es gänzlich über XSL-Stylesheets<sup>5</sup> definiert wird.

---

<sup>5</sup> Siehe: W3C Working Draft (2000). Extensible Stylesheet Language (XSL) Version 1.0, 27 March 2000, <http://www.w3.org/TR/xsl/>



**Abbildung 5:** Suchmaske des WebGateways



**Abbildung 6:** Kurzanzeige der Trefferliste

Queries können derzeit wahlweise über einfache HTML-Forms oder eine den populären Internet-Suchmaschinen (Altavista „Advanced Search“) nachempfundene Query-Language eingegeben werden. Sie ermöglicht Boole'sche Algebra, Feldsuche, Trunkierung und beliebig viele Klammerebenen. (vgl. Abbildung 5). Die Trefferlisten können mit Namen versehen werden, die wiederum mit neuen Suchbegriffen kombiniert werden können.

## 5. Behandlung semantischer Heterogenität

Neben der strukturellen Vielfalt ist im Bibliotheksbereich mit einer bereiten Palette verwendeter Sacherschließungssysteme zu rechnen. Ohne geeignete intelligente Verfahren zum Transfer zwischen diesen muss das Rechercheergebnis zwangsläufig unzureichend bleiben.

### 5.1 Sacherschließungen

Im Projekt Virtuelle Fachbibliothek Sozialwissenschaften ist eine Vielzahl von Sacherschließungen vertreten. Das Spektrum reicht von allgemeinen Regelwerken, wie der Schlagwortnormdatei (SWD), über fachspezifische, wie dem Thesaurus Sozialwissenschaften (IZ-Thesaurus), bis zu nicht reglementierten freien Schlagwörtern. Gleiches gilt für die Klassifikationen, die mit verschiedenen Ausprägungen der Basisklassifikation (BK) als Allgemeinklassifikation, der Klassifikation Sozialwissenschaften (IZ-Klassifikation) als Fachklassifikation und den weniger spezifischen Aufstellsystematiken im Projekt vertreten sind. Die Sacherschließungen der Beteiligten bieten also einen guten Querschnitt der in der

Bibliothekslandschaft vorkommenden Sacherschließungen. Eine große Übertragbarkeit der Ergebnisse des Projekts ist somit gewährleistet.

Im Bereich der Verbalerschließung bilden die Schlagwortnormdatei, der Thesaurus Sozialwissenschaften und das von der FES verwendete Regelwerk einen Schwerpunkt. Im Hinblick auf eine Integration sind hier vor allem das Zusammenspiel bzw. die Gegensätze zwischen der SWD als allgemeinem Sammelwerk und dem IZ-Thesaurus als fachspezifischem Sammelwerk von Interesse. Auch das System der FES, als Sonderfall eines eigenen, auf die Bedürfnisse einer einzelnen Institution abgestimmten Systems, benötigt besonderes Augenmerk. Im Bereich der Klassifikation bilden die Basisklassifikation, mit seinen beiden Ausprägungen PICA/GVK und PICA/HEBIS, sowie die Klassifikation Sozialwissenschaften den zweiten Schwerpunkt. Auch hier findet sich wieder das Problem der sehr groben Allgemeinklassifikation gegenüber der detailreichen, stark fachbezogenen Klassifikation.

## **5.2 Transfermodule**

### **5.2.1 Problembereich**

Im vorherigen Abschnitt wurde die Vielfältigkeit der Sacherschließungen innerhalb des Projekts vorgestellt. Diese Vielfalt an Systemen für die inhaltliche Beschreibung von Dokumenten erzeugt für den Benutzer Probleme bei der Suche nach Information. Möchte er in allen beteiligten Beständen eine tiefergehende Suche durchführen, so war er bisher gezwungen, alle verwendeten Sacherschließungssysteme zu erlernen. Er kann nicht einfach das ihm vertraute System auf die anderen Bestände übertragen. Würde zum Beispiel ein Benutzer der USB Köln mit dem SWD Schlagwort 'Berufliche Fortbildung' in der Datenbank SOLIS des IZ-Sozialwissenschaften suchen, so würde er keine Treffer erhalten, denn dort ist der Begriff als 'Berufliche Weiterbildung' vergeben worden. Ebenso verhält es sich mit vielen Eigennamen und Bezeichnungen, wie z.B. der 'Sowjetunion' die in SOLIS unter 'UdSSR' zu finden ist. Besonders problematisch sind mehrdeutige Begriffe wie 'Haushalt'. Er kann einerseits das Budget andererseits den Haushalt im Privaten bezeichnen. In der USB wird er als Haushalt im Privatbereich verstanden. Um diese Lesart abzudecken, muss beim IZ-Sozialwissenschaften mit dem Schlagwort 'Privathaushalt' gesucht werden, andernfalls erhält der Benutzer nicht die gewünschte Information. Bei steigendem Recall sinkt die Precision.

Aufgabe des zu entwickelnden Systems ist die Unterstützung des Benutzers bei der "Übersetzung" von einem Sacherschließungssystem zum anderen. Idealerweise geschieht diese Übersetzung innerhalb des Systems, so dass der Benutzer zwar auf Wunsch darüber informiert wird, nicht aber selbst eingreifen muss. Er kann dann seine Anfrage an das Gesamtsystem in seiner gewohnten Weise formulieren, und die Umsetzung bzw. Anpassung an die verschiedenen beteiligten Systeme erfolgt automatisch.

### **5.2.2 Cross-Konkordanzen**

Eine Möglichkeit eine solche Umsetzung zu realisieren ist die Verwendung von Cross-Konkordanzen. Dabei vergleichen Fachreferenten intellektuell jeweils zwei Sacherschließungssysteme und definieren Übereinstimmungen zwischen diesen. Somit können neben den Synonymierelationen auch Oberbegriffs- / Unterbegriffs- und Ähnlichkeitsrelationen berücksichtigt werden.

Im Projekt wird eine (partielle) intellektuelle Cross-Konkordanz zwischen der SWD und dem Thesaurus Sozialwissenschaften sowie der Basisklassifikation und der Klassifikation Sozialwissenschaften erarbeitet.

### **5.2.3 Quantitativ-Statistische Verfahren**

Eine andere Möglichkeit einer solchen Umsetzung ist der Einsatz quantitativ-statistischer Verfahren, die den Transfer zwischen der Sacherschließung der Benutzereingabe zu den Sacherschließungen der anderen Datenquellen berechnen. Dabei werden die jeweiligen Begriffspaare nicht nach qualitativen Maßstäben erzeugt, wie bei der intellektuellen Erstellung, sondern nach ihrer Quantität bezogen auf einen Korpus, also nach der Häufigkeit ihres Vorkommens. Vereinfacht ausgedrückt: Je häufiger ein Begriffspaar in einem Korpus vor kommt, desto wahrscheinlicher ist es, dass es sich um eine sinnvolle Verbindung handelt. Hinzu kommen Parameter wie die Größe des Korpus oder die Verteilung der Begriffe innerhalb dessen.

Voraussetzung für diese Verfahren ist ein Korpus, in dem die Dokumente nach beiden Sacherschließungen indexiert sind (Parallelkorpus). Er entsteht durch den Abgleich zweier Korpora und die Extraktion von Paaren gleicher Dokumente. Jedes Dokument (-paar) ist dann nach zwei Erschließungssystemen indexiert.

Die Entwicklung der Verfahren zur Erstellung der Parallelkorpora für ViBSoz ist weitgehend abgeschlossen. Zur Zeit liegt ein Parallelkorpus USB Köln – IZ- Sozialwissenschaften mit ca. 15 Tausend Einträgen vor.

Aus einem solchen Korpus werden dann Relationen zwischen einzelnen oder Gruppen von Schlagwörtern / Klassifikationen abgeleitet. Als Verfahren hierfür werden in ViBSoz vorwiegend statistische Verfahren eingesetzt, experimentell aber auch Neuronale Netze.

#### **5.2.3.1 Statistische Verfahren**

Statistische Verfahren analysieren die Häufigkeit des gemeinsamen Auftretens zweier Schlagwörter<sup>6</sup> innerhalb des Parallelkorpus und übertragen die

---

<sup>6</sup> Im Weiteren wird das Vorgehen bei der Analyse von Paaren einzelner Schlagwörter betrachtet. Gleiches gilt aber äquivalent für Klassifikationen und Gruppen von Schlagwörtern / Klassifikationen.

gewonnenen Relationen später auf neue Dokumente. Vergleicht man dabei zwei Schlagwörter aus unterschiedlichen Sacherschließungen, so lassen sich statistische Abhängigkeiten zwischen den einzelnen Begriffen und somit zwischen den Sacherschließungen ermitteln.

Beispielsweise wurden für das Dokument *"Gysi, Jutta: Familienleben in der DDR, zum Alltag von Familien mit Kindern, Akademie Verlag Berlin, 1989, ISBN 3-05-000771-0."* an der USB Köln die Schlagwörter 'Deutschland <DDR>' und 'Familie' vergeben. Im IZ-Sozialwissenschaften erhielt das gleiche Dokument die Schlagwörter 'Arbeitsteilung', 'Ehe', 'Familie', 'DDR' und 'Partnerschaft'. Bei diesem Dokument treten also unter Anderem das SWD Schlagwort 'Deutschland <DDR>' und das IZ Schlagwort 'DDR' gemeinsam auf. Ist dieses gemeinsame Auftreten auch bei anderen Dokumenten zu beobachten, so handelt es sich hierbei höchst wahrscheinlich um eine sinnvolle Transferbeziehung. Tatsächlich tritt dieses Paar im derzeitigen Parallelkorpus von USB Köln und IZ Sozialwissenschaften noch bei 272 weiteren Dokumenten auf.

Wie aus der Statistik bekannt, reicht diese absolute Häufigkeit jedoch nicht aus, um die Güte einer solchen Transferbeziehung zu bestimmen. Sie muss z.B. mit der Häufigkeit der Vorkommen der Terme oder der Gesamtzahl der Relationen in Bezug gesetzt werden. Dazu werden in ViBSoz die bedingte Wahrscheinlichkeit und der Äquivalenzindex getestet.

Die bedingte Wahrscheinlichkeit ist die Wahrscheinlichkeit mit der ein Begriff B aus der Sachschließung X auftaucht, wenn auch der Begriff A aus der Sacherschließung Y aufgetaucht ist. Bezogen auf unser Beispiel ist es dann die Frage: Wie wahrscheinlich ist es, dass der IZ Term 'DDR' vergeben wurde, wenn die USB Köln den Term 'Deutschland <DDR>' vergeben hat. Verwendung findet dieses Vorgehen z.B. im System AIR/PHYS (Biebricher/Fuhr/Lustig et al. 1988), welches sich mit automatischer Indexierung befasst.

Der Äquivalenzindex hat seinen Ursprung in der Kookkurrenzanalyse und findet heute vielfältige Verwendung. Er bezieht, im Gegensatz zur bedingten Wahrscheinlichkeit, auch die Vorkommenshäufigkeit des zweiten Terms in die Berechnung mit ein.

Beiden Verfahren gemeinsam ist das Problem, dass der richtige Schwellwert gefunden werden muss, ab dem eine statistische Konkordanz als sinnvoll angesehen werden kann. Dieser Wert ist abhängig von dem verwendeten Verfahren und dem zugrundeliegenden Korpus. Er muss also für alle Sacherschließungspaare einzeln gefunden werden.

Zur Zeit werden die beiden Verfahren anhand des Parallelkorpus USB Köln — IZ-Sozialwissenschaften geprüft und es werden Vorgehen entwickelt, die entsprechenden Schwellwerte zu bestimmen. Erste Ergebnisse, wie bei der IuK-Tagung im März 2000 präsentiert, sind sehr vielversprechend.

### **5.2.3.2 Neuronale Netze**

Die Ausgangsbasis für statistische Verfahren und neuronale Netze ist die gleiche: ein Parallelkorpus. Im Gegensatz zu statistischen Verfahren arbeiten neuronale Netze aber nicht mit mathematischen Formeln, die statistische Relationen berechnen, sondern sie entstehen durch das Training eines neuronalen Netzes mit den Daten des Parallelkorpus. Dabei werden die Sacherschließungen paarweise an ein Backpropagation-Netz angelegt. Das Netz ‚lernt‘ bei jedem Anlegen die Verbindungen der einzelnen Begriffe und Begriffspaare. Durch häufiges Anlegen bilden sich so zwischen einzelnen Neuronen des Netzes stärkere Verbindungen heraus als bei anderen. Durch die unterschiedliche Stärke der neuronalen Verbindungen ist es in der Lage sich zu ‚merken‘, welche Begriffe mit welchen anderen korrespondieren. Legt man dann später Begriffe aus einem Sacherschließungssystem an das Netz an, so kann es die zugehörigen Begriffe eines anderen Sacherschließungssystems bestimmen.

Das in ViBSoz erprobte System basiert auf dem COSIMIR Modell (Mandl 1998) und zeigt in ersten Voruntersuchungen ermutigende Leistungen.

## **6. Ausblick**

Nächster Schritt bei der Umsetzung der Projektziele ist die rasche Integration und Ergänzung der vorliegenden Teile zu einem funktionsfähigen Prototypen. Auf seiner Basis können die erarbeiteten Verfahren getestet, verbessert und wo möglich generalisiert werden.

Im Bereich des Brokers liegen die Schwerpunkte zum Einen bei der Integration der noch fehlenden Datenbestände des IZ Sozialwissenschaften, der FES und des WZB. Zum anderen müssen die Verfahren zur Ergebniszusammenführung umgesetzt werden.

Im Bereich der Transferkomponenten stehen die Implementierung und Einbettung der erarbeiteten Cross-Konkordanzen und statistischen Transferbeziehungen in das Gesamtsystem an nächster Stelle. Die Verfeinerung der statistischen Verfahren zur Erstellung letzterer und ihre Anwendung auf die noch verbleibenden Sacherschließungen zählen zu den nächsten Schritten.

Somit ist in den kommenden Monaten mit der Fertigstellung eines funktionsfähigen Prototypen zu rechnen, der dann zur Evaluation zur Verfügung steht.

Informationen zum Projekt sind über den WWW Server des Projekts unter dem URL <http://vibsoz.bonn.iz-soz.de/> zu finden. Neben kurzen Beschreibungen der Zielsetzungen des Projekts sind auch Publikationen online abrufbar.

## Literatur

Biebricher, P., Fuhr, N., Lustig, G., et al. (1988). *'The Automatic Indexing System AIR/PHYS - From Research to Application'*. 11th International Conference on Research & Development in Information Retrieval. 1988, Grenoble, France.

Grievet, L., Mutschke, P. und Polanco, X. (1995). *'Thematic Mapping on Bibliographic Databases by Cluster Analysis: A Description of the SDOC Environment with SOLIS'*. Knowledge Organisation 22: 8.

Krause, J. (1997). *'Das WOB-Modell'* Vages Information Retrieval und graphische Benutzeroberflächen: Beispiel Werkstoffinformation. Konstanz: Universitätsverlag. pp. 59--88.

Mandl, T. (1998). *'Learning Similarity Functions in Information Retrieval'*. 6th European Congress on Intelligent Techniques and Soft Computing. , Aachen, Germany.

Marx, J. (1996). *'Bidirektionale Sprache. Faktenrecherche und Informationsdarstellung durch dynamische Erzeugung korrigierbarer Zustandsanzeigen in natürlicher Sprache'*. Hildesheim: Olms Verlag.

Meier, W., Müller, M.N.O. und Winkler, S. (2000). *'Virtuelle Fachbibliothek Sozialwissenschaften: Problembereich und Konzeption'*. Bibliotheksdienst 34: 1236-1244.

Müller, M.N.O. (1999). *'Konsistenzerhaltung in multimodalen Benutzungsoberflächen an Beispiel KO-NEXIS'*. Fachbereich Informatik, Universität Koblenz-Landau.



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 195 – 206

# **Das Dynamische Modell der Informationskompetenz (DYMIK) als Grundlage für bibliothekarische Schulungen**

**Benno Homann**

Universitätsbibliothek Heidelberg, Plöck 107-109,  
69117 Heidelberg, homann@ub.uni-heidelberg.de

## **Zusammenfassung**

Bibliotheken sind Institutionen mit umfangreichem Expertenwissen auf dem Gebiet der neuen Informationssysteme. Sie können mit eigenen Schulungen mitwirken bei der Vermittlung von Informationskompetenz. Der Umfang der zu vermittelnden Kompetenzen erfordert mehrere Schulungsveranstaltungen. Diese müssen inhaltlich und methodisch aufeinander abgestimmt werden. Hierfür eignen sich umfassende Modelle, die die wesentlichen Strukturen und Einflussfaktoren von Informationsprozessen repräsentieren. Ein solches Modell ist DYMIK, das an der Universitätsbibliothek Heidelberg entwickelt wurde, um das modulare Schulungsangebot didaktisch und methodisch zu verbessern und um internetbasierte Tutorials zu erweitern.

## **Abstract**

Libraries are institutions with an extensive cumulation of specialized knowledge in the field of information systems. They are qualified to participate in the mediation process concerning information literacy. The amount of competencies to be taught requires several lessons which are to be coordinated in content and methodology. Comprehensive models representing the essential structures and factors of the information process may become the basis for coordination. A model of this kind is DYMIK which has been developed at the university library of Heidelberg as an instrument for the didactic and methodical improvement of the modular training lessons and the addition of internet-based tutorials concerning information literacy.



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>



## **1. Einleitung**

Durch die elektronische Informations- und Kommunikationstechnologie vollzieht sich derzeit ein Funktionswandel der Bibliothek: die bestandsorientierte Archivbibliothek mutiert zu einem Zentrum für Informationsdienste. Die bisherige Hauptfunktion, Bücher bzw. Informationsmaterialien zu erwerben und für ihre primär lokalen Kunden bereitzustellen, verliert zunehmend an Bedeutung. Die Digitalisierung von Medien und die weltweite Vernetzung reduzieren die Abhängigkeit der Informationssuchenden von lokalen Bibliotheken. Informationen lassen sich möglicherweise schneller von einer weit entfernten Bibliothek als von der lokalen Bibliothek beschaffen. Vor diesem Hintergrund müssen die Bibliotheken auf alternative „Stärken“ und Fähigkeiten rekurrieren und neue Aufgabenfelder entwickeln.

Eine wesentliche Stärke ist das methodische Expertenwissen der Bibliotheken im Umgang mit Informationsfluten. Bibliotheken verfügen über Erfahrungen, Medien in unterschiedlicher Weise zu erfassen, zu selektieren und inhaltlich sowie formal zu erschließen. Diese Erfahrungen beschränken sich nicht mehr nur auf konventionelle Druckmedien, sondern beziehen sich auch auf das riesige Informationsangebot im Internet. Durch den täglichen Umgang mit unterschiedlichen, komplexen, sich häufig ändernden Informationssystemen sind sie mit Suchstrategien und Handhabungstechniken vertraut und können diese effizient nutzen. Durch die Fachreferenten verfügen wissenschaftliche Bibliotheken zusätzlich - zumindest potentiell - über Expertenkenntnisse zur Lösung fachwissenschaftlicher Informationsprobleme.

Eine weitere Stärke der Bibliotheken ist die Kumulation des Informationsangebotes auf unterschiedlichen Medienträgern in einem Gebäude. Damit sind Bibliotheken prädestiniert als ideale Lernorte für das selbständige aber auch für das unterstützte Erarbeiten neuen Wissens. Das selbständige und das selbstgesteuerte Lernen sind zentrale Merkmale des "Lebenslangen Lernens", das von Mitgliedern einer sich ständig wandelnden Informationsgesellschaft erwartet wird. In einer Bibliothek können die Besucher und Lerner aus dem Angebot selbständig auswählen, welche neuen Kompetenzen sie sich aneignen möchten und welche Vermittlungsformen, Buch, Internet, CD, etc., sie für ihr Lernen nutzen wollen.

Die Übernahme von Schulungsfunktionen als neue Bibliotheksaufgabe bedingt allerdings die Entwicklung eines entsprechenden Konzepts. Es ist erforderlich, da die zu vermittelnden Fähigkeiten auf mehrere Veranstaltungen verteilt werden müssen und eine Abstimmung mit anderen Schulungseinrichtungen zu empfehlen ist.

## **2. Anforderungen an eine "Teaching Library"**

Mit der Übernahme von Lehr-/Lernfunktionen entwickeln sich die Bibliotheken zu „teaching libraries“, die insbesondere in den USA schon seit Jahren zum Lehr-

/Lernalltag gehören. Das umfangreiche Angebot und das Potential an Informationsexperten reicht jedoch nicht aus, um diese neue gesellschaftliche „Schulungsfunktion“ zu realisieren. Wie die umfangreiche amerikanische Literatur zu diesem Themenkomplex und die ersten Erfahrungen an einzelnen deutschen Bibliotheken zeigen, müssen weitere Voraussetzungen erfüllt sein, um den Anforderungen einer „teaching library“ gewachsen zu sein. Hierzu zählen insbesondere

- Entwicklung eines Gesamtkonzepts für Schulungen
- Anwendung informationsdidaktischer Methoden
- Bereitstellung von Schulungsräumen
- Qualifizierung von Schulungspersonal

Eine fundamentale Voraussetzung ist die Entwicklung eines Gesamtkonzepts, in dem die Inhalte und Ziele eines Schulungs- oder Lernangebotes festgelegt werden. Die Präzisierung dieser Inhalte ermöglicht auch eine wirksame Außenpräsentation bzw. Werbung für das neue Schulungsangebot einer Bibliothek. Von grundlegender Bedeutung sind hier Modelle der Informationskompetenz und die modulare Strukturierung des gesamten Veranstaltungsangebotes.

Eine weitere Voraussetzung bildet die Anwendung informationsdidaktischer Methoden, wie Mind-Mapping, Suchprotokolle, etc., die sich für die Vermittlung von Fähigkeiten zur effizienten Nutzung von Informationen besonders eignen.

Von zentraler Bedeutung ist auch die Bereitstellung von Schulungsräumen mit entsprechender Größe und Ausstattung sowie die Integration dieser Räume in das Medienangebot einer Bibliothek. Dies bedeutet z.B., dass Möglichkeiten der Gruppenarbeit, der gemeinsamen oder individuellen Arbeit an PCs mit Internetzugang, der Ergebnispräsentation mit modernen Schulungsmethoden und die Nähe zu dem Medienbestand gewährleistet sein sollten.

Eine weitere wichtige Voraussetzung ist die Bereitstellung und Qualifizierung von Schulungspersonal. Expertenkenntnisse sind wichtig, aber sie sind bei unreflektierter Einbringung in Schulungs-/Lernprozesse auch hinderlich bzw. für die Teilnehmer einer Schulung frustrierend. Häufig unterschätzen Experten den Umfang des von ihnen beherrschten Wissens und überfordern die Teilnehmer von Schulungen. Erforderlich sind hier methodisch-didaktische Kenntnisse und Fertigkeiten, die eine sinnvolle Reduktion der Inhalte und lernpsychologische, effektive Vermittlung von Informationskompetenzen ermöglichen.

### **3. Das „Dynamische Modell der Informationskompetenz“ (DYMIK)**

#### **3.1. Allgemeine Anforderungen**

Die bibliothekarischen Schulungsangebote bzw. Benutzerschulungen waren bisher weitgehend objektorientiert<sup>1</sup>. Dies führte dazu, dass meist eine Fülle technischer Detailkenntnisse und Handhabungsfertigkeiten vermittelt wurden, deren Relevanz für die Teilnehmer häufig schwer nachvollziehbar war.

Erforderlich ist jedoch eine sinnbezogene Vermittlung notwendiger Methodenkompetenzen, die Kenntnisse und Fertigkeiten in einen aus Teilnehmersicht funktionalen Zusammenhang stellt. Dies setzt eine Vergegenständlichung und Strukturierung der komplexen Prozesse des Informationsprozesses aus Teilnehmerperspektive voraus. Sie lässt sich mittels konzeptioneller Modelle der Informationskompetenz<sup>2</sup> erreichen. Komplexe Sachverhalte lassen sich hierbei auf wesentliche Merkmale reduzieren und geben mentale Orientierungshilfen für den individuellen Lernprozeß.<sup>3</sup>

#### **3.2. Modelle der Information Literacy**

Erste Modelle, die sich mit der Strukturierung des Gegenstands Informationskompetenz befassen, wurden in anglo-amerikanischen Ländern erstellt. Es handelt sich hierbei um Modelle der "Information Literacy", die die Grundlage für die bibliothekarischen Schulungsaktivitäten bilden.<sup>4</sup> Gemeinsames Merkmal dieser Modelle ist die handlungs- und prozessorientierte Sichtweise. Informationssuche und -verarbeitung werden aus der Perspektive von Individuen mit Informationsdefiziten bzw. Informationsproblemen gesehen. Im Mittelpunkt steht somit nicht ein Informationssystem und dessen Handhabung, z.B. Suchinstrumente im lokalen OPAC, sondern ein konkretes Handlungsproblem, z.B. Informationen über die Integration von Ausländern in Heidelberg.

---

<sup>1</sup> Nähere Ausführungen zur Unterscheidung zwischen objektorientierten und problem- bzw. funktional orientierten Schulungskonzepten finden sich in den Aufsätzen von Homann, Benno: Benutzerschulungen für CD-ROM-Angebote an Wissenschaftlichen Bibliotheken - Bedarf und Schulungskonzepte, in: In Zukunft scheibchenweise / Hessischer Bibliothekstag 1995, 10. Mai 1995 in Rüsselsheim, Offenbach: Deutscher Bibliotheksverband, Landesverband Hessen, 1995, S. 67-91; Schulungen als Aufgabe einer benutzerorientierten Bibliothek, in: ZfBB 43 (1996) 6, S. 569 - 613

<sup>2</sup> Ein Modell der Informationskompetenz (IK) und die daraus ableitbaren individuellen Kenntnisse und Fähigkeiten ist dabei umfassender als häufig verwandte ähnliche Modelle bzw. Begriffe, wie Bibliotheks-, Medien- und Informatikkompetenz.

<sup>3</sup> Siehe hierzu ausführlicher: Gage, Nathaniel/Berliner, David: Pädagogische Psychologie, Weinheim 1996, S. 324ff. Mit der Bedeutung von Schemata, die die Grundlage von Modellen bilden, befassen sich Mandl, Heinz/Friedrich, Helmut F./Hron, Aemilian; Theoretische Ansätze zum Wissenserwerb, in Mandl, Heinz/Spada, Hans (Hrsg): Wissenspsychologie, Weinheim 1988, S. 123-160

<sup>4</sup> Einen umfassenden Überblick über die verschiedenen Modelle und Ansätze der „Information Literacy“ mit Schwerpunkt auf dem Ansatz von Kuhstau gibt Thomas, Nancy P.: Information Literacy and Information Skills Instruction, Englewood (Col): Libraries Unlimited, 1999.

Im wesentlichen prägen zwei Modelle die Schulungskonzepte anglo-amerikanischer Bibliotheken: das Modell der "Six Big Skills" und das Modell des "Information Searching Process".

Das Modell der „SIX BIG SKILLS“ wurde von Eisenberg und Berkowitz entwickelt.<sup>5</sup> Wesentliche Merkmale dieses Ansatzes sind:

- übersichtliche Strukturierung des Informationsprozesses
- lineare Sichtweise des Informationsprozesses
- weitgehende Beschränkung auf kognitive Faktoren

Von zentraler Bedeutung in diesem Modell ist die übersichtliche Untergliederung des Informationsprozesses in sechs Schritte: "Task Definition", "Information Seeking Strategies", "Location and Sccess", "Use of Information", "Synthesis", "Evaluation". Diese Unterteilung wurde zum Markenzeichen des Ansatzes.

Die einzelnen Schritte in dem Informationsprozess werden in einer linearen Abfolge dargestellt. Am Beginn steht die Task-Definition, in der das Informationsproblem und der Informationsbedarf geklärt werden; am Ende steht die Phase der Evaluation, in der das Ergebnis und der Informationsprozess als Suchstrategie beurteilt werden. Eine Rekursion bzw. Umkehrung der Phasenabfolge wird in diesem Modell nicht berücksichtigt.

Der Focus des Modells liegt in der Erfassung kognitiver Aspekte des Informationsprozesses. Deutlich zeigt sich dies bei der Aufzählung operationalisierbarer Schulungsziele bei den einzelnen Prozessphasen, z.B. die Auflistung der potentiellen Informationsquellen in der Phase der „Information seeking strategies“. Die Beschränkung auf kognitive Aspekte des Informationsprozesses unterscheidet diesen Ansatz maßgeblich von dem zweiten Modell.

Das Modell des „Information Searching Process“ (ISP) wurde von Kuhltau entwickelt<sup>6</sup>. Wesentliche Merkmale dieses Ansatzes sind:

- Focusierung des Informationsbedarfs
- Rekursivität der Phasenabfolge

---

<sup>5</sup> Über diesen Ansatz liegen inzwischen zahlreiche Publikationen vor. Als Grundlagenwerk gilt die Veröffentlichung von Eisenberg, Michael B./Berkowitz, Robert E.: Information Problem-Solving: The Six Big Skills Approach to Library and Information Skills Instruction, Norwood, NJ: Ablex 1990. Daneben sind umfangreiche Informationen über diesen Ansatz und konkrete Anwendungen im Internet unter der URL <http://big6.com> zu finden.

<sup>6</sup> Von grundlegender Bedeutung sind folgende Publikationen von Kuhltau, Carol C.: Inside the Search Process: Information Seeking from Users's Perspective, in: Journal of the American Society for Information Science, 42 (1991) 5, S. 361-371, Seeking Meaning, Norwood, N.J.: Ablex

- Integration emotionaler Faktoren
- Ableitung lernpsychologisch basierter Vermittlungsstrategien

Die Ermittlung des Informationsbedarfs und die Erstellung neuer Informationsprodukte bilden den Schwerpunkt des ISP Modells. Deutlich zeigt sich dies in der Phasenaufgliederung des Informationsprozesses: Die vier ersten Informationsphasen, „Task Initiation“, „Topic Selection“, „Topic Exploration“ und „Focus Formulation“ entsprechen der ersten Phase des „Six-Big-Skills“-Modell. Die beiden weiteren Phasen, „Collection“ und „Presentation“ eignen sich nur in geringem Maße zur Thematisierung von Informationskompetenzen, die sich auf bibliothekarische Informationssysteme beziehen.

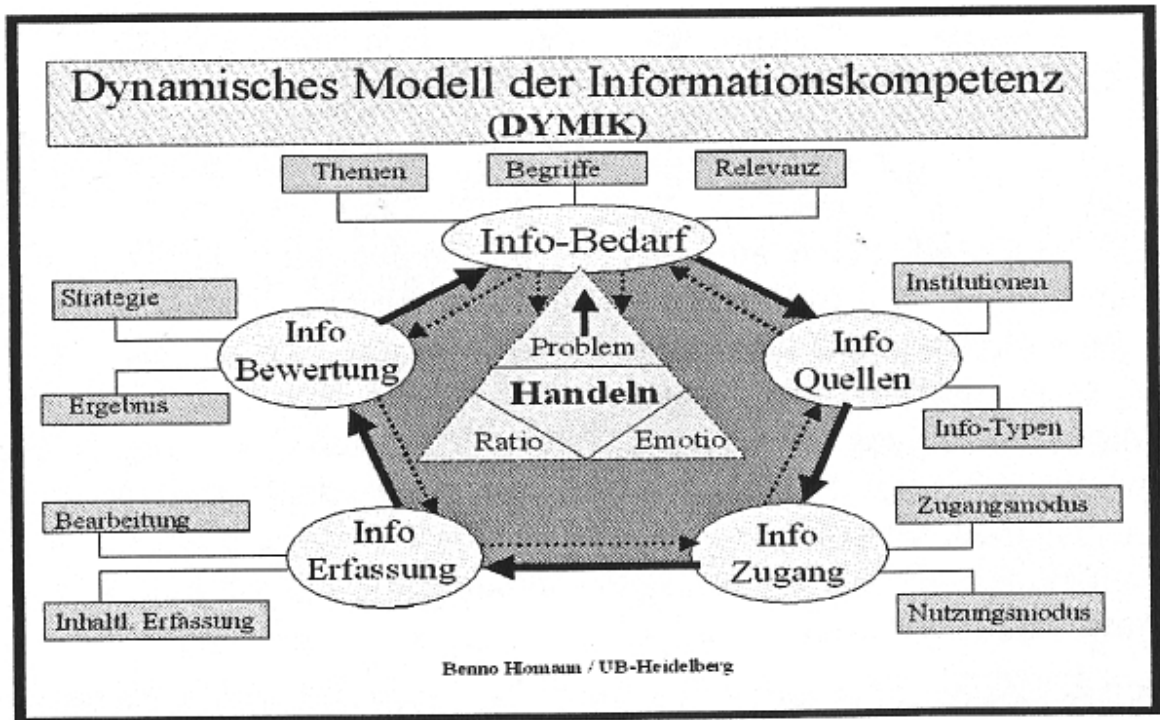
Das Modell berücksichtigt eine rekursive Phasenabfolge bzw. die Rückkehr in eine vorher schon durchlaufene Phase des Informationsprozesses und die erneute Initiation eines Informationsprozesses nach der Evaluation der Ergebnisse. Es erfasst damit einen wichtigen Aspekt realer Informationsprozesse.

Die Integration emotionaler Faktoren, wie z.B. Angst, Unsicherheit, Zufriedenheit sind ein markantes Merkmal dieses Modells. Kuhltau ermöglicht damit eine ganzheitliche Perspektive und Analyse von Informationsprozessen. Subjektive Gefühlszustände, die maßgeblich das individuelle Handeln und den Ablauf eines Informationsprozess bestimmen, werden damit thematisierbar.

Die emotionalen Faktoren sind von wesentlicher Bedeutung für die Ableitung von Vermittlungsstrategien. Hier orientiert sich Kuhltau an dem entwicklungspsychologischen Ansatz von Wigotsky und gibt Empfehlungen zur pädagogischen Intervention von Bibliothekaren in Schulungsveranstaltungen.

### **3.3. Funktionen und Merkmale von DYMIK**

Das "Dynamische Modell der Informationskompetenz" (DYMIK) baut auf den beiden hier skizzierten Modellen der "Information Literacy" auf.



Ziel ist die Erstellung eines Modells, das folgende Funktionen erfüllt:

- Didaktische Relativierung der funktionalen Bedeutung der bibliothekarischen Informationssysteme
- Methodische Veranschaulichung der Phasen des Informationsprozesses, der funktionalen Beziehungen zwischen den Einflussfaktoren sowie der Bedeutung konkreter Informationssysteme
- organisatorische Grundlage zur Abgrenzung inhaltlicher Zuständigkeiten zwischen involvierten Schulungsinstitutionen und für die Ermittlung von Kooperationsmöglichkeiten im methodischen sowie technischen Bereich

Die wesentlichen Merkmale dieses Modells sind

- subjekt-/handlungsbasierte Perspektive
- 5-Phasen-Struktur des Informationsprozesses
- dynamische Sichtweise des Informationsprozesses
- funktionale Zuordnung der Informationsobjekte und Einzelkompetenzen zu den einzelnen Informationsphasen

Von grundlegender Bedeutung ist die subjekt-/handlungsorientierte Perspektive des Modells. Dabei wird angenommen, dass ein Informationsprozess durch eine

subjektive Handlungsambivalenz bzw. ein Problem initiiert wird und neben rationalen insbesondere auch emotionale Faktoren wirksam sind. Eine solches Problem könnte sich z.B. aus dem Ziel „Kauf eines Mobiltelefons“ ergeben. Der Käufer sieht sich nach seiner Grundsatzentscheidung mit mehreren Netzanbietern und Geräten konfrontiert. Die rational erkannten Handlungsalternativen führen dabei zu einem Gefühl der Unsicherheit, welches Mobiltelefon nun das geeignetste ist. Es entsteht ein subjektives Problem, das sich möglicherweise durch zusätzliche Informationen beseitigen lässt.

Mit seiner Entscheidung das beste Mobiltelefon kaufen zu wollen, initiiert der Käufer den Informationsprozess. Er befindet sich damit innerhalb des Modells in der Phase der Ermittlung des „Info-Bedarfs“. Bevor die weiteren vier Phasen des Informationsprozesses durchlaufen werden, müssen hier die inhaltlichen Anforderungen der zu ermittelnden Informationen präzisiert werden. Diese Phase ist von zentraler Bedeutung für den Erfolg des weiteren Informationsprozesses. Bei komplexen wissenschaftlichen Problemen werden in dieser Phase umfassende kognitive und sozial-kommunikative Fähigkeiten benötigt.

Daran schließt sich die Phase der Ermittlung der potentiellen und relevanten „Info-Quellen“ an, in der Kenntnisse über potentielle Informationseinrichtungen und potentielle Informationstypen erforderlich sind.

Der konkrete Zugriff auf die einzelnen Informationen in den ausgewählten Institutionen und Informationssystemen erfolgt in der nächsten Phase, die im Modell mit „Info-Zugang“ bezeichnet wird. Benötigt werden hier technische Kenntnisse und Fähigkeiten zur Nutzung der teilweise komplexen Suchinstrumente.

Die inhaltliche Erschließung und Bearbeitung einer gefundenen Informationen erfolgt in der Phase der „Info-Erfassung“. Hier sind methodische Fähigkeiten der Textanalyse und Textbearbeitung für spezifische Verwendungszwecke wie Referate und Vorträge erforderlich. Dies impliziert auch technische Fertigkeiten, z.B. in der Nutzung von Literaturverwaltungssystemen, Textverarbeitungs- und Präsentationsprogrammen.

Die letzte Phase „Info-Bewertung“ bezieht sich sowohl auf den Informationsprozess als auch auf die konkreten Ergebnisse des Informationsprozesses. Hier findet eine Bewertung der Ergebnisse statt, wobei der Informationsbedarf und die dabei formulierten Informationsziele als zentrale Beurteilungskriterien fungieren.

Die Dynamik von Informationsprozessen wird in dem Modell repräsentiert durch die zirkulare Anordnung der Phasen des Informationsprozesses und den Einbau einer rekursiven Beziehung zwischen den einzelnen Phasen. Die zirkulare Anordnung verbindet die Phasen der „Info-Bewertung“ und „Info-Bedarf“. Damit soll deutlich gemacht werden, daß Informationsprozesse meist nicht sofort zu einem eindeutigen Ergebnis führen, sondern erneute Probleme die Initiierung eines weiteren Informationsprozesses und ggf. die Berücksichtigung zusätzlicher

Informationsinstrumenten erforderlich machen.

Die rekursive Beziehung zwischen den einzelnen Phasen wird im Modell durch die gestrichelten Pfeile symbolisiert, die in entgegengesetzter Richtung zur idealtypischen Abfolge der fünf Phasen verlaufen. Sehr häufig ist eine Wiederholung der vorangegangenen Informationsphase erforderlich, wenn z.B. in der Phase der "Info-Quellen" das Thema und die relevanten Begriffe noch nicht ausreichend geklärt sind.

Die einzelnen Informationssysteme und Fertigkeiten zu ihrer Nutzung werden in diesem Modell in einer funktionalen Beziehung zum individuellen Informationsprozess gesehen. Umfang und Niveau der zu vermittelnden Kenntnisse lassen sich ableiten aus ihrer konkreten Bedeutung für eine bestimmte Informationsphase. Entsprechend werden die für den Informationsprozess relevanten Informationsobjekte und Handlungsfähigkeiten den einzelnen Phasen des Informationsprozesses zugeordnet. So werden z.B. die Kenntnisse über verschiedenen Info-Typen als relevant für die Phase der Ermittlung potentieller Informationstypen gesehen und deshalb in dem Modell der Phase "Info-Quellen" zugeordnet.

Die Informationsphasen übernehmen so eine doppelte Repräsentationsfunktion. Sie fungieren einerseits als Knotenpunkt im Informationsprozess und andererseits als Ausgangspunkt für eine differenzierte Erfassung von zuordenbaren Einzelkompetenzen und Informationsobjekten.

#### **4. Der Nutzen von DYMIK für die Gestaltung bibliothekarischer Schulungsangebote**

Das Dynamische Modell der Informationskompetenz lässt sich in vielfältiger Weise bei der Planung und Durchführung von bibliothekarischen Schulungsveranstaltungen nutzen.

Informationskompetenzen und Informationssysteme bilden ein sehr komplexes Themenfeld, das für die Erstellung eines umfassenden Schulungsangebotes strukturiert werden muss. DYMIK eignet sich dabei als Strukturierungsinstrument. Die Informationsphasen erleichtern die Grobstrukturierung und die Herausarbeitung von zusammengehörigen Themenkomplexen, die auf die verschiedenen Veranstaltungen aufzuteilen sind. Die subjektbezogene Problemperspektive relativiert die Bedeutung einzelner Kompetenzen und bietet in Verbindung mit Lernzieltypologien zusätzliche Strukturierungsmöglichkeiten.

So führte die unterschiedliche funktionale Bedeutung von Informationssystemen an der Universitätsbibliothek Heidelberg zur Einführung einer neuen Informationstypologie. Unterschieden wird dabei zwischen Referenzinformationen, denen sich die Kataloge, Lexika und bibliographische Datenbanken zuordnen lassen und Originalinformationen, unter die sich z.B. Bücher, elektronische Volltexte, etc. subsumieren lassen. Diese Typologie wird in DYMIK der Phase "Info-Quellen" zugeordnet und bildet die immer wiederkehrende Grundlage für die



Systematisierung der unterschiedlichen Informationssysteme in den verschiedenen Grundlagen- und Aufbauveranstaltungen.

Die Phasenstruktur von DYMIK bietet zusätzlich die Möglichkeit einer inhaltlichen Abgrenzung und Abstimmung mit Institutionen, die sich ebenfalls mit der Vermittlung von Informationskompetenzen befassen. So beschränkt sich die Universitätsbibliothek Heidelberg in ihren Veranstaltungen primär auf die Phasen "Info-Quellen" und "Info-Zugang" während die restlichen Phasen mit den zugehörigen Einzelkompetenzen in regulären Lehr- oder Schulveranstaltungen vermittelt werden. Das integrative Element zwischen diesen Angeboten soll zukünftig das Modell DYMIK bilden, das in allen entsprechenden Veranstaltungen zu thematisieren ist. Voraussetzung für diese planerische Nutzung von DYMIK ist eine verstärkte Kooperation der involvierten Einrichtungen und eine systematische lernzielbasierte Abstimmung der verschiedenen Veranstaltungen. Die Grundlagen hierfür sind an der Universitätsbibliothek Heidelberg durch die Lernzielorientierung des gesamten curricularen Schulungsangebotes vorhanden.

Die subjektbasierte Perspektive von DYMIK fördert auch eine adressatenbezogene Differenzierung des Schulungsangebotes. Da sich die Veranstaltungen auf potentielle Informationsprobleme der Teilnehmer beziehen sollen, erfordert dies eine Unterteilung nach Niveau und Adressatenkreis.

An der Universitätsbibliothek Heidelberg führte dies zur Differenzierung in Grundlagen- und Aufbauschulungen mit mehreren modular aufeinander abgestimmten Einzelveranstaltungen. Eine weitere adressatenbezogene Ausdifferenzierung bildet das neue Schulungsangebot für ausländische Studenten und Wissenschaftler. In diesen sollen spezielle sprachliche Schwierigkeiten bei der methodischen Gestaltung stärker berücksichtigt werden.

Weitere Vorteile bietet DYMIK bei der konkreten methodischen Gestaltung und Durchführung von Schulungsveranstaltungen. Sie resultieren zum einen aus der problembasierten Perspektive und zum weiteren aus der Visualisierbarkeit von DYMIK.

Der Einsatz von DYMIK als didaktisch-methodisches Konzept erfordert den Einsatz problemorientierter und aktivierender Methoden. Dies führte an der Universitätsbibliothek zu einer Verstärkung der Übungsteile und der Integration möglichst realitätsbezogener Problemfällen in den einzelnen Veranstaltungen. Auch in den komplementären internetbasierten Tutorials soll dieser problemorientierte methodische Ansatz verstärkt umgesetzt werden.

Von fundamentaler Bedeutung ist die Visualisierbarkeit von DYMIK in einer Graphik. Dies erleichtert den Lernprozeß, indem die komplexen Faktoren des Informationsprozesse auf die wesentlichen Elemente und ihre Beziehungen reduziert werden. DYMIK fungiert hier als kognitives Schema, das die Vermittlung weiterer Kompetenzen und das Bewußtmachen der individuellen psychischen Vorgänge bei Informationsprozessen erleichtert.

Eine graphische Darstellung von DYMIK soll zukünftig bei allen Schulungsveranstaltungen und bei den internetbasierten Tutorials der Universitätsbibliothek präsent sein und den potentiellen Bezugsrahmen bilden, auf den bei Bedarf rekuriert werden soll. Auf diesem Weg lassen sich z.B. methodische Defizite thematisieren, die dem Themenkomplex "Info-Bedarf" zuordenbar sind, aber bei der Vermittlung von Kompetenzen im Themenkomplex "Info-Zugang" deutlich werden.

## **5. Folgerungen und Perspektiven**

Derzeit besteht noch ein beträchtlicher Entwicklungsrückstand der deutschen gegenüber den anglo-amerikanischen Bibliotheken. Die Nutzung des Dynamischen Modells der Informationskompetenz oder vergleichbarer Modelle bietet allerdings Möglichkeiten zur Reduzierung dieses Rückstandes.

Positiv sind auch die zunehmenden Schulungsaktivitäten an vielen Bibliotheken zu sehen. So entwickeln sich derzeit die öffentlichen Bibliotheken zu Bildungseinrichtungen für die Vermittlung von Internetkompetenzen und die wissenschaftlichen Bibliotheken arbeiten verstärkt am Aufbau umfassender Schulungskonzepte.

Damit werden grundlegende Voraussetzungen gelegt für den Aufbau pädagogischer Kompetenzen an den Bibliotheken und die Erschließung eines neuen zukunftssträchtigen Aufgabenfeldes der Bibliotheken.



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 207 – 226

## **Leistungsvergleiche zwischen wissenschaftlichen Bibliotheken**

### **Eine betriebswirtschaftliche Untersuchung von Universitätsbibliotheken im deutsch- und im englischsprachigen Raum**

**Gerhard Reichmann**

Karl-Franzens-Universität Graz

Institut für Informationswissenschaft

Universitätsstraße 15 / F 3, A-8010 Graz

Telefon: +43/316/380-3563, Fax: +43/316/380-9575

Email: gerhard.reichmann@kfunigraz.ac.at

#### **Zusammenfassung**

Gegenstand des vorliegenden Beitrages ist die Durchführung von Leistungsvergleichen zwischen verschiedenen Gruppen von Universitätsbibliotheken. Im Rahmen dieser Untersuchung, die sich auf Universitätsbibliotheken in Österreich, Deutschland, der Schweiz, den USA, Australien und Kanada bezieht, werden Leistungskennzahlen unterschiedlicher Komplexität als Vergleichsbasis herangezogen, wobei unter anderem eine Spitzenkennzahl zur Darstellung der Gesamtleistung einer Bibliothek zum Einsatz kommt. Während in einzelnen Teilleistungsbereichen durchaus signifikante Leistungsunterschiede zwischen Universitätsbibliotheken im deutschsprachigen und im englischsprachigen Raum, zwischen kleinen und großen Universitätsbibliotheken, zwischen Bibliotheken an renommierten und an weniger renommierten Universitäten sowie zwischen Bibliotheken an staatlichen und an privaten Universitäten auftreten, ergeben Vergleiche auf Basis der Spitzenkennzahl keine signifikanten Leistungsunterschiede zwischen diesen betrachteten Gruppen von Bibliotheken.



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

## **Abstract**

In this paper we analyse whether there are significant differences in performance between academic libraries in German speaking countries and English speaking countries as well as between 'small' and 'large' academic libraries. For this purpose we examine the performance of 132 randomly chosen academic libraries in Austria, Germany, Switzerland, Australia, Canada and the U.S. The results show that there is no significant difference in performance between libraries in different linguistic areas as far as total factor productivity is concerned, although there is evidence that in case of particular single factor productivities libraries in English speaking countries come off better. Comparing the performance of 'small' and 'large' libraries approximately leads to the same results.

## **1. Vorstellung der Untersuchung**

Betriebswirtschaftliche Aspekte finden im Bereich des wissenschaftlichen Bibliothekswesens zunehmend Beachtung. Ausgehend vom englischsprachigen Raum, spielen derartige Überlegungen mittlerweile auch im deutschsprachigen Raum eine größere Rolle. Innerhalb der Gruppe der wissenschaftlichen Bibliotheken betrachten sich vor allem Universitätsbibliotheken verstärkt als moderne Dienstleistungseinrichtungen, die alle an sie gestellten Aufgaben möglichst gut erfüllen wollen. Hinsichtlich der Bewertung der Güte dieser Aufgabenerfüllung ist dabei stets zwischen einer subjektiven und einer objektiven Komponente zu unterscheiden. Während sich die subjektive Komponente nur durch empirische Untersuchungen erheben läßt, kann die objektive Komponente mit Hilfe von Leistungsindikatoren, basierend auf den in Bibliotheksstatistiken gewöhnlich bereits vorhandenen Daten, gemessen werden. Die im Rahmen der vorliegenden Untersuchung durchgeführten Leistungsvergleiche zwischen Universitätsbibliotheken im deutsch- und im englischsprachigen Raum beziehen sich ausschließlich auf diese objektive Komponente, wobei als Vergleichsobjekte nicht die einzelnen Bibliotheken, sondern die nach verschiedenen Kriterien gebildeten Gruppen von Bibliotheken herangezogen werden.

Im 2. Kapitel werden zunächst die zu prüfenden Hypothesen formuliert. Die Definition der ausgewählten Leistungsindikatoren, die sich in einfache Kennzahlen, zusammengesetzte Kennzahlen und eine Spitzenkennzahl unterteilen lassen, sowie die Charakterisierung der verschiedenen Gruppen von Universitätsbibliotheken erfolgt ebenfalls in diesem Kapitel. In Kapitel 3 wird die der Hypothesenprüfung dienende Stichprobe, in der 132 Universitätsbibliotheken enthalten sind, vorgestellt. Anschließend werden die einzelnen Bibliotheken entsprechend der in Kapitel 2 festgelegten Gruppencharakteristika den einzelnen Gruppen zugeordnet. In der Folge wird beschrieben, wie bei der Erhebung der für die Leistungsvergleiche benötigten Daten vorgegangen wurde. Im fünften Kapitel werden schließlich die formulierten Hypothesen anhand der erhobenen Daten überprüft und mögliche Ursachen für Leistungsdifferenzen genannt. Abschließend werden die wichtigsten Untersuchungsergebnisse nochmals kurz zusammengefasst.

## **2. Zielsetzungen und Vorgangsweise**

Die gegenständliche Arbeit hat in erster Linie die Überprüfung der folgenden vier Hypothesen zum Ziel:

Hypothese 1: Sprachraumvergleich

Ho: Es gibt keine Leistungsunterschiede zwischen Universitätsbibliotheken im deutschsprachigen und Universitätsbibliotheken im englischsprachigen Raum.

Hypothese 2: Größenvergleich

Ho: Es gibt keine Leistungsunterschiede zwischen kleinen und großen Universitätsbibliotheken.

Hypothese 3: Rankingvergleich - nur für den englischsprachigen Raum

Ho: Es gibt keine Leistungsunterschiede zwischen Bibliotheken an renommierten und Bibliotheken an weniger renommierten Universitäten.

Hypothese 4: Trägervergleich - nur für die USA

Ho: Es gibt keine Leistungsunterschiede zwischen Bibliotheken an staatlichen und Bibliotheken an privaten Universitäten.

Bei der Überprüfung der Hypothesen 1 bis 4 wird so vorgegangen, dass jeweils jene in der Stichprobe enthaltenen Bibliotheken, die in den Geltungsbereich der Hypothese fallen, entsprechend der Hypothese zu zwei Vergleichsgruppen zusammengefasst werden. Während die Gruppeneinteilungen für den Sprachraumvergleich (Hypothese 1) und für den Trägervergleich (Hypothese 4), der auf Bibliotheken an renommierten<sup>1</sup> Universitäten in den USA beschränkt ist<sup>2</sup>, geläufigen linguistischen und organisatorischen Gegebenheiten folgen und daher nicht weiter erklärungsbedürftig sind, liegen den beiden anderen Gruppeneinteilungen individuell festgelegte Abgrenzungskriterien zugrunde. So basiert die Gruppeneinteilung für den Größenvergleich (Hypothese 2) auf der Zahl der Bibliotheksmitarbeiter in VZÄ<sup>3</sup>: all jene Bibliotheken, die eine Beschäftigtenzahl von mehr als 175 VZÄ aufweisen, sind demnach der Gruppe „groß“ zugeordnet.<sup>4</sup> Etwas komplexer ist die Gruppeneinteilung für den

---

<sup>1</sup> Vgl. dazu die nachfolgenden Ausführungen zu Hypothese 3.

<sup>2</sup> Da es sich bei sämtlichen Bibliotheken an privaten Universitäten, die im Rahmen der Stichprobe ausgewählt wurden, um solche an renommierten Universitäten handelt, können dieser Gruppe zur Untersuchung der Frage, ob es zwischen Bibliotheken an staatlichen und privaten Universitäten wesentliche Leistungsunterschiede gibt, sinnvollerweise nur jene Bibliotheken an staatlichen Universitäten gegenübergestellt werden, die ebenfalls der Gruppe „Universitätsbibliotheken an renommierten Universitäten“ angehören. Somit kommt es beim Einsatz dieses Gliederungsmerkmals neben der Beschränkung auf die USA noch zur weiteren Beschränkung auf die Gruppe der „Universitätsbibliotheken an renommierten Universitäten“.

<sup>3</sup> VZÄ steht für Vollzeitäquivalente.

<sup>4</sup> Es wurde ursprünglich versucht, die Abgrenzung zwischen „klein“ und „groß“ differenzierter zu gestalten, indem neben der Beschäftigtenzahl in VZÄ auch noch weitere „Größenkennndaten“ einer Universitätsbibliothek, wie etwa Buchbestand oder Fläche, herangezogen wurden. Aber auch derartig zusammengesetzte „Größenkennzahlen“ ergaben

Rankingvergleich (Hypothese 3), der ausschließlich für den englischsprachigen Raum vorgenommen wird. Hier werden die untersuchten Universitätsbibliotheken nach dem Merkmal „Ranking“ in die zwei Gruppen „Bibliotheken an renommierten Universitäten“ und „Bibliotheken an weniger renommierten Universitäten“ untergliedert, wobei mangels eines alle drei betrachteten Staaten umfassenden Rankings und aufgrund der stark unterschiedlichen Gesamtanzahl an Universitäten in den drei Staaten differenziert vorgegangen wird: in den USA sind Universitätsbibliotheken dann der Gruppe der „Bibliotheken an renommierten Universitäten“ zugeordnet, wenn die jeweilige Universität, der sie angehören, zu den 50 im Rahmen nationaler Rankings<sup>5</sup> bestplazierten Universitäten zählt. Australische Universitätsbibliotheken gehören der Gruppe „Bibliotheken an renommierten Universitäten“ unter der Voraussetzung an, dass die jeweils zugehörige Universität zu den 50 bestplazierten „allgemeinen“ Universitäten bzw. zu den 30 bestplazierten Technischen Universitäten im Rahmen von internationalen Universitätsrankings, die sich auf den ostasiatischen, australischen und ozeanischen Bereich beziehen<sup>6</sup>, zählt. Für Kanada gilt, dass Bibliotheken nur dann zur Gruppe der „Bibliotheken an renommierten Universitäten“ gezählt werden, wenn die zugehörige Universität im Zuge nationaler Universitätsrankings<sup>7</sup> einen Platz unter den 10 bestplazierten Universitäten einnimmt.

Für alle vier Vergleiche gilt, dass die Hypothesenprüfung sowohl hinsichtlich ausgewählter Teilleistungsbereiche als auch bezüglich der Gesamtleistung von Universitätsbibliotheken, ausgedrückt durch die entwickelte Spitzenkennzahl, erfolgt. Vergleichsbasis sind dabei einfache und zusammengesetzte Kennzahlen bzw. die entwickelte Spitzenkennzahl. Zu den einfachen Leistungskennzahlen im Sinne der im Rahmen der vorliegenden Arbeit verwendeten Definition zählen sämtliche Input- und Outputdaten einer Bibliothek, wie etwa Ausgaben, Buchbestand, VZÄ, Entlehnungen, Öffnungszeiten oder Buchzuwachs. Da diese einfachen Kennzahlen aber hauptsächlich von der jeweiligen Größe einer Universitätsbibliothek abhängen und für einen Leistungsvergleich zwischen Bibliotheken nur bedingt aussagekräftig sind, werden einfache Leistungskennzahlen hier mit Ausnahme der Regelöffnungsstunden pro Woche<sup>8</sup>, für die das eben Gesagte nicht notwendigerweise gilt, nicht berücksichtigt.

Zusammengesetzte Kennzahlen haben zwar ebenso wie einfache Kennzahlen eine Aussagekraft, die sich lediglich auf Teilleistungsbereiche von Universitätsbibliotheken beschränkt, doch sind sie insofern aussagekräftiger, als die einzelnen Kennzahlen weitgehend größenunabhängige Leistungsindikatoren

---

im Großen und Ganzen ein ähnliches Einteilungsergebnis, sodass letztendlich auf eine solche Vorgangsweise verzichtet wurde.

<sup>5</sup> Als Quelle wurde ein von U.S. News veröffentlichtes Universitätsranking für das Jahr 1999 verwendet; URL:

[http://www.usnews.com/usnews/edu/college/rankings/natunivs/natu\\_a2.htm](http://www.usnews.com/usnews/edu/college/rankings/natunivs/natu_a2.htm) (11.8.1999)

<sup>6</sup> Es musste auf internationale Rankings zurückgegriffen werden, da für Australien keine entsprechenden Rankings existieren bzw. offen zugänglich sind. Als Quelle wurde ein von Asiaweek veröffentlichtes Universitätsranking für das Jahr 1998 verwendet; URL: <http://www.sources-asie.tm.fr/repertoire/siteinternet/cdocbibl/bestuniversitesasie.html> (11.8.1999)

<sup>7</sup> Als Quelle wurde ein von Maclean veröffentlichtes Universitätsranking für das Jahr 1998 verwendet; URL: <http://www.macleans.ca/pipeline/unimag/98introessay.html> (11.8.1999)

<sup>8</sup> Gemeint sind die Regelöffnungsstunden des Hauptlesesaales.

darstellen.<sup>9</sup> In dieser Untersuchung werden die ausgewählten Universitätsbibliotheken anhand der folgenden sechs zusammengesetzten Kennzahlen beurteilt: „Buchbestand pro Kernnutzer“<sup>10</sup>, „Buchzuwachs pro Kernnutzer“, „Anzahl der Kernnutzer pro Leseplatz“, „Anzahl der Kernnutzer pro VZÄ“, „Anzahl der Entlehnungen pro Kernnutzer“ und „Erwerbungsausgaben pro Kernnutzer“.

An oberster Stelle eines Kennzahlensystems sollte eine Spitzenkennzahl stehen, anhand derer ein Gesamtleistungsvergleich zwischen den einzelnen Untersuchungseinheiten möglich ist. Im vorliegenden Fall handelt es sich bei dieser Spitzenkennzahl um die mit Hilfe der DEA<sup>11</sup> ermittelte relative Effizienz von Universitätsbibliotheken, bei der die Gesamtproduktivität der jeweils betrachteten Bibliothek zu jener der Bibliothek mit der höchsten beobachteten Gesamtproduktivität in Relation gesetzt wird. Die Qualität von Gesamtproduktivitätswerten bzw. relativen Effizienzwerten, die mit Hilfe der DEA berechnet werden, hängt natürlich sehr stark von der sinnvollen Auswahl entsprechender Input- und Outputfaktoren ab. Im vorliegenden Fall werden für die Ermittlung der relativen Effizienz zwei Input- und vier Outputfaktoren verwendet: die Anzahl der Beschäftigten in VZÄ und der Buchbestand als Inputfaktoren, die Anzahl der laufend gehaltenen Zeitschriften, die Anzahl der Entlehnungen, die Öffnungszeiten pro Woche und der Buchzuwachs als Outputfaktoren.

Die nachfolgende Tabelle fasst das eben beschriebene und in der Folge eingesetzte Kennzahlensystem zusammen.

<b>Einfache Kennzahl</b>	<b>Zusammengesetzte Kennzahlen</b>	<b>Spitzenkennzahl</b>
Regelöffnungsstunden pro Woche	Buchbestand pro Kernnutzer	Relative Effizienz
	Buchzuwachs pro Kernnutzer	<u>basierend auf:</u>
	Kernnutzer pro Leseplatz	VZÄ (Input)
	Kernnutzer pro VZÄ	Buchbestand (Input)
	Entlehnungen pro Kernnutzer	Zeitschriften (Output)
	Erwerbungsausgaben pro	Entlehnungen (Output)
		Öffnungszeiten (Output)
		Buchzuwachs (Output)

**Tabelle 1:** Kennzahlensystem

<sup>9</sup> Es handelt sich bei diesen Kennzahlen um Verhältniszahlen, bei denen Größeneinflüsse durch die Berücksichtigung sowohl im Zähler als auch im Nenner der Verhältniszahl relativiert werden; als Beispiel sie hier etwa der Buchbestand pro Kernnutzer genannt: da große Universitätsbibliotheken im Durchschnitt sowohl einen höheren Buchbestand als auch eine höhere Anzahl an Kernnutzern aufweisen als kleine Universitätsbibliotheken, ist die Verhältniszahl „Buchbestand pro Kernnutzer“ letztlich weitgehend größenunabhängig.

<sup>10</sup> Der Begriff „Kernnutzer“ umfaßt alle Studierenden und Universitätslehrer einer Universität.

<sup>11</sup> DEA steht für Data Envelopment Analysis: es handelt sich dabei um ein nichtparametrisches Optimierungsverfahren

### 3 Zusammensetzung der Stichprobe

Bei der vorliegenden Untersuchung handelt es sich um eine stichprobenartige Betrachtung von Universitätsbibliotheken in Österreich, Deutschland, der Schweiz, den USA, Kanada und Australien. Aus der Grundgesamtheit sämtlicher Universitätsbibliotheken in diesen sechs Staaten wurde eine nach Ländern geschichtete Stichprobe im Ausmaß von ca. 20 % gezogen. Tabelle 2 gibt einen zahlenmäßigen Gesamtüberblick über die ausgewählten

Untersuchungseinheiten: in der Stichprobe sind 12 österreichische, 50 deutsche, 8 Schweizer, 40 US-amerikanische, 12 australische und 10 kanadische Universitätsbibliotheken enthalten. Somit befinden sich 70 Untersuchungseinheiten im deutschsprachigen und 62 Untersuchungseinheiten im englischsprachigen Raum. 52 dieser insgesamt 132 ausgewählten Bibliotheken gehören der Gruppe der großen Universitätsbibliotheken an, 80 jener der kleinen Universitätsbibliotheken. Demnach fallen 61 % der Untersuchungseinheiten in die Gruppe „klein“ und nur 39 % in die Gruppe „groß“.

Anzahl der untersuchten Universitätsbibliotheken			
Land	groß	klein	Summe
Österreich	1	11	12
Deutschland	4	46	50
Schweiz	0	8	8
USA	36	4	40
Australien	4	8	12
Kanada	7	3	10
Summe	52	80	132

**Tabelle 2:** Geschichtete Stichprobe: Länderbetrachtung und Größenbetrachtung<sup>12</sup>

In den Tabellen 3a und 3b erfolgt eine länderweise Aufzählung aller Untersuchungseinheiten inklusive weiterer untersuchungsrelevanter Merkmale. Aus Tabelle 3b geht hervor, dass sich jeweils 50 % der 62 in der Stichprobe enthaltenen Universitätsbibliotheken im englischsprachigen Raum an renommierten Universitäten bzw. an weniger renommierten Universitäten befinden.

Von den 40 Bibliotheken in den USA gehören wiederum jeweils 50 % einer staatlichen bzw. privaten Universität an. Diese Gleichverteilung im Hinblick auf die Kriterien „Ranking“ und „Träger“ ist kein Zufall, sondern das Ergebnis einer entsprechenden Schichtung.

<sup>12</sup> Tabelle 2 gibt einen Überblick über die länderspezifische Schichtung der gezogenen Stichprobe. Die aus dieser Tabelle ebenfalls ersichtliche Einteilung in kleine und große Universitätsbibliotheken stellte zwar kein Schichtungskriterium für die Stichprobe dar, ist aber für die Beantwortung der Frage, ob sich kleine und große Universitätsbibliotheken hinsichtlich ihrer Leistung unterscheiden, von zentraler Bedeutung.



#### 4. Datengewinnung

Die gegenständliche Untersuchung ist eine Querschnittsuntersuchung, die sich ausschließlich auf Daten des Jahres 1998 bezieht. Zur Gewinnung der notwendigen Daten kam ein zweistufiges Erhebungsverfahren zum Einsatz, das aus einer Grunderhebung und aus einer ergänzenden Erhebung bestand. Ziel der Grunderhebung war es, für jede in der Stichprobe enthaltene Universitätsbibliothek in möglichst kurzer Zeit und mit möglichst geringem Aufwand möglichst viele der für die Berechnung der in Kapitel 2 definierten Kennzahlen benötigten Daten zu sammeln. Zu Beginn der Grunderhebung wurde für jede Bibliothek ein Datenblatt, in das alle für die Auswertungen benötigten Leistungsdaten eingetragen werden sollten, angelegt. Ideale Basis für eine Grunderhebung der genannten Art wäre natürlich eine internationale Bibliotheksstatistik, in der neben aggregierten Daten auch die Einzeldaten aller untersuchten Universitätsbibliotheken enthalten sind. Eine derartige Statistik gibt es bis dato noch nicht und wird es nicht nur aus Kostengründen, sondern auch aufgrund der weltweit doch sehr unterschiedlichen Organisationsstrukturen des (Universitäts-)bibliothekswesens vermutlich auch in absehbarer Zeit nicht geben. Deshalb stand von Anfang an die Suche nach nationalen Bibliotheksstatistiken der im Rahmen der vorliegenden Untersuchung betrachteten sechs Staaten im Vordergrund.

das heißt, zum Zeitpunkt der Datenerhebung waren noch keine Daten für das Jahr 1998 verfügbar. So mußte für Österreich im Rahmen der Grunderhebung auf die äußerst aufwendige Notlösung der Datenerhebung auf Basis von Jahresberichten der Universitätsbibliotheken zurückgegriffen werden.

Bibliothek	Größe	
	klein	groß
UB Graz	X	
UB der TU Graz	X	
UB Innsbruck	X	
LIB Klagenfurt	X	
UB Leoben	X	
UB Linz	X	
UB Salzburg	X	
UB Wien		X
UB der TU-Wien	X	
UB der WU Wien	X	
UB Boku Wien	X	
UB der VETMED Wien	X	

UB Augsburg	X	
LIB Bamberg	X	
UB Bayreuth	X	
UB Berlin (Humboldt)		X
UB Bielefeld	X	
Staats- und UB Bremen	X	
UB der TU Chemnitz	X	
UB der TU Cottbus	X	
UB Dortmund		X
LB und UB Dresden	X	
UB und LB Düsseldorf	X	
UB Duisburg	X	
UB Eichstätt	X	
UB und Forschungsbibliothek Erfurt	X	
UB Essen	X	
Zentrale Hochschulbibliothek Flensburg	X	
UB Frankfurt/Oder	X	
UB der TU Freiberg	X	
UB Greifswald	X	
UB der Fernuniversität Hagen	X	
UB und LB Halle/S		X
UB der TU Hamburg	X	
LIB der Universität der Bundeswehr Hamburg	X	
UB und Technische Informationsbibliothek Hannover	X	
UB Hildesheim	X	
UB der TU Ilmenau	X	
UB und LB Jena	X	
LIB Kaiserslautern	X	
UB Kassel	X	
UB Koblenz	X	
UB Konstanz	X	
UB Landau (Universität Koblenz-Landau)	X	
UB Leipzig	X	
Zentrale Hochschulbibliothek Lübeck	X	
UB Lüneburg	X	
LIB Magdeburg	X	
UB Mannheim	X	
UB der Universität der Bundeswehr München-Neubiberg		
UB Oldenburg	X	
UB Osnabrück	X	
UB Paderborn	X	
LIB Passau	X	
UB Potsdam	X	
UB Regensburg	X	
LIB Rostock		X
UB Siegen	X	
UB Trier	X	
UB Ulm	X	
UB Weimar	X	
UB Wuppertal	X	
UB Basel	X	
Stadt- und UB Bern	X	
Kantons- und UB Freiburg	X	
Kantons- und UB Lausanne	X	
UB der ETH Lausanne	X	
öffentliche UB und UB Neuenburg	X	
Zentralbibliothek Zürich	X	
UB der ETH Zürich	X	

**Tabelle 3a:** Geschichtete Stichprobe: Einzelbetrachtung  
(deutschsprachiger Raum)

# Leistungsvergleiche zwischen wissenschaftlichen Bibliotheken

Bibliothek	Größe		Ranking		Träger	
	klein	groß	renommierte Universität	weniger renommierte Universität	Staat	Privat
The University of Alabama Libraries		X		X	X	
University at Albany Libraries	X			X	X	
The University of Arizona Library		X		X	X	
Arizona State University Libraries		X		X	X	
Auburn University Libraries	X			X	X	
University of California Library at Berkeley		X	X		X	
Brown University Library		X	X			X
The University of Chicago Library		X	X			X
University of Cincinnati Libraries		X		X	X	
The University of Connecticut Libraries		X		X	X	
University of California - Davis Library		X	X		X	
Duke University Libraries		X	X			X
Florida State University Libraries		X		X	X	
Georgetown University Library		X	X			X
Georgia Institute of Technology Library	X		X		X	
University of Houston Libraries		X		X	X	
University of Illinois at Chicago Library		X		X	X	
University Library - Univ. of Illinois at Urbana-Champaign		X	X		X	
Iowa State University Library		X		X	X	
University of California, Irvine Libraries		X	X		X	
Kent State University Libraries		X		X	X	
University of Kentucky Libraries		X		X	X	
University Library - The University of Michigan		X	X		X	
Northwestern University Library		X	X			X
University Libraries of Notre Dame		X	X			X
Ohio University Libraries		X		X	X	
Oklahoma State University Library		X		X	X	
Penn State University Libraries		X	X		X	
University of Pittsburgh - University Library		X		X	X	
Princeton University Library		X	X			X
Purdue University Libraries		X		X	X	
University of California, Riverside Library	X			X	X	
University of California San Diego Libraries		X	X		X	
University of California - Santa Barbara Library		X	X		X	
Stanford University Libraries		X	X			X
Temple University Libraries		X		X	X	
Texas Tech University Libraries		X		X	X	
Tulane Libraries		X	X			X
University of Virginia Library		X	X		X	
Yale University Library		X	X			X
<hr/>						
The Australian National University Library	X			X		
Edith Cowan University Library	X				X	
The University of Melbourne - University Library			X	X		
University of New South Wales Library			X	X		
Northern Territory University Library	X				X	
The University of Queensland Library			X		X	
Queensland University of Technology Library	X			X		
University of South Australia Library	X				X	
Swinburne University of Technology Library	X				X	
The University of Sydney Library			X	X		
University of Technology Sydney - University Library	X			X		
University of Tasmania Library	X				X	
<hr/>						
The University of British Columbia Library			X	X		
The University of Guelph Library	X			X		
Bibliothèque De L'Université Laval			X		X	
McGill University Libraries			X	X		
McMaster University Libraries	X			X		
The University of Manitoba Libraries			X		X	
Queen's University Libraries			X	X		
University of Saskatchewan Libraries	X				X	
The University of Western Ontario Library System			X		X	
York University Libraries			X		X	

**Tabelle 3b:** Geschichtete Stichprobe: Einzelbetrachtung (englischsprachiger Raum)

Im Hinblick auf nationale Bibliotheksstatistiken ist Österreich als Negativbeispiel zu erwähnen. In Österreich gibt es überhaupt keine offizielle und öffentlich

zugängliche nationale Bibliotheksstatistik. Die jährlich in der österreichischen Bibliothekarszeitschrift<sup>13</sup> veröffentlichte „Österreichische Bibliotheksstatistik“ ist zwar ein positiver Ansatz, doch vermögen die elf darin enthaltenen Datensätze pro Bibliothek bestenfalls einen groben Überblick zu geben. Zudem werden die Daten mit einer zweijährigen Verspätung veröffentlicht;

Aus erhebungstechnischer Sicht weitaus günstiger ist die Situation in Deutschland. Hier existiert mit der jährlichen, jeweils zur Jahresmitte des Folgejahres erscheinenden Deutschen Bibliotheksstatistik eine äußerst detaillierte nationale Bibliotheksstatistik. Leider ist mit der 1999 beschlossenen Auflösung des Deutschen Bibliotheksinstituts, das für die Erstellung der Deutschen Bibliotheksstatistik zuständig ist, der Fortbestand dieser zu Beginn der 70-er Jahre eingeführten und ständig weiterentwickelten Statistik – zumindest in ihrem derzeitigen Ausmaß – in Frage gestellt. Die Schweiz verfügt mit den jährlich vom Bundesamt für Statistik herausgegebenen Statistischen Übersichten für Schweizerische Bibliotheken über eine nationale Bibliotheksstatistik, die sich zwar hinsichtlich des Umfanges keinesfalls mit der Deutschen Bibliotheksstatistik messen kann, aber doch einen guten Einblick in das Schweizerische Universitätsbibliothekswesen gibt.

Für die USA und Kanada kann aufgrund der Beschränkung der Stichprobe auf ARL<sup>14</sup>- Mitgliedsbibliotheken und der daraus resultierenden Verwendung der ARL-Statistik zwar keine für das gesamte Universitätsbibliothekswesen gültige Aussage hinsichtlich des Umfanges und der Qualität nationaler Statistiken getätigt werden, aber für den Bereich der ARLMitgliedsbibliotheken sind umfangreiche und perfekt aufbereitete Daten kostenlos über das Internet verfügbar. Der Aufwand für die Grunderhebung verringerte sich hier sogar noch weiter, da man aus den ARL-Statistik mit Hilfe von gezielten Datenbankabfragen eine exportfähige Datenliste erstellen kann, in der ausschließlich die benötigten Daten enthalten sind. Die einzigen Daten, die den Internetseiten der untersuchten Bibliotheken entnommen werden mußten, waren die Angaben zu den Öffnungszeiten. Für Australien gestaltete sich die Grunderhebung ebenfalls eher einfach, da mit der CAUL<sup>15</sup>-Statistik eine umfassende und über das Internet abrufbare nationale Bibliotheksstatistik existiert. Lediglich die Erstellung individueller exportfähiger Datenlisten wird im Gegensatz zur ARL-Statistik nicht angeboten. Die Internetseiten der einzelnen Bibliotheken mußten ebenfalls nur zur ergänzenden Ermittlung der Öffnungszeiten herangezogen werden.

Als Ergebnis der Grunderhebung lagen 132 mehr oder minder vollständig ausgefüllte Datenblätter vor. Ziel der ergänzenden Erhebung war es nun, einerseits die vorhandenen Lücken zu schließen und andererseits die bereits gewonnenen Daten einer Überprüfung zu unterziehen. Zu diesem Zweck wurde in der zweiten Oktoberhälfte 1999 allen Leitern der in der Stichprobe enthaltenen Bibliotheken das jeweilige Datenblatt mit der Bitte um Ergänzung, Überprüfung und gegebenenfalls Korrektur bzw. Aktualisierung via Email zugesandt. Die dabei

---

<sup>13</sup> Mitteilungen der vöb.

<sup>14</sup> Association of Research Libraries (Nordamerikanische Bibliotheksvereinigung)

<sup>15</sup> Council of Australian University Librarians

erzielten Rücklaufquoten betrugen 67 % in Österreich, 38 % in Deutschland, 50 % in der Schweiz, 33 % in den USA, 42 % in Australien und 50 % in Kanada, woraus sich eine Gesamtrücklaufquote von 41 % ergab.

Das theoretische Optimalergebnis der Erhebungsphase wären 132 vollständig und richtig ausgefüllten Datenblätter; das heißt, für jede benötigte Datenart wären 132 Datensätze vorhanden. In welchem Umfang bei der vorliegenden Untersuchung trotz zweistufiger Vorgangsweise Abstriche von diesem Optimalergebnis in Kauf zu nehmen waren, faßt Tabelle 4 zusammen. Dabei ist jedoch zu betonen, dass von manchen Bibliotheken, die im Zuge der ergänzenden Erhebung überhaupt nicht antworteten, aufgrund einer relativ erfolgreichen Grunderhebung mehr Daten vorlagen als von anderen Bibliotheken, über die in beiden Stufen Daten gewonnen werden konnten. Was die Qualität der erhobenen Daten betrifft, so ist darauf hinzuweisen, dass im Falle von Differenzen der ergänzenden Erhebung Priorität eingeräumt wurde.

Aus Tabelle 4 wird ersichtlich, dass auf das Optimalergebnis von 1188<sup>16</sup> für die Auswertung zur Verfügung stehenden Datensätzen lediglich 89 Datensätze fehlen; dies bedeutet, dass das Optimalergebnis zu 93 % erreicht wurde. Somit ist eine beinahe ideale Datenbasis für die nachfolgende Hypothesentestung vorhanden.

<b>Datenart</b>	<b>Vorhandene</b>
Ausgaben für Bestandsaufbau (Erwerbung)	131
Personal (VZÄ)	132
Buchbestand	124
Zeitschriften (laufend gehaltene)	129
Arbeitsplätze (Leseplätze)	62
Entlehnungen	129
Regelöffnungszeiten/Woche: Lesesaal	130
Buchzuwachs	130
Kernnutzer	132

**Tabelle 4:** Datensätze für die Auswertungen

## **5. Ergebnisse**

Im Zuge der Hypothesenprüfung werden in einem ersten Schritt für jede Gruppe von Bibliotheken die entsprechenden Kennzahlenmittelwerte für alle betrachteten Teilleistungsbereiche sowie der Kennzahlenmittelwert für den Gesamtleistungsbereich berechnet. Diese Kennzahlenmittelwerte werden einander gegenübergestellt, um zu überprüfen, ob sich die Gruppen in bezug auf die einzelnen Kennzahlen für die Teilleistungsbereiche bzw. in bezug auf die Kennzahl für den Gesamtleistungsbereich signifikant voneinander unterscheiden. Um ein höchstmögliches Maß an Übersichtlichkeit zu gewährleisten, erfolgt die

<sup>16</sup> Diese Zahl ergibt sich aus 9 (Datensätzen) multipliziert mit 132 (untersuchten Universitätsbibliotheken).

Gegenüberstellung der Mittelwerte bzw. die Darstellung der Ergebnisse der Mittelwertvergleiche in Form von Tabellen. Die Mittelwertvergleiche werden nach folgendem Schema vorgenommen: in einem ersten Schritt wird mit Hilfe des Kolmogorov-Smirnov Anpassungstests überprüft, ob hinsichtlich der Ausprägungen der einzelnen Leistungskennzahlen eine Normalverteilung vorliegt; wenn dies der Fall ist, schließt daran eine Prüfung auf Varianzhomogenität mittels Levene - Test an. Liegen sowohl Normalverteilung als auch Varianzhomogenität vor, kommt der parametrische t - Test für den Vergleich von zwei unabhängigen Stichproben zum Einsatz. Ansonsten wird auf den nichtparametrischen Mann - Whitney - Test zurückgegriffen. Sämtliche Tests erfolgen unter Einsatz des Programmes SPSS für Windows, Version 7.5.

## 5.1 Sprachraumvergleich

Beim Sprachraumvergleich geht es um die Beantwortung der Frage, ob es Leistungsunterschiede zwischen Universitätsbibliotheken im deutschsprachigen und solchen im englischsprachigen Raum gibt. Dabei stehen im Rahmen der gezogenen Stichprobe 70 Bibliotheken im deutschsprachigen Raum 62 im Bibliotheken im englischsprachigen Raum gegenüber.

Tabelle 5 zeigt, dass bezüglich des Leistungsvergleiches auf Basis der einfachen Kennzahl „Regelöffnungsstunden pro Woche“ bzw. hinsichtlich dreier Leistungsvergleiche auf Basis zusammengesetzter Kennzahlen („Kernnutzer pro Leseplatz“, „Kernnutzer pro VZÄ“, „Erwerbungsausgaben pro Kernnutzer in ATS“) signifikante Leistungsunterschiede zwischen den Universitätsbibliotheken in den beiden Sprachräumen vorliegen. Dabei schneiden die Universitätsbibliotheken im englischsprachigen Raum durchwegs besser ab: sie weisen deutlich längere Öffnungszeiten auf, die infrastrukturelle und personelle Ausstattung ist besser und auch die Erwerbungsausgaben pro Kernnutzer sind höher.

Kennzahl	Anzahl der Bibliotheken		Mittelwert		Signifikanz
	Deutschsprachiger Raum	Englischsprachiger Raum	Deutschsprachiger Raum	Englischsprachiger Raum	
Regelöffnungsstunden pro Woche	70	60	66	99	<0,001
Buchbestand pro Kernnutzer	69	55	128	178	0,054
Buchzuwachs pro Kernnutzer	68	62	3,8	3,4	0,103
Kernnutzer pro Leseplatz	37	25	29	12	<0,001
Kernnutzer pro VZÄ	70	62	178	92	<0,001
Entlehnungen pro Kernnutzer	70	59	39	33	0,232
Erwerbungsausgaben pro Kernnutzer in ATS	69	62	3.039	4.160	0,011
Relative Effizienz	66	52	66,93	68,01	0,751

**Tabelle 5:** Sprachraumvergleich

Trotz zahlreicher Indizien für eine generelle Leistungsunterschied zwischen Universitätsbibliotheken im deutsch- und im englischsprachigen Raum auf Ebene der Teilleistungsbereiche ergibt ein Leistungsvergleich auf Basis der Spitzenkennzahl keine signifikanten Leistungsunterschiede, womit diesbezüglich

die Nullhypothese, dass es keine Leistungsunterschiede zwischen den Bibliotheken in den zwei betrachteten Sprachräumen gibt, anzunehmen ist. Dieses Ergebnis ist natürlich sehr stark von der getroffenen Auswahl der für die Berechnung der relativen Effizienz herangezogenen Input- und Outputfaktoren abhängig. Diese Auswahl begünstigt offensichtlich die Bibliotheken im deutschsprachigen Raum, da sich unter den vier verwendeten Outputfaktoren mit den beiden Faktoren „Buchzuwachs“ und „Anzahl der Entlehnungen“ immerhin zwei befinden, bezüglich derer die Bibliotheken im deutschsprachigen Raum besser abschneiden. Zudem kann die weitaus bessere infrastrukturelle Ausstattung der Bibliotheken im englischsprachigen Raum, die sich höchstwahrscheinlich sowohl in einer stärkeren Präsenznutzung als auch in einer höheren Benutzerzufriedenheit widerspiegeln dürfte, im Rahmen des Gesamtvergleiches nicht berücksichtigt werden.

## 5.2 Größenvergleich

Im Rahmen des Größenvergleiches wird untersucht, ob Leistungsunterschiede zwischen kleinen und großen Universitätsbibliotheken existieren, wobei — wie bereits ausgeführt — eine Beschäftigtenzahl von 175 VZÄ als Abgrenzungskriterium zwischen den beiden Gruppen herangezogen wird. Zu beachten ist dabei, dass die Einteilung in kleine und große Bibliotheken jener in Bibliotheken im deutsch- und im englischsprachigen Bereich relativ ähnlich ist, da ein Großteil (93 %) der Bibliotheken im deutschsprachigen Raum zur Gruppe der kleinen Bibliotheken gehört, während mehr als 75 % der Bibliotheken im englischsprachigen Raum der Gruppe der großen Bibliotheken zuzurechnen sind.

Kennzahl	Anzahl der Bibliotheken		Mittelwert		Signifikanz
	Groß	Klein	Groß	Klein	
Regelöffnungsstunden pro Woche	51	79	100	69	<0,001
Buchbestand pro Kernnutzer	51	73	188	124	0,018
Buchzuwachs pro Kernnutzer	52	78	3,8	3,5	0,453
Kernnutzer pro Leseplatz	15	47	17	24	0,048
Kernnutzer pro VZÄ	52	80	86	172	<0,001
Entlehnungen pro Kernnutzer	51	78	35	37	0,927
Erwerbungsausgaben pro Kernnutzer in ATS	51	80	4.537	2.953	<0,001
Relative Effizienz	48	70	65,93	68,43	0,499

**Tabelle 6:** Größenvergleich

Aus Tabelle 6 ist ersichtlich, dass sich kleine und große Universitätsbibliotheken hinsichtlich der Regelöffnungsstunden pro Woche signifikant voneinander unterscheiden: große Bibliotheken haben im Durchschnitt um 31 Stunden pro Woche länger geöffnet. Längere Öffnungszeiten sind natürlich aufgrund der, sowohl absolut als auch relativ gesehen, besseren Personalausstattung in großen Bibliotheken leichter realisierbar. Die relativ bessere Personalausstattung großer Bibliotheken zeigt sich unter anderem in der signifikant geringeren Anzahl an Kernnutzern pro VZÄ. Signifikante Leistungsunterschiede zwischen kleinen und großen Bibliotheken bestehen bei den zusammengesetzten Kennzahlen weiters

bezüglich der Kennzahlen „Buchbestand pro Kernnutzer“, „Kernnutzer pro Leseplatz“ und „Erwerbungsausgaben pro Kernnutzer“: hier weisen die großen Bibliotheken im Durchschnitt wesentlich bessere, das heißt im ersten und dritten Fall höhere und im zweiten Fall geringere, Werte auf. Insgesamt entsprechen die Unterschiede zwischen kleinen und großen Bibliotheken in etwa jenen zwischen Bibliotheken im deutsch- und im englischsprachigen Raum. Während dabei die Differenzen hinsichtlich der Kennzahlen „Buchbestand pro Kernnutzer“ und „Erwerbungsausgaben pro Kernnutzer“ stärker auf Größendifferenzen zurückzuführen sind, beruht die Differenz hinsichtlich der Kennzahl „Kernnutzer pro Leseplatz“ eher auf Sprachraumunterschieden. Die übrigen zwei signifikanten Leistungsdifferenzen lassen sich nicht näher zuordnen. Der Gesamtleistungsvergleich auf Basis der mit Hilfe der DEA ermittelten Spitzenkennzahl ergibt wiederum keinen signifikanten Leistungsunterschied, womit auch für den Ländervergleich die Nullhypothese anzunehmen ist.

### 5.3 Rankingvergleich

Im Zuge des Rankingvergleiches wird der Frage nachgegangen, ob zwischen Bibliotheken an renommierten und Bibliotheken an weniger renommierten Universitäten im englischsprachigen Raum Leistungsunterschiede vorliegen.

Kennzahl	Anzahl der Bibliotheken		Mittelwert		Signifikanz
	Bibliotheken an renommierten Universitäten	Bibliotheken an weniger renommierten Universitäten	Bibliotheken an renommierten Universitäten	Bibliotheken an weniger renommierten Universitäten	
Regelöffnungszeiten pro Woche	30	30	99	98	0,809
Buchbestand pro Kernnutzer	27	28	254	105	<0,001
Buchzuwachs pro Kernnutzer	31	31	4,6	2,2	0,003
Kernnutzer pro Leseplatz	10		14	11	0,304
Kernnutzer pro VZÄ	31	31	78	107	0,009
Entlehnungen pro Kernnutzer	29	30	40	26	0,002
Erwerbungsausgaben pro Kernnutzer in ATS	31	31	5.591	2.728	0,001
Relative Effizienz	25	27	69,77	66,39	0,379

**Tabelle 7:** Rankingvergleich

Dabei ist zu beachten, dass Bibliotheken an renommierten Universitäten in der Regel größer sind als solche an weniger renommierten Universitäten. Deshalb könnten etwaige Leistungsunterschiede auch größenbedingt sein. Allerdings sind die in Tabelle 7 enthaltenen Leistungsdifferenzen bezüglich der zusammengesetzten Kennzahlen „Buchbestand pro Kernnutzer“, „Buchzuwachs pro Kernnutzer“ und „Ausgaben für Erwerbung pro Kernnutzer“ jedenfalls so hoch, dass sie keinesfalls alleine auf Größendifferenzen zurückgeführt werden können. Bibliotheken an renommierten Universitäten weisen zwar ein signifikant personalintensiveres Verhältnis von Kernnutzern pro VZA als solche an weniger renommierten Universitäten auf, aber die diesbezügliche Leistungsdifferenz ist deutlich geringer als bei einem Größenvergleich. Dies weist auf einen sehr überlegten Ressourceneinsatz in Bibliotheken an renommierten Universitäten hin.



Signifikant ist weiters die Differenz bei den Entlehnungen pro Kernnutzer. Ob Bibliotheken an renommierten Universitäten neben der intensiveren Inanspruchnahme der Entlehnung auch eine intensivere Präsenznutzung aufweisen, wäre eine interessante Fragestellung für weiterführende Untersuchungen. Trotz dieser zahlreichen signifikanten Leistungsdifferenzen in verschiedenen Teilleistungsbereichen gibt es auch hier keine signifikante Gesamtleistungsdifferenz, die Nullhypothese ist daher anzunehmen.

#### **5.4 Trägervergleich**

Der Trägervergleich, welcher auf den Bereich der US-amerikanischen Bibliotheken an renommierten Universitäten beschränkt ist, dient der Klärung der Frage, ob es zwischen Bibliotheken an staatlichen und Bibliotheken an privaten Universitäten signifikante Leistungsunterschiede gibt. Nachdem bereits der Rankingvergleich gezeigt hat, dass Bibliotheken an renommierten Universitäten jenen an weniger renommierten Universitäten in einigen Teilleistungsbereichen deutlich überlegen sind, läßt sich anhand von Tabelle 8 erkennen, dass es innerhalb der Gruppe der Bibliotheken an renommierten Universitäten — allerdings beschränkt auf die USA — nochmals große Leistungsdifferenzen zwischen Bibliotheken an staatlichen und solchen an privaten Universitäten gibt. Bibliotheken an privaten Universitäten weisen einen mehr als doppelt so hohen Buchzuwachs pro Kernnutzer auf, was den ohnedies bereits vorhandenen großen Abstand hinsichtlich des Buchbestandes pro Kernnutzer noch weiter wachsen lässt. Dementsprechend höher sind auch die Ausgaben für Erwerbung pro Kernnutzer in Bibliotheken an privaten Universitäten. Was die personelle und infrastrukturelle Ausstattung betrifft, so weisen die Bibliotheken an privaten Universitäten derartig „günstige“ Werte auf, dass in weiterführenden Analysen zu hinterfragen wäre, ob diesem hohen Input auch ein entsprechend hoher Output hinsichtlich der im Rahmen dieser Untersuchung nicht einbezogenen Outputfaktoren, wie etwa Präsenznutzung oder Benutzerzufriedenheit, gegenübersteht.

Sämtliche Teilleistungsvergleiche weisen ein Ergebnis zugunsten der Bibliotheken an privaten Universitäten auf. Dass es beim Gesamtleistungsvergleich zu einem umgekehrten Ergebnis kommt, ist kein systematischer Fehler, sondern liegt daran, dass es sich dabei um einen Vergleich auf Basis der durch die gewählten Input- und Outputfaktoren definierten Gesamtproduktivität handelt, der insoweit von den Ergebnissen der Teilleistungsvergleiche unabhängig ist, als es sich dabei um keine Vergleiche auf Basis von Teilproduktivitäten handelt.<sup>8</sup> Die im gegenständlichen Fall signifikant unterschiedlichen Teilleistungsbereiche sind ausschließlich solche, die nicht auf Basis von Teilproduktivitäten miteinander verglichen werden. Aufgrund mangelnder Signifikanz muss trotz großer zahlenmäßiger Differenz an der Nullhypothese, gemäß der es keine Leistungsunterschiede zwischen Bibliotheken an staatlichen und solchen an privaten Universitäten gibt, festgehalten werden.

Kennzahl	Anzahl der Bibliotheken		Mittelwert		Signifikanz
	Staatlich	Privat	Staatlich	Privat	
Regelöffnungszeiten pro Woche	10	10	104	108	0,587
Buchbestand pro Kernnutzer	10	10	174	433	0,002
Buchzuwachs pro Kernnutzer	10	10	3,7	8,6	0,006
Kernnutzer pro Leseplatz	2	1	18	5	0,604
Kernnutzer pro VZÄ	10	10	74	33	<0,001
Entlehnungen pro Kernnutzer	9	10	28	53	0,006
Erwerbungsausgaben pro Kernnutzer in ATS	10	10	3.978	10.433	<0,001
Relative Effizienz	9	10	80,23	65,98	0,072

**Tabelle 8:** Trägervergleich

## 6. Resümee

Zusammenfassend läßt sich sagen, dass teilweise deutliche Leistungsdifferenzen zwischen den in dieser Arbeit untersuchten Gruppen von Universitätsbibliotheken vorliegen. Bei den auf Basis der einfachen und

zusammengesetzten Kennzahlen untersuchten Teilleistungsbereichen zeigen sich umfassende Leistungsdifferenzen zwischen großen und kleinen Universitätsbibliotheken, zwischen Bibliotheken an renommierten und solchen an weniger renommierten Universitäten sowie zwischen Bibliotheken an staatlichen und Bibliotheken an privaten Universitäten. Es liegen hier jeweils in fünf Teilleistungsbereichen signifikante Leistungsunterschiede vor. Beim Sprachraumvergleich zwischen Universitätsbibliotheken im deutschsprachigen und solchen im englischsprachigen Raum treten signifikante Differenzen in vier Teilleistungsbereichen auf. Die jeweils festgestellten Leistungsunterschiede könnten als Basis für detaillierte Analysen herangezogen werden. Im Hinblick auf die Spitzenkennzahl bestehen keine signifikanten Leistungsunterschiede. Eine Überprüfung der zentralen Hypothesen auf Basis der mit Hilfe der DEA unter Berücksichtigung der Mitarbeiterzahl und des Buchbestandes als Inputfaktoren und der Anzahl der laufend gehaltenen Zeitschriften, der Anzahl der Entlehnungen, der Öffnungszeiten pro Woche und des Buchzuwachses als Outputfaktoren berechneten Gesamtleistung führt somit stets zu einer Annahme der jeweiligen Nullhypothese.

## **Literatur**

Banker, R./Charnes, A./Cooper, W.: Some models for estimating technical and scale inefficiencies in data envelopment analysis, in: *Management Science* 30 (1984), 1078-1093

Bannister, M./Rochester, M.: Performance measures for NSW TAFE libraries: what can we learn from the literature?, in: *AARL* December/1997, 281-296

Boekhorst, P.: Leistungsmessung in wissenschaftlichen Bibliotheken: Neue Initiativen, in: *NfD* 46 (1995), 121-126

Brophy, P.: Performance measurement in academic libraries: a polytechnic perspective, in: *British Journal of Academic Librarianship* 2/1989, 99-110

Cantner, U./Hanusch, H.: Effizienzanalyse mit Hilfe der Data Envelopment Analysis, in: *WiSt* 5/1998, 228-237

Chen, T.: An evaluation of the relative performance of university libraries in Taipei, in: *Library Review* 3/1997, 190-200

Deutsches Bibliotheksinstitut: Qualität und Leistung — Bibliotheken auf dem Prüfstand, Berlin : Deutsches Bibliotheksinstitut, 1996

ISO 11620: Information and documentation – Library performance indicators, Geneve : International Organization for Standardization, 1998

King Research Ltd (Hrsg.): Keys to success: performance indicators for public libraries, London : HMSO, 1990

Kommission des Deutschen Bibliotheksinstituts für Organisation und Betrieb, Expertengruppe Controlling in Öffentlichen Bibliotheken: Controlling für Öffentliche Bibliotheken: Wegweiser zu einer effizienten Steuerung, Berlin : Deutsches Bibliotheksinstitut, 1994

McDonald, J./Micikas, L.: Academic libraries – The dimension of their effectiveness, Westport, Connecticut/London : Greenwood Press, 1994

Moore, N.: Measuring the performance of public libraries – A draft manual, Paris : UNESCO, 1989

Poll, R./Boekhorst, P.: Measuring quality – International guidelines for performance measurement in academic libraries, München u.a. : Saur, 1996

Pritchard, S.: Determining quality in academic libraries, in: *Library Trends* 3/1996, 572-594

Pröhl, M./Windau, B. (Hrsg.): Betriebsvergleich an Öffentlichen Bibliotheken – Band 1: Empfehlungen und Arbeitsmaterialien für ein output-orientiertes Berichtswesen, Gütersloh : Verlag Bertelsmann Stiftung, 1997

Schäfer, S./Seibt, D.: Benchmarking – eine Methode zur Verbesserung von Unternehmensprozessen, in: BFuP 4/1998, 365-401

Schulte-Nölke, P.: Das Problem der betrieblichen Leistungsmessung von Bibliotheken, in: Bibliothek 1/1993, 7-28

Sumsion, J.: Practical performance indicators – 1992, Loughborough : Library and Information Statistics Unit (LISU), 1993

Van House, N./Weil, B./McClure, C.: Measuring academic library performance – A practical approach, Chicago/London : American Library Association, 1990

Ward, S. u.a.: Library performance indicators and library management tools, Brussels/Luxembourg : European Commission, 1995



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 227 – 243

## Informationskompetenz europäischer Volkswirtschaften

Willi Bredemeier, Wolfgang G. Stock  
Institute for Information Economics (IIE)  
Erzberger Str. 11+15, D-45527 Hattingen  
E-Mail: 101346.3675@compuserve.com

Fachhochschule Köln - Fachbereich Bibliotheks- und Informationswesen  
Claudiusstr. 1, D-50678 Köln  
E-Mail: Wolf.Stock@uni-koeln.de

### Zusammenfassung

#### Problemstellung

Es wird versucht, quantitative Werte zur Messung der Informationskompetenz von Volkswirtschaften zu entwickeln. Zudem sollen Korrelationen zwischen der Informationskompetenz und der wirtschaftlichen Leistungsfähigkeit darüber Auskunft geben, wie Gesamtwirtschaft und Informationswirtschaft eines Landes zusammenhängen. - *Operationalisierung*. "Informationskompetenz" wird durch den Umsatz auf nationalen Business-to-Business-Märkten operationalisiert. Abgeleitete Größen sind Pro-Kopf-Import- und Exportvolumen, Informationsproduktion und -konsumtion (insgesamt und pro Kopf) sowie der Anteil der Informationswirtschaft am jeweiligen nationalen Bruttoinlandsprodukt (BIP). *Datenbasis*. Im Rahmen der "Member States' Study" der Europäischen Kommission wurden für den gesamten Europäischen Wirtschaftsraum (EWR; Staaten der Europäischen Union sowie Norwegen und Island) für die Jahre 1994 und 1997 Umsatzzahlen der 17 Länder erhoben. Diese Zahlen wurden auf Plausibilität überprüft und unserer Fragestellung zugrundegelegt. *Ergebnisse*. Der Anteil der Informationswirtschaft am BIP ist derart gering (EWR: 0,070% im Jahr 1997), daß er nahezu ignoriert werden kann. Die Korrelation zwischen dem BIP pro Kopf und der Informationskonsumtion pro Kopf ist jedoch äußerst hoch (+0,715 im Jahr 1994 und +0,642 im Jahr 1997), ebenso die Korrelation zwischen BIP pro Kopf und Informationsimport pro Kopf (+0,588 im Jahr 1994 sowie +0,583 im Jahr 1997). Zwischen dem BIP pro Einwohner und Informationsexport besteht demgegenüber kein Zusammenhang. *Interpretation*. Eine Volkswirtschaft ist umso leistungsfähiger, je größer Informationskonsum und -import sind. Export und Leistungsfähigkeit hängen - ganz entgegen dem Trend auf anderen Gütermärkten - nicht zusammen. Zur Steigerung der wirtschaftlichen Leistungsfähigkeit eines



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

Landes kommt es demnach vor allem darauf an, sowohl die im eigenen Land für das eigene Land bestimmten Informationen als auch im Ausland produzierte Informationen optimal zu nutzen. Hierin liegt eine Besonderheit der sog. "Neuen Ökonomie", die das Wirtschaften mit Informationen in Netzen thematisiert. Klar ist auch: Um die Informationskompetenz einer Volkswirtschaft zu erhöhen, muß man deren Mitglieder qualifizieren.

## 1. Informationskompetenz

Kann man die Informationskompetenz der Wirtschaft eines ganzen Landes erfassen oder sogar messen? Werden Vergleiche zwischen Volkswirtschaften möglich? Lassen sich Korrelationen zwischen der Ausprägung der Informationskompetenz und der wirtschaftlichen Leistungsfähigkeit eines Landes errechnen? Alle Fragen sind zu bejahen, vorausgesetzt, man akzeptiert gewisse - durchaus willkürliche - Operationalisierungen der Grundbegriffe sowie einige Vorgaben methodischer Natur.

Wir operationalisieren den Term "Informationskompetenz einer Volkswirtschaft" für unsere Zwecke als Produktion und Konsumtion elektronischer Informationsdienste im Geschäftsleben. Damit beschränken wir unsere Analyse auf Business-to-Business-Märkte und vernachlässigen die Consumer-Märkte. Es geht also nur um die nationale Informationskompetenz innerhalb der professionellen Märkte in Wirtschaft, Wissenschaft und Verwaltung. "Elektronische Informationsdienste" definieren wir als elektronische Produkte, die entweder online (über spezifische Datennetze wie X.25 oder das Internet bzw. via Videotex) oder offline (als CD-ROM oder Diskette) verteilt werden und bei denen der Informationsinhalt ("Content", Wissen) im Mittelpunkt steht, also als Gesamtheit der Produkte, die die Branche der Informationswirtschaft ("Content-Industry") in kommerzieller Absicht auf dem Markt anbietet (vgl. Stock 2000, S. 20 ff.). Unser Indikator für die Ausprägung der Informationskompetenz ist der in der empirischen Wirtschaftsforschung üblicherweise herangezogene (nationale) Umsatz auf den Märkten unserer Branche. Nicht erfassen können wir mit diesem Ansatz Informationsprodukte, die subventioniert werden wie z.B. "Universaldienste", die kostenlos im World Wide Web aufliegen (etwa die "MEDLINE").

## 2. Datenbasis

Datenmaterial liegt zur Informationswirtschaft aller Länder des Europäischen Wirtschaftsraumes (das sind die Länder der Europäischen Union sowie Norwegen und Island) für die Jahre 1994 und 1997 vor. Erhoben wurden sie von 17 nationalen Projektpartnern in ihren jeweiligen Heimatländern im Rahmen der sog. "Member States' Study" der Europäischen Kommission. Die Erhebungsmethode war einheitlich, geregelt durch ein Handbuch (vgl. Schwuchow/Bredemeier/Graumann 1998). Die Ergebnisse auf europäischer Ebene wurden vom "Institute for Information Economics" aggregiert. Für das Berichtsjahr 1994 liegen die Ergebnisse ("MS-Study I") in mehreren

Aufbereitungen vor (Hauptbericht: DG XIII/E 1996; Kurzfassungen: Bredemeier 1996; Schwuchow 1996; Schwuchow/Graumann/Bredemeier 1997). Eine kurze methodische Diskussion fand im "Journal of Information Sciences" statt (vgl. White 1998; Schwuchow 1998). Die Ergebnisse wurden auch von ökonomischer Seite wahrgenommen (vgl. Handtke 1997). Die Daten des Jahres 1997 sind Anfang 2000 aufgearbeitet worden ("MS-Study II"; vgl. Bredemeier 2000).

Einige Werte sind von den nationalen Partnern nicht erhoben worden, so daß an deren Stelle Schätzungen treten mußten. Von der (zum Teil kleinen) Stichprobe ist jeweils auf den gesamten nationalen Informationsmarkt hochgerechnet worden. Durch solche Schätzungen und Hochrechnungen sind in den Zahlen gewisse Unsicherheiten enthalten. Da die hier vorliegende Studie ausschließlich auf hochaggregierte Daten (auf nationaler sowie europäischer Ebene) aufbaut, dürften die Ergebnisse - dank des Gesetzes der großen Zahl - zumindest in der Tendenz zutreffen.

### **3. Volumen der Informationsmärkte**

Elektronische Informationen sind ein Wirtschaftsgut besonderer Art. Es ist potentiell überall und jederzeit konsumierbar, egal, wo es produziert wurde und wo es physisch gespeichert ist. Der Online-Informationsmarkt ist demnach stets der Weltmarkt. So gesehen, kann man kaum Exporte und Importe bestimmen. Elektronische Informationen befinden sich jedoch im Eigentum bestimmter Unternehmen. Der Hauptsitz dieser Unternehmen läßt sich einem Land zuordnen. Alle Informationen, die solche Unternehmen vorhalten, werden dem Sitzland des Unternehmens als "Informationsproduktion" zugerechnet. Die Umsätze der deutschen Reuters AG in Frankfurt gelten demnach als britische Informationsproduktion, ist der Hauptsitz von Reuters doch London. Dieses sog. "Headquarter-Prinzip" wurde in MS-Study I und II zugrundegelegt. "Informationsproduktion" wird gemessen durch den Umsatz im Inland sowie aller Exporte von Unternehmen, die im betreffenden Land ihren Hauptsitz haben. "Informationskonsumtion" erfassen wir durch den Umsatz informationswirtschaftlicher Produkte *in* dem betreffenden Land, also durch die Summe aus Inlandsumsatz und Importen. Der Gesamtmarkt eines Landes ist die Summe aus Inlandsumsatz, Importen und Exporten.

Ein methodisches Problem des "I-Commerce" (Handel mit Informationen) gilt es zu beachten. Bei Sachgütern sind Umsatz und Produktion direkt von-einander abhängig. Bei festen Stückkosten können wir aus dem Umsatz die Anzahl der produzierten (und verkauften) Güter errechnen. Bei Informationen ist dies anders. Einmal produzierte Informationen können ohne großen Aufwand mehrfach verkauft werden, was sich nur auf den Umsatz, nicht aber auf die Produktionszahl auswirkt. Unser Begriff von "Informationsproduktion" umfaßt auch die Mehrfachverkäufe gleicher Informationen.

Tabelle 1 zeigt die Basiszahlen für die Informationsmärkte der Länder des Europäischen Wirtschaftsraumes sowie für den EWR insgesamt für die Jahre 1994 und 1997 jeweils in konstant gehaltenen 1994er ECU (d.h. für das Jahr 1997

wurde die Inflation seit 1994 herausgerechnet). Der gesamte EWR-Informationsmarkt verbesserte sich von 5.446,6 Mio. ECU 1994 auf 9.262,7 Mio. ECU 1997, was einer Steigerungsrate von 70,1% über die drei Jahre entspricht. Stark zugenommen haben (oder besser erfaßt wurden) die Importe, die von 291,2 Mio. ECU auf 2.108,7 Mio. ECU anstiegen.

Bei den nationalen Informationsmärkten des EWR dominiert Großbritannien; hier steigen die Umsätze um 31,1% von 3,6 Mrd. ECU (1994) auf 4,7 Mrd. ECU (1997). Der französische Informationsmarkt - mit 0,8 Mrd. ECU im Jahr 1994 noch der zweitgrößte Markt - wächst nur um 8% auf knapp 0,9 Mrd. ECU (1997) und fällt damit hinter Deutschland und den Niederlanden auf Platz vier zurück. Der deutsche Informationsmarkt steigt in etwa genauso stark wie der EWR-Wert und erreicht 1997 ein Volumen von 1.198 Mio. ECU. Einen Riesensprung machen die Niederlande und verdreifachen ihr Marktvolumen auf 1.150 Mio. ECU. Große Steigerungsraten (allerdings bei schwachen Ausgangswerten) realisieren Spanien und Griechenland.

Märkte mit starken Anteilen inländischer Produktion und Konsumtion sind im Jahr 1997 Griechenland (72,3% Inlandsmarkt), Finnland (68,9%), Frankreich (67,4%), Italien (64,1%) und Norwegen (62,3%). Schwache Inlandsmärkte bestehen in Island (5,3%), Luxemburg (12,4%), Irland (20,4%) und - dort jedoch auf einem hohen Niveau - Großbritannien (20,5%).

Gab es 1994 noch drei Länder mit einer positiven Außenhandelsbilanz (Großbritannien, dann mit großem Abstand Frankreich und die Niederlande), so schrumpft diese Anzahl 1997 auf zwei (Großbritannien und die Niederlande) zusammen. Unternehmen mit einem Headquarter im Vereinigten Königreich (darunter an erster Stelle die Firma Reuters) haben 1997 für 2.698,3 Mio. ECU Umsätze außerhalb des Landes; niederländische Unternehmen (darunter etwa der Verlag Elsevier mit dem Online-Archiv Lexis-Nexis) exportieren Informationsdienste im Wert von 403,3 Mio. ECU.



	1994				1997			
	Gesamt	Inland	Importe	Exporte	Gesamt	Inland	Importe	Exporte
Großbritannien	3615,6 100,0 %	554,9 15,3 %	737,4 20,4 %	2323,3 64,3 %	4740,1 100,0 %	970,5 20,5 %	1071,3 22,6 %	2698,3 56,9 %
Frankreich	821,7 100,0 %	624,2 75,9 %	79,5 9,7 %	118,0 14,4 %	887,5 100,0 %	597,8 67,4 %	149,1 16,8 %	140,6 15,8 %
Deutschland	683,2 100,0 %	278,7 40,8 %	315,8 46,2 %	88,7 13,0 %	1198,0 100,0 %	417,7 34,9 %	700,1 58,4 %	80,2 6,7 %
Italien	359,2 100,0 %	258,7 72,0 %	94,7 26,4 %	5,8 1,6 %	690,0 100,0 %	442,3 64,1 %	238,5 34,6 %	9,2 1,3 %
Niederlande	349,0 100,0 %	187,9 53,8 %	71,5 20,5 %	89,6 25,7 %	1150,0 100,0 %	500,4 43,5 %	246,3 21,4 %	403,3 35,1 %
Schweden	182,0 100,0 %	108,4 59,6 %	67,5 37,1 %	6,1 3,3 %	191,7 100,0 %	110,5 57,6 %	74,1 38,7 %	7,1 3,7 %
Dänemark	167,5 100,0 %	117,0 69,8 %	39,5 23,6 %	11,0 6,6 %	188,0 100,0 %	112,1 59,6 %	68,8 36,6 %	7,1 3,8 %
Norwegen	135,2 100,0 %	101,2 74,9 %	32,4 23,9 %	1,6 1,2 %	155,0 100,0 %	96,5 62,3 %	54,5 35,1 %	4,0 2,6 %
Finnland	102,4 100,0 %	78,1 76,3 %	22,6 22,0 %	1,7 1,7 %	167,0 100,0 %	115,0 68,9 %	48,4 29,0 %	3,6 2,1 %
Belgien	91,7 100,0 %	26,3 28,7 %	48,9 53,3 %	16,5 18,0 %	132,2 100,0 %	34,6 26,2 %	76,3 57,7 %	21,3 16,1 %
Spanien	78,8 100,0 %	56,1 71,2 %	18,6 23,6 %	4,1 5,2 %	318,2 100,0 %	173,5 54,5 %	131,8 41,4 %	12,9 4,1 %
Portugal	30,2 100,0 %	9,7 32,1 %	19,3 63,9 %	1,2 4,0 %	49,8 100,0 %	10,6 21,3 %	37,7 75,7 %	1,5 3,0 %
Österreich	69,4 100,0 %	26,7 38,5 %	39,5 56,9 %	3,2 4,6 %	106,7 100,0 %	53,0 49,7 %	49,4 46,3 %	4,3 4,0 %
Luxemburg	43,4 100,0 %	6,3 13,9 %	36,0 82,9 %	1,4 3,2 %	76,0 100,0 %	9,4 12,4 %	64,4 84,7 %	2,2 2,9 %
Griechenland	50,5 100,0 %	28,7 56,8 %	12,2 24,2 %	9,6 19,0 %	122,6 100,0 %	88,7 72,3 %	19,5 15,9 %	14,4 11,8 %
Irland	31,5 100,0 %	9,8 31,1 %	21,2 67,3 %	0,5 1,6 %	53,8 100,0 %	11,0 20,4 %	42,1 78,3 %	0,7 1,3 %
Island	9,04 100,0 %	0,6 6,6 %	8,4 93,0 %	0,04 0,4 %	10,82 100,0 %	0,57 5,3 %	10,22 94,4 %	0,03 0,3 %
EWK*	5446,6 100,0 %	3846,8 70,6 %	291,2 5,4 %	1308,6 24,0 %	9262,7 100,0 %	4718,0 50,9 %	2108,7 22,8 %	2436,9 26,3 %

**Tabelle 3:** Elektronische Informationsdienste in den Ländern des Europäischen Wirtschaftsraumes: Inlandsmarkt, Importe und Exporte 1994 und 1997 (*Quellen:* MS-Study I und II)

\* EWR als Einheit verstanden; Inland: EWR-Binnenmarkt; Import, Export: Außenhandel mit Nicht-EWR-Ländern

	Importe pro Kopf 1994	Importe pro Kopf 1997	Exporte pro Kopf 1994	Exporte pro Kopf 1997
Großbritannien	13,0	18,3	41,0	46,0
Frankreich	1,4	2,6	2,0	2,4
Deutschland	3,9	8,5	1,1	1,0
Italien	1,7	4,2	0,1	0,2
Niederlande	4,7	15,8	5,9	25,9
Schweden	7,8	8,4	0,7	0,8
Dänemark	7,6	13,2	2,1	1,4
Norwegen	7,5	12,4	0,4	0,9
Finnland	4,4	9,3	0,3	0,7
Belgien	4,9	7,5	1,7	2,1
Spanien	0,5	3,3	0,1	0,3
Portugal	1,9	3,8	0,1	0,2
Österreich	5,1	6,2	0,4	0,5
Luxemburg	90,0	161,0	3,5	5,5
Griechenland	1,2	1,9	0,9	1,4
Irland	6,1	11,7	0,1	0,2
Island	28,0	34,1	0,1	0,1
E W R	0,8	5,6	3,5	6,4

**Tabelle 4:** Pro Kopf Import- und Exportvolumen elektronischer Informationsdienste in den Ländern des Europäischen Wirtschaftsraumes 1994 und 1997 (Quelle: MS-Study II)

\* Importe aus Nicht-EWR-Ländern; Exporte in Nicht-EWR-Länder

Über die durch die Ländergröße bereinigten Indikatoren Import bzw. Export pro Einwohner gibt Tabelle 2 Auskunft. Bei den Importen pro Kopf dominiert Luxemburg; 1994 wurden Informationen im Wert von 90,0 ECU pro Einwohner und 1997 161,0 ECU pro Einwohner importiert. An zweiter Stelle steht Island, gefolgt von Großbritannien. Bei den Exporten pro Kopf beherrschen zwei Volkswirtschaften das Bild: das Vereinigte Königreich (mit einem schwachen Anstieg von 41,0 ECU im Jahr 1994 auf 46,0 ECU 1997) und die Niederlande (mit einem rasanten Anstieg von 5,9 ECU auf 25,9 ECU).

#### **4. Produktion und Konsumtion elektronischer Informationsdienste**

Schauen wir uns nunmehr Produktion und Nutzung elektronischer Informationsdienste genauer an! Über die Aspekte der Informationsproduktion gibt Tabelle 3 Auskunft.

Produktion Land	Umsätze 1994	Marktanteil 1994 (%)	Umsätze 1997	Marktanteil 1997(%)	Wachstum 1994 / 1997
Großbrit.	2878,2	55,8 %	3668,8	51,3 %	27,5 %
Frankreich	742,2	14,4 %	738,4	10,3 %	- 0,5 %
Deutschland	367,4	7,1 %	497,9	7,0 %	35,5
Italien	264,5	5,1 %	451,5	6,3 %	70,7 %
Niederlande	277,6	5,4 %	903,7	12,6 %	225,5 %
Schweden	114,5	2,2 %	117,6	1,6 %	2,7 %
Dänemark	128,0	2,5 %	119,2	1,7 %	- 6,9 %
Norwegen	102,8	2,0 %	100,5	1,4 %	- 2,2 %
Finnland	79,8	1,5 %	118,6	1,7 %	48,6 %
Belgien	42,8	0,8 %	55,9	0,8 %	30,6 %
Spanien	60,2	1,2 %	186,4	2,6 %	209,6 %
Portugal	10,9	0,2 %	12,1	0,2 %	11,0 %
Österreich	29,9	0,6 %	57,3	0,8 %	91,6 %
Luxemburg	7,4	0,1 %	11,6	0,2 %	56,8 %
Griechenland	38,3	0,7 %	103,1	1,4 %	169,2 %
Irland	10,3	0,2 %	11,7	0,2 %	13,6 %
Island	0,64	0,0 %	0,60	0,0 %	- 6,3 %
E W R	5155,4	100,0 %	7154,9	100,0 %	38,8 %

**Tabelle 5:** Produktion elektronischer Informationsdienste in den Ländern des Europäischen Wirtschaftsraumes 1994 und 1997 - Umsätze (im Inland sowie Export) in Mio. ECU (konstante 1994er ECU); Anteil an der europäischen Gesamtproduktion; Wachstum 1994/1997 - (Quellen: MS-Study I und II)

Die europäische Informationsproduktion ist stark konzentriert. Inlandsumsatz und Export allein aus Großbritannien halten über 50% der gesamten europäischen Produktion. Drei Viertel der Produktion elektronischer Informationsdienste findet in drei Ländern statt (1994: Großbritannien, Frankreich und Deutschland; 1997: Großbritannien, Niederlande und Frankreich).

Bei der Informationskonsumtion (Tabelle 4) ist die Konzentration etwas schwächer als bei der Produktion. In den Jahren 1994 und 1997 entfallen drei Viertel der europäischen Gesamtinformationsnutzung auf immerhin fünf Staaten (Großbritannien, Deutschland, Niederlande, Frankreich und Italien); die - auch hier durchaus zu beobachtende - Vormachtstellung Großbritanniens ist nicht so stark ausgeprägt.

Absolutwerte sagen über die Leistungsfähigkeit einer Volkswirtschaft weniger aus als bereinigte Werte. In Tabelle 5 errechnen wir Werte für die Produktion und Konsumtion elektronischer Informationsdienste pro Einwohner. Wenn wir den EWR-Gesamtwert als Schwellenwert betrachten, können wir die Länder jeweils in zwei Gruppen klassieren: in diejenigen über den EWR-Schnitt ("Gruppe 1 Länder") und diejenigen unter dem europäischen Durchschnitt ("Gruppe 2 Länder").

Konsumtion Land	Umsätze 1994	Marktanteil 1994 (%)	Umsätze 1997	Marktanteil 1997 (%)	Wachstum 1994/1997
Großbrit.	1292,3	31,2 %	2041,8	29,9 %	58,0 %
Frankreich	703,7	17,0 %	746,9	10,9 %	6,1 %
Deutschland	594,5	14,4 %	1117,8	16,4 %	88,0 %
Italien	353,4	8,5 %	680,8	10,0 %	92,6 %
Niederlande	259,4	6,3 %	746,7	10,9 %	187,9 %
Schweden	175,9	4,3 %	184,6	2,7 %	4,9 %
Dänemark	156,5	3,8 %	180,9	2,6 %	15,6 %
Norwegen	133,6	3,2 %	151,0	2,2 %	13,0 %
Finnland	100,7	2,4 %	163,4	2,4 %	62,3 %
Belgien	75,2	1,8 %	110,9	1,6 %	47,5 %
Spanien	74,7	1,8 %	305,3	4,5 %	308,7 %
Portugal	29,0	0,7 %	48,3	0,7 %	66,6 %
Österreich	66,2	1,6 %	102,4	1,5 %	54,7 %
Luxemburg	42,0	1,0 %	73,8	1,1 %	75,7 %
Griechenland	40,9	1,0 %	108,2	1,6 %	164,5 %
Irland	31,0	0,7 %	53,1	0,8 %	71,3 %
Island	9,0	0,2 %	10,8	0,2 %	19,9 %
E W R	4138,0	100,0 %	6826,7	100,0 %	65,0 %

**Tabelle 6:** Konsumtion elektronischer Informationsdienste in den Ländern des Europäischen Wirtschaftsraumes 1994 und 1997 Umsätze (im Inland sowie Import) in Mio. ECU (konstante 1994er ECU); Anteil an der europäischen Gesamtkonsumtion; Wachstum 1994/1997 - (Quellen: MS-Study I und II)

Bei der Informationsproduktion fallen unter die "Gruppe 1 Länder" sowohl 1994 als auch 1997 (allerdings in anderer Reihenfolge) Großbritannien, Niederlande, Luxemburg, Dänemark, Norwegen und Finnland; am unteren Ende der "Gruppe 2 Länder" rangieren Portugal, Island, Irland und Spanien.

Die "Gruppe 1 Länder" bei der Informationskonsumtion sind 1994 und 1997 (in unterschiedlicher Reihenfolge) Luxemburg, Niederlande, Island, Großbritannien, Dänemark, Norwegen, Finnland und Schweden. Frankreich war 1994 "Gruppe 1 Land", fiel aber 1997 unter den EWR-Durchschnittswert. Innerhalb der "Gruppe 2 Länder" konnte Deutschland seinen 13. Rangplatz in der europäischen "Hitparade" (1994) auf den zehnten Platz (1997) verbessern. Die Schlußlichter bei der Nutzung elektronischer Informationen sind Portugal, Spanien, Griechenland, Belgien und Italien. Ein weiterer wichtiger Wert ist der Anteil einer Wirtschaftsbranche an der gesamtwirtschaftlichen Leistung. Den Anteil der Informationswirtschaft an der wirtschaftlichen Wertschöpfung zeigt Tabelle 6.

Prod. / Kons, Land	Pro Kopf Kon- sumtion 1994	Pro Kopf Kon- sumtion 1997	Pro Kopf Pro- duktion 1994	Pro Kopf Pro- duktion 1997
Großbritannien	22,8	34,8	50,9	62,6
Frankreich	12,2	12,8	12,8	12,6
Deutschland	7,3	13,6	4,5	6,1
Italien	6,2	11,9	4,6	7,9
Niederlande	17,1	47,9	18,3	57,9
Schweden	20,5	21,0	13,3	13,4
Dänemark	30,1	34,8	24,6	22,9
Norwegen	31,1	34,3	23,9	22,8
Finnland	19,7	31,4	15,6	22,8
Belgien	7,5	10,9	4,3	5,5
Spanien	1,9	7,7	1,5	4,7
Portugal	2,9	4,9	1,1	1,2
Österreich	8,5	12,8	3,8	7,2
Luxemburg	105,0	184,5	18,5	29,0
Griechenland	4,0	10,3	3,7	9,8
Irland	8,9	14,8	2,9	3,3
Island	30,0	36,0	2,1	2,0
EWK	11,1	18,1	13,8	18,9

Tabelle 7: Pro Kopf Produktion und Konsumtion elektronischer Informationsdienste in den Länder des Europäischen Wirtschaftsraumes 1994 und 1997 in konstanten 1994er ECU - (Quelle: MS-Study II)

"pro Kopf Konsumtion" = Umsätze durch Import und Inlandsmarkt : Anzahl der Einwohner

„pro Kopf Produktion“ = Umsätze durch Export und Inlandsmarkt : Anzahl der Einwohner

Im Europäischen Wirtschaftsraum insgesamt wächst der Anteil der Informationswirtschaft von 0,057% (1994) auf 0,070% (1997). Damit liegt die Wertschöpfung der Informationswirtschaft in Regionen, die volumenmäßig nahezu zu vernachlässigen sind. Nur drei europäische Länder erreichen 1997 ein Volumen mit einem Anteil von mehr als einem Drittel Prozent des jeweiligen Bruttoinlandsprodukts: Luxemburg (0,55%), Großbritannien (0,45%) und die Niederlande (0,37%). Die Einteilung nach "Gruppe 1" und "Gruppe 2 Ländern" bringt hier eine Liste von vier Staaten, die 1997 "unterm Strich" liegen: Portugal, Österreich, Belgien und Deutschland. Italien und Spanien, 1994 noch "Gruppe 2 Länder", können sich 1997 in die "Gruppe 1" verbessern. Die Rangplätze eins und zwei der "Gruppe 1 Länder" tauschen 1994 und 1997 Großbritannien (Spitzenreiter 1994) und Luxemburg (Spitzenreiter 1997). Es folgen die Niederlande und die skandinavischen Länder.

Infwirt / BIP Land	Infwirt / BIP 1994	Rangplatz 1994	Infwirt / BIP 1997	Rangplatz 1997
Großbritannien	0,42 %	1	0,45 %	2
Frankreich	0,07 %	9	0,08 %	12
Deutschland	0,04 %	16	0,07 %	14
Italien	0,04 %	13	0,08 %	11
Niederlande	0,12 %	7	0,37 %	3
Schweden	0,11 %	8	0,10 %	9
Dänemark	0,13 %	5	0,13 %	7
Norwegen	0,15 %	3	0,12 %	8
Finnland	0,12 %	6	0,16 %	5
Belgien	0,05 %	12	0,06 %	15
Spanien	0,02 %	17	0,08 %	13
Portugal	0,04 %	15	0,06 %	17
Österreich	0,04 %	14	0,06 %	16
Luxemburg	0,35 %	2	0,55 %	1
Griechenland	0,06 %	11	0,14 %	6
Irland	0,06 %	10	0,08 %	10
Island	0,14 %	4	0,17 %	4
E; WR	0,057 %	-	0,070 %	-

**Tabelle 8:** Anteil des Umsatzes der Informationswirtschaft am Bruttoinlandsprodukt der Länder des Europäischen Wirtschaftsraumes 1994 und 1997 in Prozent (*Quelle:* MS-Study II)

Indikator: Bruttoumsatz der Informationswirtschaft (Import + Inlandsnachfrage + Export) : BIP (in konstanten 1994er ECU)

Der Top-Rangplatz von Luxemburg wird in der Literatur durch die Konzentration von Banken und internationalen Institutionen im Großherzogtum zu erklären versucht. "Luxembourg ... ranks far ahead of all other countries of the European Economic Area and therefor represents a special case. The explanation is that Luxembourg has risen to become an international banking metropolis and therefor real-time financial information ... is used extensively there" (DGXIII/E 1996, 46). "The use of information services by these banks and by the European institutions, out of all proportion to the economic environment of the rest of the country, creates this ... result" (White 1998, 64). Ein ähnliches Erklärungsmuster könnte man für Großbritannien - mit dem Finanzzentrum London - benutzen. Hinzu kommt als Argument die Dominanz der englischen Sprache auf dem globalen Informationsmarkt (vgl. DGXIII/E 1996, 47).

## 5. Informationskompetenz und wirtschaftliche Leistungsfähigkeit einer Volkswirtschaft

Informationskompetenz einer Volkswirtschaft zeigt sich auf mehreren Dimensionen. Wir erinnern an:

- Größe der nationalen Informationswirtschaft relativ zur gesamtwirtschaftlichen Wertschöpfung (BIP)
- Informationsproduktion und -konsumtion pro Kopf

- Informationsimport und -export pro Kopf.

Als Indikatoren für die wirtschaftliche Leistungsfähigkeit einer Volkswirtschaft ziehen wir das Bruttoinlandsprodukt (BIP) sowie das BIP pro Einwohner heran. Letzteres gilt in der empirischen Wirtschaftsforschung als der führende Indikator auf die wirtschaftliche Leistungsfähigkeit eines Landes. Unsere Vergleichswerte haben wir internationalen Statistiken (u.a. Eurostat und OECD) entnommen (vgl. Statistisches Bundesamt 1999).

Leitendes Erkenntnisinteresse dieses Kapitels ist: Korrelieren unsere Indikatoren der Informationskompetenz von Volkswirtschaften mit Indikatoren der wirtschaftlichen Leistungsfähigkeit? Errechnet wurde jeweils die Korrelation (nach Pearson) für zwei Variablen.

Tabelle 7 arbeitet mit der relativen Größe der nationalen Informationswirtschaft als Indikator für die Informationskompetenz der betreffenden Volkswirtschaft. Es geht hier um den direkten Einfluß der Branche der Informationswirtschaft auf die Gesamtwirtschaft. Es gibt keinen statistisch darstellbaren Zusammenhang sowohl mit dem nationalen Bruttoinlandsprodukt (BIP) als auch mit der Größe des Landes (gemessen durch die Einwohnerzahl).

Eine positive Korrelation besteht zwischen der relativen Größe der Informationswirtschaft in einem Lande und der jeweiligen Arbeitsproduktivität. Je mehr Umsätze die Unternehmen der Informationswirtschaft erwirtschaften, desto höher ist die Produktivität der Mitarbeiter. Dies wurde in der MS-Study I bereits vermutet, nunmehr konnte die Vermutung bestätigt werden. Werner Schwuchow bietet als Erklärung an: "Einmal ist anzunehmen, daß es auch im EIS-Sektor (EIS: Elektronische Informations-Dienste; die Verf.), wie in anderen Wirtschaftsbereichen, sog. 'economies of scale' gibt, d.h. daß bei größeren Produktmengen Kostendegression (auch im Personalkostenbereich) auftreten. Anders ausgedrückt: Erst von einer bestimmten Produktionsmenge an 'lohnt sich' die Produktion von EIS.

### Korrelationskoeffizienten (Pearson)

a)

Korrelation zwischen dem Anteil der Informationswirtschaft am BIP und dem nationalen BIP (alte 17 EWR-Länder)	
1994	-0,059
1997	<u>-0,061</u>

b)

Korrelation zwischen dem Anteil der Informationswirtschaft am BIP und der Einwohnerzahl der Länder (alle 17 EWR-Länder)	
1994	-0,036
1997	-0,065

c)

Korrelation zwischen dem Anteil der Informationswirtschaft am BIP und dem Anteil katholischer Einwohner in den Ländern (alle 17 EWR-Länder)	
1994	-0,232
1997	<u>-0,094</u>

Korrelation zwischen dem Anteil der Informationswirtschaft am BIP und dem Anteil katholischer Einwohner in den Ländern (16 EWR-Länder; ohne Luxemburg)	
1994	-0,497
1997	<u>-0,420</u>

d)

Korrelation zwischen dem Anteil der Informationswirtschaft am BIP und der Arbeitsproduktivität innerhalb der Informationswirtschaft	
1994 *	+0,621
1997 **	<u>+0,633</u>

**Tabelle 9:** Zusammenhang zwischen der Größe der nationalen Informationswirtschaft und dem nationalen Einkommen, der Ländergröße, der vorherrschenden Religion und der Arbeitsproduktivität innerhalb der Informationswirtschaft

\* 16 EWR-Länder (ohne Island)

\*\* 14 EWR-Länder (ohne Luxemburg, Irland und Island)

Zum zweiten ist anzumerken, daß mit einem höheren Standardisierungsgrad der Produkte (Realtime-Dienste sind ein Beispiel dafür) die 'Arbeitsproduktivität' steigt" (Schwuchow 1996, 284).

Eine interessante (negative) Korrelation hat die Größe der Informationswirtschaft mit dem Anteil katholischer Einwohner eines Landes. Die Informationswirtschaft ist umso größer, je kleiner der Anteil der Katholiken ist. Vernachlässigen wir wegen dessen Sonderstatus (das katholische) Luxemburg, so liegt mit rund -0,5 (1994) bzw. -0,4 (1997) ein nicht zu übersehender Zusammenhang vor. Man erinnert sich an den von Max Weber beschriebenen Einfluß der protestantischen Ethik auf den "Geist" des Kapitalismus. Gibt es einen hemmenden Einfluß des Katholizismus auf den "Geist" der Informationswirtschaft?

In Tabelle 8 betrachten wir die Zusammenhänge zwischen der wirtschaftlichen Leistungsfähigkeit eines Landes (gemessen durch BIP pro Kopf) und weiteren Aspekten der Informationskompetenz. Auf "normalen" Gütermärkten gilt der Export als nicht zu unterschätzender Einflußfaktor. In der Informationswirtschaft trifft dies nicht zu. Hier besteht sogar eine leicht negative Korrelation zwischen Informationsexport pro Kopf und wirtschaftlicher Leistungsfähigkeit.

Eine hohe Korrelation stellen wir zwischen Informationskonsumtion und BIP pro Kopf fest. Bei Werten zwischen +0,642 und +0,715 kann der Zusammenhang als



sehr wahrscheinlich angesehen werden. Die Rangfolge der europäischen Länder nach Informationskonsumtion pro Kopf und nach BIP pro Kopf verläuft nahezu parallel. Aber es gibt Ausnahmen: Deutschland rangiert 1994 bei der wirtschaftlichen Leistungsfähigkeit auf Platz 4, aber bei der Informationskonsumtion auf Platz 13. 1997 liegt die Rangdifferenz für Deutschland nur noch bei fünf Plätzen. Für Großbritannien sind die Verhältnisse anders herum; das Land liegt 1997 beim BIP pro Kopf auf Rang 13, bei der Informationskonsumtion jedoch auf Rang 4.

Eine ebenso hohe Korrelation besteht zwischen dem BIP pro Kopf und dem Informationsimport pro Kopf. In beiden Beobachtungsjahren errechnen wir einen Korrelationskoeffizienten von rund +0,6.

Der Zusammenhang zwischen der Informationsproduktion pro Kopf und dem BIP pro Kopf ist weniger stark ausgeprägt. Da wir keine Korrelation zwischen dem Export und dem BIP pro Kopf feststellen konnten, liegt der positive Effekt der Informationsproduktion wohl ausschließlich in der Produktion für das eigene Land begründet.

Wir können festhalten: Es gibt eindeutige Hinweise darauf, daß die Informationskompetenz einer Volkswirtschaft mit der wirtschaftlichen Leistungsfähigkeit des Landes korreliert. Wichtig für eine Volkswirtschaft sind

- die Informationskonsumtion
- der Import von Informationen und
- die Produktion von Informationen für die Konsumtion im eigenen Land.

Der Export von Informationen hat keine feststellbare Bedeutung für die Leistungsfähigkeit des exportierenden Landes.

a)

Korrelation zwischen BIP pro Kopf und Informationsproduktion pro Kopf (alle 17 EWR-Länder)	
1994	+0,296
1997	<u>+0,223</u>

b)

Korrelation zwischen BIP pro Kopf und Informationskonsumtion pro Kopf (alle 17 EWR-Länder)	
1994	+0,715
1997	+0,642

c)

Korrelation zwischen BIP pro Kopf und Informationsexport pro Kopf (alle 17 EWR-Länder)	
1994	-0,090
1997	-0,069

d)

Korrelation zwischen BIP pro Kopf und Informationsimport pro Kopf (alle 17 EWR-Länder)	
1994 *	+0,588
1997 **	+0,583

**Tabelle 10:** Zusammenhang zwischen nationaler Wirtschaftskraft und Informationsproduktion bzw. -konsumtion sowie Import und Export von Informationen

## 6. Informationskompetenz und "Neue Ökonomie"

Die Bedeutung des Imports sowie der einheimischen Produktion in der Informationswirtschaft verhält sich analog positiv für eine Volkswirtschaft wie der "Import" (bzw. die Immigration) oder die einheimische Ausbildung hochqualifizierter Fachleute. Thomas Straubhaar stellt fest, "the New Growth Theory argues that human capital produces positive knowledge externalities that spill over the economy in which they occur. Therefore, countries which have high levels of human capital, grow more quickly" (Straubhaar 2000, 20). Es geht jeweils darum, Wissen in einer Volkswirtschaft zu nutzen - egal, wo es herkommt. Das Trägermedium - elektronische Informationsdienste hier, menschliche Gehirne dort - ist dabei wenig relevant, zentral ist das Wissen.

Einzuordnen ist dieser - im Sinne "klassischer" ökonomischer Betrachtungen - überraschende Sachverhalt der hohen Bedeutung des Imports bei vernachlässigbarer Bedeutung des Exports in die Diskussionen um eine "Neue Ökonomie", eine Ökonomie, in der Informationen bzw. Wissen, Informationstechnik sowie Netzwerke im Zentrum der Betrachtung stehen. Stephen S. Cohen, J. Bradford DeLong und John Zysman charakterisieren die "Neue Ökonomie" als "E-conomy": "The term 'E-conomy' points at the fact that today's economic transformation is driven by the development and diffusion of modern electronics-based information technology" (Cohen/DeLong/ Zysman 2000, 3). In der "E-conomy" gelten andere Regelmäßigkeiten als in der "bisherigen" Wirtschaft (vgl. DeLong/Froomkin 1999; dazu kritisch: Kalmbach 2000); wir konnten eine weitere "Neuerung" hinzufügen.

Durch die überragende Rolle der Nutzung von Informationen für eine Volkswirtschaft ist eine entsprechende breite Qualifikation der Mitglieder nötig. Hier vollziehen wir einen - nicht mehr an Korrelationen, sondern am common sense orientierten - Sprung von der Empirie und der Theorie in die Politik. Will man die Informationskompetenz einer Volkswirtschaft erhöhen, muß man deren

Mitglieder adäquat qualifizieren. Mit dieser trivialen, aber dennoch sehr wichtigen These stehen wir in Übereinstimmung mit führenden Informationswissenschaftlern. "Das ist die zentrale Herausforderung an das Bildungssystem auf allen Ebenen und mit Blick auf jeden Umgang mit Wissen und Information: Informationskompetenz zu vermitteln", schreibt Rainer Kuhlen (Kuhlen 1999b, 107; vgl. auch: Kuhlen 1999a, 182 ff.). Für Norbert Henrichs führt die Anwendung der Informationswissenschaft "zu besserer Ressourcenverteilung und rationalerem Handeln", allerdings mit dem Risiko der Partizipationsgerechtigkeit in zwei Ausprägungen, "einmal der Nord-Süd-Unterschied zuungunsten der Entwicklungsländer, zum andern wird die Schere zwischen 'Wissenden' und 'Nicht-Wissenden' - egal, wo auf der Welt - immer größer". Für Henrichs ist klar: "Es ist Aufgabe des Bildungswesens, dieses Problem zu minimieren" (Henrichs 2000, 24).

## **7. Literatur**

Bredemeier, W. (1996): Die europäischen Märkte für elektronische Informationsdienste aus Anbietersicht: Umsätze, Beschäftigung, Marktanteile, Ranking in 17 europäischen Ländern. In: W. Neubauer; R. Schmidt (Hrsg.): 18. Online-Tagung der DGD. "Information ohne Grenzen - Wissensvermittlung im Zeitalter der Datennetze", Frankfurt: DGD, 1996, S. 313-339.

Bredemeier, W. (2000): Innereuropäisches Benchmarking. Password 7+8/ 2000, S. 2-10.

Cohen; S.S.; DeLong, J.B.; Zysman, J. (2000): Tools for Thought: What is New and Important About the "E-conomy". Berkeley: Univ. of California, 2000. BRIE Working Paper #138.

DeLong; J.B.; Froomkin, A.M. (1999): Speculative Microeconomics for Tomorrow's Economy. Berkeley: Univ. of California; Miami: Univ. of Miami School of Law, 1999. Version C5. (<http://econ161.berkeley.edu>).

DG XIII/E (1996): The Markets for Electronic Information Services in the European Economic Area. Supply, Demand and Information Infrastructure. European Report of the Member States' Study (MSSTUDY). Luxembourg: European Commission / DG XIII/E, October 1996.

Handtke, U. (1997): Der Markt für elektronische Informationsprodukte und -dienstleistungen. Dipl.-Arb. im Studiengang Volkswirtschaftslehre; Humboldt-Univ. zu Berlin, 1997.

Henrichs, N. (2000): Zur Bedeutung der Informationswissenschaft. Password 7+8/2000, S. 24+35.

Kalmbach, P. (2000): Eine neue Wirtschaft im neuen Jahrtausend? Wirtschaftsdienst 80 (2000), S. 210-217.

Kuhlen, R. (1999a): Die Konsequenzen von Informationsassistenten. Frankfurt: Suhrkamp, 1999.

Kuhlen, R. (1999b): Medienkompetenz durch Informationsassistenten. In: N.Szyperski (Hrsg.): Perspektiven der Medienwirtschaft. Lohmar; Köln: Eul, 1999, S. 105-107.

Look, H. (1998): The Markets for Electronic Information Services in the European Economic Area. Supply, Demand and Information Infrastructure (Review). Program - Electronic Library and Information Systems 32 (1998), S. 442-443.

Schwuchow, W. (1996): Die Studie über elektronische Informationsdienste in siebzehn europäischen Ländern (MSSTUDY). Nachrichten für Dokumentation 47, 1996, S. 277-286.

Schwuchow, W. (1998): European information services markets (reply). Journal of Information Sciences 24, 1998, S. 205-206.

Schwuchow, W.; Bredemeier, W.; Graumann, S. (1998): Manual on Assessing the Situation of the Markets for Electronic Information Services in the European Economic Area (MSSTUDY II). Hattingen; Cologne: IIE Institute for Information Economics; Munich: Infratest Burke, August 1998.

Schwuchow, W.; Graumann, S.; Bredemeier, W. (1997): A study of the European information services markets. Journal of Information Sciences 23, 1997, S. 249-259.

Statistisches Bundesamt (1999): Statistisches Jahrbuch für das Ausland 1999. Stuttgart: Metzler-Poeschel, 1999.

Stock, W. G. (1993): Der Markt für elektronische Informationsdienstleistungen. ifo Schnelldienst Nr. 14, 1993, S. 22-31.

Stock, W. G. (2000): Informationswirtschaft. Management externen Wissens. München; Wien: Oldenbourg, 2000.

Straubhaar, T. (2000): International Mobility of the Highly Skilled: Brain Gain, Brain Drain or Brain Exchange. Hamburg: Hamburgisches Welt-Wirtschafts-Archiv (HWWA), 2000. (HWWA Discussion Paper; 88).

White, M. (1998): European information services market. Journal of Information Sciences 24, 1998, S. 63-64.



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 245 – 264

# **Internetportal, Unternehmensportal, Wissenschaftsportal**

## **Typologie und Funktionalität der wichtigsten Portalkonzeptionen**

**Hermann Rösch**

Fachhochschule Köln

Claudiusstr. 1

50678 Köln

Tel.: 0221 / 8275-3376

E-Mail: Hermann.Roesch@fh-koeln.de

*Die Schwelle ist der Platz der Erwartung.  
Goethe: Wilhelm Meisters Lehrjahre*

## **Zusammenfassung**

Internetportale haben sich aus Suchmaschinen und Webkatalogen entwickelt. Mit dem Unternehmensportal ist die Konzeption auf den betrieblichen Kontext übertragen worden. Eine weitere Adaption für den Wissenschaftssektor zeichnet sich ab. Zu den Kernfunktionen der Portale gehören neben dem zentralen Einstieg, der Simplizität und der Leistungsfähigkeit der Suchwerkzeuge vor allem ein hohes Integrations- und Personalisierungspotential sowie Tools zur Kommunikation, Kollaboration und Validierung von Informationen. Mit diesen Leistungsmerkmalen sind Portale prädestiniert, zu den Marktplätzen des E-Commerce sowohl für Endverbraucher- als auch für Wertschöpfungsketten zu werden. Die weitere typologische Diversifizierung ist absehbar.

## **Abstract**

Internet portals developed from search engines and web directories. With the enterprise information portal the conception was transmitted to the business context. A further adaptation to the science sector is in the initial stages. As basic functions portals provide a single point of access, an easy handling and a powerful search engine combined with a directory. Its core features consist in tools that provide personalization and integration as well as communication, collaboration and validation. With these features portals are predestined, to become the market places of E-Commerce both for end users and B2B-transactions. A further



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

diversification of portal types is foreseeable.

## 1. Ausgangslage

Amerikanische Analysten setzen die Bedeutung der Portale für die Informationsgesellschaft gleich mit jener der Eisenbahnen für die Industriegesellschaft des 19. Jahrhunderts [Delphi Group 1999b]. Die Portalmetapher und die damit bezeichneten Konzeptionen erleben zur Zeit einen wahren Siegeszug. Portalvarianten werden vor allem eingesetzt im Internet und im Unternehmensbereich; die Adaption von Portalstrukturen für den Wissenschaftssektor ist absehbar. Klare Definitionen, die Charakterisierung typologischer Varianten und eine eindeutige terminologische Unterscheidung fehlen bislang.

## 2. Zum Begriff "Portal"

Bezeichnet wird mit dem lateinischen Lehnwort "Portal" der meist repräsentative Eingang eines bedeutenden Baues von oft beträchtlichen Ausmaßen. [Brockhaus 1998, S. 364]. Neben der primären Funktion des Einlasses kommt dem Portal als Grenze und Vermittler zwischen Außenwelt und Innenraum in allen Epochen magische, allegorische oder symbolische Bedeutung zu. [Lexikon der Kunst 1989, S. 241].

Die Windowsmetapher durch die Portalmetapher zu ergänzen, erweist sich retrospektiv als naheliegende Konsequenz der jüngeren DV-technischen und telekommunikativen Innovationen. Cyberspace und virtuelle Welten nur durch Fenster, von außen also, betrachten zu können, musste sich als unbefriedigend erweisen. Solche Welten und Räume machen nur dann Sinn, wenn sie, zumindest virtuell, auch betreten werden können. Eine Möglichkeit also musste geschaffen werden, um sich diese neuen Räume bedeutenden Ausmaßes erschließen zu können, um Zutritt zu dieser ansonsten hermetischen Cyberwelt zu erlangen. Die Sinnfälligkeit des Portals als des "Ortes", an dem der Einlass in die begehrte andere Welt gewährt wird, liegt damit auf der Hand. Im Kampf um Kunden wird in der Werbung von Portalentwicklern und -betreibern neben dieser primären Bedeutung auch das allegorische, symbolische und magische Potential des Portalbegriffes ausgeschöpft.

## 3. Entstehungskontext der Portalkonzeption

Es ist ein Grundparadox der Moderne, dass jede neue Erkenntnis neue Wissenslücken offenbart, jede neue Errungenschaft nicht nur Erleichterung, sondern auch neue Probleme erzeugt. Das Auto beispielsweise verhilft zu Mobilität in zuvor kaum geahntem Umfang, führt aber unter bestimmten Bedingungen zu neuer Immobilität. Der Stau im Berufs- oder Ferienverkehr lässt sich nur umgehen oder mildern durch verbesserte Infrastrukturen und Regeln. Im Informationssektor wird in diesem Zusammenhang vom Informationsdilemma gesprochen [vgl. u.a. Kühlen 1999, S. 22-26]: neue Techniken und Methoden der Sammlung, der Speicherung, des Zugriffs auf, der Distribution von und des

Generierens von Informationen erzeugen die oft beklagte Gefahr des Information Overload, der Informationsflut, des Informationschaos. Navigationshilfen, Selektionsfilter und Bewertungsraster können den Informationsstau vermeiden und für Orientierung sorgen.

In diesem Kontext sind auch Neudefinition und —konzeption des Portals entstanden. Das horrende Wachstum des Internet nach der Einführung des WWW ließ bald die Notwendigkeit von Web-spezifischen Such- und Erschließungswerkzeugen offenbar werden. Den Anfang machte Yahoo! als Prototyp des intellektuell erzeugten Directories oder Web-Katalogs. Wenig später traten mit Lycos, InfoSeek, Excite und AltaVista die ersten Suchmaschinen als Vertreter des zweiten Web-spezifischen Suchwerkzeuges hervor. In manchen Darstellungen werden Web-Kataloge und Suchmaschinen als Portale der ersten Generation bezeichnet, die freilich mit dieser Bezeichnung erst einige Jahre später in Verbindung gebracht werden [Delphi Group 1999a]. Im Dezember 1996 offerierte Yahoo! seinen Kunden Personalisierungsmöglichkeiten, die sich zum späteren "MyYahoo!" entwickelten. In der Personalisierungsoption liegt ein markantes Merkmal, das Internetportale von bloßen Suchmaschinen und anderen Diensten unterscheidet. Die nach Suchmaschinen bzw. Web-Katalogen zweite Gruppe von Portalanbietern bilden die Internet Service Provider. Im Sommer 1997 boten Planet Direct, einet und andere den Providern Software an, die es erlaubte, nutzerspezifische Homepages zu erzeugen. Die Softwareanbieter nannten diese individuellen Homepages Portale und brachten Begriff und Konzept erstmals zueinander [Portal Time 1999]. AOL und Compuserve zählen zu den klassischen Service Providern, die auf den Portalmarkt drängten. Im darauffolgenden Herbst griffen InfoSeek, Excite und Lycos, das Portalkonzept auf, indem sie ihre Suchmaschinen und Kataloge um weitere Dienste anreicherten und ebenfalls erste Personalisierungsangebote integrierten. Als dritte Gruppe folgten Browserhersteller wie Netscape („MyNetscape“) und Microsoft kurz darauf diesem Vorbild. Eine vierte Gruppe von Portalanbietern bilden Medienkonzerne. Der Disneykonzern etwa hat den Portalmarkt als zukunftssträchtiges Aktionsfeld entdeckt und den Einstieg durch eine Kooperation mit der Suchmaschine InfoSeek, später Go gewählt. In Deutschland ist das Nachrichtenmagazin Focus auf dem besten Weg, sein Internetangebot zu einem Portal auszubauen.<sup>1</sup> Große Rundfunk- und Fernsehanstalten werden sicher in Kürze folgen. Abzuwarten bleibt, in welchem Umfang weitere Betreiber von Internetportalen aus anderen Branchen auftreten werden. Prädestiniert sind dafür in erster Linie aktive Umschlagszentralen des E-Commerce, also etwa Amazon.com und andere virtuelle Kaufhäuser. Aber auch Handelskonzerne, Vertriebsketten usw. könnten bald auf die Idee kommen, dass die eigene Mali attraktiver wird, wenn die Besucher mit den Features eines Internetportals angelockt werden.

War von Portal anfangs ausschließlich im Kontext des Internet die Rede, so tauchte der Begriff im Sommer 1998 erstmals in einem neuen Umfeld auf: Plumtree und Scribe, beides Anbieter von diversen Software-Tools zum

---

<sup>1</sup> Angeboten werden Suchtools, Chat-Foren, subscribierbare Newsletter, Personalisierung. Wichtige Merkmale eines Internetportals sind damit zumindest in Grundfunktionen vorhanden; der Terminus "Portal" wird von den Betreibern jedoch noch nicht in Anspruch genommen. Vgl. <http://focus.de>

betrieblichen Informationsmanagement, verwendeten in ihren Pressemitteilungen den Begriff Corporate Portal. Die Initialzündung zur Entwicklung eines neuen Portaltypus für den Unternehmensbereich erfolgte im November 1998 durch eine Studie der Investmentbank Merrill Lynch [Shilakes/Tylman 1998]. Bereits im Dezember 1998 veröffentlichte das Knowledge Management Magazine eine Fallstudie zu den jetzt als Corporate Portals bezeichneten Produkten [Roberts-Witt 1998]. Im Februar 1999 publizierten führende IT-Magazine erste Artikel über Corporate Portals. Marktforschungsfirmen wie International Data Corporation (IDC), Delphi Group, Meta Group und andere präsentierten umfassende Analysen und White Papers zur Situation und zu Entwicklungsperspektiven auf dem Markt für Corporate, Business oder Enterprise Information Portals. Seit Sommer 1999 veranstalten diese und andere Einrichtungen in regelmäßigen Abständen internationale Konferenzen zum Thema Portale im betrieblichen Umfeld.<sup>2</sup> Mittlerweile kristallisiert sich neben einer Erweiterung des Funktionsumfangs

auch eine Spezialisierung weiterer Typen von Unternehmensportalen auf besondere Anwendungsfelder heraus. Initiativen zur Übertragung des Portalkonzeptes auf weitere gesellschaftliche Subsysteme wie das der Wissenschaften sind aktuell konstatierbar, befinden sich im Herbst 2000 allerdings noch in embryonalem Zustand.

#### 4. Grundfunktionen und —intentionen des Portals

Insbesondere unter die Internetportale hat sich inzwischen eine nicht überschaubare Zahl kaum seriöser Konkurrenten gemischt, die nicht mehr zu bieten haben, als eine mit wenig Aufwand und oft noch weniger Sachverstand erzeugte thematische Liste "relevanter" Links. Gleichwohl reklamieren diese Seiten das Attribut "Portal" für sich, weil dies im Kampf um Werbeeinnahmen Vorteile verspricht.<sup>3</sup> Um so wichtiger ist es, eindeutige Kriterien zu gewinnen, die es erlauben, solche "Blendportale" von wirklichen Portalen unterscheiden zu können.<sup>4</sup> Darüber hinaus bieten derartige Kriterien die Grundlage zu typologischer Unterscheidung und vergleichender Evaluierung. Die nachfolgende Zusammenstellung der Grundfunktionen von Portalen fußt auf einer Analyse vorhandener Portale, einer Auswertung der Konzeptbeschreibungen und –bewertungen seitens der Anbieter und Analysten sowie auf anwendungsbezogenen oder theoretischen Beiträgen in einschlägigen Periodika.<sup>5</sup>

Bei Portalen handelt es sich zunächst einmal um die Einstiegsseiten, von denen aus z.B. Internetsurfer ihre Suche starten oder Mitarbeiter von Firmen Informationen empfangen, suchen oder verbreiten. Portale zielen darauf, den Nutzern einen möglichst intuitiven, selbsterklärenden Zugang zu den Bezugsquellen über das Internet bzw. ein Intra- oder Extranet zu bieten.

<sup>2</sup> So fand z.B. vom 17.-19.5.2000 in London die von der Delphi Group organisierte „Corporate-Portal Conference Europe“ statt [Zimmermann 2000, S. 36]

<sup>3</sup> Robin Peek hat diesen Trend bereits Mitte 1998 konstatiert: "If you are not a portal, you are not cool." [Peek 1998].

<sup>4</sup> Als Beispiel für solche Schein- oder Blendportale seien genannt das Arzt-, das Wein-, das Bier-, das Reise- oder das Zeitungsportal, die anzusteuern sind unter <http://www.epl.de/webkatalog/begin.asp>; nicht auszuschließen ist freilich, dass derartige simple Linklisten um entsprechende Funktionalitäten ergänzt und so zu wirklichen Portalen ausgebaut werden.

<sup>5</sup> Stellvertretend genannt sei [Roberts-Witt 1999].



Grundlage und Messlatte bilden die gebräuchlichen Internetbrowser von Microsoft oder Netscape. Des weiteren bieten Portale in der Kombination von Suchmaschinen und Katalogen einen Zugriff auf die Bezugsquellen, der Recall und Precision in möglichst hohem Umfang zu erzielen imstande ist. Für die Recherchen sollen große und möglichst attraktive Informationsmengen bereitstehen. Konstitutiv für Portale aber ist die Personalisierungsoption. Die Genese des Portalkonzeptes stellt dies eindeutig unter Beweis. Die Kunden oder Benutzer erhalten damit die Möglichkeit, eine eigene Einstiegsseite nach individuellen Informationsbedürfnissen zu konfigurieren. Schließlich bieten Portale Features, die als „Communitybuilding-services“ kommunikative und kollaborative Aktivitäten stimulieren. Damit sind die Grundfunktionen eines Portals beschrieben, die je nach Umgebung in unterschiedlicher Gewichtung und mit typenspezifischen Ergänzungen auftreten.

## **5. Portalkonzeptionen für Internet, Unternehmen und Wissenschaft**

Seit die Betreiber weiterentwickelter Suchwerkzeuge für Internetrecherchen die Portalmetapher aufgegriffen haben, vor allem aber seit eben diese Metapher zur Bezeichnung neuer Methoden und Techniken des Informations- und Wissensmanagements im betrieblichen Umfeld Verwendung gefunden hat, ist eine beständig wachsende Flut von Begriffsvarianten zu beobachten. Da die neue Verwendung des Portalbegriffes nirgends klar festgelegt war, eine eindeutige und allgemein akzeptierbare Definition nicht von Beginn an vorlag (und wohl auch nicht vorliegen konnte), kam und kommt es zu divergierenden Interpretationen des Begriffes. Die Folge ist eine babylonische Begriffsverwirrung, die freilich von einigen Beteiligten zusätzlich und nicht ohne Absicht verstärkt wird. Insbesondere Entwickler und Anbieter von Portalsoftware für betriebliche Zwecke sind darauf erpicht, den Portalbegriff zu besetzen und ihn mit dem Anspruch auf Verbindlichkeit so zu definieren, dass die mit dem Begriff konnotierte Funktionalität identisch ist mit dem Leistungsumfang des eigenen Produktes: den Anbietern bietet sich damit die Chance, schon bei der Begriffsbildung Produktwerbung zu betreiben. Nicht wesentlich anders sind die Motive der zahlreichen Analysten, jeweils eigene Begriffsvarianten und Kompositaverbindungen zu prägen: wem es nämlich gelingt, den eigenen terminologischen Festlegungen zu allgemeiner Akzeptanz zu verhelfen, der verstärkt so die Bedeutung seiner Aussagen und steigert damit den eigenen Marktwert. Dieser Trend führt mittlerweile dazu, dass - wie ja auch in anderen Branchen üblich - Kunstworte gebildet werden, die dann entweder mit dem jeweiligen Produkthersteller oder aber dem Marktforschungsinstitut identifiziert werden und die Konkurrenz auf die Plätze verweist.<sup>6</sup>

Um Ordnung in dieses begriffliche Chaos zu bringen, bietet es sich an, das Begriffsfeld nach Anwendungsfeldern zu strukturieren. In Betracht kommen zum

---

<sup>6</sup> Jüngstes Beispiel bot die Delphi Group, die den Begriff "Vortal" prägte, um damit den vielversprechenden Typus der vertikalen Portale zu charakterisieren: "A vortal is a vertical portal for a specific community." [Koulopoulos 2000]; vgl. auch [Peek 1999].

einen das Internet bzw. der dort frei zugängliche, zumeist distributiven und konsumtiven Zwecken dienende Sektor und zum anderen das inner- und zwischenbetriebliche Umfeld, der produktive Sektor. Ein dritter Bereich ist im Wissenschaftssektor zu identifizieren. Zunächst sollen die diversen Portalbegriffe den drei aufgezählten Anwendungsfeldern zugeordnet werden. Mögliche Überschneidungen lassen sich dabei nicht vermeiden, zumal manche Portalkonzeptionen tatsächlich auf mehr als einen der genannten Anwendungsbereiche zielen.

Demnach verteilen sich die wichtigsten über einschlägige Suchmaschinen und relevante Publikationen feststellbaren Begriffsvarianten wie folgt:<sup>7</sup>

- Internet: B2C Portal, Consumer Portal, E-Business Portal, E-Commerce Portal, E-Portal, General Interest Portal, Horizontal Portal, Internetportal, Marktportal, Niche Portal, Public Portal, Web Portal
- Unternehmen: Business Portal, B2B Portal, B2E Portal, Collaborative Portal, Commerce Portal, Corporate Portal, Data Portal, Decision Processing Portal, Enterprise Application Portal, Enterprise Information Portal, Enterprise Knowledge Portal, Enterprise Portal, Enterprise-Web Portal, Expertise Portal, Information Portal, Intelligent Enterprise Portal, Intranet Portal, Knowledge Portal, Unternehmensportal, Vertical Portal, Vortal
- Wissenschaft: Bibliotheksportal, Information Portal, Informationsportal, Knowledge Portal, Library Portal, University-wide Information Portal, Wissenschaftsportal, Wissensportal

Die Aufzählung der bereits verwendeten "P-Words" ließe sich vervollständigen. Unter den deutschsprachigen Äquivalenten bieten sich die Begriffe "Internetportal", "Unternehmensportal" und "Wissenschaftsportal" zur funktionalen Beschreibung und Unterscheidung der drei Grundtypen an.

### 5.1. Internetportal

Entstanden sind Internetportale als "places for nonexperts to find things on the Internet (...) But they're quickly evolving into full-service hubs of electronic

commerce, mail, online communities and customized news." [Essex 1999] In dieser knappen Formulierung sind wesentliche funktionale Charakteristika genannt, die im einzelnen analysiert und vervollständigt werden müssen:

- Einstieg: Das Portal bietet von einem zentralen Punkt aus den Zugriff auf möglichst alle für den Kunden attraktiven Anwendungen.

---

<sup>7</sup> Aufgezählt werden hier ausschließlich englisch- und deutschsprachige Varianten. Nicht berücksichtigt wurden übrigens weitere Synonyme oder Quasisynonyme, die unter Umgehung des Begriffes Portal auf ähnliche Phänomene zielen wie z.B. Channel, Doorway Page, Entry Page, Gateway, Start Page usw.

- Simplizität: Internetportale zielen auf ein Massenpublikum und müssen daher möglichst selbsterklärend und intuitiv anwendbar sein. Sie basieren i.d.R. auf den Standards verbreiteter Internetbrowser.
- Leistungsfähige Suchwerkzeuge: Für die Recherche stehen sowohl eine Suchmaschine als auch eine thematische Liste zur Verfügung. Besondere Features bilden die dateitypspezifische Suche nach Foto-, Audio- oder Videodateien aus dem Web.
- Aggregation von Informationen: Gegenüber bloßen Suchmaschinen und Webkatalogen präsentieren Portale zusätzliche Informationsangebote. Attraktive Inhalte binden sie mittels strategischer Allianzen ein. Entsprechende Angebote können sich beziehen z.B. auf Aktienkurse, sparten- und interessenspezifische Nachrichten, Preisinformationen, Wetterdaten, Fahrpläne, Hoteldatenbanken, Fort- und Weiterbildungsangebote, Fernsehprogramme, aktuelle Informationen zu Telefon- oder Energietarifen usw. Zum "Content" rechnen die Internetportale auch Einkaufsinformationen, die zusätzlich durch Programme zur vergleichenden Produktsuche erschlossen werden.
- Personalisierung: Voraussetzung bildet die Registrierung durch ID und individuelles Passwort. Meist gibt der Kunde zusätzliche Daten ein wie Name, Anschrift, Alter, Geschlecht. Sein individuelles Interessenprofil kann er dann über eine Check-Box definieren und im Portal hinterlegen. Außerdem kann er festlegen, welche Features des Portals auf seiner Einstiegsseite auftauchen und welche Informationen ihm unmittelbar, in regelmäßigen Abständen oder bei Erreichen bestimmter kritischer Werte mitgeteilt werden. Das Check-Box-Verfahren wird oft ergänzt, z.T. auch ersetzt durch „Collaborative-Filtering-Verfahren“ und regelbasierte Segmentation [vgl. dazu Janetzko 1999].
- Integration von Zusatzfunktionalitäten: Die Attraktivität des Internetportals für den einzelnen Benutzer wird durch Tools gesteigert wie Adressbücher, Bookmarkverwaltungsprogramm e, Kalender, Übersetzungsprogramme usw.
- Kommunikation: Über das Portal können die Kunden einen kostenlosen E-Mail Account erhalten und eine eigene Homepage platzieren; ferner bemühen sich Internetportale darum, die Bindung ihrer Kunden dadurch zu verstärken, dass sie Zugang zu z.T. sogar vom Portal organisierten themenspezifischen Chatrooms oder elektronischen Flohmärkten offerieren.

Für die Benutzer ist die Inanspruchnahme der Portaldienste kostenlos. Die Internetportale finanzieren sich wie auch Suchmaschinen durch Werbeeinnahmen. Portale bieten jedoch den Werbepartnern insbesondere durch die

Personalisierung viel interessantere und lukrativere Werbemöglichkeiten. Nutzerspezifische Profile, von den Klienten selbst definiert, ergänzt durch das im Internetportal durchgeführte und aufgezeichnete reale Suchverhalten, schaffen für die Werbepartner den Idealzustand des One-to-One-Marketing. E-Commerce-Kaufhäuser können ihre kundenspezifischen Angebote automatisch generieren und lancieren. Der Klient wird also mit hoher Wahrscheinlichkeit Werbung von Anbietern auf seiner Einstiegsseite finden, die mit den Gegenständen korreliert, nach denen er im Web gesucht hat. Erweitern kann man das Szenario, wenn die Möglichkeit der vergleichenden Produktsuche in Kombination mit Preisinformationen ("price alerts") in die Betrachtung einbezogen wird. Das Portal fungiert hier als Mall. Die Wahrscheinlichkeit, dass Produktsuche und Preisvergleiche entweder unter besonderer oder gar ausschließlicher Berücksichtigung solcher Angebote erfolgen, die von Werbepartnern des Portals stammen, dürfte hoch sein. Hier deutet sich im Sinne des Verbraucherschutzes Regelungsbedarf seitens der Gesetzgeber an. In jedem Falle aber wird die Unwissenheit oder auch Bequemlichkeit einer nennenswerten Zahl von Portalnutzern die Stellung der Internetportale bei der Distribution von Informationen, welche die Kaufentscheidungen bestimmen, zusätzlich stärken.

Internetportale zielen mit allen ihren Aktivitäten darauf, Werbepartner anzuziehen, die bislang in Fernsehen, Radio, der Presse oder über andere Kanäle geworben haben und die sich die strategischen Vorteile des E-Commerce zu Nutze machen wollen. Sie sind daher bestrebt, die Zahl der Portalbesucher zu maximieren, um über hohe Nutzerzahlen möglichst hohe Preise für die Werbung und Dienste bei den Portalpartnern erzielen zu können.<sup>8</sup> Das strategische Konzept der Internetportale setzt aber darüber hinaus voraus, dass Benutzer oder Klienten ihre Besuche möglichst oft wiederholen, so lange wie möglich verweilen und im Idealfall eine portalspezifische "Online Audience", die Zielgruppe des One-to-One-Marketing bilden.<sup>9</sup> Untersuchungen der Suchmaschine Excite haben ergeben, dass Kunden durch Personalisierungen fünfmal häufiger zurückkehren als ohne [Frenko 1998]. Vor allem diesem Ziel also dienen aus Sicht der Portalbetreiber die Personalisierungsoptionen.

Internetportale bieten damit die Voraussetzungen dafür, zu den Scharnierstellen des E-Commerce, den elektronischen Marktplätzen zu werden. Sie stellen ein neues Instrument, einen bisher unbekannten Kanal für Handel und Vertrieb dar. Allein mit dieser Erwartung ist zu erklären, dass der Aktienwert großer Portalbetreiber wie etwa Yahoo! den gegenwärtig realen Gegenwert bei weitem übersteigt.

## 5.2. Unternehmensportal

In dem für Merrill Lynch verfassten White Paper werden die wesentlichen Merkmale eines Enterprise Information Portals bzw. Unternehmensportals

---

<sup>8</sup> Zu den Berechnungsvarianten für nutzer- und nutzenabhängige Bezahlung von Werbeauftritten in Internetportalen [Schumacher/Schwickert 1999, S. 19-21].

<sup>9</sup> [Peek 1998, S. 37] brachte dies auf die Formel: "Give the people what they want, and they will hang around Tonger — which today is the name of the Web game."

folgendermaßen beschrieben:

*"Enterprise Information Portals are applications that enable companies to unlock internally and externally stored information, and provide users a single gateway to personalized information needs to make business decisions. (...) (They are) an amalgamation of software applications that consolidate, manage, analyze and distribute information across and outside of an enterprise (including Business Intelligence, Content Management, Data Warehouse & Mart and Data Management applications)." [Shilakes/Tylman 1998, S. 1]*

Mit der Betonung des einheitlichen Einstiegspunktes, der Aggregation großer, relevanter Informationsmengen, der Integration heterogener Anwendungen, der Personalisierung und in abgeschwächter Form des Aspektes der Kommunikation werden Funktionen genannt, die auch bei Internetportalen identifizierbar sind. Nicht übersehen werden darf jedoch, dass Unternehmens- und Internetportal letztlich unterschiedliche Zweckbestimmungen haben. Firmen entscheiden sich natürlich nicht für den Einsatz eines Portals, damit dieses möglichst oft und lange benutzt wird, sondern damit es seine instrumentelle Funktion optimal erfüllt. Und die beschreibt Mohsen Rezayat ebenso schlicht wie ambitioniert mit "providing the right information to the right person at the right time and in the right format anywhere within the extended enterprise" [Rezayat 2000, S. 85]. Diese Rahmenbedingungen führen im Vergleich zu Internetportalen sowohl zu weiteren Funktionalitäten sowie zu einer anderen Gewichtung der gemeinsamen Merkmale, die genauer betrachtet werden müssen, um den Grundtypus des Unternehmensportals beschreiben zu können:

- Einstieg: Im Vordergrund steht die Absicht, einen zentralen Einstieg zu bieten für alle im Unternehmenskontext relevanten Informationen.
- Simplizität: Der Vorteil möglichst selbsterklärender und intuitiv anwendbarer Software liegt in der Zeitersparnis bzw. dem damit verbundenen Rationalisierungseffekt.
- Leistungsfähige Suchwerkzeuge: Die von beständigem Wandel und Wachstum geprägte Informationsmenge im Unternehmen wird adäquat nutzbar durch die Kombination leistungsfähiger Suchmaschinen und Directories.
- Aggregation handlungsrelevanter Informationen: Unter diesem Aspekt hat das Portal im betrieblichen Kontext erheblich mehr Aufgaben zu erfüllen als das Internetportal. In einem Pool zusammenzuführen sind hier sowohl strukturierte Daten z.B. aus Datenbanken als auch unstrukturierte Daten aus Texten, E-Mails, technischen Zeichnungen, Projektplänen, Produkthandbüchern, Audio- oder Videodateien, sowohl unternehmensintern erzeugte wie relevante externe Informationen. Im Unternehmensportal werden Daten und Informationen weder präkoordiniert, weil sie thematisch korrelieren, noch physisch fragmentiert, weil sie ein gemeinsames Format aufweisen, sondern in einem Master Directory integriert. Unabhängig von

Format, Provenienz und Inhalt sind Daten und Informationen zum Zeitpunkt der Recherche oder der Interaktion nach anwendungs- und aufgabenspezifischem Bedarf kombinierbar.

- Strukturierung und Aufbereitung von Informationen: Entscheidend für das Unternehmen ist, ob die Mitarbeiter die handlungsrelevanten Informationen ohne Zeitverzug identifizieren und weiterverarbeiten können. Neben dem Master Directory müssen also im Portal Applikationen vorgehalten werden, welche die Strukturierung und Bewertung von Daten und Informationen ermöglichen.<sup>10</sup> Mit Business Intelligence, Content Management, Data Warehouse & Mart, Data Mining oder Data Management werden entsprechende Tools und Funktionen bezeichnet, die am Markt gegenwärtig als isolierte Anwendungen angeboten werden.
- Integration: Aufgabe des Portals ist es, bislang lediglich unverbunden einsetzbare Programme wie etwa die Tools zur Aufbereitung von Informationen in eine einheitliche Informationsumgebung zu integrieren. Das Handling wird vereinfacht, weil der Benutzer statt der spezifischen Syntaxsysteme auf die standardisierte und intuitive Browseroberfläche des Portals zurückgreifen kann; diese Harmonisierung führt zu weiterem Zeitgewinn durch Synchronisierung und ermöglicht schließlich synergetische Effekte durch das Zusammenspiel der zuvor isolierten Anwendungen. Mit dieser Funktionalität entspricht das Portal dem häufig postulierten Paradigmenwechsel von "application-centric" zu "user-centric software", von "Content" zu "Context" bzw. von fragmentiertem zu integriertem Informationsmanagement [vgl. Appsolute 1999, S.5]. Damit erst können jene ineffizienten "islands of automation" überwunden werden, die für redundante und inkonsistente Daten verantwortlich sind.

Die bisher aufgeführten Merkmale des Unternehmensportals rücken diesen Typus in die Nähe des klassischen Informationsmanagements, dessen besondere Eigenschaften mit Objektorientierung und Integration zu bezeichnen sind. Damit die Qualitäten eines Portals tatsächlich erreicht werden können, muss die bloße Objektorientierung um Aspekte der Subjektorientierung ergänzt werden. Mit den Kriterien Personalisierung, Kommunikation und Kollaboration sowie Validierung von Informationen übernimmt das Portal Aufgaben des Wissensmanagements. Es verbindet damit nicht nur "Needs" und "Skills" sondern auch "Peers".

- Personalisierung: Personalisierung erfolgt im betrieblichen Kontext aus Sicht der Mitarbeiter und des Managements. Über das Portal wird sichergestellt, dass jeder Mitarbeiter nur Zugriff auf solche Informationen hat, die seiner betriebsinternen Rolle und Funktion angemessen sind. Sensible Daten werden so vor unbefugtem Zugriff und Missbrauch geschützt. Der einzelne Mitarbeiter kann seine Portalseite so einstellen, dass er nur mit aufgabenspezifischen Informationen versorgt wird. Die

---

<sup>10</sup> Nicht unwichtig ist in diesem Zusammenhang natürlich die durch die Migration nach XML in Aussicht gestellte stärkere Strukturierung von Internetquellen mittels Metadaten.

Personalisierung ermöglicht eine intelligente Kombination von Push- und Pulltechnologien. Voreingestellt auf der Einstiegsseite werden Informationsangebote, die vom Mitarbeiter häufig bei Bedarf selbst gesichtet und ausgewertet werden. Intelligente Agenten werden darüber hinaus eingesetzt, um, zugeschnitten auf den individuellen Informationsbedarf, weitere Quellen regelmäßig auszuwerten und gegebenenfalls auf relevante Zielinformationen aufmerksam zu machen.

- Kommunikation und Kollaboration: Der Wechsel von fordistischen Arbeitsformen zu kollaborativen Teamstrukturen wertet die innerbetriebliche Kommunikation enorm auf. In der Praxis taucht häufig das Problem auf, dass Informationen lediglich innerhalb geschlossener Gemeinschaften wie Abteilungen, Projektteams usw. zirkulieren. Da unternehmensweite Kommunikationskanäle fehlen, bleiben diese Zirkel oft hermetisch; Redundanzen und sonstige Reibungsverluste erschweren die Verwirklichung der Unternehmensziele [vgl. auch Schmidt 1999]. Das Unternehmensportal bildet eine adäquate unternehmensweite Infrastruktur für kollaborative Arbeitsprozesse und stimuliert die einzelnen Mitarbeiter zur Kommunikation. Portale beschränken sich also nicht auf die Informationsversorgung nach individuellem Bedarf, sondern bieten gleichzeitig die Möglichkeit, Informationsangebote zu platzieren. Es kann sich dabei um fertige Produkte handeln wie auch um Hypothesen, die zur Diskussion gestellt werden. Das Portal offeriert festen Teams oder dynamischen Usergruppen virtuelle Projekträume, Kommunikationskanäle, die es ermöglichen, Informationen bestimmten Adressaten mitzuteilen, in der definierten Öffentlichkeit zu diskutieren und zu bewerten. Die zugehörigen Infrastrukturen und Instrumente eröffnen Perspektiven, die auch in Konzepten des Wissensmanagements von zentraler Bedeutung sind. Gemeint ist die Absicht, individuell akkumuliertes, jedoch nicht dokumentiertes Wissen ("tacit knowledge") öffentlich zu machen, zu erschließen und damit in den unternehmensweiten Informationspool zu integrieren.
- Validierung von Informationen: Information Sharing als Resultat der kollaborativen Tools von Unternehmensportalen erlaubt in potenziertem Maße Bewertung und Validierung von Informationen. Für den einzelnen Mitarbeiter wird damit die Möglichkeit entscheidend erleichtert, aus Informationen Wissen zu generieren, zu lernen bzw. dieses in neuerliche Informationsprozesse einfließen zu lassen. Unter diesem Gesichtspunkt bezeichnen manche Autoren derartige Portale als Enterprise Knowledge Portale [vgl. etwa Firestone 2000].

Das Unternehmensportal ist aus mehreren Wurzeln entstanden. Vorbild war zweifellos das Internetportal. Eingeflossen sind zudem prägende Elemente des Knowledge Management wie Data Mining, Reporting, Dokument Management, Groupware usw. sowie Elemente aus Executive Information Systemen. Der besondere Mehrwert des Unternehmensportals besteht darin, die Fragmentierung des Unternehmens zu überwinden, einen konsistenten Blick auf das Unternehmen

zu ermöglichen, die Versorgung mit benötigten Informationen aus internen wie externen Quellen zu gewährleisten, leistungsstarke Such-, Verarbeitungs- und Bewertungsmöglichkeiten für Informationen bereit zu stellen, alle Angebote nach individuellen Bedarf zuschneiden zu können sowie infrastrukturelle Voraussetzungen für optimale Kommunikation und Kollaboration zu schaffen.

Neben diesem Grundtypus sollen als Funktionsvarianten des Unternehmensportals Subportale und B2B-Portale erwähnt werden. In beiden Fällen handelt sich um vertikale Portale, die entweder spezifische unternehmensinterne Wertschöpfungsketten umfassen oder aber unternehmensübergreifende Wertschöpfungsketten bzw. Produktmärkte erschließen. Bei den vertikalen Subportalen stellt sich die Gefahr erneuter Fragmentierung ein. Wenn sich in einem Unternehmen mehrere Subportale etablieren, bedarf es des Metaportales, um Integration und Supervision zu erzeugen. Business-to-Business Portale, kurz B2B Portale genannt, verbinden Zulieferer, Dienstleister, Hersteller und Distributionspartner miteinander. Dieser Typus wird in manchen Publikationen mit dem Kunstwort "Vortal" bezeichnet und folgendermaßen definiert:

"Vortals (are) Internet based business-to-business market communities that automate the transactions required to build a value chain. Vortals eliminate process latency and transaction time by providing a single point of contact, coordination, and execution of B2B transactions. Vortals (1) expand markets by opening them to all possible suppliers and purchasers; (2) accelerate the velocity of markets by speeding innovation; (3) increase market efficiency by eliminating lock-in and allowing fast interchange of both buyer and seller; (4) create demand driven markets." [Vortals.Com]

Zum Typus eines solchen branchenspezifischen vertikalen Portals gehört z.B. das "Global Petroleum Portal Resource Center" [Oilfield Directory]. In Deutschland sind vergleichbare Portale im Aufbau. Krankenhäuser z.B. bündeln ihre Nachfrage bei rund 5000 Lieferanten von Waren und Dienstleistungen im Krankenhausportal und können durch Markttransparenz und Konsortienbildung erhebliche Preisvorteile erzielen [Krankenhausportal; vgl. auch Portal AG]. Es gehört wenig Phantasie dazu, diesem Portaltypus einen erheblichen Bedeutungszuwachs zu prognostizieren. In dem Maße, in dem Unternehmen die Rationalisierungspotentiale entdecken, die darin bestehen, die Kontakte zu Geschäftspartnern über das Web abzuwickeln, werden branchenspezifische vertikale Portale im B2B-Bereich eine ähnliche Schlüsselstellung einnehmen wie Internetportale für den endnutzerorientierten E-Commerce.

### **5.3. Wissenschaftsportal**

Zugriff und Aufbereitung von Informationen, Selektion und Navigation sowie Organisation von Kooperation und Kommunikation stellen im Wissenschaftssektor ähnliche Anforderungen wie im betrieblichen Sektor. Auch die im Wissenschaftssektor etablierten Informationssysteme und — strukturen sind in der Informationsgesellschaft einem umfassenden und andauernden Wandel



unterworfen. Manche Bibliotheken etwa haben in diesem Rahmen über Automatisierung, Digitalisierung und Virtualisierung einen Stand erreicht, der von Flexibilität zeugt, aber dennoch nicht befriedigen kann [vgl. auch Rösch 2000]. Die gesamte wissenschaftliche Publikationskette steht auf dem Prüfstand. Autoren, Verlage, Bibliotheken und Rezipienten fühlen sich in ihren traditionellen Rollen verunsichert. In manchen Prognosen etwa ist die Rede vom Untergang der gedruckten wissenschaftlichen Zeitschrift. Ersatz sollen Preprint-Server oder E-Journals bieten, die gegebenenfalls die Zwischenglieder Verlag und Bibliothek zumindest partiell entbehrlich machen [Hilf 2000].

Es gibt bereits Aktivitäten und Initiativen, die sich darum bemühen, portalähnliche Konstrukte für den Wissenschaftsbetrieb zu entwickeln. Weit fortgeschritten sind die Vorarbeiten des Boston College, die in Kooperation mit knapp zwei Dutzend US-amerikanischen Universitäten erfolgen. Ziel der im Dezember 1999 gegründeten Initiative ist die konzeptionelle und technische Entwicklung eines "University-wide Information Portal", das folgendermaßen definiert wird:

*"(University portals) integrate campus-specific information, which is stored in the campus electronic vaults (i.e., databases, file systems and existing application systems) with unstructured data (text) from on and off campus. Providing a single, personalized interface to all information resources in a secure, consistent and customizable manner is the objective of the Boston College University-wide Information Portal design..."* [Gleason 2000, S. 2].

Die Präsentation eines Prototyps ist für Herbst 2000 angekündigt. Besonderer Nachdruck wird auf die Personalisierungsoptionen gelegt:

*"The most important concept is that all relevant information and services will be delivered in a personalized and coherent form to the individual, not in the traditional hierarchical structure."* [Gleason 2000, S. 5]

Die Bedeutung kommunikativer Funktionalitäten wird zwar erwähnt, scheint jedoch gegenüber dem Aspekt der Informationsversorgung deutlich geringer gewichtet zu sein. Ganz unberücksichtigt bleiben hingegen kollaborative Aspekte, die doch für den akademischen Bereich eminente Bedeutung besitzen.

Diese wie auch andere amerikanische Portalkonzepte für den Wissenschaftssektor [vgl. etwa Cohen et al. 2000] sind deutlich zugeschnitten auf dortige Spezifika: Der Campus als sozialer Kosmos, die Hochschule als eigenständiges kulturelles und ökonomisches Kraftzentrum oder der Wettbewerb der Hochschulen untereinander schaffen besondere Rahmenbedingungen, die es prinzipiell erleichtern, die Hochschule auch als Unternehmen zu begreifen und etwa im Konzept des Enterprise Information Portals ein adaptionsfähiges Konzept zu sehen. Ausdrücklich und ohne Scheu betont wird z.B. die kommerzielle Verwertbarkeit des "University-wide Information Portals" nach dem Vorbild der Internetportale. Unterentwickelt bleiben jedoch die in der Portalkonzeption akzentuierbaren Aspekte des hochschulinternen Wissensmanagements und der hochschulübergreifenden Kooperation. Warum die Unterstützung der

akademischen Kommunikation und wissenschaftlichen Produktion auf Universitätsebene nicht fokussiert wird, ist kaum erklärlich. Die Konzentration auf den Bezugsraum "eigene Hochschule" hingegen dürfte auf die genannte Spezifika der amerikanischen Hochschullandschaft zurückzuführen sein.

Für den deutschen Kontext ist ein solches universitätsweites Portal (zumindest ohne integrierendes Metaportal) mittelfristig suboptimal. Die Beschränkung auf eine Institution kollidiert zudem mit der Tatsache, dass die Informationsräume, in denen Wissenschaftler als Nehmende wie Gebende wirken, kaum in zufriedenstellender Weise isoliert, d.h. von einer Institution zu erschließen und bereitzustellen sind. Bibliothekarische Projekte wie die arbeitsteilige, themenzentrierte Erschließung relevanter Internetquellen weisen in diese Richtung und lassen sich ideal in den Typus des Wissenschaftsportals integrieren. Zu nennen ist etwa das von OCLC (Online Computer Library Center, Dublin/Ohio) betriebene Projekt CORC (Cooperative Online Resource Catalog) oder die in Großbritannien unter dem Begriff "Subject Gateway" rubrizierten Aktivitäten. In eine ähnliche Richtung zielt das funktional breiter angelegte Konzept der "Virtuellen Fachbibliothek", das von der Deutschen Forschungsgemeinschaft entwickelt und gefördert wird.

Das Wissenschaftsportal soll die portaltypischen Merkmale aufweisen und darüber hinaus in besonderer Weise institutionsinternes wie -übergreifendes Wissensmanagement unterstützen. Besonderes Gewicht kommt daneben der vertikalen Kumulation der am wissenschaftlichen Kommunikations-, Erkenntnis- und Verwertungsprozess beteiligten Personen und Institutionen zu. Das Wissenschaftsportal ist damit typologisch verwandt mit dem Unternehmensportal, es zeigt zudem ausgeprägte Merkmale eines B2BPortals. Idealerweise integriert das Wissenschaftsportal disziplinspezifische Subportale und erlaubt den zugehörigen Institutionen Anpassungen und Modifikationen nach jeweiligem Bedarf. Das vom nordrhein-westfälischen Hochschulbibliothekszenrum (HBZ) getragene Projekt "Digitale Bibliothek NRW" weist zahlreiche Ansätze auf, die eine Weiterentwicklung zu einem Wissenschaftsportal vorstellbar werden lassen. Unter einer einheitlichen Oberfläche werden heterogene Datenbanken angeboten, die parallel durchsucht werden können. Externe Pressedatenbanken sind ebenso in das Angebot integriert wie elektronische Verlagspublikationen. Die Implementierung weiterer Funktionalitäten insbesondere der Personalisierung ist angekündigt. Bislang liegt noch keine elaborierte Konzeption eines Wissenschaftsportal vor. Die potentielle Bedeutung und prinzipielle Adaptionfähigkeit des Portalkonzeptes für den Wissenschaftssektor steht jedoch außer Frage.

## 6. Entwicklungsperspektiven

Mit ihrem Funktionsumfang und ihrem Leistungsvolumen erweisen sich Internetportale und Unternehmensportale als Instrumente und Techniken, die es in den jeweiligen Anwendungsfeldern erst erlauben, steigenden Ansprüchen gerecht zu werden. Aufgrund ihres Integrations- und Personalisierungspotentials sind Portale prädestiniert, zu den Marktplätzen des E-Commerce zu werden und den

adäquaten Rahmen für Lernprozesse auf individueller, institutioneller oder überinstitutioneller Ebene zu bilden. Es zeichnet sich ab, dass neben die zur Zeit im Internet dominierenden horizontalen Portale eine Vielzahl spezialisierter, vertikaler Portale oder auch Nischenportale treten wird. Im Unternehmenssektor deutet sich mehrere Trends an. Zum einen scheint möglich, dass in großen Umgebungen ein Unternehmensportal mehrere vertikale interne Portale umfasst und zusammenführt. Zum anderen werden Portale im betrieblichen Kontext immer stärker den Business-to-Business-Aspekt integrieren. Denkbar erscheint auch, dass die Business-to-Consumer-Relation, die eher den Internetportalen zuzuschreiben ist, hinzutritt, B2B und B2C Portale also partiell konvergieren. Die Entwicklung von Portalkonzeptionen für weitere Bereiche, insbesondere den Wissenschaftssektor ist zu erwarten.

## **Literatur**

[Appsolute 1999] The Next Generation Enterprise Portal. Integrating the Knowledge and Processes of the Enterprise. Appsolute Software White Paper. 1999. <http://www.appsolut.com>.

[Brockhaus 1998] Brockhaus Enzyklopädie in 24 Bänden. Bd. 17. 20. Aufl. Mannheim 1998.

[Delphi Group 1999a] Enterprise Portals Shape Emerging Business Desktop. The Delphi Group Inc. 26.1.1999. <http://www.delphigroup.com>.

[Delphi Group 1999b] Excerpts From Delphü's Corporate Portal Report. Delphi Group 1999. [www.delphigroup.com/pubs/Corporate-portalexcerpt.htm](http://www.delphigroup.com/pubs/Corporate-portalexcerpt.htm).

[Essex 1999] Essex, David: Get Into Web Portals. In: Computerworld. 15.3.1999. <http://computerworld.com>.

[Firestone 2000] Firestone, Joseph M.: The Enterprise Knowledge Portal revisited. Executive Information Systems Inc. White Paper No. 53. 15. März 2000. <http://www.dkms.com>.

[Frenko 1998] Frenko, Andreas: Internet-portal-Sites. Aktueller Hype oder mehr? <http://www.webagency.de/infopool/internetwissen/ak981123.htm>.

[Gleason 2000] Gleason, Bernard W.: Boston College University-Wide Information Portal. Concepts and Recommended Course of Action. January 26, 2000. S. 2. [www.mis2.udel.edu/ja-sig/whitepaper.html](http://www.mis2.udel.edu/ja-sig/whitepaper.html).

[Hilf 2000] Hilf, Eberhard: Wissenschaftliche Kommunikation. [www.physik.uni-oldenburg.de/~hilf/vortraege/dgi-2000](http://www.physik.uni-oldenburg.de/~hilf/vortraege/dgi-2000)

[Janetzko 1999] Janetzko, Dietmar: Statistische Anwendungen im Internet. München 1999.

[Koulopoulos 2000] Koulopoulos, Thomas: Ask the Expert., Thomas Koulopoulos, President and founder of The Delphi Group. March 6<sup>th</sup> 2000. <http://www2.cio.com/CIO/epert/viewganda/cfm?ID=23>.

[Krankenhausportal] Krankenhausportal. <http://www.krankenhausportal.de>.

[Kuhlen 1999] Kuhlen, Rainer: Die Konsequenzen von Informationsassistenten. Frankfurt a.M. 1999.

[Lexikon der Kunst 1989] Lexikon der Kunst. Malerei, Architektur, Bildhauerkunst. Bd. 9. Freiburg 1989.

[Oilfield Directory]. Oilfield Directory. <http://www.oilfielddirectory.com>.

[Peek 1998] Peek, Robin: Wanting to be Everything to Everyone. Web search engines/directories battle to be your portal of choice. In: Information today. 15, September 1998, S. 37.

[Peek 1999] Peek, Robin: When is a Web Site a 'Vortal'? Vertical Portals seek to focus content to attract more visitors. In: Information today. 16, 1999, 8, S. 36-37.

[Cohen et al. 2000] Cohen, Suzanne, John Fereira, Angela Horne, Bob Kibbee, Holly Mistlebauer, Adam Smith: MyLibrary. Personalized Electronic Services in the Cornell University Library. In: D-Lib Magazine. April 2000, Volume 6, Number 4. <http://www.dlib.org/dlib/april100/mistlebauer/04mistlebauer.html>

[Portal Time 1999] Portal Time: Five Years That Shook the World? In: Knowledge Management Magazine. July 1999. <http://kmmag.com/kmmagn2/km199907/featurea2.htm>.

[Portal AG] Portal AG. Der Kölner Start-Up bringt Branchen ins Web. [http://www.portal-ag.de/pr/802\\_894.html](http://www.portal-ag.de/pr/802_894.html).

[Rezayat 2000] Rezayat, Mohsen: The Enterprise-Web portal for life-cycle support. In: Computer-Aided Design. 32, 2000, S. 85-96.

[Roberts-Witt 1998] Roberts-Witt, Sarah L.: Corporate Portal. In: Knowledge Management Magazine. December 1998. <http://kmmag.com/km199812/home.htm>.

[Roberts-Witt 1999] Roberts-Witt, Sarah L.: Making Sense of Portal Pandemonium. In: Knowledge Management Magazine. July 1999. <http://kmmag.com/kmmagn2/km199907/featureal.htm>.

[Rösch 2000] Rösch, Hermann: Spezialbibliothek mit neuen Aufgaben. In: Bibliothek. Forschung und Praxis. 2000, 1, S. 93-99.

[Schmidt 1999] Schmidt, Carsten: Konzeption eines Groupware-basierten Enterprise Knowledge Portals'. 1999. <http://www.enterprise-portal.de/einleitung.html>.

[Schumacher/Schwickert 1999] Schumacher, Mark und Axel C. Schwickert: Web-Portale. Stand und Entwicklungstendenzen. Mainz 1999.

[Shilakes/Tylman 1998] Shilakes, Christopher C, Julie Tylman: Enterprise Information Portals. In-depth Report. Merrill Lynch Enterprise Software Team. 16 November 1998. [http://www.sagemaker.com/company/downloads/eip\\_indepth.pdf](http://www.sagemaker.com/company/downloads/eip_indepth.pdf).

[Vortals.com] The Vortals.com. <http://thevortals.com>.

[Zimmermann 2000] Zimmermann, Anke: Das Tor zum Wissen. In: Screen. Business Online. Juni 2000, S. 34-38.



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 265 – 284

# Regionale Elektronische Zeitungen: Qualitätskriterien und Evaluierung

**Marc Rittberger / Wolfgang Semar**

Informationswissenschaft

Universität Konstanz

{Marc.Rittberger, Wolfgang. Semar} @uni-konstanz.de

## Zusammenfassung

Regionale und lokale Zeitungen sind durch besondere Qualitätsmerkmale gekennzeichnet, die im vorliegenden Beitrag diskutiert werden. Ausgehend von dem Konstanzer Ansatz zur Bewertung von Informationsdienstleistungen werden zeitungsspezifische Kriterien eingearbeitet und den fünf Hauptkennzahlen Information, Präsentation, Interaktion und Kommunikation, Technik und Methodik sowie Organisatorisches und Soziales zugeordnet. In einer ersten Untersuchung wurde die Tauglichkeit des Ansatzes anhand der regionalen Tageszeitungen Südkurier und Schwäbische Zeitung, die beide ihr Verbreitungsgebiet am Bodensee haben, untersucht.

## 1. Einleitung

Kaum waren die ersten Ausgaben elektronischer Zeitungen erschienen, wurde auch schon wieder ihr Untergang prophezeit. Der "Tod der elektronischen Zeitung", der noch vor ein paar Jahren durch die Massenmedien [Polatschek 1996] geisterte, hat sich aber inzwischen verabschiedet und ist der allgemeinen Begeisterung für die Internet Economy folgend zum "Zeitungshäuser sehen Boom durchs Internet" geworden [o.V. 2000a]. Gestützt wird dieser Optimismus durch die rasanten Zuwächse insbesondere im Werbebereich auf dem World Wide Web. Zeitungs- und Zeitschriftenverleger erhoffen, sich von dem immer größer werdenden Kuchen ein erhebliches Stück abschneiden zu können, da Werbeeinnahmen zur Zeit die Hauptfinanzierungsquelle für die Online-Angebote von Zeitungen und Zeitschriften sind [Lehr 1999], S. 90. Auch vom direkten Verkauf von Produkten und Dienstleistungen im Internet erhoffen sie sich eine Refinanzierung ihrer Angebote. Aber reicht es, die Printversion ins Netz zu bringen [Neuberger 1999], oder sind die Ansprüche nicht größer, die man in diesem Zusammenhang an Online-Zeitungen stellen muss? Die Finanzierung von Werbeangeboten und der Verkauf von Produkten und Dienstleistungen lässt sich wie schon das klassische Anzeigengeschäft nur über eine möglichst große Anzahl von Nutzern realisieren. Also gilt es auch hier, attraktive Angebote zu realisieren, die zielgruppengerecht angeboten werden. Allerdings steht dem die Entwicklung



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

im Print-Geschäft entgegen, welches für die interessante Werbegruppe der 14-19jährigen eine markant niedrige Leseratte mißt [Stürzebecher 1997; Schönbach & Peiser 1997], Tendenz weiter sinkend [o.V. 2000b]. Dem Dilemma der sinkenden Leseratte belegend empfehlen [Schönbach & Peiser 1997] Zeitungen, sich entweder an erfolgreiche Medien wie etwa Fernsehen anzupassen oder eigene Vorzüge und Stärken besser herauszuarbeiten oder der traditionellen Konzeption treu zu bleiben. Angesichts der rasanten Entwicklung der letzten Jahre, insbesondere auf dem Internet, bleiben für Zeitungen wohl nur die ersten beiden Ziele als realistisch erhalten.

Waren Ende 1996 erst 66 [Riefler 1995a], Ende 1997 schon 141 [Klettke et al. 1998], so sind laut Homepage des Bundesverbandes Deutscher Zeitungen<sup>1</sup> im Mai 2000 immerhin 184 Zeitungen mit ihren Angeboten online im Web. [Rada 1999] stellt die Vor- und Nachteile der Onlineversionen von Zeitungen den gedruckten Ausgaben gegenüber. Nutzer schätzen dabei insbesondere die kostenlose Nutzung, die gezielte Beitragssuche, die Aktualität, das Archiv, den Zugriff auf alle Angebote, die Nutzung von Links und die Interaktivität, wohingegen die fehlende Portabilität, eingeschränkte Zugangsbedingungen, schlechte Lesbarkeit und unvollständige Berichterstattung als besonders negativ genannt wurden. Relaunches der Internet-Auftritte sind an der Tagesordnung, um das Publikum besser anzusprechen und damit die Zugriffsraten zu erhöhen. Aber sind die Zeitungen damit auch erfolgreich? Während die überregionalen Angebote mit vielen Mehrwerten aufwarten, eigene Online-Redaktionen aufbauen und dem Medium angepasste Webauftritte entwickeln, sind regionale Zeitungen durch knappere Budgets gezwungen, ihre Internet-Angebote auf andere Qualitätskriterien zu fokussieren. Sie, deren Hauptattraktion regionale Nachrichten und Angebotsvermittlung, also Anzeigengeschäft, sind, müssen einen kleineren Kundenkreis ansprechen als die überregionalen Zeitungen und deren durch aufwendige, interaktive Angebote der großen Verlagshäuser verwöhnte Kundschaft weiter am regionalen Angebot interessieren.

Nach einer Einführung in Qualitätsbetrachtungen im Medienbereich, speziell dem Zeitungs- und Zeitschriftenangebot, wollen wir im Folgenden einen Kriterienkatalog für den erfolgreichen Auftritt von Tageszeitungen formulieren und mittels eines in Konstanz entwickelten Verfahrens überprüfen. Zur Überprüfung des Verfahrens soll der Webauftritt zweier regionaler, deutscher Tageszeitungen am Bodensee genutzt werden, um damit die Tauglichkeit des Kriterienkatalogs zu veranschaulichen.

## **2. Qualität von Zeitungen im World Wide Web**

Zeitungen werden von Menschen gemacht und hauptsächlich von Journalisten geprägt. So verwundert es nicht, dass im Zuge der Diskussion über die Qualität von Zeitungen vornehmlich die journalistische Qualität und ihr Einfluss auf die Qualität der von ihnen mit Informationen und Beiträgen versorgten Medien [Bramme et al. 1993], hier Zeitungen, eine Rolle spielt. Eine immer relevante, nicht

---

<sup>1</sup> <http://www.bdzv.de> am 15.5.2000

sehr originelle Betrachtung der Qualitätsabhängigkeit fußt auf dem Zusammenhang, dass die Qualität von Zeitungen durch einen erhöhten Einsatz von Kapital, insbesondere Personal, verbessert werden kann. In [Turner 1995] werden diese These und ihre Zusammenhänge detailliert diskutiert und ein mathematischer Zusammenhang zwischen der Auflagenstärke und der Anzahl der Mitarbeiter hergestellt. Im allgemeinen wird aber davon ausgegangen, dass die Qualitätseinschätzung schwierig ist ("Qualität im Journalismus definieren zu wollen, gleicht dem Versuch, einen Pudding an die Wand zu nageln" [Ruß-Mohl 1994] oder "Die Journalistische Qualität schlechthin kann es nicht geben" [Wallisch 1995]). Neben klassischen handwerklichen Qualitäten, wie Beherrschen der W-Fragen, der Recherche, Faktentreue, Exklusivität oder der Trennung von Information und Meinung, beschreibt [Wallisch 1995] in Anlehnung an [Saxer & Kuli 1981] sechs Normen journalistischer Qualität:

- Die formale Bearbeitung beschreibt die Fähigkeit, die journalistischen Darstellungsformen wie Reportage, Feuilleton zu beherrschen.
- Die thematische Bearbeitung erlaubt, den Neuigkeitswert eines Themas darstellen zu können und zu erkennen, inwieweit es für das vorgesehene Publikationsorgan und die Leserzielgruppe relevant ist.
- Institutionelle Normen geben Sicherheit, den vorgesehenen Beitrag als Journalist in der gewünschten Art und Weise schreiben und bspw. die dafür notwendigen Recherchen durchführen und Hintergrundinformationen sammeln zu können.
- Die erfolgreiche Platzierung eines Beitrags in seinem Umfeld, so dass er auch publiziert wird.
- Aktualität, Inhaltswahl und Universalität, d.h. dass das Publikum sich durch die Berichterstattung angesprochen fühlt, sind Grundvoraussetzungen journalistischer Arbeit, da die Fähigkeit, sein Publikum anzusprechen, erst einen "guten Journalisten" ausmacht.
- Die Erfüllung der Publikumserwartungen durch die Medien und damit durch die Beiträge ist eine Notwendigkeit, da damit die wirtschaftliche Grundlage für die Medienunternehmen gelegt wird.

Neben der geringen Akzeptanz und dem mangelnden Bewußtsein für Qualitätsfragen in Redaktionen überhaupt [Ruß-Mohl 2000a; Ruß-Mohl 2000b] und der in der Literatur relativ einseitigen Qualitätsdiskussion im Hinblick auf journalistische Belange gilt es, bei der Betrachtung von Zeitungsprodukten zu beachten, dass hier zwei wesentliche Zielgruppen relevant sind [Wolff 1998]. Geht man von einem kundenorientierten, allgemeineren Qualitätsbegriff für Zeitungen aus, so gilt es, sowohl die Ansprüche des Lesers als auch der Anzeigenkunden zu berücksichtigen. Geht der unbedarfte Leser zunächst davon aus, dass Zeitungen vornehmlich für ihn gemacht werden, irrt er. Kostendruck und Qualitätsverbesserungen in den letzten Jahren gingen meistens zu Lasten der



Redaktionen und zu Gunsten der Anzeigenbereiche von Zeitungen, also kamen weniger der lesenden Kundschaft, denn den Anzeigen schaltenden Kunden zu Gute.

Betrachtet man die Online-Versionen von Zeitungen und nimmt die Erkenntnis mit, dass zumindest zur Zeit die Durchsetzung von Gebühren beim Leser unmöglich ist, so basieren die Einnahmen nur noch auf Werbe- und Anzeigenkunden [Neuberger et al. 1997]. Somit kommt der Reichweite im Internet und den Internet-spezifischen Werbeformen eine noch größere Bedeutung für die Einnahmen der Zeitungen zu. Natürlich sind die Zeitung als Werbeträger und die Zeitung als Informationsträger nicht getrennt betrachtbar, die Zusammenhänge sind komplex. Die Verbreitung von Tageszeitungen über das Internet wird die Anzeigenabteilungen weiter stärken, da die Aussage "bringen die Pfennige der Leser die Markstücke der Anzeigenkunden" [Wolff 1998] S.264, noch mehr gilt.

Angesichts dieser Ausgangslage sind zur Messung der Qualität von Zeitungen auf dem Internet regionale und lokale Blättern von besonderem Interesse, da die Kundenbindung bei diesen Publikationen in hohem Maß durch die Berichterstattung über die regionalen und lokalen Informationen geprägt ist. Betrachten wir hierzu die Nutzungszahlen einer nichtrepräsentativen Befragung von 1313 Lesern von lokalen/regionalen Zeitungen und 638 Lesern überregionaler Blätter auf dem Internet [Neuberger et al. 1997] S.657.

Tabelle 1 zeigt für die wichtigsten Rubriken insbesondere bei lokalen und regionalen Nachrichten und Kleinanzeigen eine deutlich höhere Nutzung bei den lokalen und regionalen Zeitungen. Die überregionalen Leser favorisieren dagegen insbesondere die Themen Politik und Wirtschaft. Beide Nutzertypen bestätigen damit die schon aus dem Printbereich bekannten Vorlieben. Interessant ist in diesem Zusammenhang, dass bei der genannten Umfrage 22% der Beteiligten als Hauptwohnsitz einen Ort ausserhalb Deutschlands angaben, 72% dieses knappen Viertels hatten zuvor im Verbreitungsgebiet der entsprechenden Zeitung gewohnt.

	<b>Lokale/Regionale (N=1313)</b>	<b>Überregionale (N=638)</b>
Politik	55,4	88,5
Wirtschaft	59,6	76,9
Lokales/Region	71,7	34,4
Wissenschaft/T	58,7	65,5
News-Ticker	43,3	35,1
Sport	33,0	38,7
Kultur	35,4	45,1
Buntes	24,9	21,1
Archiv	31,4	32,8
Kleinanzeigen	26,1	8,2

**Tabelle 1:** Nutzungszahlen von Lesern lokaler und regionaler Zeitungen nach [Neuberger et al. 1997]

Qualität regionaler Zeitungen auf dem Internet ist natürlich durch die Potenziale des neuen Mediums in bezug auf das alte Medium, die gedruckte Zeitung, charakterisiert, inwieweit sie also eher die Printversion als "Zeitung pur" übernehmen oder "Special Interest" Angebote hinzufügen oder wirklich eine Art "lokales Informationszentrum" aufbauen möchten [Riefler 1995a], S.90ff., und [Klettke et al. 1998]. Geprägt werden sollten Darstellungen von regionalen Zeitungen durch ihre Stärken, die Lokal- und Regionalnachrichten in Politik, Wirtschaft, Sport, etc. auf ihre Zielgruppe, die Leser und Anzeigenkunden aus der Region hin zu optimieren [Flöpper 1995]. Zusätzlich werden Informationen über Bereitschaftsdienste, bspw. der Apotheken, Öffnungszeiten, Ämter und vieles mehr zugänglich, so dass ein lokales Informationszentrum oder Portal entstehen kann, der Informationsanbieter der Region [Riefler 1995b]. Begleitet werden kann dieser Informationskern durch Serien, historische Beiträge (Was war vor 50 Jahren in unserer Stadt?), Vorstellung von Vereinen, Restaurantkritiken, Newsticker, Veranstaltungskalender, Hörproben für Konzerte o.ä.. Hier zeigt sich auch die Mächtigkeit des Mediums, welches zum einen keine Beschränkung auf eine bestimmte Seitenzahl bzw. Umfang der Zeitung mit sich bringt, und auch in der Form mehr Freiheiten zulässt. So kann eine Öffnung natürlich auch in Richtung einer selbsttätigen Berichterstattung von Vereinen und Institutionen gehen, denen damit eine Plattform zur Kommunikation und Informationspräsentation gegeben wird. Sie können dann aktuelle Vereinsnachrichten, wie bspw. Spielberichte aus dem Sport, direkt über die Online-Zeitung ins Netz legen. Zeitungen können sich als regionale Kommunikationszentren etablieren, welche die klassische massenmediengeprägten 1:m Kommunikation aufbrechen.

Zeitungsangebote können individualisiert werden, um so aus der großen Menge des redaktionellen Angebots die für den einzelnen oder für typische Lesergruppen relevanten Beiträge zu filtern. Nicht zuletzt sollte das Internet viel stärker als Kommunikationsplattform zwischen Leser und Redakteuren sowie zwischen den Lesern selbst genutzt werden. Dabei steht nicht nur das Diskussionsforum als Mittel der Kommunikation der Nutzer untereinander im Mittelpunkt, sondern auch die unmittelbare Kontaktaufnahme mit dem Redakteur eines Artikels oder die Ausweitung des klassischen Leserbriefs.

Zeitungen finden bei Frauen und Jugendlichen häufig nicht die Anerkennung wie bei Männern [o.V. 2000b], da sie von Männern für Männer in mittleren Jahren gemacht sind [Naumann 1995]. Andere inhaltliche Schwerpunkte, andere Aufbereitung und Schreibstile, klarere Standpunkte und Meinungen sind in diesem Zusammenhang gefragt, um auch diese Lesergruppen anzusprechen. Marktchancen bietet das Internet regionalen Zeitungen durch den möglichen Verkauf zusätzlicher Leistungen, Produkte und Angebote. Es können dabei auf den regionalen Informationsportalen neue Formen der Öffentlichkeit geschaffen werden, die mehr Transparenz in unser wirtschaftliches und gesellschaftliches Zusammenleben bringen können.

### 3. Kriterienkatalog und Kennzahlensystem

Die Bewertung von Dienstleistungsqualität kann nach unterschiedlichen Ansätzen erfolgen. [Haller 1995; Bruhn 1997] beschreiben Methoden zur Beurteilung von Dienstleistungsqualität und unterscheiden in kundenorientierte und unternehmensorientierte Ansätze. Ein Kennzahlen-system zu verwenden, erlaubt eine Diskusssion auf beiden Ebenen, da ein Kriterienkatalog erstellt wird und die Gewichtung der einzelnen Kriterien je nach Blickwinkel erfolgen kann. Kennzahlensysteme werden den multiattributiven Verfahren zugeordnet, bei denen die Gesamtqualität eines Produktes oder Dienstleistung durch die Kombination der Qualitätsurteile der einzelnen Kriterien erlangt wird.

Ausgehend von einer Spitzenkennzahl können die Kriterien für die Bewertung von regionalen Zeitungen in verschiedene Hauptkennzahlen unterteilt werden. In bezug auf [Kuhlen 1998] und [Rittberger 1999] unterteilen wir in fünf Hauptkennzahlen oder Ebenen:

- In der **Informationsebene** werden Inhalte und inhaltliche Aspekte einer elektronischen Zeitung beschrieben.
- Die **Präsentationsebene** umfasst die genutzten Stilmittel. Es gelten designerische Grundsätze bspw. bei der Gestaltung von Bildschirmoberflächen.
- Die **Interaktions- und Kommunikationsebene** verdeutlicht die eingesetzten Such- und Navigationsverfahren und andere Interaktionsformen, und sie beschreibt die kommunikativen Möglichkeiten, welche zum Austausch mit Redakteuren, anderen Lesern etc. angeboten werden.
- Die **soziale und organisatorische Ebene** umfasst Kriterien zur Vertrauensbildung und Geschäftspolitik, wie etwa Datensicherheit oder Angebotsvielfalt.
- Die **technische und methodische Ebene** bezieht sich auf Kriterien, die die benutzte Technik und ihre Ausprägungen beschreiben wie bspw. die Robustheit einer Kommunikationsphase.

Mit diesem Ansatz bewegen wir uns durchaus im Kontext anderer Bewertungen. Bspw. wurden in [o.V. 2000c] die sechs Kriterien Content, Interaktion, Konzept, Nutzerfreundlichkeit, Technik und visuelle Umsetzung für die Bewertung überregionaler Zeitungen unterschieden.

Tabelle 2 gibt Auskunft, welche Unterkriterien auf der ersten Hierarchieebene zur Beurteilung der einzelnen Hauptkennzahlen beitragen. Auf eine erschöpfende Darstellung und Beschreibung der einzelnen Kriterien wurde verzichtet, da sie den Umfang der Arbeit sprengen würde. Bei der Auswahl der Kriterien wurde auf die Erfahrungen mit dem Aufbau von Kennzahlensystemen in Konstanz

zurückgegriffen [Rittberger 1999] und spezifische Qualitätsdiskussionen von Zeitungen aufgenommen [Schönbach 1997; Neuberger & Tonnemacher 1999; Rada 1999; Rittberger 2000].

Information	Präsentation	Interaktion/ Kommunikation	Sozial/ Organisation	Methodik/ Technik
Breite/Vielfalt	Ausgewogenheit	Suchmöglichkeit	Datenschutz	Zuverlässigkeit
Zielgruppen-orientierung	Benutzergruppenorientiert	Auffindbarkeit	Datensicherheit	Transaktionen
Einzigartigkeit	Kognitive Prinzipien	Verknüpfungsdichte	Vertrauensmanagement	Ladezeit
Vollständigkeit	Hypertextstruktur	Klassifikation	Metainformation	Robustheit
Regionalität	Konsistenz	Navigationshilfen		Technikeinsatz
Verlässlichkeit	Lesbarkeit	Diskussions-		
Aktualität	Werbung	Offenheit		
Nützlichkeit/ Relevanz	Schreibstil/ Aufbereitung	Anzeigen		
Genauigkeit				
Vielfalt				
Archiv				

**Tabelle 2:** Die fünf Hauptkennzahlen und die ihnen zugeordneten Kriterien zur Bewertung.

Anhand des vorgestellten Kriterienkatalogs, der an die Belange elektronischer Zeitungen angepasst wurde, werden wir im Folgenden die Tauglichkeit des Ansatzes für die Online-Angebote des Südkuriers und der Schwäbischen Zeitung prüfen. Beide Zeitungen haben ihre Verbreitung im Bodenseegebiet und können als typische regionale Blätter angesehen werden. In der Betrachtung der Informationsebene wird dabei nicht auf die eigentliche inhaltliche Qualität der Blätter eingegangen, die wie schon ausgeführt nur schwer messbar ist und mit der Qualität der redaktionellen Arbeit eng zusammenhängt. Wir fokussieren unsere Bewertung auf die Unterschiede, die sich zwischen Print- und Online-Version ergeben, wobei beim Vergleich das klassische Print-Medium als Ausgangsposition dient.

Die beiden regionalen Tageszeitungen sind seit mehr als einem Jahr „online“ und daher etabliert, der Südkurier ging allerdings fast ein Jahr vor der Schwäbischen Zeitung ans Netz. Der Südkurier ist seit dem 01. Mai 2000 mit einem Relaunch seiner Webdarstellung präsent, die hier gemachten Angaben beziehen sich ausschließlich auf die neue Version. Beide Zeitungen haben nur teilweise das gleiche Verbreitungsgebiet, in dem aber ein harter Wettbewerb um die führende Marktposition herrscht. Um eine bessere Vergleichbarkeit bei der Bewertung der Online-Versionen zu erreichen, wurde deshalb nur der Überschneidungsbereich, der sich im wesentlichen auf den Bodenseekreis und

somit auf die Stadt Friedrichshafen und ihre Umgebung beschränkt, betrachtet. Die Untersuchungen wurden vom 11.-13. Mai 2000 durchgeführt.

## 4. Südkurier<sup>2</sup>

Das Erscheinungsbild der Online-Version ist völlig losgelöst von dem der Print-Version und ganz auf die Eigenschaften des World Wide Web zugeschnitten. Dementsprechend firmiert die Web-Seite mit einem eigenen Logo unter „SKOL“ (Südkurier Online). Dass eine inhaltliche Verbindung zum Südkurier als Nachrichtenlieferant besteht, ist auf der Web-Site an dem Untertitel: „Mit Nachrichten aus dem SÜDKURIER“ zu erkennen. Was die Nachrichten betrifft, so sind fast alle regionalen Artikel aus der Print-Version in SKOL vorhanden. Darüber hinaus besticht die Online-Version durch zusätzliche aktuelle Agenturmeldungen, die nicht in der Print-Version vorhanden sind und laufend neu eingestellt werden. Ruft man die Startseite auf, so erscheint eine nüchterne, aber sehr klar strukturierte Bildschirmseite (siehe Abbildung 1). Am oberen Rand befindet sich die Haupt-Menüleiste mit den Rubriken „Nachrichten, Lokales, Marktplatz, Service, Sport, Treffpunkt, Magazin und Reise“. Jeder dieser Rubriken ist eine eigene Bildschirmhintergrundfarbe zugeordnet, um so eine optische Orientierung herzustellen. Unter dieser Menüleiste erscheint eine Kurzmeldung, rechts daneben drei Überschriften von aktuellen Meldungen mit der Angabe der Uhrzeit, zu der sie veröffentlicht wurden. Unter diesen Meldungen befindet sich ein Werbebanner, darunter ein Link zum Wetterbericht und zu ausgewählten regionalen Artikeln<sup>3</sup>. Auf der linken Seite sind in Höhe der Navigationsleiste drei Buttons, die allerdings nicht selbsterklärend sind. Beim Überfahren mit der Maus wird jedoch ein Kurztext<sup>4</sup> eingeblendet. Unter diesen Buttons ist ein weiteres Navigationselement vorhanden, mit dem Link zur „Frage der Woche“, darunter zwei kleine Werbebanner und abschließend das SKOL-Logo, das immer zur Startseite führt. Navigiert man in die einzelnen Rubriken, so erscheint am Bildschirmkopf die Haupt-Navigationsleiste und rechts daneben die aktuellsten Artikelüberschriften, diesmal jedoch von mehr als drei Artikeln. Auf der linken Seite hat sich das bereits angesprochene Navigationselement etwas verändert. Der Link zur „Frage der Woche“ ist nicht mehr vorhanden, dafür werden zusätzlich Unterrubriken, passend zur jeweils gewählten Hauptrubrik, eingeblendet. Bei „Nachrichten“ erscheinen die Unterrubriken „Politik, Wirtschaft, Kultur, 7-Tage Rückblick, Vermischtes, Baden Württemberg, Aktuelle Meldungen, und Computer“. Klickt man auf eine Unterrubrik, so erhält man die dazugehörigen Meldungen. Besonders gut gelungen ist die Rubrik „Lokales“. Auf einen Blick erhält man in dem linken Navigationselement eine Liste aller Städte<sup>5</sup>, zu denen es Lokalausgaben gibt. Klickt man eine Lokalausgabe an, so erscheint wieder die gewohnte Aufteilung, jedoch gibt es hier keine aktuellen Meldungen am rechten Rand, dafür sind an dieser Stelle jetzt zwei Werbebuttons vorhanden. Das linke Navigationselement enthält nun die Unterrubriken „Alle Artikel, Sport, Wetter, Kultur, Termine und Bodensee Zeitung“.

---

<sup>2</sup> <http://www.suedkurier.de>, <http://www.skol.de>

<sup>3</sup> Die Auswahl dieser Artikel erscheint den Autoren willkürlich.

<sup>4</sup> Die Buttons führen zum Impressum, Gästebuch und zur Suche.

<sup>5</sup> Bad Säckingen, Donaueschingen, Friedrichshafen, Furtwangen, Konstanz, Markdorf, Pfullendorf-Messkirch, Radolfzell, Singen, St. Georgen-Triberg, Stockach, Villingen, Waldshut-Tiengen, Überlingen

**Information:** Die Artikel in der Rubrik „Nachrichten“ werden laufend aktualisiert. Ab 17.00 Uhr werden alle Artikel komplett neu eingestellt, es handelt sich dabei um die Meldungen, die am nächsten Tag in der Print-Version erscheinen. Vergebens sucht man jedoch nach weiterführenden Links zu anderen relevanten Beiträgen am Ende eines Artikels, jeder Artikel steht somit für sich alleine. Die Breite, Vielfalt und Einzigartigkeit der online verfügbaren Nachrichten lassen jedoch keine Wünsche offen. In der Rubrik „Baden-Württemberg“ sind alle Artikel aus der Print-Version vorhanden, in „Lokales — Friedrichshafen“ sind die Kultur- und Sportnachrichten aus der Print-Version ebenfalls alle in der Online-Version vorhanden. Betrachtet man die Nützlichkeit und Relevanz der Artikel im Hinblick auf die Regionalität, so kann man den Verantwortlichen nur die Bestnoten vergeben. Unter der Rubrik „Magazin“ sind in den Unterrubriken „Lifestyle, Fahrrad, Thema des Monats und Bauen & Wohnen“ jeweils interessante Artikel zu diesen Themen zusammengestellt. So auch in der Rubrik „Reise“, hier findet man in den Unterrubriken „ Bodenseeferien, Reiseziele, Reiseangebote und Leserreisen“ Interessantes Incl. einer Suchmaske für die Hoteldatenbank der Bodensee-Ferienzeitung.

Zeitungsfremdes, aber durchaus onlineangemessenes findet man in der Rubrik "Marktplatz - Büchershop". Hier wird in Erweiterung der Printversion eine Verbindung zu dem Online-Buchhändler <http://www.booxtra.de> aufgebaut. Man erhält die Suchmaske des Buchhändlers, die aber in das Layout von SKOL integriert ist. Führt man eine Suche nach einem Buch durch, so wird das Ergebnis in einem neuen Browserfenster unter der Webadresse von Booxtra dargestellt. Dies stellt keinen wirklichen Mehrwert dar.

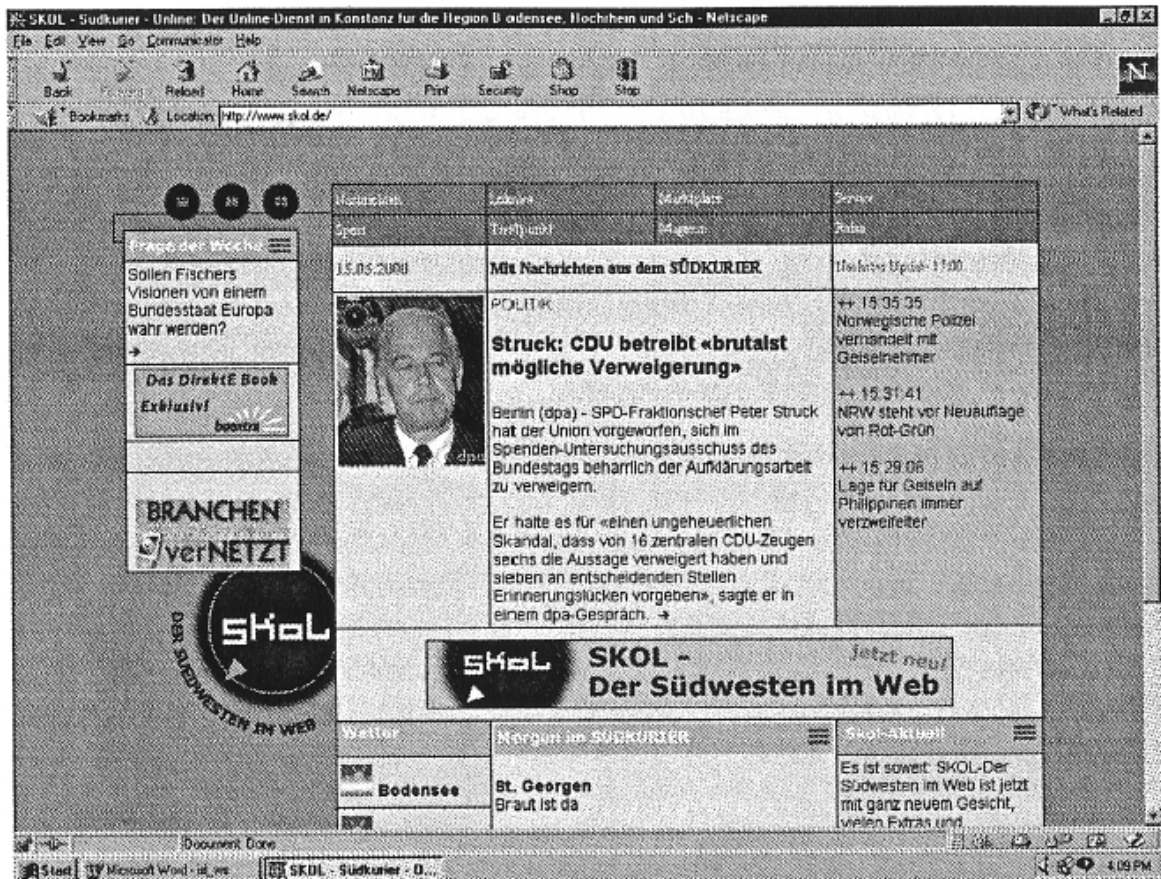


Abbildung 1: Startseite des Südkuriers

**Präsentation:** Der Erst-Nutzer findet sich durch den klar strukturierten Aufbau des Bildschirms sehr schnell zurecht. In der Rubrik „Nachrichten“ werden hauptsächlich dpa-Artikel veröffentlicht, die in ihrer Lesbarkeit, Aufbereitung sowie dem Schreibstil alle dem Standard entsprechen. Die eigenen Regionalberichte entsprechen augenscheinlich einer guten redaktionellen Qualität.

**Interaktion/Kommunikation:** Eine Interaktions-Funktion ist in der Rubrik „Kleinanzeigen“ vorhanden. Hier kann man mit Hilfe einer Suchmaske in den einzelnen Unterrubriken „KFZ, Immobilien, Partnerschaft, Stellen, An- und Verkauf und Verschiedenes“ gezielt nach den Anzeigen suchen. Allerdings ist die Menüführung, z. B. bei der Immobiliensuche, etwas gewöhnungsbedürftig. Die Zahl der insgesamt vorhandenen Anzeigen ist mit fast 2500<sup>6</sup> groß genug, um darin vernünftig suchen zu können. Das Aufgeben von Kleinanzeigen für die Print-Version ist in der Rubrik „Marktplatz — Kleinanzeigenannahme“ möglich. Das Aufgeben der Anzeigen erfolgt über ein Formular. Leider muß man sich durch drei verschiedene Formulare durcharbeiten, bis man endlich die Anzeige abschicken kann.

**Sozial/Organisation:** Das Thema Datenschutz, Datensicherheit oder Vertrauensmanagement ist bei dieser Web-Site völlig unberücksichtigt.

<sup>6</sup> Stand 14.05.2000

Persönliche Daten bzw. Angaben zu Konto- und Bankverbindungen, werden unverschlüsselt und ohne die Benutzung eines Sicherheitsservers über das Internet übertragen. Ob eine Überprüfung der angegebenen Angaben durch Dritte erfolgt, ist nicht ersichtlich. Dem Missbrauch beim Aufgeben von Anzeigen ist somit Tür und Tor geöffnet.

**Methodik/Technik:** Während der Testphase lief der Server konstant und ohne Ausfall. Es kam jedoch vor, dass einige Links ins Leere führen. Die Darstellung der HTML-Seite kommt ohne die Verwendung von Java- oder Javascript-Code aus und basiert auf der CSS-Technologie (Cascading Style Sheets). Dies setzt allerdings zur korrekten Darstellung der HTML-Seiten einen Browser voraus, der diese Technologie unterstützt. Leider lassen sich auf Basis des HTML-Codes keine Rückschlüsse auf die verwendete Datenbanktechnologie machen.

## 5. Schwäbische Zeitung<sup>7</sup>

Das Erscheinungsbild der Online-Version versucht durch die Farb- und die Logoverwendung eine Identität mit der Printversion herzustellen. Ruft man die Startseite auf, so erscheinen die Schlagzeilen des Tages in der Mitte des Bildschirms. Am linken Rand wird man durch zwei sehr auffällige Werbebanner (relativ großflächig und mit Animation versehen) eher abgelenkt als aufmerksam gemacht. Die rechte Seite flankiert die Schlagzeilen durch eine sehr farbige Gestaltung verschiedener Grafiken, die sich erst bei wiederholtem Hinsehen als Navigationsgrafiken, die zu speziellen Rubriken wie z. B. „Thema der Woche, SZ-Boulevard oder Web-Tip der Woche“ führen, entpuppen. Ganz am oberen Rand der Bildschirmseite befindet sich die Haupt-Navigationsleiste mit den Rubriken „News, Kalender, Speis & Trank, Bürger aktiv, Anzeigen, Fun und Auktionen“, ihr folgt am rechten daneben das Logo „Schwäbische Zeitung Online“, das auf die Startseite linkt. Darunter befindet sich eine weitere Navigationsleiste mit den Rubriken „Meine Schwäbische, Abo & Co, Verlag, Shop sowie Inhalt & Hilfe“. Diese Einheit aus Haupt-Navigationsleiste, Logo und Zwischen-Navigationsleiste befindet sich durchgängig auf jeder einzelnen Seite. Auf Erstbenutzer wirkt die Startseite überladen und man findet sich in Folge der vielen Navigationsmöglichkeiten zunächst einmal nicht zurecht. Folgt man dem Hinweis „Die Lokalnachrichten finden Sie unter NEWS“, so gelangt man auf die Seite „Die News des Tages“. Wie auf der Startseite befinden sich die Artikel in der Mitte des Bildschirms, umrahmt von diversen Navigationsmöglichkeiten. Auf der linken Bildschirmseite ist ein weiteres Auswahlmenü hinzugekommen, während sich bei der Bilderleiste auf der rechten Seite „nur“ die Grafiken und deren Links geändert haben. Das Auswahlmenü auf der linken Seite unterteilt sich in die Rubriken „Allgemeine News, Politik, Wirtschaft, Kultur und Sport“, die der Printversion entnommen sind. Klickt man die Rubrik „Allgemeine News“ an, so erscheint zunächst einmal folgende Fehlermeldung: "Leider konnte kein Ergebnis angezeigt werden. Entweder haben Sie bisher noch keine Region ausgewählt, oder es liegen in der von Ihnen ausgewählten Region keine Informationen vor." Was lief falsch?

---

<sup>7</sup> <http://www.schwaebischezeitung.de>, <http://www.schwaebische-zeitung.de>,  
<http://www.szonline.de>



Der mit dieser Navigation nicht vertraute Nutzer hat übersehen, dass sich in dem Auswahlmenü noch zusätzlich ein kleines, scrollbares Aufwahlfenster mit den zuvor auszuwählenden Regionen<sup>8</sup> befindet. Wählt man nun z. B. die Region „Friedrichshafen“ und klickt auf „Aktuelle News“, so ist man am Ziel angekommen und erhält die entsprechenden regionalen Artikel<sup>9</sup>. Diesem Prinzip folgend erhält man dann die Lokalnachrichten zu den entsprechenden Rubriken. Relativ unauffällig ist eine kleine Veränderung des Auswahlmenüs. Erscheinen die Lokalnachrichten auf dem Bildschirm, so erfährt das Auswahlmenü eine Erweiterung um die Punkte „Archiv und Special“. „Archiv“ führt zu einer Suchmaske, mit deren Hilfe man nach Stichworten in der Überschrift und/oder im Artikel suchen kann. Die Suche ist nur dann erfolgreich, wenn der Region jemals ein Artikel zugeordnet wurde, ist das nicht der Fall, erhält man die bereits bekannte Fehlermeldung. Der Klick auf „Special“ führt nur in den Regionen „International und Deutschland“ zum Erfolg, für alle anderen Regionen erscheint wiederum die bekannte Fehlermeldung.

In der Rubrik „Meine Schwäbische“ können Artikel ganz nach den eigenen Wünschen zusammengestellt werden. In einer umfänglichen Matrix kann man nach Angabe seiner persönlichen Daten bestimmen, aus welcher Region und aus welcher Rubrik man die Artikel erhalten möchte. Der Vorteil, den man gegenüber der „normalen“ Version erhält, ist, dass die Auswahl der Region entfällt. Durch Eingabe seines User-Namens und seines Passwortes auf der Startseite wird dem Nutzer dann „seine Schwäbische Zeitung“ angezeigt. In der Auswahlmatrix ist es möglich, die Nachrichten aus einer Rubrik und mehreren Regionen zusammenzustellen.

**Information:** Die Artikel in der Rubrik „News“ sind alle auch in der Print-Version vorhanden. Jedoch ist teilweise der Wortlaut im Vergleich zur Online-Version verschieden. Leider erhält man online nur „einige ausgesuchte“<sup>10</sup> Artikel. Vereinzelt sind am Ende eines Artikels Links vorhanden, die zu weiteren relevanten Artikeln führen, hiermit wird dem Leser die Möglichkeit gegeben, sich ausführlicher über ein Thema zu informieren. Die Breite, Vielfalt und Einzigartigkeit der online verfügbaren Artikel lassen jedoch zu wünschen übrig. Eine klare Aussage über die Aktualität der einzelnen Artikel zu machen, ist nicht möglich. So erscheinen an manchen Tagen<sup>11</sup> die tagesaktuellen Nachrichten erst nach 12.00 Uhr auf dem Bildschirm, wer sich vorher einloggt, muss sich mit den Meldungen

---

<sup>8</sup> Folgende Optionen stehen zur Verfügung: „International, Deutschland, Aalen, Biberach, Eilwangen, Friedrichshafen, Laichingen, Laupheim, Lindau, Markdorf, Ravensburg, Rottweil, Schramberg, Tettnang, Tuttlingen, Ulm, Wangen, ALLE“.

<sup>9</sup> Am Samstag, den 13.05.2000 (11.50 Uhr), handelt es sich dabei um zwei Artikel (Erscheinungsdatum der Artikel: 10.05., 09.05.). In den Rubriken „Politik (09.05), Wirtschaft (10.05), Kultur (09.05) und Sport (10.05)“ erscheint jeweils ein Artikel! Wählt man als Region „Markdorf“, erhält man die bereits bekannte Fehlermeldung: „Leider konnte kein Ergebnis angezeigt werden. Entweder haben Sie bisher noch keine Region ausgewählt, oder es liegen in der von Ihnen ausgewählten Region keine Informationen vor.“ Diesmal ist der Grund für das Auftreten der Meldung das Nichtvorhandensein von Artikeln, die der Region „Markdorf“ zugehören

<sup>10</sup> Die Auswahlkriterien sind nicht bekannt

<sup>11</sup> Insbesondere samstags

der vergangenen Tage zufriedengeben. Diese Tatsache trifft nur für die Rubrik „Allgemeine News“ zu, wer tagesaktuelle regionale Nachrichten haben möchte, wird auf dieser Web-Site teilweise gar nicht fündig<sup>12</sup>. Sonntags erscheint im Regionalteil generell nur die bereits bekannte Fehlermeldung. Betrachtet man die Nützlichkeit und Relevanz dieser Artikel im Hinblick auf die Regionalität, so kann man dieser höchstens ein „Mangelhaft“ vergeben. Ähnliche Zustände findet man in der Rubrik „Kalender“, hier sollten sich eigentlich die Kulturtermine aus der Region befinden. Aber wer diesen Kalender nutzt, um seine Abendgestaltung zu planen, bekommt nur eine sehr bescheidene Auswahl angeboten.

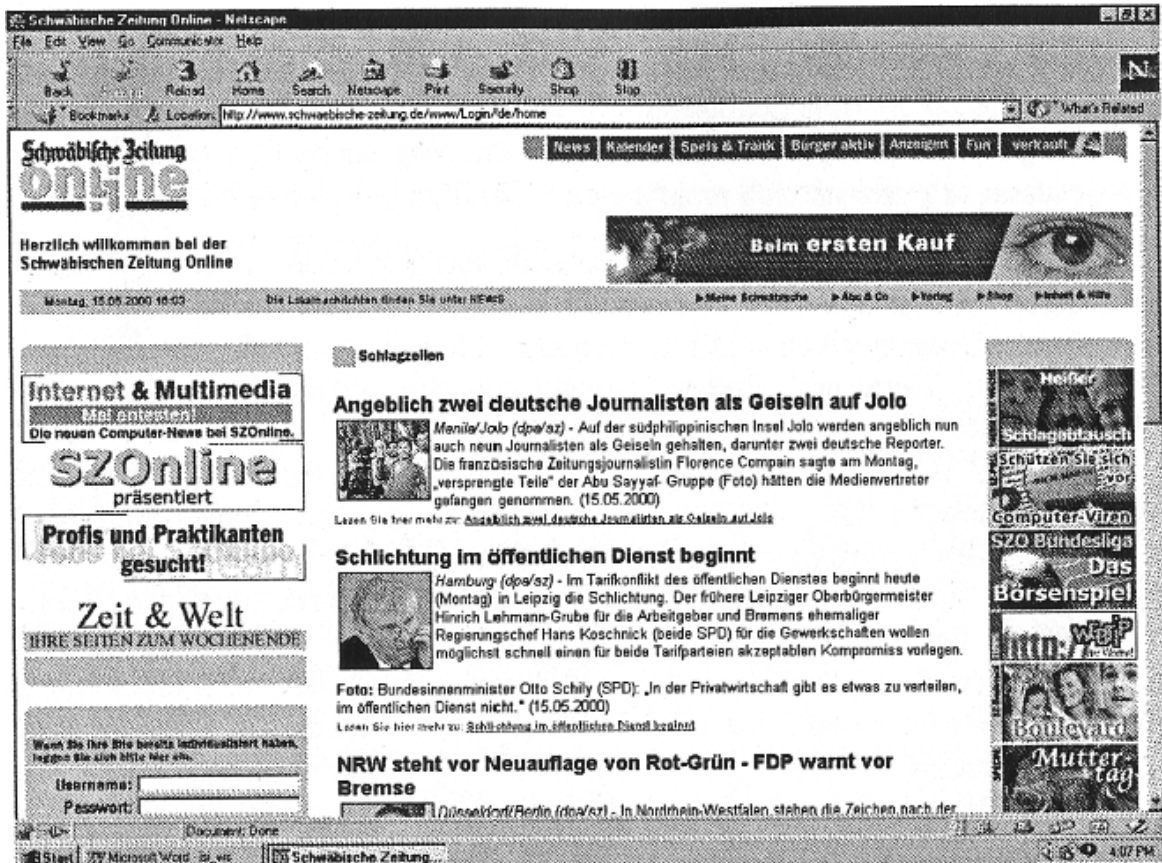


Abbildung 2: Startseite der Schwäbischen Zeitung

Zeitungsfremdes, aber durchaus onlineangemessenes findet man in Erweiterung der Print-Version in den Rubriken "Shop und Auktionen". Hinter „Shop“ verbirgt sich die Möglichkeit, online einkaufen zu gehen. Eine Übersichtsseite für die Produkte „Parfüms, Weine, Möbel und Alessi-Küchengeräte“, die von dem Online-Händler '4c The E-Commerce Services AG' in München<sup>13</sup> stammt, wird in das Layout der Web-Site integriert. Klickt man auf eine der dargestellten Produktgruppen, so öffnet sich ein neues Browserfenster unter der Webadresse des Händlers. An dieser Stelle hätte man eher erwartet, dass ein elektronischer Marktplatz mit Produkten und Dienstleistungen der regionalen Anbieter vorhanden bzw. angebunden ist. Gleich verhält es sich mit der Rubrik „Auktionen“, in der ebenfalls

<sup>12</sup> Für die Region „Friedrichshafen“ wurden am 15.05.2000 keine Nachrichten angezeigt

<sup>13</sup> <http://www.4-cee.de>

nur auf die Auktionsplattform der Firma "Itrade"<sup>14</sup> verlinkt wird, wünschenswert wäre auch hier eine eigene regionale Auktionsplattform.

**Präsentation:** Der Erst-Nutzer wird durch die unangemessenen grafischen Spielereien der Werbebuttons sowie der unterschiedlichen grafischen Navigationsleiste vom Wesentlichen, also den Nachrichten, stark abgelenkt. Die optische Präsentation ist dadurch insgesamt unruhig und unausgewogen. Die Lesbarkeit, die Aufbereitung sowie der Schreibstil der einzelnen Artikel entsprechen, da es sich hauptsächlich um dpa-Artikel handelt, dem Niveau einer Tageszeitung. Eigene sz-Berichte, die im „Regionalteil“ vorhanden sind, entsprechen augenscheinlich einer guten redaktionellen Qualität.

**Interaktion/Kommunikation:** Eine Interaktions-Funktion ist in der Rubrik „Anzeigen“ vorhanden. Hier ist es möglich, mit Hilfe einer Suchmaske in den einzelnen Unterrubriken „KFZ, Sport, Elektro, Möbel, Immobilien, Stellenmarkt, Tiere und Reisen“ gezielt zu den einzelnen Anzeigen zu gelangen. Die Anzahl der vorhandenen Anzeigen z. B. im Bereich „Immobilien“ ist jedoch so gering<sup>15</sup>, dass die Benutzung eines Suchformulars nicht angebracht ist. Der herkömmliche Weg durch anklicken der entsprechenden Navigationsbuttons führt hier sehr viel schneller zum Ziel. Weiter bietet die Rubrik „Anzeigen“ auch die Möglichkeit, Anzeigen online aufzugeben. Dabei wird unterschieden in Anzeigen für die Online-Version und Anzeigen für die Print-Version. Beidemale erfolgt die Anzeigenaufgabe mit Hilfe eines Formulars. Die Online-Anzeigen sind kostenlos, während man in dem Formular für die Print-Anzeigen die eigene Kontonummer und die Bankleitzahl zusammen mit anderen persönlichen Daten eingeben muss. Eine weitere Interaktionsmöglichkeit ist unter der Rubrik „Fun“ vorhanden. Hierunter befindet sich der „Chat der Schwäbischen Zeitung Online“. In diesem Chat halten sich die üblichen Chatter auf, leider hat keines der dort besprochenen Themen etwas mit der Schwäbischen Zeitung bzw. deren Nachrichten zu tun. Ein Diskussionsforum, in dem man über bestimmte Themen diskutieren könnte, wäre an dieser Stelle sinnvoller. Die mehrfach angesprochenen Navigationshilfen bieten dem Erst- bzw. mit der Navigation in dieser Web-Site unerfahrenen Nutzer keine Hilfe. Im Gegenteil, man verbringt erst einmal einige Zeit damit, sich mit dieser Menüsteuerung vertraut zu machen.

**Sozial/Organisation:** Von Datenschutz, Datensicherheit oder gar einem Vertrauensmanagement ist auf dieser Web-Site nichts zu erkennen. Die persönlichen sowie die Angaben zur Konto- und Bankverbindung, werden unverschlüsselt und ohne die Benutzung eines Sicherheitsservers über das Internet übertragen. Ob eine Überprüfung der angegebenen Angaben erfolgt, wird nicht veröffentlicht. Dem Missbrauch beim Aufgeben der Anzeigen für die Print-Version ist auch auf dieser Web-Site Tür und Tor geöffnet.

**Methodik/Technik:** Die Web-Site hat sich während der Testphase als sehr robust gezeigt, es kam zu keinem Absturz der HTML-Seiten auf dem Client

---

<sup>14</sup> <http://www.itrade.de>

<sup>15</sup> Es sind gerade mal sechs Anzeigen vorhanden

und zu keinem Absturz des Servers. Die Zuverlässigkeit der eingesetzten Technik kann somit nur gelobt werden. Leider lassen sich auf der Basis des HTML-Codes keine Rückschlüsse auf die verwendete Datenbanktechnologie machen. Auffällig ist nur, dass der ein oder andere Artikel mehrmals auf der gleichen Bildschirmseite erscheint, was auf eine Fehlfunktion beim Auslesen der angehängten Datenbank oder auf einen Fehler im Script schließen lässt.

## **7. Fazit**

Zeitungen sind komplexe Informationsprodukte, die durch eine große Anzahl Kriterien geprägt sind. Mit eines der wichtigsten Qualitätsmerkmale ist sicherlich die Qualität der journalistischen Arbeit von Zeitungen. Daneben profilieren sich Regional- und Lokalzeitungen auch über die Region, über die sie berichten. Ziel dieser Untersuchung war es, ein Verfahren aufzubauen, welches die Bewertung von Zeitungen zulässt und dabei sein Hauptaugenmerk weniger auf die journalistische Qualität richtet, sondern auf den Auftritt von Regional- und Lokalzeitungen im Internet, speziell dem World Wide Web. Anhand eines Kriterienkatalogs wurden die Schwäbische Zeitung und der Südkurier untersucht. Die Unterteilung in fünf Bewertungsbereiche hat sich dabei als sinnvoll und nachvollziehbar erwiesen. Der Aufbau eines hierarchischen Kennzahlensystems ist gut einsetzbar und sehr flexibel handhabbar. Ein Kennzahlensystem bietet neben der Chance, eine objektive Bewertung durchzuführen, auch die Möglichkeit, individuelle Sichten auf den Bewertungsgegenstand zu generieren. Dazu können die einzelnen Hauptkennzahlen und deren Kriterien individuell gewichtet werden, so dass bspw. ein anzeigenorientierter Nutzer der Zeitungen die sozial/organisatorische Ebene höher bewerten kann. Im nächsten Schritt sollen anhand der gewonnenen Kriterien mithilfe eines in Java entwickelten Bewertungstools auf Basis des Kennzahlensystems umfangreiche Benutzertests durchgeführt werden.

## **8. Literatur**

Bramme, A., Kotzmann, E. und Reschenberg, H. (1993). Publizistische Qualität: Probleme und Perspektiven ihrer Bewertung. München: Profil

Bruhn, M. (1997). Qualitätsmanagement für Dienstleistungen. Grundlagen, Konzepte, Methoden. Springer: Heidelberg

Flöpper, B. L. (1995). Das lokale Prinzip schafft neue Ordnung in der Zeitung. In: Zeitung der Zukunft. Zukunft der Zeitung. Flöpper, B. L. & Raue, P.-J. (eds.); Bonn: ZV Zeitungsverlag-Service, Kap. 12, 127-134

Haller, S. (1995). Beurteilung von Dienstleistungsqualität: dynamische Betrachtung des Qualitätsurteils im Weiterbildungsbereich. Wiesbaden: Deutscher Universitäts-Verlag

Klettke, S., Link, P., Remberg, S., und Wöbking, M. (1998). Der digitale Zeitungskiosk. Eine Typologie von Online-Tageszeitungen. In: Das Netz-Medium:

kommunikationswissenschaftliche Aspekte eines Mediums in Entwicklung. Neverla, I., Schäfers, A., und Siegle, D. (eds.); Opladen: Westdeutscher Verlag, 263-276

Kuhlen, R. (1998). Organisationsformen und Mehrwertleistungen elektronischer Märkte. In: Linguistik und neue Medien. GLDV-Jahrestagung vom 17. - 19. März 1997. Heyer, G. & Wolff, C. (eds.); Deutscher Universitäts-Verlag: Wiesbaden

Lehr, T. (1999). Tageszeitungen und Online-Medien: elektronisches Publizieren als produktpolitisches Instrument der Verlage. Wiesbaden: Deutscher Universitäts-Verlag

Naumann, A. (1995). Die Zukunft der Zeitung wird Kompromisse brauchen. In: Zeitung der Zukunft. Zukunft der Zeitung. Flöpper, B. L. & Raue, P.-J. (eds.); Bonn: ZV Zeitungsverlag-Service, Kap. 14, 153-166

Neuberger, C. (1999). Vom Papier auf den Bildschirm. Die Zeitung in der Metamorphose. In: Online - Die Zukunft der Zeitung? Das Engagement deutscher Tageszeitungen im Internet. Neuberger, C. & Tonnemacher, J. (eds.); Opladen: Westdeutscher Verlag, 39-46

Neuberger, C. & Tonnemacher, J. (1999). Online - Die Zukunft der Zeitung? Das Engagement deutscher Tageszeitungen im Internet. Opladen: Westdeutscher Verlag

Neuberger, C., Tonnemacher, J., Bietl, M., und Duck, A. (1997). Die deutschen Tageszeitungen im World Wide Web: Redaktionen, Nutzer, Angebote. Media Perspektiven Nr.12, 652-662

o.V. (2000b). Übersicht: Schaubilder 1999. Zur wirtschaftlichen Entwicklung. Lage der deutschen Zeitungen. Bundesverband Deutscher Zeitungsverleger, Datum: 2000b, <http://www.bdzv.de/presse/m/marktdaten/zeitungen99/pics.htm>

o.V. (5.4.2000c). Website Trend. Tageszeitungen. Horizont.Net, Nr. 17, Datum: 5.4.2000c, [http://www.horizont.de/interactive/websitetrend/trend17\\_00/index.html](http://www.horizont.de/interactive/websitetrend/trend17_00/index.html)

o.V. (26.4.2000a). Zeitungshäuser sehen Boom durchs Internet. Die Welt, Rubrik Webwelt, Datum: 26.4.2000a, <http://www.welt.de/daten/2000/04/26/0426w1164443.htx>

Polatschek, K. (1996). Tod der elektronischen Zeitung! Die ZEIT Nr.30

Rada, H. (1999). Von der Druckerpresse zum Web-Server: Zeitungen und Magazine im Internet. Berlin: Wissenschaftlicher Verlag

Riefler, K. (1995a). Zeitung online: Neue Wege zu Lesern und Anzeigenkunden. Bonn: ZV, Zeitungsverlags Service

Riefler, K. (1995b). Zurück aus der Zukunft. Eine Reise durch amerikanische Zeitungshäuser. In: Zeitung der Zukunft. Zukunft der Zeitung. Flöpper, B. L. &

Raue, P.-J. (eds.); Bonn: ZV Zeitungsverlag-Service, Kap. 13, 135-152

Rittberger, M. (1999). Certification of information services. In: IQ 1999. Proceedings of the Conference on Information Quality. Lee, Y. W. & Tayi, G. K. (eds.); Boston: MIT, 17-37

Rittberger, M. (2000). Quality measuring with respect to electronic information markets and particularly online databases. Kent, A. (ed.); Marcel Dekker: New York, NY, Vol. 31, Nr.68

Ruß-Mohl, S. (1994). Der I-Faktor: Qualitätssicherung im amerikanischen Journalismus. Modell für Europa? Zürich: Edition Interfrom

Ruß-Mohl, S. (2000a). Hollywood ist überall. Nach dem Fall Kummer: Warum sich der Journalismus so schwer tut, die Qualität seiner Produkte zu sichern. Die ZEIT Nr.23, 42

Ruß-Mohl, S. (2000b). Und was lernen wir daraus? Welche Strategien gewählt werden müssen, um journalistische Qualität in Deutschland besser zu sichern als bisher. Die ZEIT Nr.24, 39

Saxer, U. & Kuli, H. (1981). Publizistische Qualität und journalistische Ausbildung. Zürich: Publizistisches Seminar der Universität Zürich

Schönbach, K. (1997). Zeitungen in den Neunzigern: Faktoren ihres Erfolgs: 350 Tageszeitungen auf dem Prüfstand. Bonn: ZV Zeitungs-Verlag Service

Schönbach, K. & Peiser, W. (1997). Was wird aus dem Zeitungslesen? In: Zeitungen in den Neunzigern: Faktoren ihres Erfolgs: 350 Tageszeitungen auf dem Prüfstand. Schönbach, K. (ed.); Bonn: ZV Zeitungs-Verlag Service, 9-21

Stürzebecher, D. (1997). Was macht Tageszeitungen erfolgreich? Bertelsmann-Briefe Nr.138, 45-47

Turner, G. (1995). A quantitative approach to quality in Australian newspapers. Gazette, 55 Nr.2, 131-144

Wallisch, G. (1995). Journalistische Qualität. Definitionen - Modelle - Kritik. Konstanz: UVK Medien/Ölschläger

Wolff, V. (1998). Wettbewerb und Qualität bei Presseprodukten. Publizistik: Vierteljahreshefte für Kommunikationsforschung, 43 Nr.3, 260-272



# **Der Autor im digitalen Medium — ein notwendiges Konstrukt?**

## **Eine systematische Analyse des Autorenbegriffes und seine Rekonstruktion oder Auflösung im Kontext neuer Präsentations- und Kommunikationsformen**

**Stephan Werner**

Informationswissenschaft

Universität Konstanz

Stephan.Werner@uni-konstanz.de

### **Zusammenfassung**

In der vorliegenden Arbeit wird dargestellt, inwieweit das Internet mit seinen Mehrwerteffekten, Vernetzung und Interaktion, das bisherige Verständnis vom Autorenbegriff verändert. Die im Spannungsfeld von Oralität und Schriftlichkeit entwickelte Einheit von Autor und Werk erhält in der ‚Gutenberg-Galaxis‘ soziale, kommerzielle und juristische Dimensionen. Das Internet erweitert den bisherigen Raum der Schrift- und Druckkultur in zwei Richtungen und stellt in seinen Interaktionsmöglichkeiten diese Einheit in Frage. Zum einen kommen durch Inter- und Hypertextualität neue Textformen auf, deren Integrität nicht mehr festzustellen ist. Zum anderen entstehen neue Kommunikationsformen, welche eine neue Schriftlichkeit in den bisherigen oral geprägten Raum bringen. Die Annullierung der Einheit von Autor und Text bedingt eine graduelle Auflösung des Autorenbegriffs, der nur mit erhöhtem technischen und administrativen Aufwand entgegengewirkt werden kann.

### **Abstract**

This work shows the extent to which the Internet, with its value-added effects, network and interaction capabilities, changes the conventional understanding of the term „author“. Historical consideration shows that the author was firstly called upon to classify and authenticate texts before, in the age of print technology, receiving authority and social relevance. The unit of author and work, which was developed from the area of contention between speech and the written form, receives social, commercial and legal dimensions in the „Gutenberg-Galaxis“. The Internet expands the previous scope of both author and print culture in two directions, and by way of its interaction capabilities, calls this very unit into question. On the one hand, through intertextuality and hypertextuality, new forms



of text are appearing whose integrity can no longer be established. On the other hand, new forms of communication are bringing a new written form into an area previously characterised by oral tradition. The annulment of the unit author and work causes the gradual break down of the term author, which can only be counteracted upon with increased technical and administrative effort/expense.

## Einführung in das Thema

Das Buch, als Träger und Mittler von Wissen und Information, hatte schon durch das Aufkommen der elektronischen Medien seine Dominanz eingebüßt. Nun hat es in seiner eigenen Domäne, der Verschriftlichung von Wissen und Information Konkurrenz durch das digitale Medium - Internet - bekommen. Dieses verfügt durch seine vernetzte Struktur und Interaktionsmöglichkeiten über bedeutende Mehrwerte, welche eine Verdrängung des Buches als Wissensträger bewirken könnten. Nicht nur das Ende der Buchkultur, die McLuhan bildhaft als Gutenberg-Galaxis [McLuhan 68] bezeichnet, sondern das Ende des Buches als Kulturobjekt steht zur Diskussion. In der modernen Sichtweise ist die Vorstellung ausgeprägt, daß ein Buch oder ein Text durch einen schöpferischen Akt einer Person entsteht. Mit der Vorstellung, daß das Buch in naher oder ferner Zukunft keine gesellschaftliche Rolle mehr spielt, erhebt sich die Frage, ob dann auch der Autorenbegriff, wie wir ihn in unserer Kultur jetzt verstehen, noch die gleiche Ausprägung besitzt.

Die gemeinsame Basis der Schriftlichkeit zwischen Print- und digitalen Medien läßt nur in begrenztem Maße eine Analogie entstehen. Die Übertragung einer Buchmetapher in das Internet ist nur zum Teil möglich, da im Internet selbst neue Formen der Schriftlichkeit entstanden sind. Ebenso scheint auch der Autorenbegriff, der seit seinem Entstehen im antiken Griechenland eng an den Textbegriff geknüpft ist, nicht so einfach in das neue Medium zu transformieren zu sein. Das Ziel dieses Beitrages ist es nun, den aktuellen Autorenbegriff, wie er auf der Basis der Printmedien zur Diskussion steht, darzustellen. Es gilt dann aufzuzeigen, wo und wie seine Grenzen durch das digitale Medium verändert wurden und damit die Frage zu beantworten, ob eine Auflösung des bisherigen Begriffes unumgänglich oder eine Neukonzeptionierung notwendig und möglich ist.

## Autorenschaft in der Gutenberg-Galaxis

Der Begriff Autor wird auf den altrömischen Begriff *auctor oris* zurückgeführt, welcher mit „Förderer, d.h., der, der etwas noch nicht Vorhandenes unmittelbar oder mittelbar ins Dasein fördert oder schon Vorhandenes zur vollen Kraft bringt, ihm Bestand, Gedeihen, Anerkennung und Dauer leiht“, zu übersetzen ist [Georges 83,703]. Erst im 16. Jahrhundert bildeten sich mit Verbreitung der Drucktechnologie die Grundelemente von typographischen Netzen, wie Autoren, Verleger, Druckereien, Buchhandel und Käufer, heraus und deren Strukturen, wie sie noch heute Bestand haben [Giesecke 91,400]. Die Notwendigkeit, eine Autorenschaft festzuhalten, lag nicht zuletzt daran, daß das typographische Netz über eine Vielzahl von Anschlußstellen verfügte. Es wurde deshalb notwendig,



Information zu adressieren und damit ihren Ursprung festzuhalten. Um Adressierung innerhalb des Mediums zu erreichen, bot sich das Mittel einer strukturierten Zuweisung durch Paginierung. Ebenso wurde durch das Festhalten eines Autors die personelle Zuweisung bewerkstelligt.

Zu Beginn der digitalen Revolution steckte der Autorenbegriff in einer Krise. Erst in den 30er Jahren des 20. Jahrhunderts wurde der Begriff ‚Autor‘ seit mehr als 200 Jahren wiederverwendet [Schwenger 81,93]. Mit der Unterordnung und Eingliederung der Autorentätigkeit in die organisatorische und technische Maschinerie der Massenmedien, veränderte diese auch die Produktionsweise dieser Autoren. Im Herstellungsprozeß eines Werkes wird mit der zunehmende Zahl beteiligter Personen, zum Beispiel durch angestellte Autoren (Redakteure, Lektoren), die originäre Leistung eines Autors zurückgedrängt [Schwenger 81,95]. Durch die Einflußnahme, zum Beispiel durch inhaltliche Festlegung, werden die vorherigen Mittler zu Miturhebern. Das Schwinden der schriftstellerischen Originalität durch die Vergesellschaftung der Literaturproduktion führt letztendlich zur Hinterfragung des Autorenbegriffs bzw. dessen möglicher Auflösung.

Von den Vertretern des Strukturalismus wird die Auffassung, daß ein Autor durch seine Autorität auch bestimmt, wie sein Text oder Werk zu lesen sei, bezweifelt oder vollends negiert. Ein Text ist nicht in bezug auf seinen Autor, sondern in bezug auf die ihn umgebenden Strukturen zu interpretieren. Eine inhaltliche Bedeutung oder Interpretation eines Textes liegt in der Macht des Lesers, der aufgrund seiner individuellen Lebens- und Leseerfahrungen unterschiedliche Assoziationen zum Inhalt herstellt. Mit der Geburt des Lesers, so Barthes, muß notwendigerweise der Tod des Autors zugelassen werden [Barthes 77,148]. Nach Foucault definiert sich ein Autor nicht durch die Einheit mit einem Text, sondern als eine Funktion, welche Wirkungen beim Rezipienten hervorruft. [Foucault 74,21] Eine Person ‚Autor‘ ist dabei nicht notwendigerweise existent, ihr Verschwinden kann deshalb in Kauf genommen werden. Nach Eberle verschwindet nicht der Autor, sondern nur seine Autorität, d.h. die ihm zugeschriebene Originalität oder das ihm zugeschriebene Schöpferum von Neuem [Eberle 95,74]. Die Autorität, d.h. die zugewiesene Originalität kann nur durch die Rezipienten ausgesprochen werden. Es stellt sich die Frage, ob diese Krise durch die eingangs formulierte Behauptung, dass mit dem Internet eine Verdrängung der Buchkultur einhergeht, verstärkt oder gar überflüssig wird.

## **Der digitale Raum**

Die neuen medialen Formen schaffen neue Realitäten und erfüllen somit McLuhans These, daß das Medium die Botschaft ist [McLuhan 92,17]. Dies bedeutet aber auch, daß nicht der Inhalt per se wichtig ist, sondern dessen Übermittlung. Deshalb haben, so Jones, „nur wenige Interesse an dem, was wir meinen, als daran, wie wir übermitteln, was wir sagen und meinen“ [Jones 97,133]. Um das Neue zu beschreiben, werden alte und bewährte Verfahren und Begriffe angewendet, mit denen das Unbekannte beschrieben wird und sich demzufolge auch vergleichen lassen muß. So wird das Internet als neues Medium erst einmal danach bewertet, wie es die Funktionen der vorhandenen Medien und ihr

Aussehen nachbildet.

Die gutenbergsche Erfindung mußte sich Mitte des 16. Jahrhunderts dadurch beweisen, daß sie die Schreibprodukte möglichst adäquat nachbilden konnte. In ähnlicher Weise läßt sich dieses auch bei der elektrischen Schreibtechnik feststellen. Sie wird in der Regel dazu verwendet, Texte, welche eigentlich für die Printmedien konzipiert wurden, in das neue Medium zu transformieren. Zugeständnisse an die Mehrwerte des neuen Mediums werden insofern gemacht, da diese Texte mit anderen Elementen verlinkt werden. Durch die starke Analogie zum Printmedium dieser Form von Texten ist die Anforderung recht hoch, die Besonderheiten dieses Mediums, wie Beständigkeit, Unveränderlichkeit und Sicherheit, auch auf die digitalisierte Version zu übertragen [Klostermann 98,99].

Während in der Druckkultur mit der Annahme eines Manuskriptes durch einen Herausgeber oder Verleger die Voraussetzung für eine Publikation gegeben war, kann im digitalen Medium durch den einfachen Zugang die Veröffentlichung unabhängig von einer verlegerischen Tätigkeit geschehen.

Die etablierten Schritte des Zugänglichmachens, der Begutachtung und Kritik, der Entscheidung über Annahme oder Ablehnung sowie der dauerhaften Fixierung eines Textes machen gerade diese Tätigkeit des Verlegers aus. Dieser stellt damit eine Autorität, welche die Autorenschaft glaubwürdig gestaltet. Durch Veröffentlichungen im Netz, jenseits von verlegerischen Kriterien, werden nicht nur der Verleger in seiner Tätigkeit, sondern auch der Wert einer Autorenschaft hinterfragt. Dieses hat dann Konsequenzen, wenn mit der Autorenschaft auch ein sozialer Status verknüpft ist, wie es bei wissenschaftlichen Publikationen der Fall ist. Mit der Möglichkeit der Internettechnologie, Texte sofort einer Öffentlichkeit zugänglich zu machen, hat jeder Autor die Gelegenheit einer schnellen Veröffentlichung. Mit der Darbietung seiner Arbeit auf einer eigenen Webseite oder denen seiner Institution ist dem Veröffentlichenden keine Garantie gegeben, daß sie von einer gewünschten Öffentlichkeit registriert wird. Der Verfasser ist gezwungen, innerhalb einer unbestimmten Masse von Informationen seine relevante Öffentlichkeit anzusprechen, d.h. deren Aufmerksamkeit zu erregen.

Das eigentliche Problem liegt dabei in den Selektionsmöglichkeiten der Rezipienten, welche Rötzer, den „Flaschenhals der Aufmerksamkeit“ [Rötzer 98,63] nennt. Die Wahrscheinlichkeit, Aufmerksamkeit im Netz zu finden, ist somit auch für den allein agierenden Wissenschaftler gering. Die gedruckten Zeitschriften konnten die Aufmerksamkeit durch ihr Prestige gewährleisten, welches sie durch die qualitative Arbeit ihrer Herausgeber erreicht hatten. Dieses Prestige der einzelnen Zeitschriften etabliert eine Qualitätshierarchie, welche gleichsam als Informationsfilter diente. Zusätzlich gewährte, nach Meinung von Zimmer, die Schwerfälligkeit der Print-Journale den Urkundencharakter der Aufsätze. Das Internet bietet mit seinen Distributionsmöglichkeiten eine Chance, die Zeitschriftenversorgung im wissenschaftlichen Bereich zu verbessern und zu erneuern [Jäger 98,57]. Die Integrität der Dokumente wird zum einem dadurch gewährleistet, daß sie als Printversion ebenfalls vorliegen und zum andern, daß sie meistens nur mit Überschrift und ggf. Abstract auf den Webseiten des Verlages

erscheinen. Die Texte werden dann auf Anforderung und nach Entrichtung eines Preises zum Herunterladen zugänglich gemacht. Um die Authentizität beim Transport zu gewähren, werden meistens pdf- und zip-Formate verwendet. Durch seine Autorität garantiert der Verlag die Integrität der Dokumente sowohl in technischer als auch inhaltlicher Sicht.

Hier wird deutlich, daß die Analogie in der Endfassung möglichst nahe an einer Printversion liegen soll. Die eigentlichen Mehrwerte des Internets, wie Vernetzung und Interaktion, werden nicht in die Texte integriert, wie Wegner feststellt [Wegner 99]. Eine Ursache hierfür liegt sicherlich in der sozialen Relevanz, die sich aus der Autorenschaft für den Schreibenden ergibt. Sie leitet sich im wesentlichen aus einer Integrität der Texte, die eine klare Trennung zwischen Text und Kontext aufweisen müssen, damit der Autor seine eigenständigen Fähigkeiten darlegen und nachweisen kann. Auch mag ein Grund für die Beibehaltung der bewährten Version darin liegen, daß die in der Tradition des wissenschaftlichen Argumentierens entstandenen linear und hierarchisch geordneten Aufsätze als intellektuell reiner und klarer empfunden werden [Bolter 97,47].

Durch die Vernetzung von Computern werden neue Kommunikationsmöglichkeiten erschlossen, welche vorher ansatzweise nur durch eine sekundäre Oralität zum Beispiel per Telefon erreicht wurden [Ong 82,18]. Durch die Verlinkung untereinander eröffnet sich ein Raum der telemediatisierten Textualität, welche „die bisherigen linearen Strukturen der buch(druck)orientierten Konfiguration der Literalität auflösen“ [Krämer 97,93]. Die bisherigen Modalitäten von oraler und schriftlicher Kommunikation werden neuen Kriterien unterworfen. Im Vergleich zur oralen Modalität der Kommunikation geschieht die Kommunikation im Internet häufig anonym und unpersönlich. Und im Unterschied zur schriftlichen Modalität von Kommunikation treten an die Stelle von stabilen, generalisierten Texten kontingenzbehaftete und damit singuläre Texte.

Die illokutionären, performativen Aspekte, also die moralischen, politischen und rechtlichen Verankerungen in unserer ‚gewöhnlichen‘ Kommunikation, werden gerade durch die prinzipielle Anonymität des digitalen Mediums nicht mehr angesprochen [Krämer 97,96]. Die Teilnehmer im Internet agieren nicht mehr als Personen sondern unter symbolhaften, frei wählbaren, selbstgeschaffenen Namen. Die personelle Zuweisung eines Textes ist durch die Möglichkeit, daß Identitäten künstlich konstruiert werden können, nicht mehr eindeutig. Die Einheit von Autor und Text implizierte auch den moralisch-rechtlichen Sachverhalt, daß Autoren auch für das verantwortlich gemacht werden können, was sie schreiben. Mit den technisch gegebenen Freiheitsgraden einer frei wählbaren Anonymität ist eine Autorenschaft nicht mehr ohne weiteres nachvollziehbar und nachweisbar. Im eigentlichen Sinne findet dort auch keine Intersubjektivität statt, da dieser Begriff, so Krämer, „auf eine Identität im Unterschied zielt, welche in einem idealisierten und technisch unverstellten wie unverzerrten Zwischenraum wechselseitiger Bezugnahme anwesender Subjekte entsteht“ [a.a.O.,99]. Statt einer Wechselwirkung zwischen Subjekten findet im Internet nur eine zwischen Texten, mit anderen Worten, eine Intertextualität, statt.

Als neues Medium wirkt das Internet auch im künstlerischen Bereich innovativ und läßt neue Formen von Literatur, welche sich mit den technischen neuen Möglichkeiten dieses Mediums beschäftigen, entstehen. Nach Idensen geht es darum, "Neuland im telematischen Raum zu vermessen, Textlandschaften anzulegen, Schreiben und Lesen als einen nomadischen Akt des Umherschweifens durch Text-Netzwerke zu begreifen" [Idensen 99]. Mit diesem programmatischen Anspruch wird schon deutlich, daß die Beziehungen zwischen Text, Autor und Leser in einer digitalen Form der Literatur neu zu bestimmen sind. Hierbei ist zwischen einer ‚Literatur im Netz‘ und einer ‚Netzliteratur‘ zu unterscheiden [Ortmann 99b]. Während die erste Definition die Transformation von gedruckten Texten ins Netz umschreibt, wird mit dem Begriff ‚Netzliteratur‘ der Anspruch erhoben, eine neue „Kunstform“ zu kreieren [Zimmer 99].

Betrachtet man die „Literatur im Netz“, so wird die Integrität der Texte auch im digitalen Medium gewahrt. Allenfalls die Möglichkeit der Vernetzung wird genutzt, um zwischen den Texten Verbindungen (Links) herzustellen und eine Intertextualität zu schaffen. Ähnlich wie der Gebrauch von Fußnoten gehen diese Links über die Abgeschlossenheit des eigentlichen Textes hinaus. Der Mehrwert im Netz entsteht jedoch dadurch, daß hier nicht ein aktueller und ein erinnernder Text aufeinander bezogen werden, sondern zwei gleichermaßen präsente Texte, die gegebenenfalls auf verschiedenen semantischen Ebenen existieren [Wirth 97,234]. In der Intertextualität bleibt jeder Text aber autonom und kann für sich alleine betrachtet und als Ganzes ausgedruckt werden. Bei näherer Betrachtung wird deutlich, daß die Literatur im Netz keinen hohen Qualitätsanforderungen standhält. Da jeder, der Literatur als eigenständige Schöpfung ins Netz stellt, sich als Autor bezeichnen kann, befürchtet Ortmann zurecht eine Inflation und damit Verwässerung dieses Begriffes [Ortmann 99a]. Allein durch die Möglichkeiten im Internet wird ein der Printkultur entsprechender Status ‚Autor‘, der im Sinne Foucault eine kohärenzstiftende Funktion ist, nicht erreicht.

Netzliteratur zeichnet sich gegenüber der oben beschriebenen ‚Literatur im Netz‘ dadurch aus, daß sie die komplette Palette der Mehrwerte des Internets integriert und sie sogar notwendigerweise voraussetzt. Der Schreibende entwickelt ein netzartiges Gefüge, welches aus einer Pluralität unterschiedlicher Pfade und Verweisungen besteht. Durch die Interaktion des Lesenden mit dieser verlinkten Struktur entstehen jeweils neue, individuelle Texte, welche der Autor nicht vorhersagen kann. Sein hypertextuelles Schreiben, d.h. die Einbindung beliebiger Verlinkungen in das Selbsterschaffene, bringen den Verfasser in unmittelbare Interaktion mit dem Denken und Schreiben anderer Menschen. So bilden, nach Ibensen, Text, Schreiber, Leser, Welt und Gesellschaft ein Rhizom, welches einen neuen Raum für textuelle, konversationelle und diskursive Austauschprozesse liefert, die nicht mehr nach dem Sender-Message-Empfänger ablaufen [Ibensen 99]. Eine autarke Einheit von Autor und Werk ist dabei nicht mehr gefragt.

Ebenso wird im geschlossenen Hypertext ein Dominanzverlust attestiert, der den von Barthes postulierten Tod des Autors Wirklichkeit werden läßt [Hautzinger 99,28]. Die von ihm vorgegebene Ordnung, die beim Buch in der Regel linear war, wird nun durch die Interaktionsmöglichkeiten des Lesers in Frage gestellt. Der

Leser entzieht sich der räumlichen Anordnung des Textes durch eine eigenständige Selektion. Gegenüber dem Autor wird somit die Rolle des Lesers gestärkt, er wird zum Mit-Autor. Durch die aktive Mitgestaltung verliert das ‚Werk‘ seine Form und realisiert sich erst im Akt des Lesens. Gleichzeitig wird gefordert, daß der Leser auch die Struktur hinter dem Hypertext erkennen muß. Hier zeigen sich die Grenzen der vorherigen Behauptungen, da es immer noch in der Macht des Autor liegt, Verlinkung her- und bereitzustellen. Der Leser bewegt sich auf vorgegebenen Bahnen, wobei er keine Möglichkeit hat, vorab die Relevanz oder Irrelevanz, welche der Autor geplant hat, festzustellen. Er kann nur die Art seiner Betrachtung wählen, entweder als „umherschweifender Daten-Dandy oder herumsuchender Daten-Detektiv“ [Wirth 97,326]. Beide Male ist er der Macht des Autors ausgeliefert, da die Textkohärenz von diesem festgelegt wird. Nach Hasecke hat Netzliteratur ohne Literatur, also schon geschriebene Prosa, keine Existenzberechtigung [Hasecke 99].

Mit der Einführung der Schriftlichkeit in einer von Oralität geprägten Welt hatte sich die Möglichkeit ergeben, Kommunikation asynchron und disloziert zu betreiben, womit diese aus der Interaktion der Teilnehmer gelöst wurde. Der Verlust der interaktiven Elemente konnte nur unzureichend kompensiert werden. Indem nun die Interaktionsmöglichkeiten des Internets zur Verfügung stehen, stellt sich die Frage, inwieweit eine Analogie zur oralen Kommunikation wiederhergestellt werden kann. Hierbei ist zu berücksichtigen, daß die Kommunikationsform weiterhin auf Schriftlichkeit basiert, nur eben im Kontext des digitalen Mediums, was die Einbindung von non-verbalen Elemente der mündlichen Kommunikation ausschließt. Bildlich gesprochen, werden aber mit der Möglichkeit, auf der Basis des Internets interaktiv zu kommunizieren, der bisherige Raum der Schriftlichkeit erweitert und Bereiche erschlossen, die zuvor nur der Oralität vorbehalten waren. Aus dieser Zwitterstellung der digitalen Schriftlichkeit ergeben sich Potentiale zu neuen Formen der Kommunikation. Da die Schriftlichkeit weiterhin die Grundlage dieser Kommunikationsformen ist, stehen deren Mehrwerte weiterhin zur Verfügung. Die Äußerungen können fixiert werden und damit archiviert und einer Person als Urheber zugewiesen werden. Dieses würde in der Konsequenz nichts anderes bedeuten, als daß eine Autorenschaft in einer interaktiven Situation, welche den Charakter einer oralen Kommunikation hat, festzustellen und zuzuweisen ist. Es gilt im folgenden festzustellen, unter welchen Bedingungen Kriterien existieren und geschaffen werden, welche innerhalb einer internetbasierten Kommunikation eine Autorenschaft definieren. Zu diesem Zweck werden wir im folgenden zwei Kommunikationsformen im Internet näher betrachten. Zum einen handelt es sich um den Chat, der am stärksten die orale Situation nachbildet und damit am weitesten von allen Kommunikationsformen des Internets in den Raum der Oralität eindringt. Zum anderen wenden wir uns den Kommunikationsforen zu.

## **Autorenschaft im Chat?**

Wie der Name es ausdrückt, findet im Chat ein Schwatz statt, der vom Charakter her zwischen Diskussion und sinnlosem Geplauder steht [Kuhlen 98,52]. Allgemein versteht man unter einem Schwatz eine Kommunikation einer

begrenzten Anzahl Personen, welche sich situationsgebunden und spontan austauschen, wobei dieser Austausch zeitlich begrenzt ist. Der Zugang zu einem elektronischen Chat findet üblicherweise durch Angabe eines Chat-Namens und eines Paßwortes statt. In der Regel kann eine Person einerseits beliebig viele Chat-Namen und Chat-Identitäten annehmen. Doch hat andererseits eine Chat-Existenz eine Eindeutigkeit, so daß eine Wiedererkennung möglich ist [Sturm/Wundrak 99]. Dieses bedeutet auch, daß einer Chat-Identität und die damit hinter ihr stehende Person für Äußerungen zur Rechenschaft gezogen werden kann.

Bei näherer Betrachtung wird deutlich, daß die Chat-Teilnahme auf Spontaneität ausgerichtet ist, welche nach Ong einen wesentlichen Aspekt der oralen Kommunikation ausmacht [Ong 82,136]. Gleichzeitig findet aufgrund der technischen Beschränkungen nur eine quasi synchrone Kommunikation statt, da es zwischen der ‚spontanen‘ Teilnehmeraktion und der Aktivität aller Teilnehmer im Chat zu Verzögerungen kommt. In den einzelnen Beiträgen werden bewusste Textkürzungen und –änderungen der geschriebenen Sprache verwendet, wobei der Teilnehmer sich an Lautanalogien in der gesprochenen Sprache orientiert. Allgemein ist deshalb das Chat-Verhalten eher in der Analogie zur Oralität zu sehen, in der ein Autor *per se* nicht zu definieren ist. Trotz der vorhandenen Schriftform werden die wesentliche Bedingungen, um eine Autorenschaft festzustellen, nicht erfüllt. Das entscheidende Kriterium ist hierbei nicht die Kürze der Texte, wie vergleichsweise z.B. die Aphorismen L. Wittgensteins belegen, sondern die Kürze der Aufbewahrungszeit. Durch die kurze Visualisierungsphase ist der User gezwungen, sich in der Fortführung des ‚Gespräches‘ schon nach kurzer Zeit auf sein Gedächtnis zu verlassen. Das vorhandene Potential der Schriftlichkeit durch Fixierung von Textinhalten wird somit in einer Chat-Kommunikation nicht genutzt. Ein Autor ist damit nicht definierbar.

## Der Autor im elektronischen Forum

In einem elektronischen Forum ist *per definitionem* die Verwahrung der Beiträge zur Kommentierung durch andere Teilnehmer vorgesehen. Diese Möglichkeit inhaltlicher Bezugnahme setzt voraus, daß die Beiträge unverändert bleiben, da sonst der Ansatzpunkt nicht mehr gegeben ist und der Verweis sinnlos wird. Nach Kuhlen ist die Kommunikationsform der Dialog, der „den wechselseitigen Austausch von Positionen verschiedener Kommunikationspartner [widerspiegelt], wobei seine Minimalform aus Rede und Gegenrede (z.B. Frage und Antwort) [besteht]“ [Kuhlen 98,59]. Während der Dialog die Form der Kommunikation ist, ist die Interaktion ihr Mittler. Um ein Forum oder seine Diskussionsfäden (Threads) aktiv zu halten, müssen nicht nur die Themen attraktiv sein, sondern die Beiträge selbst Aufmerksamkeit erregen. Die Beiträge sollten deshalb eine Subjektivität aufweisen, welche Reaktionen provozieren, d.h. sie müssen zum Widerspruch, zu Ergänzungen oder allgemein zu Kommentaren reizen. Erst durch die Interaktion zwischen den Teilnehmern erhält ein Forum Leben und bleibt präsent.

Mit den Begriffen ‚Dialog‘ und ‚Interaktion‘ werden prägnante Merkmale der mündlichen Kommunikation angesprochen. Die Analogie zwischen Oralität und elektronischen Foren kann nicht darüber hinwegtäuschen, daß im ersten Fall

Intersubjektivität vorliegt, während im zweiten nur eine Intertextualität herrscht. Im Gegensatz zum Chat ist die spontane Reaktion auch in den elektronischen Foren nicht üblich, da sie in der Regel asynchron gestaltet sind und über längere Zeiträume geführt werden. Dadurch bieten sie den Vorteil der Schriftlichkeit, indem sie Beiträge mit höherem Komplexitätsgrad ermöglichen. Folglich kann erwartet werden, daß die Herstellung der Beiträge je nach Forum auch eine längere Herstellungszeit benötigt und damit eine Voraussetzung schafft, daß die eigenständige Leistung der Schreiber dem Charakter nach ein schöpferischer Akt sein kann.

In der Kombination von mündlichen und schriftlichen Eigenschaften läßt sich der Autorenbegriff nicht so ohne weiters bestimmen. Einerseits kann bezüglich eines einzelnen Beitrages eine inhaltliche Autorenschaft dann definiert werden (formal juristisch steht sie hier nicht in Frage), und andererseits sind die Beiträge in einen Dialog eingebettet, d.h. sie erhalten erst im Kontext mit den anderen Beiträgen ihren eigentlichen Sinn. Versucht man die elektronischen Foren als Form der digitalen Schriftlichkeit zu fassen, so sind sie in ihren kommunikativen Anwendungsmöglichkeiten zwischen dem Chat und der elektronischen Publikation einzuordnen. Wobei das Chat in Analogie zum oralen Medium und die Publikation in Analogie zum Printmedium jeweils die extreme Position der digitalen Schriftlichkeit stellen. Im ersten Fall ist eine Autorenschaft nicht gegeben, während sie im zweiten Fall erwünscht ist. Eine Analyse vorhandener Forumstypen zeigt, daß ein Anspruch an eine Autorenschaft in einem Forum nur bewerkstelligt werden kann, wenn gleichzeitig ein hoher administrativer Aufwand betreiben wird, um einen Autor in einem Forum zu definieren.

## **Kontinuität und Diskontinuität**

Von einer Kontinuität des Autorenbegriff im Übergang zu den digitalen Medien kann nur dann die Rede sein, wenn die Wertschätzung des Autors und seiner Tätigkeit Gegenstand der gesellschaftlichen Urteile ist. Konkret lassen sich hier durchaus Ansätze erkennen. So ist mit der Übernahme der dekonstruktivistischen Position eine weiterhin fallende Wertschätzung zu attestieren. Auch werden die bisherigen Wertvorstellungen zum Autorenbegriff, welche sich in der Gutenberg-Galaxis entwickelt haben, nach wie vor bestehen bleiben, da das Printmedium (Buch) auf absehbare Zeit nicht durch das digitale Medium verdrängt werden wird. Alternativ würde bei einer Diskontinuität unterstellt, daß das digitale Medium derart stark in die bisherigen Wertestrukturen eingreift, daß man nicht mehr von einer Fortführung des bisherigen Autorenverständnisses reden könnte. Oder alternativ, daß es zwar einen neuen Autorenbegriff gibt, der jedoch keine oder nur geringe Gemeinsamkeiten mit dem bisherigen aufweist. Mit der Möglichkeit, die virtualisierbaren Elemente des kulturellen Gedächtnisses - also hier Texte, aber auch Bilder, Musik usw. - in das digitale Medium zu transformieren, können diese viel leichter erschließbar sein als durch die bisherigen Distributionsmöglichkeiten. Inwieweit dann das potentielle Interesse am Schöpfer der Werke noch bestehen bleibt, kann momentan nur tendenziös prognostiziert werden. Dabei zeigt sich, daß es durchaus auch für das digitale Medium relevante Kriterien gibt, aus denen ein Interesse am Autorenbegriff abzulesen ist. Diese basieren im wesentlichen auf

Abwägungen zwischen zwei Zuständen, wobei der eine stützend auf die Autorenschaft wirkte, während der andere sie eher in Frage stellt. Der Übergang ist meistens fließend, und nicht disjunkt. Im einzelnen lassen sich folgende Kriterien(-paare) ansprechen.

## Schriftlichkeit

Ein Autor kann in oralen Kulturen nicht definiert werden. Dennoch entstehen im Übergang von der Oralität zur Schriftlichkeit semi-orale Gesellschaften, bei denen die Schriftkunde entweder als elitäres Merkmal oder als untergeordnete handwerkliche Tätigkeit betrachtet wird. Im ersten Fall wird innerhalb der Elite der Autorenschaft auch ein gesellschaftlicher Wert beigemessen.

Oralität  Schriftlichkeit

## Dislozierung

Ein wesentliches Kriterium, welches einen Mehrwert der Schriftlichkeit gegenüber der Oralität ausmacht, ist die Möglichkeit, Informationen unabhängig von einer sonst notwendigen Intersubjektivität, die begrenzte lokale Räume bedingt, weiterzugeben. Die fehlende Intersubjektivität erfordert eine personelle Zuweisung des Textes, welche auf den Autor verweist und es ermöglicht, daß Informationen disloziert genutzt werden können. Durch die gegenseitige Abhängigkeit bedingen sich beide Kriterien gleichermaßen.

lokale Räume  Dislozierung

Intersubjektivität  personelle Zuweisung

## Archivierung

Aus einer Autorenschaft lassen sich Ansprüche auf kommerzielle und soziale Interessen ableiten, welche durch gesellschaftlich-etablierte Rechte gesichert sind. Um einen Anspruch stellen zu können, muß ein Autor einen Nachweis auf seine Autorenschaft erbringen. Zu diesem Zweck muß die Vergänglichkeit der Produkte möglichst lange durch eine Archivierung des Textes ausgeschlossen werden.

Vergänglichkeit  Archivierung

## Identifikation

Ebenso wie das Werk muß auch der Autor als handelndes Subjekt erkennbar und nachweisbar sein. Dieses ist durch eine Anonymität nicht zu erreichen. Die Identität des Autor sollte hierbei gewährleistet sein, wobei es vom



Anspruch der Rezipienten abhängt, welche Informationen über den Autor hinreichend sind.

Anonymität  Identifikation

## **soziale Relevanz**

Eine Autorenschaft hat immer eine soziale Relevanz, da sie den Autor in seiner Tätigkeit in einen sozialen Kontext rückt. Die sozio-kulturellen Anforderungen an eine Autorenschaft können hierbei eine unterschiedliche Gewichtung haben. Fehlende soziale Relevanz ist dabei gleichbedeutend mit Anonymität

 Soziale Relevanz

## **Integrität**

Ein Dokument, auf das sich eine Autorenschaft bezieht, muß definierbar und von seinem Kontext abgrenzbar sein. Mit zunehmender Interaktion verschwimmt diese Trennungslinie, so daß eine Autorenschaft innerhalb eines heterogenen Interaktionssystems nicht mehr belegt werden kann.

Interaktion  Integrität

## **Authentizität**

Die Authentizität der Dokumente ist ein Korollar zur Integrität. Der Rezipient benötigt die Garantie, daß die Integrität eines Dokumentes auf dem Weg vom Autor zu ihm auch erhalten bleibt, und er die vom Autor erstellte Information unversehrt erhält. Diese Authentizität sinkt mit zunehmender Bedeutungslosigkeit des Dokumentes, womit auch das Interesse am Autor abnimmt.

Bedeutungslosigkeit  Authentizität

## **Öffentlichkeit**

Ein Autorenschaft definiert sich in der Öffentlichkeit. Um Öffentlichkeit herzustellen, bedarf es Selektionskriterien, die eine kultivierte Aufmerksamkeit in einer steigenden Informationsflut gewährleisten. Vorselektionen werden in der Regel durch gesellschaftlich autorisierte und institutionalisierte Zwischeninstanzen wie Verlage durchgeführt. Sie gewähren dadurch eine Norm, welche durch die reine private Tätigkeit nicht geleistet werden kann.



Mit der obigen Zusammenstellung dieser Kriterien besteht nicht die Absicht, Autorenschaft oder den Begriff des Autors neu zu definieren. Schon die dynamischen Übergänge der einzelnen Punkte machen deutlich, daß nicht alle Eigenschaften im rechten Bereich vollständig erfüllt werden müssen. Mit diesen Kriterien soll deshalb nur die Notwendigkeit von Autorenschaft bzw. der Definition eines Autors überprüft werden.

So stellt das Internet *per definitionem* ein verteiltes System dar, so daß das Kriterium der Dislozierung theoretisch erfüllt ist. Jedoch wird gerade im Chat die Simulationen eines Raumes und einer intersubjektiven Handlung in Analogie zu einer mündlichen Situation erzeugt. Damit ist dieses Kriterium weitgehend nicht erfüllt, obwohl hier eigentlich eine Dislozierung vorliegt und auch keine Intersubjektivität, sondern eine Intertextualität. Ebenso ist die Veröffentlichung eines Textes auf einer Webseite meines Erachtens in erster Linie ein privater Akt, da der Handelnde an keine Normen gebunden ist. Er kann seine Seite und die dort veröffentlichten Texte jederzeit ändern, ohne daß er dafür gegen gesellschaftliche Abmachungen verstößt. Aus diesem Grund ist die Öffentlichkeit im Sinne einer normgebenden Instanz eigentlich nicht angesprochen, folglich ist hier auch kein Anspruch einer Autorenschaft abzuleiten.

Der springende Punkt in der gesamten Diskussion ergibt sich aber nicht direkt aus diesen Kriterien. Es ist die notwendige Einheit von Autor und Text, welche sowohl den Autor, sein Werk und die Autorenschaft definiert. Im Internet ist diese Einheit nicht mehr gegeben, sowohl Texte als auch Personen sind eigenständige Entitäten. Wie in jedem abstrakt konstruierten System verlangt schon die Logik die notwendige Wohldefiniiertheit der Elemente, da sonst keine eindeutige Zuordnung oder Verbindung hergestellt werden kann. Der virtuelle Raum des Internets ist ebenfalls ein abstraktes System, obwohl dies gerne vergessen wird. Um eine Verlinkung zwischen Elementen herstellen zu können, muß es auch klar definierte Elemente geben, wie es hier mit den Forderung nach Integrität und Identität ausgesprochen wird.

Die Autorenschaft muß im Internet zunächst einmal auch durch eine Verlinkung zwischen Autor und Text hergestellt werden, da sie keine implizit gegebene Einrichtung mehr ist. Um dieses zu erreichen, ist ein technischer und ein administrativer Aufwand notwendig. So kann die Integrität der Texte durch die technisch aufwendige Herstellung einer digitalen Signatur erreicht werden. Werden auch die Autoren im Netz dargestellt, so erfordert dieses einen hohen verwaltungstechnischen Aufwand durch eine Userverwaltung, wie es die bei dem elektronischen Diskussionsforum praktiziert wird. Die obige Kriterien können nun dahingehend benutzt werden, daß vorab geklärt werden kann, ob die Notwendigkeit besteht, eine Autorenschaft zu präsentieren. Je nach Anforderung entscheidet sich dann, welcher Aufwand letztendlich für die Konstruktion eines Autors aufgewendet werden muß.

## **Literatur**

[Barthes 77] Barthes, R., The Death of the Author, IN: Heath, S. (Hrsg.) Image-Music-Text, London 1977, S.142-148.

[Bolter 97] Bolter, J.D., Das Internet in der Geschichte der Technologien des Schreibens, IN: Münker, S. u. Roesler, A. (Hrsg.), Mythos Internet, Frankfurt 1997, S. 37-55.

[Foucault 74] Foucault, M., Was ist ein Autor?, IN ders., Schriften zur Literatur, übers. von Hofer, v. K., München 1974, S.7-31.

[Georges 83] Georges, K.E., Ausführliches lateinisch-deutsches Handwörterbuch, 8.Auflage, Darmstadt 1983.

[Giesecke 91] Giesecke, M., Der Buchdruck in der frühen Neuzeit, Eine historische Fallstudie, Frankfurt 1991.

[Hasecke 99] Hasecke, J.U., Was ist Netzliteratur?, ONLINE: [http://www.koeln.netsurf.de/~JanUlrich.Hasecke/Thesen\\_juh.html](http://www.koeln.netsurf.de/~JanUlrich.Hasecke/Thesen_juh.html)

[Hautzinger 99] Hautzinger, N., Vom Buch zum Internet?, Eine Analyse der Auswirkungen hypertextueller Strukturen auf Text und Literatur, Mannheimer Studien zur Literatur- und Kulturwissenschaft, Bd.18, St. Ingbert 1999.

[Idensen 99] Idensen, H., Die Poesie soll von allen gemacht werden! Von literarischen Hypertexten zu virtuellen Schreibräumen der Netzwerkkultur, ONLINE: <http://www.uni-kassel.de/interfiction/projekte/pp/poesie.htm>

[Jäger 98] Jäger, W., Vom Text der Wissenschaft, IN: Jochum, U., Wagner, G. (Hrsg.) Am Ende das Buch, Semiotische und soziale Aspekte des Internets, Konstanz 1998, S. 55-82.

[Jones 97] Jones, S., Kommunikation, das Internet und Elektromagentismus, IN: Münker, S. u. Roesler, A. (Hrsg.), Mythos Internet, Frankfurt 1997, S.131-146.

[Klostermann 98] Klostermann, V., Text und Hypertext, IN: Jochum, U., Wagner, G. (Hrsg.) Am Ende das Buch, Semiotische und soziale Aspekte des Internets, Konstanz 1998, S. 83-102.

[Krämer 97] Krämer, S., Vom Mythos ‚Künstliche Intelligenz‘ zum Mythos ‚Künstliche Kommunikation‘, oder: Ist eine nicht-anthropomorphe Beschreibung von Internet-Interaktionen möglich?, IN: Münker, S. u. Roesler, A. (Hrsg.), Mythos Internet, Frankfurt 1997, S.83-107.

[Kuhlen 98] Kuhlen, R., Die Mondlandung des Internet, Der Bundestagswahlkampf 1998 in den elektronischen Kommunikationsforen, Konstanz 1998.

[McLuhan 68] McLuhan, Die Gutenberg-Galaxis, Das Ende des Buchzeitalters, übers. v. Nänny, M., Düsseldorf 1968.

[McLuhan 92] McLuhan, M., Die magischen Kanäle, Neuaufl. Düsseldorf 1992.

[Minnis 88] Minnis, A.J., Medieval Theory of Authorship, Scholastic literary attitudes in the later Middle Ages, 2. Auflage, Worchester 1988.

[Ong 82] Ong, W., Oralität und Literalität, Die Technologisierung des Wortes, Opladen 1982.

[Ortmann 99a] Ortmann, S. Elektronische Literatur - Kreativität oder Chaos, In: Berliner Zimmer, ONLINE:  
<http://www.berlinerzimmer.de/eliteratur/chaos/default.htm>.

[Ortmann 99b] Ortmann, S., Netzautoren, Erscheinungsformen der Literatur im Internet, IN: Berliner Zimmer, ONLINE:  
<http://www.berlinerzimmer.de/eliteratur/netzautoren/referat2.html>.

[Pape 54] Pape, W., Griechisch-Deutsches Handwörterbuch, 3.Auflage 1954, Sp.399.

[Rötzer 98] Rötzer, F., Digitale Weltentwürfe, Streifzüge durch die Netzkultur, München 1998.

[Schwenger 81] Schwenger, H., Der Medienautor oder der vergessene Schriftsteller, IN: Kreuzer, H. (Hrsg.), Der Autor, Zeitschrift für Literaturwissenschaft und Linguistik, Heft 42, Göttingen 1981, S.93-101.

[Sturm/Wundrak 99] Sturm, B. und Wundrak, R., Identität und Internet, ONLINE:  
<http://www.univie.ac.at/Soziologie-GRUWI/richter/ende9798/rixta.htm/untersuchung.htm>.

[Wegner 99] Wegner, B., Kostenmodelle für wissenschaftliche Zeitschriften - Ergebnisse einer Umfrage durch die IuK-Fachgruppe E-Journals ONLINE:  
<http://www.bibliothek.uni-regensburg.de/iuk/kosten.htm>.

[Wirth 97] Wirth, U., Literatur im Internet. Oder: Wen kümmert's, wer liest?, IN: Münker, S. u. Roesler, A. (Hrsg.), Mythos Internet, Frankfurt 1997, S.319-337.

[Zimmer 99] Zimmer, D.E., Web-Literatur: Realität? Gerücht? Verheißung? Sackgasse?, Die digitale Bibliothek (V) — Eine Artikelserie für Nutzer und Verächter der Computernetze, IN: ZEIT-Archiv, ONLINE:  
<http://www2.zeit.de/zeit/tag/digbib/digbib5.html>



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 303 – 310

## **Access to information in electronic age — situation in Slovakia**

**Dr. Jaroslav Šušol**

Department of LIS, Faculty of Arts,  
Comenius University Bratislava,  
Gondova 2, 818 01 Bratislava, Slovakia  
e-mail: susol@fphil.uniba.sk

### **Abstract**

Electronic publishing brings about some substantial shifts in technological as well as socio-political patterns that were in place during the era of paper publications. The tension between producer and consumer side of the information cycle is the most significant feature of this change. Authors and publishers try to maximise their revenues from the marketing of information and backed by the possibilities of technology they force major changes in legislation governing the area of (public) information access. In this environment, the society's interest in a wide-spread accessibility of electronic data must be balanced by substantial financial investments – a solution that only the rich economies of the world can feel comfortable with.

### **Introduction**

Electronic publishing in general, and network publishing in particular, has become a reality on a world-wide scale. Approaches to definitions and classifications of this phenomenon are abundant, as are the articles enumerating advantages of the new communication technology over the traditional one.

Potentials connected with the usage of internetworking for the purposes of communication are really enormous – but how is the reality? Although people may live in illusion of having huge archives of computer data at the reach of their hands, the more you work with the internet the better you know how difficult it can be to find a really valuable piece of information. It is very often hidden before the eyes of a user who is not willing or able to register and/or pay.

What is even more important, electronic publishing as a technological change in communication cycle brings about some substantial shifts in socio-political patterns (author/publisher — user relationship) that were in place during the era when paper dominated the world of publishing.

Technology, economics and underlying legislation represent three intertwining principal aspects that one has to take into account when trying to assess state of



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

the art in the area of access to information. We will concentrate mostly, but not exclusively, on the topic of public access — a set of principles, interests, policies and regulations whose primary aim is to guarantee that every citizen has a free (preferably free of charge) access to certain type and amount of information, usually via specific state/community-supported institutions. The issue is of primary importance to the democratic society and state willing to have its citizens actively participating at its development. In the framework of this study the topic will be seen from two different angles:

1. overview of the problematic issues on the global level — what particularities are hidden behind electronic access and what is their impact on libraries' functioning
2. specific situation in Slovakia seen mostly from the viewpoint of present copyright legislation.

## **Electronic access**

First, predominantly technological, change lies in the way the information is being published. Its definition is very often based on a contradiction "Access versus ownership". This 3-word conjunction is becoming a sacred expression of the electronic era — it is heavily used and heavily significant. The transition from ownership of materialized publication to access-over-the-network pattern is one of the most representative changes that define new paradigm of information work in electronic environment.

In the traditional, printing stages of publishing system development the information (publication) was sold as a "can", a discrete product and once you bought it you (usually) owned it eternally and used happily ever after! Now the system is approaching a service model where the publisher provides the client with a network access to the electronic data.

This service approach is directly based on the traditional technology of providing access to large databases via database centers (Dialog etc.) which was popular mostly during 70s and 80s. Terminological question, however, comes up immediately — does providing access to databases equal publishing? Most of the information specialists will surely agree that in case of full-text databases we can give a positive answer.

## **Libraries and public access**

Archiving and providing access to information — these are two basic functions that libraries are most often connected with. Depending on their type, access function of libraries can be defined more specifically:

- in case of public libraries it is their social mission, the task to provide access to information for the people who otherwise, due to mostly economic

reasons, would not have an opportunity to get to a wide range of publications available on the market.

- in case of school/academic/research (special) libraries their primary function is the aggregation of diverse information resources fulfilling the information needs (both potential and real) of a specific patron community defined more or less institutionally.

Changes in publishing chain are inevitably bringing changes to the work of information institutions, including libraries. Traditionally, libraries were bound with material, mostly paper publications, later on adding other media and objects such as records, paintings, cassettes, even toys... Hence lending was the principal method of providing access to these objects.

Situation is changing and boundaries are being blurred — even public libraries nowadays work with modern information media and provide their patrons with electronic access to publications available over the networks. In accordance with their social function there is a strong movement — not only within the European Union — to turn libraries into public access points that could be used by those who cannot afford their own internet connection or access to certain web publications. Problems that the libraries are facing in connection with their involvement in new type of publishing market can be grouped into several categories, predominantly bound with the legislative/economic side of the author-user relationship:

#### **A. fair use**

- most of the publishers and authors believe that the traditional concept of exceptions from the "basket" of exclusive rights of authors is no longer valid in electronic environment - principal argument being the ease of quality copying.

#### **B. licensing**

- as publishers feel that present copyright legislation does not protect their interests in electronic publishing sufficiently they force their clients (i.e. also libraries) to enter into a contractual relation with them. They use "licence agreement to restrict access, define use and assure protection of their rights." (Davis, Reilly, 1998). Most of the authors see this type of arrangement as legitimate and natural (Okerson, 1998) although some stress the fact that signing an individual contract with each information provider can be quite a job for a library. Other library experts underline that no license agreement should go below the provisions of the valid copyright act, that it should address the question of archiving and perpetual access to data once paid for.

### **C. definition of the patron community**

- interlibrary loan and document delivery are not very popular among publishers as they consider these activities interfering with their own economic interests. That is why it is extremely important for a library to define precisely its patron community and the legitimate ways of accessing library resources (in-house, campus, external) in order to avoid future conflicts with the provider.

### **D. price models**

- subscription to the traditional and electronic versions of a journal are quite often bound together. This model usually requires a library to pay some extra amount of money (10–50 per cent, depending on the publisher – see, for example Tilburg..., 1997) for the electronic data which is a factor that libraries would like to eliminate. Other way of finding a reasonable price level for electronic publications is the creation of library consortia for common access to data - in this case publishers are often willing to negotiate the conditions acceptable for both sides.

## **Network publishing in Slovakia**

Publishing data on CD-ROM is becoming a common business activity on the Slovak publishing market – ranging from traditional texts (books) and multimedia publications to statistical data in a form more or less interactive. Nevertheless, as this type of publishing resembles traditional sell-and-buy (ownership) trade model we are going to concentrate on network publications.

This type of products, again, ranges from simple presentation of basic functions and missions of persons, institutions and companies to more sophisticated pages and sites containing what we could call a full-text scholarly/economic information. Quite a complete picture of the present situation can be found at the "Dictionary of Slovak internet", a site dedicated to mapping and "cataloguing" information resources available on the servers situated in Slovak republic ([www.zoznam.sk](http://www.zoznam.sk)). The sites can be divided into several groups:

#### **1. governmental and non-governmental organizations**

- mostly presenting missions, structures, projects, outputs as well as some full-text data in the form of reports, statistics...

#### **2. economic and business data**

- providing information useful for the subjects looking for the possibilities of basic trade relation or economic involvement either in Slovakia or abroad - banks, insurance companies, big enterprises... There is also certain amount of resources that are generally produced by commercial institutions as a



result of direct order by or with the support of state or various support funds. To name just one of them – Foreign Trade Support Fund ([www.fpzo.sk](http://www.fpzo.sk)) offers a complex site of this type, including a concise handbook on investing in Slovakia entitled "Doing Business in the Slovak republic".

### 3. scientific journals

- electronic versions of traditional scholarly journals – full-texts or abstracts with a possibility of subsequent ordering full-text in a paper or electronic format – published by institutes of Slovak Academy of Science or universities. Probably the most representative site at the moment can be found at [www.elis.sk](http://www.elis.sk)

### 4. newspapers and journals

- originally concentrated on the "historical" site of the Slovak Academy of Science, nowadays every major daily or journal in Slovakia has its own website where the electronic version of the paper is displayed. Besides, there is a small number of journals that are published only in cyberspace ([www.inzine.sk](http://www.inzine.sk)) – their survival, however, is problematic due to the small size of Slovak internet community and, consequently, small advertising potentials.

The division between free and paid resources reflects, as anywhere else in the world, the impeccable logic of market mechanisms. In general — free is everything that is not expected to bring any profit. Similar logic works specifically in the area of scholarly data — if there is a journal that is able to earn some money on subscription fees it is difficult to persuade the editor to give it for free. As most of the scholarly journals in Slovakia are published by scientific institutions, produced with a financial support of scientific or state grants and have a limited circulation, the push towards profit is not very strong, even in electronic form.

## Copyright

The provisions of a national copyright act define basic, political principles governing the overall atmosphere on the information/publishing market. Of course, fundamental features of this type of act are guided by international agreements — Czechoslovakia signed the Berne convention relatively shortly after the foundation of the state, in 1928. Modern, post-1945 practice of the copyright in the country was governed by the "socialist" act 35/1965 on literary, scientific and artistic works. At that time copyright was not a very hot topic in Slovakia — neither for the authors nor for the libraries. Settled in the information and social system of the country they offered their service to the patrons and did not ask about the legal side of the "industry".

In connection with the aspirations towards the European Union the process of the approximation of legal system of Slovakia started very early after the break-up of Czechoslovakia and one of the first acts to be created on the principles of a

European directive was the copyright act 383/1997. It passed in the Parliament on December 5, 1997.

Here are some of the act's provisions, crucial from the point of view of library and information user activities. Some of them are quite common, some are not:

- *the content of copyright* (art. 16) — the author has an exclusive right to authorise any use of the work..., in particular distribution of the work or its copy onto the public by sale, rental, lending or any other form of distribution of work or transfer of ownership
- *distribution of work after first sale* (art. 20) — the right of author to authorise a lending of the original of work or its copy does not cease to exist by sale or any other act of dissemination
- *reproduction of work for personal purposes* (art. 21) — physical person can make without the authorisation of author a copy of a work which is made public for his/her personal purposes and is not obliged to provide a remuneration to author. In this case the remuneration is paid by the producers or importers of reprographic or other technological devices or media
- *reproduction for teaching purposes* (art. 23) — the authorisation of author is not required and the remuneration is not paid in case of the reproduction of a short part of work which is made public and is done for teaching purposes. This includes the reproduction of a published work by the use of special methods for needs of visually impaired persons
- *reprographic reproduction by libraries and archives* (art. 24) — these institutions can make copies without the authorisation of the rightholder and without any remuneration if they are short works or short extracts and the copies are done either on behalf of the patron (study or research purposes) or as a replacement, archiving or preservation copies.

Among the apologists of public access rights the act is seen as very restrictive, limiting traditional fair use exceptions. What is more, even the principal traditional library service — lending was put into a strange position (art. 16, 20), as it is absolutely impossible and even illogical for every single library in the country to ask an author for the permission each time when his/her book is being borrowed. Although there are tendencies to apply the principles of the collective administration of authors' rights, it does not solve the absurd contradiction of the permission as such.

Unfortunately, the overall image of the world's copyright legislation seems to be moving in this direction — let us name just the U.S. Digital Millennium Copyright Act or the above mentioned EU copyright directive. EBLIDA, the umbrella association for library, information and documentation associations and all kinds of libraries throughout Europe, have several times expressed their concerns connected with narrowing traditional exceptions (fair use) in the new copyright legislation model. EBLIDA's objections reflect the fact that proposed provisions recognise only "use for the sole purpose of illustration for teaching or scientific

research" which excludes other educational purposes and other non-scientific research, including private copying (Save..., 1998).

The difference between well-established copyright systems and emerging systems of Central and Eastern Europe (Slovakia included) is very well visible when it comes to a mere activity of library lending. Various types of *public lending schemes* have been started in the countries of EU but basically their role is to pay financial remuneration to the authors for the works that are lent by libraries. In most cases it is the state who takes care of the economic side of the scheme and thus takes the responsibility for the public access to information for the citizens. In Slovakia there is no such a system for the moment but there is a pressure from the libraries to establish one – in accordance with the EC Directive 92/100/EEC of 19 November 1992 on rental right and lending right and on certain rights related to copyright in the field of intellectual property.

## **Conclusions**

The aim of this article was to give a short overview of the situation in the area of access to information on the territory of Slovakia, especially in the context of electronic data and recent changes in economic environment and legislation.

Technological aspects of the situation are improving and private internet market is slowly growing bigger. According to the latest survey of NetProjekt 10% of the population of Slovakia uses internet more or less regularly ([www.prieskum.sk](http://www.prieskum.sk)). As for the public access to network resources, libraries (mostly on the regional and city level) are assuming the position of public access points.

Valuable network resources can be found both on domestic and foreign servers, free of charge but, mostly, paid. Especially when it comes to "serious" scholarly information (Journals) one has to turn to foreign paid resources. As the economic situation of libraries is quite complicated at present, usage of this type of data is very rare – if it occurs it is mostly with the support of various scientific and cultural grants (universities cooperation programmes, Open society foundation grants). Legislative controversies within the copyright act should be solved with adding principles of public lending into the act on libraries or new version of copyright act that is going to be on the parliamentary agenda later this year. It should reflect the idea that if the society wants to keep free public access to information then it is the state who should be paying money for the remuneration of authors and it would be more efficient if it was done on a centralized level rather than on a level of individual library, as it works in many other European countries.

With the support of international programmes like CECUP (Central and Eastern European Copyright User Platform) the Slovak Library Association raises the copyright awareness among librarians, library patrons and seeks to sup

.....



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 17 – 30

# **Usability Engineering Methods for the Web**

## **Results From a Usability Study**

**PD Dr. Ilse Harms and Werner Schweibenz**

Fachrichtung Informationswissenschaft

Universität des Saarlandes

D-66041 Saarbrücken

Fon: +49(0)681/302-3542

E-Mail: {harms, w.schweibenz}@rz.uni-sb.de

### **Abstract**

The paper presents the results of a study on usability methods for evaluating Web sites. It summarizes the "Heuristics for Web Communications," and reports the practical experiences with these heuristics, contrasting them with the "Keevil Index" and combining them with user testing with thinking aloud. It concludes that working with the "Heuristics for Web Communications" takes more time and effort than working with the "Keevil Index," but produces more consistent results. The heuristics proved to be applicable both in heuristic evaluation and in combination with user testing.

### **Zusammenfassung**

Der Beitrag präsentiert eine Studie über Evaluationsmethoden zur Web-Usability. Er beschreibt die "Heuristics for Web Communications" und berichtet von den praktischen Erfahrungen mit den Heuristiken, die mit dem "Keevil Index" verglichen und mit Benutzertests mit lautem Denken kombiniert werden. Das Ergebnis zeigt, dass eine Evaluation mit den beschriebenen Heuristiken gegenüber dem "Keevil Index" mehr Zeit und Aufwand erfordert, aber konsistentere Ergebnisse bringt. Die Heuristiken haben sich sowohl in der experten-zentrierten Evaluation als auch in Kombination mit dem Benutzertest insgesamt als geeignete Evaluationsmethode erwiesen.



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

## **1. Usability Engineering for the Web**

Usability engineering is one of the key concepts of human-computer interaction. Currently, the focus of usability engineering is on Web usability. The reason is that the World Wide Web is more and more used for commercial purposes and applications and therefore requires usable Web sites because "Usability mries the Web. Simply stated, if the customer can't find a product, then he or she will not buy it." (Nielsen 2000: 9). This explains the recent interest in evaluation methods for Web sites.

Dealing with Web usability, one has to consider two aspects. One aspect is the concept of usability and the other aspect is the complex graphical user interface of the World Wide Web and how it is applied by users. The demand for usability evokes the question what usability means for Web sites. In general, the usability of a product can be defined, according to EN ISO 9241-11: 1998, as the extend to which a certain user can use it to reach certain aims in a way which is effective, efficient and satisfying (for definitions of these terms see Krömker 1999: 23). EN ISO 9241 is a standard that refers to various kinds of ergonomic requirements for office work with visual display terminals and therefore represents more an ideal than a concrete list of requirements. In the context of the World Wide Web, "usability refers to how easy it is to find, understand and use the information displayed on a Web site" (Keevil 1998: 271). A general problem of usability is that it is "soft" concept, i.e. that it is hard to measure as it is based on subjective factors as user performance or the mental effort and attitude of users and that it has to take into account the context in which the product is used, e.g. the class of users being studied, the tasks they perform and the environment they work in. Another problem is that usability engineering is surrounded by an intimidation barrier and hardly used in practice due to perceived costs and intimidating complexity (Nielsen 1994). To describe the amount of complexity, the next section gives a survey of methods for usability engineering.

## **2. A Survey of Methods for Usability Engineering**

According to Krömker (1999: 25), usability engineering is a set of methods to design user-friendly products and a process in which users, developers, and technical writers work together in order to integrate the knowledge and experience of all participants and enhance the quality of the product. The methods of usability engineering can be categorized in expert-focused and userfocused methods. Among the expert-focused methods are several variations of heuristic evaluation. According to Nielsen (1997a, 1543) "heuristic evaluation is a way of finding usability problems in a design by contrasting it with a list of established usability principles". The established usability principles are listed in guidelines or checklists like Keevil's Usability Index (Keevil 1998) or Molich and Nielsen's nine principles for human-computer dialogue (Molich & Nielsen 1990) or the "Heuristics for Web Communications." In the evaluation process, experts compare the product with these guidelines and judge the compliance of the interface with recognized usability problems. The advantage of expert-focused evaluation is that it is a

relatively simple and fast process. A comparatively small number of five evaluators can find some 75 per cent of the usability problems of a product in a relatively short time (for details see Levi & Conrad 1996). The disadvantages are that experts have to do the evaluation and that experts cannot ignore their own knowledge of the subject, i.e. they cannot "step back behind what they already know." So they will always be surrogate users (expert evaluators who emulate users) as Kantner & Rosenbaum (1997: 154) call them. In contrast to expert-focused methods, user-focused methods rely on actual users to test the usability of a product. This process is called user testing, and according to Nielsen (1997a: 1543) it "is the most fundamental usability method and is in some sense irreplaceable, since it provides direct information about how people use computers and what their exact problems are with the concrete interface being tested." There are various methods for user testing. One of the most popular and most effective method is the laboratory test with the thinking aloud method (Nielsen 1993: 195) which was used in our case study. The advantage of user-focused evaluation is that the tests supply a huge amount of qualitative data that show how actual users handle the product. The disadvantages are that the tests take place in a laboratory situation and that a lot of equipment and coordination is necessary to conduct the test which makes it laborintensive.

### **3. Description of the Usability Study**

#### **3.1 The Process**

In the winter term 1999/2000, the department of information science at the University of Saarland decided to evaluate different methods of usability engineering for the Web. In cooperation with the Stiftung Saarländischer Kulturbesitz, a foundation for cultural heritage, the Web site of the Saarland Museum (<http://www.saarlandmuseum.de>), the art museum of the federal state of Saarland, was evaluated. It is a graphically designed Web site of the third generation (Siegel 1999: 15) that went online in summer 1999. The project was carried out by sixteen graduate students, who all had already received a training in usability engineering in a research class, and two lecturers as coordinators. The study had two aims. The first aim was to evaluate some evaluation methods, especially the "Heuristics for Web Communication," the second was to improve the usability of the Web site of the Saarland Museum. Therefore we decided to use a combination of heuristic evaluation and user testing as suggested in the research literature (Nielsen 1993; 1997a; Kantner & Rosenbaum 1997). Figure 1 illustrates how the two methods were applied.

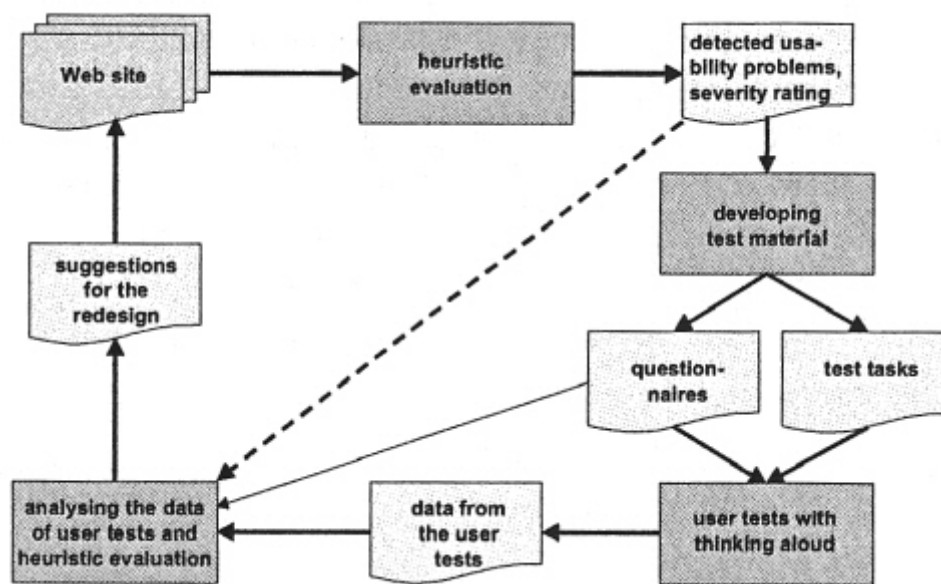


Figure 2: The evaluation process of the usability study

### 3.2 Heuristic Evaluation With the Heuristics for Web Communications

The first step was a heuristic evaluation. As mentioned above, there is a multitude of heuristics. Heuristics can be specific for a certain domain or generally applicable. They can be design-oriented or evaluation oriented or both. They can be based on research or on experience of practitioners. Therefore, heuristics vary in extent and quality. In our case study we used the "Heuristics for Web Communications," developed at the International Summer Workshop *Exploring a Communication Model for Web Design*, Seattle, WA, July 10-17, 1999 (the heuristics used during the workshop are available on the home page of the workshop). The workshop was organized by the faculty of the Departments of Technical Communication of the University of Washington and the University of Twente, the Netherlands. The heuristics are based on research findings in technical writing and psychology. During the workshop the heuristics were tested by 40 participants and professional Web developers from various Web design companies in the Seattle area. The heuristics were revised according to the feedback of the participants and developers and were published in a special issue of the listed in a special issue of the Journal of Technical Communication in August 2000.

The "Heuristics for Web Communications" consist of five different heuristics, all based on profound research in technical communication. The five heuristics deal with all important aspects of Web sites: displaying information, navigation, text comprehension, role playing (i.e. author-reader relationship), and data collection for analyzing interaction. The content of the heuristics can be summed up as follows:

The heuristic *Displaying information on the Web* consists of guidelines for visuals, e.g.

- how to design and arrange display elements,
- how to ensure that text is readable
- how to use pictures, illustrations, icons and motion.

The *Heuristic for Web Navigation* deals with hypertext theory and offers guidelines for navigation and orientation, e.g.

- how to design orientation information on each page,
- how to coordinate navigation devices,
- how to design site-level orientation information.

The heuristic *Text Comprehension and the Web: Heuristics for Writing Understandable Web Pages* focuses on text comprehension and issues of text quality, e.g.

- how to select, design, and organize content,
- what style to use,
- what makes Web pages credible and trustworthy.

The heuristic *Role Playing on the Web* discusses the typical rhetorical roles of the implied author and reader of the Web pages and their rhetorical roles, e.g.

- how rhetoric is used to describe author roles and reader roles, and
- what kind of relation exists between author roles and reader roles.

The heuristics *Web Data Collection for Analyzing and Interacting with Your Users* focuses on analyzing the audience of a Web site and building a relationship between either between you and your users or among the users themselves using for example

- server log data for analyzing the use of Web pages and their audience, and
- means to build a relationship and create a sense of community with the audience.

The four content-oriented heuristics (except the heuristics on *Web Data Collection* which was not applied due to access restrictions to log files) were applied according to Kantner & Rosenbaum (1997: 155). The graduate students, who had received an introduction to the heuristics in a research class, worked in teams of four. Each team conducted an evaluation of the same selected number of pages from the chosen Web site. In a two hour session of individual evaluation, the team members applied one of the four heuristics on the Web site. During the evaluation process they took notes of usability problems according to the various points listed in the heuristics. Then the team members gathered and discussed their findings. The usability problems detected in the heuristic evaluation were graded in a severity rating according to Nielsen (1997b) by each team. The rating ranged from 0 (no usability problem) to 4 (usability catastrophe and was conducted with respect



to the frequency and persistence of the problems and the impact they have on users. At the end of the evaluation, the four teams presented their findings in a plenary meeting. The evaluation process took about five hours.

The most frequent usability problems were navigational and orientation problems as described in the *Heuristic for Web Navigation*, followed by general design problems as named in the heuristic *Displaying information on the Web*. Afterwards the two lecturers collected the written findings of the team members and the compiled lists of problems of the plenary session in order to draw up a list of all usability problems. The findings were used to design tasks for a user test in the laboratory.

### 3.3 User Testing in the Usability Laboratory With Thinking Aloud

The next step of the evaluation was a user test in the usability laboratory. Figure 2 shows a sketch of the laboratory in which the tests take place.

In the lab, real users have to work on tasks while thinking aloud, i.e. they verbalize their thoughts and comment on their actions while they handle the computer. This "allows a very direct understanding of what parts of the dialogue cause the most problems" (Nielsen 1993: 195). During the test users work on standardized test task and are supervised by a test manager. The tests are recorded on video by a technical assistant who operates two video cameras. One of the camera is focused on the face and the hands of the participant, the other one on the computer screen. The recordings of the two cameras are blended together on the video. In order to catch the details of interaction a digital screencam records the actions on the screen. In a laborintensive process, the findings are transcribed and categorized.

As we evaluated a museum Web site, it suggested itself to recruit participants with an interest in art. Therefore we asked students of the arts and science department and art teachers to participate. The teachers were chosen to increase the average age. The number of participants was arranged according to Virzi (1992: 468), who suggests at least 15 participants. In our study, 17 users participated. Five of them were teachers, 12 were students of the arts and science department. Seven participants were male, ten female. The youngest user was 19, the oldest 48, the average age being 27.

It takes some time and effort to design the test task scenario for the user test of a large informational Web site (cf. Kantner & Rosenbaum 1997: 154). The test tasks should be as representative as possible of the use to which the System will be put in the field and small enough to be completed, but not so small that they are trivial (Nielsen 1993: 185f). The test scenario, which had been discussed with the client of our case study, consisted of nine tasks that represented potential usability problems detected in the heuristic evaluation. Table 1 shows a selection of the test tasks.

Potential Usability Problem	Tasks (abbreviated)
Links are hidden in graphical design (images).	1) Go from the splash screen to the core page.
Insufficiently linked information, the exhibition is not linked to the opening hours.	3) Look for the opening hours of a futur exhibition.
Insufficiently linked information, the exhibition is not linked to the service section where tours are offered.	4) Look for guided tours to the current exhibition.
All pages are titled the same. There are no individual title-tags on the different pages.	7) Use bookmarks to go back to certain page.

**Table 1:** A selection form the test tasks of the user tests

The user tests revealed that all assumed problems were usability problems for test users. The findings were grouped into several categories, e.g. problems handling the splash screen (task 1), insufficient use of links between related information (tasks 3 and 4) etc., and illustrated by lively quotes from the test protocols. For example, after finishing task 1 a participant stated: *"The first screen only shows a headline, a picture and an address but no link. (break) I click on the picture. It works!"*, another remarked: *"As an Internet beginner I honestly have a problem to get to the next page. I'm a little helpless because I prefer big arrows and buttons that say 'next page.' But I made it."* While struggling with task 4 the patience of a participant snapped and he said: *"Now I would try to get in touch with someone and write an email, if they offer a guided tour because all the searching takes far to long."* Trying to use the bookmarks in task 7 a participant said: *"Usually I would use the bookmarks. (clicks on bookmarks) Well, now I see that all the bookmarks have the same name. "*, while another participant used sarcasm: *"That's really funny, because I have five bookmarks all named Saarland Museum Saarbrücken. That helps a lot. Great! (laughs)."*

Quotes like these make usability problems come alive and show the reactions of real users. From our experience, these reactions are much more convincing as dry statements of experts no matter how profound these statements might be and no matter on what kind of theory they are based.

The test time and performance of the users was influenced by their Computer literacy, data which were collected in a questionnaire. The average test time was about 20 minutes and corresponds with the expected test time derived from three pre-tests. The duration of the test shows some connection between the Computer literacy and the amount of time needed to complete the tasks. The shortest test of 14 minutes was conducted with a user who used the Web for two years or longer and several times a week. The longest test of 30 minutes was conducted by a participant with little Computer literacy. Further analysis did not seem necessary to

us because the duration of a test is also influenced by other factors like interest in the subject and the medium and the method of thinking aloud.

#### **4. Practical Experiences With the Different Methods**

In our case study the theoretical foundations of the heuristic evaluation were the Heuristics for Web Communications. At the time of our evaluation the Heuristics for Web Communications were brand new. Revised versions, incorporating the experiences from the summer workshop, had been published for a limited audience just a few month before. Only little practical experience existed in applying the Web heuristics.

Therefore we thought it useful to contrast the "Heuristics for Web Communications" with another heuristic evaluation tool, "Keevil's Usability Index." According to Keevil (1998: 271) the usability index is a "measure, expressed as a per Cent, of how closely the features of a Web site match generally accepted usability guidelines."

The Usability Index consists of five categories (Keevil 1998: 273):

- Finding the information: Can you find the information you want?
- Understanding the information: After you find the information, can you understand it?
- Supporting user tasks: Does the information help you perform a task?
- Evaluating the technical accuracy: Is the technical information complete?
- Presenting the information: Does the information look like a quality product?

Judging from the 203 questions, Keevil's Usability Index seems focused on commercial Web sites. But it seemed adequate to use it for a cultural institution Web sites also, because Keevil (1998: 275) points out that the Usability Index is generally applicable: "Information Developers can use the checklist to measure how easy it is to find, understand, and use information displayed on a Web site."

In our case study, fifteen students (one didn't hand in the index) used the Keevil Usability Index to evaluate the Web site of the Saarland Museum. The results reached from 29% to 55%, the arithmetic mean being 47%. The deviation in the percentage of usability is remarkable. It is due to the bias of the evaluators in interpreting the questions. One reason is that certain sets of questions from the Usability Index did not fit for the Web site under evaluation. In this case some evaluators choose to vote for not applicable, while others voted for no. This explains the big differences between the 'N/A' votes and the 'No' votes and in the overall usability because the usability is calculated from the total number of yes answers divided by the total numbers of yes and no answers. Another reason is that the checklist, like every checklist, is open to interpretation. Keevil (1998: 275) was aware of this problem and tried to reduce it by only allowing the answers 'not applicable,' 'yes,' and 'no.' But still, there is considerable room for interpretation. Apart from the problem of interpretation, there is another disadvantage of the Usability Index. It gives a number in per Cent and some hints on usability

problems that can be derived from the 203 questions but hardly identifies concrete usability problems.

The identification of concrete usability problems and suggestions on how to improve usability are the practical advantages of the "Heuristics for Web Communications." The heuristics are not simply checklists that can be answered by 'not applicable,' 'yes,' and 'no'. They are guidelines designed as statements and questions that guide the evaluator to identify concrete usability problems by asking guiding questions like "Can you decipher all of the elements in the display easily? If not, consider making them larger." or "Which of the organization's values should be emphasized?" or giving hints like "Make sure the most important links appear high enough on the page to be visible without scrolling, regardless of the resolution of the user's monitor. When pages must scroll, provide visual cues to encourage users to scroll down to links that are below the scroll line." By contrasting these established usability principles with the Web site under evaluation the evaluator or information designer can decide if and what kind of usability problems exist and how they can be removed. This is the big advantage of the heuristics.

The disadvantage of the Heuristics for Web Communications is that they are very detailed and complicated compared with general heuristics like the ones of Molich & Nielson. Molich & Nielson suggest nine basic items of usability (Molich & Nielson 1990: 338) (Table 2).

1 Use simple and natural language	6 Provide clearly marked exits
2 Speak the user's language	7 Provide shortcuts
3 Minimize the user's memory load	8 Provide helpful error messages
4 Be consistent	9 Prevent errors
5 Provide feedback	

**Table 2:** Molich & Nielson's nine basic items of usability

Although it is possible to do a successful evaluation with these nine basic items of usability, users might need more guidance in the evaluation as it is offered in the Heuristics for Web Communications. These heuristics support the evaluator by providing a structured "guided tour" for the evaluation process that takes into consideration both the big picture and important details. They help the evaluator to consider all substantial usability issues and to focus on the important points. Moreover they generate a profound impression of the overall quality of a Web site. This makes the Heuristics for Web Communications a valuable tool in Web usability engineering.

The four content-oriented heuristics (the fifth heuristic was not applied due to restrictions on log file access) are very different as far as their ease of application and the level of background knowledge are concerned. We found that the heuristic *Displaying information on the Web*, the *Heuristic for Web Navigation*, and the heuristic *Text Comprehension and the Web* can be successfully applied if the evaluators have an average level of knowledge in information design and Web design. The evaluators in our case study, all graduate students of information

science, had no difficulties to apply them. The heuristic *Role Playing on the Web* requires some special knowledge in hypertext theory as it is based on the quite complicated author-reader relationship in hypertext (Michalak & Coney 1993). Although it is very interesting and provides promising results, the evaluators in our case study had some difficulties to apply it.

As expected from research literature (Nielsen 1992: 378f), the heuristics detected a great number of minor usability problems, which is no disadvantage at all because user testing is not an adequate means to detect such minor problems. Minor problems were, for example, inconsistent use of link colors, no text messages for graphic links, complicated sentences, deficits in page structure and organization, lack of informative titles, overuse of bold and italics, meaningless animation, flaws in the author-reader relationship etc. Although being real usability problems, such minor problems are not observable in user testing, because average users do not realize that this kind of deficiencies cause problems because they lack the background knowledge in Web design and technical communication. The user tests in the usability lab were very laborintensive for several reasons: the technical equipment had to be arranged, the test scenario had to be designed and tested, participants had to be recruited and tests had to be conducted with two experimenters who had to be present all the time to supervise the participants and the technical equipment. The analysis and evaluation of the test data was also time-consuming because the data had to be transcribed and categorized. The big advantage of this method was that the recordings, especially the screencan flies, show cursor movements that help to identify problems in navigation and orientation. This is especially helpful when discussing the findings and suggestions for the redesign with the client. There is also a simplified method of thinking aloud testing, in which the experimenters simply take notes of their observations is less laborintensive than videotaping and transcribing the tests. But from our experience, it is difficult for one or two experiments to follow the course of the test and take notes at the same time if the test consists of more than some basic functions. Therefore video taping or screencam recording is essential. An alternative to transcribing whole test sessions would be to transcribe only the most important sequences.

An important point we noticed when comparing answers in the questionnaires with courses of the tests is that answers about the test experience are often not very reliable. The answers about the satisfaction with the Web site did often not correspond with the actual experience of the participants. For example, several participants stated that they had no problems with navigation and orientation although they had had serious problems during the test. The reasons for this gap between behavior and statements cannot be discussed here. From our experience it is important to remark that a questionnaire alone cannot provide reliable results. This is not new but confirms the phenomenon that impressions of the own behavior and the behavior in the situation of social reality show a certain deviation. Despite of this fact, from our experience questionnaires or interviews are necessary to give the participants the opportunity to comment on the course of the test. Test user appreciate this opportunity and the results can be used to derive additional information about the acceptance of the Web site.

## **5. Conclusion**

The evaluation method used in usability engineering depends of the subject that is evaluated and the goals of the evaluation. Although the combination of heuristic evaluation and user testing provides good results, it is costly as far as time and resources are concerned. With respect to the cost-benefit ration, in many cases the heuristic evaluation is sufficient to detect a reasonable number of minor and major usability problems.

In our case study, the Heuristics for Web Communications proved to be applicable tools for heuristic evaluation. The heuristics support a structured evaluation and help both to find and to solve usability problems. In contrast to simple checklists, they give the evaluators some scope for interpretation while offering guidance at the same time. The drawback of the heuristics is that they cannot be applied by novices. The evaluators need some background knowledge in Web design and evaluation. The heuristics were helpful in pointing out critical points in the Web site that were evaluated in the user test. Compared to user testing, the heuristic evaluation was less laborintensive. Nevertheless, user testing is a very valuable tool for usability engineering because actual users give and impression how the Web site will be used in practice. This focus on the actual users and the vivid and expressive statements they give justifies the much higher expense in certain cases. From our experience, the combination of both heuristic evaluation with the Heuristics for Web Communications and user testing with thinking aloud is a very useful method of usability engineering.

## **6. References**

EN ISO 9241-11:1998 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 11: Guidance on Usability. Geneve, CH: ISO.

Heuristics for Web Communications. Special Issue of the Journal of Technical Communication, 47 (3) August 2000.

Keevil, Benjamin (1998): Measuring the Usability Index of Your Web Site. In: Proceedings of the CHI '98 Conference, 18-23 April 1998, Los Angeles, CA. New York, NY: ACM Press. 271-277. Also available online: Internet, URL <http://www3.sympatico.ca/bkeevil/sigdoc98/index.html>. Version: 09/98. Visited: 08/17/00.

Kantner, Laurie/Rosenbaum, Stephanie (1997): Usability Studies of WWW Sites: Heuristic Evaluation vs. Laboratory Testing. In: Proceedings of the 15th International Conference on Computer Documentation SIGDOC '97: Crossroads in Communication. 19-22 October 1997, Snowbird, UT. New York, NY: ACM Press. 153-160.

Krömker, Heidi (1999): Die Welt der Benutzerfreundlichkeit. In: Hennig, Jörg/Tjarks-Sobhani, Marita (Hrsg.): Verständlichkeit und Nutzungsfreundlichkeit von technischer Dokumentation. (tekomp Schriften zur technischen Kommunikation Bd. 1) Lübeck: Schmidt-Römhild. 22-33.

Levi, Michael D./Conrad, Frederick G. (1996): A Heuristic Evaluation of a World Wide Web Prototype. In: interactions, 07/1996. 51-61.

Michalak, Susan/Coney, Mary (1993): Hypertext and the Author/Reader Dialogue. In: Proceedings of Hypertext '93, November 14-18, 1993, Seattle, WA. New York, ACM. 174-182.

Molich, Rolf/Nielsen, Jakob (1990): Improving A Human-computer Dialogue. In: Communications of the ACM, 33 (3) 1990. 338-348.

Nielsen, Jakob (1992): Finding Usability Problems Through Heuristic Evaluation. In: Proceedings of the CHI '92 Conference, 3-7 May 1992, Monterey, CA. New York, NY: ACM Press. 373-380.

Nielsen, Jakob (1993). Usability Engineering. Boston: Academic Press.

Nielsen, Jakob (1994): Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier. Internet, URL [http://www.useit.com/papers/guerrilla\\_hci.html](http://www.useit.com/papers/guerrilla_hci.html). Version: 1994. Visited: 08/17/00.

Nielsen, Jakob (1997a): Usability Testing. In: Salvendy, Gavriel (1997, ed.): Handbook of Human Factors and Ergonomics. 2nd edition. New York: John Wiley & Sons. 1543-1568.

Nielsen, Jakob (1997b): Severity Ratings for Usability Problems. Internet, URL <http://www.useit.com/papers/heuristic/severityrating.html>. Version: 01/11/99. Visited: 08/17/00.

Nielsen, Jakob (2000): Designing Web Usability. Indianapolis, IN.: New Riders Publ. .

Siegel, David (1999): Web Site Design. Killer Web Sites der 3. Generation. Frankfurt/Main: Zweitausendeins.

Virzi, Robert A. (1992): Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough? In: Human Factors, 34 (4) 1992: 457-468. Web site of the Department of Technical Communication at the University of Washington. Internet, URL <http://www.uwtc.washington.edu/international/workshop/1999/post-workshop/heuristics/default.htm>. Version: 08/02/99. Visited: 08/17/00.



# Effektivität von Recherchen im WWW

## Vergleichende Evaluierung von Such- und Metasuchmaschinen

**Christian Wolff**

Universität Leipzig, Institut für Informatik,  
Abt. Automatische Sprachverarbeitung  
Augustusplatz 10/11, 04109 Leipzig  
Wolff@informatik.uni-leipzig.de

### Zusammenfassung

Der vorliegende Beitrag befasst sich mit der Informationssuche im World Wide Web und ihrer Evaluierung. Ausgehend von einer Zusammenschau der wesentlichen Merkmale des World Wide Web als heterogener Dokumentkollektion (Kap. 1) werden Aufbau und Merkmale von Suchmaschinen und Metasuchmaschinen eingeführt sowie die Problematik der Evaluierung von Suchmaschinen und eine Übersicht bisheriger Ergebnisse diskutiert (Kap. 2). In Kap. 3 werden Aufbau, Durchführung und Ergebnisse einer Vergleichsstudie vorgestellt, bei der mit Hilfe eines *paper-and-pencil*-Experiments ausgewählte Such- und Metasuchmaschinen evaluiert wurden. Schließlich zieht Kap. 4 Schlussfolgerungen aus dieser Studie und gibt einen Ausblick auf Optimierungsmöglichkeiten für Suchmaschinen.

### Abstract

This paper is concerned with the evaluation of web search engines. In particular, an empirical comparison of search engines with meta search engines is described. Starting from some introductory notes on basic characteristics of the web as a document collection (Ch. 1), the main features of search and meta search engines are given and relevant literature on search engine evaluation is presented (Ch. 2). Design, execution, and results of this empirical study follow in Ch. 3. Finally, Ch. 4 discusses major implications from these results and gives some hints for further study and possible optimisation strategies for search engines.





## 1. Einleitung: Charakteristika des World Wide Web

Das World Wide Web als größter elektronisch verfügbarer Informationsbestand stellt neue Herausforderungen für die Informationserschließung: Suchmaschinen als Information Retrieval-Systeme stellen neben *Web Directories* wie *Yahoo*, die auf intellektuell erstellten thematischen Hierarchien aufbauen, das wichtigste Werkzeug für die Informationssuche dar und werden auch in vielen Fällen als WWW-Portale, d. h. als primäre Zugangsseiten genutzt. Man kann Suchmaschinen zwar mit traditionellen IR-Systemen vergleichen, sie unterscheiden sich von diesen aber in besonderer Weise durch die Heterogenität und Variabilität der von ihnen indexierten Datenbestände im World Wide Web. Diese sind durch folgende Merkmale gekennzeichnet:

- Anzahl und Umfang der Informationseinheiten wachsen sehr dynamisch; gleichzeitig sind die Daten relativ unbeständig, d. h. ihre Verfügbarkeit über einen längeren Zeitraum ist nicht gesichert (vgl. OCLC 1999).
- Die Daten ein sehr hohes Volumen, derzeit ca.  $10^9$  Dokumente.
- Die Daten sind in der Regel semi- oder unstrukturiert, d. h. sie können deklaratives Markup in gewissem Umfang enthalten (HTML-Dateien); es existiert eine Vielfalt von Medientypen und Formaten.
- Die Ressourcen liegen in unterschiedlichen Sprachen vor, bei eindeutiger Dominanz des Englischen: Dessen Anteil liegt bei etwa 50 – 80 %, gefolgt von Japanisch und Deutsch mit je etwa 4 %, vgl. PIMIENTA et al. 1998.
- Durch die allgemein verfügbaren Publikationsmöglichkeiten im World Wide Web, bei denen in vielen Fällen keine Qualitätskontrolle stattfindet, sind die Inhalte unterschiedlicher (inhaltlicher, formaler) Qualität.

Eine einheitliche Methodologie für die Bestimmung der wesentlichen Merkmale des World Wide Web beginnt sich – unter dem Einfluss der *Web Characterization Activity* des *World Wide Web Consortium* – erst herauszubilden (vgl. OCLC 1999, LAVOIE & FRYSTYK NIELSEN 1999). PITKOW 1998 gibt einen Literaturüberblick zur den bisherigen Ansätzen der Kenngrößenerfassung und Charakterisierung des World Wide Web.

## 2. Suchmaschinen

Die Informationserschließung im World Wide Web kann grundsätzlich auf unterschiedliche Arten erfolgen:

- Durch direkte Anwahl einer bekannten Adresse (URL), ggf. in Verbindung mit einer lokalen Suche auf einem Webserver,

- durch die Verwendung eines Verzeichnisdienstes wie etwa *Yahoo* (vgl. <http://www.yahoo.com>, <http://www.yahoo.de>), der Ressourcen aus dem World Wide Web klassifiziert und ordnet,
- durch die Verwendung einer Suchmaschine, die über einen Index eines Teils der im World Wide Web verfügbaren Ressourcen verfügt und auf Anfrage Nachweise (d. h. Adressen und Kurzbeschreibungen) von Dokumenten im World Wide Web liefern kann oder
- durch Metasuchmaschinen, die zwar selbst keinen Index verwalten, aber die Ergebnismengen einer Mehrzahl von Suchmaschinen bündeln.

Es besteht dabei das grundsätzliche Problem, dass die (leistungsfähigeren) Suchmaschinen im World Wide Web kommerzielle Softwaresysteme sind, über deren inneren Aufbau nur wenige Details bekannt sind. In Kenntnis der Verfahren des Information Retrieval und durch Analyse der von den Suchmaschinen erzielten Ergebnisse und der dabei verwendeten Anfragelogik lassen sich aber Rückschlüsse bezüglich ihrer Funktionsweise ziehen.

## **2.1 Aufbau von Such- und Metasuchmaschinen**

Eine Suchmaschine für das World Wide Web ist in der Regel aus folgenden Komponenten aufgebaut:

1. Einem oder mehreren Erfassungsagenten (spider, crawler), die Dateien aus dem World Wide Web zur Indexierung laden,
2. einer Indexierungskomponente, die die Daten auswertet und den Index aufbaut und verwaltet,
3. der *Abfrageschnittstelle*, über die der Benutzer seine Anfrage an die Suchmaschine formulieren kann, und
4. der *Anfrageverarbeitung*, die die Anfrage mit dem Index abgleicht (Retrievalfunktion) und die Ergebnisse an die Benutzerschnittstelle der Suchmaschine weiterleitet.

Im Unterschied zu „einfachen“ Suchmaschinen verwalten Metasuchmaschinen keinen eigenen Index. An die Stelle der Crawler-Komponente tritt bei ihnen die Schnittstelle zu verschiedenen Suchmaschinen, die von der Metasuchmaschine abgefragt wird.

Die verschiedenen Suchergebnisse werden anschließend von der Metasuchmaschine unifiziert. Das Ziel ist die Optimierung der Retrievaleffektivität, da anzunehmen ist, dass aufgrund der geringen Überlappungsraten in den Inhalten einzelner Suchmaschinen eine solche Kombination von Ergebnissen die

Verbesserung der Suchresultate nach sich zieht. Dazu fehlen allerdings bisher empirische Studien. Eine Metasuchmaschine verfügt über Suchagenten für jede Suchmaschine und übersetzt die Anfrage des Benutzers in die jeweilige Abfragesprache der Einzelsuchmaschine, was zur Folge hat, dass Metasuchmaschinen oft nur eine im Vergleich zu einzelnen Suchmaschinen eingeschränkte Mächtigkeit der Anfragesprache aufweisen. Zu den bekannten Metasuchmaschinen zählen *SavvySearch* (<http://www.savvysearch.com>, vgl. DREILINGER 1996), *MetaCrawler* (<http://www.metacrawler.com>) und *MetaGer* (<http://www.metager.de>, für den deutschsprachigen Bereich).

## 2.2 Indexierung und Abdeckungsraten

Bei der Indexierung des World Wide Web durch Suchmaschinen entsteht analog zur Architektur eines Information-Retrieval-Systems eine invertierte Datei (Index). Die Analyse erfolgt auf der Basis des Volltexts jeder Webseite; zusätzlich können Strukturmerkmale (HTML-Marken für Schlagwörter, Überschriften, Hervorhebungen) bei der Indexierung berücksichtigt sein. Die invertierte Datei enthält als Beschreibungsmerkmale für jedes Dokument in der Regel nicht nur die Adresse (URL) selbst, sondern darüber hinaus eine Kurzbeschreibung des Inhalts (Textanfang, Überschrift, Textextrakt) und formale Angaben (z. B. letztes Aktualisierungsdatum). Zu den Leistungsmerkmalen von Suchmaschinen gehören ihre

- Datenbankgröße und ihr Abdeckungsgrad hinsichtlich des WWW,
- die Qualität und Aktualität der gespeicherten Daten,
- die Anfragemerkmale und das Retrievalmodell sowie
- die Art der Ergebnisaufbereitung.

	Anzahl der indexierten Webseiten			
	Search Engine Showdown		Search Engine Watch	
Suchmaschine	XI/99	VII/00	XI/99	VI/00
Northern Light	200	282	189	265
Fast Search	192	327	200	340
AltaVista	191	331	250	350
Google!	126	355	85	560
Anzwers	79	108	—	—
j Won	78	356	—	—
Excite	71	159	150	250
Snap	50	278	—	—
HotBot	39	280	—	—

**Tabelle 1:** Datenbankumfang bekannter Suchmaschinen (Millionen URLs)

Aktuelle Statistiken zeigen, dass die umfangreichsten Indizes von Suchmaschinen eine Größenordnung von mehreren hundert Millionen Einträgen erreichen. Tab. 1 stellt dabei die Daten aus zwei bekannten Datenquellen für den Suchmaschinenvergleich (NOTESS 1999f, SULLIVAN 1990 für den Zeitraum des Testdesigns (XI/99) sowie aktuelle Daten für Mitte 2000 gegenüber.

Untersuchungen von NOTESS 1999f und LAWRENCE & GILES 1998 ergeben, dass diese Indizes *weitgehend disjunkt* sind, d. h. der indexierte Datenbestand differiert stark zwischen den verschiedenen Suchmaschinen. LAWRENCE & GILES 1998, 1999 kommen zu folgenden Ergebnissen: Derzeit deckt keine Suchmaschine mehr als 16 % des World Wide Web ab; generell werden Sites, auf die viele externe Verknüpfungen verweisen, bei der Indexierung berücksichtigt. Hinsichtlich der Domänen gilt dies auch für kommerzielle Sites (. com) sowie in den USA betriebene Webserver. Eine ebenfalls hohe Schwankungsbreite existiert bei den „toten Links“, d. h. nachgewiesenen HTML-Seiten, die aufgrund des Alters des Index bereits nicht mehr existieren. Je nach Suchmaschine ist von einem Anteil zwischen 3 % und 20 % auszugehen (vgl. NOTESS 1999f, LEIGHTON & SRIVASTAVA 1999: 880 f., Appendix B, C). Mangelnde Abdeckungsraten bei gleichzeitiger Verschiedenartigkeit der indexierten Dokumentmengen ergeben eine starke Grundplausibilität für das Potential von Metasuchmaschinen hinsichtlich der Optimierung der Rechercheeffektivität im WWW. Diese Beobachtung ist daher die primäre Motivation für die in Kap. 3 beschriebene Evaluierungsstudie.

## 2.3 Retrievalmodelle und Anfragesprachen

Die von den Suchmaschinen verwandten Retrievalmodelle bauen auf <sup>B</sup>ooleschen und statistischen Verfahren auf; in der Regel besteht für den Benutzer sowohl die Möglichkeit, Operatoren der Booleschen Logik bei der Anfrage zu verwenden, als auch eine natürlichsprachliche Eingabe (mit praktisch beliebiger Länge) ohne Verwendung von Suchoperatoren zu verwenden. Tabelle 2 fasst Merkmale der Anfragesprachen und ihre Ausprägungen zusammen (nach NOTESS 1999f vgl. auch SCHWARTZ 1998: 975 f.).

<i>Merkmal</i>	<i>Mögliche Ausprägungen</i>
Defaultstrategie	Konjunktion, Disjunktion
Boolesche Operatoren	AND, OR, NOT, (), sowie die vereinfachte Syntax: +, -
Abstandsoperatoren	NEAR, exakte Phrase
Trunkierung	Rechtstrunkierung, Einzelzeichenmaskierung
Morphologische Expansion	v. a. automatische Pluralbildung im Englischen (selten)
Suche in Feldern	Titel, URL, Website
Inhaltliche und sprachliche	Dokumenttyp, Sprache, Datum (Sprachheuristik i. d. R. über <i>top level domains</i> )
Stoppworteliminierung	ja/nein

**Tabelle 2:** Recherchemöglichkeiten und Anfragesprachen von Suchmaschinen

Betrachtet man die Ausprägungen der in Tab. 2 genannten Merkmale, so fällt auf, dass sich bisher kein Konsens hinsichtlich der Interpretation von Benutzeranfragen *ohne Operatorenverwendung* herausgebildet hat: Die Suchmaschinen verwenden unterschiedliche Default-Strategien bei der Anfrageinterpretation. Dies ist vor allem deswegen bemerkenswert, weil Benutzer in der Regel kurze Anfragen ohne Operatoren stellen. Soweit verschiedene Suchmaschinen dies logisch unterschiedlich interpretieren, kann sich kein allgemein akzeptiertes Interpretationsmodell für Suchanfragen herausbilden, mit der Folge, dass Operatoren vielfach falsch eingesetzt werden (vgl. JANSEN, SARACEVIC & SPINK 2000).

## 2.4 Evaluierung von Suchmaschinen: state-of-the-art

Die Mehrzahl der bisher vorliegenden Studien zur Retrievaleffektivität von Suchmaschinen beschränkt sich auf einfache Maße wie die Anzahl der nachgewiesenen Treffer hinsichtlich einer Anfrage, ohne aber eine

Effektivitätsbewertung im engeren Sinn durchzuführen.<sup>1</sup> Recherchen im World Wide Web stellen im Kontext des IR insofern ein Novum dar, als ein wesentlich größerer Nutzerkreis ohne technische Vorkenntnisse mit Suchmaschinen als Information Retrieval-Systemen interagiert, vgl. dazu KIRSCH 1998, JANSEN et al. 1998 und SILVERSTEIN 1999. KIRSCH 1998: 4

berichtet beispielsweise, dass für die Suchmaschine *Infoseek* nur 1 % aller Anfragen mit Hilfe von *advanced search features* erfolgen. Intelligente Anfrageunterstützungsmechanismen, wie sie zeitweilig von *AltaVista* oder *Lycos* angeboten worden waren, sind mittlerweile wieder aus dem Angebot verschwunden, vgl. SCHWARTZ 1998: 977, Abb. 1: *Lycos Pro Power Panel* und 980: Abb. 2: *AltaVista Refine*. Nach ersten Studien lässt sich ein „typischer Suchmaschinennutzer“ wie folgt charakterisieren:

- Es werden kurze Anfragen gebildet (weniger als drei Suchbegriffe im Mittel, vgl. JANSEN et al. 1998, JANSEN, SPINK & SARACEVIC 2000).
- BOOLEsche Operatoren werden kaum verwendet (weniger als 10 % der Anfragen), positive und negative Auftretensoperatoren (+/-) in derselben Größenordnung. Zudem enthalten zahlreiche Anfragen Fehler hinsichtlich der Verwendung von Operatoren (JANSEN, SPINK & SARACEVIC 2000: 217).
- Benutzer betrachten nur selten mehr als zwei Ergebnisseiten (ein *cut-off-Wert* bei der Bewertung der Suchergebnisse von weniger als 30, meist wird nur die erste Ergebnisseite betrachtet). SILVERSTEIN et al. 1999: 10 analysieren mehrere hundert Millionen Anfragen und berichten, dass bei etwa für 85% aller Anfragen nur die erste Ergebnisseite betrachtet wird, bei JANSEN, SPINK & SARACEVIC 2000: 215 liegt dieser Wert bei „nur“ 58%.

Eine umfassende benutzerbezogene Studie zur Retrievaleffektivität haben GORDON & PATHAK 1999 vorgelegt. Sie führen in einem Experiment mit 33 Testpersonen eine vermittelte Recherche an acht verschiedenen Suchmaschinen durch, wobei die Suchergebnisse von den Testpersonen nach einer vierstufigen Skala hinsichtlich ihrer Relevanz bewertet werden. Die Ergebnisse zeigen eine hohe Schwankungsbreite der Suchmaschinen sowohl hinsichtlich der *precision* als auch des *recall*. Bei einem cut-off-Wert von 20 betrachteten Dokumenten erreicht die qualitativ beste Suchmaschine (*AltaVista*) einen *recall* von nur 15 %, selbst bei 200 betrachteten Dokumenten steigt der Wert nicht über 23 % an (GORDON & PATHAK 1999: 160 ff.).

Auch in dieser Studie wird die bereits erwähnte Beobachtung geringer Überlappung der Ergebnismengen verschiedener Suchmaschinen bestätigt: Von insgesamt 160 Treffern (die ersten 20 Treffer von acht Suchmaschinen) wurden 150 Dokumente jeweils nur von einer Suchmaschine nachgewiesen (GORDON &

---

<sup>1</sup> Ein Überblick über bisherige Evaluierungsstudien für Suchmaschinen im World Wide Web findet sich bei GORDON & PATHAK 1999: 145 ff., insb. 148: Tabelle

PATHAK 1999: 170 ff.). Die bisherigen Ergebnisse zur Bewertung von Suchmaschinen und ihrer Retrievaleffektivität geben eine starke Anfangsplausibilität für die Verwendung von Metasuchmaschinen sowie für die Weiterentwicklung der Retrievalfunktionalität hinsichtlich besserer Benutzerschnittstellen und der Auswertung weiterer Strukturmerkmale der indexierten Texte. Die nachfolgende vergleichende Evaluierung widmet sich daher u. a. der Frage, ob mit Metasuchmaschinen bessere Effektivitätswerte erreicht werden können als mit "einfachen" Suchmaschinen.

### 3. Evaluierung von Such- und Metasuchmaschinen

Ausgehend von den bisherigen Ergebnissen zur Evaluierung von Suchmaschinen wurde im WS 1999 / 2000 im Rahmen einer Vorlesung zum Thema Information Retrieval am Institut für Informatik der Universität Leipzig eine empirische Untersuchung zur Evaluierung von Suchmaschinen im World Wide Web durchgeführt. Die Versuchspersonen waren Teilnehmer dieser Lehrveranstaltung, d. h. Informatik-Studenten im Hauptstudium mit vertieftem Hintergrundwissen zum Information Retrieval, insbesondere auch zur Problematik der Such- und Metasuchmaschinen. Die Kenntnisse im IR-Bereich umfassen dabei auch traditionelle IR-Systeme und deren Abfragesprachen (z. B. *stn* / *INSPEC* / *messenger*), allerdings kaum als Erfahrung aus *aktivem Umgang* mit solchen Systemen.

#### 3.1 Aufbau und Durchführung der Untersuchung

Ein wesentliches Erkenntnisziel der Untersuchung war der Vergleich der *Retrievaleffektivität* von Such- und Metasuchmaschinen. Ausgehend von den Beobachtungen zur Abdeckungsrate von Suchmaschinen (vgl. Kap. 2.2), ist es naheliegend, zu untersuchen, ob Metasuchmaschinen durch Verbindung verschiedener Suchmaschinen — und damit deutlich unterschiedlicher indexierter Untermengen des World Wide Web — zu besseren Retrievalergebnissen kommen als Suchmaschinen.

Eine zweite Fragestellung ist die Unterscheidung von Anfragetypen nach fachspezifischen Anfragen (hier: mit Bezug zu Fragestellungen aus der Informatik) und nach Fragen von allgemeinem persönlichen Interesse der Testpersonen. Da es bisher keine differenzierte Theorie der Typisierung von Informationsbedürfnissen im World Wide Web gibt, erscheint eine solch einfache Aufteilung zulässig. Unterscheidungen wie sie BAEZA-YATES und & RIBEIRO-NETO 1999: 91 f. treffen (*specific queries*, *broad queries*, *vague queries*), haben kaum eine ausreichende Trennschärfe. Erste Benutzerstudien für das Web Retrieval, wie GORDON & PATHAK 1999, nehmen ebenfalls keine Typisierung der Anfragen vor, sondern zeigen lediglich die thematische Bandbreite ihrer Testanfragen auf (vgl. GORDON & PATHAK 1999: 150, Tabelle 2 und 177 (Beispielanfrage) und LEIGHTON & SRIVASTAVA 1999).

### 3.1.1 Testkonzeption

Die Evaluierungsstudie ist als zweistufiges paper-and-pencil-Experiment angelegt: Ähnlich der von GORDON & PATHAK 1999 vorgelegten Studie orientiert sich das Versuchskonzept dabei an der — für die Recherche im World Wide Web an sich nicht üblichen — vermittelten Recheresituation (vgl. auch MIELKE 2000: 108ff): Den Versuchspersonen lag dabei eine aus BERGMANN et al. 1999:3f entlehnte Übersicht von Hinweisen zur Formulierung von Suchanfragen für WWW-Suchmaschinen vor.

In einem ersten Schritt wurden die Versuchspersonen gebeten, unter Anwendung eines vorgegebenen Operatoreninventars (+, -, \*, ,, ") Suchanfragen schriftlich zu formulieren. Dabei war je eine Anfrage von *fachlichem Interesse* (z. B. mit Bezug zu einer gerade besuchten Lehrveranstaltung) sowie eine Anfrage *mit persönlichem Interessenschwerpunkt* zu formulieren. Neben der Formulierung von Suchanfragen mit Standardoperatoren für Suchmaschinen wurden die Versuchspersonen auch gebeten, Suchformulierungen als natürlichsprachlichen Text sowie als Boolesche Anfrageausdrücke (für den Test erweiterter Recherchemodi der Suchmaschinen vorzunehmen). Die ausformulierte Variante der Anfragen diente dabei der inhaltlichen Kontrolle der formalisierten Anfragefassungen, d. h. für Korrekturen bei der Testdurchführung. Zusätzlich zu den Suchanfragen wurden auf einem Fragebogen einige allgemeine Parameter erfasst (Erfahrungen im Umgang mit Suchmaschinen, bevorzugte *search engines*, durchschnittliche wöchentliche Online-Zeit, etc., s. u. Kap. 3.2.1).

### 3.1.2 Recherchedurchführung

Für die Evaluierung wurden vier Suchmaschinen ausgewählt, je zwei einfache Suchmaschinen und zwei Metasuchmaschinen, von denen je eine seit längerem verfügbar ist bzw. jüngeren Datums ist. Für die Auswahl der Suchmaschinen wurden folgende Kriterien herangezogen:

- Abdeckungsrate bzw. Umfang des Index (vgl. oben Tab. 1)
- Kompatibilität der für die Suchmaschine verwendbaren Operatoren
- Anzahl der angesprochenen Suchmaschinen (nur bei Metasuchmaschinen).

Nach diesen Kriterien wurden folgende Suchmaschinen ausgewählt:

Suchmaschinen	Metasuchmaschinen
AltaVista, <a href="http://www.altavista.com">www.altavista.com</a>	MetaCrawler, <a href="http://www.metacrawler.com">www.metacrawler.com</a>
Northern Light, <a href="http://www.northernlight.com">www.northernlight.com</a>	C4, <a href="http://www.c4.com">www.c4.com</a>

**Tabelle 3:** Für den Test ausgewählte (Meta-)Suchmaschinen

Bei der Konfiguration der Metasuchmaschinen vor Durchführung der Recherchen wurde die maximale Zahl einzelner Suchmaschinen berücksichtigt. Dabei waren



jeweils auch die im Test als „Einzelsuchmaschinen“ berücksichtigten Recherchedienste *AltaVista* und *Northern Light* enthalten.

Die formalisierten Anfragen der Testpersonen wurden vom Testleiter — ggf. unter Bereinigung offensichtlicher Fehler<sup>2</sup> — für die Abfrage von Suchmaschinen verwendet. Auf eine gesonderte Evaluierung der Anfragen mit erweiterter Boolescher Logik musste verzichtet werden, da diese nicht in hinreichender Zahl vorlagen. Von den Ergebnismengen der einzelnen Suchmaschinen gingen jeweils nur bis zu 30 Suchergebnisse in qualitätsorientierter Sortierung (*ranking*) in die Auswertung ein. Der *cut-off*-Wert 30 ist für vergleichbare Studien üblich (vgl. MIELKE 2000: 132ff) und lässt sich auch mit Blick auf bisherige Studien zur Suchmaschinennutzung begründen, nach denen selten mehr als zwei Ergebnisseiten zu je 10 Treffern betrachtet werden (vgl. oben Kap. 2.3).

Während bei den Suchmaschinen angenommen werden kann, dass die Ergebnismengen bezüglich einer Anfrage über einen gewissen Zeitraum hin stabil sind, hängt die Ergebnismenge einer Metasuchmaschine deutlich stärker von kurzfristig variablen Parametern wie Übertragungsgeschwindigkeit, Verfügbarkeit der (*primären*) Suchmaschinen und eingestellten *time-out*-Werten ab. Daher wurde für die Metasuchmaschinen jede Anfrage solange wiederholt, bis sich ein Maximum der Treffermenge für eine Anfrage beobachten ließ, d. h. jede Anfrage wurde bis zu 15 mal wiederholt, bis die jeweils höchste Trefferzahl mehrfach aufgetreten war. Von dieser umfangreichsten Treffermenge wurden anschließend die ersten 30 Dokumente zur Auswertung herangezogen.

### 3.1.3 Effektivitätsbewertung

Die Ergebnismengen der Recherchen, also die Trefferlisten der einzelnen Suchmaschinen, wurden jeweils für die einzelnen Anfragen zusammengefasst, einheitlich formatiert und den Testpersonen als Ausdruck zur Auswertung vorgelegt. Grundlage der Bewertung war dabei *nicht* das nachgewiesene Dokument selbst, sondern die auf den Ergebnisseiten der Suchmaschinen angegebene *Metainformation* (i. d. R. eine URL, Dokumenttitel sowie z. T. Textausschnitte aus dem Trefferdokument). Diese nicht unerhebliche Einschränkung der Auswertungssituation rechtfertigt sich zum einen aus dem großen Aufwand, der mit Ausdrucken und Verteilen von bis zu 4400 Webseiten verbunden gewesen wäre, zum anderen ist die Fragestellung, welche Effektivitätsergebnisse sich aus der Bewertung dieser Metainformation ableiten lassen, für sich genommen legitim.

Als Konsequenz aus dieser Einschränkung wurde nicht die traditionelle binäre Bewertung *relevant (rel.)* — *nicht relevant (n. rel.)* für die Effektivitätsbewertung verwendet, sondern als dritte Kategorie die Bewertung als *vielleicht relevant (v. rel.)* hinzugenommen, vgl. dazu SALTON & MCGILL 1987: 151f. Letztere Kategorie sollte dabei in denjenigen Fällen angewandt werden, in denen sich die

---

<sup>2</sup> Z. B. Korrektur von Rechtschreibfehlern, Eliminierung redundanter Klammerung und Korrektur von Operatoren. Zur Motivation solcher Eingriffe diene dabei der natürlichsprachliche „Volltext“ der Anfragen, der die intendierte Interpretation deutlich werden lässt.

Testpersonen zu einer sicheren Relevanzbeurteilung in Ermangelung des Trefferdokuments nicht in der Lage sahen, aber ggf. zu der entsprechenden Seite im WWW navigieren würden. Die nachfolgenden Ergebnisse sind auf der Basis dieser Vorgehensweise konservativ zu interpretieren, da anzunehmen ist, dass sich bei Bewertung der Dokumente selbst die Relevanzbewertungen verschlechtern würden.

### 3.2 Testergebnisse

Die Ergebnisse der Evaluierung umfassen zunächst die Daten zur Selbsteinschätzung der Versuchspersonen sowie die Relevanzbewertungen für die einzelnen Suchanfragen.

#### 3.2.1 Fragebogenauswertung

Tab. 4 fasst die Auswertung des Fragebogens zu Vorkenntnissen und Erfahrungen der Versuchspersonen zusammen:

<i>Parameter</i>	<i>Ergebnis</i>
Anzahl Versuchspersonen	25
Mittlere wöchentliche Online-Zeit	0 10h, Minimum: 1h, Maximum 30h
Nutzung des WWW (0=nie ... 5=täglich)	0 4,14
Nutzung von Suchmaschinen (0=nie ... 5=täglich)	0 2,98
Bevorzugte Suchmaschinen (mehr als einmal genannt, Anzahl der Nennungen)	Altavista (21), Yahoo (15), Fireball (6), Google (4), Lycos (4), Metacrawler (3), Excite (2), HotBot (2), Metager (2), Web.de (2)
Kenntnisse von HTML (0=keine ... 5=sehr gute)	0 3,04
Eigene Homepage	14 (= 56 %)

**Tabelle 4:** Ergebnisse der Auswertung des Fragebogens

Die Ergebnisse des Fragebogens bestätigen die Annahme, dass es sich bei der ausgewählten Personengruppe um relativ erfahrene Benutzer von Suchmaschinen handelt, denen deren Funktionsweise gut vertraut ist. Nimmt man hinzu, dass die Versuchspersonen gezielt in die Arbeitsweise von Suchmaschinen und die verfügbaren Suchoperatoren eingeführt wurden, so sind die nachfolgenden Ergebnisse dahin zu bewerten, dass man deutlich schlechtere Effektivitätswerte für „durchschnittliche Websurfer“ erwarten kann. Dafür spricht auch die Analyse

der Suchanfragen hinsichtlich der Anzahl der verwendeten Konzepte<sup>3</sup> und Operatoren (Tab. 5): Sie unterscheidet sich deutlich von den Ergebnissen der Auswertung umfangreicher *query logs* von Suchmaschinen, wie sie etwa SILVERSTEIN et al. 1999 vorgelegt haben; dort kommen Operatoren in weit weniger als 10% aller Anfragen zum Einsatz.

Mittlere Anzahl Suchkonzepte	3,96
Anwendung des +-Operators	93%
Anwendung der Rechtstrunkierung (*)	41%
Anwendung des —Operators	20%
Kennzeichnung von Phrasen	20%

**Tabelle 5:** Ergebnisse der Analyse der Anfragen

### 3.2.2 Auswertung der Relevanzbewertungen

Bei der Auswertung der Relevanzbewertungen konnten 41 Anfragen berücksichtigt werden, davon 21 fachlich motivierte Anfragen und 20 mit persönlichen Interessenschwerpunkt. Sie enthalten insgesamt 4409 Einzelbewertungen (*rel.* / *v. rel.* / *n. rel.*). Aus ihnen lassen sich *precision* und *recall* als elementare Effektivitätskennzahlen errechnen. Aufgrund der relativ

schwach<sup>en</sup> Datenbasis kommt dabei der Abschätzung des *recall* eine untergeordnete Bedeutung zu. Tab. 6 a/b zeigt jeweils Werte für die *precision* als Anteil ermittelter relevanter an allen nachgewiesenen Dokumenten.<sup>4</sup>

	AltaVista			C4			MetaCrawler			Northern Light		
	Anz.	%	%	Anz.	%	%	Anz.	%	%		%	%
<i>n. rel.</i>	693	63,23	63,23	836	69,10	69,10	882	77,64	77,64	700	69,17	69,17
<i>v. rel.</i>	261	23,81	(v + r)	248	20,5	(v + r)	156	13,73	(v + r)	198	19,57	(v + r)
<i>rel.</i>	142	12,96	36,77	126	10,41	30,91	98	8,63	22,36	114	11,26	30,83
$\Sigma$	1096	100	100	1210	100,01	100	1136	100	100	1012	100	100

**Tabelle 6a:** Gesamtergebnisse precision nach Suchmaschinen

<sup>3</sup> Als *Suchkonzept* werden hier *inhaltlich unterschiedliche* Suchbegriffe innerhalb einer Anfrage gewertet, nicht aber die Verwendung von Synonymen oder unterschiedlicher flektierter Formen eines Suchbegriffs.

<sup>4</sup> Es werden *recall* und *precision* als die klassischen Maße der Effektivitätsbewertung<sup>ver</sup>wendet, vgl. dazu TAGUE 1981:66ff, WOMSER-HACKER 1989:34ff.

Alle Suchmaschinen			
	Anzahl	%	%
nicht relevant (n. rel.)	3111	70,56	70,56
vielleicht relevant (v rel.)	818	<b>18,55</b>	<b>(v + r) 29,44</b>
relevant (rel.)	480	<b>10,89</b>	
Summe	4409	100	100

**Tabelle 6b:** Gesamtergebnisse precision (addiert)

Tabelle 6 a/b zeigt keinen klaren Vorteil für Metasuchmaschinen; das mit Abstand beste Ergebnis liefert mit 13% bzw. 37% *AltaVista*, d. h. bei weiter Interpretation werden zwei von fünf Dokumentennachweisen als *relevant* oder *vielleicht relevant* eingestuft. Angesichts der Tatsache, dass die ausgewählten Metasuchmaschinen auch jeweils beide im Test verwendeten „Einzelsuchmaschinen“ abfragen, deutet dies darauf hin, dass es den Metasuchmaschinen bei der Zusammenführung der Einzelergebnislisten nicht gelingt, über das Qualitätsniveau der Einzellisten hinaus zu kommen.

Die Tabellen 7 und 8 zeigen die Differenzierung der Auswertung zur *precision* nach Interessenschwerpunkt der Anfrage und bei cut-off-Wert 10, d. h. bei Betrachtung jeweils nur der obersten 10 Dokumente jeder Ergebnismenge. Bei der Differenzierung nach Interessenschwerpunkte zeigt sich kein wesentlicher Unterschied, was — auch gestützt durch den für die Inhalte des WWW günstigen fachlichen Hintergrund der Versuchspersonen — wenigstens andeutungsweise den Schluß zuläßt, dass die im WWW recherchierbaren Inhalte tatsächliche für eine sehr weite Bandbreite inhaltlicher Fragestellungen geeignetes Material bereithalten. Dafür spricht auch die Tatsache, dass zu *allen* Anfragen mindestens ein wenigstens *vielleicht relevanter* Nachweis ermittelt werden konnte; bei drei Anfragen konnte keine Suchmaschine ein als *relevant* eingeschätztes Dokument nachweisen.

	Gesamt			fachliches Interesse			privates Interesse		
	Anzahl	%	%	Anzahl	%	%		%	%
n. rel.	3111	70,56	70,56	1575	71,66	69,17	1536	69,47	69,47
v. rel.	818	<b>18,55</b>	<b>(v+r) 29,44</b>	419	<b>19,06</b>	<b>(v+r) 28,34</b>	399	<b>18,05</b>	<b>(v+r) 30,53</b>
rel.	480	<b>10,89</b>		204	<b>9,28</b>		276	<b>12,48</b>	
Summe	4409	100	100	2198	100	100	2211	100	100

**Tabelle 7:** precision-Werte für fachliche / private Interessenschwerpunkte der Anfragen

Der Anstieg der *precision* bei *cut-off* 10 um etwa 50 % gegenüber dem Wert bei *cut-off* 30 belegt einerseits das Funktionieren des *ranking* (relevante Dokumente befinden sich an den vorderen Plätzen in der Ergebnismenge), erreicht aber insgesamt dennoch keinen zufriedenstellenden Wert (Tab. 8).

Alle Anfragen cut-off 10			
		%	%
nicht relevant (n. rel.)	1005	64,46	69,47
vielleicht relevant (v. rel.)	319	<b>20,46</b>	<b>35,53</b>
relevant (rel.)	235	<b>15,07</b>	
Summe	1559	100	100

**Tabelle 8:** precision-Werte für cut-off-Wert Wert 10

Die Berechnung des mittleren *recall* als Anteil der **nachgewiesenen relevanten** in Bezug auf **alle relevanten** Dokumente der Kollektion, d. h. des WWW, ist nicht direkt möglich, da die Anzahl aller relevanten Dokumente zu einer Anfrage nicht bekannt ist.<sup>5</sup> Daher wird nach einer pooling-Methode vorgegangen, bei der die Gesamtzahl der zu einer Anfrage ermittelten relevanten Dokumente als Bezugsgröße verwendet wird (vgl. HARTER 1996: 37).<sup>6</sup> Aufgrund der geringen Anzahl untersuchter Systeme pro Anfrage ist diese Versuchsgröße vermutlich noch relativ weit von der tatsächlichen Gesamtmenge der relevanten Dokumente entfernt. Die ermittelten *recall*-Werte können daher ausschließlich für den Vergleich der Systeme *untereinander* herangezogen werden.

Wie Tab. 9 zeigt, ist hinsichtlich der Unterscheidung von Suchmaschinen und Metasuchmaschinen auch mit Hinsicht auf den *recall* keine klare Tendenz zu erkennen: Wie bei den Ergebnissen für die *precision* liefern je eine Such- und eine Metasuchmaschine (*AltaVista* / *C4*) die jeweils besten Ergebnissen. Die *recall*-Werte stehen zudem in Einklang mit den in anderen Evaluierungsstudien beobachteten Werten.<sup>7</sup>

Suchmaschine	<i>recall</i> (nur rel.)	<i>recall</i> (rel. + v. relevant)
<i>AltaVista</i>	0,235	0,298
<i>Northern Light</i>	0,231	0,212
<i>C4</i>	0,256	0,296
<i>MetaCrawler</i>	0,204	0,193
<b>Gesamt</b>	0,233	0,251

**Tabelle 9:** recall-Werte bei cut-off-Wert 30

<sup>5</sup> Dies wird durch die Tatsache gestützt, daß die von einzelnen Suchmaschinen erfassten Teilmengen des WWW weitgehend disjunkt sind, vgl. oben Kap. 2.2.

<sup>6</sup> Eine Feindifferenzierung jeweils nach einzelnen Anfragen wäre hier wünschenswert.

<sup>7</sup> Die Werteberechnung für den *recall* erfolgt durch Makromittelung, vgl. WOMSER-HACKER 1988: 67ff, MIELKE 2000: 135. s

## 4. Fazit & Ausblick

Die Ergebnisse der vorliegenden Untersuchung decken sich von ihrer Tendenz her mit bisherigen Evaluierungsstudien zur Retrievaleffektivität von Suchmaschinen im World Wide Web und machen deutlich, dass sich diese Ergebnisse auch auf die Betrachtung von Metasuchmaschinen übertragen lassen. Dabei hat in der vorliegenden Studie auch die verstärkte Verwendung von Suchoperatoren keinen wesentlichen Beitrag zur Effektivitätserhöhung bewirkt.<sup>8</sup> Die Annahme, durch Verwendung von Metasuchmaschinen zu besseren Ergebnissen zu gelangen, ließ sich gleichfalls nicht bestätigen. Dies bedeutet, dass die Verfahren der Zusammenführung einer Mehrzahl von Ergebnismengen, wie sie Metasuchmaschinen durchführen, zwar theoretisch den Zugriff auf eine größere Teilmenge der im WWW verfügbaren Daten möglich machen, für überschaubare Ergebnismengen (cut-off 30) aber zu keiner effektiven Verbesserung führen.

Dabei ist allerdings zu beachten, dass die schwerwiegenden methodischen Einschränkungen dieser Untersuchung (paper-and-pencil-Experiment, Bewertung der Rechercheergebnisse nur auf der Basis von Metainformation, geringe Zahl von Anfragen, Testpersonen und untersuchten Systemen) die Testergebnisse nur als erstes Indiz für die Effektivität von Metasuchmaschinen erscheinen lassen. In weiteren Evaluierungen auf einer breiteren empirischen Basis wäre neben den hier vorgestellten Ergebnissen auch zu untersuchen, wie stark sich die Ergebnismengen der Suchmaschinen überlappen und ob Metasuchmaschinen hier ein anderes Verhalten aufweisen. Wünschenswert wären dabei auch Untersuchungen mit Testpersonen ohne entsprechenden fachlichen Hintergrund.

Die insgesamt relativ schlechten Effektivitätswerte bei einer angenommenen hohen "Dunkelziffer" zusätzlicher relevanter Dokumente, die mit weiteren Suchmaschinen nachzuweisen wären, lassen es erforderlich erscheinen, die Arbeitsweise sowie die Benutzerschnittstellen von Suchmaschinen und Metasuchmaschinen zu verbessern. Denkbare Ansätze hierfür sind

- der verstärkte Einsatz von Visualisierungstechniken, z. B. um Zusammenhänge zwischen Suchbegriffen darzustellen,
- die verstärkte Individualisierung der Recherche durch Adaption des Suchwerkzeugs an typische Informationsbedürfnisse des Benutzers und
- der Einsatz von *relevance feedback-Verfahren*, wie sie von einigen Suchmaschinen bereits angeboten werden.

---

<sup>8</sup> Hier ist darauf hinzuweisen, dass Studien zur Operatorenverwendung i. d. R. auf Auswertung großer *query logs* ohne *empirische Evaluierung* beruhen, d. h. ein Zusammenhang zwischen Operatorenverwendung und Effektivität ist bisher nicht untersucht.

## 5. Literatur

BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier (1999). Modern Information Retrieval. Harlow et al.: The ACM Press/The MIT Press.

BERGMAN, Michael et al. (1999). Search Tutorial: Guide to Effective Searching of the Internet. Vermillion/SD: VisualMetrics Corporation, Juli 1999, <http://thewebtools.com/tutorial/index.htm>.

CAGLAYAN, Alper K.; HARRISON, Colin G. (1998). Intelligente Software-Agenten. Grundlagen, Technik und praktische Anwendung im Unternehmen. München & Wien: Hanser.

DREILINGER, Daniel E. (1996). "Description and Evaluation of a Meta-Search Agent." M. Sc. Thesis, Colorado State University, Department of Computer Science, Fort Collins/CO, Oktober 1996.

GORDON, Michael; PATHAK, Praveen (1999). „Einding Information on the World Wide Web: the Retrieval Effectiveness of Search Engines." In: Information Processing & Management 35 (1999), 141-180.

HARTER, Stephen P. (1996). „Variations in Relevance Assessments and the Measurement of Retrieval Effectiveness." In: Journal of the American Society for Information Science 47 (1996), 37-49.

JANSEN, Bernard J. et al. (1998). „Real Life Information Retrieval: A Study of User Queries On the Web." In: SIGIR Forum 32(1) (1998), 5-17.

JANSEN, Bernard J.; SPINK, Amanda; SARACEVIC, Tevko (2000). "Real Life, Real Users, and Real Needs: A Study and Analysis of User Queries on the Web." In: Information Processing & Management 36(2) (2000), 207-227.

KIRSCH, Steve (1998). „Infoseek's Experiences Searching the Internet." In: SIGIR Forum 32(2) (1998), 3-7.

LAVOIE, Brian; FRYSTYK NIELSEN, Henrik (1999). Web Characterization Terminology & Definitions Sheet. World Wide Consortium Working Draft, Mai 1999, <http://www.w3.org/1999/05/WCA-terms/>.

LAWRENCE, Steve; GILES, C. Lee (1998). „Searching the World Wide Web." In: Science 280 (1998), 98-100.

LAWRENCE, Steve; GILES, C. Lee (1999). „Accessibility and Distribution of Information on the Web." In: Nature 400 (1999), 107-109 [<http://www.wwwmetrics.com/>].

LEIGHTON, H. Vernon; SRIVASTAVA, Jaideep (1999). „First 20 Precision among World Wide Web Search Engines." In: Journal of the American Society for Information Science 50(10) (1999), 870-881.

MIELKE, Bettina (2000). Evaluierung juristischer Informationssysteme. Köln et al.: Heymanns [= Ius Informationis Bd. 11].

NOTESS, Greg (1999). Search Engine Showdown. The Users' Guide to Web Searching. Dezember 1999 / September 2000, <http://www.notess.com/search/>.

OCLC (1999). June 1999 Web Statistics. Online Computer Library Center, Inc, Web Characterization Project, <http://www.ocic.org/ocic/research/projects/webstats/statistics.htm>.

PIMIENTA, Daniel et al. (1998). L4. The Fourth Study on Languages and the Internet. The Place of Latin Languages and Cultures on the Internet. September 1998. Santo Domingo: Funredes Networks and Development Foundation, <http://funredes.org/LC/english/L4prologue.html>.

PITKOW, James E. (1998). "Summary of WWW Characterizations." In: Proc. Seventh World Wide Web Conference, Brisbane, April 1998, <http://www7.scu.edu.au/program/me/fullpapers/1877/com1877.htm>

SALTON, Gerard; MCGILL, Michael J. (1987). Information Retrieval – Grundlegendes für Informationswissenschaftler. Hamburg et al.: McGraw-Hill.

SCHWARTZ, Candy (1998). „Web Search Engines.” In: Journal of the American Society for Information Science 49(11) (1998), 973-982.

SILVERSTEIN, Craig et al. (1999). "Analysis of a Very Large Web Search Engine Query Log." In: SIGIR Forum 33(1) (1999), 6-12.

SULLIVAN, Danny (1999). Search Engine Watch. Dezember 1999 / September 2000, <http://www.searchenginewatch.com>.

TAGUE, Jean (1981). „The Pragmatics of Information Retrieval Experimentation.” In: SPARCK JONES, Karen (ed.) (1981). Information Retrieval Experiment. London et al.: Butterworths, 59-102.

WOMsER-HACKER, Christa (1989). Der PADOK-Retrievaltest. Zur Methode und Verwendung statistischer Verfahren bei der Bewertung von Information-Retrieval-Verfahren. Hildesheim: Olms [= Sprache und Computer Bd. 10].





## **TransRouter revisited - Decision support in the routing of translation projects**

**Rainer Hammwöhner**

Department of Information Science

University of Regensburg

93040 Regensburg

rainer.hammwoehner@sprachlit.uni-regensburg.de

### **Abstract**

This paper<sup>1</sup> gives an outline of the final results of the TransRouter<sup>2</sup> project. In the scope of this project a decision support System for translation managers has been developed, which will support the selection of appropriate routes for translation projects. In this paper emphasis is put on the decision model, which is based on a stepwise refined assessment of translation routes. The workflow of using this system is considered as well.

### **1. Introduction**

Quality and efficiency of translation services depends on the appropriate routing of translation projects. The choice of a translation route requires knowledge about the technology at hand, the resources available and the relevant project parameters. For several reasons this information is difficult to obtain for translation managers.

- Translation technology is undergoing a swift development process. As a consequence new approaches to translation support or new systems will not be considered in decision making.

---

<sup>1</sup> A more comprehensive description of the prototype is provided by [Hammwöhner 00].

<sup>2</sup> TransRouter [King 99] is project LE4-8345 in the Telematics Applications of Common Interest programme of the Fourth Framework Programme, supported by the Commission of the European Communities and by the Swiss Federal Office for Education and Science. The members of the consortium are:

Berlitz, Dublin (Charles Hughes, John Micks), CST, Copenhagen: (Bart Jongejan, Nancy Underwood), University of Edinburgh (Jo Calder), TIM/ETI, University of Geneva (Margaret King, Sandra Manzi), Lernaut and Hauspie, Munich (Johannes Ritzke), LRC, Dublin (Keith Brazil, Conor McDonagh, Reinhard Schäler), University of Regensburg (Rainer Hammwöhner, Jürgen Reischer).



- The need of an organisational memory covering the resources (e.g. translation memories, term banks) at hand is not recognised by a majority of translation agencies.
- Relevant project parameters may not be estimated easily. Is the text repetitive and to which degree? Is special terminology employed? Is the text too complex to use a machine translation system?

The objective of the TransRouter project is the design of a decision support system providing the translation manager with the relevant information. Since no widely accepted notion of decision support in this application field can be built on, the general approach of TransRouter is based on the development of a sequence of prototypes that were presented to the public. These prototypes share the following features:

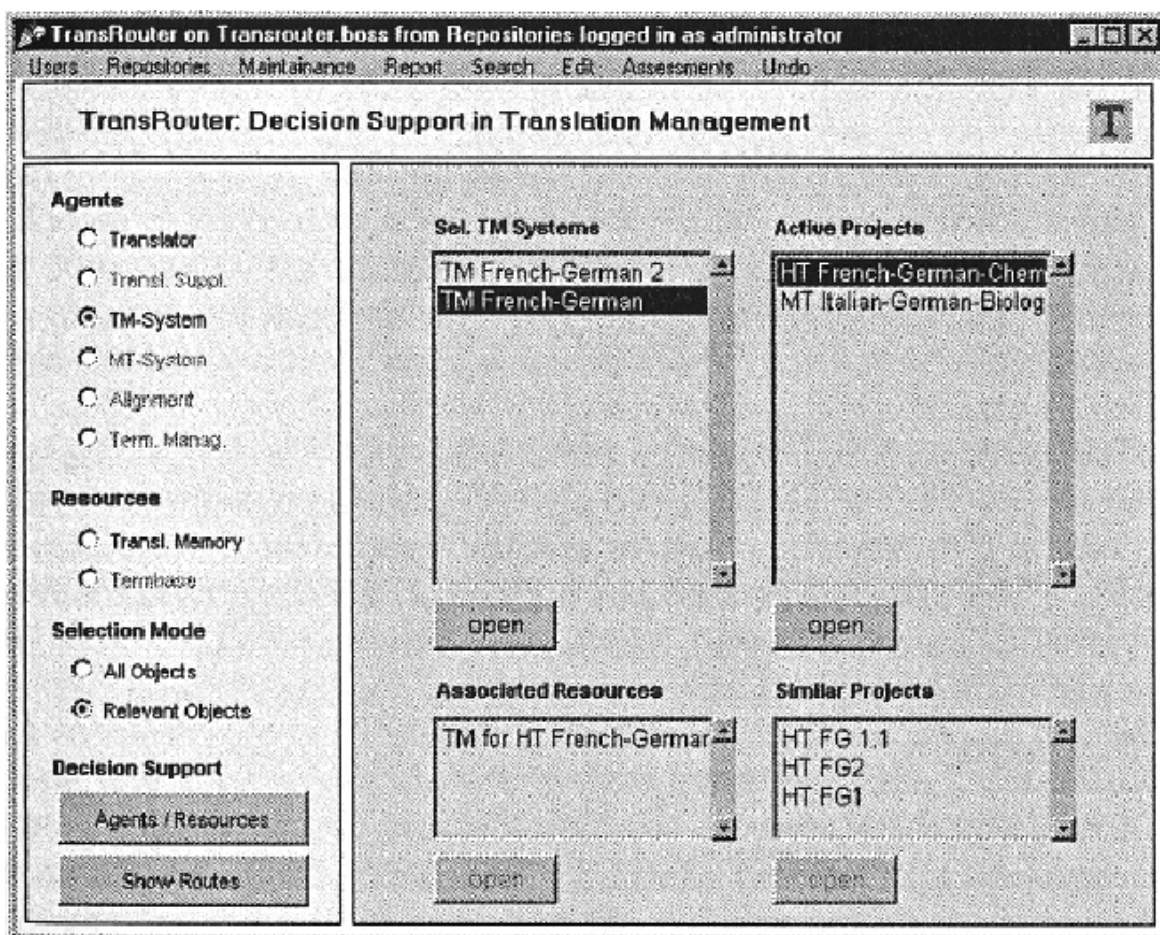
- Several profiles contain the relevant features of agents (translators or translation tools) and resources (translation memories or term banks)
- The features of translation projects are covered by a different type of profiles.
- A set of tools is developed, which allow the (semi)-automatic acquisition of project data (e.g. text size, terminology, complexity, repetitiveness).
- A decision kernel computes translation routes based on agent, resource and project profiles.

The general approach of the TransRouter project has — in an early stage of the project — been presented at the least ISI conference [Hammwöhner 98]. This paper will give an overview on the TransRouter prototype as developed at the University of Regensburg and thus give an account of the final results of the project. The task of route construction and selection consists of a couple of steps, which need specific support each. Thus, a first part of this paper will give an overview on the workflow which may be supported by the system and which is needed to operate it.

Before a detailed discussion can be started, the nature of the intended support should be pointed out<sup>3</sup>. The goal of TransRouter is not to find *the* optimal route for a translation automatically. This approach would need formalised rules of decision making in translation projects, which are not at hand. Furthermore it seems to be questionable whether translation managers would accept a system that would seem to take over the responsibility of decision making. Thus, TransRouter will support the manager in decision making by pointing out alternatives in agent and resource selection and the resulting route choice. TransRouter will give support in the acquisition of the relevant project data, the resource assessment and the assessment of routes with respect to time, costs and quality.

---

<sup>3</sup> An overview on approaches to decision support may be found in [Turban 95]



**Figure 1:** Initial page of TransRouter— lists of projects, agents and resources are provided

## 2. Workflow

This section will provide a short overview of the general workflow in Trans-Router. A discussion of more detailed workflow will be given then in the following sections.

- Data acquisition: The translation manager will expect agent and re-source data to be available in the system when he starts to use it. Nevertheless there will be a need to update the system at regular intervals.
- Furthermore there should be some means to enter project data conveniently. TransRouter offers appropriate tools for these tasks, but this pa-per will not go into detail here.
- Information retrieval: The user might want to extract data from the system without using the inference mechanism.
- Agent and resource selection: The first step of decision making is the selection of agents and resources, which may be used to process a given

project. This step may be performed automatically since it is based on formally defined criteria. Some manual editing may follow the automatic selection step.

- **Route construction:** The selected agents and resources are used to construct possible routes, which are based on a built-in route model. The system will not generate all possible routes but a representative set of routes covering all useful route types. The user can refine these routes afterwards.
- **Route assessment:** The routes, which are created by the system, have to be assessed by the user. He will choose one of several evaluation functions, a cost function, time or quality function or a combination of those. The system will then sort the routes according to those criteria. The manager will then pick one or more routes which seem to be promising. He then may modify the set of tools and resources assigned to these routes.
- **Route selection and route processing:** when the route assessment and refinement is completed a route can be selected for the further processing of the translation project. At this stage the job of the TransRouter system seems to be done. It will be shown that some additional documentation job will be helpful for the appropriate handling of further projects.

The steps as mentioned above give an outline of a macroscopic workflow in the use of TransRouter.

### **3. Information Retrieval**

The whole process of decision support of course may be seen as some kind of information retrieval. If a user wants to access the TransRouter repository directly he may choose from two retrieval options. Matching oriented retrieval is based on the formulation of a query. This query will be processed by the system, which then will come up with a list of relevant objects fitting the request. Browsing uses implicitly or explicitly given references between objects that can be navigated via the TransRouter interface.

The approach to matching oriented retrieval is based on the construction of sample objects. A user, who wants to find a machine translation system by its specific features, will have to construct an agent profile describing such a system. Truncation symbols ('#' and '\*') substituting an arbitrary character or sub-string are available. If attributes are numeric some kind of fuzzy match is performed.

Browsing oriented retrieval employs relationships between objects, which are made explicit by the interface. A profile of a translation memory may provide not only its data storage format but also list translation memory systems, which can process this data format.

One of the most promising applications of browsing in TransRouter seems to be the use of organisational experience. The translation manager will check, whether a machine translation system has already been used in successful projects. Did these projects have anything in common with the project currently in progress? Were there any projects at all, which were similar to the current one? What can be learned from their performance?

This leads to the question of which notion of similarity should be employed. Should the similarity of projects be based on the basic project attributes only or should the chosen routes be considered too? Since the similarity measure should be applicable to all projects — even if newly defined— there must be at least one notion of similarity, which does not take routes into account. The similarity measure should allow a ranking of projects. It is reasonable to have similarity values between 0 (no match) and 1 (full match). The numeric attributes of a project can be used easily to compute such a measure. Even for symbolic attributes some distance measure can be defined.

One more application of browsing oriented retrieval is the exploration of tools able to process a given resource — e.g. a translation memory or a term-bank — or to find the available resources that can be used when operating a given tool — e.g. a dictionaries used by a machine translation system (see figure 1).

#### **4. Agent and resource selection**

The first step of route construction is the selection of agents and resources, which are relevant with respect to the project profile. TransRouter is capable of handling several kinds of selection rules, which will be described in this section.

Type specific rules do not apply to individual objects but to object classes. Such a rule may express the fact that machine translation systems in general are of no use for projects with certain features. Currently the following type specific rules are built into the system:

- If there is no previous version of a project and no further version is to be expected and if the repetitiveness of the text is below a certain threshold, then no translation memory should be built.
- A machine translation system should not be used if the complexity of the text exceeds a certain threshold.

Most of the selection rules implemented within TransRouter apply to individual objects.

- The agent or resource must support the language pair required by the project.
- Human agents may act in different roles (translator, reviser etc), which represent individual translation services. The translator must know the required language pair (target language in case of revising) and be able to provide one of these services at the required quality level.

- Dictionaries and term banks must cover the same content domain as the project's text.
- A translation memory must be derived from a prior version of the project. This will prevent the system from doing costly assessments on memories, which probably will not have a sufficient coverage.
- Machine translation system must be able to provide the required quality.
- An agent must be at hand that can process the resource being in question.

Weak selection rules cover phenomena which are mere obstacles in the use of a resource or system. Examples are licences being outdated or format not matching. These obstacles will lead to the exclusion of a system or resource if and only if an alternative is at hand. Otherwise the rule will be suppressed in order to get some operational routes. A comment on the problems with these systems / resources is provided.

Licence must be up to date. If no other tool is available a licence can be updated easily.

- Tools must be able to process the storage format of the project's text. In most cases it should be possible to convert the formats with some reasonable effort.
- Tools must be able to produce the desired destination storage format (see above).
- Resources must be approved by a translation manager.
- Human translators must be knowledgeable in the content domain of the text and know about the relevant text styles. If nobody is at hand who has this knowledge, somebody knowing the languages should be able to do at least a low quality translation. A good reviser can sort out quality problems in the last step of the route.

After the completion of the automatic selection process the translation manager may want to reduce the relevance set further. He may know that some translator is occupied by other projects or for some reason does not want to use a specific tool etc. Removing agents or resources at this stage of the decision process will make the task of route construction and assessment less complicated.

## **5. Route construction**

The explanation of the route construction process first needs some introduction to the route model of TransRouter. Then it can be shown how a route will be furnished with agents and resources.

### **5.1 The route model of TransRouter**

The route model of TransRouter is comparatively simple. A route basically consists of three processing steps. Each step is performed by one main agent using a set of tools operating on a set of resources associated with this step. The pre-processing step covers all activities, which are necessary to prepare text and or resources — initial proofreading, enhancing dictionaries. The main step contains the translation process whereas the post-processing step deals with all activities following the translation until the end of the project — e.g. proofreading, formatting. Pre- and post-processing steps are performed by humans. The main agent of the translation step may be a tool as well (e.g. a machine translation system). According to the nature of the main agent the type of translation step and route will be defined. Because each type of main agent has its own requirements regarding pre- and post-processing, there are special subtypes for these steps too. The route type will have consequences on the time, cost and quality estimations as well. The agent of a translation step can make use of tools — e.g. a terminology management system or a translation memory system — which themselves will operate on data derived from special resources — term banks or translation memories. Currently TransRouter supports the following route types:

- Translation by a translator who is employed by the agency
- Translation by a freelancer
- Translation by a machine translation system
- Translation by a translation memory system (automatic mode)

Human translators can be assigned to various roles in the translation process. They can be the main agents of the translation step if the respective sub-profile covering the translation performance is filled out. Another role would be that of a reviser or post-editor — main agents for the post-processing step — requiring further sub-profiles.

### **5.1 Generation of routes, assigning tools and resources to routes**

TransRouter has some basic understanding of which kinds of agents and resources can be combined and which kinds of route steps they may be assigned to. The system will not try to generate and assess all possible combinations of main agents, tools and resources but to find some reasonable equipment for each step. This process starts with the main translation step. The system will select resources first, because the content of a term bank or a translation memory is assumed to be

prior to the effects of handling the software. In a following step the system will find the optimal tool for each of the selected resources. Finally those tools which do not need any resources (e.g. an alignment tool) will be assigned to a translation step. This approach has two implications. The system must provide the means to find a ranking of resources and tools in order to find the best fit. The solution found may not be the optimal one, because a slightly less optimal resource may be processed by some more user friendly or efficient tool, which could not be used for the resource selected. The equipment of the auxiliary steps will follow almost the same procedure with the only exception that, if possible, the same resources and tools will be used as assigned to the main step.

Not every combination of translation steps and agents is possible. A selection of rules applying is listed here:

- The agents of pre- and post-processing steps are human translators.
- The step type defines the agent type of the main step.
- If the main agent of the translation step is a tool, the main agents of the pre- and post-processing steps must be experienced in using the tool.
- The profile of a human translator must indicate that he may take the appropriate role — translator, reviser, pre- or post-editor — in the route step. This means that a processing performance greater than 0 must be assigned to this specific activity.

The most straightforward approach to the sorting of resources implies the use of resource assessments.

- Translation memories are sorted according to the coverage of the project's text.
- The sorting of term banks makes use of information on the number of unknown terms within the text.

Unfortunately the assessment of resources is time consuming. Therefore it can't be assumed that all resources which are of some relevance to the project are assessed. Thus, TransRouter has to employ two sorting strategies. If all resources of some type are assessed, TransRouter will use the assessments for sorting. If this is not the case, TransRouter has to use an alternative strategy using basic resource features for sorting.

- Translation memories are sorted according to their position in the version hierarchy. The translation memory, which is most recent with respect to the current project, will be preferred. It is most likely that this memory will have the best coverage.
- A good indicator to estimate the quality of a term bank is its size. This



largest term bank probably will have the best terminology coverage.

The sorting of agents imposes fewer problems than the sorting of resources. Basic features of the respective profiles may be used.

- Translating (translators, machine translation systems) agents are sorted according to translation quality and performance.
- The relevant features of freelancers are quality and costs.
- Translation memory systems judged according to their performance (average access and storage time).

Nevertheless, some of these data can be fully estimated only with knowledge of the complete route data. The performance of a translator for instance depends on the tools at hand, the quality of a machine translation is related to the quality of the resource being used. TransRouter will feed as much information into the sorting and ranking process as is available within the current state of decision making. In an early step only the agent profile will be available, in a next one a project profile will be added. Finally all data of the route and route steps currently being elaborated are available and can be used for agent assessment.

## **5.2 Dependencies between agents**

The sorting process as described above does not take any dependencies between agents into account. Nevertheless it seems to be quite obvious that a terminology management system which is an integral part of some other tool being used (machine translation system, translation memory system) is to be preferred to others which are not. The same applies to alignment tools or even translation memory systems. TransRouter distinguishes three levels of integration (built in, add on, compatible output). Human agents or freelancers on the other hand are more experienced in the use of some tools compared to others. These dependencies are represented in the agent's profiles and will be used in the construction of routes as follows:

- If the main agent is human, TransRouter will prefer tools that are familiar to the translator. Furthermore it will prefer tools which are able to mutually cooperate. The level of integration will be considered only if there is no severe lack of performance compared to some other tools.
- If the main agent is a tool — e.g. a machine translation system — it is requested that all tools assigned to the main translation step allow some integration with the main agent.

## **5.3 Manual modification of routes**

The system, as already has been mentioned, will not necessarily find the optimal route. But even an optimal route could be of little use, if the agents of the route

were occupied with other projects. In this case there is a need for the translation manager to modify routes suggested by TransRouter manually. He may delete entire routes or copy routes to try out different versions of the same general approach. Possible modifications of a route include the re-placement of the main agent of a step, the removal of tools or resources from a step or the assignment of additional or alternative ones. This process is governed by a set of simple rules.

- The main agent may only be replaced by an agent of the same type. The route type will be unaffected. Changing the main agent will trigger a consistency check on tools and resources assigned to that step. It is checked whether the new agent may use them. If this is not the case the resource or tool will be replaced as well.
- If an agent is removed from a step, the corresponding resource will be removed too and vice versa. This will prevent the user from constructing inconsistent translation steps containing useless tools or resources lacking an agent.
- If a new resource is added to a step, TransRouter will remove an equivalent resource (same type) from the route if present. If the agent corresponding to the replaced resource is not able to process the new resource it will be replaced too. The optimal tool, which can process the needed data format, will be chosen automatically. An equivalent process will take place if an agent is replaced. Since all steps of a route should have the same equipment if possible, these exchange processes are performed on all steps simultaneously if the new agents or re-sources are valid for all of them. Otherwise the manipulation is restricted to the explicit manipulation of a single step.

## **6. Route assessment**

The step of route construction is followed by the assessment of routes. This step will be performed automatically. The user can guide this process only by adjusting the criteria which have to be used. Generally the assessment of routes can serve different purposes (see also figure 3):

- Ranking of routes will help to find the best route with respect to a set of criteria.
- Estimation of time, cost or quality numbers will be helpful for the final planning processes. Setting a frame for time, cost and quality is an important task at the very start of a project.

In the course of the workflow supported by TransRouter the ranking of routes would be the first step. An exact estimation could be restricted to the best routes one should consider implementing. Even from the viewpoint of the de-signer of a decision support system this sequence seems to be reasonable. Where the

ranking of routes is a rather feasible task, the exact estimation of route features imposes severe methodological problems:

- Some of the relevant criteria — this is true especially for quality — are not well defined.
- The nature of translation processes up to now is not well understood. The effects of the environment — features of projects, agents and re-sources — on the translation process with respect to time, costs and quality can — in many cases — be quantified only by very rough approximations.
- There are aspects of the handling of a translation project, which are idiosyncratic to any translation agency.
- Some cost relevant issues can be discussed only on a larger scale than a single project. What is the benefit of a high quality translation memory? Which share of a software licence has to be charged for?
- An exact estimation of cost, time and quality would require rather de-tailed data about projects, agents and resources. It is questionable whether the result would justify the effort of data acquisition.

As a consequence TransRouter will provide mechanisms for the ranking of projects according to cost, time and quality. Additionally means for the computation of total values are provided. These depend on domain specific knowledge that cannot be provided in the scope of the project. Means for the acquisition of these data, however, are present.

A full description of the route assessment mechanisms of TransRouter would exceed the scope of this paper. As a consequence the general mechanism will be explained at the example of time assessment. Some remarks on quality will follow. Cost estimation will not be dealt with at all.

## **6.1 Translation time**

The most basic understanding of translation time can be defined in a few sentences.

- The time needed to process a translation route is the sum of the processing time of all of its steps.
- The translation time of a route step is computed from the number of words of the text times the agent's (translator, reviser etc) performance as contained in his profile (measured in words per hour).

Obviously this formula is a crude abstraction because there is no single translation performance of a translator. Performance in translation is dependent on a number of factors the most important of which will be named here:

- The first important factor seems to be the language pair. A translator may be competent in several languages but the translation performance will vary.
- High translation speed will have a negative impact on quality. Thus translation performance will decrease in projects with higher quality requirements.
- Subject domain and text style will also be influential. The knowledge of a specialised vocabulary or of conventional rules of text structure and formulation might be necessary.
- If the text contains a great deal of new terminology, this will affect translation performance adversely.
- Certain text types, for example legal text containing quotations which must be quoted rather than re-translated, will have an adverse effect on translation performance (because of the time required to search out the quotations): mitigated of course, if an appropriate translation memory is at available.
- Translation performance probably will depend on the readability or complexity of the text. TransRouter provides a tool for the estimation of text complexity. Since the notion of text complexity is not well understood up to now, this estimate can only be heuristic.

The influence of these factors seems to differ between individual translators. Thus, an exact translation performance measure would require the empirical acquisition of a huge matrix of interdependencies. Since this is not feasible a sufficient approximation must be found. The level of detail, which is required or asked for, will differ between translation agencies. Thus, TransRouter will support the definition of rather general and quite detailed models as well.

The mutual dependencies between translation performance and project features are of major importance for the time estimates of TransRouter. TransRouter uses an associative access method based on keys of variable length (see also figure 2). This mechanism will be described on the base of performance mapping as an example. The same mechanism will be used for other complex features — translation quality and translation costs — as well.

### **Definition of absolute translation speed**

Each profile of a translator or machine translation system contains mappings from project features to translation performance values (speed, quality, cost). The current implementation uses all relevant project features (source and target language, subject domain, text style and complexity, formats etc.) for keyed performance access. Additionally information about the translation route may be used (tools being used, features of resources being used). To avoid data acquisition overhead, a partial definition of access keys is possible. A fully

unspecified key will retrieve a default value from the system. An access key may contain the following 'wildcards' instead of true project or re-source data:

- '\*': Matches any value. This is useful especially if a default value is to be defined that is valid for any project constellation.
- 'some': Matches any value other than the empty object. This is useful for instance if a default value for pre-editing for machine translation is to be defined. In this case at least some machine translation system must be present within the route.
- 'none': Matches if only the empty object is present. This is useful if the use of specific system or agent types should be excluded.

Using this specification method the following phenomena can be expressed easily:

- A machine translation system can handle the following six language pairs at an average speed with given quality. The language pairs will be defined in the profile. No specific performance keys will be used. System performance will be defined as a default.
- The system will translate English to German at a higher speed. A specific key (*source*: English, *target*: German) concerning this language pair will be entered.
- A translator is responsible only for some very specific cases (e.g. scientific reports about biology). A specific key covers the respective translation performance. The default translation performance will be set to zero, thus prohibiting the assignment of other projects.

To access performance data for a project or route the following steps will be performed:

- Derive an access key from the project's (route's) features.
- Sort the access keys of the profile according to the number of features specified in descending order.
- Select those keys which subsume the access-key for the chosen project or route.
- Map the selected keys to their performance values.
- Compute a single value from the selected ones. In the case of translation speed and translation quality this means using the smallest value. Translation speed, thus, is defined by upper bounds, which are valid for specific situations. In case of cost computation every key will contribute a cost factor adding to the total costs.

### **Definition of factors affecting translation performance - attached procedures**

Obviously this matching algorithm, which prefers the most specific access keys, does not allow the use of general rules. More general effects are not described by absolute values but by numeric factors or even attached procedures. TransRouter will use a specific key structure (subject domain, text structure, text complexity) to access these data using the following algorithm:

- Derive an access key from the project's features.
- Find all keys matching the search key.
- Compute the product of all factors, which are associated to these access keys.
- Compute product of the resulting cumulative factor and the already known total performance value.
- Additionally attached procedures may be retrieved. These code fragments are sorted according to an inherent precedence value and then arranged as a pipe. The performance value found so far is used as input to this pipe. The output of the pipe is the final performance value, which will be further used. The attached procedures accept three parameters: 1<sup>st</sup> is the translation step currently being elaborated (or nil), 2<sup>nd</sup> is the project profile and 3<sup>rd</sup> is the translation performance value that has been computed so far.

The screenshot shows a window titled "Machine Translation System". At the top, there is a menu bar with "Undo" and a toolbar with "ID" (containing "MT Italian-German-Biology") and "Machine Translation System" (with a "T" icon). The main area is divided into several sections:

- Skills:** A list of radio buttons for "Address", "Language", "Formats", "Tools", and "Projects".
- Performance:** A section with radio buttons for "Transl. Rate", "Quality", and "Coding / Text".
- Environment:** A section with a radio button for "Requirements".
- Costs:** A section with a radio button for "Transl. Costs".
- Machine translation: Language Pairs to Performance:** A section containing:
  - A list of language pairs: "source: German target: Italian" and "source: Italian target: German".
  - A "Text Properties" table:
 

Source Language	Target Language	Domain	Style	Compl.
German	Italian			
  - Radio buttons for "text prop.", "text formats", "resource data", and "tools used".
  - A "Performance (words/h)" section with two input fields: "raw translation" (2500) and "preprocessed translation" (4000).
  - Radio buttons for "values", "factors", and "algorithms".
  - "update" and "remove" buttons.

At the bottom left are "ok" and "cancel" buttons.

**Figure 2:** Profile of a machine translation system defining the mapping of project data to performance values

The expressive power of TransRouter now is extended to phenomena like the following:

- A translator translates scientific texts by an excess of 30% of average translation time. Assign a factor of 1.3 to scientific texts.
- The use of a specific tool increases productivity by 15%.
- Do not even consider using a specific machine translation system to translate texts of legislation. Assign a performance factor of 0 to a general key (*domain*: legislation).
- The translation performance will not exceed a certain threshold, if the text is very complex. In this case a code fragment will check the threshold.
- The performance of teams is computed by some algorithm, which is defined as default (see below). As a consequence this algorithm can be modified easily according to the specific needs of an organisation.

## **Defaults**

A final extension of the data model allows the definition of default parameters for agent classes. A default performance profile for each type of agent concerned with translation (translators, freelancers, and machine translation systems) and the related activities (revising, pre- and post-editing) is available. In a commercial environment a system like TransRouter would probably be delivered with agent profiles (except translators) and defaults being set.

Default profiles have the same structure as those of individual agents. Thus, the same phenomena can be expressed. Nevertheless, default profiles will contain only a few absolute values but most of the general factors and attached procedures of the system. Individual profiles on the other hand will contain absolute values, which will be modified by factors or procedures derived from the default profiles.

Default profile and individual profile will be merged on access time. Each key and value pair of the default will be moved to the individual profile if and only if there is not a similar key already in existence in the individual profile. Thus, definitions in individual profiles take precedence over those of default profiles.

## **Preinstalled defaults**

Some defaults concerning attached procedures will be defined automatically by the system at installation time. They may be modified later on according to the specific needs of an organisation or user.

- Team performance: If a translator is member of a team, his translation performance will be reduced by a certain amount to cover organisational overhead.
- The performance of a post-editor depends on the difference between the quality value of the main translation step and the degree of quality expected from the project.

## **Stepwise refinement of access keys**

The flexible size of access keys does not only allow the choice of an adequate level of detail in the definition of data but also the stepwise refinement of access within the decision process. In the beginning only project data are available. Later on additional information about possible routes and their resource assignment is at hand. Especially knowledge about the tools being used within a translation step can influence the translation performance and therefore will be included into the key structure. As a consequence TransRouter's estimate of translation performance (and quality) will be improved when the user enters additional information — especially about routes.



## Translation quality

Quality is one of the most problematic concepts within translation evaluation [Marx et al 98]. There is neither a clear definition of the concept of text or translation quality nor a sufficient understanding of the interrelation between the translation process and its outcomes.

The screenshot shows the 'Route Elaboration Tool' window. At the top, there's a title bar and a menu bar with 'Undo'. Below the menu bar, there are two text boxes: 'Route' containing 'R2 translate by HT French-German-Chemistry' and 'Project' containing 'HT French-German-Chemistry'. To the right of the 'Project' box is a button labeled 'T'. On the left side, there's a vertical panel with three radio buttons: 'Preetit', 'Translation' (which is selected), and 'Postedit'. Below these is a 'compute' button. At the bottom of this panel are 'ok' and 'cancel' buttons. The main area is divided into two sections. The top section is titled 'Step Data' and contains a 'Type' dropdown set to 'translate', an 'Agent' dropdown set to 'HT French-German-Chemistry' with an 'open' button, and two columns labeled 'Tools' and 'Resources'. The 'Tools' column has a dropdown set to 'TM French-German 2' with 'open' and 'remove' buttons. The 'Resources' column has a dropdown set to 'TM for HT French-Germ' with 'open' and 'remove' buttons. To the right of the 'Resources' column is a checkbox labeled 'enhance TM' which is checked. Below these sections are two rows of data. The first row is labeled 'Time' with a value of '75.18', 'h', 'Costs' with a value of '751.81', 'Euro', and 'Quality' with a value of 'publication'. The second row is labeled 'Route Data' and contains 'Time' with a value of '137.68', 'h', 'Costs' with a value of '2889.31', 'Euro', and 'Quality' with a value of 'publication'.

**Figure 3:** Assessment of a translation route with emphasis on the translation step

Within a conventional production process quality is described as the probability that an individual product has the required features. A clear definition of product features, which is usually given in a product description (design, modes of operation etc.), is a prerequisite of this approach. Additionally a sufficient number of similar objects must be produced in order to be able to compute probabilities. Neither of these conditions is fulfilled in the case of text translation. Every translation is a very individual product presumably not allowing the estimation of fault probabilities. Most of the quality criteria that can be agreed on can not be formalised in a way that it can be used by a decision support system.

The lack of an exact quality measure is a common problem for service providers. One solution of this problem is to define quality not primarily as a feature of a

product but as a feature of the process of product construction or service delivery. From this point of view a tool like TransRouter is a major means of quality management since TransRouter will show possible translation routes and name the quality effects that can be expected. Quality requirements can be defined with respect to specific translation tasks [White/Taylor 98, Pavlsen et al 98]. This is a major step forward even if a general quantitative model of quality can not be provided.

Quality values are communicated to the user by four symbolic values:

- No use: The translation will probably be in such a bad shape that it can not be used at all.
- Browsing quality: The reader will be able to identify what the text is about.
- Information dissemination quality: The reader will identify the arguments and major facts of the text.
- Publication quality: The translation fully meets the quality standards of the original version with respect to content and form.

The quality of a translation will primarily result from the competence of the translator or translation system and the quality effects of translation tools and resources. Thus, TransRouter allows the definition of a detailed quality profile for translation agents. A quality profile generally is a mapping from project data to quality values. The access method is the same as described earlier in this paper. Similarly quality effects can be described using absolute values, factors or algorithms. Similar to translation performance all relevant project and route features — language pair, domain, text style, TM coverage, unknown terminology etc. - are covered.

## **7. Project documentation and learning from data**

Since TransRouter is a tool for supporting a translation manager in the appropriate choice of a route for a translation project. Issues concerning the project management and documentation seem to be outside the scope of the project. This is only true from the point of view of a single project. The decisions taken in a project and the outcome of these decisions however constitute valuable information for a new project if the project features are comparable in some way.

This is especially true as most of the data initially fed into TransRouter can only by approximations or even guesses. If projects that are processed using TransRouter are documented, the translation manager can get a notion of the quality and value of the decisions taken by TransRouter.

Thus, TransRouter will ask the translation manager for the following information — if not already present - on every project and put it into the archive:

- The project profile
- The route that is finally chosen after the decision process, including the agent, tool and resource assignments and the calculated time, cost and quality values
- The route as it is implemented within the course of the translation project, including the agents, tools and resources that were finally assigned to the project and the time that was needed for each translation step and the quality which was achieved.

On the basis of this information TransRouter can assess its own performance and calculate the averages of the deviation between predicted and true values for all projects and — more decisive — for all projects similar to a new project which is to be tackled by TransRouter. Furthermore all tools or resources may be identified, which although chosen by TransRouter on the basis of their pro-files are often skipped by the managers in the real implementation of the project. These simple but useful features are already implemented in the Trans-Router prototype.

The next step in the evolution of TransRouter, which cannot be taken within the scope of this project, would be the learning from agent features from real world data. Every translation project represents a new case from which system parameters can be learned. The first step would be the acquisition of very specific access keys to performance and quality that represent the relevant project features. These keys will then be assigned to the agent profiles. If similar cases occur later on, the data may be adapted to get a best fit to all similar cases. Later on, when a fair sized pool of cases is at hand, generalisation processes may be run on the profiles. They will isolate those project and route features, which contribute significantly to the project outcome and skip those that don't. Thus, the general predictive quality of the system will be gradually enhanced.

## References

James R. Evans and William R. Lindsay: The management and control of quality. South-Western College Publishing, 1999.

Entscheidungsunterstützung bei der Planung von Übersetzungsprojekten. In: Zimmermann / Schramm (eds.), Knowledge Management and Communication Systems, Proc. of the 6<sup>th</sup> Int. Symposium on Information Science, pp. 47-57, 1998  
Hammwöhner 00 Decision support in the routing of translation projects. Re-port on the Regensburg prototype of TransRouter. Technical Report, 1999.

The TransRouter consortium (contact author: Margaret King): TransRouter: a decision support tool for translation managers. In. Proc. MT Summit Conference, September 13-17 1999, Singapore, 1999

Jutta Marx, Nancy Smith and Bernhard Staudinger: Some Problems in the

Evaluation of the Russian-German Machine Translation System MIROSLAV. In: Rubio et al (eds.), Proc. of the First Int. Conf. On Language Resources & Evaluation, Granada, 1998, Vol. 2, pp. 1219-1226, 1998.

Claus Pavlsen, Nancy Underwood, Bradley Music and Anne Neville: Evaluating Text-type Suitability for Machine Translation a Case Study on an English-Danish MT System. In: Rubio et al (eds.), Proc. of the First Int. Conf. On Language Resources & Evaluation, Granada, 1998, Vol.1, pp. 27- 31, 1998.

Efraim Truban: Decision Support and Expert Systems. Management Support Systems. Prentice Hall, 1995.

John S. White and Kathryn B. Taylor: A Task-Oriented Evaluation Metric for Machine Translation. In: Rubio et al (eds.), Proc. of the First Int. Conf. On Language Resources & Evaluation, Granada, 1998, Vol. 1, pp. 21- 25, 1998.



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 71 – 87

## **Wortmodell und Begriffssprache als Basis des semantischen Retrievals**

**Gerhard Rahmstorf**

Oberer Rainweg 57

69118 Heidelberg

Tel. 06221-808129

Fax. 06221-802682

rahmstorf@regio-info.de

Technische Universität Darmstadt

### **Zusammenfassung**

Der heutigen Retrievaltechnik wird das Projekt eines semantisch basierten Suchsystems gegenübergestellt. Es soll genauer und vollständiger arbeiten sowie systematische Zusammenhänge zwischen Themen unterstützen. Bei diesem Ansatz wird ein umfassendes Wörterbuch mit einer einfachen begrifflichen Darstellung der Wortbedeutungen benötigt. Das Wortmodell bildet Wort, Wortmerkmale, Lemma, Wortbedeutungen (Lesarten), Lesartenmerkmale und Begriffe ab. Begriffe sind formale Ausdrücke einer Begriffssprache. Entsprechend dieser Differenzierung wird Lemmaindexierung, Lesartenindexierung und Begriffsindexierung unterschieden. Begriffe werden mit dem Programm Concepto grafisch konstruiert und erfasst.

### **Abstract**

The requirements for a more perfect retrieval are described from the user's viewpoint: thematic search, queries in the form of natural language phrases, basic interpretation of word meanings etc. A word model for such a system supports three different types of objects: words, readings and concepts. Concepts are expressions of a concept language. A software system called Concepto supports the acquisition of lexical data and the graphical construction of formal concept expressions.



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

## 1. Einleitung

Schon früh wurde in der Geschichte der Informationstechnik vom semantischen Retrieval gesprochen<sup>1</sup>. In der Praxis hat sich jedoch der schon in den frühen sechziger Jahren etablierte Ansatz des zeichenkettenbasierten Recherchierens durchgesetzt. Dieser Ansatz dominiert in verschiedenen Varianten die heutige Suchtechnik. Die Varianten unterscheiden sich z. B. bezüglich der Funktionen bei der Abfrage (Boolesche Operatoren), der für die automatische oder intellektuelle Dokumentcharakterisierung verwendeten Basis (Trigramme, "Wörter", Deskriptoren eines kontrollierten Vokabulars), der textuellen Vorverarbeitungen (Erkennung von Synonymen und Wortformen) und weiterer nachgeschalteter Funktionen (Relevanzfeedback). Das Spektrum der Methoden wird verfeinert und auch für die Besonderheiten der Suche im Web weiterentwickelt (Mladenic 1999). Diese Verfahren können in differenzierter Weise Relevanz von Dokumenten bestimmen, allerdings auf der Basis des Vorkommens von Zeichenketten und der formbezogenen Eigenschaften der Anfrage und der Dokumenttexte. Erst wenn ein Thesaurus bei der Relevanzanalyse verwendet wird, können wenige semantische Beziehungen bei der Relevanzanalyse berücksichtigt werden.

Die eingeführte Retrievaltechnik hat den großen Vorteil, dass mit ihr schnell und kostengünstig große Mengen von Texten automatisch indexiert werden. Sie ist robust, gerade weil sie ohne linguistische Analysen und komplexe Begriffsstrukturen auskommt. Das Schlagwort "Volltextrecherche" signalisiert, dass ohne nennenswerte intellektuelle Zusatzarbeit jeder beliebige Text beliebiger

---

<sup>1</sup> Im Arbeitsgebiet Information und Dokumentation war von Beginn an klar, dass es beim Dokumentieren, Ordnen und Recherchieren primär um Inhalte und Themen geht. Daher wurden im Verlaufe der letzten 30 Jahre immer wieder Retrievalsysteme entworfen und erprobt, bei denen natürliche Sprache, semantische Analysen, begriffliche Repräsentationen, Weltwissen und Inferenztechniken eine Rolle spielen. Eine angemessene Würdigung dieser Entwicklungen würde den Rahmen dieser Arbeit sprengen. Das semantische Retrieval in der hier dargestellten Form ist nur eine Variante unter den Entwicklungen, die das klassische Schema verlassen. Es geht auf Überlegungen des Verfassers aus den frühen siebziger Jahren zurück. Zu dieser Zeit wurden schon roles und links als Instrumente zur Entwicklung einer genaueren Indexierung und Recherche diskutiert. Die Linguistik hatte mit der Kasusgrammatik (Fillmore 1968) einen inspirierenden Beitrag zur semantischen Strukturierung geliefert. Tiefenkasus ließen sich in semantischen Netzen verwenden. Die Idee des Indexierens und Recherchierens mit Phrasen und der Gedanke der Relevanzanalyse auf der Grundlage der begrifflichen Phrasenrepräsentationen wurden in (Rahmstorf 1978) skizziert. Die Bezeichnung "Semantic Information Retrieval" hat schon Raphael 1968 für sein System verwendet, das jedoch nicht die hier dargestellte Methodik verfolgt. Schank bezeichnete mit "Conceptual Information Retrieval" eine sehr viel weitergehende Aufgabenstellung, zu der die Analyse von Texten, die konzeptuelle Repräsentation von Fakten aus diesen Texten, die Pflege der Daten und die Beantwortung von frei formulierten Anfragen gehören (Schank, Kolodner, DeJong 1980). Die unbeschränkte Verwendbarkeit der natürlichen Sprache stellt bis heute eine schwer zu nehmende Hürde dar. Dokumentation und Retrieval haben sich weitgehend auf Wörter (Deskriptoren) beschränkt. Der vielfältige Einsatz von linguistischen Mitteln (Klavans 1994) und die Versuche, Methoden der Künstlichen Intelligenz und Expertensystemtechnik für bestimmte Aufgaben des Retrievals anzuwenden, führten bisher nicht zu einem neuen praktikablen Paradigma des Retrievals.

Sprachen maschinell recherchierbar gemacht werden kann. Das ist daher -soweit betrachtet- die passende Methode für den immensen Zuwachs an maschinell verfügbarem Material im Internet. Alternative Methoden haben nur deshalb eine Chance, in Zukunft verwendet zu werden, weil die Genauigkeit und Vollständigkeit der Recherchen unbefriedigend sind, und zwar insbesondere dann, wenn nach spezifischen Themen gesucht wird, die nicht durch Eigennamen oder andere eindeutige Codes gesucht werden können.

Ein weiterer Einwand gegen die heutige Recherchetechnik ist die fehlende Ordnung und Durchschaubarkeit des Verfahrens und der Ergebnispräsentation. Zumindest das Fachpublikum möchte Zusammenhänge, Kontexte und Ordnung in der Masse der angebotenen Dokumente erkennen. So wie man von der Buchwelt herkommend wohlgeordnete Lehrwerke, Fachwörterbücher, Handbücher, Enzyklopädien und systematische Vertiefungen in Form von Monographien kennt, so möchte man auch in der elektronischen Welt bessere Hilfen zur Orientierung haben und ein systematisch gestaltetes Wissenssystem benutzen können.

Alternative Retrievalverfahren unterscheiden sich von dem etablierten Ansatz nicht nur durch die Einführung von Begriffen als Basis für die Bestimmung der Relevanz eines Dokuments bezogen auf einen Text, sondern auch durch die Einführung einer Informationssprache für Recherche und Indexierung. Diese Sprache umfasst Wörter, feste Phrasen und eine offene Menge regelhaft gebildeter Benennungen, vorwiegend vom Typ Nominalphrase (Tabelle 1).

<b>Informationssprache</b>	<b>ohne Semantik: Zeichenketten</b>	<b>mit Semantik: Begriffsdarstellung</b>
sublexikalische Einheiten	n-gramme	-
lexikalische Einheiten: kontrolliert	Thesaurus: Deskriptoren	(Notationen, Thesaurus online für Recherche)
lexikalische Einheiten: beliebige Wörter	Gegenwärtige Technik: "Freitextsuche"	
syntaktische Einheiten: kontroll. Phrasenmenge	(Thesaurus) begrenzt möglich	
syntaktische Einheiten: freie Phrasensprache		semantisches Retrieval

**Tabelle 1:** Retrievaltechniken

Ordnung wird bei diesem Ansatz dadurch erzielt, dass die Begriffsdarstellung soweit möglich und zweckmäßig dem folgt, was die Benennungen besagen. Die Relevanz wird auf der Basis von Beziehungen zwischen Begriffen bestimmt. Der Entwicklungsaufwand muss sich lohnen. Wird ein qualitativ besseres Retrieval wirklich dringend benötigt? Die Benutzer können das heute ebenso wenig beantworten, wie sie vor 15 Jahren die Frage hätten beantworten können, ob sie ein Internet benötigen würden? Wir müssen uns daher selber in die Rolle eines Benutzers versetzen und den Ansatz des semantischen Retrievals aus den

Interessen der informationssuchenden Anwender, insbesondere der Fachleute aus Wissenschaft, Technik und Wirtschaft, begründen.

Die Anwender haben sich ebenso wie die Informationsexperten an die Arbeitsweise von Datenbanken und Suchmaschinen gewöhnt. Sie kennen nichts anderes. Sie wissen auch oft, wie die Programme ihre Suchresultate erzielen. Sie kommen daher gar nicht erst auf den Gedanken, von einem System die perfekte Bereitstellung von gesuchtem Wissen zu fordern und an dieses System wie an einen fachkundigen Bibliothekar heranzutreten. Bei Gesprächen mit Bibliothekaren werden sie nicht verzichten wollen auf thematische Suche, freie Wahl der Themenbenennung, Unabhängigkeit von festgesetzten Strukturierungen und veralteten Systematiken u. a.

## **2. Anforderungen 2.1 Thematische Suche**

Benutzer suchen nach Literatur zu einem Thema. Diese thematische Suche unterscheidet sich von anderen Arten der Suche, z. B. der spontanen Entscheidung für angebotene Links in einem Hypertext, der Suche nach einem bestimmten, schon bekannten Dokument oder der Suche nach Dokumenten, in denen bestimmte Wörter oder Zeichenketten vorkommen. Bei der thematischen Suche sind die Dokumente für den Benutzer relevant, die das von ihm angegebene Thema behandeln, und zwar unabhängig von inhaltlichen Einzelheiten, Stil, Gestaltung, Gliederung und Wortwahl der möglichen Texte.

## **2.2 Unabhängigkeit von Strukturfestlegungen**

Benutzer möchten nicht unbedingt in einer bestimmten Struktur suchen, die in Form eines Menübaums oder eines bibliothekarischen Klassifikationssystems vorgefertigt wurde. Ordnungssysteme können dem Benutzer Orientierung vermitteln oder auch zur Navigation dienen. Sie sollten aber möglichst die bedeutungsgemäßen Beziehungen zwischen den Klassenbenennungen wiedergeben.

## **2.3 Natürliche Sprache für die Anfrage**

Benutzer möchten das Thema mit den differenzierenden Mitteln der eigenen Sprache ausdrücken können. Sie wollen die Wörter verwenden, die ihrer Meinung nach der Sache angemessen sind. Sie wollen sich nicht auf bestimmte Wörter oder Codes beschränken, die die Suchlogik des Systems unterstützt. Thesauren und andere kontrollierte Vokabulare stellen eine Erschwernis dar, die insbesondere dem Internetbenutzer nicht ohne Not zugemutet werden sollte.

Zur Beschreibung eines Themas sollte man die sprachlichen Konstruktionen verwenden können, die das jeweilige Thema so präzise wie nötig gegen andere Themen abgrenzen. Zur Themenbenennung sollte die Retrievaltechnik daher diejenigen Attributformen unterstützen, die zur Benennung für Titel in der fachlichen Literatur verwendet werden. Das sind z. B. Adjektivattribute, Genetivattribute und Präpositionalattribute, die ein substantivisches Bezugswort



näher bestimmen. Erst wenn solche Phrasen als Ausdrücke der Informationssprache zugelassen werden, wird klar, welches Thema den Benutzer eigentlich interessiert. Die Relevanz von Dokumenten bezüglich einer Anfrage kann nur dann genauer bewertet werden, wenn die Anfrage explizit ausformuliert ist und sich damit von anderen Anfragen, die mit den gleichen Wörtern gebildet wurden, unterscheidet. Durch die sprachliche Formulierung werden die Wörter nicht nur in eine bestimmte Reihenfolge gebracht, sondern auch syntaktisch strukturiert. Mit Hilfe dieser Struktur wird der Ausdruck interpretiert. Aus den drei Wörtern *Ausbildung*, *Arbeitsloser* und *Computertechnik* können z. B. die folgenden Phrasen unterschiedlicher Bedeutung gebildet werden (Rahmstorf 1994, Lein 1994):

- Ausbildung der Arbeitslosen in der Computertechnik
- Ausbildung in der Computertechnik durch Arbeitslose
- Computertechnik zur Ausbildung für Arbeitslose
- Arbeitslose mit Ausbildung in Computertechnik
- Arbeitslose in der Ausbildung für Computertechnik

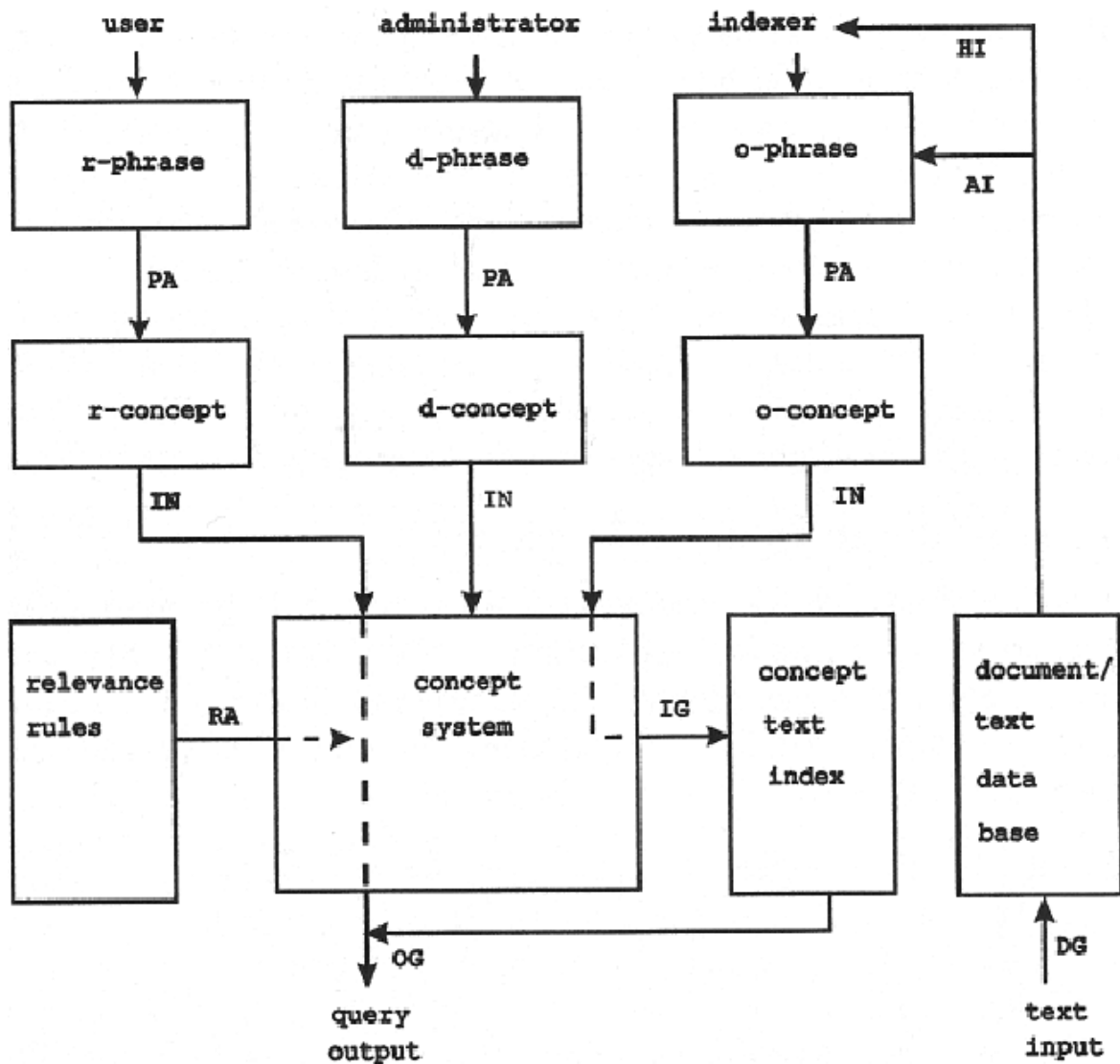
Phrasen stellen nur einen Teil der Ausdrucksmöglichkeiten der natürlichen Sprache dar. Zu dieser Informationssprache gehören z. B. keine vollständigen Sätze.

### **3. Komponenten des semantischen Retrievals**

Mit den natürlichen Benutzeranforderungen wird der Rahmen für das semantische Retrieval festgelegt. Diese Retrievaltechnik wird kurz durch drei Eigenschaften gekennzeichnet: sie bietet die Möglichkeit, mit Phrasen zu indexieren und zu recherchieren, sie stellt die Phrasen als begriffliche Strukturen dar und sie bestimmt die Relevanz von Texten für eine gegebene Anfrage auf der Basis der begrifflichen Repräsentation des Anfragethemas und der Themen der Informationsangebote. Dabei werden gebietsunabhängige Relevanzregeln verwendet.

Abbildung 1 stellt den symmetrischen Aufbau für den Datenfluß im semantischen Retrieval dar. Die anzubietenden Dokumente oder Texte werden in eine Dokumentdatenbank geladen (Prozeß DG). Den Texten oder Kapiteln eines Textes müssen Themen durch jeweils eine Phrase zugewiesen werden. Diese Indexierung muß ein Indexierer machen (HI), solange noch keine Programme vorhanden sind, die ein automatisches Phrasenindexieren (AI) ermöglichen.

Wortmodell und Begriffssprache als Basis des semantischen Retrievals



**Abbildung 1:** Systemübersicht zum semantischen Retrieval

Der als offering phrase (o-Phrase) bezeichnete Ausdruck gehört einer der vom System unterstützten natürlichen Sprachen an. Er muss lexikalisch, syntaktisch und semantisch analysiert werden (PA), um als Begriff dargestellt werden zu können. Der Begriff wird als Struktur in das Netz der schon vorhandenen Begriffe eingefügt (IN). Zu diesem Begriff wird ein Verweis bzw. Index auf den zugehörigen Text generiert (IG). Die Verweise bilden zusammen einen Index für eine gegebene Textdatenbank. Bei der Recherche kann der Benutzer seinen Informationsbedarf mit einer Phrase ausdrücken. Diese request phrase (r-Phrase) wird danach wie die o-Phrase analysiert (PA). Die Anfrage wird als Begriff im vorhandenen Begriffsnetz integriert (IN). Die Relevanzanalyse (RA) bestimmt, welche o-Begriffe für die gegebene Anfrage relevant sind. Für die relevanten Begriffe werden die zugehörigen Texte zusammengestellt (OG) und dem Benutzer ausgegeben.

Die begriffliche Basisstruktur, die mit dem Wortschatz einer Sprache verbunden ist, muss in dem Retrievalsystem intellektuell erfasst werden. Ausgangspunkt jeder Bedeutungsbeschreibung bilden einfache verbale Definitionen der einzelnen Wortbedeutungen (Lesarten). Auch die Definitionen sind syntaktisch betrachtet Phrasen. Durch diese d-Phrasen (definition phrase) werden Begriffe auf andere Begriffe zurückgeführt. Alle Begriffe bilden ein Netz, das zur Relevanzanalyse verwendet wird. Die Beziehungen zwischen Begriffen legen Zugriffswege fest. Wenn z. B. *Sehen* als *visuelles Wahrnehmen* definiert wird, dann können Texte, die sich mit dem Thema *Sehen* befassen, auch über die Begriffe *Wahrnehmung* und *visuell* erreicht werden. Die Basisstruktur wird zusammen mit dem Wörterbuch betreut und aktualisiert. Diese Aufgabe ist mit der Pflege eines Thesaurus vergleichbar. Die Benutzer müssen sich mit dieser Struktur nicht befassen.

### **3.1 Anforderungen an das Wörterbuch**

Wörterbücher für semantisches Retrieval müssen u. a. folgende Aspekte der Benennungen unterstützen: Lesarten (Homonymität), Synonymität, Wortäquivalenz (Übersetzung) und begriffliche Strukturen. Dazu muss ein entsprechend differenziertes Wortmodell für die lexikalischen Daten aufgebaut werden.

Wenn die Benutzer ihre Themen frei benennen können, aber ein beträchtlicher Teil der Wörter natürlicher Sprachen mehrdeutig ist, müssen die verschiedenen Lesarten der Wörter im Wörterbuch des Retrievalsystems unterschieden werden. Es gehört aber auch zu den Eigenarten aller natürlichen Sprachen, dass ein bestimmtes Thema mit verschiedenen Wörtern und Ausdrücken benannt werden kann. Die Benutzer wollen sich aber nicht auf bestimmte Vorzugsbenennungen, die die Maschine erkennt, festlegen lassen. Daher müssen verschiedene Ausdrücke, die die gleiche Bedeutung haben, durch eine einzige begriffliche Repräsentation im Retrievalsystem dargestellt werden.

Auch in einer globalisierten Welt muss die Telekommunikation und das Recherchieren auf der Basis der verschiedenen Einzelsprachen durchführbar sein. Wenn mehrsprachiger Zugriff möglich sein soll, muss das Retrievalsystem über Wörterbücher für jede unterstützte natürliche Sprache verfügen. Äquivalente Wörter und Benennungsausdrücke von verschiedenen Sprachen müssen als gleichbedeutend erkannt und dargestellt werden.

Jeder Benutzer weiß, dass zwischen den unendlich vielen Themen, zu denen Information angeboten und gesucht werden kann, Beziehungen bestehen. In der Kommunikation zwischen Benutzer und Bibliothekar sind diese Beziehungen Grundlage des gegenseitigen Verstehens. Im Retrievalsystem sollten sie annähernd abgebildet werden. Bedeutungsstrukturen sollen nicht unnötig komplex werden, damit die begriffliche Repräsentation in technische Hinsicht und bezüglich der Erstellung von Strukturen beherrschbar bleibt. Das Wortmodell wird daher durch eine verhältnismäßig einfache Begriffssprache ergänzt. So wie in der natürlichen Sprache die Grammatik den Rahmen für die Satzbildungen festlegt,

legt die Begriffssprache den Rahmen für die begrifflichen Strukturen fest, die den einzelnen Lesarten eines Wortes zugewiesen werden können.

### 3.2 Wortmodell

Aus den genannten Anforderungen leitet sich eine differenzierte Modellierung der sprachlichen Gegebenheiten ab. In dem Retrievalsystem müssen vier Arten von linguistischen Objekten unterschieden werden: Sprachen, Wörter, Wortbedeutungen (Lesarten) und Begriffe. Diese Objekte haben jeweils eigene Eigenschaften (Attribute). Zwischen den Objekten bestehen bestimmte Arten von Beziehungen. Sprachen sind hauptsächlich durch ihren Wortschatz und ihre Grammatik bestimmt. Für die Modellierung ist zu beachten, dass sich die unterschiedlichen Grammatiken der Einzelsprachen auch darin niederschlagen, dass die Kategorien der Wortbeschreibung in den Einzelsprachen voneinander abweichen. Wir können daher nicht von einem einheitlichen Wortbeschreibungsschema für alle Sprachen ausgehen. Die den Wörtern zuzuweisenden Attribute, z. B. Genus, Deklinationsklasse, und die jeweils möglichen Attributwerte sind sprachspezifisch. Mehrsprachig betreibbare Retrievalsysteme, die auch für Übersetzungsfunktionen verwendet werden, müssen daher eine an die jeweilige Sprache anpassbare Merkmalstruktur für die Wortbeschreibung aufweisen.

Zu den Einheiten des Wortschatzes gehören nicht nur die "Wörter", sondern auch Mehrwortausdrücke, die eine spezifische, nicht regelhaft rekonstruierbare Bedeutung haben. Jedes Wort gehört zu nur einer Einzelsprache. Wörter können in verschiedenen Formen vorkommen. Die Wortidentität wird durch die Benennungsform des Wortes und durch weitere Merkmale bestimmt. Die Zeichenkette *Kiefer* als Lemma reicht nicht aus, um das Wort eindeutig zu identifizieren. In diesem Fall ist der Wert des Merkmals Genus zusätzlich anzugeben: *die Kiefer* bezeichnet etwas anderes als *der Kiefer*. Im Englischen muss die Zeichenkette *capital* durch einen Wert für die Kategorie Wortart - Nomen oder Adjektiv - identifiziert werden. Das Lemma als die Zeichenkette, mit der ein Wort benannt wird, ist im Deutschen nach der Orthographiereform durch die inflationäre Zulassung von Schreibvarianten kein eindeutiges Wortmerkmal mehr. Eine datenbankmethodisch saubere Abbildung dieser weiteren 1:n-Beziehung wird unzumutbar komplex. Man wird daher in einem Computerwörterbuch für Retrieval und Übersetzung entweder benutzungserschwerende Verweise oder redundante Einträge für die Schreibweisen *Fotografie*, *Photographie* und eventuell weitere Varianten haben. Dies ist deshalb ein Problem, weil das Lemma für die Praxis verschiedene Funktionen erfüllt. Es ist Aufrufbenennung eines Wörterbucheintrags, Grundlage der alphabetischen Einsortierung der Wörter, gemeinsame Kennzeichnung aller Lesarten eines Wortes in der Strukturvisualisierung u. a.

Wörter können mehrere Lesarten haben. So hat z. B. das englische Substantiv *capital* vier Lesarten. Jede Lesart wird durch Lesartenmerkmale beschrieben. Die in der Praxis wichtigsten Lesartenmerkmale sind: verbale Definition, Markierungen bezüglich Sprachebene, Metaphorik u. a., Beispielsatz zur Charakterisierung von

Bedeutung und Gebrauchsweise, Fachgebietsangabe, kommentierende Hinweise, Lesartenidentifikation. Die Kategorien, mit denen Lesarten beschrieben werden, sind für alle Sprachen gleich. Zumindest sind derzeit keine Gründe erkennbar, warum für die informationstechnische Anwendung sprachspezifische Unterschiede in der Lesartenkategorisierung vorgesehen werden sollten. Die Lesartenmerkmale werden im Datenmodell sauber von den Wortmerkmalen unterschieden. Diese Gegebenheiten der Sprache machen die Benutzungsoberfläche bei der Wortschatzerfassung komplexer. Die Vorteile dieser lexikographischen Oberfläche möchten aber die Entwickler des Wörterbuchs nach einer kurzen Einarbeitungszeit nicht mehr missen.

Ein wichtiger Modellierungsaspekt ist, dass jede Lesart ihre eigene Identität hat. Dies soll am Beispiel der Lesarten der Wörter *Anfang* und *Beginn* gezeigt werden. Jedes dieser beiden Wörter hat nur eine Lesart. Die Wörter sind synonym, d. h. ihre Bedeutungen stimmen überein. Dennoch sind die Beschreibungen der beiden Lesarten unterschiedlich. Der Beispielsatz muss unterschiedlich sein. Die verbale Definition kann unterschiedlich formuliert sein. Die wichtigsten Daten zur Charakterisierung der Lesarten sind textueller Art und damit sprachspezifisch. Deutsche Wörter werden normalerweise durch deutsche Definitionen beschrieben, französische Wörter durch französische Bedeutungsumschreibungen. Daher können zwei Lesarten, die Wörtern verschiedener Sprachen angehören, nicht identisch sein, selbst wenn sie ihrer Bedeutung nach äquivalent sind.

Jeder Lesart kann mindestens ein Begriff zugewiesen werden. Die Lesart ist ein sprachspezifisches Objekt, der Begriff ist sprachinvariant. Lesarten können nicht ohne Wörter, aber ohne Begriffe spezifiziert werden. Begriffe können durch formale Ausdrücke spezifiziert werden, ohne einer Lesart eines Wortes zugewiesen sein zu müssen. Ein gegebener Begriff kann verschiedenen Lesarten zugewiesen sein. Dies wird immer dann der Fall sein, wenn diese Lesarten von Wörtern einer bestimmten Sprache als synonym gelten. Entsprechendes gilt für Lesarten, die Wörtern unterschiedlicher Sprachen angehören. Man spricht dann von einer interlingualen Synonymität oder auch von semantischer Äquivalenz. Die Begriffe werden durch die noch zu beschreibende Begriffssprache charakterisiert. Zwischen Wörtern und Begriffen bestehen semantische Beziehungen. Zwischen Begriffen bestehen begriffliche Beziehungen.

Das hier beschriebene Wortmodell wurde in dem Programm Concepto implementiert. Dieses Programm ist kommerziell verfügbar von Antje Rahmstorf Sprachsysteme, 69118 Heidelberg, Oberer Rainweg 57. Es unterstützt auch die Begriffssprache und die dazugehörigen Operationen der grafischen Konstruktion von Begriffen. In dem Wortmodell werden folgende Objekte, Relationen und Attribute unterschieden:

1. Sprache
2. Benennungseinheit (Wort, idiomatische Phrase, frei formulierte Benennung)
3. Merkmale des Wortes für sprachspezifische Kategorien

4. Lemma als besonderes Merkmal des Wortes
5. Lesarten
6. Merkmale der Lesarten
7. Begriffe, die Lesarten zugewiesen werden
8. Beziehungen zwischen Begriffen

Die Objekte müssen aus technischen Gründen identifizierbar sein. Es fällt wegen der Wortmerkmale schwer, für Wörter eine bestimmte Identifizierung, z. B. durch eine vereinbarte Wortnummer, durchzuführen. Jede Institution hat daher seine eigenen Wortnummern. Lesarten lassen sich dagegen leichter identifizieren. Eine Lesartennummer ist notwendig, um eine klare Schnittstelle zwischen Lesart und Begriff einzuführen. Die Begriffe identifizieren sich durch die formalen Ausdrücke der Begriffssprache.

### **3.3 Begriffssprache**

Mit einer Begriffssprache lassen sich Bedeutungen von Benennungsausdrücken der Einzelsprachen in einer sprachübergreifenden formalen Weise darstellen. Jede Begriffssprache beschreibt ein Spektrum von möglichen Strukturen oder Repräsentationen. Bei dem Aufbau von begrifflichen Strukturen geht man von den verbalen Definitionen der einzelnen Lesarten eines Wortes aus. Nicht jede Lesart eines Wortes kann definiert werden. undefinierbare Wörter bzw. Lesarten werden in der Begriffssprache als undefinierte Begriffe eingeführt. Sie erhalten keine definierende Struktur. Um die undefinierten Begriffe, die für den Menschen verschiedene Bedeutungen haben, in der Maschine unterscheiden zu können, kann man ihnen jeweils andere Notationen der Begriffssprache als Formalausdrücke zuweisen. Nicht jede Lesart, die verbal definiert werden kann, muss auch in der Begriffssprache formal definiert werden. Definierte Strukturen werden nur dann erstellt, wenn sie für die jeweilige Anwendung Vorteile bringen.

Begriffliche Strukturen unterstützen Recherchen, Schlussfolgerungen, Übersetzungen u. a.. Sie werden von Informatikern, Informationsfachleuten und anderen Experten benutzt. Benutzer dieser Systeme müssen die begrifflichen Strukturen nicht unbedingt verstehen. Sie können sie aber zur Navigation verwenden. Dennoch dürfen Begriffssysteme nicht in eine undurchschaubare, unkontrollierbare Komplexität ausufern. Sie müssen für den Entwickler der Struktur leicht interpretierbar sein. Hinzu kommt, dass formale Ausdrücke einer Begriffssprache in textueller Form schwerer verständlich sind. Eine Visualisierung der formalen Ausdrücke ist für die Erfassung der begrifflichen Strukturen notwendig. Das erwähnte Programm Concepto unterstützt alle notwendigen Funktionen der Editierung und grafischen Manipulation von begrifflichen Strukturen.

Unter dem Anwendungsaspekt der Rechertechnik betrachtet müssen Begriffssprachen für beliebige Benennungen aus allen denkbaren Gebieten offen sein und nicht von den Besonderheiten der jeweiligen Fachgebiete geprägt werden. Die Grenzen zwischen den Fachgebieten sind ohnehin fließend. Wer in der Biochemie tätig ist, will mit denselben Rechertechniken in der Chemie, der Biologie, der Medizin und der Informatik suchen können. Ziel einer begrifflichen Repräsentation in der Informationstechnik ist es daher, sprachinvariant und fachgebietsinvariant zu sein. Dies wird u. a. dadurch erreicht, dass die Relationstypen, die in der Begriffssprache CLF von Concepto verwendet werden, nicht aus Fachgebieten, sondern aus den Attributformen, Wortbildungsmustern und anderen universalen grammatischen Mitteln abgeleitet worden sind.

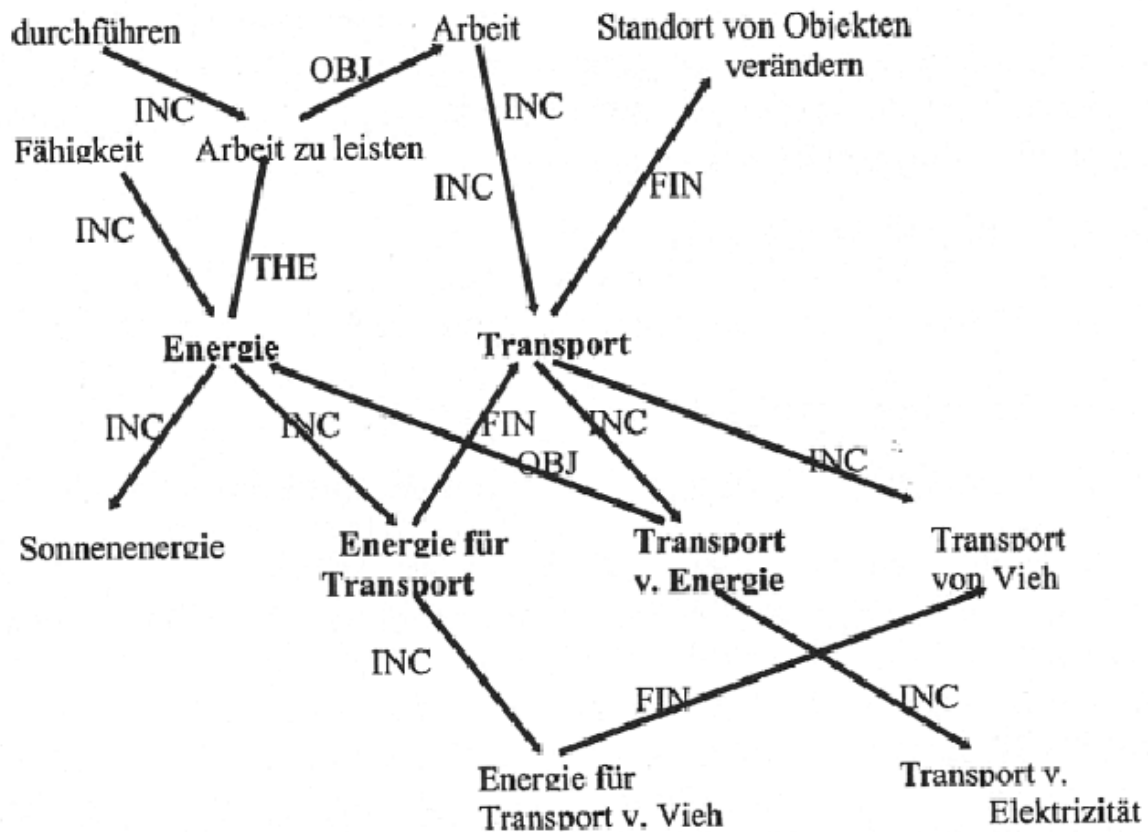
### **3.3.1 Begriffliche Repräsentation**

Abb. 4 zeigt als Beispiel eine begriffliche Repräsentation für die Wörter *Energie* und *Transport* und für weitere Begriffe im Umfeld dieser Wörter. Die Knoten stellen Begriffe dar, die Kanten symbolisieren begriffliche Beziehungen. Die Wörter *Energie* und *Transport* sind durch folgende Phrasen definiert worden:

*Energie* := *Fähigkeit, Arbeit zu leisten*

*Transport* := *Arbeit, die den Zweck hat, den Standort von Objekten zu verändern*

Die Analyse dieser Definitionsphrasen führt zu der begrifflichen Struktur, die im Diagramm oberhalb der Begriffe *Energie* und *Transport* dargestellt ist.



**Abbildung 3:** Beispiel einer begrifflichen Darstellung

In dieser Begriffsrepräsentation werden wie in einem Thesaurus nur Beziehungen aus einem begrenzten Vorrat an zweistelligen Relationen zugelassen. Die im Beispiel vorkommenden Beziehungen und ihre Paraphrasen sind:

Inklusionsrelation	INC(x,y)	y ist ein x
Finalrelation	FIN(x,y)	y ist der Zweck von x
Objektrelation	OBJ(x,y)	y ist das Objekt von x
Themarelation	THE(x,y)	y ist das Thema von x

Jeder Begriff wird hier durch einen direkten Oberbegriff und durch einen direkten Differenzbegriff als begriffliche Struktur definiert. Der Begriff *Fähigkeit* definiert mit der Inklusionsrelation INC den Begriff *Energie*. *Fähigkeit* ist daher der direkte Oberbegriff von *Energie*. *Arbeit leisten* ist der direkte Differenzbegriff von *Energie*. Zum Umfeld eines definierten Allgemeinbegriffs x gehören außer diesen zwei definierenden Positionen direkter Oberbegriff o(x) und direkter Differenzbegriff d(x) die folgenden fakultativen Positionen: direkter Unterbegriff u(x) und direkter Seitenbegriff s(x). Wenn  $y=o(x)$ , dann ist  $x=u(y)$ . Ebenso gilt: wenn  $y=d(x)$ , dann ist  $x=s(y)$ . So ist z. B. der Begriff *Energie* ein Seitenbegriff von *Arbeit leisten*, weil *Energie* mit dem Begriff *Arbeit leisten* als Differenzbegriff definiert wurde.



Die beiden Phrasen *Transport von Energie* und *Energie für Transport* können entsprechend ihrer unterschiedlichen Bedeutung durch zwei verschiedene Begriffe explizit dargestellt werden:

*Energie für Transport* =  $u(\text{Energie})$

*Energie für Transport* =  $s(\text{Transport})$  mit der Differenzrelation FIN

*Transport von Energie* =  $u(\text{Transport})$

*Transport von Energie* =  $s(\text{Energie})$  mit der Differenzrelation OBJ

Im nichtsemantischen Retrieval auf Wortbasis würden jede der beiden Phrasen auf die Menge der beiden Deskriptoren {Energie, Transport}, zurückgeführt.

### 3.3.2 Indexierung

Der symmetrische Ansatz des semantischen Retrievals fordert eine gleichartige Vorgehensweise für Anfrage und Indexierung. Wenn thematisch gesucht werden soll, wird auch eine thematische Indexierung benötigt. Man kann prinzipiell auf den Ebenen der Benennung, der Lesarten und der Begriffe indexieren. In Tab. 2 sind die diese drei Indexierungsmittel für zwei Indexierungsarten angegeben. Bei der ersten Art (Fälle A, B und C) beschränkt sich die Indexierung auf Mengen von Vokabulareinheiten. Das können Wörter oder Mehrwortausdrücke aus einer festgelegten Menge von Phrasen sein. Jede Einheit muss in einem gespeicherten Lexikon vorhanden sein. Die zweite Art (Fälle D, E und F) ist die interessantere. Hier steht für die Indexierung zusätzlich eine Benennungssprache zur Verfügung, die eine syntaktische Verknüpfung der Lexikoneinheiten zulässt.

	Indexate mit Mitteln der		
Indexierungssprache	Ausdrucks- formen	Lesarten- beschreibung	Begriffs- darstellung
nur Vokabular (Wörter und lexi- kalisierte Phrasen)	A Lemmata	B Lesarten-Nummer	C formale Ausdrücke
zusätzl. regelhaft gebildete Benennungen (Phrasen)	D. Phrase als Zeichenkette	E -	F formale Ausdrücke

**Tabelle 2:** Möglichkeiten der Indexierung

Der Fall A entspricht der vokabularbezogenen Indexierung mit den Lemmas, wie sie aus der gegenwärtigen Technik bekannt ist. Der Fall B berücksichtigt die Mehrdeutigkeit der Wörter bei der Indexierung. Da gegenwärtig keine

Lesartennummern in der Praxis eingeführt sind, werden bei kontrollierten Vokabularen Vorzugsbenennungen oder Lemmata mit Zusatzsymbolen verwendet, die für die jeweiligen Lesarten zu vereinbaren sind. Statt dieser Mittel kann man für eine Dokumentbeschreibung durch Lesartenmengen auch die Lesartennummern verwenden, sobald eine Verständigung über ein solches System der Lesartenidentifikation erfolgt ist. Der Fall C stellt die Indexierung mit formalen Ausdrücken für Begriffe dar. Er hat gegenüber dem Fall B den Vorteil, dass die Indexierung nicht in einer bestimmten Einzelsprache (Deutsch, Englisch usw.) erfolgt.

Der Fall D besagt, dass einem Dokument als Indexat ein beliebiges Thema in Form einer sprachlichen Benennung als Zeichenkette zugewiesen wird. Damit kann zwar wesentlich präziser als mit Wortmengen indexiert werden, aber die Wiederauffindbarkeit solcher Dokumente kann nur durch eine entsprechende linguistische Analyse des Benennungsausdrucks erzielt werden. Mit der Zuweisung einer das Thema treffenden Phrase ist die informationswissenschaftlich kritische Indexierungsleistung erbracht, aber die informationstechnische Recherchierbarkeit noch nicht gelöst. Durch den nachfolgenden Analyseprozess muss der Zeichenkette die begriffliche Repräsentation als formaler Ausdruck entsprechend dem Fall F zugeordnet werden. Dafür stehen heute noch keine praxisreifen Lösungen zur Verfügung. Dessenungeachtet ist die Indexierung beliebiger Themen durch formale Ausdrücke der Begriffssprache das Endziel des semantischen Retrievals. Der Fall E stellt keine brauchbare Alternative dar, weil nicht alle Ausdrücke einer Indexierungssprache, die das große Vokabular der natürlichen Sprachen verwenden, durch Lesartennummern identifiziert werden können.

Mit dem Wortmodell und der Begriffssprache sind notwendige Voraussetzungen für eine neue Recherchetechnik beschrieben worden. Diese können mit dem Programm Concepto verwendet werden. Es geht jetzt darum, weitere Erfahrungen mit der Strukturierung von großen Mengen von Wörtern zu sammeln. Die von Concepto zugelassenen Möglichkeiten der Begriffssprache werden für den Aufbau eines umfassenden Begriffssystems für Retrieval in einer methodisch kontrollierten Weise eingesetzt. Dazu gehören Beschränkungen beim Aufbau von Begriffsstrukturen, Festlegung auf ein bestimmtes Relationeninventar, Regeln für die Wortarteninterpretation u. a.

Die etablierte Retrievaltechnik hat die Vorteile, robust und voll operationalisierbar zu sein. Beim semantischen Retrieval wird eine intellektuelle Indexierung nötig sein, solange eine automatische Themenzuweisung für Texte noch nicht in einsatzfähigem Zustand entwickelt ist. Die kontinuierlichen Aktivitäten der linguistischen Komplementierung der etablierten Recherchetechnik durch Stammformenbildung, Phrasenextraktion u. a. (Perez-Carballo und Strzalkowski), die Einführung von Sacherschließungsinformation bei Metadaten für Webseiten, die Möglichkeiten, mit XML Dokumente "ontologiebasiert" inhaltlich zu markieren (Rabarijaona et. al.) und anderes deuten aber auf eine Wende hin. Man erkennt, dass man für qualifizierte Information auch einen kleinen intellektuellen Aufwand treiben sollte. Der professionelle Benutzer wägt schließlich ab, ob seine

Rechercheergebnisse in einem angemessenen Verhältnis zu seinem Zeitaufwand und seinen Kosten für die Informationssuche stehen.

## **Literatur**

Klavans, Judith L.: Visions of the Digital Library: Views on Using Computational Linguistics and Semantic Nets in Information Retrieval. Current Issues in Computational Linguistics: In Honour of Don Walker. Ed. By Antonio Zampolli, Nicoletta Calzolari, Martha Palmer. Pisa 1994, p. 227-236.

Lein, Hendrik: Aspekte der Realisierung des semantischen Retrievals. In: Blick Europa! Informations- und Dokumenten-Management. Deutsche Dokumentartag 1994. Trier, 27.-30.9.1994. Herausgegeben von der Deutschen Gesellschaft für Dokumentation, Frankfurt 1994.

Mladenic, Dunja: Text-Learning and Related Intelligent Agents: A Survey. IEEE Intelligent Systems. July/August 1999, p. 44-54

Perez-Carballo, Jose; Tomek Strzalkowski: Natural Language Information Retrieval: Progress Report. In: Information Processing and Management 36 (2000), p. 155-178

Rabarijaona, Auguste; Rose Dieng, Oliver Corby; Rajae Quaddari: Building and Searching an XML-Based Corporate Memory. IEEE Intelligent Systems, May/June 2000, p. 56-63.

Rahmstorf, Gerhard: Use of Semantic Networks for Information Retrieval. In: G. Rahmstorf; M. Ferguson (ed.): Proceedings of a Workshop on Natural Language for Interaction with Data Bases, IIASA (International Institute for Applied Systems Analysis), January 10-14, 1977, Laxenburg 1978.

Rahmstorf, Gerhard: Semantisches Information Retrieval. In: Blick Europa! Informations- und Dokumenten-Management. Deutsche Dokumentartag 1994. Trier, 27.-30.9.1994. Herausgegeben von der Deutschen Gesellschaft für Dokumentation, Frankfurt 1994.

Raphael, Bertram: SIR, A Computer Program for Semantic Information Retrieval. In: Minsky, M. (ed.): Semantic Information Processing. Cambridge, MA, 1968, p. 33-145.

Schank, Roger; Janet Kolodner; Gerald DeJong: Conceptual Information Retrieval. New Haven, Connecticut. Yale University, Research Report 190, December 1980.



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 89 – 111

# **Informationskompetenz am Beispiel einer szientometrischen Untersuchung zum Informationsmanagement**

**Christian Schlögl**

Karl-Franzens-Universität Graz  
Institut für Informationswissenschaft  
Universitätsstraße 15/F3, A-8010 Graz  
Tel. 0043(0)316/380-3566  
email: christian.schloegl@kfunigraz.ac.at

## **Zusammenfassung**

In diesem Beitrag wird eine szientometrische Studie zum Informationsmanagement vorgestellt. Unter Verwendung von Science Citation Index<sup>®</sup> und Social Science Citation Index<sup>®</sup> wurde die Literatur zum Informationsmanagement nach verschiedenen Kriterien (Fachgruppen, in denen publiziert wird; Herkunft der Autoren; Publikationssprachen; zeitliche Verteilung) ausgewertet. Darüber hinaus wurde eine Zitatenaanalyse durchgeführt. Auf Basis einer Autoren-Kozitationsanalyse wurde schließlich die formale Wissenschaftskommunikation im Bereich des Informationsmanagements abgebildet. Neben den Ergebnissen werden aber auch die Problembereiche aufgezeigt, die mit szientometrischen Untersuchungen verbunden sind.

## **Abstract**

This paper presents a scientometric study about information management. Both the Science Citation Index<sup>®</sup> and the Social Science Citation Index<sup>®</sup> were used to analyse the literature about information management by various criteria like journal subject category, origins of the authors, publication languages, chronological distribution. Furthermore, a citation analysis was conducted. On the basis of an author co-citation analysis, it was tried to map scholarly communication in information management. Besides presenting the results, special emphasis is given to all the problems encountered during the scientometric study.



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

## 1. Was ist Szientometrie

In einem Artikel zum State-of-the-Art definiert Van Raan Szientometrie<sup>1</sup> folgendermaßen: „Scientometric research is devoted to quantitative studies of science and technology“<sup>2</sup>. Die Szientometrie befasst sich also mit der Anwendung exakter Messmethoden bei der Untersuchung der Wissenschaft.<sup>3</sup>

Die Anwendungsgebiete sind äußerst vielfältig. Sie lassen sich grob in folgende Bereiche einteilen:

- Auswahl und Beurteilung von Buch- und Zeitschriftenbeständen,
- Ermittlung unterschiedlicher Charakteristika der Literatur (zum Beispiel einer bestimmten wissenschaftlichen Disziplin): Wachstum, Halbwertszeit, länderweise Vergleiche, ...
- wissenschaftshistorische, -soziologische und sonstige Anwendungen.<sup>4</sup>

Von Bedeutung ist die Szientometrie weiters für Forschungsförderung und -evaluation, wenn auch bei der Interpretation der Ergebnisse Vorsicht geboten ist und noch weitere Faktoren berücksichtigt werden sollten.<sup>5</sup>

## 2. Konzeption der szientometrischen Studie

Hauptziel der Untersuchung war es, mehr Klarheit über die Literatur zum Informationsmanagement (IM) zu gewinnen. Unter anderem sollten durch die szientometrische Untersuchung folgende Fragen beantwortet werden:

- In welchen wissenschaftlichen Disziplinen wird der Begriff Informationswissenschaft verwendet?
- In welchen Sprachen wird hauptsächlich publiziert? Aus welchen Ländern kommen die publizierenden Autoren?
- Handelt es sich bei Informationsmanagement um ein Modethema?
- Welche Autoren haben den höchsten „Impact“ („Citation Classics“)?
- Wie sieht die formale Wissenschaftskommunikation im Bereich des Informationsmanagements aus?

---

<sup>1</sup> Der Begriff wurde ursprünglich von den beiden sowjetischen Wissenschaftlern Nalimov und Mul'cenko geprägt.

<sup>2</sup> Van Raan 1997, S. 205.

<sup>3</sup> Vgl. Schmidmaier 1983, S. 17 f.

<sup>4</sup> Vgl. Lawani 1981, S. 302 ff.

<sup>5</sup> Vgl. zum Beispiel Daniel 1988, S. 238; Daniel/Fisch 1988, S. 148; Fröhlich 1999, S. 31 ff u. S. 36.

Das Forschungsdesign basierte auf folgenden Anforderungen:

1. Aus auswertungstechnischen Gründen und dem Stand der Technik entsprechend sollten elektronische Medien (Datenbanken) als Basis für die Durchführung der szientometrischen Untersuchung dienen.
2. Die Untersuchung sollte möglichst breit angelegt werden. Dadurch sollte keine Disziplin ausgeschlossen werden, die sich (auch) mit dem Gebiet des Informationsmanagements auseinandersetzt.<sup>6</sup>
3. Die zugrunde liegende(n) Datenbank(en) sollte(n) auch „erweiterte“ szientometrische Analysen, speziell Zitatensanalysen ermöglichen.

Die Zitationsdatenbanken des Institute for Scientific Information® (ISI®) werden allen diesen Anforderungen gerecht<sup>7</sup>. Science Citation Index® (SCI®), Social Science Citation Index® (SSCI®) und Arts & Humanities Citation Index® (A&HCh®) decken mit Natur-, Sozial- sowie Geisteswissenschaften und Kunst das gesamte Spektrum wissenschaftlicher Forschung ab. Es handelt sich also um multidisziplinäre Datenbanken.<sup>8</sup>

Bei der szientometrischen Untersuchung wurde folgende Vorgehensweise gewählt: Nach der Festlegung des Untersuchungszeitraums wurden die IM-Kernpublikationen ermittelt, die nach verschiedenen Kriterien (disziplinenmäßige Verteilung, Publikationssprachen, ...) ausgewertet wurden. Dem Grundgedanken von Zitierindizes entsprechend wurde dann eine Zitatensanalyse durchgeführt. Bei dieser wurden die (von den IM-Kernpublikationen) am öftesten zitierten Autoren einer Autoren-Kozitationsanalyse unterzogen.

In der Literatur wird vielfach die Meinung vertreten, dass der 1980 von der US-Bundesregierung erlassene Paperwork Reduction Act bzw. die ihm vorausgegangenen Arbeiten der „Federal Paperwork Commission“ die Anfänge des Informationsmanagements markieren<sup>9</sup>. Da die ISI®-Zitationsdatenbanken auf den meisten Hosts erst ab den siebziger Jahren verfügbar sind,<sup>10</sup> wurde bei der Suche keine zeitliche Einschränkung gemacht.

Nachdem ein Hauptziel der Studie darin bestand, die Verwendung der Bezeichnung Informationsmanagement in einzelnen wissenschaftlichen Disziplinen zu untersuchen und in weiterer Folge verschiedene IM-Ansätze

---

<sup>6</sup> Wäre zum Beispiel die Datenbank BLISS szientometrisch ausgewertet worden, so wäre die Studie von Beginn an primär auf betriebswirtschaftliche Informationsmanagement-Ansätze fokussiert gewesen.

<sup>7</sup> Eine grundlegende Einführung findet sich bei Garfield (1983).

<sup>8</sup> Vgl. Garfield 1995.

<sup>9</sup> Vgl. zum Beispiel Herget 1997, S. 785.

<sup>10</sup> In DIALOG®: SCI® ab 1974 (siehe <http://library.dialog.com/bluesheets/html/b10034.html>), SSCI® ab 1972 (siehe <http://library.dialog.com/bluesheets/html/b100007>), A&HCI ab 1980 (siehe <http://library.dialog.com/bluesheets/html/b10439.html> — Stand jeweils 2. September 1999).

herauszuarbeiten, sollte zunächst keine Zitationssuche durchgeführt werden. Bei einer solchen hätte man von einigen Standardpublikationen bzw. –autoren ausgehen müssen. Wenn nun die Auswahl von Standardwerken nicht repräsentativ gewesen wäre, dann hätte die Gefahr bestanden, dass letzten Endes nicht alle Dimensionen des Informationsmanagements ermittelt worden wären.<sup>11</sup> Dies vor allem dann, wenn einzelne Bereiche der IM-Literatur nicht oder kaum aufeinander Bezug nehmen.

Da sich die zu untersuchende Dokumentenmenge *primär* mit IM-Inhalten auseinander setzen sollte, wurde eine Titelsuchung verwendet. Eine Suche nach Schlagworten kam aufgrund der schlechten Qualität der Beschlagwortung<sup>12</sup> und ihrer Verwendung erst ab 1991 nicht in Frage. Auch Abstracts werden in SCI® und SSCI® erst ab 1991 bzw. 1992 erfasst.<sup>13</sup> Die Suche nach Stichworten im Abstract wurde auch deshalb nicht durchgeführt, weil dadurch weniger relevante IM-Publikationen bei der szientometrischen Untersuchung berücksichtigt worden wären.

Nachdem davon ausgegangen wurde, dass sich verschiedene Disziplinen mit dem Informationsmanagement auseinander setzen bzw. dass es unterschiedliche IM-Ansätze gibt, sollten bei der Titelsuchung möglichst "neutrale" Bezeichnungen verwendet werden. "Informatik-Management" hätte beispielsweise eher technikzentrierte Ansätze bevorzugt. Aus diesem Grund wurde nach allen Publikationen gesucht, die im Titel entweder die Bezeichnung "INFORMATION MANAGEMENT" oder "INFORMATION RESOURCE? MANAGEMENT"<sup>14</sup> enthielten. Ausgeschlossen werden sollten Artikel mit "INFORMATION MANAGEMENT SYSTEM?" im Titel, da die Vorabrecherche ergab, dass es bei Artikel mit diesem Titel primär um Software und weniger um inhaltliche Aspekte des Informationsmanagements geht.

Bei den Quellenartikeln<sup>16</sup> unterscheidet das ISI® zwischen verschiedenen Dokumenttypen: Artikel, Review-Artikel, Buchbesprechung, Editorial, ... Da nur (wissenschaftliche) Aufsätze berücksichtigt werden sollten, und auch in Hinblick auf die später geplanten Zitatensanalysen, wurde die Abfrage auf Artikel

---

<sup>11</sup> Vergleiche dazu McCain, die speziell in Hinblick auf Autoren-Kozitationsanalysen meint: „If the authors are not chosen to capture the full range of variability in subject specializations, methodologies, political orientations, etc., these aspects of structure cannot be demonstrated.“ (McCain, S. 433)

<sup>12</sup> Zur Beschlagwortung werden „Author Keywords“ und „Keywords Plus“ verwendet. Die „Author Keywords“ werden von den Autoren der einzelnen Artikel selbst vergeben und sind dementsprechend inkonsistent. „KeyWords Plus“ enthalten Worte oder Phrasen, die häufig im Titel der Referenzen vorkommen. Sie werden dann angegeben, wenn sie weder im Titel noch in den „Author Keywords“ vorkommen.

<sup>13</sup> Ein Kurzreferat ist allerdings nicht bei allen Artikeln vorhanden. Im A&HCI® werden Kurzreferate nach wie vor nicht angeboten.

<sup>14</sup> Beim Fragezeichen handelt es sich bei DIALOG® um ein sogenanntes Trunkierungssymbol. Es ist eine Art Platzhalter für null, ein oder mehrere beliebige Zeichen.

<sup>15</sup> Bei den Abfragen handelt es sich um eine sogenannte Phrasensuche. Es werden also nur solche Dokumente gefunden, bei denen die Wörter INFORMATION MANAGEMENT bzw. INFORMATION RESOURCE? MANAGEMENT im Titel unmittelbar hintereinander stehen.

<sup>16</sup> In den Zitationsdatenbanken des ISI® werden ausschließlich Periodika erfasst.

eingeschränkt. Die nun so ausgewählten Artikel werden im restlichen Teil dieser Arbeit IM-ARTIKEL<sup>17</sup> bezeichnet.

Die Online-Recherchen wurden mit DIALOG durchgeführt. Die Hauptrecherche fand am 30. Juni 1999 statt. Ergänzende Daten wurden am 27. Juli 1999 recherchiert. Die Rohdaten für die Kozitationsanalyse wurden am 1. September 1999 ermittelt.<sup>18</sup>

Wie anschließend gezeigt wird, muss bei der Durchführung von szientometrischen Untersuchungen eine Reihe von auftretenden Problemen beachtet werden. Diese beziehen sich sowohl auf die methodische Vorgehensweise als auch auf die verwendeten „Instrumente“ (Datenbanken) und wirken sich unterschiedlich stark auf die Ergebnisse aus. Mit entsprechender Informationskompetenz können derartige Problembereiche teilweise umgangen, zumindest aber bei der Interpretation der Ergebnisse berücksichtigt werden. In den folgenden Ausführungen werden die im Rahmen dieser szientometrischen Analyse aufgetretenen Problembereiche kursiv dargestellt.<sup>19</sup>

Nachfolgend werden nun die Ergebnisse vorgestellt. Es erfolgt dabei eine Trennung zwischen Auswertungen, die auf Häufigkeitsverteilungen beruhen, und Zitatenanalyse.

### **3. Ergebnisse**

#### **3.1 Häufigkeitsverteilungen**

##### **1) IM-ARTIKEL in Geistes-, Sozial- und Naturwissenschaften**

Wie bereits erwähnt wurde, sollte die Recherche zunächst möglichst viele wissenschaftliche Disziplinen einschließen. Aus diesem Grund wurden alle drei Zitationsdatenbanken des ISI® (SCI®, SSCI® und A&HCI®) in die Abfrage einbezogen. Dies ermöglichte eine erste grobe Abschätzung, wie sich die IM-ARTIKEL auf Geistes-, Sozial- und Naturwissenschaften verteilen.

---

<sup>17</sup> Der entsprechende DIALOG-Suchbefehl zur Bestimmung dieser Artikel lautet: SELECT (( INFORMATION 0 MANAGEMENT (NOT W) SYSTEM?) OR INFORMATION 0 RESOURCE? 0 MANAGEMENT » ART, TI

<sup>18</sup> Für die Durchführung der Online-Recherche bedanke ich mich sehr herzlich bei Frau Dr. Mayerl von der Informationsvermittlungsstelle der Universitätsbibliothek Graz.

Wertvolle Anregungen bei der Konzeption der Studie und Diskussion der Ergebnisse lieferten Herr Dr. Gerhard Fröhlich und Herr Dr. Juan Gorraiz, bei denen ich mich ebenfalls vielmals bedanke.

<sup>19</sup> Die hier vorliegende Publikation erhebt keinen Anspruch darauf, alle möglichen Problembereiche, die bei szientometrischen Untersuchungen mit den ISI-Zitationsdatenbanken auftreten können, aufzuzeigen.



	IM	IRM	IM+IRM
A&HCI <sup>®</sup>	9	2	11
SSCI <sup>®</sup>	250	26	276
SCI <sup>®</sup>	234	28	262
$\Sigma^{20}$	493	56	549

**Tabelle 1:** Anzahl IM-ARTIKEL (IM+IRM) in SCI<sup>®</sup>, SSCI<sup>®</sup> und A&HCI<sup>®</sup>

Aus Tabelle 1 kann entnommen werden, dass die Bezeichnung „Information Management“ (IM) wesentlich häufiger als „Information Resource Management“ bzw. „Information Resources Management“ (IRM) verwendet wird. Weiters ist ersichtlich, dass in den Geisteswissenschaften nur wenige IM-ARTIKEL publiziert wurden. Aus diesem Grund wurde der A&HCI<sup>®</sup> in den folgenden Auswertungen nicht mehr berücksichtigt. In SCI<sup>®</sup> und SSCI<sup>®</sup> sind in etwa gleich viele IM-ARTIKEL enthalten. Demnach könnte man die Vermutung anstellen, dass IM-Aufsätze entweder technisch/naturwissenschaftlich oder eher sozialwissenschaftlich orientiert sind.

*Es muss aber beachtet werden, dass sich A&HCI<sup>®</sup>, SCI<sup>®</sup> und SSCI<sup>®</sup> teilweise überlappen.* Im konkreten Fall sind 43 IM-ARTIKEL sowohl im SCI<sup>®</sup> als auch im SSCI<sup>®</sup> enthalten.

## **2) IM-ARTIKEL in unterschiedlichen (Zeitschriften)Fachgruppen und wissenschaftlichen Disziplinen**

Die grobe Zuordnung von IM-ARTIKEL zu den Natur- bzw. Sozialwissenschaften soll nun etwas genauer untersucht werden. Dazu können die sogenannten (Zeitschriften)Fachgruppen des ISI<sup>®</sup> herangezogen werden. Bei diesen handelt es sich um eine Zusammenfassung von jeweils fachlich verwandten Zeitschriften.<sup>21</sup>

---

<sup>20</sup> Diese Werte enthalten noch Duplikate (zum Beispiel, wenn derselbe IM-ARTIKEL sowohl im SCI<sup>®</sup> als auch im SSCI<sup>®</sup> enthalten ist).

<sup>21</sup> Alle in SCI<sup>®</sup>, SSCI<sup>®</sup> und A&HCI<sup>®</sup> erfassten Quellenzeitschriften werden vom ISI jährlich unter dem Namen „Journal Citation Reports“ (JCR<sup>®</sup>) veröffentlicht und mittlerweile auch schon auf CDROM vertrieben.

Rang	IM-Artikel	(Zeitschriften)Fachgruppe
1	183	INFORMATION SCIENCE & LIBRARY SCIENCE
2	44	COMPUTER APPLICATIONS & CYBERNETICS
3	22	MANAGEMENT
4	19	EDUCATION & EDUCATIONAL RESEARCH
5	16	INSTRUMENTS & INSTRUMENTATION
6	14	COMPUTER SCIENCE, INFORMATION SYSTEMS
7	10	CHEMISTRY, ANALYTICAL
8	9	HEALTH POLICY & SERVICES
9	8	COMPUTER SCIENCE, INTERDISCIPLINARY APPLIC.
9	8	ENGINEERING
9	8	ENGINEERING, ELECTRICAL & ELECTRONIC
9	8	MEDICINE, MISCELLANEOUS
9	8	PUBLIC ADMINISTRATION
14-101	7-1	Verschiedene

**Tabelle 2:** Anzahl IM-ARTIKEL in Zeitschriftenfachgruppen (mit mehr als acht IM-ARTIKELN)

Aus Tabelle 2 geht hervor, dass IM-ARTIKEL in zahlreichen unterschiedlichen Fachgruppen publiziert werden. Mit Abstand am häufigsten finden sich IM-ARTIKEL allerdings in der Informations-/Bibliothekswissenschaft.

*Wie aus obiger Tabelle zu erkennen ist, entsprechen die Zeitschriftenfachgruppen im System der Wissenschaft unterschiedlich hierarchischen Ebenen. Teilweise decken sie eine wissenschaftliche Disziplin (z. B. INFORMATION SCIENCE & LIBRARY SCIENCE), meistens aber nur Spezialgebiete davon (z. B. COMPUTER SCIENCE, INFORMATION SYSTEMS) ab.<sup>22</sup> Ein direkter Vergleich der Zeitschriftenfachgruppen ist daher problematisch.*

*Weiter ist zu beachten, dass eine Zeitschrift mehreren Fachgruppen zugeordnet sein kann. Im konkreten Fall wurden die 495 IM-ARTIKEL 562 Fachgruppen zugewiesen. Aus diesen Gründen können die Zeitschriftenfachgruppen nur der groben fachlichen Orientierung dienen.*

<sup>22</sup> Die Größe der einzelnen (Zeitschriften)Fachgruppen, d. h. die Anzahl der in ihnen enthaltenen Quellenartikel, variiert teilweise stark, z. B. CHEMISTRY: 1.680.538 Artikel, INFORMATION SCIENCE & LIBRARY SCIENCE: 156.907 Artikel, COMPUTER APPLICATIONS & CYBERNETICS: 158.069 Artikel, COMPUTER SCIENCE, INFORMATION SYSTEMS: 24.528 Artikel (Stand: 27. Juli 1999).

### 3) Sprachliche und länderweise Verteilung der IM-Artikel

Eine weitere Problematik der ISI<sup>®</sup>-Zitationsdatenbanken bringt die Reihung der IM-ARTIKEL nach der Publikationssprache zutage. Demnach wurden 90 % der Artikel in englisch veröffentlicht. Damit liegt der Anteil der in englischer Sprache publizierten IM-ARTIKEL aber noch immer unter dem Durchschnittswert des SCI<sup>®</sup> (95 % für das Jahr 1997).<sup>23</sup>

Rang	IM-ARTIKEL	Sprache
1	449	englisch
2	41	deutsch
3	5	Sonstige

**Tabelle 3:** Anzahl IM-ARTIKEL in verschiedenen Sprachen

Die Herkunft der Autoren ist ebenfalls stark ungleich-verteilt. 50 % der Autoren arbeiteten in Organisationen (in der Regel Universitätsinstitute) mit Sitz in den USA, ein Fünftel kommt aus Großbritannien. Diese Anteile sind selbst für die Zitationsdatenbanken des ISI<sup>®</sup> recht hoch. Beispielsweise kam 1997 im Durchschnitt die Hälfte der Artikel des SCI<sup>®</sup> (im Web of Science<sup>®</sup>) aus Ländern mit der Muttersprache englisch.<sup>24</sup>

Im Bereich des Informationsmanagements ist die anglo-amerikanische Dominanz, zumindest was die Herkunft der Autoren betrifft, also noch größer. Aus diesem Grund beziehen sich die Ergebnisse der hier durchgeführten szientometrischen Untersuchung primär auf den anglo-amerikanischen Raum.

Rang	IM-ARTIKEL	Land
1	223	USA
2	87	Grossbritannien
3	42	Deutschland
4	14	Kanada
4	14	Niederlande
6	12	Australien
7-39	39	Diverse

**Tabelle 4:** Anzahl IM-ARTIKEL, länderweise (Organisation des Autors) Verteilung

Bei obiger Auswertung muss allerdings berücksichtigt werden, dass nur 431 (der 495) Datensätze eine Landesbezeichnung hatten. Von einer völlig korrekten Erfassung der Landesbezeichnung kann auch nicht ausgegangen werden. Beispielsweise wurden Autoren aus Deutschland unter folgenden

<sup>23</sup> [http://165.123.33.33/eugene\\_garfield/papershnapsciworld.html](http://165.123.33.33/eugene_garfield/papershnapsciworld.html) (Stand: 29. 1. 1999)

<sup>24</sup> [http://165.123.33.33/eugene\\_garfield/papers/mapsociworld.html](http://165.123.33.33/eugene_garfield/papers/mapsociworld.html) (Stand: 29. 1. 1999)

Bezeichnungen gefunden: „Federal Republic of Germany“ (16 Mal), „Germany“ (15 Mal), „Fed Rep Ger“ (6 Mal), „FedRepGer“ (5 Mal).

### 3) Zeitliche Verteilung der IM-ARTIKEL

Wie aus Abbildung 1 ersichtlich ist, steigt die Anzahl der IM-ARTIKEL pro Jahr nach dem Inkrafttreten des Paperwork Reduction Act von 1980 stark an und bleibt in den Folgejahren relativ konstant. Dies deutet darauf hin, dass es sich bei Informationsmanagement um kein Modethema handelt.

#### 3.2 Zitatenanalyse

Bei der Zitatenanalyse soll nun anhand der Referenzen der IM-ARTIKEL die formale Wissenschaftskommunikation im Bereich des Informationsmanagements genauer durchleuchtet werden. Dabei wird folgendermaßen vorgegangen:

- Zunächst werden jene Autoren ermittelt, die von den IM-ARTIKELN am öftesten zitiert wurden.
- Danach wird mit diesen eine Autoren-Kozitationsanalyse<sup>25</sup> durchgeführt.

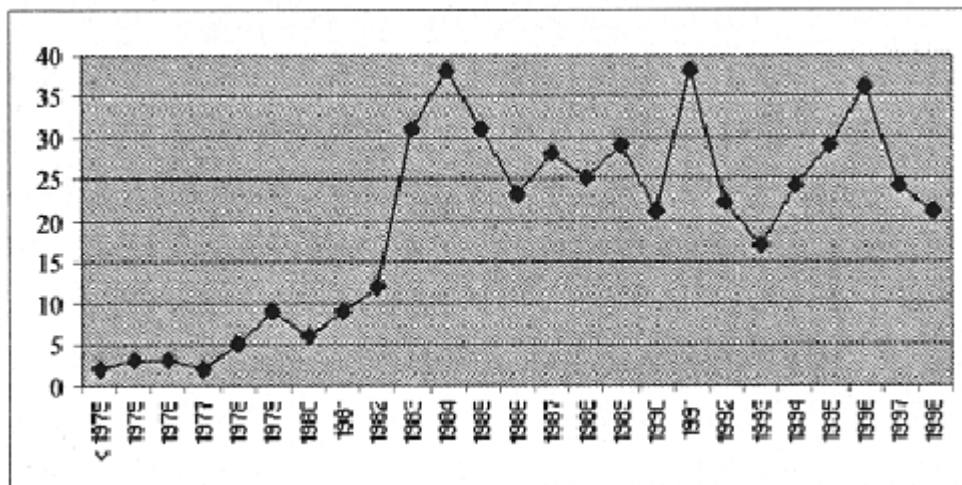


Abbildung 1: Zeitliche Verteilung der IM-ARTIKEL

Eine Zitatenanalyse baut auf der Grundannahme auf, dass ein Zitat eine inhaltliche Beziehung zwischen zitiertem und zitierendem Dokument darstellt. Demnach werden Dokumente zitiert, „... that support, provide precedent for, illustrate, or elaborate on what the author has to say“.<sup>26</sup> Von einer Homogenität der Zitate darf allerdings nicht ausgegangen werden. Tatsächlich gibt es eine Reihe von unterschiedlichen Motiven für die Verwendung von Zitaten. Weinstock unterscheidet zum Beispiel 15 Gründe. Das Spektrum reicht von Anerkennung bis zu Kritik einer früheren Arbeit.<sup>27</sup> Zitate sind somit als ein allgemeines Maß für den

<sup>25</sup> Die Abbildung der formalen Wissenschaftskommunikation ist nur ein Anwendungsgebiet der Zitatenanalyse. Weitere Anwendungen werden in Smith (1981, S. 94 ff.) beschrieben.

<sup>26</sup> Garfield 1983, S. 1.

<sup>27</sup> Vgl. Weinstock 1971, entnommen aus Cronin 1984, S. 30.

*Beitrag eines Forschers im Rahmen der „Wissenschaftspraxis“ zu sehen.<sup>28</sup> - Nicht mehr und nicht weniger Unabhängig davon, wieviele Zitate eine negative oder positive Semantik haben, entscheidend ist letztendlich, wieviele Wissenschaftler sich mit einem Autor (gemeinsam mit anderen) in Summe auseinander gesetzt haben.<sup>29</sup>*

*Gegen Zitatensanalysen werden darüber hinaus noch folgende Kritikpunkte<sup>30</sup> vorgebracht: Weglassen von Zitaten; Verzerrungen als Folge von Einverleibung, Gefälligkeitszitaten oder Zitierung von bekannten (Review)Artikeln anstatt der Originalarbeit<sup>31</sup> oder Selbstzitierungen<sup>32</sup>. Ein weiteres Problem stellen Zitate dar, die ohne Konsultation der Originalquellen anderen Veröffentlichungen entnommen wurden.<sup>33</sup> Diese Kritikpunkte wirken sich auch auf Zitatensanalysen aus. Teilweise können sie umgangen werden (z. B. bei Selbstzitierungen), teilweise halten sich die negativen Einflüsse in gewissen Grenzen, da szientometrische Analysen in der Regel auf stark aggregierten Daten basieren.*

### **1) Autoren mit dem größten „Impact“<sup>34</sup>**

Von den 495 IM-ARTIKEL enthielten 145 keine Referenzen. Die restlichen 350 Artikel zitierten insgesamt 5872 verschiedene Publikationen. Die durchschnittliche Länge der Literaturliste liegt damit bei rund zwölf Referenzen je IM-ARTIKEL. Eine Reihung der am öftesten zitierten Autoren, die für die Autoren-Kozitationsanalyse ausgewählt wurden, zeigt Tabelle 5.

---

<sup>28</sup> Vgl. Garfield 1979, S. 362.

<sup>29</sup> Vgl. Stock 1985, S. 312; White 1990, S. 89 f.

<sup>30</sup> Eine umfassende Darstellung der Problembereiche von Zitatensanalysen findet sich u. a. bei Gorraiz (1992, S. 197 ff.) und MacRoberts/MacRoberts (1989, S. 343 f.).

<sup>31</sup> Dieses Phänomen ist laut Garfield in der Literatur unter der Bezeichnung Matthäuseffekt bekannt (vgl. Garfield 1989, S. 30). Ursprünglich wurde der Begriff von Merton geprägt, der damit die ungleiche Würdigung von Wissenschaftlern bei Kooperationen oder gleichzeitig gemachten Entdeckungen charakterisierte (vgl. Merton 1973, S. 443 ff.).

<sup>32</sup> Selbstzitate machen einen relativ hohen Anteil an den Gesamtzitaten aus. Beispielsweise errechnete Vinkler bei einer Analyse von 20 Chemie-Aufsätzen den Anteil von Selbstzitierungen mit 22,3 % (vgl. Vinkler 1987, S. 52).

<sup>33</sup> Vgl. Broadus 1983.

<sup>34</sup> Impact wird in dem Sinn verwendet, in wie weit sich andere Wissenschaftler (durch Zitate) mit einem Autor auseinander gesetzt haben (siehe oben).

Rang	erhaltene Zitate	Autor	Rang	erhaltene Zitate	Autor
1	31	HORTON F(W) <sup>35</sup>	11	10	MINTZBERG H
2	17	CRONIN B	14	9	NOLAN RL
3	15	PORTER ME	14	9	SYNNOTT WR
3	15	MARCHAND	14	9	CASH J(I)
5	14	MCFARLAN	17	8	DICKSON GW
6	13	DRUCKER PF	17	8	ROBERTS N
6	13	ROCKART	17	8	TRAUTH E(M)
8	12	SIMON HA	20	7	HAMMER M
9	11	EARL M(J)	20	7	IVES B
9	11	WILSON T(D)	20	7	KUHLEN R
11	10	LUCAS HC	20	7	VICKERS P
11	10	MARTIN J	20	7	WISEMAN C

**Tabelle 5:** Anzahl der (von IM-ARTIKELN) erhaltenen Zitate (Autoren, die öfter als sechs Mal zitiert wurden)

Dabei muss allerdings berücksichtigt werden, dass in der Online- und CDROM-Version von SCI® und SSCI® bei den zitierten Publikationen nur der erstgenannte Autor erfasst ist.

Das Ranking nach den zitierten Autoren brachte auch die Synonymproblematik zutage. Bei Autoren mit zwei oder mehreren Vornamen wurde in einigen Fällen nur der Anfangsbuchstabe des ersten festgehalten. Die Untersuchung einer Stichprobe (alle Autoren mit mehr als drei erhaltenen Zitaten) ergab, dass von den 184 Autoren immerhin acht doppelt (mit einer und zwei Initialen) angesetzt wurden.

Auch Homonyme stellen beim Science Citation Index® ein Problem dar. Selbst bei der in der vorliegenden Untersuchung relativ geringen Zahl an Zitaten, konnten Ansetzungen gefunden werden, die sich auf verschiedene Autoren beziehen, zum Beispiel MARTIN J.

## 2) Autoren-Kozitationsanalyse

Nachfolgend wird die Wissenschaftsstruktur im Bereich des Informationsmanagements genauer untersucht. Es soll also eine Art „Soziogramm von Druckerzeugnissen“<sup>36</sup> erstellt werden, das die Struktur der Informationsbeziehungen innerhalb des Informationsmanagements darstellt. Dies erfolgt in Form einer Autoren-Kozitationsanalyse.

<sup>35</sup> Die Klammer soll zum Ausdruck bringen, dass der jeweilige Autor in den Zitaten sowohl mit einer als auch mit zwei Initialen angesetzt wurde (Synonymproblematik).

<sup>36</sup> Vgl. Stock 1985, S. 308.

Diese geht von der Grundannahme aus, dass ein Zusammenhang zwischen zwei Autoren dann besteht, wenn sie gemeinsam von einem Drittautor zitiert werden. Autor darf allerdings nicht im buchstäblichen Sinn verstanden werden, sondern als eine Menge von Publikationen, die allgemein mit einem bestimmten Autor in Zusammenhang gebracht wird. White verwendet dafür die Begriffe „body of writings“ bzw. „oeuvre“. Aus der Anzahl der Kozitate von Autorpaaren können die Beziehungen zwischen den einzelnen Autoren (je mehr Kozitate desto größer der Zusammenhang (aus der Sicht der zitieren-den Autoren)) und letztendlich gemeinsame Forschungsbereiche ermittelt werden.<sup>37</sup>

Bei der hier durchgeführten Autoren-Kozitationsanalyse wurden die von den ausgewählten Autoren erhaltenen Kozitate nicht auf IM-ARTIKEL eingeschränkt, und zwar aus folgenden zwei Gründen: Es wird davon ausgegangen, dass es sich bei Informationsmanagement um keine eigene Wissenschaftsdisziplin handelt, sondern dass es in verschiedene Disziplinen eingebettet ist. Ziel ist es nun, diese Disziplinen herauszuarbeiten, wodurch die „Zitationsquellen“ nicht nur auf IM-ARTIKEL eingeschränkt sein müssen, sondern (auf die ihnen zugrunde liegenden Disziplinen) ausgeweitet werden können (siehe Abbildung 2). Ein weiterer Grund liegt darin, dass die Anzahl der von den IM-ARTIKELN erhaltenen Zitate (siehe Tabelle 5) nicht ausreicht, um eine sinnvolle Kozitationsanalyse durchführen zu können.

Die Kozitationshäufigkeiten der ausgewählten 24 Autoren wurden auch in DIALOG recherchiert.<sup>38</sup> Für diesen Zweck stellt DIALOG den CA-Befehl<sup>39</sup> zur Verfügung. Beispielsweise werden durch den Befehl „SELECT CA = CRONIN B“ alle Artikel gefunden, die CRONIN B zitiert haben. Dabei wurde die Synonymproblematik berücksichtigt.<sup>40</sup> Die Kozitationshäufigkeiten der einzelnen Autorenpaare erhält man durch eine UND-Verknüpfung, zum Beispiel „CA = CRONIN B AND CA = KUHLEN R“. Bei der Ermittlung der Kozitationshäufigkeiten werden (fast) alle Homonyme ausgeschieden, da die Wahrscheinlichkeit sehr gering ist, dass ein gleichlautender Autor (der in der Regel nicht in derselben Wissenschaftsdisziplin tätig ist) ebenfalls mit demselben Drittautor gemeinsam zitiert wird.

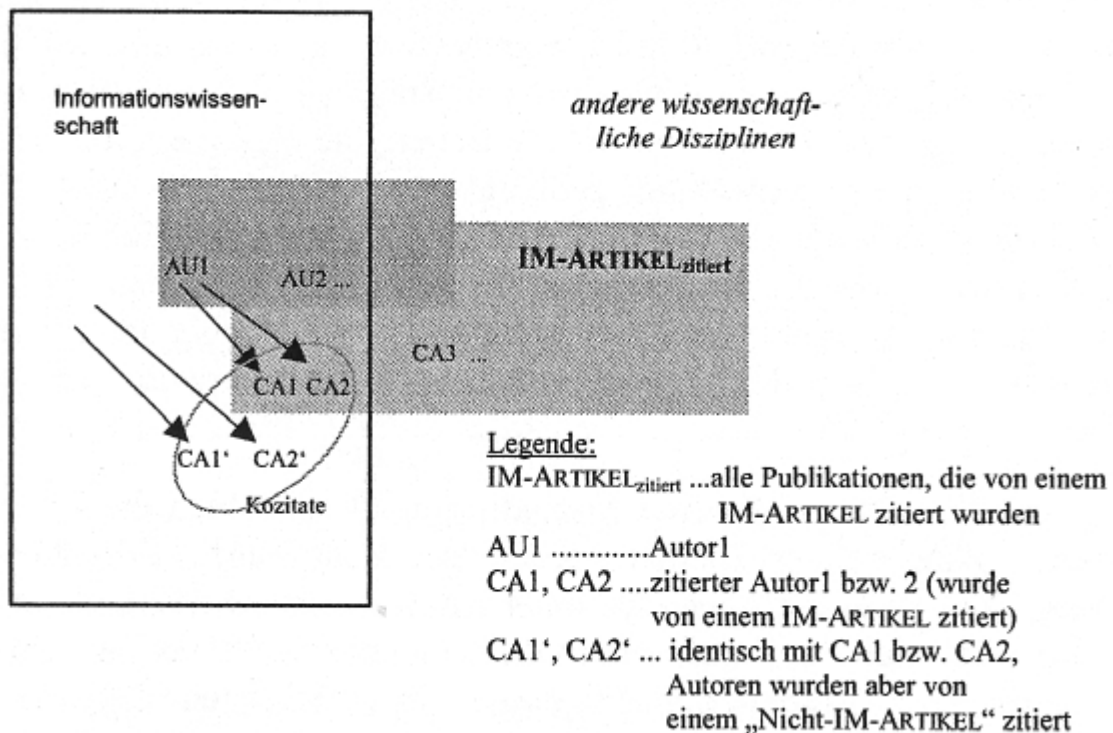
---

<sup>37</sup> Vgl. White/Griffith 1981, S. 163.

<sup>38</sup> Die Recherche fand am 1. 9. 1999 statt

<sup>39</sup> CA = Cited Author

<sup>40</sup> Zum Beispiel wurden alle Publikationen, die Horton zitiert haben, folgendermaßen ermittelt: SELECT CA = HORTON F OR CA = HORTON FW.



**Abbildung 2:** Informationsmanagement, in wissenschaftliche Disziplinen (z. B. Informationswissenschaft) eingebettet

Tabelle 6 zeigt die Kozitationsmatrix, die aus der Recherche der Kozitationshäufigkeiten resultierte. Im nächsten Schritt wurden die Rohdaten<sup>41</sup> in Ähnlichkeitsmaße transformiert. Dazu wird der Korrelationskoeffizient nach Pearson verwendet. Das Ergebnis ist die Korrelationsmatrix (aus Platzgründen in diesem Beitrag nicht dargestellt). Diese drückt die Ähnlichkeit von Kozitationsprofilen aus. Beispielsweise haben Autoren, die mit bestimmten Drittautorens oft „kozitiert“ werden, mit anderen aber seltener, eine stark positive Korrelation und dadurch aus der Sicht der zitierenden Autoren eine große Ähnlichkeit. Der Korrelationskoeffizient beseitigt auch große Unterschiede zwischen Autoren mit einer sehr hohen Zahl an erhaltenen Zitaten und solchen, die zwar weniger Zitate erhalten haben, dafür aber ein ähnliches Kozitationsprofil aufweisen.<sup>41</sup>

<sup>41</sup> Um mögliche „Instabilitäten“ aufgrund von geringen Kozitationshäufigkeiten zu vermeiden, können Grenzwerte definiert werden, die Autoren überschreiten müssen, um für die weiteren Analysen berücksichtigt zu werden. Bei der vorliegenden Kozitationsmatrix wären auf jeden Fall Kuhlen und Vickers von einem derartigen Schwellenwert betroffen gewesen. Da das Weglassen dieser beiden Autoren auf die restlichen Ergebnisse keine Auswirkungen hatte (es wurden zwei Verarbeitungsläufe durchgeführt), wurden sie in den nachfolgenden Auswertungen weiter berücksichtigt. Die geringe Kozitationshäufigkeit von Kuhlen dürfte primär an der bereits diskutierten Bevorzugung des anglo-amerikanischen Sprachraumes liegen.

<sup>41</sup> Vgl. McCain 1990, S. 436.



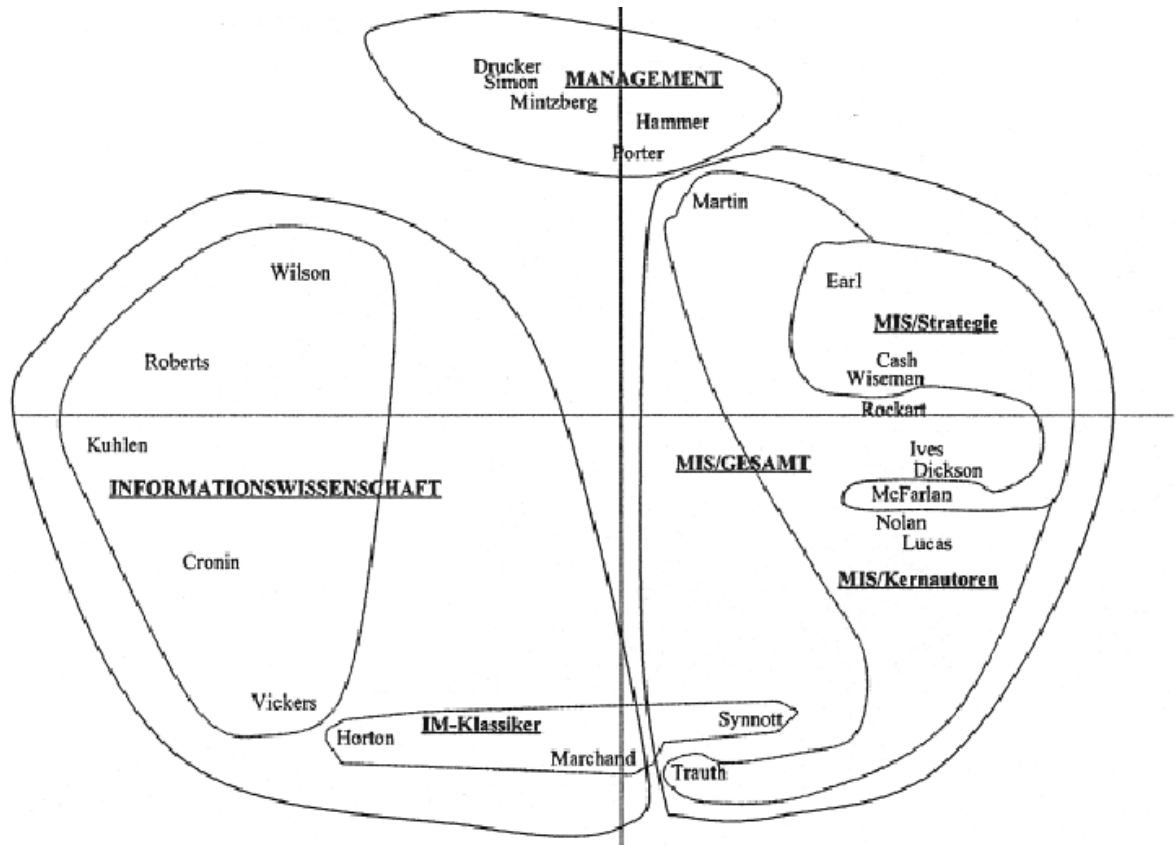
Cash																								
Cronin	1	Cronin																						
Dickson	38	1	Dickson																					
Drucker	31	14	19	Drucker																				
Earl	65	5	30	40	Earl																			
Hammer	42	1	15	126	86	Hammer																		
Horton	3	20		11	6	0	Horton																	
Ives	110	6	159	29	58	35	3	Ives																
Kuhlen	0	2	0	3	0	1	3	0	Kuhlen															
Lucas	47	5	176	18	33	21	14	296	0	Lucas														
Marchand	5	8	4	9	8	3	21	4	2	16	Marchand													
Martin	33	2	42	63	42	85	10	71	3	63	7	Martin												
McFarlan	132	4	64	37	100	44	9	182	0	122	14	64	McFarlan											
Mintzberg	47	2	63	392	92	114	8	100	0	98	6	136	67	Mintzberg										
Nolan	50	2	54	27	45	34	7	108	0	122	14	66	127	52	Nolan									
Porter	176	19	37	343	120	151	8	181	1	64	17	99	224	820	80	Porter								
Roberts	0	18	1	9	2	0	7	0	2	4	3	4	1	13	1	7	Roberts							
Rockart	92	6	83	74	87	75	11	177	0	115	15	114	154	135	127	170	3	Rockart						
Simon	16	12	90	322	22	53	9	76	2	108	3	136	38	798	44	384	13	80	Simon					
Synnott	10	7	8	10	12	3	14	15	0	19	14	14	16	7	17	21	3	20	2	Synnott				
Trauth	5	4	7	2	6	5	9	8	0	9	8	8	5	6	5	7	2	12	4	7	Trauth			
Vickers	1	8	0	0	2	1	10	1	0	3	3	2	4	1	1	3	3	1	1	3	2	Vickers		
Wilson	5	24	5	9	16	7	7	4	2	7	3	10	11	31	5	20	35	11	48	0	2	5	Wilson	
Wiseman	64	3	25	18	31	14	3	71	0	19	5	27	96	28	33	114	1	56	16	11	2	2	6	

**Tabelle 6:** KozitationsmatrixAbb. 3: Wissenschaftslandkarte zum Informationsmanagement

## Multidimensionale Skalierung

Mit Hilfe der multidimensionalen Skalierung werden die Autoren auf der Basis von Ähnlichkeitsmaßen in einem mehrdimensionalen Raum dargestellt. Dies erfolgt nach folgenden Grundsätzen:

- Autoren mit hohen Kozitationshäufigkeiten werden in geringen Abständen zueinander positioniert. Autoren, die große Unterschiede aufweisen, werden im Wissenschaftsraum in großer Entfernung voneinander gezeichnet.



**Abbildung 3:** Wissenschaftslandkarte zum Informationsmanagement

- Autoren, die "Beziehungen"<sup>43</sup> zu vielen anderen Autoren haben, werden im Zentrum lokalisiert. Hingegen werden Autoren, die zu den meisten anderen Autoren keine Verbindung haben, peripher dargestellt. Dem entsprechend kann man auch zwischen zentralen und peripheren Fachbereichen oder Fachgruppen unterscheiden.<sup>44</sup>
- Mit Hilfe der multidimensionalen Skalierung können also auch die hervorspringenden Dimensionen identifiziert werden, die für die Positionierung der Autoren ausschlaggebend sind.

Abbildung 3 zeigt die resultierende Wissenschaftslandkarte zum Informationsmanagement.<sup>45</sup> Die Gruppierung der Autoren erfolgte auf Grundlage der durchgeführten Faktoren- und Clusteranalyse, die aus Platzgründen im Rahmen dieser Arbeit nicht dargestellt wurden.

<sup>43</sup> im Sinne der Kozitationsanalyse.

<sup>44</sup> im Sinne der Kozitationsanalyse.

<sup>45</sup> Für die Berechnungen wurde das ALSCAL-Modul von SPSS verwendet. Das Ergebnis genügt den Anforderungen von McCain: Stress 1 = 0,2; R Square = 0,85 (vgl. McCain 1990, S. 438).

Die Wissenschaftslandkarte zeigt sehr klar die beiden Abstufungen auf der horizontalen Achse, die die „Ingredienzen“ des Informationsmanagements ausmachen. Auf der linken Seite handelt es sich um informationsorientierte Aspekte des Informationsmanagements. Diese werden von den Informationswissenschaften wahrgenommen. Auf der rechten Seite geht es um Informationstechnologie, speziell dessen (effektiven und effizienten) Einsatz in Organisationen und die sich daraus ergebenden Möglichkeiten (Management Information Systems). Schwieriger ist die Interpretation der Ausprägungen auf der vertikalen Achse. Am oberen Pol befinden sich primär Autoren, die allgemeine (betriebswirtschaftliche) Grundlagen, primär für die MIS-Autoren, erarbeitet haben. Am unteren Ende wurden durch Horton, Marchand und Synnott hingegen umfassende Informationsmanagementkonzepte erarbeitet. Besonders interessant ist, dass alle Cluster um das Zentrum herum angeordnet sind, der Ursprung des Koordinatensystems hingegen völlig unbesetzt ist. Der Grund dürfte darin liegen, dass es zwischen dem informationswissenschaftlichen und dem MIS-Ansatz nur geringe „Berührungspunkte“ gibt. Am ehesten nehmen Horton, Marchand und Synnott diese Brückenfunktion wahr.

#### 4. Validität der Ergebnisse

Wie bereits erwähnt wurde, war die Anzahl der IM-ARTIKEL (und die Referenzen davon) wesentlich geringer als dies ursprünglich erwartet wurde. Da es sich bei der Autoren-Kozitationsanalyse zudem um ein stark filterndes Verfahren handelt, war es nicht möglich, detailliertere Erkenntnisse über die (formale) Wissenschaftskommunikation im Bereich des Informationsmanagements<sup>46</sup> zu gewinnen.

*Auffallend war die geringe Konzentration auch bei den zitierten Autoren. Dies könnte negative Auswirkungen auf die Autoren-Kozitationsanalyse insofern gehabt haben, weil die Grenze zur Auswahl der Autoren nicht scharf genug gezogen werden konnte. Es wäre also möglich, dass weitere wichtige Autoren bei der Autoren-Kozitationsanalyse nicht berücksichtigt wurden.* Der Hauptgrund für das relativ homogene Zitierverhalten dürfte in der interdisziplinären Natur des Informationsmanagements liegen. Dadurch "dominieren" (durch die Anzahl der erhaltenen Zitate) wenige Autoren nicht das gesamte Informationsmanagement, sondern jeweils nur einen Teilbereich (zum Beispiel den informationswissenschaftlichen IM-Ansatz) davon. Für das gesamte Informationsmanagement fällt dann die Konzentration geringer aus.

Auch die Diskrepanz zwischen dem Ranking der IM-ARTIKEL nach der Zeitschriftenfachgruppe (primär von den Informationswissenschaften dominiert) und den Ergebnissen der Autoren-Kozitationsanalyse (der Großteil der Autoren kommt aus der MIS- bzw. Managementliteratur) dürfte mit der Interdisziplinarität des Informationsmanagements zu tun haben. Um dies zu verifizieren, wurden die Referenzen der IM-ARTIKEL einer weiteren Auswertung unterzogen.

---

<sup>46</sup> Zum Beispiel unterschiedliche informationswissenschaftliche IM-Ansätze.

Rang	Zitate	Zeitschriftl Monographie <sup>47</sup>	Zeitschriftenfachgruppe <sup>48</sup>	Klasse <sup>49</sup>	Impact-Faktor <sup>50</sup>
1	56	HARVARD BUS REV	BUSINESS MANAGEMENT	MIS	2,5
2	46	INFORM MANAGE	COMPUTER SCIENCE, INFO SYSTEMS MANAGEMENT INFORMATION SCIENCE & LIBRARY SCIENCE (LIS)	MIS	0,7
3	39	MIS Q	MANAGEMENT LIS	MIS	1,6
4	29	COMMUN ACM	COMPUTER SCIENCE	Comp	1,3
5	26	INT J INFORM MANAGE	LIS	LIS	0,4
5	26	SLOAN MANAGE REV	MANAGEMENT BUSINESS	MIS	1,8
7	23	INFORMATION RESOURCE	- (Monographie)	-	-
7	23	INFORMATION SYSTEMS	- (Monographie)	-	-
9	19	MANAGEMENT INFORMATI	- (Monographie)	-	-
9	19	J AM SOC INFORM SCI	LIS	LIS	1,3
9	19	ACM T DATABASE SYST	COMPUTER SCIENCE, INFO SYSTEMS	Comp	0,4
12	16	INFORMATION TECHNOLO	- (Monographie)	-	-
12	16	ASLIB P	LIS COMPUTER SCIENCE, INFO SYSTEMS	LIS	0,2
12	16	J INFORM SCI	LIS COMPUTER SCIENCE, INFO SYSTEMS	LIS	0,4

**Tabelle 7:** Zeitschriften, die von IM-ARTIKELN mehr als 15 Mal zitiert wurden

Wie aus Tabelle 7 ersichtlich ist, wurden MIS-Zeitschriften von den IM-ARTIKELN wesentlich öfter zitiert als LIS- Zeitschriften. Folgende Gründe wären dafür denkbar:

<sup>47</sup> Es werden die in der Datenbank benutzten Kurzbezeichnungen verwendet.

<sup>48</sup> Die Zeitschriftenfachgruppe wurde den Journal Citation Reports® aus dem Jahre 1997 (CDROM-Version) entnommen. Dabei muss beachtet werden, dass die Zuordnung der Zeitschriftenfachgruppe bei Zeitschriften, die sowohl im SCI® als auch im SSCI vorkommen, nicht immer eindeutig ist.

<sup>49</sup> Die Klassifikation (MIS, LIS, Comp(uter Science)) wurde vom Autor aufgrund der persönlichen Einschätzung des Hauptschwerpunkts der jeweiligen Zeitschrift vorgenommen.

<sup>50</sup> Der Impact-Faktor wurde ebenfalls den Journal Citation Reports® entnommen.

1. LIS-Autoren zitieren MIS-Autoren wesentlich häufiger als dies umgekehrt der Fall ist.
2. Im Bereich von MIS werden IM-Publikationen oft nicht als solche tituliert. MIS-Publikationen im Bereich des Informationsmanagements sind daher in den IM-Artikeln unterrepräsentiert.
3. Das Zitierverhalten ist im Bereich von MIS stärker ausgeprägt als in den Informationswissenschaften.

Während das erste Argument nicht zutreffen dürfte, spricht einiges dafür, dass im Bereich von MIS eine Reihe von Publikationen erscheint, die zwar Teilaspekte des Informationsmanagements behandeln, aber nicht als solche tituliert werden. Das letzte Argument kann in den ISI<sup>®</sup>-Zitationsdatenbanken mit Hilfe des sogenannten "Impact Factor" untersucht werden. Dieser gibt an, wie oft ein Artikel einer bestimmten Zeitschrift im jeweils betrachteten Jahr durchschnittlich zitiert wurde.<sup>51</sup> Wie Tabelle 7 zu entnehmen ist, ist der Impact-Faktor in den MIS-Zeitschriften<sup>52</sup> deutlich höher als in den LIS-Zeitschriften<sup>53</sup>. Das Zitierverhalten ist in den beiden Disziplinen also sehr unterschiedlich. *Wie diese szientometrische Studie also zeigt, ist eine Zitationsanalyse, die sich über mehrere Disziplinen erstreckt, nicht unproblematisch, da das Zitierverhalten in unterschiedlichen Disziplinen, ja selbst Unterdisziplinen stark variieren kann.*

*Das Ranking nach den zitierten Werken<sup>54</sup> brachte auch deren schlechte Ansetzung zutage. In der Stichprobe aller Werke, die öfter als drei Mal zitiert wurden (109 Werke), waren elf Zeitschriften uneinheitlich angesetzt (zum Beispiel „MIS Q“, „MIS Q Mar“, MIS Q Sep“, „MIS Quart“). Für diese elf zitierten Werke gab es insgesamt 30 unterschiedliche Ansetzungen. Damit wurde schon in dieser sehr kleinen Stichprobe zirka ein Fünftel der zitierten Werke falsch angesetzt. Dies zeigt, dass die Qualität der Zitate sehr schlecht ist.*

---

<sup>51</sup> Vgl. Garfield 1994.

Der Impact-Faktor (IF) berechnet sich für eine bestimmte Zeitschrift (ZS) in einem bestimmten Jahr (i) nach folgender Formel:  $zslf_i = \frac{\text{Anzahl der im Jahr i für Artikel in ZS erhaltene Zitate, die in den Jahren i-1 und i-2 publiziert wurden}}{\text{DIVIDIERT DURCH die Gesamtzahl der in ZS in den Jahren i-1 und i-2 veröffentlichten Artikel}}$

Bei der Berechnung ist Folgendes zu beachten: Bei der Gesamtzahl der veröffentlichten Publikationen werden nur die Dokumenttypen „article“ und „review“ berücksichtigt. Hingegen werden die von allen Dokumenttypen erhaltenen Zitate gezählt. Zeitschriften mit einem zum Beispiel höheren Anteil an „letters“ werden dadurch bei der Berechnung des Impact Factor bevorzugt.

<sup>52</sup> Harvard Business Review: 2,5; Sloan Management Review: 1,8; MIS Quarterly: 1,6 ...

<sup>53</sup> Journal of the American Society for Information Science: 1,3; International Journal of Information Management: 0,4; Journal of Information Science: 0,4 ...

<sup>54</sup> Dabei kann es sich auch um Monographien handeln.

## **5. Resümee**

Wie gezeigt wurde, können bereits recht einfache Auswertungen, im Wesentlichen RANKING-Befehle nach unterschiedlichen Feldern, zu nützlichen Ergebnissen führen. So ergab die zeitliche Analyse ein sprunghaftes Ansteigen des Publikationsvolumens am Beginn der achtziger Jahre, das sich seither in einem bestimmten Rahmen bewegt. Man kann also davon ausgehen, dass es sich bei Informationsmanagement um kein Modethema handelt.

Weiter wurde gezeigt, dass Informationsmanagement in verschiedenen Anwendungsbereichen, hauptsächlich in Technik, Wirtschaft und öffentlicher Verwaltung, Medizin sowie im Erziehungswesen ein Thema ist. Mit Abstand am meisten IM-ARTIKEL werden in der Informations-Bibliothekswissenschaft, gefolgt von der Informatik, verfasst. Aufgrund der Ergebnisse der Autoren-Kozitationsanalyse sind MIS und Informationswissenschaft jene zwei Disziplinen, die sich primär mit Fragen des Informationsmanagements auseinander setzen.

Im Rahmen der durchgeführten szientometrischen Untersuchung trat aber auch eine Reihe von „Unzulänglichkeiten“ zutage: Zumindest für den Bereich des Informationsmanagements konnte die starke „Bevorzugung“ des anglo-amerikanischen Raumes aufgezeigt werden. Dies lässt sich sowohl aus der Herkunft der publizierenden Autoren (über 70 % aus den USA oder Großbritannien) als auch dem Erscheinungsort der meisten IM-Zeitschriften schließen. Alle hier vorliegenden Ergebnisse gelten daher primär nur für den anglo-amerikanischen Raum.

Als ein Vorteil von szientometrischen Untersuchungen wird ihre Objektivität ins Treffen geführt. Dies ist auf der Ebene des einzelnen Wissenschaftlers zutreffend, obwohl auch hier eine subjektive Einflussnahme nicht ausgeschlossen werden kann.<sup>55</sup> Eine objektive Darstellung des gesamten Systems der Wissenschaft bzw. eines Teilbereichs davon ist aber aus obigen Gründen ebenfalls nicht gegeben. Bei der Autoren-Kozitationsanalyse zeigte sich, dass ihre Durchführbarkeit wesentlich von den zugrunde liegenden Daten abhängt. Da es sich um ein hoch aggregierendes Verfahren handelt, müssen gewisse „Datenmengen“ vorliegen, um zu sinnvollen Ergebnissen zu kommen. Weiter muss bei einer „interdisziplinären“ Autoren-Kozitationsanalyse berücksichtigt werden, dass unterschiedliche Disziplinen ein unterschiedliches Zitierverhalten aufweisen und die Ergebnisse dementsprechend verfälscht werden können.

Trotz dieser Problembereiche können szientometrische Analysen von Nutzen sein, sofern diese bei deren Durchführung und der Interpretation der Ergebnisse berücksichtigt werden. Einer entsprechenden Informationskompetenz kommt also bei szientometrischen Untersuchungen eine zentrale Bedeutung zu.

## **Literaturverzeichnis**

Broadus R. N. (1983): An Investigation of the Validity of Bibliographic Citations. In: Journal of the American Society for Information Science, 34(1983)2, S. 132 - 135.

Cronin B. (1984): *The Citation Process: The Role and Significance of Citations in Scientific Communication*. Taylor Graham, London.

Daniel H.-D. (1988): Methodische Probleme institutsvergleichender Analysen der Forschungsproduktivität. In: Daniel/Fisch 1988, S. 215 – 241.

Daniel H.-D., Fisch R. (Hg.) (1988): *Evaluation von Forschung: Methoden – Ergebnisse – Stellungnahmen*. Universitätsverlag Konstanz, Konstanz.

Fröhlich G. (1999): Das Messen des leicht Meßbaren. Output-Indikatoren, Impact-Maße: Artefakte der Szientometrie? In: Becker J., Göhring W. (Hg.): *Kommunikation statt Markt. Zu einer alternativen Theorie der Informationsgesellschaft*. Forschungszentrum Informationstechnik GmbH, St. Augustin, S. 27 - 38.

Garfield E. (1979): Is Citation Analysis a Legitimate Evaluation Tool? In: *Scientometrics*, 1(1979)4, S. 359 - 375.

Garfield E. (1983): *Citation Indexing — Its Theory and Application in Science, Technology, and Humanities*. Nachdruck. ISI' Press, Philadelphia.

Garfield E. (1989): Citation Classics and Citation Behavior Revisited. In: *Current Contents*, 12(1979)4, S. 3 - B.

Garfield E. (1994): The Impact Factor. In: *Current Contents*, 20. Juni 1994. Elektronische Version unter: <http://www.isinet.com/hot/essays/7.html> (Stand: September 1999).

Garfield E. (1995): Monitoring Complex Literature: The Advantages of Using a Multidisciplinary Database. <http://www.isinet.com/hot/essays/18.htm1> vom 27. März 1995 (Stand: September 1999).

Gorraiz J. (1992): Die unerträgliche Bedeutung der Zitate. In: *Biblos* 41 (1992)4, S. 193 — 204.

Herget J. (1997): Informationsmanagement. In: Buder M., Rehfeld W., Seeger T., Strauch D. (Hg.) (1997): *Grundlagen der praktischen Information und Dokumentation*. 4. völlig neu gefaßte Auflage, Band 1. Saur, München et al., S. 781 — 794.

Lawani S. M. (1981): Bibliometrics: Its Theoretical Foundations, Methods and Applications. In: *Libri*, 31(1981)4, S. 294 — 315.

MacRoberts M. H., MacRoberts B. R. (1989): Problems of Citation Analysis: A Critical Review. In: *Journal of the American Society for Information Science*, 40(1989)5, S. 342 — 349.

McCain K. (1990): Mapping Authors in Intellectual Space: A Technical Overview. In: Journal of the American Society for Information Science, 41(1990)6, S. 433 — 443.

Merton R. K. (1973): The Sociology of Science: Theoretical and Empirical Investigations. The University of Chicago Press, Chicago.

Schmidmaier D. (1984): Zur Geschichte der Bibliometrie. In: Zentralblatt für Bibliothekswesen, 98(1984), S. 404 — 406.

Smith L. C. (1981): Citation Analysis. In: Library Trends, Summer 1981, S. 83 - 105.

Stock W. (1985): Die Bedeutung der Zitatenanalyse für die Wissenschaftsforschung. In: Zeitschrift für allgemeine Wissenschaftstheorie, 16(1985) 2, S. 304 — 314.

Vinkler P. (1987) A Quasi-Quantitative Citation Model. In: Scientometrics, 12(1987)1 - 2, S. 47 - 72.

Van Raan A.F.J. (1997): Scientometrics: State-of-the-Art. In: Scientometrics, 38(1997)1, S. 205-218.

Weinstock M. (1971): Citation Indexes. In: Encyclopaedia of Library and Information Science. Band 5, New York, S. 16 - 40.

White H. D., Griffith B. C. (1981): Author Cocitation: A Literature Measure of Intellectual Structure. In: Journal of the American Society for Information Science, 32(1981)5, S. 163 — 171.

White H. D. (1990): Author Co-Citation Analysis: Overview and Defense. In: Borgman C. L. (Hg.) (1990): Scholarly Communication and Bibliometrics. Sage, Newbury Park, USA, S. 84 - 106.





In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 113 – 125

# **Integrativer Ansatz zur Wissensbewertung**

**Otto Krickl, Elisabeth Milchrahm**

Institut für Organisations- und Personalmanagement  
Karl-Franzens-Universität Graz  
Universitätsstraße 15 / E4  
A-8010 Graz  
email: otto.krickl@kfunigraz.ac.at

Institut für Informationswissenschaft  
Karl-Franzens-Universität Graz  
Universitätsstraße 15 / F3  
A-8010 Graz  
email: elisabeth.milchrahm@kfunigraz.ac.at

## **Zusammenfassung**

Im Gegensatz zu den traditionellen Rechnungslegungssystemen gibt es im Rahmen des Wissensmanagements keine erprobten Messverfahren auf die man zur Messung und Bewertung organisationalen Wissens zurückgreifen kann. Wissen kann selten auf eine Messdimension zurückgeführt werden und häufig ist der Messaufwand sehr hoch, jedoch liegt im Bereich der Wissensbewertung ein bisher stark vernachlässigtes Potential. Das Ziel der vorliegenden Arbeit liegt in der kritischen Analyse eines Ansatzes zur Wissensbewertung. Im Zuge dieser Untersuchung wird die Methodik der Balanced Scorecard zur Eignung der Wissensbewertung hinterfragt, indem auch die betriebliche Praxis anhand eines konkreten Fallbeispiels untersucht wird. Abschließend wird ein mehrstufiges Indikatorenmodell zur Wissensbewertung vorgestellt.

## **Abstract**

Currently, within business, intangible assets are growing in significance in relation to traditional tangible assets. There is a need of alternative approaches in order to measure the value of knowledge. This paper looks at the problem of knowledge measurement and argues for the importance of metrics to the

overall process of knowledge management. The concept of Balanced Scorecard is critically analyzed in respect of its approaches to broaden the range of metrics for organisational performance. Finally, a framework for measuring knowledge is developed.



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

## 1. Problemstellung

### 1.1 Messung und Bewertung von Wissen

Zur Unternehmensbewertung gewinnt die Entwicklung qualitativer Indikatoren zunehmend an Bedeutung. Qualitative Indikatoren umfassen u.a. auch den Wert des Wissens<sup>1</sup> der Unternehmung. Indikatoren zur Messung und Bewertung von Wissen zielen daher darauf ab, Veränderungen der organisationalen Wissensbasis<sup>2</sup> aufzuzeigen (Wissensmessung, d.h. welches Wissen ist in welcher „Menge“ vorhanden) und in Bezug auf definierte Wissensziele zu interpretieren (Wissensbewertung, d.h. was ist der „Wert“ des Wissens im Kontext des Unternehmens). Eine systematische Entwicklung von wert-steigerndem Wissen im Unternehmen erfordert eine entsprechende Abstimmung der zur Wissensmessung ausgewählten Kennzahlen mit den jeweiligen Wissenszielen.<sup>3</sup> Andernfalls sind die im Zuge der Wissensbewertung getroffenen Interpretationen nicht schlüssig.

### 1.2 Festlegung von Wissenszielen

Im Rahmen des Wissensmanagements<sup>4</sup> stellen Wissensziele eine Sonderkategorie von klassischen Unternehmenszielen dar: sie legen fest, wie die organisationale Wissensbasis weiterentwickelt werden soll. Wissensziele sind somit als bewusste Ergänzung herkömmlicher Planungsaktivitäten zu sehen. Sie übersetzen herkömmliche normative, strategische und operative Unternehmensziele in eine wissensorientierte Sprache. Probst<sup>5</sup> und Romhardt<sup>6</sup> beschreiben Wissensziele auf diesen drei Ebenen wie folgt: Mit der Festlegung normativer Wissensziele soll der Bedeutung der Ressource Wissen für den Unternehmenserfolg durch die Schaffung einer entsprechenden Unternehmenskultur Rechnung getragen werden. Strategische Wissensziele werden für jene langfristigen Aktivitäten formuliert, die den Aufbau von organisationsrelevantem Wissen unterstützen sollen. Operative Wissensziele schließlich sollen die systematische Steuerung und Kontrolle des Wissensaspektes unternehmensweit ermöglichen.

### 1.3 Begriff „Wissenskapital“

Verfahren zur Messung von Wissen basieren auf der Annahme, dass sich der Marktwert eines Unternehmens aus seinem finanziellen Wert und einem bisher nicht näher bestimmbareren Wert zusammensetzt, der als „Wissenskapital“ oder „Intellectual Capital“ bezeichnet wird. Im Rahmen der unternehmerischen Wertgenerierung ist es sinnvoll, das Modell des Wissenskapitals differenziert zu

---

<sup>1</sup> Kuhlen (1989, S 13) definiert Wissen als die Summe aller bisherigen individuellen oder kollektiven Erfahrungen oder Erkenntnisse, die in gewissem Ausmaß verallgemeinerbar sind.

<sup>2</sup> Siehe dazu Kapitel 3

<sup>3</sup> Zur Problematik der Messgrößenauswahl siehe Horvath/Gaiser (2000, S 27f.). Einen

<sup>4</sup> Überblick über das Konzept des Wissensmanagements geben zum Beispiel Nonaka/Takeuchi (1997), Krcmar/Rehäuser (1996).

<sup>5</sup> Vgl. Probst/Raub/Romhardt (1998), S. 68 - 90.

<sup>6</sup> Vgl. Romhardt (1998), S. 72.

betrachten und eine Unterscheidung zwischen Humankapital und Strukturkapital zu treffen. Dieser Differenzierung liegt die Überlegung zugrunde, dass das Unternehmen über keine Eigentumsrechte am Humankapital (z.B. Kompetenz der Mitarbeiter) verfügt, das Strukturkapital (z.B. Infrastruktur) sich jedoch in größerem Umfang in den Händen des Unternehmens befindet.<sup>7</sup>

Die unterschiedlichen Begriffsabgrenzungen und Zielsetzungen im Rahmen der Wissensmessung und -bewertung führen zu entsprechend divergierenden Ansätzen, die im folgenden Kapitel im Überblick dargestellt werden.

## **2. Ansätze zur Messung und Bewertung organisationalen Wissens**

Die einfachste Messgröße zur Feststellung immateriellen Vermögens beschränkt sich auf die Bezifferung des Unterschiedes zwischen Marktwert (Börsenwert) und Buchwert (entnommen aus der Bilanz) eines Unternehmens. Diese auf Marktwert-Buchwert-Relationen basierende Verfahren werden auch als „deduktiv-summarische Ansätze“<sup>8</sup> bezeichnet. Zur operativen und strategischen Steuerung eines Unternehmens unter dem Gesichtspunkt der Wissensbewertung sind diese Ansätze jedoch nicht geeignet.

„Induktiv analytische Ansätze“<sup>9</sup> hingegen haben als Zielsetzung, finanzielle und nicht-finanzielle Indikatoren zu entwickeln, die eine gezielte Steuerung der organisationalen Wissensbasis ermöglichen. Zu diesen Ansätzen zählen die Balanced Scorecard (BSC) und das Indikatorenmodell „Skandia Navigator“, die in den folgenden Abschnitten daraufhin untersucht werden, welchen Beitrag sie tatsächlich zur unternehmensweiten Wissensmessung und -bewertung leisten.

### **2.1 Balanced Scorecard**

Die von Kaplan und Norton<sup>10</sup> entwickelte Balanced Scorecard ist ein Instrumentarium zur Steigerung der Aussagefähigkeit traditioneller finanziell orientierter Kennzahlen. Finanzielle Kennzahlen werden dabei durch treibende Faktoren für zukünftige Wettbewerbsfähigkeit ergänzt.

Das aus dieser Aktionsforschung abgeleitete Modell betrachtet ein Unternehmen aus vier Perspektiven: der Kundenperspektive, der finanziellen Perspektive, der Perspektive der internen Geschäftsprozesse und der Lern- und Entwicklungsperspektive.

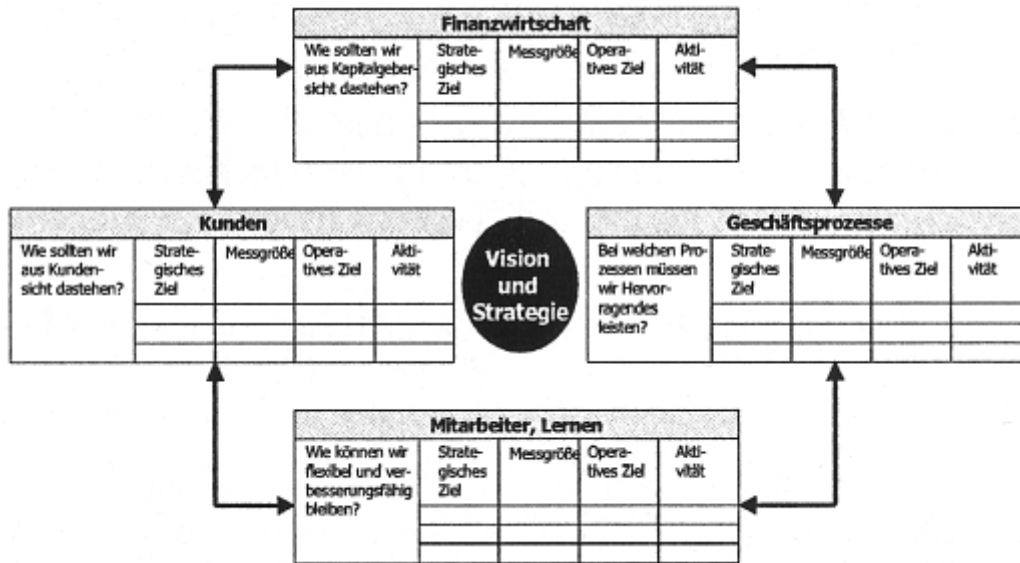
---

<sup>7</sup> Eine detaillierte Beschreibung der Komponenten von Humankapital und Strukturkapital gibt Reinhardt (1998), S. 153 — 158.

<sup>8</sup> Weitere Beispiele dafür sind „Tobin's Q“ und „CIV“ (Calculated Intangible Value).

<sup>9</sup> Vgl. North/Probst/Romhardt (1998) S. 161ff.

<sup>10</sup> Vgl. Kaplan/Norton (1997), S. 7ff



**Abbildung 1:** Grundmodell der Balanced Scorecard<sup>11</sup>

Die unterschiedlichen Perspektiven des Unternehmens werden nicht nur nebeneinander gestellt, sondern auch in ihren Ursache-Wirkungs-Zusammenhängen dargestellt. Für jede der Perspektiven werden strategische Ziele, Messgrößen, operative Ziele und Maßnahmen (Initiativen) festgelegt (siehe Abb. 1). D.h. aus der Unternehmensstrategie werden jene Prozesse abgeleitet, die wesentlich zum Erfolg der Unternehmung beitragen. Die Prozessmodelle liefern wiederum die Anforderungen an die Qualifikation der Mitarbeiter. Die Verbesserungspotentiale der Prozesse bilden den Ausgangspunkt für die Infrastruktur, die die Organisation schaffen muss um langfristig die Wettbewerbsfähigkeit zu sichern. Die Lern- und Entwicklungsperspektive wird daher nicht nur systematisch analysiert, sondern auch mit den Wechselwirkungen zu den anderen Perspektiven in Zusammenhang gebracht.

Die Lern- und Entwicklungsperspektive entwickelt Ziele und Kennzahlen zur Förderung einer wissensorientierten Organisation. Damit werden die treibenden Faktoren für hervorragende Ergebnisse in den anderen drei Perspektiven identifiziert und dem Management auch ein Instrumentarium für die Beurteilung von Investitionen in Aktivitäten des Wissensmanagements geliefert. Beispiele für Kennzahlen dieser Perspektive sind die strategische Aufgabendeckungsziffer<sup>12</sup>, Relation umgesetzter zu den vorgeschlagenen Verbesserungsideen und Half-life-Kennzahl.<sup>13</sup>

Eine konkrete Operationalisierung der Wissensperspektive mit entsprechenden Wissensindikatoren ist innerhalb des BSC-Konzeptes nicht zu finden. Zentral

<sup>11</sup> Zeichnung der Verfasser nach Vorlage von: Kaplan/Norton (1997), S. 9

<sup>12</sup> Siehe dazu Kapitel 3.

<sup>13</sup> Die Half-life-Kennzahl misst den Zeitraum, innerhalb dessen eine Prozessleistung (Durchlaufzeit, Kosten, etc.) um 50% verbessert werden kann. Vgl. Schneidermann (1988), S 51ff.

erscheint die Erkenntnis, dass sich jede Organisation ihr eigenes, maßgeschneidertes und kontextspezifisches Indikatorenset erarbeiten muss. Nur so kann sie die für sie relevanten Dimensionen erfassen und steuern. Gleichzeitig unterstützt die BSC die enge Verbindung von Wissenszielen und Wissensbewertung, was idealer Weise zu schnellen Feedback-Prozessen führen wird<sup>14</sup>. Der Versuch einer Umsetzung ist der „Skandia Navigator“.

## **2.2 Indikatorenmodell (Skandia)**

Ein Indikatorensystem, das darauf abzielt, Wissenskapital zu messen und zu bewerten, wurde von dem schwedischen Finanzdienstleistungskonzern Skandia entwickelt und erstmals 1994 als Anlage<sup>15</sup> zum traditionellen Geschäftsbericht des Tochterunternehmens „Skandia AFS“ (Assurance and Financial Services) herausgegeben. Skandia hat damit als eines der ersten Unternehmen Indikatoren zur Quantifizierung von Wissenskapital entwickelt, die der externen Berichtlegung dienen.

Zielsetzung des im Zuge einer Umstrukturierung eingeführten Funktionsbereiches „Intellectual Capital“<sup>16</sup> ist die Erfassung, der Aufbau und die Nutzbarmachung von Humankapital<sup>17</sup> und Strukturkapital<sup>18</sup>. Zur Messung des Intellectual Capital wurde ein Indikatorenset („Skandia Navigator“) entwickelt, welches in den folgenden Kapiteln anhand der in den Jahren 1994 bis 1998 veröffentlichten Geschäftsberichte (inklusive Anhang) auf seine Eignung zur Wissensmessung und -bewertung überprüft wird.

### **2.2.1 Indikatorenentwicklung**

Die Entwicklung von Indikatoren zur Messung von Wissen setzt die Identifikation jener Erfolgsfaktoren voraus, die erfüllt werden müssen, um das Geschäft erfolgreich abwickeln zu können. In einem nächsten Schritt wird herausgearbeitet, mit welchen Wissenskapitalindikatoren sich diese kritischen Erfolgsfaktoren am besten messen und bewerten lassen.

---

<sup>14</sup> Vgl. Romhardt (1998) S. 302.

<sup>15</sup> Die Veröffentlichung erfolgt halbjährlich unter dem Titel „Intellectual Capital – Supplement to Skandia 's interim report“.

<sup>16</sup> Skandia definiert „Intellectual Capital“ wie folgt: „IC arises in a value-creating process based on the interaction of Human and Structural Capital, where continuous renewal, innovations transform and refine the individual's knowledge into enduring value for Skandia“ (<http://www.skandia.com/capital/valuescheme.htm>).

<sup>17</sup> Humankapital wird von Skandia folgendermaßen definiert: „The accumulated value of investments in employee training, competence and future. Might also be described as the employee's competence, relationship ability and values.“ ([www.skandia.com/flashnot/index\\_upgrade\\_flash.htm](http://www.skandia.com/flashnot/index_upgrade_flash.htm)).

<sup>18</sup> Die (plakative) Beschreibung von Strukturkapital lautet: „That is everything that remains when the staff goes home, i.e. databases, manuals etc.“ ([http://www.skandia.com/flashnot/index\\_upgrade\\_flash.htm](http://www.skandia.com/flashnot/index_upgrade_flash.htm)).

Anhand dieser Vorgangsweise entwickelte Skandia - basierend auf der Balanced Scorecard - die Indikatorenklassen Kundenfokus („Customer Focus“)<sup>19</sup>, Mitarbeiterfokus („Human Focus“)<sup>20</sup>, Prozessfokus („Process Focus“)<sup>21</sup>, Finanzieller Fokus („Financial Focus“)<sup>22</sup> und den für den Konzern besonders wichtigen Erneuerungs- und Entwicklungsfokus („Renewal & Development Focus“). Mit diesem Indikator soll die Entwicklung neuer Produkte und Dienstleistungen abgebildet und die Umwandlung von Humankapital zu Strukturkapital dargestellt werden.<sup>23</sup>

Für das Versicherungsunternehmen DIAL - ein Tochterunternehmen des Skandia-Konzerns - setzt sich der Erneuerungs- und Entwicklungsfokus folgendermaßen zusammen:

- „Increase in gross premiums written (%)“,
- „Share of direct payments in claims assessment system (%)“
- „Number of ideas filed with Idea Group“.<sup>24</sup>

Die Aussagekraft dieser Indikatoren ist jedoch aus folgenden Gründen kritisch zu sehen:

- Den ausgewählten Kennzahlen fehlt der klare Bezug auf definierte Wissensziele: Indikatoren wie zum Beispiel „Anstieg im Prämienvolumen“ oder „Anteil der Zahlungen in Schadensbewertungsverfahren“ sind aus der Perspektive der Wissensmessung bzw. -bewertung schwer interpretierbar.
- Die Kennzahl „Anzahl der von der „Idea Group“ registrierten Ideen“ sagt nichts über realisierte bzw. erfolgreich umgesetzte Ideen aus.<sup>25</sup>

Anhand der diskutierten Beispiele wird deutlich, dass die bisher entwickelten Kennzahlen teilweise schwer nachvollziehbar und zu wenig ausgereift sind, um Hinweise auf eine gezielte Entwicklung der organisationalen Wissensbasis geben zu können. Das angestrebte Ziel, durch Vermittlung von „wissensbasierten“

---

<sup>19</sup> Diesem Indikator werden u.a. folgende Kennzahlen zugeordnet: „telephone accessibility (%)“, „number of individual policies“.

<sup>20</sup> Entsprechende Kennzahlen sind u.a.: „average age“, „ number of employees“

<sup>21</sup> Eine Kennzahl dieses Indikators ist zum Beispiel „IT-employees/total number of employees (%)“.

<sup>22</sup> Der finanziellen Perspektive zugeordnet wird zum Beispiel die Kennzahl „gross premiums written/employee“.

<sup>23</sup> Siehe dazu Kapitel 2.2.2.

<sup>24</sup> Die Entwicklung dieser Kennzahlen im Jahr 1997 (1996) zeigt sich wie folgt: increase in gross premiums written (%): 7,9 (6,3); share of direct payments in claims assessment system (%): 21,5 (20,5); number of ideas filed with Idea Group: 595 (175). Vgl. Intellectual Capital Report (1998), S. 21.

<sup>25</sup> Zum hier angesprochenen Themenbereich Wissensentstehung und -verteilung im Rahmen der Produktentwicklung siehe Nonaka/Takeuchi (1997), S. 74 - 87.

Kennzahlen der interessierten Öffentlichkeit Informationen zur Verfügung zu stellen, die zur Messung der in den abgelaufenen Perioden erbrachten Managementleistung und zur Einschätzung zukünftiger Entwicklungen geeignet sind, wird mit den vorliegenden Indikatoren nur ansatzweise erreicht.

### **2.2.2 Transformation von Humankapital in Strukturkapital**

Auf Basis des „Human Ressource Accounting“<sup>26</sup> versucht Skandia weitere Indikatoren zu entwickeln, die strategische Investitionen in Humankapital und dessen effiziente Nutzung und Entwicklung hervorheben. Der Ansatz von Skandia stützt sich nun darauf, zu dokumentieren, in welchem Ausmaß innerhalb einer Periode Humankapital in Strukturkapital umgewandelt wird. Darüber hinaus soll aus den errechneten Werten („Human capital relationship value per employee“ und „Structural capital in relation to value added“) die zukünftige Entwicklung von Humankapital abgeleitet werden können. Grundsätzliche Bewertungsprobleme hinsichtlich Qualität und Konstanz von Humankapital führen jedoch dazu, dass bereits bei der Berechnung eines Basiswertes („Human capital base value“) aufgrund der Überbetonung des Kostenaspektes<sup>27</sup> Schwachstellen auftreten.

Die Quantifizierung von Humankapital und die Visualisierung des Umwandlungsprozesses von Humankapital in Strukturkapital erfordern ausgereifte Indikatoren, die Entwicklungspotentiale von Humankapital aufzeigen können. Zum jetzigen Zeitpunkt liegen solche Indikatoren noch nicht vor.

### **2.2.3 Bildung von Wissenskapitalindizes**

Die steigende Bedeutung qualitativer Indikatoren zur Bewertung von Unternehmen erfordert eine entsprechende Dokumentation des unternehmensweiten Wissenskapitals. Dabei soll die Veröffentlichung von Wissenskapitalindizes als Zusatzinformation für interne und externe Bezugsgruppen (z.B.: Aktionäre) dienen. Firmenspezifische Indizes verhindern aber den Vergleich von Unternehmen untereinander.

Die zu Indizes zusammengefassten Indikatoren werden in Abhängigkeit ihrer Relevanz gewichtet und ihre Entwicklung im zeitlichen Ablauf dargestellt. Zudem versucht das Unternehmen im Rahmen von Pilotprojekten die Indikatoren zu einem aussagekräftigen Index (IC-Index<sup>TM</sup>) zu verdichten. Dieser Index muss einen quantifizierbaren Zusammenhang zwischen ökonomischen und wissenskapitalbezogenen Messgrößen herstellen können.<sup>28</sup> Der ICIndex<sup>TM</sup> spiegelt jedoch weder einen absoluten Wert des Wissenskapitals wider, noch gibt er Auskunft über den realen Wertzuwachs. Ziel der Indexbildung ist, einen Überblick über das unternehmenseigene Wissenskapital zu schaffen und entscheidungsunterstützende Informationen zu geben.

---

<sup>26</sup> Gutschelhofer (1999, S. 44 - 56) gibt einen Überblick über die Entwicklung des Human Resource Accounting.

<sup>27</sup> Vgl. dazu Siegert (1999), S. 18 - 22.

<sup>28</sup> Vgl. Reinhardt (1998), S. 164.

## 2.2.4 Beurteilung des Ansatzes

Das von Skandia entwickelte Indikatorenmodell hat zum Ziel, das unternehmensweite Wissenskapital zu erfassen und zu quantifizieren, um darauf aufbauend entscheidungsrelevante Informationen vermitteln zu können. Wie aus den vorangegangenen Kapiteln ersichtlich, sind die Indikatoren zur Wissensmessung und -bewertung nur eingeschränkt brauchbar. Die Interpretation der Indikatoren gestaltet sich teilweise schwierig, und ihr Stellenwert zur Erreichung unternehmensweiter Wissensziele wird nicht immer deutlich. Zudem muss bei der Bewertung von Wissenskapital berücksichtigt werden, dass Investitionen in Humankapital nicht sofort, sondern erst mit zeitlicher Verzögerung zu greifen beginnen.

Aufgrund einer Analyse der in den Geschäftsberichten<sup>29</sup> der Jahre 1994 bis 1998 veröffentlichten Indikatoren kann weiters festgestellt werden, dass das Ziel, Wissenskapital im Unternehmen durch Indexbildung abzubilden, nur ansatzweise erreicht wird. Trotz der angeführten Kritikpunkte haben der Aufbau eines Indikatorensets auf Basis der BSC und die Bildung von Indizes durchaus Signalwirkung: sie versuchen festzustellen, woher jener Teil des Börsenwertes stammt, der sich nicht durch traditionelle Finanzindikatoren (Bilanz, Gewinn- und Verlustrechnung) erklären lässt.

## 3. Messung und Bewertung organisationaler Wissensbasen

Organisationale Wissensbasen setzen sich aus individuellen und kollektiven Wissensbeständen zusammen. Zu ergänzen sind diese noch um die Daten- und Informationsbestände, auf die individuelles und kollektives Wissen aufbauen. Welche Wissensbestände in eine Bewertung einzubeziehen sind, ist jedoch nicht klar abzugrenzen. Ebenso kann der Zusammenhang zwischen der Wissensbasis und dem Geschäftserfolg nur indirekt hergestellt werden. Die Ansätze zur Wissensbewertung weisen daher sehr große Unterschiede auf.<sup>30</sup>

Aufbauend auf der BSC und dem vierstufigen Indikatorensystem nach North, Probst und Romhardt wird in der vorliegenden Arbeit ein verfeinertes Messsystem für die Indikatoren vorgestellt. Dabei wird die Grundstruktur der Indikatorenklassen und deren Begriffsbestimmung übernommen (siehe Tab. 1).

Neue Ansätze werden bei der Messung der Indikatoren beschritten:

Indikatorenklasse	Begriffsbestimmung	Analysebereich
Organisationale Wissensbasis (1)	Bestand des organisationalen Wissens zum Zeitpunkt T	Qualifikationserfordernisse aus den Geschäftsprozessen, strategische Aufgabendeckungsziffer

<sup>29</sup> Siehe "Intellectual Capital - Supplement to Skandia's Interim Report".

<sup>30</sup> Vgl. North/Probst/Romhardt (1998), S. 159 f.



Interventionen (II)	Beschreibt den gesamten Prozess des Wissensmanagements	Vorhandensein und Qualität der Teilprozesse des Wissensmanagements
Zwischenerfolge und Übertragungseffekte (III)	Misst die Ergebnisse der Geschäftsprozesse anhand deren Evaluierungskriterien	Erfolg der Geschäftsprozesse
Ergebnisse der Geschäftstätigkeit (IV)	Misst alle Faktoren der finanziellen Perspektive der BSC	Erfolg der gesamten Organisation und der Verknüpfung mit den Strategien

**Tabelle 2:** Indikatoren zur Wissensmessung und -bewertung<sup>31</sup>

Grundsätzlich wird bei der Messung der Indikatoren auf die jeweiligen Geschäftsprozesse<sup>32</sup> abgestellt. In der Indikatorenklasse I werden beispielsweise die Kerngeschäftsprozesse der Organisation herangezogen. Dabei bildet die Prozessdefinition als solche einen wesentlichen Bestandteil der Wissensbasis.<sup>33</sup> Aus den Evaluierungskriterien der einzelnen Prozesse werden die notwendigen Schlüsselqualifikationen der Mitarbeiter abgeleitet (dies ist für eine vollständig definierte BSC ohnehin notwendig). Der Indikator ergibt sich dann aus dem Soll-Profil im Vergleich zu den Ist-Qualifikationen.

Als Beispiel für Indikatoren der Klasse I können selektive Logistikkennzahlen für die Auslieferung von Produkten an die Kunden herangezogen werden. Bei E-Business<sup>34</sup> stellt die Lieferzeit oft einen wesentlichen Erfolgsfaktor da, sodass dieser besondere Aufmerksamkeit geschenkt wird. Die sich daraus ergebenden Schlüsselqualifikationen der Mitarbeiter sind u.a. Wissen über die Analyse und Beseitigung von kritischen Engpässen des Logistiksystems. Dazu gehört beispielsweise Wissen über alternative Logistikwege und deren Eignungsbereich für den jeweiligen Anwendungsfall.

Ein neuer Weg ist in der Indikatorenklasse II zu beschreiten. Hier dient als Analyseobjekt der Prozess des Wissensmanagements. D.h. es wird nicht nur auf Schulungstage und deren Qualität (z.B. Ratings der Schulungsqualität) abgestellt,

<sup>31</sup> Modifiziert nach North/Probst/Romhardt (1998), S. 165.

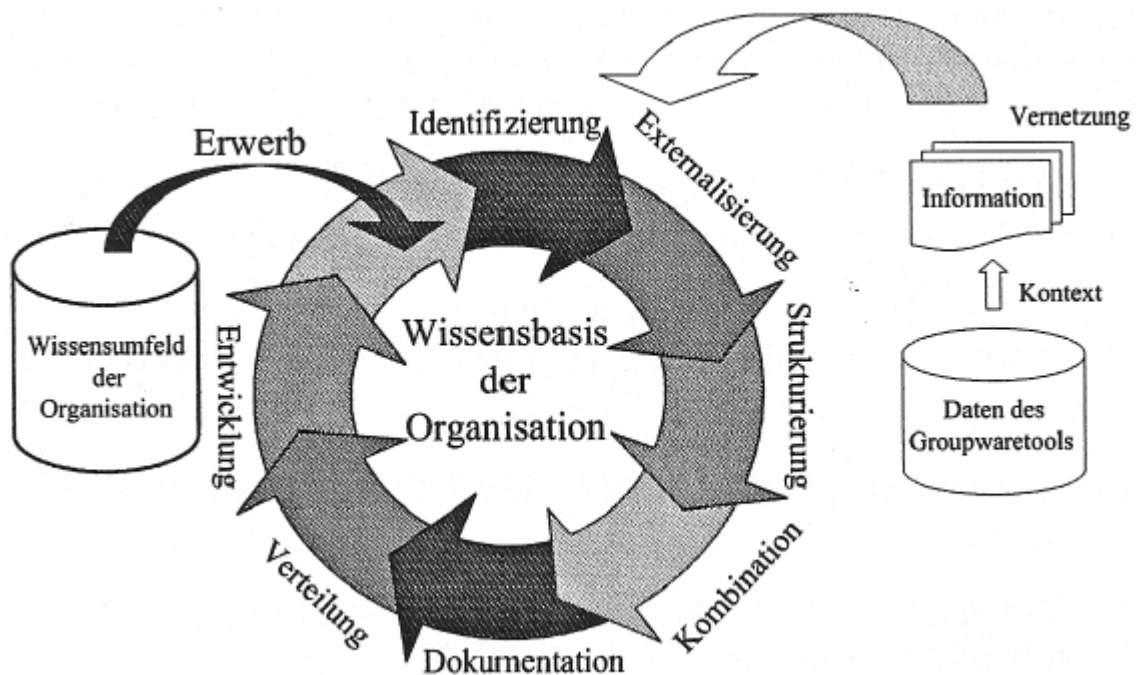
<sup>32</sup> „Ein Geschäftsprozess bezeichnet eine Abfolge von Tätigkeiten, Aktivitäten und Verrichtungen zur Schaffung von Produkten oder Dienstleistungen, die in einem direkten Beziehungszusammenhang miteinander stehen, und die in ihrer Summe den betriebswirtschaftlichen, produktionstechnischen, verwaltungstechnischen und finanziellen Erfolg des Unternehmens bestimmen.“ Striening (1988), S. 57.

<sup>33</sup> Die Struktur und der Umfang der Prozessdefinitionen bzw. die Prozessevaluierung orientiert sich dabei an der ISO 9000:2000.

<sup>34</sup> Unter „E-Business“ ist die Integration der elektronischen Transaktionen (üblicherweise als E-Commerce bezeichnet) in die Geschäftsabläufe der beteiligten Parteien zu verstehen. E-Business bezeichnet demnach die strategische Einbeziehung der elektronischen Transaktionen in alle Aspekte der Geschäftstätigkeit (Strategie etc.).

sondern es wird versucht, den gesamten Prozess von der Identifikation bis zur Entwicklung des Wissens abzubilden (siehe Abb. 2). Damit wird nicht nur die Existenz des für die Abwicklung der Geschäftsprozesse notwendigen Wissens, sondern auch die Qualität der Prozesse zu dessen Identifikation, Externalisierung, Strukturierung etc. beurteilt.

In der Indikatorenklasse III wird prinzipiell dem vorhandenen Ansatz gefolgt, aber konsequent auf die Evaluierungskriterien der Geschäftsprozesse abgestellt. Damit wird erreicht, dass sämtliche Indikatoren relevant für den Erfolg der Organisation sind und auch die laufende Erhebung und Dokumentation sichergestellt ist.



**Abbildung 2:** Phasenmodell des Wissensmanagements<sup>35</sup>

Die Indikatoren der Klasse IV umfassen sämtliche Aspekte der finanziellen Perspektive der Balanced Scorecard und stellen daher nicht nur auf die üblichen finanziellen Kennzahlen ab.

## 4. Resümee

„Um den Erfolg des Wissensmanagements messbar zu machen, ist das Unmögliche nötig: Die kontextgebundene Ressource Wissen muss objektivierbar gemessen werden.“<sup>36</sup> Im vorliegenden Beitrag konnte natürlich nichts Unmögliches realisiert, aber dennoch gezeigt werden, dass über die Verknüpfung der Balanced Scorecard mit verfeinerten Ansatzpunkten und Messverfahren für ein mehrstufiges Indikatorenmodell, eine näherungsweise Messung von Wissen möglich ist. Da dieses Messverfahren auch den direkten Konnex zur

<sup>35</sup> Vgl. Krickl /Milchrahm (1998), S. 93.

<sup>36</sup> Vgl. North/Probst/Romhardt (1998), S. 158.

Unternehmensstrategie herstellt, kann von einem integrativen Ansatz gesprochen werden, der auch in einer folgenden Pilotstudie näher untersucht wird.

## **Literatur**

Gutschelhofer A. (1999): Koordinierendes Personal-Controlling - Entwicklungslinien und Barrieren. Wiesbaden: Deutscher Universitäts-Verlag.

Horvath P. (1998): Wissensmanagement mit Balanced Scorecard. In: Bürgel H.D. (Hg.): Wissensmanagement: Schritte zum intelligenten Unternehmen. Berlin: Springer-Verlag, S. 153 – 162.

Horvath P./Gaiser B. (2000): Implementierungserfahrungen mit der Balanced Scorecard im deutschen Sprachraum - Anstöße zur Weiterentwicklung. In: BfuP, vol. 1, S 17 - 35.

Huang K.-T. (1998): Capitalizing on intellectual assets. In: IBM Systems Journal, vol. 37, Nr. 4, S. 570 - 583.

Intellectual Capital - Supplement to Skandia's Interim Report: 1994 - 1998, Schweden.

Kaplan R./Norton D. (1997): Balanced Scorecard - Strategien erfolgreich umsetzen. Stuttgart: Schäffer Poeschel Verlag.

Krcmar H./Rehäuser J. (1996): Wissensmanagement im Unternehmen. In: Conrad, P., Schreyögg G (Hg.): Managementforschung 6 - Wissensmanagement. Berlin: de Gruyter.

Krickl O./Milchrahm E. (1998): Wissensmanagement in Lernenden Organisationen - Eine neue Herausforderung für die Informationswissenschaft? In: Zimmermann, H.; Schramm, V. (Hg.): Knowledge Management und Kommunikationssysteme. Konstanz: Universitätsverlag Konstanz, S. 88 - 102.

Kuhlen, R. (1989): Pragmatischer Mehrwert von Information. Sprachspiele mit informationswissenschaftlichen Grundbegriffen. Konstanz, Bericht 1/89.

Nonaka I./Takeuchi H. (1997): Die Organisation des Wissens. Wie japanische Unternehmen eine brachliegende Ressource nutzbar machen. Frankfurt/Main: Campus Verlag.

North K./Probst G./Romhardt K. (1998): Wissen messen - Ansätze, Erfahrungen und kritische Fragen. In: Zeitschrift Führung und Organisation, S. 158 - 166.

Probst G./Raub S./Romhardt K. (1998): Wissen managen: Wie Unternehmen ihre wertvollste Ressource optimal nutzen. Wiesbaden: Gabler Verlag.

Reinhardt R. (1998): Das Management von Wissenskapital. In: Pawlowsky P. (Hg.): Wissensmanagement: Erfahrungen und Perspektiven. Wiesbaden: Gabler Verlag, S. 146 - 176.

Romhardt K. (1998): Die Organisation aus der Wissensperspektive: Möglichkeiten und Grenzen der Intervention. Wiesbaden: Gabler Verlag.

Schneidermann A. (1988): Setting Quality Goals. In: Quality Progress, vol. 21,4,551-57.

Siegert T. (1999): Humankapital: Erfolgsmessung und Partizipation. In: Bühler W., Siegert T. (Hg.): Unternehmenssteuerung und Anreizsysteme. Stuttgart: Schäffer-Poeschl Verlag, S. 17 - 47.

Stewart, T. A. (1997): Intellectual Capital. London: Bautam Doubleday Dell Publishing Group.

Streim H. (2000): Die Vermittlung von entscheidungsnützlichen Informationen durch Bilanz und GuV. In: BfuP, vol. 2, S. 111 - 131.

Striening H. (1988): Prozessmanagement. Frankfurt/Main: Peter Lang Verlag.

URLs (Mai 2000):

[http://www.skandia.com/flashnot/index\\_upgradeflash.htm](http://www.skandia.com/flashnot/index_upgradeflash.htm)

<http://www.skandia.com/capital/valuescheme.htm>



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 127 – 143

# **INSYDER - Information Retrieval Aspects of a Business Intelligence System**

**Gabriela Mußler, Harald Reiterer, Thomas M. Mann**

Universität Konstanz

Fachbereich Informatik und Informationswissenschaft,

D-78457 Konstanz

{Gabriela.Mussler, Harald.Reiterer, Thomas.Mann} @uni-konstanz.de

## **Zusammenfassung**

Dieser Beitrag beschäftigt sich mit INSYDER, einem visuell unterstützten Suchsystem im Umfeld der sogenannten Business Intelligence Systeme. Dabei liegt der Schwerpunkt dieses Beitrags in der Erläuterung der Information Retrieval- und Visualisierungsaspekte zur Unterstützung des Suchprozesses im WWW. Vorgestellt wird das Rankingverfahren, das Interaktive Relevance Feedback sowie die Beschreibung der Unterstützung des Benutzers bei der Formulierung der Suchanfrage mittels Visualisierung.

## **Abstract**

This paper presents a visual information seeking system for the WWW called INSYDER<sup>1</sup>. The aim of INSYDER is to find business information from the WWW. Information seeking - especially on the Web - is an imprecise process. Information seekers often have only a fuzzy understanding of how they can get the information they want. This paper focuses on the use of information retrieval and visualisation techniques of the INSYDER system to support the user in extracting information from the WWW, namely in formulating queries, refining and reviewing results.

---

<sup>1</sup> The research project INSYDER (INternet SYsteme DE Recherche) was funded by a grant from the European Union, ESPRIT project number 29232.



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

## 1. Introduction

The benefits of using external information for business intelligence systems<sup>2</sup> are significant. An enterprise must know more and more about its customers, its suppliers, its competitors, government agencies, and many other external factors. Valuable information about external business factors is readily available on the WWW and its amount is increasing every hour. While a few WWW resources are used as data sources, the immense resources of the Internet are largely untapped. What is needed is a continuous and systematic approach to make use of these untapped resources. [Hackathorn 1998] proposes such an approach called Web farming: *"Web farming is the systematic refining of information resources on the Web for business intelligence."* The visual information seeking system INSYDER presented in this paper can be seen as a kind of Information Assistant [Kuhlen 1999] for finding and analysing business information from the Internet.

A detailed description of the INSYDER system can be found in [Mann, Reiterer 1999], [Mußler 1999] and first results of a user evaluation of the different visualisations are presented in [Reiterer, Mußler, Mann et al. 2000], [Mann, Reiterer 2000]. This paper focuses on the use of information retrieval and visualisation techniques to support the user in extracting information from the WWW, namely in formulating queries, refining and reviewing results. Here we shall present ideas and their prototypical implementation extending and continuing the original idea of INSYDER. These ideas were mainly born when reviewing results of various surveys concerning information seeking on the WWW. Conclusions of these surveys are that users often do not know how to express their information need [Nielsen 1997], [Pollock, Hockley 1997], that Relevance Feedback is only useful, if made transparent to the user [Koenemann, Belkin 1999], and that users have problems with the current paradigm of information retrieval systems simply presenting long lists of results [Zamir, Etzioni 1998]. In the following, we present our proposed solutions to these problems, which have been included in the overall Information Assistant approach of INSYDER<sup>3</sup>.

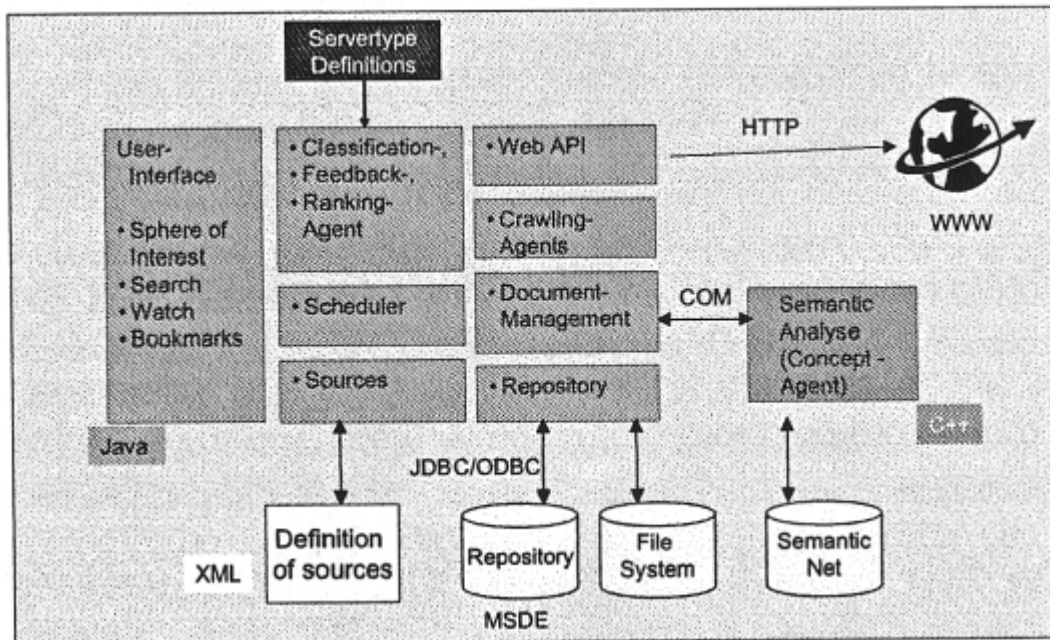
The paper is organised into the following chapters: Chapter two gives a short overview of the architecture and main components of the INSYDER system. Chapter three describes the novelty of the INSYDER system. Chapter four describes the enhancements to the existing system, which are partly finished, namely a query interface, called Visual Query, the two used ranking algorithms used in INSYDER, and the pursuit of the Relevance Feedback. Chapter five gives a summary of the ideas of this paper and an outlook on future work.

---

<sup>2</sup> "A business intelligence system ... provides a set of technologies and products for supplying users with the information they need to answer business questions, and make tactical and strategic business decisions." [IBM]

<sup>3</sup> After the official EU-project has been finished in March, the University of Konstanz is still working on the INSYDER system in close relation with the French company Arisem.

## 2. Overview of the INSYDER system



**Figure 1:** Architecture of the INSYDER system

The INSYDER Information Assistant acts on the user's behalf and is built up using different agents using Information Retrieval techniques and a synchronised visualisation approach.

Figure 1 shows the architecture of the INSYDER system consisting of several components mainly developed in Java. Only the component for the semantic analysis has been developed in C++, because it is a further development of an existing module by the project partner Arisem already written in C++.

The user-interface and all visualisations (e.g. Scatterplot, BarGraph, TileBar, Relevance Curve) have been developed in Java using the JFC (Java Foundation Classes) also called Swing.

The different agents are responsible for special retrieval tasks (e.g. crawling the WWW; clustering and ranking of the search results; preparing the relevance feedback for a new crawling).

The scheduler's is responsible for the monitoring process of the WWW (watch function). The watch function is able to check user-defined Web pages for changes regularly.

The sources are the representation of starting points of a search such as commercially available search-engines, email-servers, newsgroups or directories. All sources are defined in XML documents, which enables an easy maintenance and extension of the sources in a standardised format.

The **Web-API** is a set of functions and methods, which supports an easy access to the documents of the Web. The crawling agents use the Web-API for searching and crawling for Web documents, downloading them, and putting them into the document management component.

The **document management** component is responsible for the management of all documents and their metadata. It is the central component of our architecture and implements the classes and methods for the other components.

For every document the document management calls the **semantic analysis** (developed in C++) via a COM wrapper to get a relevance value for the specific document. The semantic analysis uses a semantic net that models the real world by a controlled vocabulary. This semantic net can be individually adapted to various application domains (e.g. building and construction; computer industry) to meet business demands. It consists of concepts (nodes) describing the semantics of the system by using typical relationships (typed links), like "is-a", "consists-of" and so on. With the help of this semantic net it is possible to find also documents, which do not contain the terms of the query, but for instance a synonym, an acronym, or broader or narrower terms (see chapter 5). Another advantage is that the results may be more precise than results from other information seeking system as homonyms can be avoided. For instance a search for "*bank*" could result in the financial institute, it could be the memory *bank* in the computer, or the *bank* at the shore. Specifying a domain-specific semantic net e.g. for the computer industry, INSYDER can determine that *bank* must have to do with the hardware of a computer and will apply this when calculating the ranking value for documents found.

The document-management component stores all retrieved web documents (only the text) in the **file-system** while the corresponding metadata are stored in the **repository** with a link to the web document in the file system. The communication between the document management component and the repository is made via the JDBC/ODBC interface.

### 3. Overview of the novelty of the INSYDER approach

The retrieval aspects of the visual information seeking system INSYDER have not been in the primary research focus. Nevertheless the system offers some retrieval features that are not very common in today's business intelligence tools used for Web searching.

INSYDER uses a *dynamic search* approach. The idea is to use an online search to discover relevant information by following links. The main advantage is that the current structure of the Web is searched, and not the index of a search engine. The dynamic search is based on special crawling agents. They use different heterogeneous sources (like search engines, Web directories, Web sites, documents) as starting points for the link traversal. For example the query terms are submitted to pre-selected search engines (like Altavista, Google etc.) and the



hyperlinks in the search results are then used for the further crawling in the Web. This way INSYDER can also be seen as a meta search engine. All documents found are downloaded, analysed incrementally to find out how good these documents match the query. This way we can guarantee that the documents presented in the result views are up-to-date. Inquirus [Lawrence, Giles 1998] does also an online analysis and an own and therefore consistent ranking of documents found by search engines, but is designed to be a mere meta search engine, as it does not any further crawling starting with the documents found.

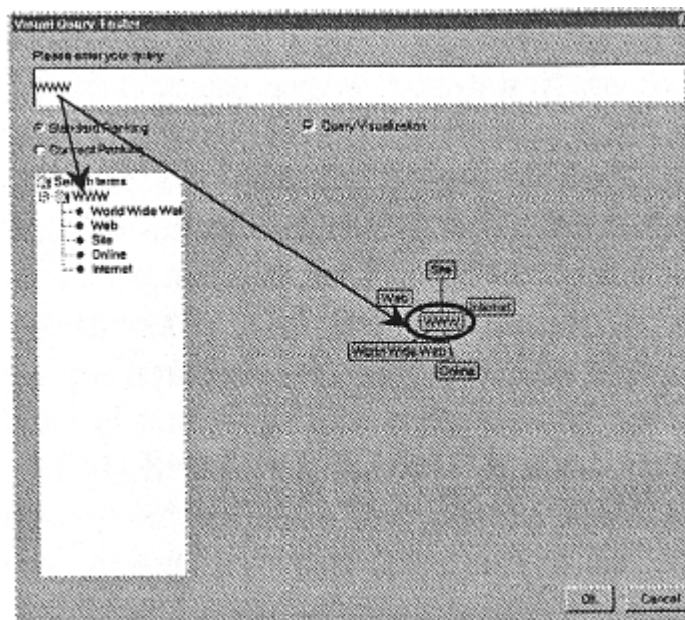
At the moment, our system has the drawback that we do not lookup the query in the own INSYDER repository. This means that it takes some time until search results are available, though they might be available in the own database. We know about this drawback and will do work on it. Still the main advantage of the system is the dynamic search. Unlike commercial search systems we do not intend to crawl all the WWW and store its contents, but only dedicated parts, which seem to be relevant for a given query by a user. By this way of specialising the search by specialising the crawling, e.g. for a specific branch like building and construction, we intend to increase the precision and recall compared to other meta search engines, which only rely on the results from the search engines indices.

Another special feature of INSYDER is that the relevance ranking uses a semantic analysis of documents, which is based on a *semantic net* provided with the system. Like a thesaurus the semantic net can be seen as a controlled vocabulary for the documents and the query. It offers important advantages such as identification of search terms with a clear semantic meaning, or retrieval based on concepts rather than on words. There is still an open debate how useful a thesaurus-based retrieval might be in the context of the Web, because a well known body of knowledge which can be associated with the documents in the Web does not exist. This was one of the reasons why the semantic nets used in INSYDER have been developed for two specific application domains (e.g. Building and Construction, CAD). By doing this we can also meet the demand by the business users from a Business Intelligence System to get more precise results than from a general search-engine.

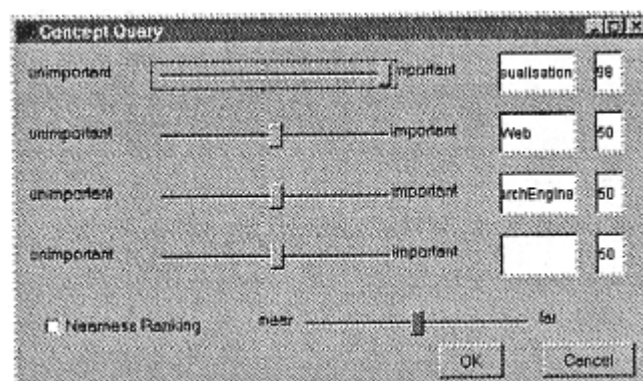
As the INSYDER System has no vector model built up from a known document source (as e.g. done by search engines using their database), other ways were needed to provide common Information Retrieval functions, as e.g. query expansion and as a special case of this the relevance feedback. Therefore INSYDER uses the above mentioned semantic net. As the WWW can not be compared to a given corpus, which can be determined by a distinct number of descriptors, a new way had to be found to describe documents from the WWW efficiently in a common way, e.g. for the relevance feedback. That is why we use the concepts of the semantic net. Using the concepts we take into account the disadvantage that it is possible that the concepts proposed are too broad for a query, e.g. information is expanded to 'wire of news', 'electronic Information', 'media'. Therefore we present the user in the Interactive Relevance Feedback a Suggestion which terms to use.

Using a weighted query based on a semantic net is new in the field of WWW, too. Though this part has not been evaluated yet, a first impression we have is that the results tend to be more precise. Still here is also the disadvantage that this is only true for concepts being well represented in the semantic net.

An important aspect of INSYDER thinking of its novelty is the fact that ideas and components from different fields were combined. It is for sure not new to combine visualisations and information retrieval aspects, but nowadays Systems which do a dynamic search with a metadata generation and the different visualisations of this metadata and document inherent data are new. Our approach aimed at getting the biggest added-value for the user combining components like dynamic search, visualisation of the query and different visualisations of the results (e.g. Scatterplot, Bargraph, Stacked Column) and information retrieval techniques (e.g. query expansion, ranking of results).



**Figure 2:** Prototype View of the Visual Query



**Figure 3:** Dialog window for the Concept Query4 Enhancements to the INSYDER System

## 4. Enhancements to the INSYDER System

### 4.1 Visual Query

The idea of the query visualisation is to help users to specify their information need more precisely using query expansion techniques and visualisation. From the literature it is well known that the average query consists of one or two query terms. This led to the demand of methods to overcome the problem of lacking knowledge to formulate queries. We propose therefore a Visual query, which will show the user terms and relationship of terms for his query (see figure 2), taking into account other successful Solutions and ideas from automatic query expansion and query visualisation, e.g. [Smeaton, Crimmins 1997], [Voorhees 1994], [Voorhees, Harmann 1998], [Zizi, Beaudouin-Lafon 1994] and [WebVibe].

First thing for the user to do is to type in his information need in a Standard text entry field. This text entry field has a wider than average size as proposed by [Shneiderman, Byrd, Croft 1997] to motivate the user to type in as many terms as he can think of. In a next step this query is visualised both in a tree view and optionally in a network graph view. The query terms the user entered are the basis for the tree view (each term is a branch, unless terms occur twice). This tree is then extended by related terms, which are inserted as leafs into the tree. These related terms are taken from a semantic net, which is a basic component of the semantic analysis of the overall system INSYDER. The semantic net models the real world by representing it with concepts. For example, the WWW is defined as a concept, Web is handled as a synonym and terms like Internet, Online, Site etc. are modelled as having a strong relationship to the term WWW. If the user searches for the term WWW, the system could propose him also the terms Internet, Online, Site etc. in the Visual Query to include these in his search. At the moment, such an inclusion is made implicitly as related terms are taken into account when calculating the quality of a document (a value for the match of a query and a document found, see section 4.2). But here the related terms are only taken partly into account, not fully (unless the term is a synonym). So it would be nice for the user to have two things: first the certainty about the query and second the possibility to include more terms (the user may not have been thinking of before) into his query. The visualisation of the query has been designed taking into account several aspects:

- *The semantic net includes various relationships (e.g. part-of broader term, narrower term):* These are not represented in the graph visualisation, just the fact that there is a relationship.
- *The user may have many input terms at first sight:* To keep the overview the system was designed having a detailed and a full view. This is simply done by taking the information from the tree view. E.g. if the user clicks on a branch of the tree view, only this branch is visualised in the graph, while clicking on the root of the tree (search terms) will result in a graphical presentation of the whole tree in the graph, which may be not easy to overview.

- *Interaction with die graph representation:* all terms represented in the graph representation can be moved keeping the relation to their base node. The elements are ordered automatically using an algorithm to make sure that when having many elements connected to a node most of them are viewable (see figure 2).

The Visual Query has been implemented prototypically. What is still missing is the connection to the semantic net to show the user directly related terms - the real added value of this feature. This problem derives from the fact that the INSYDER system brings together two programming worlds: on the one hand there is C++ for the handling of the semantic net, the semantic analysis, parts of the Relevance Feedback, and the ranking algorithms. On the other hand we have Java for the implementation of the GUI, for the assistants for the classification and visualisations and also for the Visual Query. The two worlds are communicating using a COM interface. As the access to the semantic net by COM objects has not been implemented yet and obviously appears not to be easy, this part is still missing. Meanwhile, we have thought about a workaround using a different semantic net, e.g. WordNet, but still further investigations in this direction have to be made.

## 4.2 Ranking

In the INSYDER system we have implemented two ranking algorithms, called Natural Language and Concept Query.

#### **4.2.1 Natural Language**

The Natural Language Algorithm is the default ranking algorithm of the system. Here, the user types in a query in a similar way he would express his information need. E.g. 'How many inhabitants has London?' This natural language query is then translated into a query for the crawler agents. Whereby translation means the elimination of stop words (common words in the different languages, e.g. 'and', 'or', 'many' etc.) and the elimination of punctuation. Doing this the given query would lead to a final query consisting of the following terms: 'inhabitants London'. In this step we believe to have a great added value from the visual query, because the user can see which terms are finally used for searching. This way it should be much clearer for him to derive why some documents were found. The two meaningful terms extracted from the original query are then sent to different search engines (chosen by the user), to get a first result from their indexes (for an overview on crawling agents see e.g. [Turau 1998]). From these results the crawling agents excerpt the links, which are then going to be the base for further crawling. The results coming from the crawling agents are given to the ranking assistant, which calculates a number, describing how well a document found by the crawling agent matches the query. Numbers range from 0 to 100, whereby 0 means no match at all and 100 means best match possible. Documents above a ranking value of 75 are very good documents thinking in terms of satisfaction from the user's point of view. 75 means that at least in one part of the document all query terms occur. These ranking values are calculated using a semantic analysis algorithm in the background. Documents found are represented in a meta description, consisting of information about segments (which are usually similar to the sentences) and their description in the semantic net. The query itself is also put in such a meta description. This way, the system has a basis for the comparison of the query and the document. The matching itself is done segment by segment. That is to say that a value for each segment is achieved hereby. We call this set of values the detailed rank. The values of this detailed rank are used for the visualisations, e.g. the relevance curve in the result table or the various representations in the segment view. The final ranking value is then calculated, taking into account the mean value of all segments and the maximal value reached of one or more segments. The maximum value gets a higher weight in the final calculation than the mean, to make sure that documents containing all query terms (from the system generated query) in the same segment are ranked higher than documents containing the query terms in a loosely order spread all over the document. This way, to calculate the overall relevance of a document can be compared to a Boolean 'AND' with a 'NEAR' proximity operator. The crawling agents will also find documents containing only some terms of the original query, having a much smaller value for the overall ranking value. This means an 'OR' compared to systems using Boolean logic. At the moment, there is no indication for this for the user, but this is planned for a future version of the system, e.g. as a representation in a scale. Hereby studies saying that the Boolean logic is too complex to most users shall be taken into account when designing this scale. [Shneiderman, Byrd, Croft 1997] for instance, propose to use a verbal description, e.g. 'contain all words', 'contain one or more words', and so forth.

#### 4.2.2 Concept Query

The Concept Query has been designed to give the user a higher level of influence and interaction, defining his information need. This Concept Query is not seen as a possibility for beginners in Web searching, but for advanced users. The idea behind is that the user himself defines how important the different query terms are for the satisfaction of his information need. Therefore the user weights the different concepts of his query. Speaking of query we mean a concept which must be defined in the semantic net. In figure 3 a prototype version of the implementation can be seen. The user asked here for the concepts 'Visualisation', 'Web' and 'SearchEngine'. Obviously the user is very much interested in the concept of 'Visualisation', giving this concept an 'important' value of 98, while the other two concepts are not seen as important, both given a weight of 50. In the dialog window a nearness check box with a slider can be seen as well. This is another interaction possibility for the user. Here the user has the possibility to define, if all the concepts should appear next to each other (*near* e.g. in the same segment) or could be spread all over the document (*far*). Depending on the selection the final ranking value for the documents differs. As in the *Natural Language* algorithm, the overall ranking value is derived from the detailed rank value, calculated for each segment.

For the moment, the use of the Concept Query is rather uncomfortable, as the user has to know the concept terms from the semantic net. This means there is temporarily no translation from a given query (which could be natural language) to the concepts. In the example above this would be no problem, as 'Inhabitant' and 'London' are both concepts of the semantic net.

At this point it is planned to combine the Concept Query window with the Visual Query window. Instead of having a slider for the weight, we think of having a spin button next to the concepts in the graph or where the user can define the weight then. Another idea we have in mind is the use of colour and thickness of the concepts, e.g. the user could set the weight by a slider coming up when clicking with the right mouse button or by a separate dialog box. The intensity of the concept itself (or the frame) could then be a visual indicator for the weight of a concept (fade - unimportant, strong - important). These ideas have been thought about, but there is no prototypical implementation available yet.

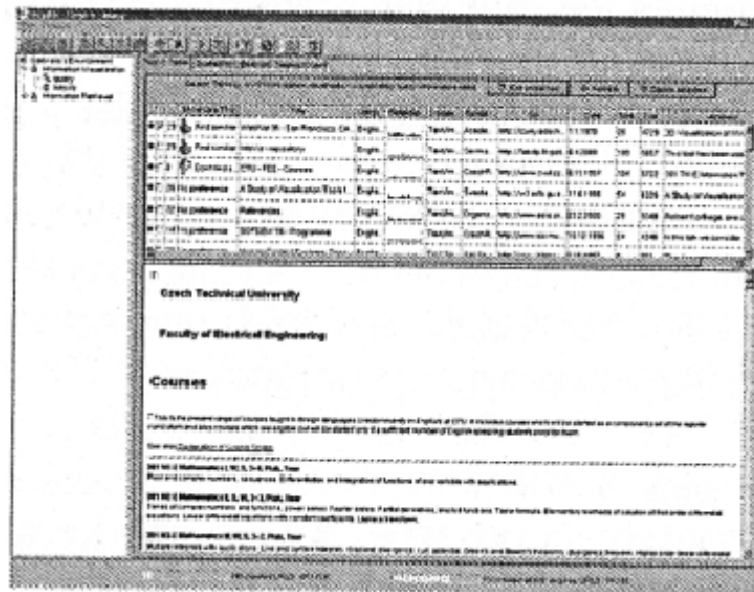
#### 4.2.3 Evaluation planned

The effectiveness of the system shall be measured by using the *tfidf* (term frequency inverse document frequency) measurement as a baseline for the ranking algorithms. Therefore the *tfidf* measurement has been integrated into the system. The basis for the *idf* measurement is the actual document collection, meaning the documents which have been found by the web crawling agents. The *tfidf* is calculated taking the formula of [Salton, McGill 1987]. Thinking about measuring the precision at different cut-off levels, we take the studies of [Buckley, Voorhess 2000] into account. They suggest that a cut-off level of 10 to 20 makes sense for the evaluation of Web search engines, but also point out that with increasing

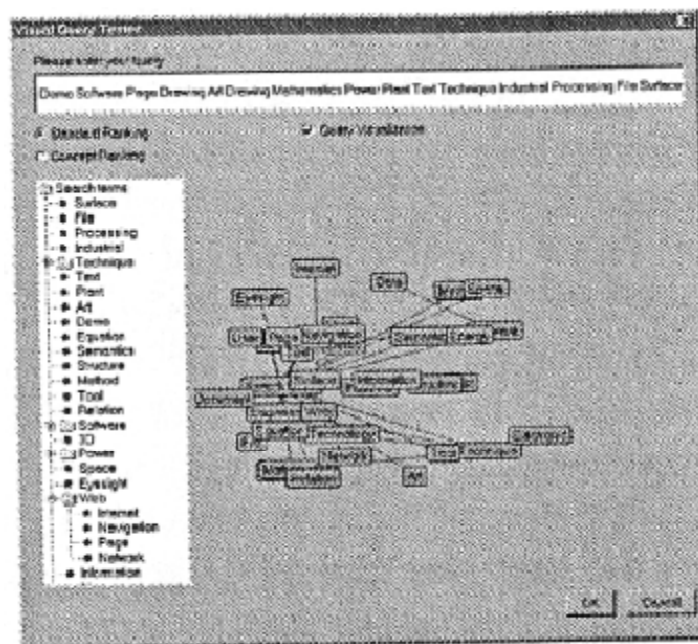
number of topics (queries to the search systems) the error level decreases. For our evaluation purpose we have therefore decided to choose 30 of the topics provided by [TREC]. The topics are described by keywords, phrases and give a narrative, what a document has to fulfil to be judged relevant. It is planned to measure the precision at a cut-off level of 20 as maximum. This is done to keep the error level at a low value and still to be able to handle the evaluation data. It is also planned to do an evaluation of different systems, similar to the evaluation of search engines by [Leighton, Srivastava 1997]. The evaluation will differ from their evaluation, as meta search engines and local search engines are going to be used. The purpose of this evaluation will mainly focus on the efficiency of the systems finding relevant information. For the time being we are working on the test settings.

### **4.3 Interactive Relevance Feedback**

As it has been said before INSYDERs retrieval functions base very much on the thesaurus. Therefore it was decided to give the user just suggestions for terms for the relevance feedback as otherwise the documents found (using a fully automatic relevance feedback) would be too broad (low in precision), covering too many different subjects. Hence INSYDER offers the function of an Interactive Relevance Feedback, meaning the use of judgements made by the user about documents (like, dislike) to derive from these new query terms, which are suggested to the user. The first approach was to extract the feature concepts of a document. Meaning that an analysis of the positive or negative judged document is made, extracting from each document those concepts describing it best. Hereby best describing means that this concept occurs most often in a document. As the concept 'proper name' might occur several times derived from different nouns (e.g. Smith, Hull, Clarke would be three times the concept proper name), these concepts are ignored, the same applies for concepts like number etc. and verbs. The remaining concepts are ordered according to their frequency. The frequency itself is calculated using a positive and a negative weight. Concepts from documents judged as relevant ('find more like this') are added while negative concepts are subtracted. In the example about a search on "information visualisation" (see figure 4) we can see that the user decided to have more like the first two documents and nothing similar to the third one.



**Figure 4:** View of the INSYDER system with relevance judgement by the user



**Figure 5:** Prototype View of the Visual Query with an overview presentation of the input from the Relevance Feedback.

A negative thing about the fact that proper names are eliminated is that if a document e.g. is about the biography of a special person, the person's name will never occur in the Relevance Feedback query. As it was not possible for technical reasons to get the real name of the concepts (e.g. instead of proper name - Smith), we decided to try the classical approach of text analysis, too. Therefore we use the Porter algorithm for English documents [Porter 1997] beside the elimination of stopwords to get the feature terms of the documents. Using the Porter Algorithm we get the principle forms of all the words in the document. By



calculating their frequency and the re-translation to the original term, we get the possibility to include proper names in the Relevance Feed-back query, too. We decided to use a mixture of these two algorithms to get the best out of the two of them: by the concepts the user widens his view of things, by using the classical approach we do not have to sacrifice the proper names. Still we found out that the use of the two algorithms is very time consuming (particularly the classical algorithm), so that we are thinking about providing only the concept algorithm instead of the hybrid version of both of them, giving the user the Option of using both of them, if he wants (e.g. for a second refinement step). The query terms created are presented to the user in the Visual Query window (see figure 5). The user has the possibility to have an overview or detailed view of the feature concepts and terms as described in chapter 5. Unfortunately, here are also the relations from the semantic net missing, as for technical reasons there is no direct connection available yet. From the Visual Query the user can select terms to go on with.

## **5. Conclusion and Outlook**

In this paper we have focused on the description of enhancements to the existing Information Assistant INSYDER. These were in particular the Visual Query and the Concept Query Ranking Algorithm. Apart from these the Natural Language Ranking Algorithm and the enhancement of the systems feature the Relevance Feedback using concepts was explained. Some of the ideas presented are still ideas and need to be implemented. Throughout the ideas presented above we are still working on the enhancement of the overall system. This includes the visualisations of the search results, the visualisation algorithms and particularly the user interface of the whole application. We have started to redesign the system using the Java Web Style Guide [Sun 2000]. Still we are working to solve the problem to get a direct connection to the semantic net in close co-operation with our French partner ARTSEM. Moreover evaluations of the ranking, the Relevance Feedback and the usability of the Visual Query have to be made soon.

## **Acknowledgements**

The authors wish to thank the following EU-ESPRIT project partners for their contributions and support: Alain Garnier, Olivier Spinelli, Laurent Dosdat (ARISEM, Paris); Guillaume Lory, Carlo Revelli (Cybion, Paris); Rina Angeletti (Innova, Rome); Flavia D'Auria (Promoroma, Rome).

## **References**

[Arisem] <http://www.arisem.com> [2000-01-18].

[Buckley, Voorhees 2000] C. Buckley, E. Voorhees. Evaluating Evaluation Measure Stability. Proceedings of the annual International ACM SIGIR Conference on Research and Development in Information Retrieval SIGIR '00, Athens 24-28 July 2000.

[Hackathorn 1998] R. Hackathorn. Web farming for the data warehouse. Morgan Kaufmann, San Francisco, 1998 .

[IBM]

<http://www.software.ibm.com/data/pubs/papers/bisoluti on/index.html>  
[2000-01-18].

[Koenemann, Belkin 1999] J. Koenemann and N. J. Belkin. A Case for Interaction: A Study of Interactive Information Retrieval Behavior and Effectiveness. CHI 96 - Electronic Proceedings. R. Bilger, S. Guest and M. J. Tauber (eds.). [http://www.uni-paderborn.de/StaffWeb/chi96/E1Pub/WWW/chi96www/papers/Koenemann/jkl\\_txt.htm](http://www.uni-paderborn.de/StaffWeb/chi96/E1Pub/WWW/chi96www/papers/Koenemann/jkl_txt.htm) [1999-11-11].

[Kuhlen 1999] R. Kuhlen. Die Konsequenzen von Informationsassistenten. Was bedeutet informationelle Autonomie oder wie kann Vertrauen in elektronische Dienste in offenen Informationsmärkten gesichert werden? Frankfurt (Suhrkamp-Verlag). 1999.

[Lawrence, Giles 1998] S. Lawrence and L. Giles. Inquirus, The NECI Meta Search Engine, Proceedings of the Seventh International World Wide Web Conference, Brisbane, Australia, 1998.

[Leighton, Srivastava 1997] V. Leighton; J. Srivastava: Precision among World Wide Web Search Services (search engines). <http://www.winona.msus.edu/library/webind2/webind2.htm> [1999-03-02]

[Mann, Reiterer 1999] T. M. Mann, H. Reiterer. Case Study: A Combined Visualization Approach for WWW-Search Results. IEEE Information Visualization Symposium. N. Gershon, J. Dill and G. Wills (eds.). 1999 Late Breaking Hot Topics Proc. Supplement to: G. Will, D. Keim (eds.): Proc. 1999 IEEE Symposium on Information Visualization (InfoVis'99). Conference: San Francisco, CA, USA, October 24-29, 1999. Los Alamitos, CA (IEEE Computer Soc. Press). San Francisco 1999: 59-62. 1999.

[Mann, Reiterer 2000] T. M. Mann, H. Reiterer. Evaluation of different Visualizations of WWW Search Results. Proc. Eleventh International Workshops on Database and Expert Systems Applications (DEXA 2000). Conference: Greenwich, UK, September 4-8, 2000 (IEEE Computer Society).

[Mußler 1999] G. Mußler. Ein Agentensystem zur Unterstützung bei der Informationssuche im WWW. Electronic Proceedings <http://wwwdb.informatik.uni-rostock.de/adi99/endversionGM.ps> [1999-10-05].

[Nielsen 1997] J. Nielsen. Search and You May Find. <http://www.useit.com/alertbox/9707b.html> [1999-03-18].

[Pollock, Hockley 1997] A. Pollock and A. Hockley. What's Wrong with Internet Searching. D-Lib Magazine, 1997, <http://www.dlib.org/dlib/march97/bt/03pollock.html> [1999-02-01].

[Porter 1997] M. Porter. An algorithm for suffix stripping. Readings in Information Retrieval. K. Sparck-Jones and P. Willett (eds.). Morgan Kaufmann, San Francisco, 1997.

[Reiterer, Mussler, Mann et al. 2000] H. Reiterer, G. Mußler, T. M. Mann and S. Handschuh: INSYDER - An Information Assistant for Business Intelligence. Proceedings of the annual International ACM SIGIR Conference on Research and Development in Information Retrieval SIGIR '00, Athens 24-28 July 2000.

[Salton, McGill 1987] G. Salton and M. J. McGill. Information Retrieval - Grundlegendes für Informationswissenschaftler. McGraw-Hill, 1987.

[Shneiderman, Byrd, Croft 1997] B. Shneiderman, D. Byrd, and W. B. Croft. Clarifying Search: A User-Interface Framework for Text Searches. D-Lib Magazine, 1997, [http://www.dlib.org/dlib/january97/retrieval/01\\_shneiderman.html](http://www.dlib.org/dlib/january97/retrieval/01_shneiderman.html) [1999-08-17].

[Smeaton, Crimmins 1997] A. F. Smeaton and F. Crimmins. Relevance Feed-back and Query Expansion for Searching the Web: A Model for Searching a Digital Library. Research and Advanced Technology for Digital Libraries. P. C. Pisa and C. Thanos (eds.). First European Conference, ECDL'97 (Springer) 1997.

[SUN 2000] <http://java.sun.com/products/j1f/dg/higtitle.alt.htm> [2000-09-04].

[TREC] TREC-8 ad hoc and small web topics  
[http://trec.nist.gov/data/topics\\_eng/topics.401-450.gz](http://trec.nist.gov/data/topics_eng/topics.401-450.gz) [2000-08-24].

[Turau 1998] V. Turau. Web-Roboter. Informatik Spektrum, 21(3): 159-160, 1998.

[Voorhees 1994] E. M. Voorhees. Query Expansion using Lexical-Semantic Relations. Proceedings of the seventeenth annual International ACM SIGIR Conference on Research and Development in Information Retrieval SIGIR '94, 3-6 July 1994. B. W. Croft, (eds.). Dublin, Ireland, London, Berlin u.a. (Springer): 61-69, 1994.

[Voorhees, Harman 1998] E. M. Voorhees and D. K. Harman (eds.): NIST Special Publication 500-242: The Seventh Text Retrieval Conference (TREC-7) Gaithersburg, Maryland (Government Printing Office (GPO)) 1998. [http://trec.nist.gov/pubs/trec7/t7\\_proceedings.html](http://trec.nist.gov/pubs/trec7/t7_proceedings.html) [1999-12-20].

[WebVibe] <http://www2.sis.pitt.edu/Lwebvibe/WebVibe/> [2000-09-06].

[Zamir, Etzioni 1998] O. Zamir and O. Etzioni. Web Document Clustering: A Feasibility Demonstration. SIGIR 1998. <http://zhadum.cs.washington.edu/zamir/sigir98.ps> [1999-03-23].

[Zizi; Beaudouin-Lafon 1994] M. Zizi; M. Beaudouin-Lafon. Accessing Hyperdocuments through Interactive Dynamic Maps. Conference on Hypertext and Hypermedia Proceedings of the 1994 ACM European conference on Hypermedia technology. 126-134.



In: Knorz, Gerhard; Kuhlen, Rainer (Hg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000), Darmstadt, 8. – 10. November 2000. Konstanz: UVK Verlagsgesellschaft mbH, 2000. S. 145 – 162

# Automatic Document Classification:

## A thorough Evaluation of various Methods

C. Goller<sup>1</sup>, J. Löning<sup>2</sup>, T. Will<sup>1</sup>, W. Wolff<sup>2</sup>

SAIL-LABS<sup>1</sup>  
Balanstr. 57  
81541 München  
Germany  
<http://www.sail-labs.de>

iXEC<sup>2</sup>  
Lilienthalstr. 25  
D-85399 Hallbergmoos  
Germany  
<http://www.ixec.com>

Email: [christoph.goller@sail-labs.de](mailto:christoph.goller@sail-labs.de)

### Abstract

(Automatic) document classification is generally defined as content-based assignment of one or more predefined categories to documents. Usually, machine learning, statistical pattern recognition, or neural network approaches are used to construct classifiers automatically. In this paper we thoroughly evaluate a wide variety of these methods on a document classification task for German text. We evaluate different feature construction and selection methods and various classifiers. Our main results are: (1) feature selection is necessary not only to reduce learning and classification time, but also to avoid overfitting (even for Support Vector Machines); (2) surprisingly, our morphological analysis does not improve classification quality compared to a letter 5-gram approach; (3) Support Vector Machines are significantly better than all other classification methods.

### 1. Introduction

*Document classification* is known under a number of synonyms such as *document/text categorization/routing* and *topic identification*. For more work on automatic document classification see e.g. the TREC-conference series [1].

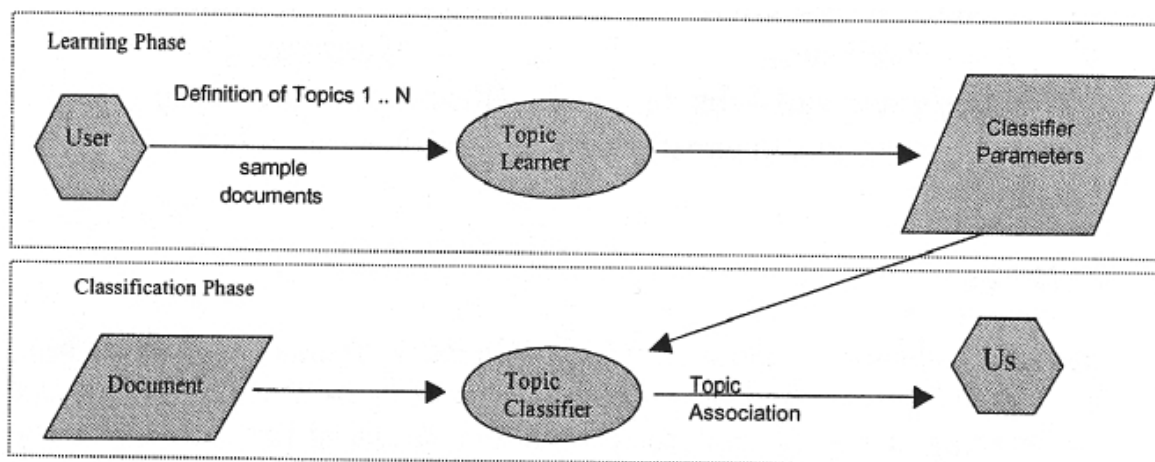
Basically document classification can be defined as *content-based assignment* of one or more *predefined categories (topics)* to documents. Document classification can be used for document filtering and routing to topic-specific processing mechanisms such as information extraction and machine translation. However, it is equally useful for filtering and routing documents directly to humans.



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:  
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

Applications are e.g. filtering of news articles for knowledge workers, routing of customer email in a customer service department, or detection and identification of criminal activities for police, military, or crete service environments.

Manual document classification is known to be an expensive and timeconsuming task. *Machine learning* approaches to classification suggest the automatic construction of classifiers using induction over *pre-classified sample documents*. In this paper we thoroughly evaluate and compare various methods for this kind of automatic document classification.



**Figure 2:** Document Classification

Section 2 defines automatic document classification and relates it to document retrieval. In Section 3 we describe the document corpus used for our experiments. In Section 4 we introduce the different approaches that we have selected and evaluate them from a theoretical point of view. The experimental comparison follows in Section 5. We summarize our results in Section 6 and present some ideas for further improving automatic document classification.

## 2. Automatic Document Classification

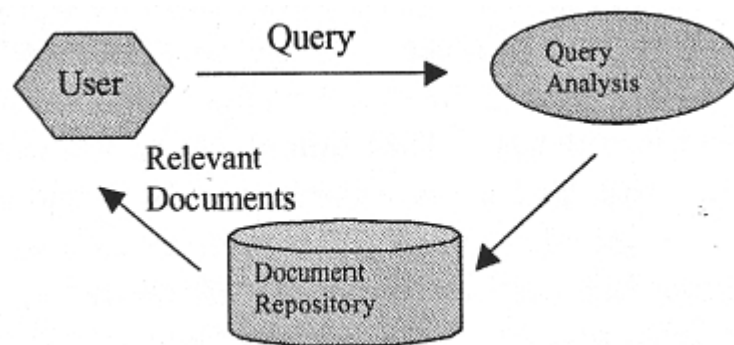
We can distinguish two phases in automatic document classification, the *learning phase* and the subsequent *classification phase* (see Figure 1). In the *learning phase* users define categories (topics) in which they are interested (their *information need*) by giving *sample documents* (*training examples*) for each of these categories. Most methods for automatic document classification also require *counterexamples* for each category that is sample documents that do not deal with the respective topic.

In a standard application for automatic document classification like news filtering users assign categories to the documents of a selected collection by hand. Documents may be assigned to more than one category if they deal with several of the topics or if one has a hierarchy of topics. The document collection used for learning should be as representative as possible for the documents that one

expects to encounter in the future. All documents that are not assigned to a category serve as counterexamples for this category.

The topic learner component analyzes the sample documents and tries to identify the content that is specific for each category. This analysis is essentially an *inductive reasoning* process which involves *generalization* and *abstraction*. The output is a *model* for each category which is represented by a set of *classifier parameters*. Most classifiers make *a priori assumptions* about the underlying model and its *complexity*. If the assumed model complexity is too high, *overfitting* can occur. This means that the model overspecializes with respect to the training examples and generalization to new previously unseen examples becomes bad. Overfitting is one of the biggest problems in machine learning. Normally one assumes that the learning phase is invoked very rarely. Therefore efficiency is not of primary importance.

In the *classification phase* new (previously unseen) documents can be given to the topic classifier which returns a topic association (a rating or classification for each topic).



**Figure 3:** Document Retrieval

Basically, it tells the user whether the document deals with the training topics and especially with which topics it deals. It is assumed that a lot of documents have to be classified. In a news filtering application new messages constantly arrive and have to be filtered or routed. Therefore efficiency is very important for the classification phase.

Automatic document classification is closely related to *document retrieval* (see Figure 2). In document retrieval the user specifies his *information need* by giving a *query*, which is analyzed and then applied to a relatively fixed and preprocessed (indexed) document corpus. The relevant documents are presented to the user.

There are two main differences between automatic document classification and document retrieval. The First difference is a rather technical one. In document retrieval the documents which are compared to the users information need are typically members of a big and relatively fixed document corpus which is preprocessed (indexed) independently of the individual queries. Normally, queries are small. They can be applied to the whole corpus in parallel and do not have to be applied to every document sequentially. Usually a lot of queries have to be answered and new queries constantly arrive. Therefore query analysis also has to

be very efficient. In document classification, the users' information need (topics) is relatively stable. However, one does not have a fixed document collection which can be preprocessed. Normally documents have to be processed sequentially by the classifier.

The second difference concerns the way how the information need is specified. In document retrieval the users' *information need* is specified by usually small queries (keywords). The users have to know exactly what they are interested in. Differences in language uses between authors of documents and users who specify queries often lead to low recall. Therefore, query expansion by using thesauri or domain models is very important. In document classification users specify their information need by selecting documents as examples and counterexamples for a topic. If there are enough sample documents and if they are representative for the kind of documents usually encountered, classification will be quite good. The users don't have to know exactly what they are looking for and the use of thesauri or domain models is less important. However, this difference between document classification and retrieval disappears if we consider document retrieval with user feedback. Here the users are allowed to refine their original queries by selecting relevant and non-relevant documents from the documents that are presented by the retrieval system as answer to their original query.

### **3. Test Corpus**

As test corpus for automatic document classification we used a subset of articles from the German newspaper "Süddeutsche Zeitung (1998)" (SZ). The CD-Rom version [2] of the SZ contains topic-labels for most of the articles. Some articles are labeled with more than one topic. From the articles from the "München" section (local news) we selected those 20 topics with the highest numbers of articles (see Table 1). The total number of articles used in the experiments is 1083. This is less than the sum of column 2 from Table 1 due to multiple topic assignments for some of the articles.



Topic	#of articles	Topic	#of articles
Raubüberfälle	112	Flughafen München	47
Abschiebung von "Mehmet"	95	Mordversuche	45
Das ganze Leben ist ein Test	79	Verkehrsunfälle	45
Diebstähle	71	Betrugsdelikte	43
Oktoberfest	62	Adventskalender der SZ	42
Mordfälle	62	S-Bahn in München	39
Umbau Olympiastadion	60	Ausländer	37
Uhl, Hans-Peter	52	Illegaler Drogenhandel	37
Brände	52	Landtagswahl 1998	36
SZ-Serie Vier für München	51	Kiest, Erich	35

**Table 1:** Test Corpus

## 4. Selected Approaches and Theoretical Evaluation

### 4.1 Feature Construction

In Section 2 we define automatic document classification as an inductive reasoning task. This of course suggests the use of classification techniques from the fields of machine learning, statistical pattern recognition and neural networks. However, almost all existing learning and classification techniques require vectors of (real) numbers as input. They cannot work directly on documents (text). Therefore, vector representations of documents have to be constructed in order to make these methods applicable. The process of constructing these vector representations is called *feature (vector) construction*. Feature construction methods generally differ in the amount of linguistic and statistic sophistication that is applied. We compared two kinds of features, viz. letter 5-gram features and features constructed by a morphological analysis.

#### 4.1.1 Letter 5-gram Features

For our first approach we generated letter 5-grams without applying any tokenization. Special characters like "!%&\$" and whitespaces were included. We did not distinguish upper and lower case characters. The reason for including inter-word 5-grams is that in this way at least some multiword phrases (e.g. 'New York' in the example sentence below) are represented. The reason for including whitespaces is that we wanted short important strings like " IBM " (that are always included in inter-word 5-grams) not to become unrecognizable because of their varying context. For the sentence ' New York ist groß.', we get the following 5-grams: 'new y' 'ew yo', 'w yor', 'yourk', 'york ' ork\_i', 'rk\_is', 'kist', ' -groß', 'groß'.

Topic	1	2	3	4	5
Abschiebung von "Mehmet"	ehmet	mehme	hmet"	met" _	"mehm
Adventskalender der SZ	ür_gu	skate	ute_w	erke	adven
Ausländer	ölker	völke	bevö	evölk	lkeru
Betrugsdelikte	etrug	betru	trugs	ädigt	aatsa
Brände	brand	bran	euerw	rwehr	erweh
Diebstähle	einbr	einb	dieb	stohl	estoh
Flughafen München	ghafe	hafen	flugh	lughaf	ughaf
Illegaler Drogenhandel	rogen	ealer	droge	drog	deale
Kiesl, Erich	lesl_	h_kie	ch ki	kiest	kies
Landtagswahl 1998	agswa	wahl	wähl	ähler	tagsw
Mordfälle	mord	mord	s_opf	pfers	es_op
Mordversuche	mord	hlags	bensg	tots	totsch
Oktoberfest	wiesn	wies	berfe	iesn-	oberf
Raubüberfälle	berfa	erfal	überf	rfall	räube
S-Bahn in München	bahn-	s-bah	-bahn	_s-ba	e_s-b
Das ganze Leben ist ein Test	_test	r_tes	prüf	klass	bewer
SZ-Serie Vier für München	ler_f	ür mü	für m	vier	r mün
Uhl, Hans-Peter	_uhl	r_uhl	er uh	_uhl	hans-
Umbau des Olympiastadions	stadi	adion	tadio	_umba	umbau
Verkehrsunfälle in München	prall	unfal	Hetz	erlet	unfa

**Table 2:** The five most relevant 5-gram features for each topic

#### 4.1.2 Morpheme Features

For our second approach we implemented a simple *morphological analyzer* for German which combines *inflectional* and *derivational stemming* and a *word compound analysis*. After *tokenization* tokens are analyzed as concatenations of the following four morphological categories:

	category	examples		lexicon entries <sup>1</sup>
		Verurteilung	aber	
v	prefix	ver		111
s	(word-)stem	urteil		6.614
e	suffix	ung		97
t	solitary stein		aber	132

An admissible analysis is subject to the following two restrictions:

words analyzed as t consist of just this morpheme (no pre- or suffixes):  $p = t$   
otherwise, the analysis of words follows the regular pattern p:

$$p = w^+ \text{ and } w^- v^+ s^+ e^+$$

In order to tune the morphological analyzer heuristics concerning the ratio between the morphological categories in an analyzed token were developed.

We tested our morphological analyzer on the Munich part of the SZ (176000 unique tokens). For our test 1000 tokens were randomly selected. The analyzer was able to analyze 77.4% of the test tokens. 91.3% of the analyzed words were classified as correctly analyzed by humans. 5.1% of the analyzed tokens were proper names or misspelled words that were analyzed nonetheless. For 3.6% the analysis was incorrect. The 22.6% of the tokens that could not be analyzed by our morphological analyzer were mainly irreducible words such as names, abbreviations, monosyllables, foreign words, misspellings, dialect words, or numbers. Only 7.7% of the tokens that could not be analyzed could be analyzed by humans.

---

<sup>1</sup> The starting point for prefix, suffix, and stein lexicons were lexicons developed for hyphenation by the Technical University of Vienna [3]. However, we considerably extended and modified these lexicons.

Topic	1	2	3	4	5
Abschiebung von	mehmet	mehmets	eltern	ausweis	türk
Adventskalender der SZ	kalend	adventkalen	advent	Süd-deutsche	hilfwerk
Ausländer	völk	bevölk	Länder	ausländer	multikultu
Betrussdelikt	betrug	tru	eeschäd	reu	zweihalb
Brände	Brand	feuer	feuerwehr	lösch	wehr
Diebstähle	dieb	stohl	eestohl	einbrech	stahl
Fluthafen München	hafen	flughafen	flus	airport	fine
Illegaler Drogenhandel	deal	drog	rauschgift	rausch	heroin
Kiesl, Rich	kiesl	erich	altober	kiesls	altob
Lantagswahl 1998	tagswahl	wähl	wahl	stimmkreis	landtag
Mordfälle	mord	töt	Mord-kommis	kommüss	opfer
Mordversuche	mord	Mord-versuch	mordsuch	stoch	totschlag
Oktoberfest	wiesn	Oktoberfest	zelt	oktober	bier
Raubüberfälle	überfall	raub	räuber	beut	tät
S Bahn in München	fahrgäste	verspät	stamm-streck	streck	bahnhof
Das ganze Leben ist ein	test	prüf	bewerb	schul	klass
SZ-Serie Vier für München	vier für	uns	ia	was	häus
Uhl, Hans-Peter	hanspete	uhl	peter	hans	csu
Umbau des Olympiastadion	stadion	o.-stadion	umbau	olympia	architekt
Verkehrsunfälle	unfall	schleuder	prall	erfaßt	verletz

**Table 3:** The five most relevant morpheme features for each topic.

From every analyzed token we generated the following features: (1) the whole token without suffix; (2) every stem with its associated prefix and without suffix; (3) every stem; (4) 2-grams of stems. Furthermore, every irreducible token became a feature by itself. In order to represent *multiword phrases*, we additionally included those inter-word 2-grams of stems which had a very high frequency in our corpus.

#### 4.1.3 From Features to Feature Vectors

The features described in Sections 4.1.1 and 4.1.2 are strings of characters. The number of unique features in the document collection determines the *dimension* of the feature vector representations for the documents and the Position of each feature in an alphabetically ordered list of all features determines its position in the feature vector representations. A *feature vector representation* for a document is

simply a vector of weights for all the features<sup>2</sup>.

We started our experiments with *binary weights* (only using the information whether a feature occurs in a document or not). However, in a first evaluation phase we found that using the *term frequency (TF)*, which is the number of occurrences of the feature in the document, improves classification quality for all classifiers. We also found that using the product of term frequency and *inverse document frequency* ( $IDF = \# \text{ documents in the collection} / \# \text{ documents that contain the feature}$ ), which is a very popular weighting scheme in document retrieval, improves classification quality for the *Centroid classifier* and the *Support Vector Machine* considerably, while it does not change results for the other classifiers<sup>3</sup>. For Support Vector Machines a refinement (the square root of TFIDF) further improves classification quality. For a brief description of the classifiers that we used see Section 4.3. For a discussion of TFIDF see e.g. [4].

#### 4.1.4 General Comparison of N-gram Features and Morpheme Features

N-gram letter features have a variety of advantages compared to morpheme features. Implementation of n-gram feature construction is very easy and *independent of language*. N-gram features automatically perform certain kinds of *stemming* and they are robust against *misspellings*. Furthermore, N-gram features automatically capture many kinds of *multiword phrases*, if one considers n-grams across word-borders (inter-word n-grams).

Morphological analysis is of course *language-specific* and it is quite an effort to implement a morphological analyzer. One has to develop the lexicons, take care of incorrectly analyzed words and ambiguities, and one has to avoid both, *over-* and *under-stemming*. For German, *word compound analysis* is also very important. Furthermore, a morphological analyzer consumes more computational resources than n-gram analysis.

However, there are also big advantages which speak for morpheme features. It seems clear that morpheme features contain more information than n-gram features, since they represent *meaningful word-constituents* even if these constituents are longer than n characters. Therefore, classifiers built from morpheme features are more *understandable* than those built from n-gram features. Normally, there are considerably fewer unique morpheme features than e.g. 5-gram features for a given document and morpheme features are less correlated than n-gram features. The *lower dimension* of morpheme feature vectors and the *lower correlatedness* make feature selection / dimensionality reduction and learning easier. In our experiments the *inverse file index* of the test collection for 5-gram features is five times bigger than for morpheme features leading to an increase in feature selection costs of a factor five. Tables 2 and 3 list the five most important features for each of our test topics determined by the mutual information measure (see Section 4.2). The higher correlatedness of 5-gram features can easily be

<sup>2</sup> If a new document (classification phase) contains features not encountered in the document collection, these features are simply ignored.

<sup>3</sup> Interestingly, even the Perceptron was not able to profit from the additional IDF-weighting.

recognized since for several topics the five most relevant 5-gram features originate from only one word. For an experimental comparison of 5-gram and morpheme features see Section 5.4.

## 4.2 Feature Selection / Dimensionality Reduction

For our test corpus about 30000 unique morpheme features and about 185000 unique 5-gram features were generated. Most classification methods cannot handle such *high-dimensional input* (computational costs for learning and/or classification become intractable). A further problem is that the model complexity for many classifiers increases with the dimension of their input. This means that high-dimensional input vectors can cause overfitting. Therefore, *dimensionality reduction* or *feature subset selection* methods have to be applied. There are a lot of methods for dimensionality reduction around in the statistics literature. One of them is *principal component analysis*, in which orthogonal (empirically independent) linear combinations of the original features are determined which have the highest variance. In the document retrieval and classification area this approach is known as *latent semantic indexing* [5]. However, the computational costs are very high.

We decided to use a feature subset selection method. This means that for each topic a subset of the most relevant features is determined. In order to determine the relevancy of a feature with respect to a topic, we used the *mutual information* measure. The definition of mutual information can be found in every textbook on information theory. Basically, the mutual information between a feature and a topic tells us how much information the feature provides for the topic (and vice versa). Of course one does not know the mutual information between features and topics. We estimated the mutual information between the property that a document belongs to the considered category and the property that the document contains the considered feature. Probabilities were estimated using empirical relative frequencies. Tables 2 and 3 list the five most relevant features found for our test topics in this way.

Furthermore, we considered two extensions of feature subset selection with mutual information: *FuCE* (Feature subset selection using Classifier Errors) and *Autofeatures*.

*FuCE* works as follows. First a set of relevant features is constructed for every topic in the way described above. The classifier is trained using these features. Those documents of the training set which the classifier does not classify correctly are used to extend the set of features. *False positives* and *false negatives* are used separately. For *false positives* we determine features that distinguish them from real positives. In the same way we determine features that distinguish *false negatives* from real negatives. For both tasks we again used mutual information.

Feature subset selection requires to determine the number of features one wants to use and we extensively explored how the classification performance depends on the number of selected features (Section 5.3). With *Autofeatures* we

implemented a method that determines the number of relevant features for feature subset selection in a *topic-dependent way*. Features are selected until the sum of their individual mutual information with the topic is greater than the entropy of the topic. This takes into account that if there are some very unique features for a topic, only these few features are selected. On the other hand, if there are no features containing much information for the topic, this method selects a bigger subset of. The problem with this method is, that features normally are not independent of each other. Therefore, their joined mutual information with a topic is normally much smaller than the sum of their individual mutual information with the topic. However, the joint mutual information cannot be computed since the statistical data are normally not sufficient. In our experiments we therefore added features until the summed mutual information exceeded the entropy of the topic times a factor (2, 4, 6, 10, 20, 30, 50).

### 4.3 Classification Methods

#### 4.3.1 Selected Classification Methods

We compared the following classification methods in our experiments:

- Perceptrons (standard Perceptron algorithm that produces a linear discrimination function),
- discrete Naive-Bayes ( $P(c|x) = \prod_i P(x_i|c)P(c)$ , if  $x_i$  independent),
- MC4 (decision tree similar to the well-known ID3),
- 3 Nearest-Neighbor (the new vector is assigned to the class of the majority of its nearest neighbors),
- Rocchio Centroid (very simple linear classifier based on the difference between the means of positive and negative examples), and
- Support Vector Machines.

Perceptron, Naive-Bayes, MC4, and Nearest-Neighbor were taken from the MLC++ environment [6] and used with their default parameter settings. The Rocchio Centroid is well-known in document retrieval [7]. For the support vector machine we used the implementation from [8] with linear kernels.

#### 4.3.2 General Comparison of Classification Methods

Perceptrons, Rocchio Centroids, and Support Vector Machines with linear kernels produce linear classifiers (hyperplanes separate the classes in feature vector space). Of course this is a strong restriction in *expressiveness* of the classifiers (a strong *model assumption*). However, experience has shown, that most document classification problems are linearly separable. As far as *learning* is concerned, the Rocchio Centroid does not have much theoretical foundation. *The Perceptron*

*learning rule* seems much more appropriate since it minimizes the number of incorrect classifications. Support Vector Machine learning has the best theoretical foundation. Support Vector Machines are less susceptible to overfitting than other learning methods since the model complexity is independent of the feature space dimension. The model assumptions of Nearest-Neighbor, Naive-Bayes (independence assumption), and decision trees (local decisions) can also be criticized. However, they are different from the assumption of linear separability and therefore a comparison is interesting.

The Nearest-Neighbor does not do any learning at all. It simply stores the training examples. *Learning costs* for the Rocchio Centroid are very low, even with very high dimensions. Perceptron and Support Vector Machine learning is tractable, even for high dimensions, though it is more expensive than Rocchio learning. Naive-Bayes and decision tree learning is tractable only for low dimensions at least in the implementations we used (see also Section 5.2).

*Classification costs* are quite low for all classifiers except for the Nearest-Neighbor where the new feature vector has to be compared to all training examples. The classifiers are all reasonably small except for the Nearest-Neighbor classifier, which grows with the number of training examples.

The decision tree has the best *explanation capabilities*, however, inspecting the other classifiers can also be helpful in order to understand how the Classification is done.

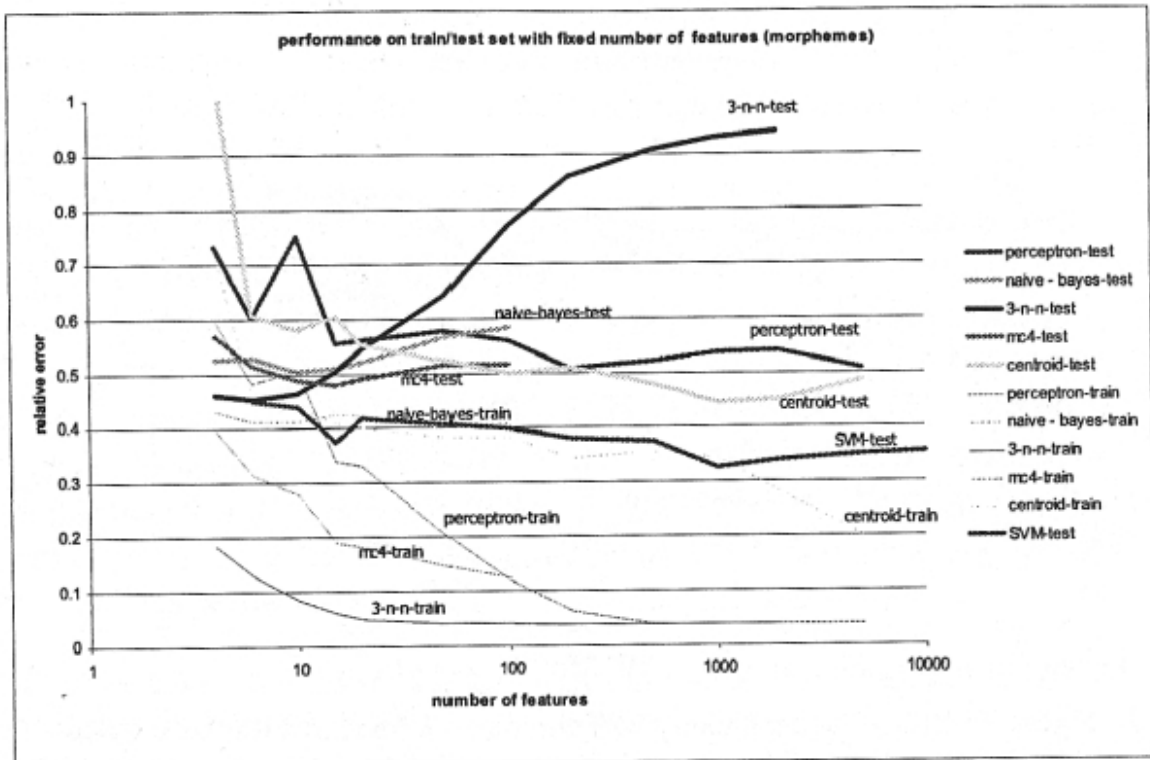
A lot of previous studies have shown that k-Nearest-Neighbor and Rocchio Centroid are both very suitable for document Classification. It can be said that they represent the State of the art in this field. However, recent studies (on Reuters benchmarks) indicate that Support Vector Machines might be superior [9].

## 5. Experimental Comparison

### 5.1 General Remarks on the Experimental Setup

In order to test the generalization capabilities of the different Classification methods, the test corpus (Section 3) was split into two disjoint sets of documents (a *training set* and a *test set*) which both contained approximately 550 documents. The split was stratified in the sense that the original ratio of the topics was preserved. Only the training set was used for feature subset selection and learning.





**Figure 3:** Classification Quality on Training and Test Set with Standard Feature Subset Selection for Morpheme Features

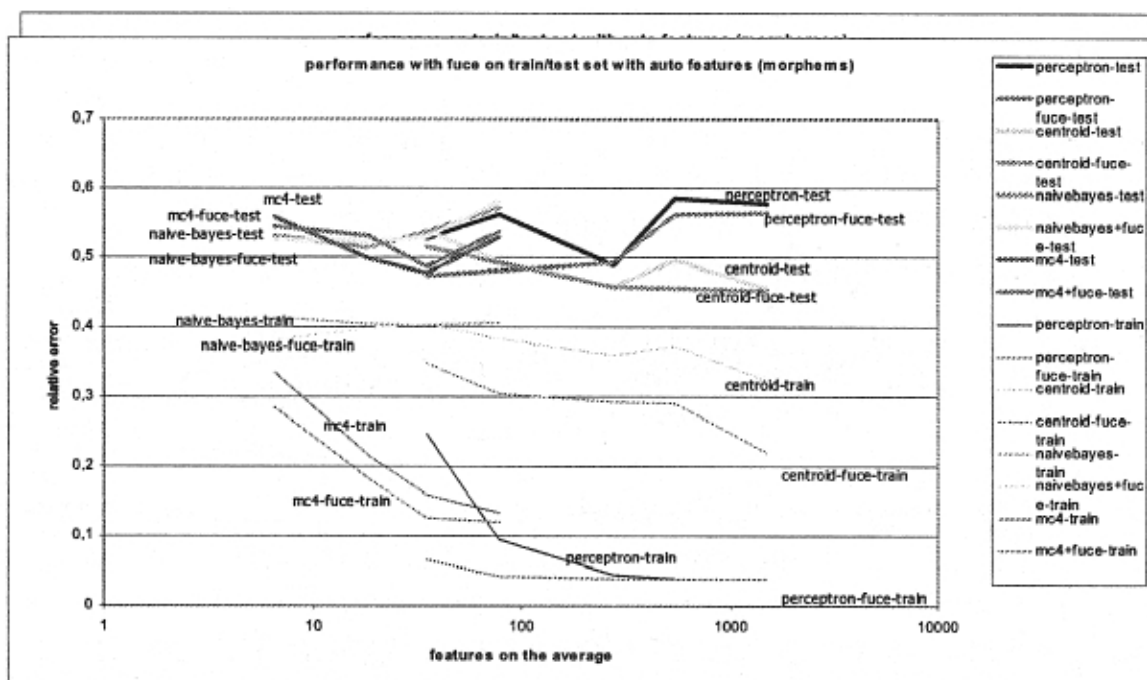
As error measure we computed the *relative error*, that is the number of falsely classified texts divided by the number of texts in the given topic (positive examples), averaged over all 20 topics (macro averaging). Since the number of positive and negative examples differs about one order of magnitude, this seems to be a more meaningful measure than accuracy without the problems of the *F-measure* [10].

In cases where very expensive experiments (like training MC4 with more than 1000 features) aren't expected to show interesting results, we omitted these experiments.

## 5.2 Classification Methods

Figure 3 shows the Classification quality on training and test sets for morpheme features with the Standard feature subset selection method. We use these results to compare the different Classification methods. The Autofeature selection method and n-gram features lead to the same results as far as the comparison of Classification methods is concerned.

Support Vector Machines produce the best results both on test and training sets, are robust with respect to the number of features and very fast at training and at Classification. The only drawback seems to be the complexity of implementation.



**Figure 5:** Classification Quality on Training and Test Set with FuCe Feature Subset Selection for Morpheme Features

The Centroid has the second best test set - Performance with over 1000 features, is very simple and fast and is the only algorithm suitable for retrieval, as it works also if there are only positive examples for a category. The Performance also seems to be relatively robust to the number of features used. A drawback is it's poor Performance on the training set.

The Perceptron does a good job both on test and training set with approximately 500 features. A disadvantage are the relative high training costs. An optimization by making use of sparseness of feature-vectors may help.

It is interesting to see that with very few features, Nearest-Neighbor and the MC4 do perform nearly as good on the test set as the Centroid with 1000 features and far better on the training set. Note that the Nearest-Neighbor even with few features is relative slow (10 seconds Classification time with 10 features, 550 texts, on a Pentium II - the centroid takes only 1 second with 500 features). MC4 is fast to train with few features and especially fast at Classification time.

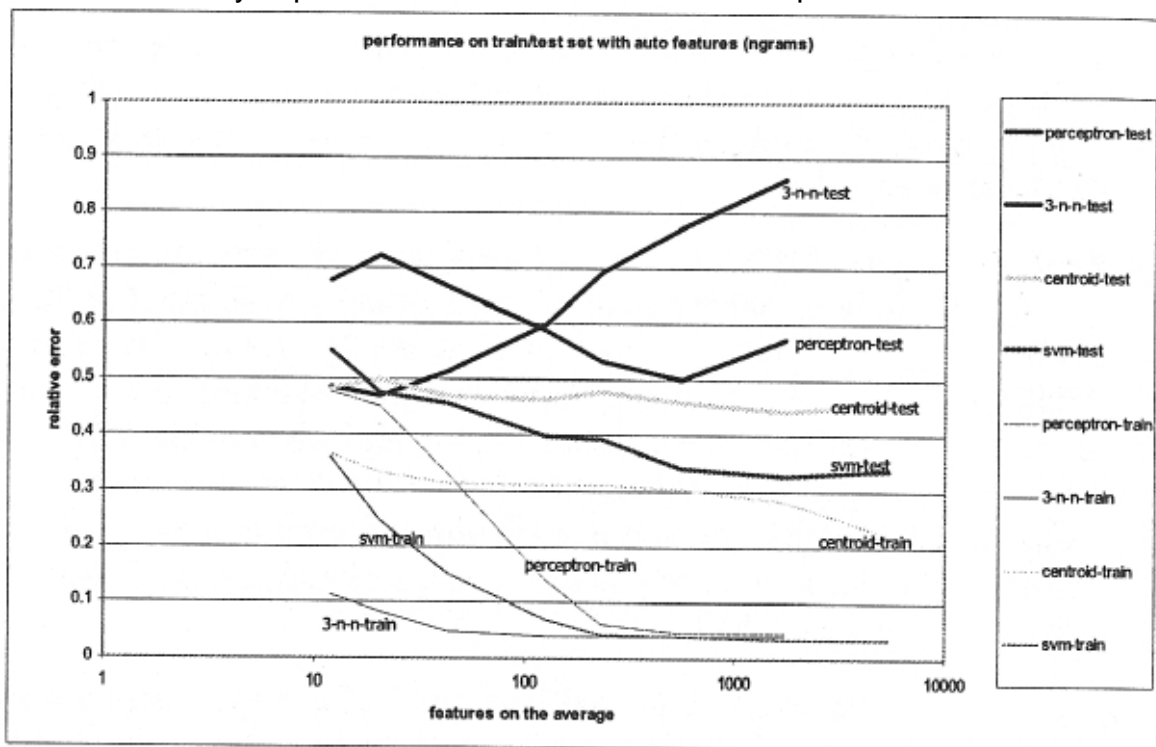
### 5.3 Feature Subset Selection

In Figure 3 we see that Classification quality on the test sets decreases for all classifiers if too many features are used. This happens with very few features for the 3-Nearest-Neighbor, the Naive- Bayes, and MC4. However, it also happens for the Centroid and the Support Vector Machine for more than 1000 features. If we look at the consistently good results on the training sets, the only explanation for the decrease on the test set is overfitting. We think that the effect of overfitting has

not been investigated enough for document Classification so far. Our results indicate that feature subset selection is essential in order to avoid overfitting.

In Figure 4 the results with the Autofeatures subset selection method are shown. The relative errors are plotted against the average number of features that this method computed for factors between 2 and 50. The vector space methods Centroid, Perceptron and Support Vector Machine are not too sensible to changing the numbers of features for high dimensions. Therefore, it does not surprise that the Autofeature method leads to similar results as the Standard subset selection method in high dimensions. On the other hand, with few features we get an improvement for some classifiers (e.g. the Centroid) with the Autofeature method.

Figure 5 shows our results achieved with the FuCe subset selection method. For each classifier we show only results with feature vector dimensions that are suitable for the classifier: higher dimensions for the vector Space methods, lower ones for MC4 and Naïve-Bayes. For the Support Vector Machine we omitted these experiments because we didn't integrate the Software fully in our environment. FuCE almost everywhere improved training set Performance. Test set Performance only improved in some cases for the Perceptron and the Centroid.



**Figure 6:** Classification Quality on Training and Test Set with Autofeatures Subset Selection for 5-Gram Features

#### 5.4 N-grams vs. Morphemes

Figure 6 shows our results achieved with 5-gram features. The Classification quality is more or less comparable to the one achieved with morpheme features (Figure 4). However, Perceptron and Support Vector Machine need less

morpheme features than 5-gram features to achieve an identical Classification quality. Our explanation for this is that morpheme features contain more Information than 5-gram features. Surprisingly, with few features the Centroid with 5-gram features produces better results than with morpheme features.

### 5.5 Significance of Results

To get a feeling for the significance of our results, we computed the probability  $P$  that one classifier is at least the observed number of times better than the other, under the assumption this number is binomially distributed (with probability of success = 0,5). Note that  $P$  can be interpreted as the probability that the performance difference between the two methods happens by chance. Here are the results for some cases from Figure 4: For the Centroid and Perceptron, the Centroid had a better test set Performance from 100 to 1500 features, this corresponds to the 4 rightmost data-points in Figure 4. For the three points with the highest difference,  $P$  was around 0,05. At the third data point from the right,  $P$  was 0,24. For the Centroid and the Support Vector Machine, we addressed the fourth data-point from the right where these two classifiers show the smallest difference. Here,  $P$  is below 0,02. Results with a *t-test* are similar. We conclude that for the given corpus, our results are sufficiently significant.

## 6. Conclusion

In this paper we have presented a thorough evaluation of different approaches for document Classification. We have confirmed recent results about the superiority of Support Vector Machines on a new test set. Furthermore, we have shown that feature subset selection or dimensionality reduction is essential for document Classification not only in order to make learning and Classification tractable, but also in order to avoid overfitting. It is important to see that this also holds for Support Vector Machines. Last but not least, we have shown that linguistic preprocessing not necessarily improves Classification quality.

We are convinced that in the long run linguistic preprocessing like a morphological analysis pays off for document Classification as well as for document retrieval. However, this linguistic preprocessing probably has to be more sophisticated than our simple morphological analysis. A big advantage of linguistic preprocessing compared to n-gram features is that integration of *thesauri*, *concept nets*, and *domain models* becomes possible.

Besides linguistic sophistication, statistics can also help to produce good features. For the future we plan to evaluate different methods for finding topic-relevant *collocations* and *multi-wordphrases*.

Furthermore, we think that feature selection with mutual information can be improved. There should be a better way for approximating the Joint mutual information of a feature set with a topic, than simply summing up the individual mutual information of every feature. Considering feature pairs probably helps a lot. Furthermore, one should include the *term frequency* into the computation of mutual information.

## Acknowledgements

At the time when this work was carried out, the authors worked for iXEC (Executive Information Systems GmbH). Writing this paper took place when two authors had changed to SAIL-LABS (Speech and Artificial Intelligence Labs) where they currently work in the content group.

## References

TREC (Text Retrieval Conference): <http://trec.nist.gov/>

Süddeutsche Zeitung

<http://www.sueddeutsche.de>, <http://www.diz-muenchen.de/cdrom.htm>

TU Wien: Sichere Sinnentsprechende Silbentrennung

<http://www.apm.tuwien.ac.at/research/SiSiSi-dt.html>

Gerard Salton, and Michael J. McGill. 1983. *Introduction to Modern Information Retrieval*. New York: McGraw-Hill.

S. Deerwester, S. T. Dumais, G. W. Furnas, T. Landauer, and R. Harshman. 1990. *Indexing by Latent Semantic Analysis*. Journal of the American Society for Information Science 41:391-407.

Kohavi and Sommerfield, 1996, "*MLC++ Machine Learning library in C++*", available at <http://www.sgi.com/Technology/mlc/>

J. Rocchio. 1971. Relevance Feedback in Information Retrieval. In G. Salton, editor, *The SMART Retrieval System: Experiments in Automated Document Processing*, pages 313-323. Prentice Hall Inc.

T. Joachims. Making large-Scale SVM Learning Practical. *Advances in Kernel Methods - Support Vector Learning*, B. Schölkopf and C. Burges and A. Smola (ed.), MIT-Press, 1999.

T. Joachims. *Text Categorization with Support Vector Machines: Learning with Many Relevant Features*. Proceedings of the European Conference on Machine Learning, Springer, 1998.

D. D. Lewis. 1995. Evaluating and Optimizing Autonomous Text Classification Systems, SIGIR Conference 95.