

IZ-Arbeitsbericht Nr. 11

**Das Verbandsinformationssystem
ELVIRA II
- Projektskizze -**

Jürgen Krause, Thomas Mandl, Maximilian Stempfhuber

April 1997

Das Projekt ELVIRA II wird vom
Bundesministerium für Wirtschaft gefördert.
Aktenzeichen IV C2-003060/22



InformationsZentrum
Sozialwissenschaften

Lennéstraße 30
D-53113 Bonn
Tel.: 0228/2281-0
Fax.: 0228/2281-120
email: krause@bonn.iz-soz.de
 mandl@bonn.iz-soz.de
 stempfhuber@bonn.iz-soz.de
Internet: <http://www.social-science-geis.de>

Dieser Arbeitsbericht entstand aus dem Projektantrag für ELVIRA II. Bei der Erstellung waren Mitarbeiter der Verbände beteiligt.

ISSN: 1431-6943
Herausgeber: Informationszentrum Sozialwissenschaften der Arbeits-
gemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI)
Druck u. Vertrieb: Informationszentrum Sozialwissenschaften, Bonn
Printed in Germany

Das IZ ist Mitglied der Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen e.V. (GESIS),
einer Einrichtung der Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL)

Inhalt

1 Gesamtziel des Vorhabens	5
2 Bezug des Vorhabens zu den förderpolitischen Zielen des Förderprogramms	8
3 Wissenschaftliche und technische Arbeitsziele.....	9
3.1 Multimodale Informationsinhalte und Text-Fakten Mischung.....	11
3.2 Vage Informationsverarbeitung	14
3.3 Verteiltes Objektmodell.....	16
3.4 Intelligente Rechercheunterstützung.....	16
3.5 Dynamische Anpassung	17
4 Erfolgsaussichten	19
5 Stand der Wissenschaft und Technik	20
5.1 Bestehende Datenbanken zur Marktinformation	20
5.1.1 Verband Deutscher Maschinen- und Anlagenbau (VDMA)	20
5.1.2 Hauptverband der Deutschen Bauindustrie (HVB).....	21
5.1.3 Zentralverband der Elektrotechnik- und Elektronikindustrie (ZVEI).....	21
5.2 Multimodalität und Text-Fakten-Mischung beim Information Retrieval.....	22
5.2.1 Integration von Texten und Fakten.....	24
5.2.2 Inhaltserschließung von Texten und Text-Fakten Integration	26
5.2.3 Sonderfall ELVIRA: Deskriptoren zur Beschreibung von Zeitreihen	27
5.2.4 Formen der Kombination von Texten und Fakten in der Wissensbasis	29
5.3 Vagheit in Informationssystemen.....	29
5.3.1 Vagheit im Information Retrieval.....	30
5.3.2 Vagheit im Faktenretrieval.....	31
5.3.3 Mangelnde Benutzerorientierung	33
5.4 Dynamische Anpassung	34
5.5 Theoretisches Konzept bei der Softwareentwicklung.....	36
5.6 GUI-Tools.....	41

6 Projektpartner	42
6.1 Verband des Deutschen Maschinen- und Anlagenbaus (VDMA) ...	42
6.2 Hauptverband der Deutschen Bauindustrie (HVB)	42
6.3 Zentralverband der Elektrotechnik- und Elektronikindustrie (ZVEI)	43
6.4 InformationsZentrum Sozialwissenschaften (IZ) der GESIS	43
6.5 Ausgewählte Unternehmen der beteiligten Verbände	44
7 Bisherige Arbeiten der Antragsteller	44
7.1 Verband Deutscher Maschinen- und Anlagenbau (VDMA)	44
7.2 Hauptverband der Deutschen Bauindustrie (HVB)	45
7.3 Zentralverband der Elektrotechnik und Elektronikindustrie (ZVEI)	45
7.4 Informationszentrum Sozialwissenschaften (IZ) der GESIS	46
8 Literaturverzeichnis	49

1 Gesamtziel des Vorhabens

Ziel des Projektes ist die Entwicklung eines integrierten **Marktinformationssystem**s das eine **Verbandsplattform** für mehrere Industrieverbände darstellt. Die Mitgliedsfirmen sollen über diese Verbandsplattform online Zugriff auf wichtige Marktinformationen haben.

Die Verbände stellen immer wieder fest, daß Unternehmen wichtige Informationen zu wenig nutzen. In einer Zeit zunehmender Internationalisierung der Märkte wird der Produktionsfaktor Information jedoch immer bedeutender für unternehmerischen und gesamtwirtschaftlichen Erfolg oder Mißerfolg. Im Zeichen der fortschreitenden Globalisierung und die Verbreitung moderer Informationstechnologie erhöhen sich die Anforderungen an Vollständigkeit und Aktualität von Informationen. Die Durchdringung mit Informationstechnologie bietet aber auch die Chance, umfassende und aktuelle Informationen zu erhalten und auszuwerten.

In kleinen und mittelständischen Unternehmen (KMU) hat sich moderne Informationstechnologie im Bereich der Informationsbeschaffung aus verschiedenen Gründen erst wenig durchgesetzt. Dazu zählen u.a.:

- Verbreitung von „**Insellösungen**“
- **Mangelnde Problemorientierung** kommerzieller Hosts
- Mangelndes Wissen bezüglich Analyse und Auswertungsmöglichkeiten von Informationen
- Schwer bedienbare **Benutzungsoberflächen**

Hier sind insbesondere die **Verbände** gefordert. Sie sind den Firmen als Lieferanten von qualitativ hochwertigen Informationsdienstleistungen bekannt. Die Kosten hierfür sind in der Regel durch den Verbandsbeitrag abgedeckt. Um die Informationsbedürfnisse der Firmen noch schneller und besser zu erfüllen, ist es wünschenswert, daß die Verbände ihren Mitgliedern Informationen online anbieten und den Zugang dazu einfach gestalten.

Dieses Ziel haben sich der **Verband Deutscher Maschinen- und Anlagenbau (VDMA)**, der **Hauptverband der Deutschen Bauindustrie (HVB)** und der **Zentralverband der Elektrotechnik und Elektronikindustrie (ZVEI)** ge-

setzt. Sie wollen ein umfassendes Marktinformationssystem entwickeln, das möglichst viele Bedürfnisse der Firmen abdeckt und online den Zugang zu aktuellen Daten erlaubt.

Der VDMA hat dafür bereits verschiedene Datenpools geprüft und aufbereitet:

- Nationale und internationale amtliche und VDMA-eigene Zeitreihen
- Statistische Verfahren zur einfachen Analyse der Daten
- Messedaten zur Verbesserung der globalen Marktbearbeitung
- Textuelle Außenwirtschaftsinformationen zur qualitativen Marktbeurteilung
- Firmendatenbank mit Informationen über Mitglieder, globale Wettbewerber und Vertriebsfirmen
- Externe Datenbanken (kommerzielle Anbieter wie FIZ Technik, Internet)

Der ZVEI verfügt mit dem System **ELVIRA** (**E**lektronisches **V**erbands**I**nformations- **R**echerche- und **A**nalyse-System) bereits über ein System, das obige Anforderungen für statistische Zeitreihen erfüllt. ELVIRA wird nach der zweijährigen Projektphase seit Beginn 1997 für alle Unternehmen angeboten und bereits bei ca. 40 Unternehmen eingesetzt. Der VDMA hat ELVIRA an seine Bedürfnisse angepaßt und setzt es im Haus ein. In Kürze soll es auch bei ausgewählten Unternehmen getestet werden. Der HVB paßt seine Daten an das ELVIRA-Schema an und will es dann möglichst bald bei seinen Regionalverbänden installieren.

Diese hohe Akzeptanz zeigt, daß der bei ELVIRA eingeschlagene Weg erfolgreich war. Der **praktische Einsatz** von ELVIRA hat aber auch die Vermutung bestätigt, daß Zeitreihen nur einen Ausschnitt der von Firmen benötigten Zeitreihen darstellen. Neben weiteren Fakten sind insbesondere Texte wichtig. Der weitere praktische Einsatz von ELVIRA in drei Verbänden bietet eine ideale Plattform für empirische Untersuchungen zu konkreten weiteren Informationsbedürfnissen.

Aufbauend auf diesen Erfahrungen soll der in der ersten Phase des Projekts der Entwurf des integrierten, benutzerfreundlichen Marktinformationssystems konkretisiert werden. In der zweiten Phase sollen die Entwürfe dann realisiert und bei den Firmen eingesetzt werden.

Im Rahmen des Projektes ist insbesondere die Problematik der **Integration von Texten und Fakten** zu behandeln. Um Informationsbedürfnisse im wirtschaftlichen Bereich zu erfüllen, kommt dieser Integration eine entscheidende Bedeutung zu. Obwohl die Begriffe Multimedia und Informationsgesellschaft bereits Allgemeinplätze sind, werden bei der Entwicklung von Informationssystemen Benutzerbedürfnisse meist nur unzureichend berücksichtigt. Die kommerzielle Entwicklung produziert vorwiegend „Insellösungen“, die jeweils nur einen Ausschnitt von Informationsbedürfnissen abdecken. Während PCs heute zwar Texte, Fakten und Grafiken übertragen und verarbeiten können, wird oft übersehen, daß diese aus unterschiedlichen Datenbanken stammen und von verschiedenen Programmen übertragen und bearbeitet werden. Aus Benutzersicht ist diese Beschränkung auf einzelne Modalitäten unbefriedigend, da Informationsbedürfnisse häufig durch die **Integration von Texten und Fakten** und der Modellierung ihrer Beziehungen besser erfüllt werden können. Ein integriertes System soll in einer benutzerfreundlichen Oberfläche einen möglichst einheitlichen Zugriff auf heterogene Datenbestände bieten und z.B. bei Nullantworten, iterativen Prozessen oder Suchen nach Meta- oder erklärender Information Anfragen auf zusätzliche Modalitäten erweitern können. In der heutigen Praxis erhalten Unternehmen Informationen aus verschiedensten Quellen, wobei die Wahl des Anbieters oft schon über die Modalität der Information entscheidet (Zeitreihen vom Statistischen Bundesamt, Texte von der Bundesstelle für Außenhandelsinformationen).

Auch mit dem Verband der Chemischen Industrie (VCI) und dem Verband der Deutschen Automobilindustrie (VDA) wurden erste Gespräche über eine mögliche Kooperation geführt; erste Absichtserklärungen liegen vor. Durch die Integration mehrerer Verbände wird das Generalisierungspotential der entstehenden Software stark erhöht. Zusätzlich sollen weitere Möglichkeiten der Kooperation zwischen den Verbänden untersucht werden, mit dem Ziel, geeignete Datenbestände wechselseitig verfügbar zu machen.

2 Bezug des Vorhabens zu den förderpolitischen Zielen des Förderprogramms

Daß die „Verfügbarkeit von hochaktuellen und qualitativ hochwertigen wissenschaftlichen und technischen Informationen für Nutzer in Industrie, Wissenschaft und Staat“ (BMBF 1996:1) der Schlüssel für effizientes Forschungs- und Wirtschaftshandeln ist, kann heute als allgemeiner Grundkonsens zwischen Politik, Forschung und Wirtschaft angesehen werden. Dies belegt auch das Programm der Bundesregierung „Information als Rohstoff für Innovationen“ (INFR). In seinem Kontext gehört das Projekt Marktinformationssystem in den Bereich der in Kap. 6 genannten „Literatur- und Faktendatenbanken“ (INFR:29), speziell zu den auf S.49 als förderwürdig genannten „Datenbanken, die ... Wirtschaftsinformationen (z.B. ... Marktinformationen) bereithalten, die über wirtschaftliche Gegebenheiten in Deutschland, Europa und der Welt unterrichten).“

In diesem Rahmen will das BMWi

„in enger Zusammenarbeit mit der Wirtschaft Konzepte und Voraussetzungen schaffen, damit die deutschen Unternehmen ...

- schnellen Zugriff auf ein umfassendes Informationsangebot haben ... soweit möglich in deutscher Sprache,
- im Rahmen rechnerintegrierter Informationsverarbeitung dem Nutzer an seinem Arbeitsplatz die Möglichkeit geben, die Information unmittelbar im Planungs-, Konstruktions- und Fertigungsprozeß eines Produkts zu verarbeiten.“ (INFR:34)

Als einen Schwerpunkt der Förderung des BMWi nennt INFR:34 zudem den „modellhaften Einsatz intelligenter Nutzeroberflächen ... zur Unterstützung der Endnutzer“ (cf. auch S.53).

Damit deckt sich das Projektziel mit mehreren zentralen Kernbereichen des Förderprogramms „Information als Rohstoff für Innovationen“.

3 Wissenschaftliche und technische Arbeitsziele

Nach Analyse der Benutzerbedürfnisse soll das System von VDMA und IZ entwickelt werden. Dabei wird auf das ZVEI-Verbandsinformationssystem ELVIRA zurückgegriffen, das die online-Recherche von Zeitreihen ermöglicht. Es basiert auf dem WOB-Modell, das sich bei der Entwicklung von ELVIRA als Plattform für eine benutzerfreundliche Systemgestaltung bewährt hat. Das WOB-Modell tritt als Mittelmodell für die Gestaltung von BOF zwischen kognitionspsychologische Erkenntnisse, die sich nicht operationalisieren lassen, und detaillierte Styleguides, die in sich widersprüchlich sind. Es ist strikt objektorientiert und ermöglicht dadurch konzeptuelle und programmtechnische Erweiterungen des bestehenden Systems. Es wurde bereits in mehreren kommerziellen Anwendungen eingesetzt und wird am IZ Sozialwissenschaften weiterentwickelt. Bei der Realisierung benutzer- und problemorientierter Systeme müssen erfahrungsgemäß die technischen Möglichkeiten der Programmierung grafischer BOF über die in GUI-Tools enthaltenen Standardmittel hinausgehend ausgereizt werden.

Ein wesentliches Problem bei Entwicklung des Marktinformationssystems unter Einbeziehung von ELVIRA-Komponenten ist, daß der bisherige Entwurf von drei Deskriptor-Facetten ausgeht, die am Bildschirm angezeigt werden (zwei davon geöffnet). Eine Erweiterung um zusätzliche Facetten zur Beschreibung anders strukturierter Datenbestände macht neue Überlegungen zur Gestaltung der Oberfläche notwendig, die nicht trivial sind. Gerade die Einschränkung auf nur wenige Facetten wurde ausgenutzt, um eine effiziente Zwei-Bildschirm Bedienung zu gestalten. Von diesen Überlegungen ist jedoch die theoretische Basis, das WOB-Modell nicht betroffen. Deshalb ist davon auszugehen, daß ein Teil der prinzipiellen Vorgehensweise und die einzelnen Werkzeuge von ELVIRA zumindest modifiziert in das Marktinformationssystem übernommen werden können. Das System wird um zusätzliche Datentypen und -bestände erweitert und parallel für VDMA, ZVEI und HVB entwickelt. Neben der Text-Fakten-Integration finden vor allem Probleme der Vagheit und der dynamischen Anpassung der Einzelkomponenten Beachtung, die in Mehrwertkomponenten (z.B. Thesaurus, Erklärungskomponenten) bearbeitet werden.

Ab dem zweiten Projektjahr soll es bei Mitgliedsfirmen installiert werden. Dort kann es sowohl für die routinemäßige Informationsbeschaffung als auch für

problemorientiertes Recherchieren in wenig bekannten Datenbeständen benutzt werden.

Während der ersten Projektphase wird der ZVEI das System ELVIRA projektbegleitend verstärkt in den Markt einführen. Es ist damit zu rechnen, daß der ZVEI die derzeitige Nutzerzahl von ca. 40 Firmen weiter steigert. Damit entsteht eine große Basis für empirische Untersuchungen. HVB und VDMA führen ELVIRA verbandsintern sowie bei jeweils drei Mitgliedsfirmen ein. Alle drei Verbände übernehmen hierbei neben der Auswahl der Pilotfirmen auch die anfallenden Hotline-Tätigkeiten (technische Beratung, Einweisung und Betreuung). Das IZ wird einen ausgewählten Kreis von Nutzern kontaktieren, um Möglichkeiten der Weiterentwicklung des Faktenzugangs sowie eventuelle Schwachstellen in ELVIRA in einem praxisnahen Umfeld zu identifizieren.

Parallel hierzu sichten die Verbände alle von ihnen in das zukünftige Marktinformationssystem einzubringenden Daten und strukturieren diese nach Datenart, Inhalt und möglichen Verbindungen zu anderen Datenbeständen. Zusätzlich werden externe Datenquellen auf Relevanz, Nutzungs- und Integrationsmöglichkeiten geprüft.

Darauf aufbauend analysiert das IZ in enger Zusammenarbeit mit den Fachexperten in den Firmen und Verbänden die Benutzerbedürfnisse an das zu entwickelnde Marktinformationssystem, entwirft ein Konzept für die Integration der Datenbestände und der von den Nutzern gewünschten Funktionalität und fixiert diese ähnlich einem Pflichtenheft. Zusätzlich werden Standardsoftwareprodukte (z.B. Indexierungs- und Retrievalsysteme) gesichtet und schwerpunktmäßig für ihren Einsatz im Projekt getestet.

Aus diesen Vorarbeiten resultieren mehrere durch das IZ erstellte Prototypen, die unterschiedliche Aspekte des geplanten Marktinformationssystems implementieren. So liegt im ersten Projektjahr der Schwerpunkt der Entwicklung auf dem Gesamtsystem. Die mit ELVIRA und seiner Evaluierung gewonnenen Erkenntnisse werden für das Marktinformationssystem genutzt. Dabei wird verstärkt die Problematik der Integration heterogener Datenbestände (z.B. Texte und Fakten) bearbeitet. Im zweiten Projektjahr werden vorwiegend die oben erwähnten Mehrwertkomponenten entworfen, die die als grundlegend identifizierte Funktionalität erweitern und zusätzlichen Aspekten der informationellen Prozesse im Bereich der Marktforschung Rechnung tragen.

Die einzelnen Abschnitte werden im Sinne einer zyklischen Entwicklung so mit Benutzertests gekoppelt, daß die empirischen Ergebnisse aus den Tests sofort bei der weiteren Vorgehensweise berücksichtigt werden können.

Die Systemerweiterungen werden so in einzelne, aufeinander aufbauende Stufen verteilt, daß den Verbänden der Zeitpunkt der Übernahme innerhalb gewisser Grenzen freigestellt ist. So wird sichergestellt, daß jeder der Verbände unter Berücksichtigung seiner Ressourcen die neuen Komponenten bei den eigenen Mitgliedsfirmen einführen kann.

Ausgehend von den erstellten Prototypen stellen die Verbände eine kommerzielle Umsetzung der Systemkomponenten sicher. Das IZ wirkt hierbei unterstützend mit.

Auf die wesentlichen Probleme bei der prototypischen Entwicklung des Marktinformationssystems wird in den folgenden Abschnitten kurz eingegangen.

3.1 Multimodale Informationsinhalte und Text-Fakten Mischung

Ein umfassendes Informationssystem muß neben numerischen Faktendaten (Zeitreihen) auch strukturierte Textdaten (Messdaten), unstrukturierte Texte (Berichte, Länderreports) und Grafiken (Abbildungen in Texten, vorbereitete Liniengrafiken mit interessanten Entwicklungen) enthalten. Benutzer sollen Anfragen sowohl auf einzelne Datentypen beschränken als auch über mehrere Modalitäten hinweg anfragen können. Somit ist es notwendig, neben gemeinsamen Repräsentationen auch mögliche Wechselbeziehungen zwischen den Modalitäten, die im Anwendungsfall Wirtschaftsinformation sinnvoll sind, zu definieren. Dabei sind folgende Fragen interessant:

- Welche Medientypen ergänzen sich bei welchen Informationsbedürfnissen gut? Relativ einfache Fälle sind etwa eine Liniengrafik zu einer Tabelle oder eine Grafik zu einem Bericht.
- Welche Modalitätsmischungen sind im Anwendungsfall sinnvoll und wie können diese für die Auswahl der Anfragemodalität ausgenutzt werden? So können Bedingungen für Liniengrafiken sehr effektiv grafisch formuliert werden.

- Welche Elemente der Oberfläche werden dem Benutzer in speziellen Retrievalsituationen als Unterstützung angeboten? (z.B. erklärende Texte zu Besonderheiten in Zeitreihen)

In Sachen gegenseitiger Ergänzung von Informationsmodalitäten, sind bei dem derzeit angestrebten Datenumfang des Marktinformationssystems Fragestellungen aus folgender Liste denkbar. Welche Informationsbedürfnisse in der Praxis besonders relevant sind, soll in der ersten Projektphase geklärt und konkretisiert werden.

- Der ZVEI veröffentlicht regelmäßig Quartalsberichte, die aktuelle Entwicklungen einzelner Branchen darstellen. Diese Berichte bestehen zum größten Teil aus Text, der mit Grafiken und statistischen Daten angereichert ist. Das enthaltene Zahlenmaterial wird i.d.R. nicht direkt der Datenbank entnommen, sondern ist das Ergebnis von Berechnungen auf Basis der in der Datenbank enthaltenen Fakten. Für eigene Analysen benötigt der Informationsnutzer in der Praxis die mit dem Thema des Berichts zusammenhängenden Faktendaten (Zeitreihen).
- Entscheidungen über das Engagement einer Unternehmung auf einem ausländischen Markt werden u.a. auf der Basis von Betrachtungen des Marktvolumens eines Produktes in einem bestimmten Land getroffen. Wird das Marktvolumen positiv beurteilt, ist zu überprüfen, welche Zollbestimmungen zu beachten sind oder auf welchen Messen das Produkt präsentiert werden kann. In beiden Fällen können ausgehend von den Fakten (Zeitreihen) im Dokumentenbestand die zugehörigen Texte recherchiert werden.
- In halbjährlichen Veröffentlichungen, z.B. dem "Konjunkturbericht Westeuropa" des ZVEI, wird in einem allgemeinen Teil die europäische und in länderspezifischen Teilen die Konjunkturentwicklung in einzelnen Ländern aufgezeigt. Um die zugrunde liegenden Fakten mit firmeninternen Daten zu vergleichen, benötigt der Marktforscher schnellen Zugriff auf die Zeitreihen. Abhängig vom betrachteten Textabschnitt können die Fakten aus der Datenbank gewonnen werden.
- Verschiedene Produkte unterliegen beim Handel mit einzelnen Ländern einer Reihe von Exportbeschränkungen. Angaben zu Exportbeschränkungen können in der Textdatenbank recherchiert werden. Bestehenden für das Produkt keine hinderlichen Exportbeschränkungen, kann mittels einer Analyse des Marktvolumens festgestellt werden, ob der Export nach betriebswirtschaftlichen Kriterien lohnend ist. Ausgehend vom Text zu den Exportbeschrän-

kungen können anhand der Land- und Produktinformationen die zugehörigen Zeitreihen aus der Faktendatenbank abgerufen werden.

- Ein Unternehmen will prüfen, ob seine Produkte in „Indonesien“ Absatzchancen haben. Durch die Anfrage an die Datenbank erhält der Marktforscher
 - die Export- und Importströme sowie die Produktionsentwicklung aus der Statistikdatenbank,
 - die allgemeinen wirtschaftlichen Informationen über seine Branche sowie relevante Presseberichte aus der Textdatenbank,
 - die in Frage kommenden Messen und Ausstellungen aus der Messedatenbank sowie
 - die Kontaktadressen von AHK, Botschaft, Herstellern, Importeuren oder Händlern aus der Adressdatenbank.

Im Gegensatz zur bisherigen Praxis muß der Benutzer nur eine Recherche über eine einheitliche Rechercheoberfläche für mehrere Datenbanken durchführen. Das System liefert ihm relevante Informationen aus allen angebotenen Datenbanken unabhängig davon, ob dem Benutzer die Datenbanken bekannt sind. Das in den Datenbanken vorgehaltene Wissen wird unabhängig von der Kompetenz des Benutzers verfügbar gemacht.

- Ein Nachfrager nach Informationen über das Bauvolumen findet im System eine Zeitreihe über die langfristige Entwicklung in den einzelnen Bausparten. Bei der Betrachtung als Liniengrafik stellt er eine starke Veränderung im Verlauf des Wohnungsbaus fest, den er sich alleine aus den Daten nicht erklären kann. Er sucht daher unter dem Thema „Einflußfaktoren auf den Wohnungsbau“ nun Textdokumente, die ihm z.B. steuerliche Rahmenbedingungen, demographische Entwicklungen etc. aufzeigen. In einem Text findet er z.B. eine Bemerkung über Fertigstellungsziffern von einzelnen Gebäudetypen. Dieser Hinweis führt ihn bei seiner weiteren Recherche wieder zurück zu Faktendaten.
- Zu den Themenbereichen, die in der Faktendatenbank recherchierbar sind, existieren parallel fast immer vielfältige Textinformationen über Tarifpolitik,

Lohnkosten, Insolvenzen, Ausbildungssituation, private Finanzierung öffentlicher Struktureinrichtungen, Baustoffe und vielem mehr.

Informationelle Prozesse auf Computern verlaufen in der Regel iterativ, d.h. zu Beginn der Anfrage stehen noch nicht alle Parameter fest, sondern sie werden aufgrund von Zwischenergebnissen verändert. Benutzer setzen also mehrstufige Problemlösungsstrategien ein, in deren Verlauf sich die ursprüngliche Problemlage wandelt. Dies ist eine besondere Herausforderung für ein multimodales System.

3.2 Vage Informationsverarbeitung

In der menschlichen Kognition werden viele Konzepte vage verarbeitet. Im Information Retrieval herrscht aufgrund der kognitiven Sichtweise auf den Suchprozeß zunehmend die Ansicht, daß der Abgleich zwischen Informationsbedürfnis und Retrievalergebnis als vager Prozeß betrachtet werden muß. Ein Teilaspekt der Vagheit im Suchprozeß sind Fälle, in denen der Benutzer ein vages Informationsbedürfnis hat oder es nur vage formulieren kann. Folgende Typen von Anfragen sind in diesem Zusammenhang denkbar:

- Ein Unternehmer recherchiert mit vagen Bedingungen, um eine gewisse Toleranz zulassen zu können. Er sucht z.B. nach *großen* Messen in einer bestimmten Region oder nach Ländern mit *niedriger* Zollrate. In diesen Fällen können vage Quantoren (z.B. *mittlere* Firmen, *kleine* Märkte, *großer* Export) für die Recherche verwendet werden.
- Ein Nutzer recherchiert in einem ihm nicht sehr vertrauten Datenbestand (z.B. potentielle Abnehmerbranchen) und muß sich erst einen Überblick über die wichtigsten Entwicklungen verschaffen. Er sucht dazu die Produkte mit dem *größten* Umsatz in diesem Industriezweig ohne daß er im Detail weiß, was *großer* Umsatz hier konkret numerisch bedeutet. Auch für Überblickssuchen können vage Quantoren eingesetzt werden.
- Auch Veränderungsdaten und damit dynamische Prozesse können über vage Quantoren beschrieben werden (*schnell* wachsende Märkte). Dies erschwert die Modellierung.

- Um Prognosen stützen zu können sucht ein Marktforscher nach einer Zeitreihe, die einen *ähnlichen* konjunkturellen Verlauf hat, wie die zu prognostizierende Reihe. Zusätzlich sollte die Reihe einen Vorlauf von zwei bis drei Monaten haben. Dazu ist ein Ähnlichkeitsmodell notwendig.
- Zu einer unerwarteten Knick in einer Zeitreihe sucht ein Nutzer *erklärende* Dokumente. Dabei kann es sich um Texte zu diesen Sachverhalten oder um andere Zeitreihen mit ähnlichem Knick handeln. Die Vagheit ist in diesem Fall besonders deutlich und kann sich auch auf die Art der Dokumente beziehen.

Durch die Integration von Vagheitskomponenten, die derartige Anfragen ermöglichen, und deren Einsatz in den geeigneten Situationen kann die Interaktion „natürlicher“ und effizienter erfolgen. Dazu müssen geeignete Modellierungstechniken gefunden werden. Für deren Realisierung muß das nötige Wissen mit den Domänenexperten erarbeitet werden. Das Marktinformationssystem sollte dann über intelligente Verfahren verfügen, um Vagheiten aufzulösen und in Datenbankabfragen umzusetzen.

Auch auf der Datenseite tritt Unsicherheit in verschiedenen Formen auf. Die Qualität und Zuverlässigkeit ist gerade bei wirtschaftlichen Informationen sehr wichtig und spielte schon bei der Entwicklung des ZVEI-Verbandsinformationssystems ELVIRA eine Rolle. Daneben bringt die Repräsentation der Daten eine gewisse Unsicherheit mit sich. Besonders bei textuellen Daten, bei denen meist einige wenige Deskriptoren einen Text repräsentieren, ist einsichtig, daß der Repräsentationsmechanismus mit einem Informationsverlust und damit mit Unsicherheit behaftet ist.

Treten nun mehrere, mit Unsicherheit behaftete Modi zusammen in einem Informationssystem auf, verschärfen sich diese Probleme. Gerade die oben genannten Abhängigkeiten zwischen Einheiten verschiedener Modalitäten werden oft nur über unsicheres Wissen modelliert werden können, wobei auch die Kontextproblematik eine Rolle spielt (d.h. unsicheres Wissen kann nur innerhalb eines Kontextes adäquat interpretiert werden). So bedeutet ein *hoher* Lagerbestand in der Verbrauchsgüterindustrie etwas anderes als in der Investitionsgüterindustrie.

3.3 Verteiltes Objektmodell

Im Sinne eines homogenen Gesamtkonzeptes zur Integration von numerischen Fakten, Texten und Grafiken scheint es essentiell, die Wechselwirkungen innerhalb des Systems zu identifizieren, die durch unterschiedliche Datenarten und Modalitäten verursacht werden. Darauf basierend soll ein globales Kommunikationskonzept entwickelt werden, das - zunächst auf den Anwendungsfall Marktinformations-system begrenzt - die Inter- und Intra-Beziehungen zwischen Datenarten und Modalitäten formalisiert. Parallel dazu wird ein Metamodell erstellt, das alle im System verfügbaren Datenbanken und die darin gespeicherten Datenbestände implementationsunabhängig beschreibt. Das Recherchesystem soll alle in diesem Modell beschreibbaren Datenbanken möglichst ohne zusätzlichen Änderungsaufwand unterstützen.

Als Ergebnis wird ein (verteiltes) Objektmodell angestrebt, das eine kontext-sensitive und dynamische Anpassung des Retrievalsystems an die aktuelle Recherchesituation erlaubt. Hierbei soll z.B. die Art der recherchierten Daten den Zustand aller Recherchewerkzeuge beeinflussen bzw. dem Benutzer alternative Recherchewege anbieten. Das verwendete Recherchewerkzeug bewirkt anhand der vom Benutzer formulierten Anfrage eine intelligente Parametrisierung anderer Werkzeuge, so daß der Benutzer beim Wechsel des Werkzeugs optimal innerhalb seiner aktuellen Strategie unterstützt wird. Hierdurch werden Eingaberedundanzen vermieden, die Recherchewege verkürzt, die Anzahl der Interaktionsschritte verringert und der auf der Benutzungsoberfläche verfügbare Platz optimal genutzt. Zusätzlich sollen Möglichkeiten untersucht werden, die Benutzungsoberfläche in Abhängigkeit von den jeweils verfügbaren Datenquellen anzupassen. Dies kann durch den Einsatz eines offenen, plattformunabhängigen Standards zur Kommunikation und Verteilungen von Objekten (z.B. CORBA) erreicht werden, der den Einsatz des Systems in lokalen Netzwerken sowie im Internet ermöglicht. Die dynamische Integration verteilter Datenbestände würde somit nicht nur verbandsübergreifend möglich, auch die Systemnutzer könnten eigene Datenbestände nahtlos in das System integrieren.

3.4 Intelligente Rechercheunterstützung

Menschen bauen aufgrund ihres gemeinsamen Vorverständnisses Verständigungsebenen auf, die es ihnen ermöglichen mit einem Minimum an intellektu-

ellem Aufwand zu kommunizieren. Ein Informationssystem muß versuchen, zumindest einige dieser Ebenen in der jeweiligen Anwendungsdomäne zu modellieren. Die dazu notwendigen intelligenten Komponenten sind Teil eines anwenderorientierten Designs und erleichtern den Umgang mit dem Werkzeug Computer.

Auf der BOF können Zugänge zu Informationen verschiedener Modalität durch verschiedene Objekte repräsentiert sein, so daß hier die Wechselbeziehungen zwischen Objekten modelliert werden müssen. Intelligente Verfahren müssen entscheiden, wann bestimmte Eingaben oder Ergebnisse in andere Objekte direkt oder nach Deduktion übernommen werden, in welchen Retrievalsituationen automatische Anpassungsvorgänge ausgelöst werden und unter welchen Bedingungen bestimmte Objekte dem Benutzer zur Verfügung gestellt werden.

Dafür muß das System über intelligente Komponenten verfügen, die aus Anfragen auf das Informationsbedürfnis des Benutzers schließen können und solche Anfragen bei Bedarf für andere Modalitäten umsetzen. Das geplante Marktinformationssystem soll Firmen auch den Zugriff auf externe Datenbanken erlauben, was das Datenvolumen potenziert. Von Interesse sind hierbei kommerzielle Hosts, Behörden (z.B. textuelle Informationen der BFAI), Wirtschaftsforschungsinstitute (z.B. DIW, IFO) bis hin zur ungeordneten Fülle von Informationen auf dem Internet. Das System muß aus dieser Menge dem Benutzer in verschiedensten Kontexten einen sinnvollen Ausschnitt anbieten und auch bei der Suche in diesen heterogenen Datenbeständen unterstützen. z.B. könnte eine intelligente Komponente ausgehend von der Art der recherchierten Zeitreihen geeignete statistische Analyseverfahren vorschlagen.

3.5 Dynamische Anpassung

Versprach der Wechsel von zeichenbasierten („Textmodus“, DOS-Programme) zu grafischen Benutzungsoberflächen (Macintosh, OS/2 Workplace Shell, MS-Windows, X-Windows) durch den nicht unerheblichen Zuwachs an Gestaltungsmöglichkeiten und nutzbarer Bildschirmfläche zunächst eine qualitative Verbesserung der verfügbaren Software, so zeigt sich in heute am Markt verfügbare Anwendungen unverkennbar die Tendenz zum Negativen. Der Grund dafür ist neben der mangelhaften - und zugegeben oft schwierigen - Umsetzung software-ergonomischer Gestaltungsregeln (z.B. Stylegui-

des) vor allem im Einsatz inadäquater syntaktischer Mittel zur Oberflächengestaltung zu suchen. Mit zunehmender Komplexität der Anwendung wird zunehmend auf eine starke Hierarchisierung der Benutzungsoberfläche (BOF) zurückgegriffen, um zum einen den Informationsgehalt der einzelnen Bildschirme auf ein für den Nutzer kognitiv verarbeitbares Maß zu reduzieren, zum anderen um durch starre Dialogleitlinien den Nutzer zu einer syntaktisch korrekten Programmbedienung zu „zwingen“. Beide Lösungsansätze lassen jedoch oft die tatsächlichen kognitiven Prozesse beim Anwender außer Acht und verhindern im Kontext der Informationssuche darüber hinaus u.U. die aus Benutzersicht „natürliche“ Formulierung des Informationsbedürfnisses als Suchanfrage in einem Recherchesystem. Besonders in langen Dialogfolgen stellt sich regelmäßig das Problem, die in den vorangegangenen Schritten getätigten Eingaben in späteren Schritten auf dem Bildschirm zu präsentieren (Statusanzeige) und dem Benutzer die Möglichkeit zu geben, frühere Setzungen jederzeit (möglichst ohne Rücksprung in der Dialoggeschichte) zu korrigieren.

Wie sich in Benutzerstudien und -interviews während der gesamten Projektlaufzeit im Projekt ELVIRA zeigte, kann durch die Ausnutzung domänenspezifischer Restriktionen und den Einsatz innovativer Gestaltungsmittel die Nutzbarkeit von Software wesentlich verbessert und die Zufriedenheit der Nutzer gesteigert werden. Von besonderer Bedeutung ist hierbei der von Recherchesystemen oft außer Acht gelassene Aspekt, daß das gefundene Dokument nicht das Endergebnis eines informationellen Prozesses darstellt, sondern i.d.R. zusammen mit anderen Dokumenten erst zum Endergebnis (z.B. vergleichende grafische Darstellung, errechneter Kennwert) weiterverarbeitet wird.

Durch die Integration von Texten, Grafiken und Fakten in einer aus Benutzersicht homogenen Datenbank stellen sich neue Probleme für die Präsentation und Interaktion mit den gefundenen Dokumenten. Anregungen können aus einer Reihe kommerziell verfügbarer, objektorientierter Benutzungsoberflächen entnommen werden. Oftmals sind hier jedoch nur triviale Interaktionen zwischen Standardobjekten realisiert (z.B. Datei, Drucker, Papierkorb). Unter dem Gesichtspunkt multimodaler Rechercheoberflächen sind jedoch Objekte und Funktionalitäten zu fordern, die u.a.

- unterschiedliche Datenarten anzeigen (Texte, Zeitreihen, Grafiken),

- unterschiedliche Dokumente zu einem neuen Dokument zusammenfügen (z.B. Texte und Grafiken),
- Datenarten ineinander überführen (z.B. Zahlen in Grafiken),
- eine gemeinsame Repräsentation unterschiedlicher Datenarten erlauben und
- eine Suchintention über verschiedene Datenarten hinweg repräsentieren.

Von besonderem Interesse ist hierbei die dynamische, kontextabhängige Anpassung der Objekte untereinander. Vor allem die Koppelung spezialisierter Suchmodalitäten für die unterschiedlichen Datenarten spielt eine wichtige Rolle, um dem Benutzer in jeder Situation eine homogene Repräsentation des aktuellen Recherchkontextes zu bieten.

4 Erfolgsaussichten

Die Erfolgsaussichten für das Projekt sind gut. Aufgrund der günstigen Ausgangslage ist zu erwarten, daß ein lauffähiger Prototyp entsteht, der die gestellten Anforderungen erfüllt und kommerziell umgesetzt werden kann.

Da die beteiligten Verbände bereits im Vorfeld begonnen haben, einen Teil der technischen Voraussetzungen für die testweise Implementierung einiger Datenbestände zu schaffen, sind die Realisierungsaussichten durchweg positiv zu beurteilen. Zudem tragen die Verbände durch den Einsatz bzw. Test von ELVIRA dazu bei, empirische Grundlagen für die Weiterentwicklung zu sammeln.

Die Projektgruppe am IZ hat bereits vor der Entwicklung von ELVIRA Erfahrungen in den Bereichen Softwareentwicklung, Oberflächendesign und der Modellierung informationeller Prozesse gesammelt. Mit ELVIRA wurde ein benutzerorientiertes Informationssystem im Bereich der Wirtschaftsinformation entwickelt. Das Wissen über informationelle Prozesse im Umfeld ökonomisch motivierter Fragestellungen und die software-technische Erfahrung bei der Entwicklung von grafischen Benutzungsoberflächen erlauben im Projekt Marktinformationssystem eine effiziente Projektarbeit ohne langwierige Einarbeitung in diese Bereiche. Durch die Einbettung am IZ, wo weitere Projekte in den Bereichen Softwareentwicklung, grafische Benutzungsoberflächen und Inhaltserschließung durchgeführt werden, und die Anbindung an den Fachbereich

Informatik an der Universität Koblenz-Landau sind starke Synergieeffekte zu erwarten.

Viele der im Projekt angestrebten Teilziele sind heute rein technisch lösbar, wie etwa die Integration von Texten und Fakten. Sie müssen aber benutzer- und aufgabenorientiert in einem Marktinformationssystem umgesetzt werden, um die Informationsbedürfnisse der Nutzer optimal zu erfüllen.

Die Integration von mindestens drei Verbänden erlaubt eine sehr generelle Aufgabenanalyse und Benutzerorientierung. Die Expertise der Informationsvermittler von drei Verbänden kann somit bei der Systementwicklung berücksichtigt werden. Durch den umfassenden Ansatz und die bereits jetzt durch VDMA, HVB und ZVEI sichergestellten Informationsinhalte wird ein für die Unternehmen sehr attraktives System entstehen, das leicht und mit kontrollierbaren Kosten nutzbar ist.

Das System ELVIRA geht in das Marktinformationssystem mit ein, so daß für die Recherche von statistischen Daten bereits eine Lösung vorliegt. Da der VDMA sich für die gleichen Entwicklungswerkzeuge entschieden hat, mit denen ELVIRA realisiert wurde, kann der vorliegende Programmcode voll übernommen werden.

5 Stand der Wissenschaft und Technik

5.1 Bestehende Datenbanken zur Marktinformation

5.1.1 Verband Deutscher Maschinen- und Anlagenbau (VDMA)

Statistikdatenbank

Ausgangspunkt für das Projekt ist die durch ELVIRA erschlossene Statistikdatenbank. Sie enthält eine Vielzahl relevanter statistischer Daten für den Maschinenbau aus Deutschland, der EU und den westlichen Industriestaaten. Diese Daten werden gegenwärtig für ELVIRA aufbereitet.

Außenwirtschaftsdatenbank

Im VDMA wurde in den letzten zwei Jahren eine Außenwirtschaftsdatenbank entwickelt. Sie enthält neben den Informationen der BfAI (Bundesstelle für Außenhandelsinformationen) auch Informationen aus anderen Quellen (z.B. Reuters Business Briefing) sowie im VDMA erzeugte Textdokumente.

Messedatenbank

Der VDMA hat die Daten des AUMA (Aussteller- und Messeausschuß der deutschen Wirtschaft) zu Messen in aller Welt in eine Messedatenbank eingespielt. Durch Ergänzungen um eigene branchenspezifische Informationen besteht hier eine wertvolle Informationsquelle um die globalen Vertriebsaktivitäten möglichst optimal zu steuern.

Adressdatenbank

In der Adressdatenbank werden relevante Kontakt-, Hersteller-, Händler- und Endkundenadressen aus dem In- und Ausland gespeichert. Besonders bei Auslandsaktivitäten sind aktuelle Adressen unentbehrlich.

5.1.2 Hauptverband der Deutschen Bauindustrie (HVB)

Der Hauptverband hat 1995 eine Rundschreibendatenbank mit der Software Lotus Notes installiert. Diese Datenbank hat reinen internen Nutzen, indem sie den regionalen Mitgliedsverbänden des HVB die zahlreichen Rundschreiben mit Volltextrecherche und direkter Weiterverarbeitungsmöglichkeit auf den Arbeitscomputer liefert. Über weitere Datenbanken, die den Mitgliedsunternehmen (nicht den Mitgliedsverbänden) Zugang zu den gewünschten Informationen bietet, verfügt der HVB bis heute nicht.

5.1.3 Zentralverband der Elektrotechnik- und Elektronikindustrie (ZVEI)

Der ZVEI hat in ELVIRA bereits wichtige statistische Informationen für den online Zugang aufbereitet. Es handelt sich um die vierteljährliche Produktionsstatistik, den deutschen und internationalen Außenhandel und die ZVEI - Konjunkturdaten. Daneben stehen in ELVIRA die Ergebnisse des IFO-

Konjunktur- und Investitionstests sowie die Volkswirtschaftliche Gesamtrechnung des DIW zur Verfügung. Dieses Angebot soll nach Anforderungen der Nutzer erweitert werden. So sollen Daten aus Liefer- und Abnehmerbranchen der Elektroindustrie, Wechselkurse und Zinsinformationen integriert werden.

Neben statistischen Informationen will der ZVEI auch textuelle Informationen online zur Verfügung stellen. Dazu gehören u.a. das Konjunkturschlaglicht (Rubrik der ZVEI-Mitteilungen), Länderberichte, der Quartalsbericht für die deutsche Elektroindustrie und jährliche Strukturberichte zur Elektroindustrie in der EU.

5.2 Multimodalität und Text-Fakten-Mischung beim Information Retrieval

Bei Multimodalität bzw. Multimedialität in Informationssystemen sind drei Varianten zu unterscheiden, die zu jeweils eigenständigen Problemklassen und Lösungswegen führen, wenn sie auch eng miteinander verbunden sind. Allen gemeinsam ist die technische Problematik, die verschiedenen Datentypen effizient zu verwalten und wiederzugeben:

a) Multimediale Darstellung

Beim elektronischen Publishing bzw. bei der Ausgabe von Dokumenten nach einer Suchanfrage (z.B. Patentrecherche, Verkaufskataloge) sollen Grafiken und Bilder - bei CD-ROM-Versionen auch Animationen und Videosequenzen - die Informationswiedergabe optimieren.

Neben der technischen Problematik spielen hier konzeptuell vor allem Fragen des Layouts eine Rolle (cf. das GMD-IPSI Projekt APALO, Reichenberger et al. 1996) und die Frage, welche Informationseinheiten am sinnvollsten mit welcher Modalität darzustellen ist und wie sinnvolle Mischung aussehen sollen (cf. z.B. André et al. 1993; André/Rist 1993).

Dieser Aspekt bleibt im folgenden weitgehend ausgeklammert.

b) Gestaltung der Benutzungsoberfläche

Der Übergang vom Paradigma der formalsprachlichen Bedienung von Computersystemen zum derzeit vorherrschenden der sog. grafischen Oberflächen wie OS/2 oder Windows '95 ist hierfür ein typisches Beispiel. Heute geht die Entwicklung hin zu weitergehenden Visualisierungen und zur Integration von Bildwelten (als Überblick cf. Krause 1997).

Ein Ansatzpunkt, der auch für das Marktinformationssystem von Interesse sein könnte, ist das ergebnisorientierte grafische Retrieval von WING-M2 (cf. Wolff 1997), bei dem Benutzer Liniengrafiken direkt manipulieren und als neue Suchformulierung abschicken können (sog. visual formalisms).

c) Verschiedene Datentypen als Wissensgrundlage und ihre Inhaltserschließung

Im Gegensatz zu a) bilden hier verschiedene Datentypen wie Fakten in Zeitreihen, Bilder und Texteinheiten die Wissensgrundlage, in der computergestützt recherchiert wird.

Idealer muß der Benutzer bei seiner Suchformulierung nicht danach unterscheiden, ob er z.B. Fakten direkt als Antwort bekommen kann, ob das System nur Literaturverweise zur angefragten Thematik besitzt oder z.B. nur auf einen Experten verweist, der die Frage beantworten könnte.

Wie schon bei der traditionellen Literaturrecherche (cf. Krause 1992) läßt sich die Problematik c) nicht von b), dem WIE der Interaktion, trennen, wenn beide Aspekte auch konzeptuell auseinander gehalten werden müssen.

Die Variante c) bildet den Schwerpunkt der folgenden Überlegungen.

Multimedialität in heutigen kommerziellen Systemen bedeutet überwiegend die technische Integration verschiedener Datenobjekte, die z.B. zum Typ Text, Bild, Animation oder Videosequenz gehören können. Alle nicht sprachlich codierten Informationen werden dabei entweder aus einem Text oder einer Faktentabelle über einen Hyperlink direkt angesprochen, oder über zugeordnete Deskriptoren (z.B. als Bildbeschreibung) inhaltlich beschrieben. Konzeptuell bedeutet dies, daß Bilder, Grafiken oder Videosequenzen für die Recherche auf den Typ Textrecherche hin standardisiert werden. Die Standardisierungs-

funktion ist in der Regel intellektuell vermittelt; Die Deskriptorenvergabe erfolgt z.B. nicht über eine Bilderkennung (von hochspezialisierten Anwendungen wie der Auswertung von Fingerabdrücken oder experimentellen Forschungsprototypen abgesehen). Multimedialität im Information Retrieval entsteht in diesen Fällen somit vermittelt über die Darstellungs- und Eingabefunktion.

5.2.1 Integration von Texten und Fakten

Eine der offensichtlichsten Formen der Multimedialität, und gleichzeitig eine der konzeptuell schwierigsten, ist die Text-Fakten-Mischung. Aufgabenanalysen verschiedenster Anwendungsfälle zeigen immer wieder, daß eine Mischung eher die Regel als die Ausnahme ist.

- Versicherungssoftware wie OASIS (NKK, Regensburg) verbindet von der Aufgabenstellung her ganz selbstverständlich strukturierte Fakteninformationen wie Vertragsdatenbanken, die in Tabellen als Faktensammlung abgespeichert werden, mit textuellen Informationen wie den Briefwechsel mit den Versicherten. Bei Vertragsverhandlungen bzw. im Schadensfall müssen den Sachbearbeitern der Versicherungen all die unterschiedlichen Informationen zur Verfügung stehen.
- Die bei der GMD-IPSI entwickelte experimentelle Forschungsdatenbank HYDRA (cf. Gu et al. 1993) beinhaltet Fakten- und Textwissen über EU-Projekte. Die beschreibenden Textteile sind verbunden mit strukturierten Fakteninformationen wie dem Projektleiter, Förderstelle oder Projektlaufzeit.

Dieses Beispiel verweist darauf, daß Textretrievalsysteme schon immer Strukturfelder enthielten (wenn auch nur wenige, die sog. „gebundenen Deskriptoren“). Am typischsten sind die bibliographischen Angaben.

- Ein experimentelles System zur Integration von Texten und Fakten, und darüber hinaus von Bildern und Tondokumenten für Büroanwendungen, ist auch MULTOS (cf. Rabitti/Savino 1990).

Eine integrierte Recherche nach Datentypen verschiedener Modalitäten wie Texte, numerische Fakten und Grafiken ist bei auf dem Markt befindlichen Systemen in der Regel nicht möglich. Hier gibt es eine deutliche Trennung zwi-

schen Datenbanksystemen für (vor allem numerische) Fakten (vorwiegend relationale Datenbanken, mit steigender Tendenz objektorientierte) und Textrecherchesystemen (z.B. Literaturdatenbanken), die im Kern auf indexsequentiellen Listen der Deskriptoren aufbauen, die maschinell oder intellektuell aus den Texten gewonnen werden. Der Grund für dieses Auseinanderklaffen liegt vor allem in der technischen Entwicklung der letzten Jahrzehnte, die für beide Datentypen jeweils verschiedene Systemarchitekturen zur effizienten Behandlung entwickelte, mit der Konsequenz, daß die einheitliche Benutzersicht auf eine Problemstellung künstlich geteilt wurde (Informationen zu x, gleich in welchem Datentyp sie codiert ist). Benutzer müssen heute in der Regel zwei Datenbanken, die weder technisch noch konzeptuell miteinander verbunden sind, befragen, um eine Problemstellung zu behandeln, die beide Datentypen benötigt.

Auch in der Forschung wurden Faktenrecherche und Textretrieval bis vor wenigen Jahren stark getrennt.

Wissenschaftliche Projekte zur Integration von Texten und Fakten, die in den letzten Jahren durchgeführt wurden, beschäftigten sich vorwiegend mit technischen Aspekten und versuchten beide Modalitäten in einem System zugänglich zu machen (cf. weitere Beispiele in Womser-Hacker 1996). Wenig bekannt sind noch die aus der Integration entstehenden Problemen und Chancen aus Benutzersicht unter Berücksichtigung informationeller Prozesse.

Womser-Hacker 1996 hat in ihrer Habilitationsschrift gezeigt, daß Unsicherheit und Vagheit keine Attribute sind, die sich zur Trennung von Textretrieval vs. Faktenretrieval eignen, auch wenn bisher in der Regel davon ausgegangen wurde, daß Fakten weitgehend präzise sind und deshalb hierfür kaum Vagheitsinstrumente zum Einsatz kamen, im Gegensatz zum Textretrieval, das eine starke Tradition quantitativ-statischer Verfahren hat. Auch bei einem Marktinformationssystem ist davon auszugehen, daß die Behandlung von Unsicherheit und Vagheit eher ein Bindeglied der Integration von Fakten und Texten ist als ein Differenzierungsmerkmal, das die Integration erschwert (cf. auch den folgenden Abschnitt).

Im kommerziellen Umfeld nimmt ORACLE eine Sonderstellung in Bezug auf die Möglichkeiten zur Text- und Faktenintegration ein. Auf der Basis von Oracle7 und dem Textserver3 lassen sich im Rahmen eines leistungsfähigen kommerziellen relationalen Datenbanksystems - neben den üblichen Tabellen für das Faktenwissen - auch Terme textuellen Dokumenten zuordnen (Bitmap-Indices), ohne daß die Grundlage des relationalen Modells verlassen wird. Ent-

sprechend gibt es in der erweiterten SQL-Anfrage eine „CONTAINS“-Funktion, mit der die üblichen Tabellenanfragen durch Bedingungen für Deskriptoren in Textteilen erweitert werden können. Damit ist eine kommerziell verfügbare Plattform für das Marktinformationssystem gegeben, die es erlaubt, eine Text-Fakten Mischung auch konzeptionell ohne allzuviel Eigenprogrammierung zu realisieren.

5.2.2 Inhalterschließung von Texten und Text-Fakten Integration

Im Kern ist die heutige Trennung der Datenmodelle für Texte und Fakten in der Problematik der Inhalterschließung begründet. Fraglos enthalten Texte Fakten (Ludwig Erhardt wurde am 4.2.1897 geboren). Die Verfahren zur automatischen Indexierung verzichten aber darauf, die Fakten zu extrahieren, sondern begnügen sich mit syntaktisch kaum verbundenen semantischen Markern, den Deskriptoren, die nicht das beschriebene Fakt selbst enthalten, sondern Hinweise darauf, daß es in einem bestimmten Text steht (Erhardt, Geburtsdatum/geboren). Diese Reduzierung der Aufgabenstellung der Inhalterschließung auf den Hinweischarakter erleichtert die Wissensextraktion aus Texten erheblich und ist wohl für größere Textmengen, die nicht auf bestimmte, eng umgrenzte Aufgabengebiete eingeschränkt sind, der einzig derzeit gangbare Weg. Er ist zudem „natürlich“, da auch ein menschlicher Berater verschiedene Stufen von Wissen zur Problemlösung hat. Auch er weiß in vielen Fällen nur, wo ein Fakt gefunden werden könnte, bzw. von wem. Wissenseinheiten, die auf den Hinweischarakter eingeschränkt sind, zu behandeln, ist weiter per se unvermeidbar, da Texte immer Wissen in Form von Hinweisen enthalten werden, die sich nicht in die Form einer Faktentabelle bringen lassen, weil hierzu Informationen fehlen.

Daß sich jedoch zumindest ein großer Teil von Textinformationen für eng umgrenzte Aufgabengebiete und Nutzer bereits bei der Inhalterschließung in faktisches Wissen umwandeln ließe, zeigt z.B. das experimentelle System KONTEXT, das aus Texten extrahierte Fakten als eine Wissensstruktur neben der textuellen abspeichert (cf. Haenelt 1996).

Ob eine umfassende Faktenextraktion aus Texten für ein Marktinformationssystem wertvoll und machbar ist, kann im geplanten Projekt nicht geklärt werden. Lösungen dieser Art führen zwangsläufig zu neuen Typen von Datenbanksystemen. Sie zu einer Reife zu entwickeln, die einen kommerziellen Einsatz ermöglichen, überschreitet den vorgegebenen finanziellen und zeitlichen Rahmen erheblich. Der Erfolg wäre auch ungewiß. In der Künstlichen Intelligenz sind semantisch basierte Ansätze der Textextraktion aus Fakten immer wieder versucht worden, ohne daß sich Lösungen außerhalb von sehr engen Aufgabenfeldern ergeben hätten.

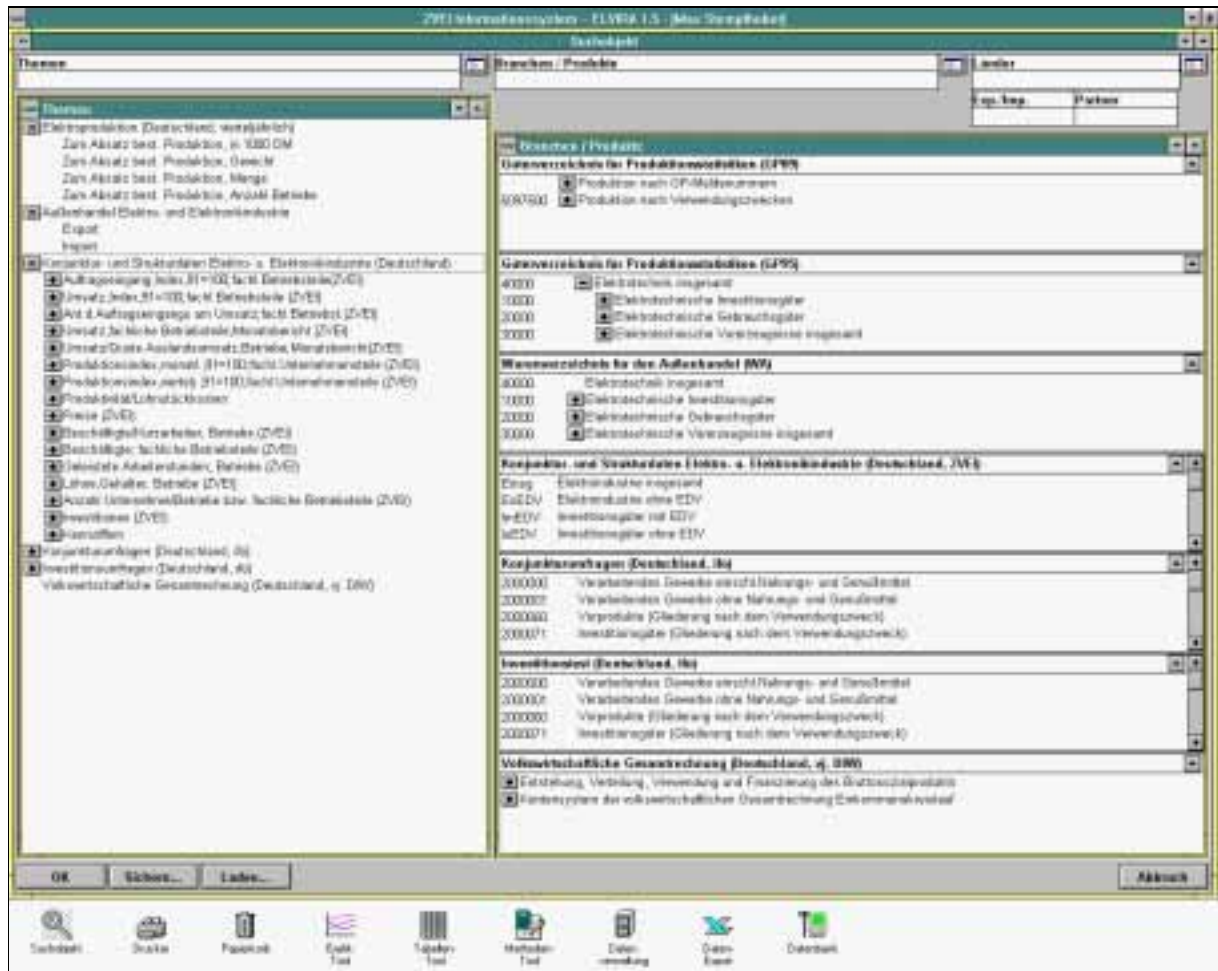
Deshalb beschränkt sich die Inhaltserschließung aus Texten im projizierten Marktinformationssystem vorerst auf die gängigen, kommerziell verwendbaren Verfahren (cf. als Überblick Womser-Hacker 1996).

5.2.3 Sonderfall ELVIRA: Deskriptoren zur Beschreibung von Zeitreihen

Bei ELVIRA handelt es sich um einen Spezialfall eines Fakteninformationssystems, der im Kontext Fakten-Text Mischung besonders interessant ist. ELVIRA basiert letzten Endes auf einer Deskriptorensuche, obwohl Faktendaten zugrunde liegen. Die Zeitreihen werden durch drei Deskriptoren aus den drei Facetten Themen, Branchen und Länder beschrieben, die intellektuell zugeordnet werden, nicht wie sonst üblich über ihre Attributwerte innerhalb der Tabelle. Damit liegen Fakten vor, die mit der für Texte spezifischen Repräsentationsform codiert werden.

Eine typische Anfrage wäre: „Gib mir die Produktionszahlen von Waschvollautomaten in Deutschland“.

Beim Außenhandel kann noch ein weiteres Land hinzukommen, so daß vier Deskriptoren erforderlich sind.



Die Fülle der prinzipiell möglichen Anfragen mit Boole'scher Algebra wird bei ELVIRA durch den Anwendungsbereich, die Inhalte der Datenbank und ihre Nutzungsbedürfnisse auf eine sehr spezifische Art eingeschränkt. Diesen Umstand macht sich der ELVIRA-Designentwurf zunutze. Ziel ist, die Boole'sche Algebra nicht explizit, sondern implizit durch die Vorgabe bestimmter Eingabemöglichkeiten festlegen zu lassen. Dieses Verfahren ist bei Faktendatenbanken mit Suchformularen durchaus üblich (cf. z.B. für WING-M2 Krause/Womser-Hacker 1997) und wurde schon von WIDAB für die ZVEI-Daten als nützlich erkannt und eingesetzt (detaillierter cf. Krause et al. 1996).

5.2.4 Formen der Kombination von Texten und Fakten in der Wissensbasis

Da Wissen prinzipiell in verschiedenen Modalitäten dargestellt werden kann, sind die bei einer Aufgabenstellung vorliegenden Mischformen vielfältig. Für das Marktinformationssystem sollen - über den Sonderfall von Abschnitt 2.1.3 hinaus - die folgenden Grundtypen geprüft und eventuell prototypisch umgesetzt werden:

- a) Bestimmte Zeilen, Spalten oder Zellen von Faktentabellen sind mit einem kurzen erläuternden Text verbunden, der z.B. Aussagen zur Meßgenauigkeit oder zu Besonderheiten der Fakten(Akkumulationen) enthält.
- b) Aus Faktentabellen bzw. aus einem Fließtext heraus gibt es Verweise auf selbständige Dokumenteinheiten wie Aufsätze, die den Fakteninhalt von Tabellen ergänzen, bzw. mit ihnen in einem verwaltungstechnischen Zusammenhang stehen (z.B. Briefwechsel zu einem Versicherungsvertrag).
- c) Text und Fakten sind in einem Dokument direkt vermischt. Der Faktenteil kann z.B. bei einer Publikation tabellarisch dargestellt und im Text erläutert, interpretiert und ergänzt werden.
- d) Die gleiche Wissensseinheit tritt innerhalb eines Dokuments bzw. eines Dokumentenbestandes gleichzeitig textuell und als numerisches Fakt codiert auf.

Einen Sinn macht die Behandlung der verschieden Formen einer Integration von Fakten und Texten allerdings erst, wenn parallel dazu ein adäquates Modell für die Behandlung des Problems bei der Rechercheformulierung durch den Benutzer gefunden wird. Eine rein technisch modellierte Ausweitung der Syntax von SQL wie in Oracle7 wird dem mit Sicherheit nicht gerecht.

5.3 Vagheit in Informationssystemen

In der menschlichen Kognition und in der zwischenmenschlichen Kommunikation spielen Vagheit und Unsicherheit eine große Rolle. Menschen können vage Informationen schnell und angemessen verarbeiten wie etwa in einer Wegbeschreibung (z.B. „kurz nach einer *scharfen* Linkskurve erscheint ein *großes* Haus“). Vagheit ist oft ein großer Vorteil, da exakte Informationen in vielen realen Situationen unangemessen sind (z.B. „ein Haus, das 9,7 m hoch ist“).

In den letzten Jahren haben vage Methoden in einigen Bereichen der Informatik einen starken Boom erlebt. Vor allem probabilistische Verfahren, Neuronale Netze und die Fuzzy Technologie werden zunehmend in hochkomplexen Problemstellungen eingesetzt, für die sie leicht handhabbar sind und oft schnell gute Ergebnisse liefern. Es setzt sich zunehmend die Einsicht durch, daß die Qualität eines Modells nicht von der Exaktheit der einfließenden Informationen abhängt, sondern u.a. von der Effizienz, Korrektheit und Benutzerfreundlichkeit (Kruse 1996).

Auch bei der Gestaltung von Informationssystemen werden Vagheit und Unsicherheit zunehmend als inhärente Eigenschaften des Suchprozesses verstanden und bei der Modellierung berücksichtigt. Im Faktenretrieval und bei Datenbanksystemen wird einerseits versucht, vage Informationsbedürfnisse zu unterstützen und andererseits auf der Wissensseite unsicheres Wissen zuzulassen.

5.3.1 Vagheit im Information Retrieval

Im Rahmen des Textretrieval beschäftigt sich die Forschung schon seit längerem mit „best-match-Verfahren“, die davon ausgehen, daß eine Anfrage nicht exakt auf eine Menge von Dokumenten abgebildet werden kann. So entstand etwa das Vektor-Space-Modell, bei dem Ähnlichkeiten zwischen Anfrage und Dokumenten aufgrund von Vektorvergleichen berechnet werden. Besonders die probabilistischen Verfahren sind bereits gut erforscht, werden aber in kommerziellen Systemen erst seit kurzem eingesetzt. Sie beruhen auf der Wahrscheinlichkeitstheorie und berechnen für Kombinationen von Anfragen und Dokumenten die Wahrscheinlichkeit, daß ein Dokument relevant ist. Die Text Retrieval Conference veranstaltet seit einigen Jahren Wettbewerbe für Information Retrieval Systeme, bei denen alle Systeme mit dem gleichen Korpus von Zeitungstexten und Anfragen arbeiten und anschließend bewertet werden. In diesen Wettbewerben (TREC I - IV) konnten probabilistische Systeme gute Ergebnisse erzielen (für einen Überblick über probabilistische Modelle und TREC cf. Womser-Hacker 1996).

Seit einigen Jahren wird auch versucht, neuronale Netze im Information Retrieval einzusetzen. Neuronale Netze sind eine Art der Informationsverarbeitung, die auf dem Vorbild des menschlichen Gehirns aufbaut. Zahlreiche parallele Prozessoren verarbeiten eingehende Signale und leiten wiederum Si-

gnale weiter. Lernfähige Modelle können Abbildungen durch die Verarbeitung subsymbolischer Merkmale aufgrund von Beispieldaten lernen und dadurch komplexe Klassifikationsaufgaben lösen. Basierend auf den Verbindungen zwischen den Neuronen wird diese Art der Informationsverarbeitung auch Konnektionismus genannt.

Besonders aufgrund der Lernfähigkeit scheinen neuronale Netze im Information Retrieval gut einsetzbar zu sein. Entsprechend sind bereits viele experimentelle Systeme entstanden (für einen Überblick cf. Doszkocs 1990). Ein besonders vielversprechender Ansatz ist das System SCALIR, das konnektionistische und traditionelle symbolische Methoden integriert (cf. Rose/Belew 1991). In ein Netz aus Gesetzestexten, Urteilen und Kommentaren können zum einen exakte Beziehungen eingegeben werden wie etwa „Urteil A revidiert Urteil B“. Daneben verfügt das Netz über weitere Verbindungen, die trainierbar sind und die im laufenden Betrieb des Systems komplexe und subtile Zusammenhänge lernen. Im Retrievalprozess wird auf beide Arten der Information zurückgegriffen.

Trotz des vermeintlich hohen Potentials spielten konnektionistische Verfahren in den TREC Wettbewerben bisher keine Rolle. Bei näherer Betrachtung zeigt sich, daß die meisten Systeme das subsymbolische Paradigma des Konnektionismus nicht ausnutzen und keine Mikroeigenschaften repräsentieren. Gerade dies machte die neuronalen Netze aber in anderen Bereichen so erfolgreich. Es müssen daher noch Architekturen gefunden werden, die dies leisten.

5.3.2 Vagheit im Faktenretrieval

Im Bereich Faktenretrieval herrschte lange die Ansicht vor, Fakten seien immer exakt. Jedoch wird in vielen Wissensgebieten mit unsicherem Wissen gearbeitet. Bei statistischen Daten führen etwa Erfassungsproblematiken zu Unsicherheit, in technischen Datenbanken treten u.a. aus Kostengründen für etwaige Messungen häufig Lücken auf. Auch Fehler finden sich in realen Datenbanken.

Selbst wenn man von exakten Daten ausgeht gibt, es auf der Benutzerseite vage Informationsbedürfnisse wie etwa eine natürlichsprachliche Problemsammlung zu einem Werkstoffsystem zeigt (cf. Womser-Hacker 1997). So formulie-

ren Benutzer ihre Wünsche mit vagen Quantoren oder fragen nach *ähnlichen* Datenbankobjekten.

Um diese Formen von Vagheit zu modellieren, wird häufig die Fuzzy Logik eingesetzt. Die Fuzzy Logik läßt im Gegensatz zur zweiwertigen aristotelischen Logik beliebige Wahrheitswerte zwischen 0 und 1 zu. So entstehen Fuzzy Mengen, zu denen Elemente mit einem bestimmten Grad gehören. Natürlichsprachliche Konzepte wie etwa „hohes Alter“ können damit besser modelliert werden als mit der klassischen Logik, in der es eine harte Grenze gibt, ab der jemand ein hohes Alter hat. Mit der Fuzzy Logik lassen sich solche Übergänge fließend modellieren, was eher dem menschlichen Denken entspricht. Die logischen Operatoren wie UND und ODER müssen darin neu definiert werden.

In zahlreichen experimentellen Systemen wurden bereits Fuzzy Erweiterungen für Datenbankabfragesprachen wie SQL vorgeschlagen. Damit werden Anfragen mit vagen Quantoren möglich:

- Finde alle Häuser, die *nah* am Zentrum sind und die einen *günstigen* Quadratmeterpreis haben.
- Finde alle *teueren* Gleichstrommotoren, die in Japan *viel* gekauft werden.

So stellen Yager/Larsen 1993 ein System vor, das eine Anfrage mit vagen Kriterien in eine breite SQL-Anfrage umsetzt, diese an eine relationale Datenbank schickt und die Ergebnisse anhand der ursprünglichen Anfrage sortiert.

Bosc/Pivert 1991 befanden, daß ihre Fuzzy Abfragesprache SQLf die in SQL bestehenden Äquivalenzen zwischen syntaktisch verschiedenen Anfragen weitgehend beibehält.

Neben vagen Anfragen wird auch auf der Datenseite Fuzzy Logik zur Modellierung eingesetzt. Medina et al. 1994 stellen mir GEFRED ein generelles Modell für ein Fuzzy Relationales Datenbanksystem vor. Darin können verschiedene Arten der Unsicherheit wie Lücken und Fuzzy Mengen repräsentiert werden. Über das gesamte Modell kann mit vagen Quantoren abgefragt werden.

Daneben gibt es Systeme, die als Spezialfall von Vagheit Ähnlichkeitsanfragen zulassen. Für den Anwender kann dies eine sehr komfortable und natürliche Art der Anfrageformulierung sein.

In einem auf Neuronalen Netzen basierenden System von Escobedo et al. 1993, das beim Design von Flugzeugteilen unterstützen soll, können anhand von vorgegebenen Designmustern ähnliche Teile recherchiert werden. In Pepperrell 1994 wird ein System beschrieben, das ähnliche dreidimensionale chemische Strukturen findet und von Experten sehr positiv beurteilt wurde.

5.3.3 Mangelnde Benutzerorientierung

Der vorhergehende Abschnitt zeigt, daß die Forschung viel zur Integration von Vagheitsmethoden in Informationssysteme geleistet hat. In kommerziellen Faktenretrievalsystemen und Datenbanken sind solche Verfahren jedoch praktisch nicht zu finden. Positive Beispiele zeigen aber, daß Vagheitskomponenten die Interaktion mit einem Fakteninformationssystem sehr wohl verbessern kann.

Die Erklärung für dieses Mißverhältnis liegt wohl in der mangelnden Benutzerorientierung vieler der vorgestellten Entwicklungen besonders im Bereich Fuzzy Datenbanken. Während die technischen Grundlagen einen guten Stand erreicht haben, findet sich als Motivation meist lediglich der Hinweis auf die vagen Konzepte der natürlichen Sprache. Die natürliche Sprache und die menschliche Kognition umfassen jedoch neben den oben genannten vagen Elementen auch exakte Zahlen und Konzepte; der Mensch setzt kontextabhängig die jeweils günstigste Strategie ein. Es ist aber sehr wenig darüber bekannt, in welchen Kontexten in Informationssystemen vage Konzepte etwa zur Anfrageformulierung besonders sinnvoll eingesetzt werden können.

Dutta 1991 schlägt vor, bei Nullantworten aus der Datenbank dem Benutzer ein ähnliches Objekt zu präsentieren, ohne dies auf ein bestimmtes Problem anzuwenden.

Neben dem Verlauf des Retrievalprozesses (große oder kleine Antwortmengen) scheint es aber auch andere situationsbedingte Einflüsse auf die Angemessenheit von vagen Methoden zu geben. Auch das Wissen eines Anwenders beeinflußt die Fragestellung. Recherchiert man etwas außerhalb des eigenen Fachbereichs oder unternimmt man Überblickssuchen, so werden vage Formulierung

gen von den Benutzern häufiger verwendet. Daneben werden Vagheitskomponenten in der Forschung besonders für sehr komplexe Datenbankobjekte entwickelt. Die oben diskutierte Integration von Text und Fakten verschärft die Vagheitsproblematik. Dadurch entstehen komplexe Datenbankobjekte und Beziehungen, die kognitiv häufig nur noch vage verarbeitet werden können.

Durch starke Nutzerorientierung bei der Entwicklung von Vagheitskomponenten kann eine deutliche Verbesserung der Interaktion erzielt werden. Im Projekt WING-IIR (Werkstoffinformationssystem mit natürlichsprachlicher/graphischer Benutzungsoberfläche - Intelligentes Information Retrieval) entstanden Prototypen von Vagheitswerkzeugen auf der Basis von Aufgabenanalyse und Benutzertests, die jetzt von der MTU in einem kommerziellen Werkstoffinformationssystem realisiert werden. Fuzzy-WING erlaubt die Verwendung vager Quantoren in Anfragen und interpretiert Bedingungen in adäquaten Kontexten vage (cf. Womser-Hacker 1997). Ein Ähnlichkeitswerkzeug auf der Basis neuronaler Netze sucht im Falle von Datenlücken nach ähnlichen Werkstoffen (Ludwig/Mandl 1997).

5.4 Dynamische Anpassung

Unterschiedlichste Erkenntnisse aus Benutzerstudien belegen starke Probleme der Benutzer beim Umgang mit grafischen Benutzungsoberflächen. Neben Störungen, die vor allem durch Verstöße gegen ergonomische Grundregeln verursacht werden (z.B. Inkonsistenzen der Oberfläche) treten mit zunehmendem Einsatz der Fenstertechnik Probleme auf, die von den Ein-Fenster Systemen her unbekannt waren. So läßt sich z.B. beobachten, daß 63% der Fenstermanipulationen in Mehr-Fenster-Applikationen auf das Wechseln zwischen den einzelnen Fenstern entfällt (Gaylin 1986). Zusätzlich wird der Benutzer durch - aufgrund der fehlenden grafischen Repräsentation der Beziehungen zwischen Fenstern durchzuführenden - Fenstermanipulationen sehr stark belastet (Bederson/Hollan 1994). Im direkten Zusammenhang stehen ferner Beobachtungen, wonach Benutzer bei zu vielen gleichzeitig geöffneten Fenstern („Windowitis“) die Orientierung verlieren, die Beziehungen zwischen Fenstern aufgrund von fehlenden optischen Hinweisen nicht mehr nachvollziehen können und unproduktiv arbeiten (Kahn/Charnock 1995). So verwundert es nicht, daß verstärkt die Forderung erhoben wird, Mehr-Fenster-Systeme bzw. deren einzelne Fenster sollten sich entsprechend ihrem Inhalt anpassen und den Benutzer vom Fenster-Management entlasten (Bly/Rosenberg 1986). Eine allge-

meine Motivation zum Einsatz neuer Gestaltungstechniken geht von der Beobachtung aus, daß Systeme i.d.R. von inhomogenen Benutzergruppen bedient werden (unterschiedliche Anforderungen und Bedürfnisse), sich die Bedürfnisse des einzelnen Benutzers ändern (vom Anfänger zum Fortgeschrittenen), sich das System selbst ändert (veränderter Systemzustand muß auf der Oberfläche dargestellt werden) und ein Benutzer verschiedene Systeme in evtl. unterschiedlichen Umgebungen einsetzt (Konsistenz über Plattformen und Applikationen hinweg) (Kühme/Schneider-Hufschmidt 1993).

Eine logische Konsequenz dieser Negativpunkte ist das Ausstatten von Softwaresystemen mit Adaptionmöglichkeiten. Während eine Vielzahl kommerzieller Systeme diese Forderung bereits erfüllt, festigt sich jedoch die Meinung, daß viele der Benutzer nicht in der Lage sind, die Adaptionfähigkeiten annähernd optimal einzusetzen. Auch die Art und Weise, in der die Anpassung Gestalt annehmen soll, ist äußerst umstritten. Allgemeiner Konsens besteht über die Verwendung in Hilfesystemen oder Erklärungskomponenten, wogegen die Verwendung zur Anpassung von Benutzungsoberflächen als zu komplexer Bereich angesehen wird (vgl. Diskussionsergebnis in Kühme & Schneider-Hufschmidt 1993). Als Folge davon beschäftigt sich der weitaus größte Teil der aktuellen Forschung mit Benutzermodellierung und Anpassung von für den Benutzer i.d.R. nicht sichtbaren Systemkomponenten. Vereinzelt wird der Benutzer jedoch auf Adaptionmöglichkeiten der Benutzungsoberfläche hingewiesen und bei ihrer Durchführung unterstützt (vgl. als Überblick hierzu Grunst et al. 1996)

Weitere, mehr auf Oberflächengestaltung ausgerichtete Arbeiten schlagen eine hierarchische Fensteranordnung, ein den verfügbaren Platz optimal ausnutzendes dynamisches Layout und Operationen vor, die sich gleichzeitig auf mehrere Fenster erstrecken (Kandogan/Shneiderman 1996). Das bereits kommerziell verfügbare Recherchesystem für Wirtschaftsdaten ELVIRA (Krause et al. 1996) nutzt hierarchische Fensteranordnungen um semantische Zusammenhänge zwischen Fenstern zu visualisieren und reduziert den Informationsgehalt übergeordneter Fenster durch das Ausblenden der in der konkreten Nutzungssituation nicht benötigten untergeordneten Fenster, wobei diese kontextabhängig wieder erscheinen. Durch die Unterordnung von Fenstern können Größenänderungen auf eine Gruppe von Fenstern in einem Bedienschritt ausgeführt werden. Zusätzlich wird der von Zustandsanzeigen benötigte Raum dynamisch variiert, so daß der Vorlagecharakter anderer Fenster stets maximiert bleibt. Um Umschaltvorgänge zwischen Fenstern zu reduzieren wurden Zugänge zu Funktionen mehrfach realisiert, so daß der Benutzer z.B. über die

Manipulation der Zustandsanzeige den Inhalt von Fenstern direkt verändern kann, ohne erst das entsprechende Fenster zu aktivieren.

Zusammenfassend existieren noch relativ wenig Erkenntnisse in bezug auf das optimale Zusammenspiel von Adaption und Adaptivität. Forschungsergebnisse aus dem Bereich der Benutzermodellierung können hier in Einzelfällen wertvolle Hinweise geben, verallgemeinerbare Schlußfolgerungen können jedoch wegen der i.d.R. sehr starken Einengung der betrachteten Domäne nicht gezogen werden. Verschiedene Systeme erreichen durch den Einsatz von Adaptivität oft nur marginale Verbesserungen z.B. in der Bediengeschwindigkeit (Browne et al. 1990). In die gleiche Richtung weisen Versuche, statische Bestandteile der Benutzungsoberfläche (z.B. Menüs) dynamisch zu verändern. Die Verringerung der Zugriffszeit auf Funktionen durch z.B. Umsortierung von Menüeinträgen unter Analyse des Benutzerverhaltens geht sehr oft mit einem Anstieg der Suchzeiten während des Erlernens der Systembedienung einher (Mitchell/Shneiderman 1989). Erfolgversprechender scheinen Maßnahmen zur kontextbezogenen Variation der Informationsdichte auf dem Bildschirm und dem intelligenten Management von Fenstergruppen.

5.5 Theoretisches Konzept bei der Softwareentwicklung

Gemäß dem empirischen Postulat der Informationswissenschaft muß der bei der Softwareentwicklung gewählte theoretische Ansatz empirisch ausgerichtet sein, d.h. daß Leistungsspektrum und Design eines zu entwickelnden Systems nicht rein analytisch bestimmbar sind, sondern durch geeignete Methoden und Verfahren in direkter Interaktion mit dem potentiellen Benutzerkreis erarbeitet werden müssen. Hierbei hat sich in der Informationswissenschaft das Rapid Prototyping-Verfahren bewährt, das durch seinen iterativen Charakter eine Alternative zum herkömmlichen Stufen-Modell in der Softwareentwicklung darstellt, indem es ein "refinement of the product based on feedback from users" (cf. Hix/Hartson 1993:250) ermöglicht. Frühe Benutzerpartizipation anhand von empirischen Prototypen erlaubt eine stringente Ausrichtung an konkreten Domänenerfordernissen und Benutzerwünschen sowie eine frühzeitige Korrektur des Designkonzeptes bei Fehlleistungen und bewirkt, daß dieser Ansatz dem pragmatischen Primat am ehesten gerecht wird.

Auf diese Weise kommen Methoden aus der Human Computer Interaction (kurz: HCI)-Forschung zur Anwendung, wie sie z.B. Preece et al. 1994:346 für den Entwurf von HCI-Schnittstellen fordern:

"... HCI design ... should:

- be user-centred and involve users as much as possible so that they can influence the design,
- integrate knowledge and expertise from the different disciplines that contribute to HCI design,
- be highly iterative so that testing can be done to check that the design does indeed meet users' requirements"

Beim Rapid Prototyping-Verfahren lassen sich folgende Entwicklungsphasen unterscheiden:

a) Vorphase A: Analyse der Anwendungsdomäne

In der sog. Knowledge Engineering-Phase können je nach Beschaffenheit der Anwendungsdomäne verschiedene Methoden zum Einsatz kommen:

- intensive Gespräche mit potentiellen Benutzergruppen über deren Aufgaben, Arbeitsweisen und Bedürfnisse
- Sichtung bzw. Rekonstruktion der interpersonellen Arbeits- und Geschäftsprozesse
- Analyse vorhandener, von den Benutzern verwendeter Werkzeuge und Hilfsmittel (Handbücher, Nachschlagewerke, Thesauri, EDV-Komponenten etc.)
- Studium domänenrelevanter Literatur zur Einarbeitung in den Gegenstandsbereich
- Fragebögen bzw. -kataloge zur Erhebung spezifischer Bedürfnisse und zur Analyse der konkreten Benutzerstruktur (Benutzerstereotype)

- Paper and Pencil-Tests bei Informationssystemen zur Generierung eines Skopus an Rechercheanfragen und typischen Problemstellungen
- explizite Regelformulierung durch Experten bei stark kausal geprägten Domänen (z.B. Expertensystemen)

b) Vorphase B: Entwicklung eines ersten Prototyps und Pretest

Gemäß den Ergebnissen aus der Vorphase wird ein erster Prototyp realisiert, der die Hauptaspekte des zu modellierenden Gegenstandsbereiches funktional abdeckt. Hierbei wird bereits auf ein ergonomisches Design geachtet (gemäß allgemeinen softwareergonomischen Erkenntnissen sowie anhand der Ergebnisse und Erfahrungen aus vorherigen Projekten), da gerade der erste Test nicht zu unterschätzende Auswirkungen auf die spätere Akzeptanz bei den Benutzern und deren Bereitschaft zur Teilnahme an weiteren Zyklen hat.

Um zu gewährleisten, daß der Prototyp nicht zu früh, d.h. in einem „unreifen“ Stadium einem Test unterzogen wird, erfolgt zunächst mit einer begrenzten Zahl von Benutzern (ca. 2-4 Personen) ein Pretest. Hierbei sollen die Versuchspersonen (VP) durch freies Explorieren oder auch schon durch Bearbeiten eines fixen Aufgabensets anhand der Thinking Aloud-Technik auf eventuell vorhandene gravierende Schwächen des Prototyps aufmerksam machen sowie den gesamten Testaufbau auf Praktikabilität hin überprüfen.

Eine solche Pretest-Phase kann bei Bedarf auch in der Hauptphase den eigentlichen Benutzertests vorgeschaltet werden.

c) Hauptphase: iterative Entwicklung und Test von Prototypen, Redesign

Die Hauptphase läßt sich bzgl. Planung und Durchführung im wesentlichen in vier Teilphasen unterscheiden (in Anlehnung an die Formative Evaluation-Methode von Hix/Hartson 1993:283ff):

- Developing the Experiment

In diese Phase fällt die Auswahl der Versuchspersonen (VP), die einen Querschnitt über die verschiedenen Abteilungen bzw. Arbeitsbereiche der jeweiligen Anwendungsdomäne darstellen sollten, um das gesamte

Funktionalitätsspektrum abdecken zu können. Bei kleinen Benutzergruppen (ca. 20 Personen) kann auch ein Volltest erfolgen.

Anschließend ist der Testaufbau zu klären. Idealerweise soll der Prototyp in vivo getestet werden, d.h. in der konkreten Arbeitssituation unter Produktionsbedingungen, wo er die bislang genutzten Instrumente für einen bestimmten Zeitraum (ca. 1 bis 2 Wochen) ersetzt bzw. ergänzt. Ist dies nicht möglich (z.B. bei sehr sensiblen Domänen oder hochkomplexen Arbeitsprozessen) muß in einem zeitlich fest umrissenen Rahmen (in der Regel 30 Minuten bis max. 1,5 Stunden) anhand eines möglichst realistischen Testaufbaus die Arbeitsatmosphäre simuliert werden.

Der Standardtestaufbau besteht hierbei meist aus einem Rechner mit dem zu evaluierenden System, an den ein zweiter Monitor angeschlossen wird. Dieser sowie die VP werden gefilmt und die Aufnahmen über ein Mischpult gekoppelt, so daß die Aktionen der VP (Gestik, Handbewegungen u.ä.) und die Bildschirmfolgen bzw. Eingaben synchron beobachtet werden können.

Bei noch nicht realisierten Komponenten, die aber für das Verständnis und die Akzeptanz des Gesamtsystems wesentlich sind, empfiehlt sich ein Simulationsaufbau, bei dem das fehlende Modul im Hidden Operator-Verfahren simuliert wird.

Zu Beginn des Tests sind die VP in die Systembedienung und -funktionalität einzuführen und mit der Testumgebung vertraut zu machen.

Nach Testende sollten Interviews zu den jeweiligen Systemen durchgeführt werden, die den VP eine abschließende Bewertung erlauben.

- Generating and Collecting Data

Idealerweise sollten beim Rapid Prototyping-Verfahren sowohl qualitative als auch quantitative Daten erhoben werden, um ein möglichst umfassendes Bild von der Eignung eines Systems für die gegebene Aufgabenstellung zu erhalten.

Als qualitativ sind hierbei alle Interaktionsschritte sowie die schriftlichen und mündlichen Aussagen der Versuchspersonen einzustufen. An

quantitativen Daten sind die Dauer der Aufgabenbearbeitung, die Anzahl der Interaktionsschritte sowie die Fehlerquote relevant.

Liegen Ton- oder Bildaufnahmen der Tests vor, so sind von diesen detaillierte Transkripte zu erstellen, die die Ausgangsbasis für die Datenanalyse bilden.

- Analyzing the Data

Für die Analyse der Benutzeraktionen und -äußerungen auf der Basis der Transkripte, Fragebögen oder Protokolle im Rahmen einer qualitativen Datenanalyse haben sich insgesamt folgende vier Kategorien bewährt (cf. Krause/Womser-Hacker 1997):

1. Bedienung: Auffälligkeiten beim Umgang mit den angebotenen Funktionselementen
2. Informationsdarstellung: Kommentare zur Präsentation des Rechercheergebnisses sowie der Funktionselemente
3. Funktionsumfang: Art und Menge der angebotenen Funktionselemente
4. Allgemeines: allgemeine Bemerkungen zur jeweiligen Such- oder Bearbeitungsstrategie sowie Anmerkungen zu Datenbankstruktur und -inhalt bei DB-basierten Anwendungen

Die quantitativen Daten sind in Form von Häufigkeiten tabellarisch zu erfassen und dienen in Kombination mit den qualitativen Daten zur Beurteilung der getesteten Systeme.

Liegt die Auswertung der Testergebnisse vor, ist der Prototyp entsprechend zu überarbeiten bzw. bzgl. der Funktionalität auszuweiten. u.U. kann es hierbei auch zu einem Verwerfen des ersten Ansatzes und zu einem völligen Neuentwurf kommen, wenn sich der ursprüngliche Systementwurf als nicht tragbar erwiesen hat. Durch den Prototyp-Charakter der Komponente bleibt der Verlust an Arbeitszeit im Vergleich zu einer Vollimplementation jedoch begrenzt.

Die einzelnen Zyklen der Hauptphase lassen sich beliebig oft wiederholen, wobei sich in der Praxis ein dreimaliges Durchlaufen als in den meisten Fällen ausreichend erwiesen hat.

d) Endphase: Erstellung eines lauffähigen Systems und Dokumentation

Ist ein hoher Deckungsgrad zwischen Benutzererwartung, Bedienkomfort und Systemleistung erreicht, ist der Prototyp in eine fehlerrobuste Vollversion überzuführen und eine Systemdokumentation bzw. ein Benutzerhandbuch zu erstellen.

5.6 GUI-Tools

Die BOF von ELVIRA wurde mit dem GUI-Tool PowerBuilder erstellt. Dabei zeigten sich bereits Schwächen, die in der Natur von GUI-Tools als generellen Werkzeugen für einen breiten Markt liegen. Sie sind auf den momentanen Umfang von Elementen von grafischen BOF fixiert und erlauben nur in eingeschränktem Maße die Integration von innovativen Gestaltungselementen. Es ist zu erwarten, daß wie bei der Entwicklung von ELVIRA teilweise auf traditionelle Programmiersprachen (C, C++, Pascal) zurückgegriffen werden muß.

Der Markt für GUI-Tools soll im Rahmen des Projektes beobachtet werden und vielversprechende Systeme sollen getestet werden (z.B. GRIT, Visual Age). Ebenso muß die Entwicklung verfolgt werden, die sich aus der weiten Verbreitung von GUI-Möglichkeiten im Internet u.a. durch die Einführung von JAVA ergibt.

6 Projektpartner

Die Projektpartner organisieren sich in einem „Steuerkreis“. In regelmäßigen Treffen werden die Projektziele abgesteckt sowie die erforderlichen Arbeitsschritte definiert und koordiniert. An den Treffen nehmen die primär mit den Arbeitspaketen befaßten Mitarbeiter teil.

6.1 Verband des Deutschen Maschinen- und Anlagenbaus (VDMA)

Der VDMA ist einer der Spitzenverbände der deutschen Wirtschaft mit zahlreichen Mitgliedern. Er verfügt über große Datenbestände, die aber teilweise noch verteilt und heterogen sind, und z.Zt zu einer Datenbank integriert werden. Weiterhin verfügt er über große Erfahrung bei der Informationsvermittlung an die Firmen und bei der Analyse der Daten. Aufgrund der starken Exportausrichtung seiner Firmen ist der VDMA besonders an internationalen Informationen interessiert.

Der VDMA verfügt bereits über breites Wissen im EDV Bereich. In der Projektlaufzeit soll die nötige Kompetenz zur Pflege von System und Datenbank aufgebaut werden, auf die dann auch andere Verbände zurückgreifen können.

6.2 Hauptverband der Deutschen Bauindustrie (HVB)

Der HVB vertritt eine Branche mit besonderer Bedeutung für die Gesamtwirtschaft. Er gibt wie andere Verbände zahlreiche, verschieden strukturierte Informationen an seine Mitglieder weiter. Für die Mitgliedsfirmen sind Informationen in regionaler Gliederung besonders wichtig.

Der HVB verfügt im Bereich EDV nur über geringe Kompetenz und Ressourcen, so daß eine Eigenentwicklung nicht möglich ist. Besonders hinsichtlich der Pflege der Datenbank soll im Verlauf des Projektes Know-How aufgebaut werden. Auf Basis der aktuellen Version von EIVIRA möchte der HVB auf jeden Fall beginnen, seine Daten in die Datenbankstruktur zu integrieren und das entwickelte System den interessierten Unternehmen zur Verfügung zu stellen.

6.3 Zentralverband der Elektrotechnik- und Elektronikindustrie (ZVEI)

Der Zentralverband der Elektrotechnik- und Elektronikindustrie (ZVEI) vertritt die gemeinsamen Interessen der Elektro- und Elektronikindustrie und der dazugehörigen Dienstleistungsunternehmen in Deutschland und damit eine der bedeutendsten Industriegruppen mit einem Umsatz von 225 Mrd. DM pro Jahr. Bereits jetzt setzt der ZVEI das online Verbandsinformationssystem ELVIRA bei einigen Unternehmen ein. Er kann laufend über Erfahrungen berichten und die Informationsbedürfnisse konkretisieren. Der ZVEI verfolgt die laufenden Entwicklungen im Projekt und testet sie auf Anwendbarkeit für seine Firmen. Einzelne Systemkomponenten werden in das bestehende System ELVIRA integriert und getestet. Dazu wird die bestehende Datenbank gepflegt und erweitert.

6.4 InformationsZentrum Sozialwissenschaften (IZ) der GESIS

Die Abteilung Forschung und Entwicklung des IZ Sozialwissenschaften (Leitung: Prof. Dr. J. Krause) hat bei der Entwicklung von ELVIRA Erfahrungen mit den informationellen Prozessen in der Wirtschaftsinformation und deren Umsetzung in eine benutzerfreundliche BOF gesammelt. Dabei griff die Projektgruppe auf mehrjährige Erfahrungen im Design von benutzerfreundlicher Software in mehreren Drittmittelprojekten und kommerziellen Projekten zurück (z.B. Werkstoffinformationssystem WING).

Am IZ Sozialwissenschaften Bonn sind zudem Synergieeffekte mit anderen Projekten des IZ zu erwarten, insbesondere mit der Entwicklung des integrierten Informationssystems GESINE, das sozialwissenschaftliche Literatur- und Forschungsinformationen beinhaltet. Dieses System wird von einer Gruppe von vier Entwicklern realisiert, wobei weitere Impulse von Gastwissenschaftlern und Stipendiaten am IZ zu erwarten sind. Damit stellt sich das Problem verschiedener Modalitäten ebenso wie in einem umfassenden Marktinformationssystem. Positive Effekte dürften sich sowohl auf konzeptueller als auch auf programmiertechnischer Ebene einstellen.

Gleichzeitig ist Prof. Dr. Krause Lehrstuhlinhaber am Institut für Informatik der Universität Koblenz-Landau und lehrt dort Software-Ergonomie und infor-

mationelle Prozesse. Von dieser Universitätsanbindung profitiert sowohl das IZ als auch die am IZ durchgeführten Projekte. Bei der Gestaltung von Benutzungsoberflächen für das Information Retrieval besteht eine Kooperation, bei der Oberflächenkonzepte generell entwickelt werden, die dann auf die Anwendbarkeit auf den Anwendungsfall GESINE geprüft werden.

6.5 Ausgewählte Unternehmen der beteiligten Verbände

Die Unternehmen setzen das Marktinformationssystem testweise ein und schaffen die dazu nötigen Hardware-Voraussetzungen. Sie stellen Mitarbeiter für Benutzertests zur Verfügung.

7 Bisherige Arbeiten der Antragsteller

7.1 Verband Deutscher Maschinen- und Anlagenbau (VDMA)

Zusätzlich zu den unter 5.1.1 genannten Datenbanken befinden sich weitere Datenbanken bereits im Aufbau, darunter Datenbanken über Zölle, Normen, technische Regelwerke und Gesetzestexte; eine Integration der einzelnen Datenbanken steht noch aus.

Diese Datenbanken stehen heute im VDMA bereits intern und in einer späteren Phase den Verbandsmitgliedern über das Internet direkt zur Recherche zur Verfügung. Die Erfahrung hat aber gezeigt, daß die unterschiedlichen Rechercheverfahren den ungeübten Nutzer überfordern. Der Grad der Unzufriedenheit - und damit der Ablehnung - steigt mit der Anzahl der Datenbanken bzw. der unterschiedlichen Benutzungsoberflächen. Hier soll durch das Forschungsvorhaben der Durchbruch erzielt werden, KMUs tatsächlich die Nutzung von Datenbanken zu ermöglichen.

Einen aktuellen Überblick über die verfügbaren Datenbanken und Recherechensysteme gibt die Broschüre „Marktinformationssystem des VDMA“ (VDMA 1997).

7.2 Hauptverband der Deutschen Bauindustrie (HVB)

Der Hauptverband hat schon Anfang der 80er Jahre zusammen mit dem Zentralverband des Deutschen Baugewerbes eine Studie über die Entwicklung einer Bauwirtschaftlichen Datenbank erarbeitet. Was die Informationsinhalte sowie die BedürfnisBedürfnisse der Nutzer angeht, sind die Ergebnisse dieser Studie durchaus heute noch aktuell. Sie vermitteln die Vielfältigkeit der Bauinformationen und die Probleme, die bei der Recherche der Daten durch die verschiedenen Nutzergruppen entstehen.

7.3 Zentralverband der Elektrotechnik und Elektronikindustrie (ZVEI)

Basis des beantragten Projektes ist ELVIRA, das elektronische Verbandsinformations-, Recherche- und Analysesystem, das der ZVEI zusammen mit den Instituten DIW, IFO und IZ als vom BMWi gefördertes Pilotprojekt (1.10.94 - 31.12.96; Förderkennzeichen II C7-003060/10) entwickelt hat. Das ZVEI-Projekt wurde hinsichtlich

- *Systemarchitektur* (Client/Server-System mit Möglichkeit der offline-Formulierung von Anfragen und offline-Nutzung der recherchierten Daten),
- *Inhalt* (verfügbar sind die Statistiken des ZVEI, die IFO Konjunktur- und Investitionsdaten sowie die Volkswirtschaftliche Gesamtrechnung des DIW), sowie
- *Benutzungsoberfläche* (komfortable grafische Benutzungsoberfläche mit dynamischer Bildschirm Anpassung abhängig von der Recheresituation)

entsprechend den ursprünglichen Zielvorgaben erfolgreich abgeschlossen. Während der Entwicklungsphase wurde ELVIRA in verschiedenen Stadien bereits bei Pilotfirmen eingesetzt und getestet, wodurch wertvolle Hinweise für die Weiterentwicklung gewonnen werden konnten. Viele Anregungen aus der

Praxis konnten bereits während der Entwicklungsphase berücksichtigt und umgesetzt werden. Inzwischen wird ELVIRA vom ZVEI für alle interessierten Unternehmen angeboten. Es wurde bereits bei ca. 40 Firmen installiert. Der Verband betreut die Anwender hinsichtlich Installation und Nutzung.

ELVIRA ist also geeignet, auf breiter Basis zur Anwendung zu kommen. Modifizierte Testimplementationen erfolgen bereits bei VDMA und HVB.

7.4 Informationszentrum Sozialwissenschaften (IZ) der GESIS

Das IZ Sozialwissenschaften, das durch Kooperationsvereinbarungen mit der Informatik der Universität Koblenz-Landau verbunden ist, fertigte in den vergangenen Jahren eine Reihe industrieller und wissenschaftlicher Softwareentwicklungen mit den Schwerpunkten Softwareergonomie (cf. Krause 1996b), Intelligentes Information Retrieval und automatische Inhaltserschließung.

- **WING-IIR:** Das vom BMWi geförderte Projekt WING-IIR (Werkstoffinformationssystem mit natürlichsprachlicher/graphischer Benutzungsoberfläche und Intelligentes Information Retrieval, bis März 1995 FG Informationswissenschaft Universität Regensburg) verfolgte die empirisch fundierte Entwicklung einer multimodalen Benutzungsschnittstelle für Werkstoffinformationssysteme der Partner mtu, München, Institute for Advanced Materials, Petten, NL und IMA Materialforschung und Anwendungstechnik GmbH, Dresden. Vages Retrieval und die Gestaltung der bedienfreundlichen Benutzungsoberfläche waren die Schwerpunkte des Projektes. Das Kernsystem basiert auf der Werkzeugmetapher und den Prinzipien der Objektorientierung (d.h., Werkzeugobjekte werden auf andere Objekte angewendet). Vagheit und verstärkte Visualisierung sind auf vielfältige Weise verschränkt, so daß sie parallel behandelt wurden. Bei der Gestaltung der Benutzungsoberfläche stellte sich besonders die Frage nach der adäquaten Verknüpfung verschiedener Modalitäten. So übernimmt die natürliche Sprache die Funktion einer manipulierbaren Zustandsanzeige, wobei an verschiedenen Systemzuständen das natürlichsprachliche semantische Pendant generiert wird. Der entstandene Prototyp wurde von der MTU in ein kommerziell einsetzbares

System umgesetzt (für einen Überblick zu WING cf. Krause et al. 1993 und Krause/Womser-Hacker 1997).

- **ELVIRA:** Das IZ entwickelte die Benutzungsoberfläche für das ZVEI-Verbandsinformationssystem ELVIRA (siehe ZVEI, cf. Krause et al. 1996; Stempfhuber 1996).
- **COGET:** Das System COGET (Computergestützte Erstellung von Themendokumentationen) wurde am IZ entwickelt, um die Erstellung der halbjährlich erscheinenden Dokumentation Sozialwissenschaftlicher Informationsdienst (soFid, umfaßt Dokumentationen zu 33 Fachgebieten) zu erleichtern. COGET unterstützt die Zuordnung von Dokumenten zu Gliederungspunkten und erstellt automatisch eine druckfertige Vorlage (cf. Riege et al. 1996).
- **GESINE:** Das Projekt GESINE (GESIS-Datenbestände integrierendes sozialwissenschaftliches Informationssystem) soll die heterogenen Datenbestände der Institute Zentralarchiv für empirische Sozialforschung (ZA, Köln), Zentrum für Umfragen, Methoden und Analysen (ZUMA, Mannheim) und Informationszentrum Sozialwissenschaften, die bisher getrennt recherchiert werden müssen, integrieren. Damit stellen sich hier ebenfalls die Probleme der Integration von Texten und Fakten. Im Projekt GESINE arbeitet am IZ eine Gruppe von vier Entwicklern an der Datenbankstruktur und der Benutzungsoberfläche (cf. Marx et al. 1995).
- **HSP-Stelle Visualisierung:** Am IZ werden im Rahmen einer HSP-Stipendiaten-Stelle fortgeschrittene Ansätze zur Visualisierung im Information Retrieval untersucht (zu Visualisierung cf. Krause 1996a; Krause 1997). Als erster Ansatz wurde dazu eine Version des InfoCrystal (cf. Spoerri 1994) implementiert und getestet.
- **Kooperationsabsprache DESIGN:** Zwischen Prof. Bürdek, Offenburg, Prof. Bauer-Wabnegg von der Firma Medien-Design, Nürnberg dem IZ und der Universität Koblenz-Landau besteht eine Kooperation, die auf die stärkere Integration von Software-Ergonomie und Design ausgerichtet ist. Dazu werden in einem ersten Arbeitsschritt softwareergonomische Oberflächenentwürfe von Designexperten geprüft und verbessert.
- **HSP-Stelle Intelligente Hilfe:** In einer weiteren Stipendiaten-Stelle wird am IZ ein adaptives Hilfesystem entwickelt, das auf der Basis der Planerkennung kontextabhängig Hilfe bietet und im Rahmen der Windows-Hilfe reali-

siert wird. Diese Arbeit basiert u.a. auf früheren Hilfesystemen aus dem Bereich der Büroautomatisierung (cf. Krause et al. 1993).

- **Internetprojekt GESIS:** Das Internetangebot der GESIS-Institute wurde nach softwareergonomischen Gesichtspunkten neu gestaltet. Zusätzlich soll ein Clearinghouse für das deutsche Angebot im Rahmen der Sozialwissenschaften eingerichtet werden.
- **Kooperation Universität Regensburg und GIRT:** Mit Frau Dr. Womser-Hacker von der Universität Regensburg, Fachbereich Informationswissenschaft, besteht eine Kooperation im Bereich automatische Indexierung und Information Retrieval Verfahren. Frau Dr. Womser-Hacker hat ihre Habilitationsschrift zu diesen Themen kürzlich abgeschlossen. In diesem Zusammenhang wird am IZ eine deutsche Testdatenbank (German Indexing and Retrieval Testdatabase - GIRT) für IR Systeme aufgebaut und gepflegt. An dieser Basis können interessierte Firmen und Forschungseinrichtungen ihre Retrieval- und Indexierungsverfahren testen und mit den am IZ eingesetzten Verfahren vergleichen.

8 Literaturverzeichnis

- André, Elisabeth; Finkler, Wolfgang; Graf, Winfried; Rist, Thomas; Schauder, Anne; Wahlster, Wolfgang (1993): WIP: The Automatic Synthesis of Multimodal Presentations. In: Maybury, Mark (Hrsg.): Intelligent Multimedia Interfaces. AAAI Press. Menlo Park California. S. 75-93.
- André, Elisabeth; Rist, Thomas (1993): The Design of Illustrated Documents as a Planning Task. In: Maybury, Mark (Hrsg.): Intelligent Multimedia Interfaces. AAAI Press. Menlo Park California. S. 94-116.
- Bederson, Benjamin; Hollan, James (1994): Pad++: A Zooming Graphical Interface for Exploring Alternate Interface Physics. In: Proceedings of the UIST'94, User Interface Software and Technology Conference. S. 17-26.
- Bosc, Patrick; Pivert, Olivier (1992): About Equivalences in SQLf, a Relational Language Supporting Imprecise Querying. In: Terano, T. et al. (Hrsg.): Fuzzy Engineering towards Human Friendly Systems. IFES '91. Amsterdam. S. 309-320.
- Bly, Sara; Rosenberg, Jarrett. (1986), A Comparison of Tiled and Overlapping Windows. In: Proceedings of the CHI'86 Conference - Human Factors in Computing Systems, ACM, New York. S. 101-106.
- Browne, Dermot; Totterdell, Peter; Norman, Mike (1990): Adaptive User Interfaces. Academic Press, London.
- Doszkocs, T.E.; Reggia, J.; Lin, X. (1990): Connectionist Models and Information Retrieval. In: Annual Review of Information Science and Technology (ARIST), vol. 25. S. 209-260.
- Dutta, Soumitra (1991): Approximate Reasoning by Analogy to Answer Null Queries. In: International Journal of Approximate Reasoning, vol. 5, 4. S. 373-398.
- Escobedo, Richard; Smith, Scott; Caudell, Thomas (1993): A Neural Information Retrieval System. In: International Journal of Advanced Manufacturing Technology, vol. 8, 4. S. 269-274.
- Gaylin, Kenneth B. (1986): How are windows used? Some notes on creating empirically-based windowing benchmark task. In: Proceeding of CHI'86 Conference - Human Factors in Computing Systems, ACM, New York. S. 96-100.
- Grunst, Gernoth; Oppermann, Reinhard; Thomas, Christoph G. (1996): Adaptive and Adaptable Systems. In: Hoschka, Peter (Hrsg.): Computers as Assistants. A New Generation of Support Systems. Lawrence Erlbaum Associates, Mahwah, New Jersey. S. 29-46.
- Gu, Junzhong; Thiel, Ulrich; Zhao, Jian (1993): Efficient Retrieval of Complex Objects: Query Processing in a Hybrid DB and IR System. In: Knorz, Gerhard; Krause, Jürgen, Womser-Hacker, Christa (Hrsg.): Information Retrieval '93. Von der Modellierung zur Anwendung. Proceedings der 1. Tagung Information Retrieval '93. Konstanz. S. 67-81.

- Haenelt, Karin (1996): Das KONTEXT-Modell und die Konzeption der textmodellbasierten Verarbeitung natürlichsprachiger Texte. Arbeitspapiere der GMD 1009. Juli 1996. Darmstadt.
- Hix, D.; Hartson, H. (1993): Developing User Interfaces. Ensuring Usability Through Product & Process. New York et al.
- Kahn, Michael; Charnock, Elisabeth (1995): How to prevent „Windowitis“ in your graphical interface? In: Proceedings of the Silicon Valley Ergonomics Conference & Exposition, ErgoCon '95. S. 211-243.
- Kandogan, Eser; Shneiderman, Ben (1996), Elastic Windows: Improved Spatial Layout and Rapid Multiple Window Operations. To appear in Proceedings of ACM AVI'96 - Advanced Visual Interfaces.
- Knorz, Gerhard; Krause, Jürgen; Womser-Hacker, Christa (1993) (Hrsg.): Information Retrieval '93. Von der Modellierung zur Anwendung. Proceedings der 1. Tagung Information Retrieval '93. Konstanz. S. 67-81.
- Krause, Jürgen (1992): Intelligentes Information Retrieval. Rückblick, Bestandsaufnahme und Realisierungschancen. In: Kuhlen, Rainer (Hrsg.): Experimentelles und praktisches Information Retrieval. Festschrift für Gerhard Lustig, Konstanz. S. 35-58.
- Krause, Jürgen (1993): An Multilayered Empirical Approach to Multimodality. Towards Mixed Solutions of Natural Language and Graphical Interfaces. In: Maybury, Mark (Hrsg.): Intelligent Multimedia Interfaces. AAAI Press. Menlo Park California. S. 328-352.
- Krause, Jürgen (1994): Das WOB-Modell zur softwareergonomischen Gestaltung grafischer Oberflächen. In: Proceedings Deutscher Dokumentartag 1994. Blick Europa! Informations- und Dokumentationsmanagement. Universität Trier, 27.-30.9.1994. Frankfurt am Main: Dt. Ges. f. Dokumentation. S. 309-325.
- Krause, Jürgen (1996): Principles of Content Analysis for Information Retrieval Systems. An Overview. In: Zuell, C.; Harkness, J.; Hoffmeyer-Zlotnik, J. (Hrsg.): Text Analysis and Computer. ZUMA-NACHRICHTEN Spezial. S. 77-104.
- Krause, Jürgen (1996a): Metaphern versus visual formalisms: zwei konträre Gestaltungskonzepte für visuelle Benutzungsoberflächen? In: Neubauer, Wolfram (Hrsg.): Deutscher Dokumentartag 1996 - Die digitale Dokumentation: Neue Universität Heidelberg, 24. bis 26. September 1996. Frankfurt am Main: Dt. Ges. f. Dokumentation. S. 201-210.
- Krause, Jürgen (1996b): Softwareergonomie. In: Buder, Marianne; Rehfeld, Werner; Seeger, Thomas; Strauch, Dietmar (Hrsg.): Grundlagen der praktischen Information und Dokumentation: Ein Handbuch zur Einführung in die fachliche Informationsarbeit; Bd.2.-4., völlig neu gefaßte Ausg. München: Saur, 1996. (DGD-Schriftenreihe), S. 688-697.
- Krause, Jürgen (1997): Visual formalisms and „natural mapping“. In: Krause/Womser-Hacker 1997. S. 237-258.

- Krause, Jürgen; Herfurth, Matthias; Marx, Jutta (1996) (Hrsg.): Herausforderungen an die Informationswirtschaft: Informationsverdichtung, Informationsbewertung und Datenvisualisierung; Proceedings des 5. Internationalen Symposiums für Informationswissenschaft (ISI '96), Humboldt-Universität zu Berlin, 17. - 19.10.1996. Konstanz: Universitäts-Verlag (Schriften zur Informationswissenschaft).
- Krause, Jürgen; Hirschmann, Astrid; Mittermaier, Eva (1993): The Intelligent Help System COMFO-HELP. Towards a Solution of the Practicality Problem for User Modeling and Adaptive Systems. In: Kobsa, Alfred (Hrsg.): User Modeling and User -Adapted Interaction. An International Journal. Kluwer Academic Publishers, Dordrecht, NL. vol. 3 S. 249-282.
- Krause, Jürgen; Mandl, Thomas; Stempfhuber, Maximilian (1996): Design des ersten Prototypen des ZVEI-Verbandsinformationssystems ELVIRA. IZ-Arbeitsbericht Nr. 4, IZ Sozialwissenschaften, Bonn.
- Krause, Jürgen; Marx, Jutta; Roppel, Stephan; Schudnagis, Monika; Wolff, Christian; Womser-Hacker, Christa (1993): Multimodality and Objektoriented Interaktion in an Intelligent Materials Information System. Part 1. In: Journal of Document and Text Management, vol. 1, 3. November 1993. S. 256-275.
- Krause, Jürgen; Wolff, Christian; Womser-Hacker, Christa (1994): Multimodality and Vagueness in the Context of a Graphical, Object-Oriented Materials Information System. In: Intelligent Multimedia Information Retrieval Systems and Management. Proceedings of the RIAO 95 (Recherche d'Information assistée par Ordinateur) Rockefeller University, New York, October 11-13, 1994. S. 585-597.
- Krause, Jürgen; Womser-Hacker, Christa (1997) (Hrsg.): Vages Information Retrieval und graphische Benutzeroberflächen - Beispiel Werkstoffinformation. Konstanz: Universitäts-Verlag (Schriften zur Informationswissenschaft).
- Kruse, Rudolf (1996): Fuzzy-Systeme - Positive Aspekte der Unvollkommenheit. In: Informatik-Spektrum, vol. 19. Berlin. S. 4-11.
- Kühme, Thomas; Schneider-Hufschmidt, Matthias (1993): Introduction to Adaptive User Interfaces. In: Schneider-Hufschmidt, Matthias; Kühme, Thomas; Malinowski, Uwe (Hrsg.): Adaptive User Interfaces. Principles and Practice, S. 1-6.
- Ludwig, Michaela; Mandl, Thomas (1997): Ähnlichkeit von Werkstoffen: Die Anwendung unterschiedlicher Wissensmodellierungstechniken für eine intelligente Komponente von WING. In: Krause/Womser-Hacker 1997. S. 169 - 184.
- Mandl, Thomas (1996): Eine Fuzzy-Retrieval Komponente in einem Wirtschaftsinformationssystem. In: Anwendungen der Intelligenten Datenanalyse; Realisierungen mit DataEngine, Frankfurt am Main. Aachen: MIT-Management Intelligenter Technologien, 1996. S. 187-194.
- Mandl, Thomas; Womser-Hacker, Christa (1995): „Softcomputing“-Verfahren zur Behandlung von Ähnlichkeit und Vagheit in objektorientierten Informationssystemen. In: Kuhlén, Rainer; Rittberger, Marc (Hrsg.) (1995): HIM'95. Hypertext, Information Retrieval,

- Multimedia. Synergieeffekte elektronischer Informationssysteme. Konstanz. 5.-7. April 95. Schriften zur Informationswissenschaft, Konstanz: Universitätsverlag. S. 277-292.
- Marx, Jutta; Mutschke, Peter; Schommler, Marcus (1995): Möglichkeiten der intelligenten Integration heterogener Datenbestände: das Projekt GESINE. IZ-Arbeitsbericht Nr. 2, IZ Sozialwissenschaften, Bonn.
- Medina, J.M.; Pons, O.; Vila, M.A. (1994): GEFRED. A Generalized Model of Fuzzy Relational Databases. In: Information Sciences, vol. 76, 1-2. S. 87-109.
- Mitchell, Jeffrey; Shneiderman, Ben (1989): Dynamic versus Static Menus: An Exploratory Comparison. In: SIGCHI Bulletin 20(4). S. 33-37.
- Pepperrell, Catherine (1994): Three-Dimensional Similarity Searching. Taunton, Somerset, England.
- Preece, J.; Rogers, Y.; Sharp, H.; Benyon, D.; Holland, S.; Carey, T. (1994): Human-Computer Interaction. Workingham et al.
- Rabitti, F.; Savino, P. (1990): Retrieval of Multimedia Documents by Imprecise Query Specification. In: Bancilhon, F.; Thanos, C.; Tsichritzis, D. (eds.): Advances in Database Technology - EDBT'90. Berlin et al. S. 203-218.
- Reichenberger, Klaus; Rondhuis, Klaas; Kleinz, Jörg; Bateman, John (1996): Effective Presentation of Information Through Page Layout: a Linguistically-based Approach. Arbeitspapiere der GMD 970. Januar 1996. Darmstadt.
- Riege, Udo; Schomisch, Siegfried; Schommler, Marcus (1996): COGET - Computergestützte Erstellung von Themendokumentationen. IZ-Arbeitsbericht Nr. 9, IZ Sozialwissenschaften, Bonn.
- Rose, D.; Belew, Richard (1991): A Connectionist and Symbolic Hybrid for Improving Legal Research. In: International Journal of Man-Machine-Studies. vol. 31, 1. S. 1-33.
- Spoerri, A. (1994): InfoCrystal: Integrating Exact and Partial Matching Approaches through Visualization. In: Proceedings of the RIAO 94 (Recherche d'Information assistée par Ordinateur). Rockefeller University. New York. S. 687-696.
- Stempfhuber, Maximilian (1996): Intelligente grafische Informationssysteme und ihre Realisierung mit 4GL-Werkzeugen. In: Krause, Jürgen; Herfurth, Matthias; Marx, Jutta (Hrsg.): Herausforderungen an die Informationswirtschaft: Informationsverdichtung, Informationsbewertung und Datenvisualisierung. S. 227-238.
- VDMA (1997): Marktinformationssystem des VDMA. Informationsbroschüre des VDMA, Frankfurt/M.
- Wolff, Christian (1997): Graphisches Retrieval in WING-IIR. In: Krause/Womser-Hacker 1997. S. 113-144.
- Womser-Hacker, Christa (1996): Das MIMOR-Modell. Mehrfachindexierung zur dynamischen Methoden-Objekt-Relationierung im Information Retrieval. Habilitationsschrift. Universität Regensburg, Informationswissenschaft.

Womser-Hacker, Chista (1997): FUZZY-WING: Ein Werkzeug für Faktenabfragen mit vagen Kriterien. In: Krause/Womser-Hacker 1997. S. 185-204.

Yager, Ronald; Larsen, Henrik (1993): Retrieving Information by Fuzzification of Queries. In: Journal of Intelligent Information Systems, vol. 2. S. 421-441.

Liste der Arbeitsberichte im Projekt ELVIRA

ELVIRA I

- | | | |
|---|-------------------------------------|--|
| 1 | Mandl | Anforderungen und Benutzereigenschaften für das ZVEI-Verbandsinformationssystem (Juni 1995) |
| 3 | Krause/Stempfhuber/
Mandl | Design des ersten Prototypen des ZVEI-Verbandsinformationssystems ELVIRA (April 1996)
2. Vers.: Dez. 96; entspricht IZ-Arbeitsbericht 4 |
| 4 | Stempfhuber | Realisierung der Benutzungsoberfläche von ELVIRA mit Power-Builder
ersetzt durch ELVIRA-Arbeitsbericht 8 |
| 5 | Krause | Das WOB-Modell. Zur Gestaltung objektorientierter, grafischer Benutzungsoberflächen (Dezember 1995)
entspricht IZ-Arbeitsbericht 1 |
| 6 | Mandl/Stempfhuber | Die Benutzertests zur ersten Version von ELVIRA (Dezember 1996) |
| 7 | Mandl | Vages Information Retrieval und die Verarbeitung unsicheren Wissens (Januar 1997)
gekürzt als Kapitel im Endbericht |
| 8 | Stempfhuber | Generalisierung des 3-Facettenzugangs (1997) |
| 9 | Stempfhuber/Mandl/
Gräßer/Hermes | Beschreibung der Benutzungsoberfläche von ELVIRA (Dezember 1996) 2. Vers.: Juli 1997; entspricht Handbuch; auch Anhang Endbericht |

ELVIRA II

- | | | |
|----|--------------------------------|--|
| 11 | Krause/Mandl/
Stempfhuber | Das Verbandsinformationssystem ELVIRA II. Projektskizze (April 1997)
aus Projektantrag; entspricht IZ Arbeitsbericht 11 |
| 12 | Krause/Mandl/
Stempfhuber | Text-Fakten-Integration in ELVIRA (Dez. 1997)
entspricht IZ Arbeitsbericht 12 |
| 14 | Mandl/Stempfhuber | Ergebnisse aus der Praxisphase von ELVIRA (Dezember 1997) |
| 15 | Mandl/Schaefer/
Stempfhuber | Exemplarische Transformationen für die Text-Fakten-Integration (Februar 1998) |
| 16 | Krause/Schaefer | Benutzungsoberfläche für Textretrieval [Arbeitstitel] (1998) geplant |
| 17 | Krause/Mutschke | Indexierung und Fulcrum Evaluierung [Arbeitstitel] (1998) geplant |
| 18 | Stempfhuber | ELVIRA II im Internet [Arbeitstitel] (1998) geplant |

In der Reihe **IZ-Arbeitsberichte** sind bisher erschienen:

- 1 Krause, J.: Das WOB-Modell. Zur Gestaltung objektorientierter, graphischer Benutzungsoberflächen (Dezember 1995)
- 2 Marx, J.; Mutschke, P.; Schommler, M.: Möglichkeiten der intelligenten Integration heterogener Datenbestände: das Projekt GESINE (Dezember 1995)
- 3 Krause, J.: Visualisierung und graphische Benutzungsoberflächen (Mai 1996)
- 4 Krause, J.; Mandl, T.; Stempfhuber, M.: Design des ersten Prototypen des ZVEI-Verbandsinformationssystems ELVIRA (April 1996, vergriffen)
- 5 Mutschke, P.: Relationale Datenbanksysteme im Vergleich: Eine Zwischenbilanz (Dezember 1995)
- 6 Krause, J.: Informationserschließung und -bereitstellung zwischen Deregulation, Kommerzialisierung und weltweiter Vernetzung - Schalenmodell - (September 1996)
- 7 Weingärtner, M.: Computergestützte Gruppenarbeit: State of the Art (September 1996)
- 8 Binder, G., Herfurth, M., Krause, J., Marx, J., Stahl, M.: Nutzerforschung, Gestaltung informationeller Abläufe und Softwareentwicklung am Informationszentrum Sozialwissenschaften (Dezember 1996)
- 9 Riege, U.; Schomisch, S.; Schommler, M.: COGET - Computergestützte Erstellung von Themendokumentationen (Dezember 1996)
- 10 Frisch, E.; Kluck, M.: Pretest zum Projekt German Indexing and Retrieval Testdatabase (GIRT) unter Anwendung der Retrievalsysteme Messenger und freeWAISsf (Juli 1997/Nachauflage Oktober 1997)
- 11 Krause, J.; Mandl, T.; Stempfhuber, M.: Das Verbandsinformationssystem ELVIRA II. Projektskizze (April 1997)
- 12 Krause, J.; Mandl, T.; Stempfhuber, M.: Text-Fakten-Integration in ELVIRA (Dezember 1997)