
online 'classic'

Datenbankrecherchen im 'Expertenmodus'

[Gerhard Fröhlich](#)

gerhard.froehlich@jku.at

© 1998

[STARTSEITE](#)

Gliederung

[1.](#) Internet-Odysseen oder gezielte (Fach-) Informationsgewinnung? Digitale Schrebergärten und professionelle Datenbanken (eine polemische Einleitung)

[1.1](#) Mangel im Überfluß? Wege aus der wissenschaftlichen Informationskrise

[1.2](#) Der Nutzen konventioneller Information und Dokumentation

[1.3](#) Online-Recherchen in Datenbanksystemen

[2.](#) Inhalte von Datenbasen

[2.1](#) Referenzinformationen

[2.2](#) Primärinformationen

[3.](#) Sozial- bzw. kulturwissenschaftlich relevante Hosts und Datenbasen

[4.](#) Suchfunktionen am Beispiel der Kommandosprache CCL (Common Command Language)

4.1 Elementare Suchfunktionen

4.2 Zugriffsmöglichkeiten über Datenfelder

4.3 Verknüpfungen mittels logischer Operatoren

4.4 Die Syntax von CCL/GRIPS

4.5 Maskierung und Kontextbestimmung

4.6 Die Kosten von Recherchen

4.7 Die Optimierung von Recherchestrategien

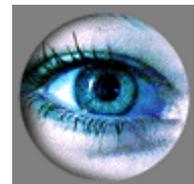
5. Schlußbemerkung

6. Literatur

7. Anhang

8. Aktualisierungen

8.1 ECHO



1) Internet-Odyseen oder gezielte(Fach-) Informationsgewinnung? Digitale Schrebergärten und professionelle Datenbanken (eine polemische Einleitung)

Online-Datenbanken (d. h. Dateien mit Zehntausenden, ja Millionen von Dokumenten oder Dokumentnachweisen, welche über Daten- bzw. Telefonleitungen abrufbar sind) gibt es schon lange. Sie fristen bis heute ein eher auf Fachkreise beschränktes, relativ bescheidenes Dasein. Alle von einschlägigen Institutionen und Unternehmen prognostizierten Wachstumsraten haben sich als viel zu optimistisch erwiesen.

Hingegen sind einige Internet-Dienste, allen voran das sog. *World Wide Web (WWW)*, zum Medienereignis und Massenphänomen geworden - zum Symbol der sog. "Informationsgesellschaft". Mit diesen "Netzen" bzw. "Netzwerken" werden zahlreiche sozialutopische und sozialromantische Hoffnungen verbunden - so die Demokratisierung

der Gesellschaft, die Abschaffung jeglicher sozialer Ungleichheit, ja Unsterblichkeit. (1)

Diese Mißverständnisse und zGtl. trügerischen Hoffnungen sind sowohl auf sensationslüsterne Medienberichte (einerseits euphorisch, andererseits dämonisierend - das Internet, voll von Kinder pornos, Neonazis, Sekten und digitalen Betrügern) als auch auf zahlreiche unkritische und mangelhaft informierte Einführungsbücher zurückzuführen. Doch selbst viele Informatiker und andere Fachleute sind hier nicht selten nur mangelhaft informiert.

Einerseits gibt es diverse keineswegs oder nur gegen hohe Gebühren für "Normalbürger" zugängliche Netze (z. B. militärische, Finanzinformationsdienste, Großrechnerverbünde). Das Internet ist bloß das in der Öffentlichkeit bekannteste Computernetz, und dasjenige, in dem manfra als "Normalbürger" die (Illusion von) Eigenaktivität haben kann. (2)

Entgegen vielfacher Behauptung ist das Internet mit allen seinen Diensten keineswegs in allen seinen Sektoren öffentlich zugänglich (d. h. ohne Paßwörter) und schon gar nicht kostenlos. Medienberichte und Internet-Bücher, -Zeitschriften etc. konzentrieren sich fast ausschließlich auf die öffentlich und (abgesehen von Pornoangeboten im WWW, und abgesehen von Telefon- und sonstigen Datenübertragungskosten inkl. Providergebühren) kostenfrei zugänglichen Angebote, welche zGtl. unter die Rubrik *"p.r."/Werbung /Infotainment* oder *"digitale Schrebergärten"* fallen.

Was ist damit gemeint? Die kostenfrei zugänglichen Internetangebote, insbesondere im aufwendigen und daher (wenn professionell gestaltet) kostenintensiven WWW, dienen konventionellen Medien, Unternehmen, Institutionen zur Werbung (für ihre sonstigen konventionellen Angebote), zur p.r. (public relations)- bzw. Fan clubs-Arbeit oder - wie z. B. die Search Engines - finanzieren sich zumindest über Werbung. Für den Wissenschaftsalltag sind diese Angebote nur teilweise brauchbar - etwa, wenn man sich über die offiziellen Selbstdarstellungen diverser soziologischer Institute interessiert. Zudem ist die Odyssee durch die Millionen Webseiten (deren Adressen sich zudem häufig ändern) selbst mit Unterstützung sog. Search Engines äußerst zeitraubend.

Ein weiterer großer Anteil an WWW-Seiten basiert auf ehrenamtlicher, sportlich-ehrgeiziger "Selbstaubeutung": Jeder Verein, ja jede Privatperson kann hier mit vergleichsweise ganz bescheidenen finanziellen Mitteln "publizieren": Was früher als Schülerzeitung oder Vereinsmitteilung in einigen Hundert Exemplaren hektographiert wurde, wird heute "ins Netz" gestellt - nicht selten allerdings mit ähnlich geringer Resonanz wie diese fast schon ausgestorbene Publikationssorte.

Daneben gibt es diverse nichtöffentliche Sektoren im Internet. Manche Informationen (z. B. astronomische Fotos) sind eine Zeitlang (hier: 1 Jahr) gesperrt und in dieser Zeit nur für Mitglieder zugänglich, oder sie bleiben es. Die schon lange vor der Internet-Euphorie aufgebauten Online-Datenbanken sind zGtl. ebenfalls über das Internet (und über sonstige Datenleitungen, wie ISDN, Datex-P etc.) zugänglich. Doch erfordert die Recherche in ihnen einen Nutzervertrag mit einem Host (=Wirtsrechner, welcher die Datenbanken, in der Regel mehrere Hundert verschiedene Datenbanken) anbietet, und zur gezielten, effizienten und - da kostenpflichtig - kostengünstigen Recherche gewisse Kenntnisse.

Während in den "digitalen Schrebergärten" viele links im Nichts enden und durch warteintensives Weiterklicken, aufwendigen Bildaufbau etc. durchaus zum Einüben von Geduld und Meditation pädagogisch wertvoll sein mögen, letztlich aber immer nur den Zugriff auf einzelne Hinweise oder Texte ermöglichen (und auch eine Trefferliste von

Search Engines muß aufwendig einzeln angeklickt werden), offerieren Datenbanken, welche kosten-, d. h. personalintensiv aufgebaut wurden, 100 000ende, oft Millionen Dokumente bzw. Dokumentnachweise, und insbesondere zu allen wissenschaftlichen Gebieten. In ihnen sind daher Recherchen, wie im weiteren demonstriert werden soll, wesentlich rascher (bei einiger Übung um ein Vielfaches) und gezielter möglich. Zur Gewinnung von Übersichtsinformation ist auch die statistische Auswertung großer Treffermengen möglich.

Damit soll nicht geleugnet werden, daß die konventionelle Online-Datenbank-Technologie von den neuen WWW- und sonstigen Internet-Entwicklungen lernen könnte und sollte. Doch jede Abweichung vom "minimalistischen" Standard (=ASCII-Text) (3) muß mit Verlangsamung, geringerem Recherchekomfort etc. bezahlt werden.

Entgegen vielfach zu lesender/hörender Vorurteile sind übrigens WWW, news groups und e-mail nicht die einzigen Internet-Dienste. So ist der *Gopher* (= der schlichte Vorgänger des WWW, in der Regel nur Text, ohne Firlefanz, in der Informationsszene - ungerechtfertigterweise - "mega-out") für die digitale "Reise" in osteuropäische oder Dritte-Welt-Länder durchaus noch immer zu empfehlen. So gibt es im Gopher z. B. kubanische elektronische Zeitschriften zu lesen - oder die Reden von Nelson Mandela auf dem Gopher des ANC.

Bei der Formulierung des folgenden Skriptums wird "Fachchinesisch" aus dem Bereich der Informatik und der wissenschaftlichen Information und Dokumentation so weit wie möglich vermieden. Es soll bewußt Teilnehmern *mit geringen EDV- und Dokumentationskenntnissen bzw. -Erfahrungen* einen Einstieg ermöglichen. Daher wird - Fortgeschrittene sollten diese Passagen überspringen - oft auch "Selbstverständliches" erläutert. Die Literaturliste im Anhang soll sowohl auf einführende Literatur als auch auf fortführende Literatur verweisen. Diese ist z. T. bereits einige Jahre alt. Aufgrund der immer kürzeren "Halbwertszeit" der wissenschaftlichen Literatur insbesondere auf diesem Gebiet steht sie bereits unter dem Verdacht, völlig veraltet zu sein. Doch hat sich in der Logik der Datenbankrecherche im Expertenmodus und auch bei den großen professionellen Datenbanken weitaus weniger Fluktuation und rasant-zielloser Fortschritt gezeigt, sodaß etwa Überlegungen zum Gebrauch von logischen Junktoren (und, oder, nicht) noch immer Gültigkeit haben. Andererseits demonstriert die dokumentierte Literatur zur Frühzeit der EDV-gestützten IuD (= Information und Dokumentation) die Rasanz der Entwicklung. Vor allem zeigt sich hier, daß die eher aus dem traditionellen Bibliotheks- und Dokumentationssektor kommende IuD (welche letztlich klassische Nachschlagewerke in Online- bzw. CD-ROM-Datenbanken transferierte) sich weitaus langsamer entwickelte (was nicht immer zum Nachteil war) und weitaus geringere Verbreitung fand als die aus der Computerbranche bzw. Informatik kommende Internet-Technologie bzw. Internet-Dienste. (4)

Gleichzeitig soll ein *nicht bloß technischer Zugang* zur Thematik versucht werden, sondern sowohl theoretische und empirische Aspekte aus Wissenschaftstheorie und Wissenschaftsforschung als auch persönliche Erfahrungen einfließen. Denn voluminöse offizielle Handbücher und bunte Glanzpapierprospekte der Anbieter von Informationsdienstleistungen gibt es bereits genug. Sie weisen nicht selten eine große Diskrepanz zur Alltagspraxis auf, die oft an Kleinigkeiten scheitert. Der informelle Erfahrungsaustausch mit erfahrenen Kolleginnen (die übrigens in der Informationsszene häufiger als in anderen wissenschaftlichen Bereichen zu finden sind) und Kollegen ist daher auch auf diesem Gebiet unverzichtbar. Manfrau sollte sich auch niemals scheuen, die jeweils eingerichteten telephonischen Hilfsdienste der Datenbanksysteme ("help desks") zu konsultieren. Allerdings wissen auch diese oft keinen - oder nur falschen - Rat.

Es erschien mir jedoch sinnvoller, ein *realistisches Bild* zu zeichnen - was oft vage Formulierungen erforderte, da es an verbindlicher Standardisierung selbst innerhalb einzelner Datenbanksysteme mangelt - als die bunte weite Welt der Datenbankprospekte zu reproduzieren, die "*Wissen auf Knopfdruck*" versprechen, bei hoffnungsfrohen neuen Nutzern hohe Erwartungen wecken, insbesondere Komfort und Bequemlichkeit betreffend, und so mitunter bittere Enttäuschungen produzieren. Trotz aller Enttäuschungen, Irr- und Umwege (und manchmal horrender Rechnungen) hat mir jedoch das Jagen, Sammeln, Fischen in Datenbanksystemen immer sehr viel Spaß und Spannung bereitet. Das wünsche ich auch den Leserinnen und Lesern dieses Skriptums bei ihren ersten praktischen Erfahrungen.

1.1 Mangel im Überfluß? Wege aus der wissenschaftlichen Informationskrise

Intersubjektivität gilt in allen heute vertretenen Wissenschaftstheorien als wesentliches Kriterium für Forschung und Wissenschaft. Experten der Wissenschaftskommunikation befürchten nun, daß die "wissenschaftliche Informationsexplosion" zu "Informationsstreß" und zu einer "wissenschaftlichen Informationskrise" führt. Die alltägliche Wissenschaftspraxis kann stichwortartig wie folgt charakterisiert werden: wachsende Zahl von Wissenschaftlern, wachsende Zahl von Disziplinen, wachsender Output an Pflichtpublikationen, aufgehäuft als Summe der individuellen Karriereanstrengungen. Ergebnis dieses Wildwuchses wissenschaftlicher Produktivität ist aber, wie man aus der Praxis weiß, oft nicht ein Mehr an Kommunikation, sondern vielfach zunehmende Unübersichtlichkeit und von seiten des Individuums *Informationsverdrossenheit* und Rückzug aus der wissenschaftlichen Kommunikation. Gleichzeitig mit der Informationsüberlastung zeigen sich jedoch auch Chancen der modernen Informationstechnik und Informationswissenschaften, mit den genannten Problemen fertig zu werden: Fachinformationssysteme ermöglichen die systematische Beschaffung und Reduktion von Information.

Die *Recherche* in Fachinformationssystemen ist metaphorisch mit dem Fischen in mehr oder weniger bekannten Gewässern zu vergleichen: Es sind "Netze" nötig, um die "Fische" einzuholen. (5) Die "Netze" bestehen aus Suchtermen, die mit logischen Junktoren (und/oder/nicht) und anderen Operatoren verknüpft werden. Als Suchterme können Wörter aus dem Titel, aus den Abstracts von Publikationen verwendet werden, ja sogar zitierte Autoren bzw. Publikationen: Es ist etwa über die Datenbank "Social Sci Search" möglich, in Minutenschnelle zu eruieren, wer in der internationalen Wissenschaftergemeinschaft im Jahre 1985 die "General Theory" von John Maynard Keynes in seiner wissenschaftlichen Publikation verwertet hat.

Die Angebote der Fachinformationssysteme beschränken sich nicht auf Literaturdokumentation (*Referenzinformationen*) im klassischen Sinn. Projektdatenbanken liefern detaillierte Informationen zu abgeschlossenen, laufenden und geplanten Forschungsprojekten; Datenarchive stellen die Datensätze von archivierten Befragungen Interessenten für Sekundäranalysen zur Verfügung. Die sogenannten Faktendatenbanken bieten *Primärinformationen* an: So gibt es etwa Terminologiedatenbanken, Wirtschaftsinformationsdatenbanken (z. B. zu Unternehmen und Produkten), graphische Datenbanken chemischer Strukturformeln, statistische Datenbanken, juristische Informationssysteme (zu Gesetzen, Verordnungen, Verwaltungsvorschriften, Rechtsprechung), sowie Volltextbanken, die in Form des sogenannten "Electronic Publishing" Zeitschriften und Forschungsarbeiten zum Ausdruck auf individuellen Abruf hin bereithalten - eine zukunftsweisende, ressourcensparende Publikationsform. So kann z. B. ein Ingenieur ein Ergebnis, das er einmal in den VDI-Nachrichten (Verein Deutscher

Ingenieure) gelesen hat, schnellsten ohne großen Aufwand und *auf Basis bruckstückhafter Erinnerung* wiederfinden. Es reichen dazu z. B. zwei bis drei thematisch einschlägige Wörter (ja Wortfragmente) aus dem Text des Artikels.

Die Möglichkeiten der Fachinformationssysteme sind noch nicht ausgeschöpft: Einerseits werden laufend neue Datenbanken aufgebaut und wird an der Entwicklung von benutzerfreundlicherer Retrieval-Software gearbeitet, andererseits sollte durch die Integration von Theorie und Praxis der Fachinformation in die Allgemeinbildung und damit in alle Studienrichtungen auch von Seiten der Einzelwissenschaftler her die nötigen Kompetenzen zur direkten Nutzung der Datenbanksysteme erweitert werden.

Die gespeicherten Rechercheergebnisse können direkt weiterverarbeitet werden. Auch nichtnumerische Rechercheresultate können nach einer Vielzahl von Kriterien sortiert und statistisch analysiert werden, mit Programmen aus der Scientometrie, der maschinellen Inhaltsanalyse, der linguistischen Datenverarbeitung. Mit Zitations-Netzwerkanalysen können die formalen Kommunikationsströme auch graphisch dargestellt und gleichsam "wissenschaftliche Landkarten" erstellt werden. Die - nicht unumstrittenen - EDV-gestützten Verfahren der sogenannten *Meta-Analytik* (darunter versteht man Methoden der statistischen Zusammenfassung von unabhängig voneinander gewonnenen Ergebnissen) ermöglichen eine gewaltige Informationsreduktion in Form der Integration Hunderter Einzelergebnisse.

Zweifellos kann durch die moderne Informationstechnik das konstruktive Denken nicht ersetzt werden. Vielmehr ist ein enormer Aufwand an Fachwissen und konzeptioneller Anstrengung erforderlich, um die Potentiale der EDV-gestützten Fachinformationssysteme für die eigene konkrete Fragestellung nützen zu können.

Wir wissen, daß Information in der sich abzeichnenden Informationsgesellschaft eine wichtige ökonomische, soziale und intellektuelle Ressource sein wird. Die Probleme, die sich hieraus für wissenschaftliche Arbeit und wissenschaftliche Organisation ergeben, sind Gegenstand neuer Entwicklungen in Wissenschaftstheorie und empirischer Wissenschaftsforschung. Dementsprechend sollte in der Lehre in allen Studiengängen Platz sein für (1) eher *praktisch orientierte Lehrveranstaltungen* (praktische Fragestellungen der Wissenschaftskommunikation, Benutzung von Internet-Diensten, e-mail etc. sowie von Fachinformationssystemen und Möglichkeiten der Informationsverdichtung bzw. -reduktion - Stichwort: "*Wissenschaftstechnik*"), (2) aber auch (nicht zuletzt im Rahmen der philosophisch-wissenschaftstheoretischen Wahlfächer) für die Vermittlung reflexiver, *theoretischer Grundlagen*, wobei insbesondere darauf abgestellt werden sollte, die bereits zahlreich vorliegenden Ergebnisse aus dem angelsächsischen Wissenschaftsbereich für den deutschen Sprachraum nutzbar zu machen - Stichwörter: Wissenschaftswissenschaften ("science of science"), Informationswissenschaften, sowie der kritischen kulturphilosophischen Reflexion (Mythos Informationsgesellschaft, Internet als Sozialutopie, Neue Kulturtechniken etc.)

1.2 Der Nutzen konventioneller Information und Dokumentation

Die *konventionelle* Information und Dokumentation, die im Zeitalter der EDV-gestützten Fachinformationssysteme an relativer Bedeutung verliert, hat eine lange Tradition. Ihre Produkte sind Ordnungssysteme (etwa die Dezimalklassifikationen), Thesauri (strukturierte, meist hierarchisierte Wörterbücher kontrollierter Terminologie), Karteien,

Referatedienste, *Nachschlagewerke* verschiedenster Art. Eine der ältesten deutschsprachigen, auch für die Sozial- und Wirtschaftswissenschaften relevanten Nachschlagewerke ist die Internationale Bibliographie der Zeitschriftenliteratur (IBZ) und die mit ihr "verwandte", vom selben Verlag herausgegebene Internationale Bibliographie der Rezensionen (IBR). Der Umgang mit Nachschlagewerken ist meist etwas umständlich und erfordert oft längere Einschulung und Übung; daher haben sich in der Vergangenheit in der Regel nur Bibliothekare und Dokumentare mit konventioneller Information und Dokumentation befaßt. Der Großteil der Studierenden und Lehrenden hat sich leider nur selten mit den konventionellen Methoden der Informationsbeschaffung befaßt.

Für Sozial- und Wirtschaftswissenschaftler relevant sind insbesondere die vom Informationszentrum Sozialwissenschaften (Bonn) sowie die vom Institut für Arbeitsmarkt- und Berufsforschung (IAB) der Bundesanstalt für Arbeit herausgegebenen Forschungs- und Literaturdokumentationen (vgl. [Übersicht 1](#)).

Obwohl sich das vorliegende LV-Skriptum mit der Informationsgewinnung aus Datenbanken, d. h. EDV-gestützten Fachinformationssystemen befaßt, erfolgen des öfteren Verweise auf konventionelle Nachschlagewerke. U.a. können durch die sorgfältige *Vorbereitung* von Recherchen in Datenbanksystemen mittels Konsultierung konventioneller Nachschlagewerke vermeidbare Umwege und Fehlschläge oft vermieden und damit die Kosten von Recherchen beträchtlich gesenkt werden: schlecht vorbereitete Recherchen werden in der Informationsszene als "*quick and dirty*" bezeichnet. Manchmal kann auch ein einziger Blick in ein Nachschlagewerk eine kostspielige Datenbankrecherche ersetzen. Überdies wird bei retrospektiven Recherchen in die Vergangenheit weiterhin auf konventionelle Zugänge zu setzen sein, denn es ist kaum anzunehmen, daß in absehbarer Zeit der immense Aufwand weltweiter Digitalisierung alter Bücher-, Zeitschriften-, Aktenbestände finanziert werden wird.

1.3 Online-Recherchen in Datenbanksystemen

Der erste Teil des Skriptums (Pkt. [2](#) und [3](#)) soll eine Einführung in die Welt der Fachinformationssysteme bieten. Im zweiten Teil des Skriptums (Pkt. [4](#)) werden die Suchfunktionen und Zugriffsmöglichkeiten beim *Information Retrieval* (Informationsrückgewinnung) aus Datenbanksystemen erläutert. Die Beispiele werden dabei in der Kommandosprache CCL/GRIPS (Common Command Language) ausgeführt.

Im folgenden befassen wir uns mit *Online-Recherchen* in Datenbanksystemen. Manche Datenbasen sind jedoch nicht über Datennetze zugänglich. Diese Datenbasen können manchmal immer noch nur indirekt, über postalische (oder telefonische) Suchanfragen an bestimmte Informationsvermittlungsstellen konsultiert werden oder sind in Form von Disketten oder - in den letzten Jahren immer häufiger - als CD-ROMs ([6](#)) käuflich erwerbbar. Diese können dann mit Hilfe von Retrievalprogrammen auf dem eigenen PC recherchiert werden; die dabei erforderlichen Suchfunktionen unterscheiden sich prinzipiell nicht von denen bei Online-Recherchen; allerdings entfällt der Zeitdruck (da die Recherchedauer keinen Kostenfaktor darstellt) und insofern auch ein Optimierungszwang.

Was benötigen wir zur Durchführung einer Online Recherche in einem EDV-gestützten Fachinformationssystem? Wir brauchen

- einen *PC bzw. Mac* mit entsprechender Software, eine *Terminal-Emulation* (z. B.

MacTerminal) bzw. Netzwerksoftware (telnet, Gopher, WWW: sog. "browser", welche das Navigieren im Netz ermöglichen, z. B. netscape),

- einen (ganzen) *Telefonanschluß* oder einen ISDN (=Breitbandanschluß) mit einem *Modem*, um die Verbindung zwischen unserem PC und den internationalen Datennetzen herstellen zu können (oder ein Funkmodem), oder (ganz veraltet) einen Akustikkoppler (nur für Übertragung von Text in geringer Quantität noch zu empfehlen), (7)

- *Benutzerverträge* mit der Post bzw. einem Internet-Provider und den gewünschten *Hosts* (Wirtsrechnern, d. h. den Datenbanksystemen mit unterschiedlichen inhaltlichen Schwerpunkten); nach Abschluß der Verträge werden uns *Paßwörter* (alphanumerische Zeichenkombinationen) mitgeteilt. Diese sollten sorgsam verwahrt werden - wenn sie in fremde Hände fallen, müssen wir (bis zu einer Sperrung) i. d. R. für die angefallenen Benutzungsgebühren aufkommen.

- Die Kenntnis von *Kommandosprachen* ist inzwischen nur mehr bei wenigen Hosts erforderlich. Viele bieten neben dem sog. *Expertenmodus*, bei dem die Recherche mit Hilfe einer Kommandosprache durchgeführt wird, die Möglichkeit *menügestützter* Recherchen an, die keinerlei Vorkenntnisse erfordern ("*Laienmodus*"). Damit sind jedoch verschiedene Nachteile verbunden, nicht zuletzt eine Verlangsamung der Recherchen. Eine menügestützte Recherche ist mit dem Bestellvorgang in einem Restaurant zu vergleichen, bei dem der Kellner vor jedem Gang beharrlich die Speisekarte erläutert, obwohl wir bereits genau wissen, welche Speisen wir zu bestellen wünschen. Weitere Nachteile: Es ist nur ein Teil der Funktionen des Expertenmodus ansprechbar. Die Recherchevorgänge sind für die NutzerInnen wenig transparent.

Da die Kommandosprachen sehr einfach sind (und nicht mit den weitaus komplizierteren und umfangreicheren *Programmiersprachen* verwechselt werden dürfen), ist es ratsam, zumindest eine oder zwei dieser Kommandosprachen zu erlernen. Manche Informationsvermittlungsagenturen offerieren gegen entsprechende Gebühren *Übersetzungsprogramme*, die von einer Retrievalsprache in eine Vielzahl anderer übersetzen. Meist erübrigt sich durch einen Vertrag mit einem solchen "Gateway" der Abschluß einer ganzen Reihe von Verträgen mit einzelnen Hosts, allerdings ist diese Aufwandsersparnis nicht kostenlos.

Leider herrscht auch im Bereich der Fachinformation - wie in vielen anderen Disziplinen - terminologische Anarchie, es mangelt an bedeutungsmäßig eindeutig festgelegten Fachtermini. In diesem Skriptum wird der Sprachgebrauch übernommen, zwischen *Datenbasen* (d. h. den einzelnen Dateien, die zu einem bestimmten Inhalt Daten gespeichert haben) und *Datenbanken* bzw. Datenbanksystemen als der übergreifenden Organisationsform zu unterscheiden (in einem Teil der Literatur wird hingegen auch die einzelne Datenbank als Datenbank bezeichnet).

Datenbasenproduzenten und Datenbasenanbieter sind meist nicht identisch, ein und dieselbe Datenbank wird oft von verschiedenen Hosts (Wirtsrechnern) angeboten. Benutzerverträge schließen wir jedoch nicht mit den einzelnen Datenbasenproduzenten (z. B. dem Informationszentrum Sozialwissenschaften, Bonn) ab, sondern mit den Hosts, den "Supermärkten des Wissens", die uns den Zugang zu mehreren oder sogar hunderten Datenbasen mittels eines einzigen Benutzervertrages und damit eines einzigen Paßwortes ermöglichen. Wir müssen uns also darüber orientieren, *welche Datenbasen* zu unseren Fachgebieten bisher aufgebaut wurden und bei welchen Hosts sie angeboten werden. Mit den entsprechenden Hosts müssen wir Benutzerverträge abschließen, um Paßwörter zu erhalten.

Übersichtsinformationen zur Informationslandschaft erhalten Sie einfach und regelmäßig durch Mitgliedschaft in *Fachvereinigungen* - etwa in der mit zahlreichen Tagungen, Ausstellungen, Informationsdiensten aktiven *DGD* (Deutsche Gesellschaft für Dokumentation)(8). Eine *Online User Group* befaßt sich speziell mit den Problemen bei Online-Recherchen in Datenbanksystemen. Die DGD gibt die zweimonatliche Zeitschrift "*Nachrichten für Dokumentation*" heraus, die neben vielen aktuellen Meldungen (z. B. Implementierung neuer Datenbanken) auch laufend die neueste Literatur zum Thema (z. B. Datenbankrezensionen) dokumentiert. Die Österreichische Gesellschaft für Dokumentation und Information (ÖGDI) mit Sitz in Wien ist leider nur sehr mäßig aktiv, zusammen mit anderen österreichischen Organisationen wird fallweise die Zeitschrift "Fakten Daten Zitate" herausgegeben, in der v. a. österreichische Produkte und Dienstleistungen vorgestellt werden, aber auch theoretische Beiträge zum Thema Information Retrieval und Informationsgesellschaft.

Als Besucher einschlägiger Tagungen und Ausstellungen (z. B. der "Dokumentartage" oder der "Infobase") wird man überdies in der Folge mit einer Vielzahl von Werbesendungen aus der Fachinformations-Szene überhäuft. Auch im WWW finden sich viele Werbeseiten der Hosts (s. die WWW-Adressen in den Übersichten), in der Regel mit einem 'link' (= automatischer Verbindungsaufbau) zu den von ihnen angebotenen Datenbasen, wobei i. d. R. (Ausnahme z. B. ECHO) ein Nutzervertrag vorliegen muß, d. h. ein geheimes Paßwort eingegeben werden muß, auch wenn die Recherche selbst bereits im WWW-Modus stattfinden kann.

Neuankündigungen von Datenbanken finden sich auch in den papierenen P. R.-Organen (z. B. *grips-News* vom Institut für Medizinische Information und Dokumentation (DIMDI) in Köln, *ECHO Facts for User*, *chronolog* vom Informationskonzern Knight-Ridder für *Dialog* und *data star*), Detailinformationen (inkl. Datenfelder, Beispiele für Suchstrategien) in den *blue sheets* (Datenbank-Infos für registrierte Nutzer) der jeweiligen Hosts, die z. t. bereits zusätzlich kostenlos über das WWW zugänglich sind; mehr oder minder kritische Besprechungen (Datenbank-Rezensionen) finden sich in den zahlreichen Fachzeitschriften (z. B. *Information World Review*), vgl. auch Übersicht 2. - Themenspezifische Samm

lungen von WWW-Seiten werden regelmäßig im Techno-Teil der Zeitschrift *tv-media* veröffentlicht; aber auch auf den Wirtschafts-, Kultur- und Wissenschaftsseiten der gehobenen Tages- und Wochenpresse finden sich laufend (allerdings nicht immer gründlich recherchierte und kritische, sondern nicht selten eher vordergründig sensationalistisch-euphorische) Berichte und Ankündigungen von CD-ROMs und WWW-Seiten.

Experten schätzen die Zahl der weltweit bereits aufgebauten Datenbasen auf über 7000. Einen systematischen Überblick über recherchierbare Datenbasen (leider meist aufgrund der rasanten Entwicklung nicht auf dem neuesten Stand) bieten verschiedene *Nachschlagewerke* bzw. Einführungen (vgl. [Übersicht 2](#)), die *Datenbasenkataloge* einzelner Hosts, die - in Kurzversion - als Werbematerial kostenlos angefordert werden können, sowie *Datenbasen der Datenbasen* (vgl. Pkt. [2.1](#)).

Allerdings ist für die beschränkten Zwecke des normalen wissenschaftlichen Alltags meist die Konsultation einer kleinen Zahl von Datenbasen ausreichend. Aufgrund der Teilnehmerstruktur der LV wird der Schwerpunkt dieses Skriptums auf die wichtigsten sechs soziologischen und psychologischen Datenbasen gelegt, und zwar auf die deutschsprachigen Datenbasen *FORIS*, *SOLIS*, *Psyndex* sowie die internationalen Datenbasen *Social Sci Search*, *Sociological Abstracts* und *Psychnfo* (Datenbasenversion der "Psychological Abstracts").

Alle diese, weiter unten vorgestellten Datenbasen können über den Host DIMDI (mit Zugriff über Postnetze oder Internet: telnet, Gopher, WWW) mit der im Skriptum vorgestellten Kommandosprache CCL/ GRIPS recherchiert werden. Überdies können diese Datenbasen - abgesehen von FORIS - in CD-ROM-Versionen an vielen deutschsprachigen Universitätsbibliotheken kostenlos recherchiert werden. (9) Die Kommandosprache CCL hat überdies große Ähnlichkeit mit dem CD-ROM-Retrievalsystem "SPIRS", mit dem der Zugriff zu mehreren dieser CD-ROMs (PsyInfo, Psyindex) erfolgt.

2. Gespeicherte Inhalte von Datenbasen

Übersicht 3 zeigt, welche Inhalte in öffentlich zugänglichen Datenbasen gespeichert werden. Dabei müssen zweierlei Informationsarten unterschieden werden: Referenz- und Primärinformationen.

2.1 Referenzinformationen

Ein Großteil der Datenbasen enthält sog. *Referenzinformationen*: Die angebotene Information ist noch nicht die von uns gewünschte Information, z. B. der volle Text eines bestimmten Zeitschriftenaufsatzes, sondern bloß der *Verweis* auf diese (z. B. in Form bibliographischer Angaben). Verwiesen wird insbesondere auf Datenbasen, Forschungsprojekte, Literatur.

Die Referenzinformation enthält in der Regel eingehende orientierende Informationen (oft z. B. eine Kurzzusammenfassung im Ausmaß von einer halben bis zwei A4-Seiten), die uns bei der Entscheidung behilflich sein können, ob sich die Beschaffung des Originaldokuments, bzw. eine Kontaktaufnahme mit den jeweiligen Forschern für uns lohnt.

Datenbasen der Datenbasen enthalten Informationen über Datenbasen. In der Regel hat jeder große Host eine Datenbank bzw. eine Hilfsfunktion, in der die *eigenen* Datenbasen knapp (meist nur durch Angabe des inhaltlichen Schwerpunktes, des Produzenten, des erfaßten Zeitraums) vorgestellt werden.

Zusätzlich gibt es Datenbasen, die auch die Informationsdienste *fremder* Hosts erfassen. Die Datenbank I'MGUIDE (10) enthält z. B. Informationen über Datenbasen, die von Hosts im EG-Raum angeboten werden (allerdings auch außerhalb Europas produziert worden sein können). Diese Datenbasen werden (abgesehen von Datenübertragungsgebühren) kostenfrei vom Host der Europäischen Gemeinschaft ECHO in Luxemburg angeboten. Ihre Nachweise sind allerdings leider nicht unbedingt vollständig. ECHO offeriert auch im Rahmen dieser Datenbank (ebenfalls sehr lückenhafte) Hinweise auf öffentliche und private *Informationsbroker*, d. h. auf Informationsvermittler, die wir gegen Gebühr bzw. Honorar mit Recherchen beauftragen können (vgl. ebenfalls Abb. 1).

Projektdatenbasen verweisen auf geplante, gerade laufende oder bereits abgeschlossene Forschungsprojekte. Sie beruhen auf freiwilligen Meldungen von Forschern bzw. von Institutionen, in denen Projekte durchgeführt werden oder von denen Projekte gefördert/in

Auftrag gegeben werden. Die Datenbase FORIS [\(11\)](#), vor kurzem mit einem Preis ausgezeichnet, basiert auf einer einmal jährlich im deutschen Sprachraum durchgeführten Erhebung, die für die BRD vom *Informationszentrum Sozialwissenschaften* in Bonn und weiteren kooperierenden Einrichtungen, in Österreich und der Schweiz von nationalen Kooperationspartnern durchgeführt wird. Dabei werden an einschlägige Institutionen (z. B. an alle sozialwissenschaftlichen Universitätsinstitute) Fragebögen ausgesendet; es ist aber auch bereits eine Onlineversion verfügbar. Auch der *Bundesminister für Arbeit und Sozialordnung* führt Erhebungen durch, bezogen auf das Gebiet der Beschäftigung.

Abb. 2 zeigt die vom IZ Bonn verwendeten *Projekterfassungsbögen* in der Internet-Version. Erfragt werden die für ein Forschungsprojekt relevanten Informationen: Institution, in der das Projekt durchgeführt wird, Bearbeiter, Kontaktadressen, Stand, Auftraggeber/Finanzgeber, Titel, Zeitraum, geographischer Raum, inhaltliche Ziele, Veröffentlichungen, methodische Anlage, Stichprobenart und -größe. Wenn Sie selbst eines Tages eine Diplomarbeit bzw. später eine Dissertation erstellen, sollten Sie auch an dieser Projekterhebung teilnehmen. Fragebögen können Sie beim Informationszentrum Sozialwissenschaften, Lennéstr. 30, D-53113 Bonn (Tel. 0228/2281-0, e-mail: post@iz-bonn.gesis.d400.de) bzw. beim Bundesminister für Arbeit und Sozialordnung, 5300 Bonn-Duisdorf (Tel. 0228/5272969) anfordern. Auch an dieser Universität werden jedem Institut einmal jährlich Fragebögen übermittelt. Nicht selten landen diese Fragebögen jedoch im Papierkorb, weil viele Universitätsangehörige noch nicht die Bedeutung der Repräsentanz eigener Forschungsprojekte und Publikationen in Datenbanken erkannt haben. Auch Diplomanden und Doktoranden werden nur selten auf diese Erhebungen aufmerksam gemacht. Zu einem FORIS-Projektnachweis vgl. *Abb. 3*.

Die Projektnachweisdatenbasen sind zu unterscheiden von den *Datenarchiven* (vgl. z. B. das Zentralarchiv für empirische Sozialforschung, Köln). Datenarchive haben Datensätze von Befragungen archiviert und stellen diese gegen Gebühr auf Datenträgern (meist auf Magnetbändern), z. t auch über Wissenschaftsnetze für Sekundäranalysen zur Verfügung. [\(12\)](#)

Wissenschaftliche Zeitschriftenaufsätze sind mittlerweile in Form von Nachschlagewerken bzw. Informationsdiensten und von Datenbasen sehr gut erfaßt. In der Regel umfassen die Dokumentnachweise bibliographische Angaben, Schlagwörter und Abstracts (Kurzzusammenfassungen). Oft ist auch die Adresse der Forscher oder ihrer Institution angegeben. Im vorliegenden Kontext sind relevant: für deutschsprachige Soziologie *SOLIS* [\(13\)](#), für deutschsprachige Psychologie *Psyndex*, für internationale, insbesondere englischsprachige Literatur multidisziplinär-sozialwissenschaftlich *Social Sci Search* bzw. multidisziplinär kulturwissenschaftlich *Arts and Humanities Search* [\(14\)](#), für Soziologie *Sociological Abstracts*, für Psychologie *PsycInfo*.

Der Originalaufsatz muß also erst beschafft werden, was oft größere Schwierigkeiten bereitet, da viele (besonders fremdsprachige) Zeitschriften nur in wenigen Bibliotheken des deutschen Sprachraums, ja Europas vorhanden sind. Über den Fernleihverkehr der Bibliotheken können gegen Gebühr Fotokopien, manchmal auch Mikroverfilmungen des betreffenden Aufsatzes bestellt werden. Verschiedene Hosts, insbesondere im technisch-naturwissenschaftlichen Bereich, sowie die TIB (Technische Informationsbibliothek Hannover) bieten auch - zu relativ hohen Gebühren - die prompte Übermittlung einer Kopie der Originalarbeiten an, meist per Eilpost oder Telefax. Die Bestellung dieser Express-Kopien kann oft bereits online durchgeführt werden. Im Prinzip ist heute natürlich auch die elektronische Speicherung der Volltexte und ihre Übermittlung insbes. per e-mail möglich; hier gibt es jedoch noch (urheber-)rechtliche Probleme, so daß diverse Informationszentren

ihre bereits digital gespeicherten Texte wieder ausdrucken lassen und faxen oder postalisch versenden, eine Irrationalität sondergleichen.

Eine normale internationale Fernleihbestellung kann u. U. mehrere Monate dauern, ebenfalls hohe Kosten verursachen oder erfolglos bleiben; sofern die Adresse der Autoren bekannt ist, lohnt sich daher die direkte Kontaktaufnahme mit Bitte um Übersendung eines sog. "*Sonderdruckes*" (des betreffenden Beitrages) bzw. der betreffenden Datei per e-mail bzw. der Langfassung eines unveröffentlichten Forschungsberichtes. Es kann auch sinnvoll sein, in den homepages der Institutionen der Verfasser zu recherchieren (das kann bei sehr mobilen Autoren eine ganze Anzahl sein), z. t werden Texte (meist jene, die nicht in Zeitschriften veröffentlicht werden konnten bzw. Langfassungen dieser i. d. R. auf Wunsch der Herausgeber höchst gekürzten Artikel) über ftp, Gopher bzw. WWW zum Herunterladen im Volltext angeboten. Möglicherweise wird die weitere Zunahme elektronischer Zeitschriften und die Übersiedlung konventioneller Zeitschriften ins WWW (z. B. eine Reihe von Zeitschriften des Springer-Verlags) die Mühen der Originaltextbeschaffung in einigen Jahren als völlig obsolet erscheinen lassen. Immer mehr Autoren, welche in herkömmlichen Papierzeitschriften publizieren, stellen zusätzlich (sicherlich des öfteren im rechtsfreien Raum bzw. im Grauzonenbereich des Urheberrechts) ihre Aufsätze auf ihre eigene Homepages, da dies eine wesentlich größere und v.a. wesentlich raschere Diffusion ihrer Beiträge bringt als die herkömmliche Papierpublikation, die allerdings immer noch höher im Prestigewert steht, zumindest außerhalb von Informatik, Informationswissenschaft etc. Die US-amerikanische Philosophengesellschaft hat auch einen Preprint-Service eingerichtet, sodaß Aufsätze dort noch vor ihrem Erscheinen im Volltext abgerufen werden können.

Sofern die Adressen der Autoren nicht in der jeweils konsultierten Datenbase verzeichnet sind, können diese auch (für den deutschen Sprachraum) über die Projektdatenbase FORIS, weltweit über die (sehr umfangreiche) internationale Zitationsdatenbase *Social Sci Search*, den Sociological Abstracts und anderen einschlägigen Datenbanken eruiert werden.

Viele Zitate (insbesondere von Zeitschriftenaufsätzen) in den Literaturlisten von Veröffentlichungen sind fehlerhaft; vor einer Fernleihbestellung kann sich daher ihre Überprüfung durch Datenbankrecherchen lohnen.

Beiträge zu Sammelbänden (Readern) sind derzeit noch eher schlecht dokumentiert. Die Bücher dokumentierenden Datenbasen erfassen in der Regel nur den Titel des Sammelwerkes, allenfalls werden einige Schlagwörter vergeben, die den mitunter äußerst heterogenen Inhalt von Readern oft nur unzulänglich repräsentieren. Dabei könnte allein die (wenig arbeitsintensive) Speicherung des Inhaltsverzeichnisses von Sammelbänden (auch der Stichwörter der Buchregister) wertvolle Informationen bieten (von manchen OPACs im angelsächsischen Raum wird dies erfreulicherweise auch bereits getan). Im soziologischen Bereich wird im Rahmen der Datenbase *SOLIS* die mühevollen Aufbereitung der Beiträge zu Sammelbänden (Abstract-Erstellung, Schlagwortvergabe) betrieben. Auch die psychologischen Datenbasen *Psyndex* und *PsycInfo* erfassen in eigenen Sektionen Beiträge aus Sammelwerken. Es ist hier also der Dokumentationsstand wesentlich besser als in vielen anderen Wissenschaften.

Der allgemeine Rückstand bei der Dokumentation von Buchbeiträgen mag auch darauf zurückzuführen zu sein, (1) daß die meisten Zeitschriftenherausgeber bereits auf die Verfassung von Abstracts (meist sogar in mehreren Sprachen) durch die Autoren selbst achten, was in Sammelbänden leider noch nicht üblich ist (und den Arbeitsaufwand des Datenbasenproduzenten erhöht), (2) daß der wissenschaftliche Status von Beiträgen zu Sammelwerken niedriger ist als der von Aufsätzen, die in "renommierten"

wissenschaftlichen Zeitschriften veröffentlicht wurden, was u.a. damit begründet wird, daß hier Aufsätze oft einer langwierigen und kritischen Filterung unterliegen (durch - meist anonyme - "Referees", d. h. Gutachter), während in Sammelbänden in der Regel "nur" die Herausgeber für die Zurückweisung bzw. Aufnahme der Beiträge zuständig sind.

Die getrennte Dokumentierung der einzelnen Beiträge hat allerdings einen Nachteil, sie führt zu scheinbar hohen Trefferzahlen. 20 angezeigte Treffer können sich u. U. nur auf den Inhalt von drei Sammelbänden beziehen.

Unter "*Grauer Literatur*" wird Literatur verstanden, die nicht in herkömmlichen Verlagen erschienen ist und daher nicht oder nur mit beträchtlichen Schwierigkeiten über den Buchhandel erhältlich ist, z. B. unveröffentlichte Diplomarbeiten und Dissertationen, Forschungsberichte, Kongreßpapiere.

Graue Literatur ist jedoch für die wissenschaftliche Forschung sehr wichtig:

- Einerseits ist sie meist weitaus *aktueller* als die in wissenschaftlichen Zeitschriften oder gar Büchern publizierten Studien, da diese (z. B. aufgrund von langwierigen Begutachtungsverfahren, Zurückweisungen durch Herausgeber/Gutachter mit Überarbeitungswünschen, "Manuskriptstau" bei den Verlagen und dem langwierigen (da meist mit bescheidenen personellen und technischen Mitteln betriebenen) desktop-publishing durch die einzelnen Wissenschaftler bzw. Institute) sowie der in der Regel seltenen Erscheinungsweise von wissenschaftlichen Zeitschriften meist erst mehrere Jahre nach der Manuskripterstellung erscheinen.
- Andererseits sind Aufsätze in Zeitschriften oder Sammelbänden (nicht zuletzt auf Drängen der Herausgeber und der Verlage) oft *sehr gekürzt* und enthalten nicht alle jene Informationen (z. B. nicht den verwendeten Fragebogen), die erst eine intersubjektive Überprüfbarkeit zumindest teilweise ermöglichen.

Dazu ist meist die Lektüre des zugrundeliegenden *vollständigen* Forschungsberichtes nötig. Die Datenbank SIGLE umfaßt Graue Literatur aus dem EG-Raum, mit dem Schwerpunkt eher auf naturwissenschaftlich-technischer Literatur. Die Datenbank SOLIS umfaßt (neben Zeitschriftenliteratur und Beiträge zu Sammelbänden) auch soziologische Graue Literatur. Sie wird aufgrund der Projekterhebungen für die sozialwissenschaftliche Datenbank FORIS (s.o.) erfaßt, bezieht sich jedoch im Gegensatz zu FORIS nur auf den soziologischen Bereich im engeren Sinne. (15) Sofern im Projekterhebungsbogen unveröffentlichte Forschungsberichte, Tagungsbeiträge etc. angegeben werden, ersucht das Informationszentrum Sozialwissenschaften um Übersendung dieser Grauen Literatur; sie wird von der Stadt- und Universitätsbibliothek Köln (Schwerpunktbibliothek für Sozialwissenschaften in der BRD) für Fernleihbestellungen bereitgehalten. Die betreffende Kölner Bibliothekssignatur ist ebenfalls in SOLIS gespeichert.

Bücher. Die Deutsche Bibliothek in Frankfurt/Main erfaßt alle im deutschen Sprachraum erschienenen Bücher, sie nimmt (nach den bibliothekarischen Regeln) die "Titelaufnahme" vor, die von anderen Bibliotheken übernommen werden kann (vgl. u.a. Datenbank *biblio data*). Auch der österreichische wissenschaftliche Verbundkatalog (BIBOS) holt sich die Informationen zur Aufnahme von Büchern aus Frankfurt. *Bibliothekarische Datenbasen* enthalten als *Bibliotheskataloge* die Bestände einzelner Bibliotheken, vgl. die Datenbank der TIB (Technischen Informationsbibliothek) Hannover, oder der LoC (Library of Congress, Washington), oder als *Verbundkataloge* die Bestände einer größeren Anzahl von Bibliotheken, vgl. z. B. den *Niedersächsischen Monographien Nachweis (NMN)* des Bibliotheksrechenzentrums Niedersachsen (BRZN) in Göttingen. In den NMN geben die großen wissenschaftlichen Bibliotheken Niedersachsens ihre Daten ein. Sie enthalten in

der Regel Standortnachweise (Bibliothekssiegel und -signaturen) zur Unterstützung des Fernleihverkehrs.

Diese Verbundkataloge haben allerdings oft den Nachteil, daß in den am Verbund teilnehmenden Bibliotheken unterschiedliche Regeln bei der Titelaufnahme angewendet werden, z. B. bei der Aufnahme des/der Vornamen der Autoren. Hohe Trefferzahlen sind daher oft bloß auf unterschiedliche Nachweise weniger Bücher zurückzuführen. Auf dem Gebiet der Verbundkataloge sind die Angebote im WWW bereits vorzüglich. Besonders empfehlenswert ist der sog. Karlsruher Virtuelle Katalog. Er ermöglicht die simultane Recherche in zahlreichen Katalogen nach Wahl, sowohl in deutsch- als auch in englischsprachigen, zudem sowohl in Bibliotheks- als auch in Buchhändlerkatalogen (denn nicht jedes neu erschienene Buch wird sofort von einer Bibliothek gekauft und damit, mit einer gewissen Verzögerung, erfaßt).

Verbund-Datenbasen enthalten allerdings oft viele *Eingabefehler*. Es ist daher sinnvoll, in den jeweiligen Wörterbüchern dieser Datenbasen nach der Schreibweise von Autoren, Titelwörtern etc. nachzusehen und die Recherche sowohl mit der richtigen als auch mit den gefundenen falschen Schreibweisen durchzuführen. Häufig zu beobachten sind beim Titelwort "Philosophie" die falschen Schreibweisen "Philosophie" oder "Philosphie". Beim Tippen bzw. bei der Dateneingabe ist offensichtlich eine "Doppel-Vokal-Hemmung" zu beobachten: Wörter mit zwei gleichen, aufeinanderfolgenden Konsonanten verführen nach meinen langjährigen Rechercheerfahrungen häufig zu Eingabe- bzw. Tippfehlern. Allerdings hat sich die Situation gebessert, als etwa beim Österreichischen BIBOS kein Bibliothekar ein Buch einfach so aufnehmen und fehlerhaft in seinen Computer eingeben darf, sondern er muß sich einer bereits erfolgten Eintragung einer anderen österreichischen Bibliothek anschließen; falls es noch keinen österreichischen Eintrag gibt, muß BIBLIO DATA der Deutschen Bibliothek, Frankfurt am Main, konsultiert werden und deren Eintragung übernommen werden. So wird die Anarchie der Titelaufnahmen und damit verbundener Fehler zumindest langfristig vermindert werden.

Wir sollten Datenbanksystemen niemals völlig vertrauen, es können stets inhaltliche Fehler oder unerwartete Probleme bei der Recherche auftreten. Beides wird man wahrscheinlich nie völlig ausschalten können, einerseits aus Kostengründen (qualifizierte Fehlerkontrolle ist sehr zeitaufwendig und kann nur von Experten der jeweiligen Fachgebiete erfolgreich betrieben werden), andererseits ist unsere Welt so komplex, daß unser Wissen und die im Computer gespeicherten Informationen notgedrungen stets eine *Reduktion von Komplexität*, eine Informationsreduktion darstellen (gleichzeitig neue "Kompliziertheit", neue Informationsüberflutung produzieren). Die Reduktion - d. h. immer auch: Selektion - ist nach unterschiedlichen Kriterien möglich, die nicht die unseren sein müssen: Die Anzahl der möglichen Fragestellungen ist nahezu unbegrenzt, die Datenbasenproduzenten können daher nicht für alle möglichen Fragestellungen Deskriptoren und sonstige Hilfsmittel im Vorhinein bereitstellen.

Ich möchte zwar vor irrationaler Computerfeindlichkeit warnen, gleichzeitig aber auch vor blindem Vertrauen vor Datenbanken oder gar den in im Internet lagernden Informationen: Nur weil etwas per Knopfdruck (bzw. Mausklick) abgerufen werden kann, ist das Gespeicherte keineswegs fehlerfrei oder gar "wahr": Die verlockenden Versprechungen der Werbeprospekte der Datenbasenproduzenten und -vertreiber wie auch der WWW-Reklamen sind leider - wie die anderer Werbeprospekte - mit Vorsicht zu genießen.

2.2 Primärinformationen

Im Gegensatz zu den Datenbasen, die bloß Verweise auf Informationen bieten, die wir uns in einem zweiten Schritt noch besorgen müssen, bieten Datenbasen, die Primärinformationen gespeichert haben, *unmittelbar* die gewünschten Informationen, z. B. als Wörterbücher in Datenbankform, als statistische Datenbasen, als juristische Informationssysteme, die Gesetze und Verordnungen im Volltext enthalten.

Neben den oben erwähnten Presse-Verweisdatenbasen werden immer mehr *Pressedatenbasen* angeboten, die Artikel einer oder bestimmter Zeitungen *im Volltext* dokumentieren. Insbesondere die Wirtschaftsseiten renommierter Zeitungen werden ausgewertet.

Nicht zuletzt aufgrund des großen Zeitungssterbens versuchen zahlreiche *Nachrichtenagenturen* sich einen neuen Kundenkreis jenseits der Medienredaktionen zu erschließen. Sie bieten die *Volltexte ihrer Meldungen*, meist nach Themenbereichen getrennt und retrospektiv, zur Recherche an.

Terminologiebasen gewinnen in der Gegenwart immer mehr an Bedeutung. Wie bereits erwähnt, führt die terminologische Inflation und Anarchie zumindest in den Wissenschaften zu großen Problemen. Auf manchen Gebieten, wie z. B. der Medizin oder in den technischen Disziplinen ist das Terminologieproblem ganz schwerwiegend, daher sind auf diesen Gebieten schon seit längerem Normierungskommissionen tätig. Allen ist sicher die DIN, die "Deutsche Industrienorm" ein Begriff: fünfstelligen Nummern, z. B. DIN 34026. Durch die deutschen Industrienormung werden nicht nur Stecker, z. B. Diodenbuchsen und -stecker bei Tonbandgeräten genormt, sog. DIN-Stecker - amerikanische und japanische Geräte haben Bananenbuchsen oder andere von der DIN abweichende Buchsen. Auch die Herstellung bzw. die Eigenschaften eines Abstracts, der Zusammenfassung eines Zeitschriftenaufsatzes oder Buches, ist in einer eigenen DIN normiert. Ebenso die Zitierweise oder die Korrekturzeichen bei der Manuskripterstellung.

Die WHO, die Weltgesundheitsorganisation, hat sich für die internationale Standardisierung von Krankheitsbegriffen eingesetzt: Bei internationalen Hilferufen anlässlich von Katastrophen oder Seuchen muß Verständigung darüber möglich sein, welche Krankheiten ausgebrochen sind und welche Medikamente, welche pharmazeutischen Wirkstoffe benötigt werden. Daher muß es international verbindliche Terminologie geben, sowohl die Krankheiten, als auch die pharmazeutischen Substanzen betreffend.

Als Beispiel für eine Terminologiebase habe ich hier die Datenbase *MeSH* angeführt, eine medizinische Terminologiebase. Hier können wir etwa die Fachausdrücke für seltene Krankheiten recherchieren und die Recherche in der medizinischen Datenbase *Medline* vorbereiten.

Auch die europäischen Vereinigungsbestrebungen schaffen einen erhöhten Bedarf nach kompatiblen Fachbegriffen in den offiziellen Sprachen der EG. Daher bietet der Host ECHO einige Terminologiebasen an, so EURODICAUTOM zu (insgesamt fast 600 000) wissenschaftlich-technischen Begriffen, Kontext-Phrasen und Abkürzungen in den offiziellen Sprachen der Europäischen Gemeinschaft, sowie eine eigene Datenbase zu Abkürzungen und *Akronymen* (16) im Kontext der EG-Forschungs- und Entwicklungsprojekte (RTD-Acronyms).

Zahlreiche Lexika und Wörterbücher liegen inzwischen als CD-ROM vor. Auch im Internet finden sich vereinzelt Wörterbücher, bietet das WWW der Uni Passau das Langenscheidt-Wörterbuch, allerdings nicht die ausführlichste Ausgabe, und aufgrund der Langsamkeit der

Antworten ist weiterhin der Blick in die Buch-Version zu vorzuziehen.

Numerische und *Statistische Datenbasen* sind u. a. für die Sozial- und Wirtschaftsforschung hilfreich. In der BRD bietet die Datenbasis STATIS-BUND Bundesstatistiken an. Wenn Sie eines Tages eine Arbeit über die Entwicklung der Landwirtschaft verfassen möchten, müssen sie nicht mühsam Statistiken in Buchform durchsehen, Tabellen kopieren und erneut eintippen. Sie können stattdessen per Direktzugriff bzw. über Disketten Daten in maschinenlesbarer Form beziehen und mit den üblichen Programmen weiterverarbeiten.

Ein Großteil an Informationen könnte prinzipiell, es wird noch zu wenig betrieben, *papierlos* übermittelt werden, ein Großteil der heutigen üblichen Mehrfach-Schreibarbeit könnte vermieden werden. Buchverlage liefern z. T. bereits auf Diskette oder auf Magnetbändern gespeichert die Inhaltsverzeichnisse der von ihnen publizierten Bücher an Datenbanken, dort wird mit spezieller Software dieses Datenmaterial für die Informations-Wiedergewinnung, das *information retrieval*, aufbereitet. Die Daten müssen nicht noch einmal manuell in die Datenbank eingegeben werden (auch die Alternative, das Scannen, d. h. das Einlesen optischer Vorlagen durch den Computer mit automatischer Texterkennung, ist bekanntlich fehlerhaft). Somit ist eine beträchtliche Fehlerquelle ausgeschaltet. Papierlose Kommunikation und Publikationsformen ersparen also Arbeit und reduzieren mögliche Fehlerquellen.

Juristische Informationssysteme: Das größte europäische juristische Informationssystem ist *Juris* in der Bundesrepublik. Es enthält die deutschen Bundesgesetze, Verordnungen, Verwaltungsvorschriften, die Rechtssprechung und auch Literatur zum Thema. Die Dokumentation der Rechtssprechung ist besonders nützlich: Richter orientieren sich bei ihren Urteilen meist an bereits gefällten Gerichtsurteilen. Die große österreichische Rechtsdatenbank nennt sich RDB. Die Datenbase "JUSLETTER" von ECHO informiert über die Entwicklung des EU-Rechts, weitere EU-Rechtsdatenbasen sind über EUROBASES, den kommerziellen Host der EU, gegen Gebührenzahlung zugänglich.

Zusätzlich werden auch eine Reihe von juristischen Spezialdatenbasen angeboten, etwa über *ärztliche Kunstfehler*. Die große Berliner Mietervereinigung, Mitglied des Deutschen Mieterbunds, hat eine eigene *Mietrechtsdatenbank* erstellt, in der sämtliche deutsche Gerichtsurteile zur Umwandlung von Miet- in Eigentumswohnungen, zu Zwangsräumungen usw. dokumentiert sind. Ein Rechtsanwalt kann daher bei Prozeßführung diese Datenbank konsultieren, um sich zu informieren, welche Gerichtsurteile bisher zum Klagegegenstand gefällt wurden und sich in seiner Argumentation anpassen. Man muß also nicht ständig neue Ideen haben, sondern kann auf juristische Datenbasen zurückgreifen.

Im Rahmen des "*Electronic Publishing*" werden in *Volltextbanken* z. B. Zeitschriften mit den Texten sämtlicher Artikel gespeichert. Eine der ersten im deutschen Sprachraum auch elektronisch dokumentierten Zeitschriften sind die *VDI-Nachrichten*, die Zeitschrift der "Vereinigung Deutscher Ingenieure", einer großen und einflußreichen Organisation. Ihre Zeitschrift wird weiterhin in Hardcopy-(=Papier-) Form, also in Papierform herausgebracht, parallel dazu aber auch als Volltext-Datenbank.

Auch immer mehr große Werke der Literatur, der Philosophie und Theologie (auch die Bibel) werden zusätzlich zu konventionellen Editionen auf Diskette oder CD-ROM aufgelegt, um Geisteswissenschaftlern die Möglichkeit systematischer Auswertungen zu bieten. So liegt z. B. die gesamte, 23 Bände umfassende Akademie-Ausgabe des Werkes von Immanuel *Kant* in einer für PC maschinenlesbaren Form vor, und sogar die Werke von Thomas von Aquin.

Verschiedene Projekte erproben auch zur Zeit die Brauchbarkeit des electronic publishing für bestimmte Formen der Grauen Literatur, z. B. die Speicherung von Dissertationen auf optischen Speicherplatten. Eine konventionelle Vorform ist die Mikroverfilmung von Dissertationen durch die nordamerikanische Firma UMI (University Microfilms International). Der Nachteil dieses Verfahrens, das anstrengende Lesen dieser Mikrofilme oder Mikrofiches (Mikrofilme im Postkartenformat) in Lesegeräten bzw. die langwierige und kostenintensive Rückerstellung von Kopien in schlechter Qualität, kann aufgrund der technischen Fortschritte bei Laser- und Tintenstrahldruckern beim electronic publishing vermieden werden. Der größte Vorteil digitaler Speicherung im Vergleich zur Mikrofilm-Archivierung ist jedoch die Möglichkeit, die verschiedenen Suchfunktionen im Volltext-Dokument anwenden zu können.

Fast als ungeplanter Wildwuchs haben sich die zahllosen Textdateien auf ftp-, gopher- und WWW-Servern entwickelt. Sie können v. a. über Search Engines und über die gezielte Suche in den gopher-/WWW-Seiten/ftp-Archiven von akademischen Institutionen gesucht werden.

Unter *Expertensystemen* versteht man *natürlichsprachliche Frage-Antwort-Systeme*. Damit überschreiten wir das Feld der Fachinformationssysteme in Richtung "künstliche Intelligenz" (KI). KI bedeutet: Maschinen bewerkstelligen etwas, wozu Menschen Intelligenz benötigen würden. Expertensysteme müssen sowohl die an sie in natürlicher Frage gestellten Fragen analysieren ("verstehen") können, auf der Basis von Expertenwissen eine Antwort finden und diese ebenfalls in natürlicher Sprache formulieren können. Viele Experten können gar nicht genau und formalisiert angeben, wie sie zu ihren Expertenurteilen kommen. Die Expertensysteme stecken daher noch in den Kinderschuhen, denn Computer haben (manche Experten meinen: vorläufig, viele hingegen: prinzipiell) größte Schwierigkeiten, "bildhaftes" und "unscharfes", "intuitives" bzw. "implizites" (Polyani) Wissen oder gar metaphorische (bildhafte) Formulierungen zu verarbeiten. Es gibt bereits Prototypen von Expertensystemen auf eng abgesteckten Gebieten, z. B. in der Medizin (Krankheitssymptome, -diagnose) sowie für Kfz-Reparaturen (Fehlerdiagnose).

3. Sozialwissenschaftlich relevante Hosts und Datenbasen

In der *Übersicht 4* werden ausgewählte sozialwissenschaftlich relevante Hosts (= "Wirtsrechner") aufgelistet:

- In der Bundesrepublik Deutschland ist der größte Host mit sozialwissenschaftlichen Datenbasen das *DIMDI* (Deutsches Institut für medizinische Dokumentation und Information). Eigentlich für biowissenschaftliche Dokumentation zuständig, hat dieser Host auch soziologische und psychologische Datenbasen in sein Programm aufgenommen, u. a. um Mediziner in Forschung und Praxis dazu anzuregen, auch psychologische und soziologische Aspekte in ihre Tätigkeit einzubeziehen.
- Der einzige größere österreichische Host ist die *EDV* GesmbH in Wien, im Besitz des Gewerkschaftsbundes, fungiert u.a. als Host für die umfangreiche sozialwissenschaftliche Dokumentation SOWIDOK (der Arbeiterkammer Wien) und die österreichische Rechtsdatenbank RDB. Daneben sind an mehreren österreichischen Universitätsrechenzentren Datenbasen zur Online-Recherche implementiert, z. B. die FoDokAustria, die österreichische Forschungsprojekte dokumentiert, oder die

"Österreichische Dissertationsdatenbank" und Datenbasen zur Erwachsenenbildung.

- Der Schweizer Host (mit britischer Beteiligung) "*Data star*" ist ein kommerzieller Host von "Radio Suisse", der tatsächlich als ein "Supermarkt" des Wissens zu bezeichnen ist: "Data star" arbeitet laufend mit Sonderangeboten. Bei zeitlicher Flexibilität von Recherchen kann man daher bei Data star günstig recherchieren bzw. Recherchen trainieren und neue Datenbasen erproben. Auch eine Reihe von lizenzgebührenfreien *Trainingsdatenbasen* wird angeboten, d. h. Auszüge aus Datenbanken zur Kostenverringerung, da hier bei Recherchen geringere CPU-Zeiten (zentrale Rechenzeiten) anfallen. Der Schwerpunkt liegt zwar im kommerziellen bzw. kommerziell verwertbaren Bereich, es werden jedoch eine Reihe von allgemein interessanten Datenbasen (z. B. Cuadra Directory of Databases) angeboten sowie psychologische und soziologische Datenbasen.

- Die Pariser *INIST Diffusion* (eine Tochtergesellschaft des nationalen französischen Forschungszentrums CNRS) bietet sämtliche *francis*-Datenbasen auf dem Gebiet der Geistes-, Sozial- und Wirtschaftswissenschaften zur Online-Recherche und auf CD-ROM an, d. h. die Datenbasenversionen der entsprechenden Nachschlagewerke der Reihe *bulletin signaletique*. Hier werden v.a. Publikationen in den romanischen Sprachen, die bei Datenbanken mit angelsächsischer Orientierung vielleicht unter den Rost fallen, nachgewiesen.

- Der bekannteste Wissenschaftshost in den USA mit internationaler Bedeutung ist zweifellos *DIALOG*.

In der *Übersicht 5* werden Datenbasen für *Soziologen* angeführt. Die Übersicht zeigt, daß ein- und dieselbe Datenbank bei mehreren Hosts angeboten wird. Die einzelnen Hosts unterscheiden sich hinsichtlich des Benutzer-komforts ihrer Retrievalsysteme, der Aufwendigkeit und Effektivität der Fehlerkontrollen sowie der Preisgestaltung, es erscheinen dazu auch in speziellen Fachzeitschriften (z. B. in "database", USA) Vergleichsstudien. Der durchschnittliche Endnutzer verzichtet jedoch meist auf aufwendige "rationale" Entscheidungsfindungen und wählt denjenigen Host, dessen Kommandosprache er beherrscht oder der die meisten verwandten Datenbasen zur gewünschten Thematik anbietet.

Übersicht 6 zeigt die im vorliegenden Kontext interessantesten Hosts, nämlich *Hosts mit der Kommandosprache CCL*. Bei den Hosts DIMDI und ECHO sind z. T. auch menügestützte Recherchen möglich, welche die Kenntnis von CCL nicht voraussetzen.

CCL steht für *Common Command Language* und war ursprünglich als gemeinsame europäische Kommandosprache entwickelt worden, konnte sich jedoch aufgrund verschiedenster Eigeninteressen nur teilweise durchsetzen. GRIPS bzw. DIRS (17) sind erweiterte und geringfügig modifizierte Formen von CCL.

Aus der Übersicht können wir entnehmen, daß wir im Bereich der Soziologie /Psychologie auch bei Beschränkung auf *eine* Kommandosprache auf fachspezifische und interdisziplinäre bzw. allgemeine Datenbasen ausreichend Zugriff haben. (18)

Wir werden uns daher bei der folgenden Einführung in die Logik von Datenbankrecherchen auf das Beispiel CCL beschränken.

4. Suchfunktionen am Beispiel der Kommandosprache CCL

Bevor wir mit der eigentlichen Recherche beginnen, müssen wir mit dem gewünschten Host in Kontakt kommen, im sog. *Eingangsdialog*. Dieser unterscheidet sich im Detail je nach Terminalprogramm und Datenübertragungsnetz. Informationen dazu erhalten wir einerseits (für Datex-P) bei den Datenfernmeldeämtern bzw. für das Deutsche Forschungsnetz bei den angeschlossenen Universitätsrechenzentren, andererseits bei den Hosts selbst.

Im Datex-P-Netz der Bundespost müssen wir den geographisch nächstgelegenen *Netzknoten* der Bundespost anwählen (bei Verwendung eines Akustikkopplers mittels Telefon, bei Verwendung eines Modems per PC). In fast jeder deutschen Großstadt befindet sich ein solcher Netzknoten. Sodann müssen wir uns mit unserer "Teilnehmerkennung" ausweisen; der erste Teil dieser Kennung ist unser Benutzername, der zweite, geheimzuhaltende, das eigentliche Datex-P-Passwort. Sodann müssen wir die NUA (Network User Address) des Hosts eingeben. Sodann müssen wir uns noch einmal, und zwar mit dem Passwort des betreffenden Hosts, ausweisen. Hat der Host unser Passwort akzeptiert, ist es soweit: wir können mit der Recherche beginnen. (19)

Bei Zugriff über das Internet können Universitätsangehörige universitäre bzw. Forschungsnetze i. d. R. (noch?) kostenlos nutzen. So stehen z. B. an der Universität Linz in allen Gebäuden öffentlich zugängliche WWW-Terminals, Dutzende von WWW-tauglichen Macs und Pcs stehen zudem jeweils in den diversen Schulungs- und Übungsräumen, und die Arbeitsplätze der MitarbeiterInnen sind mit dem EDVZ (Rechenzentrum) und damit mit E-mail- und sonstigen Internet-Servern verkabelt.

4.1 Elementare Suchfunktionen/Befehle

In *Übersicht 7* sind verschiedene elementare Suchfunktionen aufgelistet. Jedes leistungsfähige Retrievalsystem muß diese Suchfunktionen aufweisen; lediglich die konkreten Kommandos unterscheiden sich je nach Kommandosprache. Hier werden die Suchfunktionen am Beispiel der Kommandosprache *CCL* demonstriert:

- Jeder Host muß eine Möglichkeit bieten, einzelne *Datenbasen* zur Recherche auszuwählen; bei CCL mit Hilfe des Befehls **base**.
- Jedes einigermaßen benutzerfreundliche Retrievalsystem sollte eine Möglichkeit bieten, in den *Wörterbüchern* der Datenbasen zur Vorbereitung der Recherche, d. h. zur Zusammenstellung/Überprüfung der Suchbegriffe zu suchen; diese Wörterbuchfunktion wird bei CCL mittels **display** (oder abgekürzt **d**) aufgerufen. Die Wörterbuchfunktion, das Display-Kommando ermöglicht die Ausgabe von alphabetisch benachbarten Suchbegriffen.
- Jedes Retrievalsystem muß die Möglichkeit bieten, die gespeicherten Dokumente zu suchen, hier: **find** (abgekürzt: **f**).
- Nach Eingabe dieses Befehls (zusammen mit Suchbegriffen) erfolgt eine Meldung der Trefferzahl; mittels **print** (**p**)-Befehl werden die Treffer auf einem Drucker des Wirtsrechners ausgedruckt und per Post übermittelt; mittels **show** (abgekürzt **s**) werden die einzelnen gefundenen Dokumente sofort auf dem Bildschirm gezeigt und können gespeichert und

nach Ende der Recherche auf dem eigenen Drucker ausgegeben werden.

- Mit dem Befehl **tabelle (t)** kann das bisher erfolgreiche *Rechercheprofil* (= das Protokoll der Suchbegriffe mit Treffern und deren Trefferzahl) aufgelistet werden, d. h. erfolglose Suchbegriffe werden in dieses Profil nicht aufgenommen (vgl. zu Beispielen von Rechercheprofilen Abb. 7).

- Mit **stop** wird die Recherche beendet.

Zur konkreten Abwicklung einer Recherche müssen wir noch zwei Befehle kennen, welche die *Fortsetzung* von unterbrochenen Ausgaben oder Kommandos bewirken; sie werden in der Regel vom System vorgegeben; es genügt dann das Drücken der carriage return-Taste (Zeilenvorschub):

- **more** fordert die Fortsetzung der *Ausgabe* im Display- und Show-Kommando;

- **continue** fordert die Fortsetzung der *Befehlsabarbeitung*. Wenn Sie einen Befehl eingegeben haben, der sehr viel CPU-Zeit benötigt, d. h. Rechenzeit in der central processing unit (der zentralen Recheneinheit), erkundigt sich das System ab und zu, ob es noch weitermachen soll. Bei manchen CCL-Systemen muß auf diese Frage auch mit **y(es)** oder **n(o)** geantwortet werden.

Im Falle von Problemen bzw. Informationsdefiziten können auch die *Hilfsfunktion* "help" sowie die *Informationsfunktion* "news" nützlich sein:

- Mit der Eingabe von "**help**" (oder eines Fragezeichens) können Sie Unterstützung anfordern. Es ist in fast jedem System ein Hilfsprogramm eingebaut, das z. B. die Befehle erklärt. Sie können die Recherche also fast ohne Vorkenntnisse beginnen. Zu beachten ist allerdings, daß auch während der Konsultation der Hilfsfunktion Kosten anfallen (vgl. zu den Kosten Pkt. [4.6](#)).

Daher würde ich aus Kostengründen die gründliche Vorbereitung der Recherche empfehlen. Die meisten Hosts liefern auch *Kurzanleitungen*, d. h. kurze Aufstellungen der wichtigsten Funktionen und Befehle, die wir zur Gedächtnisstütze verwenden können, vgl. in Abb. 8 die GRIPS-Kurzanleitung von DIMDI.

- Mit "**news**" können Sie Neuigkeiten anfordern, es kommen jeden Tag Neuigkeiten, z. B. Seminarankündigungen, die Verfügbarkeit neuer Datenbasen. Die wichtigen Ankündigungen betreffen einerseits die Zeiten, in denen der Host abgeschaltet wird - bei Wartungsarbeiten und Systemumstellungen. Meist werden solche Arbeiten am Wochenende durchgeführt. Andererseits werden unter "news" auch Werbeaktionen (kostenfreie oder verbilligte Recherchemöglichkeiten) angekündigt (s. Pkt. [4.6](#)).

4. 2 Zugriffsmöglichkeiten über Datenfelder

Die [Übersicht 8](#) enthält Beispiele für *Zugriffsmöglichkeiten* sowie für Bezeichnungen der entsprechenden *Datenfelder* in CCL-Datenbasen.

Bei fast allen Datenbanken ist die gesamte gespeicherte Information über ein Buch oder

ein Projekt, das nennt man meist etwas irreführend "Dokument", präziser "Dokumentationseinheit", im sogenannten **Freitext** recherchierbar: In einem Recherchevorgang kann ein Suchwort in sämtlichen Datenfeldern: meist Autor, Titel, Erscheinungsjahr, Publikationsort, Verlag, Schlagwörter, Abstracts (Kurzzusammenfassung), gesucht werden.

Die Freitextrecherche ([20](#)) ist allerdings nicht unproblematisch. Dazu zwei Beispiele:

- Bei der Suche nach Projekten oder Publikationen zum Thema "Geschwister" kann es verhängnisvoll sein, wenn die Adresse eines Forschungsinstituts "Geschwister-Scholl-Straße" lautet, denn dann werden uns u. U. zahlreiche Bücher über Astrophysik und andere für uns irrelevante Themen nachgewiesen.

- Bei der Suche nach "Betrug in den Wissenschaften" haben wir im ersten Durchlauf nicht bedacht, daß in zahlreichen Abstracts Statistiken referiert werden und daher häufig das Verb "betrug" (Imperfekt von "betragen) verwendet wird. "Betrug" im Sinn von Täuschung anderer Akteure sollte daher nur als kontrollierter Deskriptor bzw. als Titelwort gesucht werden.

Solche Fehler sind nicht nur ärgerlich, sondern auch kostspielig: Wir müssen bei kostenpflichtigen Recherchen alle unsere 'Treffer' bezahlen, ob sie für uns brauchbar sind oder nicht - im Gegensatz zum Fotografen, der unsere verwackelten Bilder nicht berechnet.

Genauer gesagt können wir nicht nach allen Wörtern des Dokuments recherchieren, zumindest nicht im normalen Abfragemodus, sondern wir können in der Regel nur nach allen mit Bedeutung verbundenen, sog. "*signifikanten*" Wörtern recherchieren, die im jeweiligen Wörterbuch der Datenbanken abgespeichert werden (je Datenfeld gibt es gewöhnlich ein Wörterbuch).

Was bedeutet in diesem Zusammenhang "signifikant"? Nicht jedes Wort, das in einem Dokument vorkommt, hat eine inhaltliche Bedeutung, die bei der Wiederauffindung von Nutzen wäre. Beim Aufbau einer Datenbank wird daher (um den Umfang der Datenbasen und ihrer Wörterbücher und damit sowohl Kosten bei der Datenbasenerstellung als auch bei der Recherche zu reduzieren) in der Regel eine *Stopwortliste* erstellt. Diese enthält jene Wörter, die bei der Erstellung der datenbaseninternen Wörterbücher unterdrückt werden sollen, weil sie keine inhaltlich relevanten Informationen enthalten. Einerseits enthalten Stopwortlisten gewöhnlich Bindewörter wie "und", persönliche Fürwörter, Artikel (der, die, das). Das entlastet die Wörterbücher der Datenbase beträchtlich und führt kaum zu Informationsverlusten: Denn im Normalfall recherchieren wir nicht mit dem Ziel des Nachweises aller Dokumente, in denen ein "und" vorkommt. Eine solche Fragestellung wäre allenfalls für einen quantitativ forschenden Linguisten interessant, wenn er nachweisen möchte, daß es nur einen bestimmten Prozentsatz an Sätzen gibt, wo kein "und" vorkommt - oder ähnliches. Aber für den wissenschaftlichen Alltag in allen anderen Disziplinen ist eine solche Fragestellung vollkommen uninteressant.

Bei der Recherche in einem *bestimmten Datenfeld* werden die Gefahren der Freitextrecherche verringert:

- Wenn wir verlangen: "Geschwister" müssen im **Titel** (CCL: **TI**) vorkommen, dann ist damit ausgeschlossen, daß uns Dokumente mit Straßennamen, in dem Geschwister vorkommen, oder Autoren, deren Familienname so lauten, nachgewiesen werden. Es ist allerdings mit der Möglichkeit zu rechnen, daß das Wort "Geschwister" metaphorisch verwendet wurde, was gerade in Buchtiteln nicht unwahrscheinlich ist. So gibt es Bücher, die im Titel die verschiedenen Weltreligionen als "Geschwister" bezeichnen. Bei einer Recherche mit den

Stichwörtern "sister/ sisters" wurden uns nicht nur Geschwisterstudien, sondern auch dutzende Rezensionen eines Buches über die sieben größten US-amerikanischen Erdölkonzerne nachgewiesen (der Titel des so häufig rezensierten Buches: "The Seven Sisters") sowie Bücher über die großen Weltreligionen, die von manchen AutorInnen metaphorisch als (nicht immer friedliche) "Geschwister" bezeichnet werden.

Die metaphorische Verwendung von Begriffen ist insbesondere bei nicht- oder populärwissenschaftlichen Sachbüchern und in Philosophie/Theologie häufig zu beobachten.

Die terminologische 'Anarchie' ist im Bereich der Wissenschaften am ausgeprägtesten in Soziologie, Psychologie und anderen sogenannten *paradigmenschwachen* Fächern. In diesen gibt es dutzende Schulen nebeneinander, welche jeweils unterschiedliche Terminologien verwenden. Das bereitet bei der Datenbankproduktion und beim effizienten Information Retrieval Probleme. In den Ingenieur- und Naturwissenschaften sind m. E. diese Probleme nicht ganz so gravierend, da sich hier bestimmte Paradigmen (Lehrgebäude, Weltansichten, Kataloge anerkannter Methoden) und damit auch Sprachregelungen eindeutig durchgesetzt haben.

- Auch bei der Recherche im Titelfeld Abstracts (**AB**) ist ein gewisser Anteil an Recherchemüll kaum zu vermeiden. Wir können aber mit Recherchen in den Abstracts oft zusätzliche Treffer erzielen, etwa wenn eine behandelte Thematik nicht wichtig genug war, um sich begrifflich im Titel oder im Schlagwortfeld niederzuschlagen.

- Der treffsicherste Zugang ist die Recherche über *Deskriptoren*, über kontrollierte Terminologie. Es gibt für die meisten Datenbasen voluminöse Nachschlagewerke, die überdies alle 2 - 3 Jahre aktualisiert werden. Sie verzeichnen die bei der Datenbankproduktion verwendeten kontrollierten Schlagwörter, die *Deskriptoren*. Deskriptor ist ein Wort, das einen Gegenstand, ein Dokument *beschreibt* (deskribiert) und aus einer kontrollierten Fachterminologie stammt. Ein Verzeichnis von Deskriptoren, insbesondere wenn diese Deskriptoren in eine systematische Ordnung, eine sachliche Hierarchie gebracht wurden, nennt man *Thesaurus*. Dieses Wort stammt aus dem Griechischen und bedeutet Schatzkästlein. Bei CCL wird das Datenfeld der Deskriptoren mit **CT** (Controlled Terms) aufgerufen.

Abb.5 und 6 enthalten Auszüge aus der *Schlagwortliste Sozialwissenschaften* des IZ (Informationszentrums) Sozialwissenschaften in Bonn; sie ist ein Hilfsmittel für Recherchen in den beiden, vom IZ Bonn produzierten Datenbasen FORIS (Forschungsinformation Sozialwissenschaften) und SOLIS (Sozialwissenschaftliches Literaturinformationssystem). Zur Zeit wird diese Schlagwortliste gerade ins Russische übersetzt. Die Schlagwortliste hat zwei Teile, einen systematischen und einen alphabetischen Teil.

Der *systematische* Teil (vgl. *Abb. 5*) enthält eine Gliederung der Schlagwörter nach Fach- und Sachgebieten. Hier kann man sich ein Bild davon machen, welche Inhalte in der Datenbase überhaupt erfaßt werden. Sehr oft präzisiert sich dabei eine verschwommene Fragestellung.

Wenn man seine Vorurteile bzw. Schwellenängste gegenüber Dokumentationssystemen und Nachschlagewerken dieser Art überwunden hat, kann man solche Verzeichnisse auch für andere Zwecke benutzen, z. B. bei der Strukturierung eines Themas anlässlich der Vorbereitung der Diplomarbeit.

Es ist nicht so einfach, eine Summe von aufgehäuften Materialien sinnvoll zu strukturieren. Dokumentare befassen sich u.a. mit der Entwicklung von Klassifikationsschematas. Diese

Schemata müssen wir nicht kritiklos übernehmen, aber wir können uns von ihnen anregen lassen. Der systematische Teil enthält zur Orientierung eine Übersicht über die Systematik der Schlagworte (s. Abb.), dann die Systematik selbst, von Wissenschaftstheorie bis Wirtschaft (s. Auszug), gefolgt von den systematisch gegliederten Schlagwörtern (s. Auszug).

Bei der Vorbereitung einer Recherche über "neue Technologien und Beruf, Arbeit" finden wir die systematische Stelle "Technik, EDV und Arbeitsmarkt". Sie enthält die wichtigsten Fachausdrücke zum Thema.

Im *alphabetischen Teil* des Schlagwortverzeichnis (vgl. Abb. 6) kann ich verschiedene mir bekannte Wörter nachschlagen, bei unserem Recherchebeispiel z. B. CAD/CAM und überprüfen, ob diese Wörter als Deskriptoren verwendet werden. Fettgedruckte Wörter werden als Schlagwörter, als Deskriptoren verwendet.

In unserem Beispiel sind die Wörter **CAD** und **CAM** fettgedruckt, also Deskriptoren. Daneben oder darunter stehen in Normalschrift jene Wörter, für die der Deskriptor auch verwendet wird (UF = used for) oder ein verwandter Begriff (RT = related term). In Normalschrift gedruckte Wörter werden nicht als Deskriptor verwendet, so wird uns z. B. auf der nächsten Seite empfohlen ("USE"), statt dem Wort "Computer" den Deskriptor "EDV" zu verwenden. Oft wird auch eine Deskriptorenkombination empfohlen, z. B. statt "Computergrafik" wird uns die Verknüpfung von "EDV" und "Grafik" vorgeschlagen.

Um einen "Gegenstand" zu bezeichnen bzw. zu "konstruieren" (vieles in den Sozial- und Geistes- (bzw. Kultur-)wissenschaften wird von Kritikern als "Begriffsdichtung" bezeichnet), bestehen vielfältigste Möglichkeiten. Man kann dutzende Begriffe verwenden oder neue erfinden, nicht selten besteht wissenschaftliche Originalität gerade in der Entwicklung einer neuen Terminologie. Diese *terminologische Inflation* bereitet bei Recherchen enorme Probleme. Thesauri, also Deskriptorenlisten, sind ein Versuch, die Vielzahl bedeutungsgleicher Termini zu reduzieren, indem man verbindlich für jeden "Gegenstand" einen Hauptdeskriptor festlegt. Alle anderen synonymen Wörter kann man nur in der Freitextrecherche, bzw. in den Datenfeldern der Titel und der Abstracts verwenden. Es werden internationale Konferenzen veranstaltet, auf denen sich Dokumentare und Bibliothekare mit Experten eines Gebietes um die Vereinheitlichung der Terminologie bemühen. Trotz aller Planung und Voraussicht hinken die Klassifikationsversuche meist hinter den tatsächlichen wissenschaftlichen Entwicklungen hinterher, sodaß gerade bei Recherchen zu innovativen Ansätzen und Forschungsgebieten Freitext- bzw. Titel- und Abstract-Recherchen zu empfehlen sind.

-- *Exkurs: Datenfeld CT im Display-Kommando*

Die alphabetische Schlagwortliste ist auch online abrufbar, mit der (Wörterbuchfunktion in Verbindung mit einer Einschränkung auf das Datenfeld controlled terms. Alle Deskriptoren, die alphabetisch dem Deskriptor Technik nachfolgen (auch Wortzusammensetzungen, die mit "Technik" beginnen), erhalte ich in CCL/DIRS mit dem Kommando: **D CT=Technik**). Allerdings können wir online nur die tatsächlich verwendeten Deskriptoren (sowie ihre Häufigkeit im Datenfeld controlled terms) eruieren; die Querverweise zu nicht verwendeten Ausdrücken fehlen. Dafür wird uns die Trefferzahl angezeigt, d. h. wieviele Dokumente das jeweilige Wort mindestens einmal enthalten.

Die Schreibweise von Nachnamen und Vornamen von Autoren, allgemein von Personen, ist sehr inkonsistent, mitunter auch fehlerhaft (bei einer Recherche nach Zitationen des Sozialphilosophen Montesquieu fand ich im Social Sciences Citation Index die "abenteuerlichsten" Schreibweisen); auch hier lohnt sich eine Konsultierung des

Datenbank-Wörterbuchs zum Datenfeld Autoren (**AU**).

- Eine der interessantesten Zugangsmöglichkeiten ist die *Zitationsrecherche*, d. h. die Suche über zitierte Autoren oder Publikationen. In CCL: Datenfelder **RF** (reference), **RJ** (referenced journal) und **RP** (referenced patent no.), u.a. Datenbank *Social Sci Search* bei DIMDI.

Die Zitationsrecherche bietet einen großen Vorteil: Wir können die gesamte Terminologieproblematik elegant umgehen. Wir müssen nicht in Thesauri nachblättern und Suchwörter eruieren. Statt dessen verwenden wir bekannte Autoren bzw. Publikationen, sogenannte *Schlüsselautoren* bzw. *Schlüsselpublikationen*. Bei einer Recherche zur marxistischen Wirtschaftstheorie kann man z. B. Marx oder bekannte andere marxistische Autoren als Schlüsselautoren verwenden, denn jeder Marxist im weitesten Sinn wird wahrscheinlich zumindest an einer Stelle, Marx zitieren oder neuere bekannte marxistische Wirtschaftstheoretiker wie den Brüsseler Ökonomen Ernest Mandel oder den Weltmarkt-Theoretiker Elmar Altvater von der Freien Universität Berlin.

Wir können *Cozitationsrecherchen* durchführen und etwa Marx mit Jean Piaget kombinieren, dann weisen wir eine ganz spezifische Wissenschaftstradition nach, nämlich Autoren, die der kognitiven Entwicklungspsychologie zuzurechnen sind und gleichzeitig einen marxistischen Ansatz haben. Solche Autoren über andere Zugangsmöglichkeiten zu finden, ist äußerst aufwendig. Wir können auch einige wichtige Autoren aus der Kybernetik und einige große Soziologen koppeln und schauen, in welchen thematischen Feldern die beiden Autorengruppen zusammen zitiert werden. Gerade in den letzten Jahrzehnten wurden die größten wissenschaftlichen Innovationen von Forschern gemacht, die zwei oder mehrere bisher getrennte Wissenschaftsdisziplinen bzw. Forschungstraditionen fruchtbar miteinander verknüpft haben.

In der Regel sind jedoch nicht alle über eine Zitationsrecherche nachgewiesenen Publikationen Treffer. Z. B. kann jemand in seinem Aufsatz vermerken, eine bestimmte Buchpublikation interessiere ihn im Kontext der vorliegenden Studie überhaupt nicht (und sie gerade deshalb zitieren). Allerdings sollen systematischen Zitationsanalysen zufolge solche negativen Zitationen eher selten sein. Viele wissenschaftlichen Publikationen beginnen mit Entschuldigungen - mit einem bestimmten Thema, einem bestimmten theoretischen Ansatz könne sich der Autor in seiner Doktorarbeit nicht befassen (weil es, so die Standardformel, "den Rahmen dieser Arbeit sprengen würde") - genau deshalb zitiert nicht selten ein Autor die betreffenden Veröffentlichungen, weil er sich nicht eine Lücke seines Kenntnisstandes nachsagen lassen möchte.

4.3 Verknüpfungen

Meist ist es bei der Recherche erforderlich, *mehrere* Suchbegriffe einzugeben und diese miteinander in bestimmter Weise zu verbinden. Wie bereits erwähnt, kann man das Recherchieren in Datenbanksystemen metaphorisch mit dem *Fischen* in mehr oder weniger bekannten Fischteichen vergleichen. Wir knüpfen logische, verbale *Suchnetze*, um die gewünschten Dokumente, die gewünschten "Fische" einzufangen. Dazu benötigen wir nicht nur einzelne Stücke, Hanf- oder Plastikschnurstücke, sondern wir müssen die Stücke auch verknüpfen, damit sie ein Netz ergeben.

Verknüpfen können wir die Stücke, d. h. Suchbegriffe, mit *logischen Junktoren* bzw.

Operatoren. (21) Diese beruhen auf *angewandter Mengenlehre* (vgl. Abb. 9):

- Wenn zwei Mengen A und B sich überschneiden, können wir erstens den *Durchschnitt* dieser Mengen bilden, das sind jene Elemente, die sowohl der Menge A als auch der Menge B angehören: **A AND B**. Wobei ich betonen muß: Das logische "und" ist stärker als das alltagssprachliche "und", es impliziert keine bloße Aneinanderreihung, sondern die Forderung: *sowohl als auch*.
- Zweitens können wir die *Vereinigung* der beiden Mengen bilden: Sowohl A als auch B als auch A und B. Das ist A oder B: **A OR B** und entspricht oft dem alltagssprachlichen "und": es wird keine Alternative offeriert bzw. Entscheidung gefordert, sondern es sind beide Möglichkeiten zugelassen.
- Drittens können wir verlangen: Wir möchten zwar Elemente der Menge A, aber nur jene, die *nicht auch zugleich* Element der Menge B sind: **A NOT B**. Diese logischen Junktoren bzw. Operatoren aus der sog. Booleschen Algebra ist der für den Nutzer offen ersichtliche Beitrag der Mathematik beim Recherchieren, die übrigen mathematischen Voraussetzungen von Datenbanken sind nicht ohne weiteres ersichtlich und insofern für den Endnutzer irrelevant.

Die Untersuchung der *Zugangsbarrieren* zu Fachinformationssystemen (Fragestellung: Warum gibt es so viele Leute, die vor Datenbanken Angst haben?) erbrachte zwei Resultate: Erstens beherrschen viele Menschen nicht das Schreibmaschineschreiben und haben daher Angst vor Tastaturen. Daher hat Apple, in der Folge von vielen anderen kopiert, die "Maus" erfunden und andere sog. Mensch-Maschine-"Schnittstellen", z. B. "Joysticks", um das Tippen zu reduzieren. Die zweite Zugangsbarriere: Viele Leute haben mit dem logischen "und" und "oder" große Schwierigkeiten. Für Leute, die in der Mittelschule in den "Genuß" von Mengenlehre kamen, ist das nicht ganz nachvollziehbar.

Hilfreich ist es, sich diese logischen Operationen bildhaft vorzustellen, mit Hilfe sog. *Venn-Diagramme* (vgl. Abb. 4). Generell wird die Visualisierung geistiger Operationen in den heutigen Wissenschaften noch zu wenig und auch zu wenig genau betrieben; da jedoch nach den neueren Erkenntnissen der Hirnforscher und kognitiven Psychologen ein großer Teil des menschlichen Gehirns mittels graphischer Informationsverarbeitung funktioniert, ist dies ein schweres Manko. Doch die Macintosh-Philosophie fand inzwischen allorts Nachahmung, und auch unter den bunten Bilder des WWW finden sich (nicht immer effiziente) Versuche von Visualisierungen.

Statt "A" und "B" müssen wir bei der Recherche Suchwörter (Titelwörter, Schlagwörter, Autorennamen) einsetzen oder zitierte Publikationen, zitierte Autoren oder die Zeilennummer (statement number) bereits erzielter Treffer.

4.4 Die Syntax von CCL/GRIPS

Das *erste Zeichen* (bzw. die *erste Zeichenkette*) wird immer als *Befehl* aufgefaßt. Wenn die Befehlseingabe vergessen wurde, erfolgt daher folgerichtig eine Fehlermeldung ("Befehl unbekannt"). Ein Großteil der Befehle muß nicht ausgeschrieben werden, oft reicht die Eingabe des ersten Buchstabens. Das System akzeptiert Groß- und Kleinschreibung unterschiedslos.

Danach folgt ein Blank (Leerzeichen); hierauf folgt entweder ein Suchterm oder eine Spezifizierung des Kommandos, z. B.: **D Familie** oder **F AU**. Diesen Spezifizierungen folgt kein Blank, sondern sofort ein "=", dahinter ebenfalls ohne Zwischenraum der Suchterm oder eine numerische Angabe, z. B.: **F AU=KREUTZ**, d. h. "Suche Dokumente mit Autor 'Kreutz'", **S S=2**, d. h. "Zeige die Dokumente der Zeilennummer 2", **S F=TI**, d. h. "Zeige die Treffer (die sog. "Zielinformation") der letzten Zeilennummer, jedoch nur die Titel" (wenn keine Zeilennummer angegeben wird, wird die letzte Zeile abgearbeitet).

Wie diese Beispiele zeigen, ist dabei auf den ersten Blick verwirrend, daß

- **F** sowohl **Find** (als Befehl) als auch **Field** (Datenfeld, als Spezifizierung beim Show-Kommando),

- **S** sowohl **Show** (als Befehl) als auch **Statement number** (= Zeilen-nummer erfolgreicher Treffer, als Spezifizierung) bedeuten kann.

Sofern weitere Spezifizierungen vorgenommen werden sollen, folgt ohne Blank ein *Semikolon* (Strichpunkt), ohne Blank gefolgt von der nächsten Spezifizierung, z. B.: **S S=2;R=1**, d. h. "Zeige das erste Dokument von Zeile 2". **R = Record number**: Die Treffer einer Zeile werden automatisch durchnummeriert, wobei die nach Publikationsjahr neuesten Dokumentnachweise zuerst gereiht werden. Es spart Zeit und Geld, wenn man sich bei einer großen Anzahl von Treffern nur die Titel zeigen läßt und die Nummern der interessant erscheinenden Dokumente notiert und sich diese dann in einem zweiten Schritt in vollem Format zeigen läßt.

Mittels Verknüpfung mit Strichpunkten können längere Zeichenketten gebildet werden. Dabei ist zu beachten, daß CCL-Systeme in der Regel keine Kommas akzeptieren! Sofern in einer Suchanfrage *zwei Befehle* in einer Kommandoeingabe sofort hintereinander verarbeitet werden sollen, sind die beiden Befehle statt mit Strichpunkten ohne Blanks mit einem *Ausrufzeichen* zu verbinden, z. B.: **F AU=FROEHLICH!S**, d. h. "Suche Dokumente mit Autor 'Froehlich' und zeige uns diese".

Wie das Beispiel zeigt, müssen Umlaute in der Regel aufgelöst eingegeben werden.

Bei Zahleneingaben reicht die Eingabe der Eckwerte unter Verwendung von "to", die Zahlen müssen dann jedoch von einem Leerzeichen getrennt sein, z. B. **S S=2;R=1 TO 5**, d. h.: "Zeige die in Zeile 2 gefundenen Dokumente, und zwar die ersten 5 Dokumente".

Bei komplexeren Suchstrategien unter Verwendung mehrerer logischer Operatoren (OR; AND; NOT) ist eine *Klammerung* erforderlich. Z. B.: **F (1 OR 2) and ((3 OR 4) NOT (5 OR 6))**

4.5 Maskierung und Kontextbestimmung

Unter Truncierung (von englisch truncate = abschneiden, eine dieser in der EDV üblichen fürchterlichen Eindeutschungen) oder *Maskierung* versteht man das Abschneiden und Offenlassen von Suchtermen (s. *Übersicht 9*). Das Maskierungssymbol ist i.d.R. ein Sonderzeichen, in CCL/GRIPS das Dollarzeichen (\$), sowie (leider unnötig verwirrend) Fragezeichen (?), Doppelkreuz (#) und Prozentzeichen (%), bei anderen Systemen meist

ein Fragezeichen (?), Sternchen(*) oder das Paragraphensymbol (§).

Anmerkung: Heute ist bei DIMDI das Sternchen * das Standardsymbol fuer Maskierung!

Wir können *Frontmaskierung* oder *Endmaskierung* betreiben, d. h. ein Suchwort vorne oder hinten abschneiden, anders ausgedrückt: *linksbündig* oder *rechtsbündig* truncieren . Das Frontmaskieren ist insofern nicht ungefährlich, als sie sehr hohe Kosten verursachen kann und - in früheren Zeiten - sogar den Host lahmlegen konnte. Beim Informationszentrum Karlsruhe (INKA), dem Vorläufer von STN (Scientific Technical Network), ist das vor Jahren einmal passiert, es war auf Betriebssystemebene nicht vorgesehen, solche Jobs zu stoppen.

Daher würde ich empfehlen: Falls nicht unbedingt erforderlich, *nicht im Find-Kommando in der gesamten Datenbank, im Freitext, linksbündig maskieren.*

- Wir können relativ problemlos in einzelnen Datenfeldern linksbündig truncieren, z. B. im Autorenfeld, und fordern: **F AU=\$MAIER**. Dann bekommen wir alle - maiers: Vordermaier, Hintermaier usw. Das Autorenfeld umfaßt vielleicht 5 % der Gesamtdatenbank. Auch eine Eingrenzung mit "publication year", "language" u.ä. Datenfeldern, verkleinert die Grundgesamtheit, in der die linksbündige Truncierung vorgenommen wird, beträchtlich.

- Wir können den gesamten Datenbankinhalt durch ein Suchwort eingrenzen, und bekommen z. B. eine Trefferzahl von 500. Innerhalb dieses kleinen Pools können wir ebenfalls problemlos und kostengünstig eine linksbündige Maskierung vornehmen.

Man kann auch im *Inneren* der Suchterme maskieren und einzelne Buchstaben offenlassen, um z. B. auf englisch "Frauen" in Ein- und Mehrzahl gleichzeitig suchen zu können (**f wom\$n** bzw. **f wom%n**).

Mittels Maskierung kann auch das Problem der *Amerikanismen* in den Griff bekommen werden, d. h. die unterschiedliche amerikanische und englische Schreibweise vieler Wörter (z. B. "behaviour" auf englisch, im amerikanischen "behavior"). Mit Doppelkreuz können wir in GRIPS Maskierungen mit "maximaler" Zeichenzahl durchführen, **d behavio#r** erfaßt also sowohl behavior als auch behaviour.

Das Beispiel "**d So\$iali\$a\$ion**" zeigt, wie wir mit einer einzigen Eingabe in internationalen, mehrsprachigen Datenbanken gleichzeitig in mehreren Sprachen nach Dokumenten, in denen "Sozialisation" in der jeweiligen Schreibweise vorkommt, recherchieren könnten. Allerdings lassen nicht alle Systeme mehrere Maskierungsvorgänge gleichzeitig zu.

Bei Maskierung im normalen Find-Kommando werden zuerst sämtliche im Wörterbuch der Datenbank enthaltenen maskierten Wörter mit Trefferzahlen aufgelistet und dann erst die Gesamttrefferzahl; da sich diese Liste u.U. über mehrere Bildschirmseiten erstrecken kann, empfiehlt sich oft zur Abkürzung und Beschleunigung die Eingabe des Zusatzes ALL; es erscheint sofort das Gesamtergebnis, z. B.: **F ALL Protest\$**.

Um die Präzision der Recherchestrategie zu erhöhen, können sogenannte *Kontextoperatoren* verwendet werden. Kontextoperatoren sind Operatoren, die den Kontext, *den Zusammenhang unseres Suchwortes mit anderen Suchwörtern* definieren. In der Recherchepraxis ist es - nicht zuletzt aus Kostengründen - sinnvoll, Recherchemüll zu vermeiden, indem ich Kontextoperatoren verwende.

In CCL sind Kontextoperatoren allerdings im Vergleich zu anderen Retrievalsystemen weniger wichtig, da in CCL nebeneinanderstehende Wörter ohne Eingabe von Kontextoperationen, einfach durch Eingabe nebeneinander, gesucht werden können; in fast

allen anderen Systemen sind dazu eigene Kontextoperatoren erforderlich. Wir können also in CCL den Titel eines Aufsatzes oder Buches als Satz eingeben, ohne die einzelnen Wörter mit Kontextoperatoren verbinden zu müssen. Dies ist gerade bei bibliothekarischen Datenbasen, wo es um die Verifizierung von Zitaten bzw. um Standortnachweise bereits bekannter Publikationen geht, sehr bequem.

Bereits die Auswahl eines Datenfelds ist eine Kontextbestimmung - in der Freitextrecherche können sich die eingegebenen, z. B. mit AND verknüpften Suchwörter in völlig verschiedenen Datenfeldern befinden. Im engeren Sinn bezeichnet man als Kontextoperatoren jene Operatoren, die die *Entfernung* eines Wortes zu anderen im selben Datenfeld bestimmen: In bestimmten Datenfeldern, z. B. im Titel oder im Abstract, sind ganze Sätze oder Fachausdrücke (etwa Bezeichnungen von Institutionen), die aus mehreren Wörtern bestehen, enthalten. Ob zwei Suchwörter nahe nebeneinanderstehen oder nicht, kann daher inhaltliche Bedeutung haben. Ohne Kontextbestimmung steigt die Gefahr unerwünschter Treffer.

Überdies bestehen die Fachtermini in der heute wichtigsten Wissenschaftssprache, dem Englischen, nicht aus solchen "Wortungetümen", d. h. Wortzusammensetzungen, wie sie im Deutschen möglich sind (vgl. den berühmt-berüchtigten "Donaudampfschiffahrtsgesellschaftskapitän"), sondern aus Aneinanderreihungen einzelner Haupt- bzw. Zeitwörter (So heißt die psychologische Variable "Kontrollerwartung" im englischen Original "locus of control").

Die Verwendung von Kontextoperatoren kann auch bei der Recherche von langatmigen Namen von Gesellschaften und Kommissionen nützlich sein. Wir können ein Wort, das am Beginn des Organisationsnamens vorkommen muß und eines vom Schluß nehmen, die Wörter dazwischen offen lassen. Ich kann "International" eingeben, die dahinter kommenden Füllwörter, wie z. B. "economic" oder "social" offenlassen, und zum Schluß "commission" suchen. Dann bekomme ich alle internationalen Kommissionen, ganz beliebig, welches Ziel sie verfolgen, zu welchem Zwecke sie gegründet wurden.

In CCL-GRIPS werden die Maskierungssymbole **?**, **#**, **%**, **\$** sowie Punkte (... oder ...) freistehend, also mit Blanks von den Suchwörtern getrennt, auch als *Kontextsymbole* verwendet).

4.6 Kosten von Recherchen

Wie bereits erwähnt, fallen Kosten an, sobald man eine Datenverbindung hergestellt hat. Die Gesamtkosten einer Recherche sind erstens abhängig von der *Anschaltzeit*, diese kostet meist zwischen 70 und 150 DM pro Stunde. Die Kosten für die Anschaltzeit setzen sich zusammen aus (a) Kosten für die Datenübertragung durch die Post/TV-Kabelanbieter und/oder einen Internet-Provider, (b) Gebühren des Hosts und (c) Lizenzgebühren der einzelnen Datenbasen.

Zweitens wird die Zahl der gezeigten *Treffer* (online am Bildschirm oder als Offline-Ausdruck durch einen Drucker des Hosts) je nach *Format* der angezeigten Dokumente (z. B. ob bloß mit bibliographischen Angaben oder inklusive Abstracts) berechnet. Das vollständige Format mit allen Datenfeldern ist am teuersten. Bei den meisten Datenbanksystemen ist pro Dokumentnachweis im Vollformat mit einer Gebühr von 1 DM und mehr rechnen. Die Anzeige eines einzelnen Datenfeldes (z. B. nur der Titel) verursacht

in der Regel abgesehen von den Anschaltkosten keinerlei zusätzliche Kosten. Wie bereits erwähnt, empfiehlt es sich daher, bei größeren Trefferzahlen zuerst die Titel zu sichten und dann die interessanten Dokumente über die record numbers anzufordern.

Am langsamsten arbeiten die (wie bereits erwähnt, fast schon "steinzeitlichen" kostengünstigen Akustik-Koppler. Die Kosten für die Anschaltzeit können durch schnellere *Modems* verringert werden. Allerdings verlangen manche Hosts bei hohen Übertragungsgeschwindigkeiten auch höhere Gebühren, sodaß sich die Kostenersparnis durch schnelle Modems in Grenzen halten kann. Auch die Post berechnet nicht nur die Anschaltzeit, sondern auch die übertragene Datenmenge. Viele Hosts sind, s.o., auch über *Wissenschaftsnetze* zugänglich, wobei es z. B. möglich ist, von universitätsinternen Netzen über das jeweilige nationale Forschungsnetz in entsprechende ausländische bzw. internationale Netze zu gelangen. Diese Netze sind kostenlos oder äußerst kostengünstig.

Eine Verkürzung der Anschaltzeit kann durch die Unterdrückung der Pausen bei der Ausgabe im Show-Kommando durch das Kommando **hc (hard-copy)** erreicht werden. Alle angeforderten Dokumentnachweise werden dann ohne Pause und "more"-Vorgabe ausgegeben, es ist aber auch kaum eine Unterbrechung möglich. Daher sollte man diesen Befehl nur verwenden, wenn man sich seiner Sache sicher ist - "Stornierungen" sind nicht möglich. (22)

Manche Hosts bieten laufend, insbesondere bei der Neueinführung von Datenbasen, sog. "*happy hours*" an, d. h. bestimmte Zeitabschnitte, in denen eine lizenzgebührenfreie Erprobung dieser Datenbasen möglich ist. Diese Zeitabschnitte liegen in verkehrschwachen Zeiten, meist an Freitagnachmittagen. Üblich sind auch sog. "Mondscheintarife", d. h. verbilligte Tarife am Abend und am Wochenende. Auch die Wartezeiten im Internet sind höchst unterschiedlich - am schnellsten: am 24. Dezember - obwohl zahlreiche Ascii-Tannenbäume und -Engel kursieren.

Für die Recherche in den einzelnen Datenbasen verrechnen die Hosts unterschiedlich hohe *Lizenzgebühren*, da sich die Datenbasen nach Umfang und Herstellungskosten beträchtlich unterscheiden. Voll öffentlich finanzierte Hosts, etwa der Host ECHO der Europäischen Gemeinschaft in Luxemburg und Universitätsrechenzentren, erheben in der Regel keine Gebühren. Der Großteil der Hosts ist jedoch privat oder im öffentlich-rechtlichen Vorfeld und wird nur teilweise subventioniert. Host- und Lizenzgebühren sind auch nach *Benutzerstatus* gestaffelt. So werden meist öffentlichen Institutionen, insbesondere Bildungseinrichtungen, bei öffentlich subventionierten Datenbanksystemen günstigere Tarife verrechnet als Privatfirmen. Auch die Nationalität, ja sogar die Zugehörigkeit zu einem bestimmten Bundesland (das den jeweiligen Host subventioniert) kann die Gebührenhöhe beeinflussen. Bei kommerziellen Hosts hingegen wirkt sich z. T. ein hoher "Umsatz", d. h. eine hohe Zahl an Recherchestunden pro Jahr, gebührenermäßig aus. Es werden die Kosten getrennt nach einzelnen Benützern desselben passwords nach einzelnen Datenbasen getrennt aufgeführt. Bei Recherchen per telnet, Gopher, WWW über Wissenschaftsnetze entfallen zwar die Datenübertragungskosten, stattdessen sind u. U. aufgrund mangelndem Recherchekomfort Host- und Datenbankgebühren höher.

Zu beachten ist, daß auch bei der Konsultation der Hilfsfunktion (help) oder bei Nachdenkpausen Kosten entstehen, sowohl Datenübertragungsgebühren der Post als auch die Lizenzgebühren der jeweiligen Datenbase. Vom Host wird, wie bereits erwähnt, in der Regel die *Anschaltzeit* (und nicht oder nicht nur die eigentliche Rechenzeit) als Grundlage für die Gebührenberechnung genommen (neben Anzahl und Format der online gezeigten oder offline ausgedruckten Treffer), egal, ob Sie während der Anschaltzeit tatsächlich recherchieren, oder überlegen, in einem Nachschlagewerk nachblättern, oder die help-

Seiten lesen. Eine gründliche Vorbereitung der Recherche - sowohl was die Suchbegriffe als auch die benötigten Befehle und Datenfeldbezeichnungen betrifft - ist daher zu empfehlen (vgl. auch Pkt. 4.7). Ein kleiner Trick: Man kann die Recherche in einer kostengünstigen Datenbank aufbauen und hier auch günstig die Hilfsfunktionen in Anspruch nehmen und dann mit dem kompletten Rechercheprofil in teurere Datenbanken überwechseln (vgl. Pkt [4.7](#)).

(Reichlich veraltet anmutende) postalische Recherchen sind empfehlenswert, sofern Informationszentren für Studierende beträchtlich reduzierte Sondertarife anbieten bzw. die Recherche nur über eine unkomfortable Kommandosprache durchführbar ist oder (wie z. B. bei der österreichischen Philosophiedatenbank PHILO in Graz (kostenlos), aber auch bei den Forschungs- und Literaturdatenbanken der Bundesanstalt für Arbeit in Nürnberg; postalische Recherchen erfolgen gegen Gebühr) bis jetzt noch kein Online-Zugriff angeboten wird.

Viele kleinere Datenbanken können (aufgrund von Beschränkungen der verwendeten Hard- und Software) keine Online-Recherchemöglichkeit anbieten und offerieren daher Recherchen auf der Basis telefonischer oder postalischer Anfragen. Z. T sind diese Recherchen kostenfrei, etwa bei der Österreichischen Philosophiedatenbank in Graz.

Nachteil von Recherchen, die in Auftrag gegeben werden: geringer Einfluß auf die Suchstrategie, wenig Möglichkeiten der Fehlerkontrolle. Falls möglich, sollte man selbst Suchwörter aus Thesauri heraussuchen und im Suchantrag mitteilen sowie die Übersendung des *Suchprotokolls* verlangen (beinhaltet die Suchstrategie inkl. Verknüpfungen, Trefferzahlen, Einengungen). Leider ist eine unaufgeforderte Übersendung von Suchprotokollen kaum üblich.

Darüber hinaus gibt es auf vielen Gebieten, gerade auf dem Gebiet der Geisteswissenschaften und Kultur, noch immer zahlreiche *konventionelle Informationsdienste*, die gratis oder gegen meist geringe Gebühr Kopien von Karteikarten oder Zeitungsausschnitten übermitteln. Sehr kostenintensiv sind hingegen Medieninformationsdienste (Rundfunk- und TV-Archive). Auch über das WWW bieten Nachrichtenagenturen und Wochen- und Tagespresse archivarische Dienstleistungen an, meist nur gegen Registrierung, und dies auch oft kostenpflichtig.

4.7 Die Optimierung von Recherchestrategien

Im vorliegenden Skriptum können nur elementare Kenntnisse vermittelt werden. Es gibt jedoch zahlreiche Möglichkeiten, die Recherchestrategien zu optimieren, d. h. sie einerseits eleganter, also zeit- und kostensparender abzuwickeln, sowohl was die Befehlsstruktur als auch die Selektion der Datenbanken und der Suchwörter betrifft. Hier können nur einige dieser Möglichkeiten erwähnt werden:

- Bei einem *Datenbasenwechsel* innerhalb desselben Hosts ist es nicht erforderlich, die gesamte Recherchestrategie noch einmal einzugeben. Gleichzeitig mit der Wahl der gewünschten Datenbank können bestimmte Zeilen mit erfolgreichen Suchwörtern (oder Verknüpfungen) oder die gesamte Profiltabelle in die nächste Datenbank *übernommen* werden, diese werden automatisch abgearbeitet, z. B:

base solis;s=all oder **s=1 to 10** oder **s=1;3;7**).

Zumindest beim Datenfeld Autoren kann es aufgrund inkonsistenter Feld- Bezeichnungen (sowohl **AU** als auch **PER**) zu Problemen kommen; außerdem ist bei manchen CCL-Datenbasen eine Spezifizierung nicht nur durch Eingabe eines *Prefix* (23) vor dem Suchterm (**F AU= ACHATZ**), sondern auch durch Anhängen eines *Suffix* (24) nach dem eingegebenen Suchwort möglich, das von anderen Datenbasen nicht immer übernommen werden kann, z. B.: **F ACHATZ/per**.

- Eine gewisse Arbeitserleichterung kann auch die Speicherung der Suchstrategie bringen: Mittels **SAVE** (Kurzform: **SA**) kann die Profiltabelle auf längere Zeit für spätere Recherchen unter einem beliebigen Namen gesichert werden, mit **DELETE** (Kurzform: **DEL**) können einzelne Zeilen oder die gesamte Profiltabelle wieder gelöscht werden. (25)

- Bei der Durchführung der Recherche empfiehlt sich stets der *zeilenweise Aufbau*. Es ist zwar möglich, sämtliche Suchwörter und Verknüpfungen in einem Satz einzugeben. Doch ist es ratsam, die Suchwörter einzeln einzugeben und erst in einem zweiten Schritt die Zeilen mit Treffern zu verknüpfen. Warum? Gerade als Anfänger vertippt man sich leicht, oft ist ein Suchwort unzutreffend und muß durch ein besseres ersetzt werden, manchmal nimmt man eine zu enge Eingrenzung vor und erhält keine oder zu wenige Treffer. In diesem Fall müßte die gesamte Recherche wiederholt werden oder zumindest aufwendig (durch Substraktion und Ergänzungen) korrigiert bzw. modifiziert werden. Diesen Mehraufwand kann man durch vorsichtigen zeilenweisen Aufbau der Recherchestrategie vermeiden.

Einen Kompromiß zwischen Geschwindigkeit und Vorsicht stellt dabei das Spezialkommando **FIND STEPS** (abgekürzt **FS**) dar. Mit FS wird die gesamte Recherche in einem Satz akzeptiert, aber schrittweise dokumentiert, d. h. zeilenweise ausgedruckt. Wir erhalten die Trefferzahlen der einzelnen Suchwörter jeweils in einer Zeile, getrennt davon die Unterverknüpfungen und in der letzten Zeile die Trefferzahl der Gesamtverknüpfung.

Dem stufenweisen Aufbau der einzelnen Recherchen entspricht der stufenförmige, genauer der *spiralförmige Aufbau des gesamten Recherchevorhabens*.

Es gibt keine absolute Treffersicherheit bei Datenbankrecherchen, das Recherchieren ist tatsächlich ein "Fischen in trüben Gewässern", man muß trotz aller gründlichen Vorbereitungen mit dem Zufall leben und auch seine Intuition entwickeln. Man sollte sich nicht zum Ziel setzen, mit einer Recherche auf Anhieb alle einschlägigen Publikationen oder Projekte zu einem Thema zu finden. Das ist auch gar nicht notwendig:

- Bei der Recherche können wir *spiralförmig* vorgehen, rekursiv, wir machen eine erste, explorative Recherche und lernen dazu, wir entdecken u. a. neue Suchwörter und Schlüsselautoren; wir verwenden andere Datenbanksysteme, andere Nachschlagewerke, lesen auch einige der eruierten Originaldokumente, werten die Texte, aber vor allem die Literaturverzeichnisse aus (nach Fachausdrücken in den Titeln der zitierten Publikationen, in den Zeitschriftentiteln nach Ausdrücken nach Sachgebieten) und machen einen zweiten Durchgang.

Die Auswertung von Recherchen kann auch EDV-gestützt erfolgen: Mit dem GRIPS-Kommando **EXTRACT (EX)** können *Relevanzwerte von Suchbegriffen* auf der Basis des Häufigkeitsvorkommens in den Rechercheergebnissen (relativiert durch das Häufigkeitsvorkommen in der gesaten Datenbase) bestimmt werden. EXTRACT erlaubt auch die Ausgabe einfacher *quantifizierender Übersichten* (etwa: wieviele Studien stammen jeweils aus welchem Land?), sowohl in absoluten Zahlen als auch prozentuell

(vgl. das Recherchebeispiel in DIMDI-psyndextrain im Materialenteil, dort allerdings exemplarisch nur auf der Basis von n=4 recherchierten Dokumenten durchgeführt).

Standardmäßig wird durch das EXTRACT-Kommando eine Zufallsstichprobe von maximal 40 Dokumenten ausgewählt. Der betreffende Parameterwert kann jedoch auf maximal 2000 erhöht werden. Damit sind, direkt im Hostdialog, bereits durchaus sinnvolle Auswertungen für Zwecke der quantitativen Wissenschaftsforschung (*Scientometrie*) möglich.

Insbesondere für Anfänger oder Ungeübte empfiehlt sich die *gründliche Vorbereitung* der Recherche. Alle erforderlichen *Unterlagen* (Schlagwortverzeichnisse, Kurzübersichten der Befehle etc.) sollten griffbereit neben dem Terminal liegen. Dabei ist zu berücksichtigen, daß, wie bereits erwähnt, auch beim selben System (!) in verschiedenen Datenbasen u.U. *unterschiedliche* Datenfeldbezeichnungen verwendet werden müssen. Daher sollten auch die jeweiligen "*blue sheets*" bzw. "*Memokarten*" bereitliegen, d. h. die Datenbasenbeschreibungen des Hosts, bei dem wir recherchieren möchten (da ein und dieselbe Datenbank bei verschiedenen Hosts unterschiedliche Feldspezifikationen erfordern kann!). Besonders für Anfänger empfiehlt es sich, die Recherche in allen Einzelheiten, inkl. Befehle und korrekter Syntax zuerst "trocken", d. h. auf Papier *zu notieren* und auf Fehler zu kontrollieren. Komfortable Retrieval-Programme (z. B. *Genesys*) erlauben es auch, die einzelnen Rechenschritte (Befehle, Suchbegriffe, Verknüpfungen) in allen Einzelheiten bereits vor Beginn der Online-Recherche einzugeben, auf Eingabefehler (nicht zuletzt Tippfehler) zu überprüfen, sie während der Recherche aufzurufen und zum Host zu übertragen.

Grundsätzlich gilt: Bei Recherchen, die *keinerlei Resultate* erbringen, ist stets der Verdacht begründet, daß *wir* einen *Fehler* gemacht haben. Sofern nicht im Datenfeld "Deskriptor" (Controlled Terms) ein Nicht-Deskriptor eingegeben wurde, (26) ist ein Null-Ergebnis insbesondere in internationalen Datenbanken sehr unwahrscheinlich. Kein Suchbegriff ist zu selten oder zu abstrus, als daß er nicht zumindest in einigen Publikationen - wenn auch oft in einer anderen als der von uns gemeinten inhaltlichen Bedeutung - vorkommen würde.

Viele Probleme bei Datenbankrecherchen sind jedoch nicht oder nicht nur auf mangelnde Kompetenzen der Nutzer zurückzuführen, sondern auf Mängel der Datenbanksysteme selbst. So ist in den USA bereits eine "Gewerkschaft" der Datenbankbenutzer gegründet worden, die vehement Reformen und Verbesserungen in Richtung Benutzerfreundlichkeit fordert, insbesondere wird die Standardisierung von Kommandosprachen bzw. Datenbankfunktionen gefordert. Möglicherweise wird auch die Konkurrenz durch die bunten Angebote in der Internet-Welt den bisher nur mäßig kundenfreundlichen konventionellen Datenbanken gut tun. Manche, bisher kostenpflichtig, haben sich sogar inzwischen kostenfrei ins WWW gewagt (Medline).

Ein Problem ist die Fusionierung zahlreicher Recherchetreffer zu einer Gesamtbibliographie. Abgesehen von gewissen Copyright-Problemen hat fast jede Datenbank, sogar innerhalb eines Hosts wie z. B. DIALOG, unterschiedliche Formate bei der Aufnahme der Literatur entwickelt. Es werden allerdings bereits Programme angeboten, die die Lösung dieser Probleme versprechen (und sie müssen für jede Datenbank extra ein Makro enthalten, d. h. ein Subprogramm, welches die Transformierung der Daten in jeweils datenbankspezifischen Format in das eigene gewünschte Format vornehmen), doch ist bei diversen Literaturgattungen wie Beiträgen in Kongreß-Proceedings, Sammelbänden mit kollektiven und von der Organisation beauftragten individuellen Herausgebern und diverser Grauer Literatur oft auch ein solches Programm überfordert.

5. Schlußbemerkung

Gründliche Recherchen sind unverzichtbar, um die *Kumulativität des wissenschaftlichen Erkenntniszuwachses* zu ermöglichen. Dieser verläuft nicht konfliktlos- und geradlinig, sondern durch Kritik und Verwerfung irriger Hypothesen bzw. Theorien. Doch dazu müssen sich Theorien und potentiell falsifizierende empirische Befunde überhaupt finden können: Im Zeitalter der Informationsüberflutung ist dies keine selbstverständliche Angelegenheit. Im Sinne des wissenschaftstheoretischen Postulats der *intersubjektiven Überprüfbarkeit* ist auch die Angabe der Recherchequellen (m.E. auch der verwendeten Suchstrategien) bei *Übersichtsarbeiten* (reviews) unverzichtbar, ebenso für viele Arbeiten im Bereich der Wissenschaftsforschung. Viele Übersichtsarbeiten verdienen diese Bezeichnung eigentlich nicht, da einzelne Arbeiten willkürlich-zufällig gefunden und ausgewählt wurden, d. h. keine systematische Recherchen durchgeführt wurden. Der "Überblick" ist so oft eher ein Übersehen im negativen Sinn des Wortes (etwas nicht zu sehen) und nicht im Sinne eines Panoramablickes, einer repräsentativen Zusammenfassung.

Gute Wissenschaft ist jedoch *Informationsverdichtung*. Vom bekannten Mathematiker Hilbert stammt der vielzitierte Satz, *daß man den Wert einer wissenschaftlichen Arbeit daran ermessen könne*, wieviele andere wissenschaftliche Arbeiten zu lesen man sich bei der Lektüre dieser Arbeit erspare.

6. Literatur

Adam, L. R. (1982): Language and information retrieval in the social sciences. *Aslib Proceedings* 34, 394-405

Allischewski, H. (1976): *Bibliographiekunde*. Wiesbaden

Anderla, G. (1973): *Information in 1985. A forecasting study*. Paris

Anderla, G. (1975): Vorstellungen der Europ. Gemeinschaften über die Verwirklichung eines Informationsnetzes in Europa. *Nachrichten für Dokumentation*, 26, S. 221-226

Angione, P. V. (1975): On the equivalence of Boolean and weighted searching based on the convertibility of query forms. *Journal of the American Society for Information science* 26, 112-124

Artus, H. M. (1984): Graue Literatur als Medium wissenschaftlicher Kommunikation. *Nachrichten für Dokumentation* 35(3), S. 139-146

Atherton, P. (1978): Standards for a User System Interface Lang. Online Retrieval Research 1 (2) März 1978, S. 57-61

Atteslander P. (1977): Ist der Wissenschaftler überfordert - ? überfordert die Wissenschaft die Kommunikationsgesellschaft?, in: Stachowiak, H. (Hg.): Technologie und Zukunftssicherung. Paderborn

Austin, D. (1984): Automatisierung in der Sacherschließung der British Library. Bibliothek 8 (1), 45-57

Bärtsch, E. (1979): Die Bibliographie. München

De Bakey, L. (1965): Instruction in scientific communications. Journal for Medical Education 40

Bayer, A. E./Folger, J. (1966): Some correlates of a citation measure of productivity in science. Sociology of Education. 1966, 39

Batori, I. (1969): Automatische Textanalyse für die maschinelle Informationsverarbeitung. Nachrichten für Dokumentation 20, S.123-7

Becker, H. S. (1994): Die Kunst des professionellen Schreibens. Ffm/N.Y.

Becker, J. (1981): Euro-amerikanische Konflikte im Bereich des grenzüberschreitenden Datenflusses. Nachrichten für Dokumentation 32, 220ff.

Becker, J./Bickel, S. (1992): Datenbanken und Macht. Opladen

Beling G./Wersig G. (1972): Zur Typologie von Daten und Informationssystemen. Pullach b. München

Bellomy, F. O. (1979): Die privatwirtschaftliche Informationsversorgungszene (information brokerage). Nachrichten für Dokumentation 30 (1), S. 15ff.

Beyer, A. et al (1976): Wirtschaftswissenschaften und wissenschaftliches Isystem in der DDR. Erlangen

Beyer, W. (1981): Workshop. Informationsvermittlung und -beratung im Bereich der Fachinformation. Nachrichten für Dokumentation 32, 227ff.

Beyer, Wolfgang (1982): Informationsvermittlung in der Bundesrepublik Deutschland. Ffm.

Bezsonoff, D. (1981): L'utilisation des actuels fichiers bibliographiques et des futures banques données, en bio-médecine. Documentaliste 18 (3)

Bisco, R. L. (1966): Social science data archives: a review of developments. American political science review 93

Bisogno P. (1978): Some observations about information and documentation in the social science in Italy. Information processing and management 14 (3/4). S. 243-245

Bittner, E. (1984): Textreferat und sozialwiss. Textproduktion. Soziologie 1/1984, 13-48

BMWVK (1996a): Information und Kommunikation. Schlußbericht des Bundesministers für Wissenschaft, Verkehr und Kunst im Rahmen der Initiative der Bundesregierung "Österreichs Weg in die Informationsgesellschaft"
<http://www.bmwf.gv.at/7forsch/infkom/pginfkom.htm>

BMWVK (1996b): Forschungsbericht 1996 Schwerpunkt Informationstechnologien. Bericht des Bundesministers für Wissenschaft, Verkehr und Kunst an den Nationalrat gemäß § 8 Abs. 1 des Forschungsorganisationsgesetzes.
<http://www.bmwf.gv.at/7forsch/fober96/pgfober.htm>

Bollmann, S./Heilbach, C. (1996): Kursbuch Internet. Anschlüsse an Wirtschaft und Politik, Wissenschaft und Kultur. Mannheim

Bonnely, C./Drolet, G. (1978): Searching the social sciences literature online: SOCIAL SCI SEARCH. Database 1 (2), 10-25

Bonzi, S. (1984): Terminological consistency in abstract and concrete disciplines. Journal of Documentation 40, 4-1984, S. 247-263

Brittain, J. M. (1976): Information services in the social sciences: Development and rationalization Int Soc Sci J 28 (4)

Bruder, H. E. (1980): Nichtnumerische Datenverarbeitung. Mannheim etc.

Bürger, J. H. (1984): Jobs in der Informationsgesellschaft. Berufe und Chancen in der Kommunikation. Essen

Burrows, S./Kyle, S. (1979): Searching the MEDLARS file on NLM and BRS: a comparative. Bulletin of the Medical Library Association 67 (1) 1979, S. 15-24

Caldwell, J./Ellingson, C. (1979): A comparison of overlap: ERIC and Psychological Abstracts. Database 2/1979, 62-67

Caruso, E. (1977): Trainer for Online Systems. Online Oct. 1977, 28-34

Chubin, D. E./Moitra, S. D. (1975): Content analysis of references: adjunct or alternative to citation counting? Social Studies of Social Sciences 5, 423-441

Cole, J. R./Cole, S. (1973): Social stratification in science. Chicago/London

Cole, J. R./Cole, S. (1973): Measuring the quality of sociological research: problems in the use of the science citation index. American Sociologist 6, 1973

Conger L. D. (1980): International resources online in the social sciences: the roles in the Swiss cheese. Behavioral and Social Science Librarian 1 (4)

Cook, T. D./Leviton, L. C. (1981): Reviewing the literature: a comparison of traditional methods with meta-analysis. Journal of Personality

Cooper, H. M./Rosenthal, R. (1980): Statistical vs. traditional procedures for summarizing research findings. Psychological Bulletin 87

Corkill, C./Mann, M. (1978): Information needs in the humanities: two postal surveys.

Sheffield

Clay, K. (1979): Searching ERIC on DIALOG: the times they are a'changing. Database 3/1979

Crane, D. (1972): Invisible colleges diffusion of knowledge in scientific communities. Chicago/London

Crawford, S./Devlin, M. (1978): Communication in the health sciences. International Journal of Dermatology 17, 514-520

Cronin, B. (1982): Databanks. Aslib Proceedings 33, (6/1981), 243-50

Cronin, B. (1984): The citation process. London

Cuadra, C. A. (1975): SDC experiences with large data bases. Journal of chemical information and computer sciences 15, 48-51

Cuadra, C. A. et al. (1976): Impact of on-line retrieval services. A survey of research 1975-75, Santa Monica

Detemple, S./Heidtmann, F. (1986): Wie finde ich philosophische Literatur. Berlin

Dolan, D. R. (1978): Subject searching of the New York times Information Bank. Online 2 (2), 26-30

Diekmann, J. J (1981): Urheberrecht und Fachkommunikation. Nachrichten für Dokumentation 32 (6), 221-227

Ditfurth, C. v. /Kathöfer, U. (1997): Internet für Wirtschaftswissenschaftler. Ffm./N.Y.

Elchesen, D. R. (1978): Cost-effectiveness comparison of manual and on-line retrospective bibliographic searching. Journal of the American Society for Information Science 29/1978, S. 56-66

Engelbert, H. (1974): Informationsbarrieren. Informatik 21 (5)1974. S. 51-54

Essyad, M. F. (1974): Surveys of research in social sciences. Social Science Information 13(1) 163-178

Faibisoff, S. G. (1981): Is there a future for the end user in online bibliographic searching? Special libraries 1981, 347-355

Fidel, R. (1984): Online searching styles. A case-study-based model of searching behaviour. Journal of the American Society for Information Science 35 (4), 211-221

Fix, C. et al. (1964): Some characteristics of the review literature in eight fields of science. Washington

Foster, A. (1981): Which database? euronet/Diane news Nr. 26, S.2

Frey, D. (1981): Informationssuche und Informationsbewertung bei Entscheidungen. Bern etc.

Fricke, Meta-Analyse

Friedländer, F. (1972): Die weiße und die schwarze Forschung. Gruppendynamik 1972 2 (1)

Fröhlich, G. (1981): Kumulativer Erkenntniszuwachs? Inkonsistenz empirischer Befunde und Barrieren der wissenschaftlichen Kommunikation in der Geschwisterpositionsforschung. Wien: Grund- und Integrativwissenschaftliche Fakultät (Diss.; Modellrecherche zur Geschwisterforschung)

Fröhlich, G. (1988): Beiträge von Information und Dokumentation zur Verbesserung von Wissenschaftskommunikation und Forschungssynthetisierung, in: Kreuzt, H. (Hrsg.): Pragmatische Soziologie. Opladen.

Fröhlich, G. (1994a): Der (Mehr-)Wert der Wissenschaftskommunikation, in: Rauch, W. (Hg.): Der Mehrwert der Information. Konstanz: Universitätsverlag Konstanz. Schriften zur Informationswissenschaft 16, 84-95

Fröhlich, G. gem. mit Mörth. I. (1994b): Leitfaden durchs Labyrinth der Gelehrsamkeit Pierre Bourdieus. Eine kontextorientierte und referentielle Dokumentation, in: Mörth, I., Fröhlich, G. (Hg.): Das symbolische Kapital der Lebensstile. Zur Kulturosoziologie der Moderne nach Pierre Bourdieu, Ffm./ NY.: Campus Verlag, 271-311, ISBN 3-593-34964-7

Fröhlich, G. (1995): Demokratisierung durch Datenbanken und Computernetze? in: Becker, T. et al. (Hg.): Informationsspezialisten zwischen Technik und gesellschaftlicher Verantwortung. Stuttgart, 55-60 bzw. <http://www.uni-stuttgart.de/UNluser/hbi/publikat/hbipubl/guides/froehlic.htm>

Fröhlich, G. (1996): Surplus Values of Scientific Communication. Review of Information Science. <http://www.inf-wiss.uni-konstanz.de/RIS/>, I (2), http://www.inf-wiss.uni-konstanz.de/cgi-bin/isr10/kom/WWW/documents/pub/RIS/1996iss02_01/articles01/02?BUTTON=on&GIF=thumb

Fröhlich, G. (1996a): Netz-Euphorien. Zur Kritik digitaler und sozialer Netz(werk-)metaphern, in: Schramm, A. (Hg.): Philosophie in Österreich 1996, Wien, 292-306

Fröhlich, G. (1996b): Betrug und Täuschung in den Wissenschaften. kursiv 3(3), 26-30

Fröhlich, G. (1997a): Mythos Informationsgesellschaft? Informationsblockaden und Desinformation in Wissenschaft und Gesellschaft, in: Meleghy, T. et al. (Hg.): Soziologie im Konzert der Wissenschaften. Opladen, 328-336

Fröhlich, G. (1997b): Techno-Utopien der Unsterblichkeit aus Informatik und Physik, in: Becker, U. et al. (Hg.): Tod im Wandel. Neukirchen-Vluyn: Neukirchener Verlag

Fröhlich, G. 1997c (gem. mit Mörth, I.): Twelve Languages, Seven Disciplines, Five Continents: A Comprehensive Bibliography of Clifford Geertz, in: Fröhlich, G./Mörth. I. (Hg.): Symbolische Anthropologie der Moderne? Kulturanalysen nach Clifford Geertz. Ffm./N.Y.: Campus Verlag, 213-239 (in Druck)

Gallo, P.S.: (1978): Meta-analysis - a mixed metaphor. American Psychologist 33, 515-7

Galvin, T.J. (1984): Die Bedeutung der Informationswissenschaft für Theorie und Praxis

des Bibliothekswesens. Bibliothek 8 (2), 105-108

Gardner, T./Goodyear, M. L. (1977): The inadequacy of interdisciplinary subject retrieval. *Special Libraries* 68, (5/6), 193-197

Garfield, E. (1976): La science francias estelle grop provinciale? *La Recherche*, 7/1976, 757-760

Garfield, E. (1977): *Essays of an information scientist*. Philadelphia, 2. Bände

Garfield, E. (1979a): *Citation Indexing*. N.Y. etc.

Garfield, E. (1979b): *Citation Indexing: The Evolving Grammar of Citation Analysis*. In *Current Contents* 28/July 16 (5-6), 7-10

Garfield, E. (1979c): *Current Contents: its impact on scientific communications*. *Interdisciplinary Science Reviews* 1979

Garfield, E./Sher, I. H. (1967): *Diagonal Display - a new technique for graphic representation of complex topological networks*. Philadelphia

Garvey, W. D. (1979): *Communication: the essence of science. Facilitating information exchange among librarians*. Oxford

Garvey, W. D./Griffith, B. C. (1966) *Studies of social innovations in scientific communication in psychology*. *American Psychologist*, 21, 1019-36

Garvey, W. D./Griffith, B. C. (1971): *Scientific communication*. *American Psychologist* 26 (4), 349-362

Garvey, W. D. (1970): *Some comparisons of communication activities in the physical and social sciences*, in: Nelson C. E./Pollock, D. K. (Eds.): *Communication among Scientists and Engineers*. Lexington Mass., 61-84

Gergely, S. M. /Simmler O.A./Hauffe H. (1981): *Online-Recherchen für Fernsehjournalisten - Erste Ergebnisse eines Pilotprojektes*. *Nachrichten für Dokumentation* 32 (4/5)

Gewecke, W. (1980): *Odin - der nationale Informationsverbund*. *Das Inforum* 5/April, 6-7

Gibbs, J. M./Phillips, E. (1979): *Bessere Zeiten für graue oder nicht herkömmliche Literatur*. *Bibliothek. Forschung und Praxis* 3, 122-126

Gillespie, P. D. et al. (1979): *Problems of document delivery for the EURONET user*.

Glass, B. V. et al. (1981): *Meta-analysis in Social Research*. Beverly Hills/London

Goldsmith, Ch. A. (1977): *The teaching of sociology, 1975 - 1976. Annotated bibliography and review*. *Teaching Sociology* 4 (3), 341-368

Gordon, M. D./Santman, A. (1981): *Language barriers, literature usage and the role of reviews; an international and intedisciplinary study*. *Journal of information science* 3, 185-189

- Graus, W. et al. (1975): Menschliche Kommunikation in technischen Kommunikationssystemen. *Öff. Verwalt. u. Datenverarbeitung* 5 (1), 5-14
- Grishman, R./Hirschman, L. (1978): Question answering from natural language medical data bases. *Artificial Intelligence* 11, 25-43
- Gronlund, B./Nilsson, H. (1977): SCANNET, An I+D Data Network for the Nordic Countries. First On-Line-Information-Meeting, London 13. - 15. Dec. 1977. Oxford/N.Y., 66-70
- Groß, B./Stäber, P. (1981): Vom Mythos der Informationsgesellschaft - Zu einigen weltanschaulichen Aspekten der Informatik. *Informatik* 28 (1981), 40-45
- Gruch/Nubling (1978): Search corner ... comparison ... *Online Review* 2 (2), 122-123
- Grund, U./Heinen, A. (1995): *Wie benutze ich eine Bibliothek? Basiswissen - Strategien - Hilfsmittel.* München
- Grundlach, F. W. (1967): Die Universitäten im internationalen technischen Zeitalter. *IBM Nachrichten* 17, Nr. 184, 578-584
- Grünzig, H. J. et al. (1976): EVA - Ein Programmsystem zur maschinellen Inhaltsanalyse von Psychotherapieprotokollen. *Medizinische Psychologie* 2, 208-217
- Gudhe, E. et a. (1972): *Thesaurus Bildungsvorschung. Verzeichnis der Deskriptoren und Nichtdeskriptoren in der Literaturdokumentation des Max-Planck-Institutes für Bildungsforschung.* München/ Pullach/Berlin
- Gunzenhäuser, R. (Hg., 1968): *Nicht-numerische Informationsverarbeitung.* Wien/N.Y.
- Härder, T. (1978): *Die Implementierung von Datenbanksystemen.* München
- Halbach, U. et al. (Hg., 1981): *Information und Dokumentation in der Biologie.* Schlüchtern
- Hale, B. M. (1970): *The subject bibliography of the social sciences and humanities.* Oxford - Braunschweig
- Hall, J. L./Brown, M. J. (1981): *Online Bibliographic Databases.* London
- Halmos, P. (Ed., 1972): *The sociology of science.* Keele, Staffordshire
- Hawkins, D. T. (1981): *Online Information Retrieval Systems.* Williams, M. E. (Hg.): *Annual Review of Information Science and Technology* 16, 171-208
- Heidtmann, F. (1977): *Wie finde ich sozialwissenschaftliche Literatur?* Berlin
- Heidtmann, F./Habermann, A. (1976): *Wie finde ich soziologische Literatur?* Berlin
- Heidtmann, F./Ulrich, P. S. (1978): *Wie finde ich film- und theaterwissenschaftliche Literatur?* Berlin
- Heinrichs, N. (1970): *Philosophie-Datenbank. Bericht über das Philosophy Information Center an der Universität Düsseldorf.* *Conceptus* 4, 133-144

Henry, W. M. et al. (1980): Online searching. Sydney - London - Boston.

Herschfield, A. (1972): Information counsellors: new profession? Atherton, Pauline: Humanization of knowledge in the social sciences, a symposium. Syracuse, N.Y., 29-34

Hödl, G. et al. (1981): Maschinelle Information und Dokumentation am Beispiel der Österr. Hist. Bibliographie. Nachr. f. Dokument. 32 (1981), 135 ff.

Hofer, P. (1979): Informationsprobleme der Wirtschaft und Ansatzpunkte. Nachrichten für Dokumentation 30 (1)

Heine, M. G. (1973): ... Sources of citations and references for analysis purposes: a comparative assessment. Journal of Documentation 29, 73-74

Hennes, W. (1995): Informationsbeschaffung Online. Wettbewerbsvorteile durch weltweite Kommunikation. Ffm./N.Y.

Hitzeroth, C. et al. (1976): Leitfaden für die formale Erfassung von Dokumentation in der Literaturdokumentation. München.

Huly, H.-R./Raake, S. (1995): Marketing Online. Gewinnchancen auf der Datenautobahn. Ffm./N.Y.

Hutton, S. S./Hutton, S. R. (1981): Microcomputer data base management of bibliographic information. Sociol. Meth. Res. 9, 461-472

Hymes, D. (1965): The use of computers in anthropology. London

Jackson, G. B. (1980). Methods for integrative reviews. Review of reserach 50, 438-460

Janke, R. B. (1980): Searching the Social on BRS. Database 3 (2), 19-45

Jesse, A. (1975): Information systematisch gewinnen. Reinbek/H.

Jonas, H. E. (1979): Informationsverbund ODOOnline-Fernrecherche und Literaturbestellung. Zeitschrift f. Bibliothekswesen und Bibliographie. Sonderheft 29.

Jones, W. T. (1976): A fuzzy set characteristization of interaction in scientific research. Journal of the American Society for Information Science 27, 307-310

Kaiser, H. M. (1984): Ansätze zu einer Bibliothheorie. Bibliothek 8 (2), 92-103

Kaiberg, L. (1979): Wissenschaftliche Information in Dänemark heute - Probleme und Tendenzen. DFW-Dokumentation. Information 27, 3/4, 83-86

Kaminsky, R. (1974): Gedanken zur Verbesserung des Informationswesens eines Bundeslandes mit Hilfe eines Nachweissystems. ÖVD - Öffentliche Verwaltung u. Datenverarbeitung 7/1974

Kaminsky, R. (1977): Information und Informationsverarbeitung als ökonomisches Problem. München.

Kaminsky, R. (1983): Bibliotheken und wirtschaftliche Entwicklung. Zeitschrift für

Bibliothekswesen und Bibliographie 4/1983

Kaminsky, R. (1983): Erfahrungen eines privaten Informationsbrokers. Nachrichten für Dokumentation 34 (4-5), 195-200

Kaminsky, R. (1984): Wir informieren uns miserabel. Deutsche Universitäts-Zeitung, 157-58

Keitz, W. v./Zimmermann, H. H. (Hg., 1982): DOGE - Dokumentationsstellen Geisteswissenschaften

Kerbel, S. S. (1981): Searching Sociological Abstracts. Database, Juni 1981 4 (2), 230ff.

Ketnon, c./Scott, Y. B. (1978): MEDLINE searching and retrieval. Medical Information, 3 (3), 225-235

Knorr, K. D. (1977): Producing and reproducing knowledge: Descriptive or constructive? Social Science Information 16

Knorr, K. D. et al. (Hg., 1975): Determinants and controls of scientific development. Dordrecht.

Knorr, K. et al. (1981): Sozialwissenschaftliche Forschung in Österreich. Produktionsbedingungen und Verwertungsbedingungen. Wien.

Knorr, K. D. et al. (1981): The social process of scientific investigation. Dordrecht.

Knuth, D. E. (1973): The art of computer programming. Vol.13: Sorting and searching. Reading, Mass.

Kochen, M. (1967): The growth of knowledge. N.Y. etc.

Kommission der europ. Gemeinschaften (1977): Die Überwindung der Sprachbarrieren. München 2 Bde.

Kraft, R.-P. : (1984) Endforderungen an Microcomputer für Online-Recherchen. Nachrichten für Dokumentation, 35/1984, 4-5

Kriz, J. (1975): Datenverarbeitung für Sozialwissenschaftler. Reinbek/H.

Krohn, R.G. (1971): The social shaping of science. Westport

Kschenka, W. (1970): Benutzeranalyse und Pädagogische Literaturinformation. München

Kuhlen, R. (Hg., 1980): Datenbasen, Datenbanken, Netzwerke. 3 Bde. Bd. 1: Aufbau von Datenbasen, Bd. 2: Konzepte von Datenbanken, Bd. 3: Nutzung und Bewertung von Retrievalsystemen.

Kuhlen, R. (1982): Informationsvermittlung und Informationsmanagement. Nachrichten für Dokumentation 33 (3)

Kuhlen, R. (Hg., 1984): Koordination von Informationen. Die Bedeutung von Informations- und Kommunikationstechnologien in privaten und öffentlichen Verwaltungen. IX. Verwaltungsseminar Konstanz, 5.-7. Mai 1983. Berlin etc.

Kunz, W./Rittel H. (1972): Die Informationswissenschaften. München

Lammarsch, J./Steinweg, H. (1994): Internet & Co. Fachkommunikation auf akademischen Netzen. Bonn etc.

Land F. J./Land, M. (1982): Über den Einsatz von Mikro-Computern zur Dokumentation und Auswertung wissenschaftlicher Literatur im Forschungsprozeß. Zeitschrift für Soziologie 11 (2), 183-6

Landau, R. et al. (1980): Directory of online databases. Vol 2 (1). Santa Monica

Lange, B.-P. et al. (1982): Sozialpolitische Chancen der Informationstechnik. Frankfurt

Lehmann, K.-D. (1984): Informationsbanken für Lehre und Forschung. Zeitschriften für Buch- und Bibliothekswesen 31 (3), 215-231

Leskein, H./Müller, W. (1982) : Informationsvermittlung im Wettbewerb. Bibliothek 6 (1/2)

Light, R. J. (1979): Capitalizing on variation: How conflicting research findings can be helpful for policy. Educational researcher 8 (9)

Light, R. J. (1980) : Synthesis methods: Some judgement calls that must be made. Evaluation in education: an international review series 4

Light R.J./Smith, P.V. (1971): Accumulating evidence: procedures among different studies. Harvard Educational Review 41, 429-71

Line, M. B. (1981): The structure of social science literature as shown by a large-sale citation analysis. Social Sciences Information Studies 1 (2)

Loevinger, J. (1970): NISP: Noisy signal in psychology. Transaction 7 (7), 1

Maloney, J. (Ed., 1983): Online searching technique and management. Chicago

Marcus, R. S. (1980): Search Aids in a Retrieval Network. Communicating information; Proceeding of the 43 ASiS Annual Meeting, Washington

McCarn, D.B. (1978): MEDLINE users, usage and economics. Medical Informatics 3 (3) 177-183

McNaught, J. (1981): Terminological data banks: a model for a British linguistic data bank (ldb). Aslib Proceedings 33 (7/8), S. 297

Meadow, C. T. (1979): Online Searching and Computer Programming: Some Behavioural Similarities. Online 3 (1), 49-52

Meadow, C. T. (1979): Computer as a Search Intermediary. Online 3(3), 57-59

Meffert, H. (1983): Bildschirmtext als Kommunikationsinstrument. Stuttgart

Mechtiev, D. M. et al. (1974): Automated preparation of a citation index and development of an information retrieval system based on it. Sci. and techn.inform. process. 4, 10-17

Menk, R. (1975): Ein automatisches Dokumentationssysteme über anthropologische Literatur. HOMO XXVI, S. 261-263

Merton, R. K. (1973, 1968): The Matthew effect in science. ders.: The sociology of science. Theoretical and empirical investigations. Chicago/London, 439-459

Merton, R. K. (1970): Science, technology and society in seventeenth century england. New York

Merton, R. K. (1973): The sociology of science. Theretical and empirical investigations. Chicago/London.

Merton, R. K. (1980): Auf den Schultern von Riesen. Ein Leitfaden durch das Labyrinth der Gelehrsamkeit. Frankfurt.

Michailow, A. I. et al. (1970): Osnowy informatiki. Grundlagen der wissenschaftlichen Dokumentation und Information. Köln

Mocker, U. et al. (1993): PC-Einsatz in den Geisteswissenschaften. München

Moravsik, M. J./Murugesan, P. (1975): Some results on the function and quality of citations. Social Studies of Science 5, 86-92

Müller, H./Strauch, D: (1984): Von der Dokumentation zum Wissenstransfer. Ffm.

Mackay, R. (1985): Online-Informationendienste. Duisburg

Morgan, E. L. (1994): WAIS and Gopher servers: a guide for Internet endusers. Westport, CT/London

Moureau, M. (1976): Les aspects linguistiques ds stratégies d'interraogatin dans la recherche bibliographique sur ordinateur (1). Documentaliste 13 (Sept.-Dec), 190-195

Müller-Brentel, M./Schöpflin, U. (1984): Ergänzungen zum Artikel von Busowietz u. Schmidt-Reindl. Nachrichten für Dokumentation 6-1984

Nacke, O. (1976): Scientiometrie und Bibliometrie in Planung und Forschung. Bielefeld

Neubauer, K. W. (1984: Online-Informationsmarkt und Bibliotheken.

Nora, S./Minc, A. (1979): Die Informatisierung der Gesellschaft. Frankfurt/N.Y.

Norris, C. (1981): MeSH- The subject heading approach. Aslib Proceedings 33 (4), 153-160

Ockenfeld, M./Rehfeld, W.: Die eigene Position in einem weiten Umfeld: Erfordernisse künftiger IuD-Politik. Nachrichten für Dokumentation 34 (6), 267 ff.

Oeser, E. (1976): Wissenschaft und Information. Bd. 1: Wissenschaftstheorie und empirische Wissenschaftsforschung. Wien/München

Omer, Y. et al (1981): DOMESTIC - a minicomputer based information storage and retrieval system. Journal of Information Science 3 (2), 59-74

Online 6 (6), 1982: p.14-23 training end-user searching?

Opp, K. D. (1970): Methodologie der Sozialwissenschaften. Einführung in Probleme ihrer Theorienbildung. Reinbek/H.

Papier, L. (1971): Approaches to establishing dimensions and criteria for evaluation or review publications. Proceedings International Conference on Information Science Tel Aviv, 213-231

Peball, K. (1977): Ein neues Informationszentrum am Balkan. Gründungskonferenz des Centre International d'Information sur de l'Historie Balkanique et Mediterranee. Österreichische Osthefte 19, 207

Pelissier, D. (1980): PASCAL database file description and online access on ESA/IRS. Online Review 4 (1), 3-31

Peterßen, W.H. (1987): Wissenschaftliches Arbeiten: nicht leicht, aber erlernbar. München

Pfaffenberger, A./Echt, S. (1980): Substitution of Scisearch and Social Scisearch for their print versions in an academic library. Database 3 (1), 63-71

Philosophie-Informationsdienst (1979): Dokumentation der unselbständigen philosophischen Literatur. Düsseldorf, 2. überarb. Aufl.

Pilachowski, D. M. (1979): USPSD: United States Political Science Documents. Database 2 (4), 68-77

Platt, J. (1981): Evidence and proof in documentary research, I, II. Sociological Review 28, 31-66

Pörksen, U. (1973): Zur Terminologie der Psychoanalyse. Deutsche Sprache 3, 7-36

Pörksen, U. (1974): Vom pseudowissenschaftlichen Jargon. Neue Rundschau 85, 214-222

Pörksen, U. (1978): Zur Metaphorik naturwissenschaftlicher Sprache. Neue Rundschau 89, 63-82

Popper, K. (1976): Logik der Forschung. Tübingen. Sechste, verbesserte Auflage (c1934ff.)

Portrait: Der Informations-Broker (1981). Marketing-Journal 1/1981

Price (Solla), D. de (1974): Little Science, Big Science. Von der Studierstube zur Großforschung. Frankfurt

Price (Solla), D. de (1976): A general theory of bibliometric and of cumulative advantage processes. Journal of the American Society for Information Science 27, 292-306

Price (Solla), D. de (1976): Science since Babylon. London

Project Head Start (1970?): Bibliography on early childhood. Washington

Raben, J./Marks, G. (Eds.) (1980): Data bases in the humanities and social sciences. Amsterdam etc.

Rauch, W. (1984): Die Rolle der Bibliothek im computergestützten Informations- und Kommunikationssystem. Bibliothek 8 (2), 109-112

Rauch, W./Wersig, G. (Hg., 1978): Delphi-Prognose in Information und Dokumentation. Untersuchungen über zukünftige Entwicklungen des Bibliotheks-, Informations- und Dokumentationswesens in der Bundesrepublik Deutschland und in Österreich. München

Rase, W. (1972): FAMULUS. Angewandte Informatik 1972/10, 459-465

Raith, W. (1988): Gut schreiben. Ein Leitfaden. Ffm./N.Y.

Rehfeld, W. (1979): Die bewältigte Informationsflut - Fachinformationssysteme bieten neue Möglichkeiten. Das Inforum, Nr. 1/April 1979

Reif, F. (1961): The competitive world of the pure scientist. Science 134, 1957-62

Resch, I. (1980): Kopierpraxis bei Fachzeitschriften. Lösungsvorschläge auf der Grundlage eines allgemeinen Kopierverbots für Großkopierer. Film und Recht (FUR) 1980/2, 91 ff.

Ridley, D. D. (1996): Online Searching: A Scientist's Perspective. A Guide for the Chemical and Life Sciences. Chichester

Rischka, G. (o. J.): Uni, ich komme! Ein Ratgeber für Erstsemester und alle, die es werden wollen. Wien etc.

Robertson, E./Sparck, I. K. (1976): Relevance weighting of search terms. Journal of ASIS 27, 129-146

Roloff, H. (1976): Lehrbuch der Sachkatalogisierung. 4. Aufl. München

Rosenthal, R. (1978): Combining results of independent studies. Psychological bulletin 85, 185-193

Rosenthal, R. (1979): The 'file drawer problem' and tolerance for null results. Psychological bulletin 86, 633-641

Runge, W. (1983): Die Rolle von Reviews in den Naturwissenschaften. Nachrichten für Dokumentation 34 (2), 72-78

Salton, G. (1967): Designing automatic information systems: results obtained with the SMART programs. Social Science Information 62/3

Salton G. et. al. (1983): An evaluation of term dependence models in information retrieval. Orig. en. Lect. Notes Comput Sci, 151-173

Sardar, Z. (1981) Between GIN and TWmeeting the information needs of the Third World. Aslib Proceedings 33 (2), 53-61

Savage, G. S./Pemberton, J. (1978): SciSearch on DIALOG. Database 1 (1), 50-67

Schenkel, W. (1971): Ein Programmgenerator für die maschinelle Texterschließung. Beiträge zur Linguistik und Dantenverarbeitung. Heft 21, 58-64

- Schiro, H. (1966): Kwic Index. Ein maschinelles Dokumentationsverfahren. IBM Nachrichten 16, Nr. 177, 127-133
- Schlitt G. (1983): Elektronisches Publizieren von Volltexten. ABI-Technik 3 (2), 125-129
- Schmidt, J.(1984): Wissenschaftsladen. Nachrichten für Dokumentation 35(3), 148-151
- Schuhmacher, D. (1983): Rez. zu Beyer, W. Informationsvermittlung in der BRD. Nachrichten für Dokumentation 34, 182
- Schuller, A. (1977): Konzept für den Aufbau der wissenschaftlichen Fächer in der Medizin. MMG Mensch Medizin Gesellschaft 2, 108-113
- Schwab, H. (1984): Das luD-Programm aus heutiger Sicht. Nachrichten für Dokumentation, 32 (4/5), 165ff.
- Schwarz, C. (1982): Freitextrecherche - Grenzen und Möglichkeiten. Nachrichten für Dokumentation 33 (4/5), 228-236
- Schwarz, C. (1984): Linguistische Hilfsmittel beim Information Retrieval. Nachrichten für Dokumentation. 35 (4/5), 179-182
- Schwarz, R. (1975): Wissenschaft als interidisziplinäres Problem. Berlin, Hahn P. 1979
- Schweizerische Koordinationsstelle für Bildungsforschung (1978): On-Line-Literatur-Suchdienst zur Erschließung ausländischer Informationssysteme auf dem Gebiet der Bildung- und Sozialwissenschaften.
- Sedelow, W. A. (1978): Sedelow formalized historiography, the structure of scientific and literary texts. Journal of the history of the behavioral sciences 14
- Seelger, T. (1977): Ausbildungsgänge im Tätigkeitsbereich I u. D. München
- Seeger T./Strauch, D. (1979): Aufgaben und Möglichkeiten von Informationsunternehmen. Nachrichten für Dokumentation 30
- Seiffert H. (1970): Information über die Information. 2. Aufl. München
- Seelbach, D. (1975): Computerlinguistik und Dokumentation. München 1979 ??
- Seltzer R. et al. (1997): The ALTA VISTA Search Revolution. Osborne
- Sietmann, R. (1984): Oft kopiert, nie erreicht? Psychologie Heute März 1984, 52-56
- Sievert M. E. (1980): The Philosopher's Index. Database 3 (1), 50-61
- Sievert, M. C./Sievert, D. E. (1991): Online searching in philosophy. Online review 15, 63-79
- Simmons A./Stehr N. (1981): Language and the Growth of Knowledge in Sociology. Social Science Information 20 (4-5), 703-41
- Smith, ML. (1980): Publication bias and metaanalysis. Evaluation in education: An

international review series 4, 22-24

Smith M. L./Glass G. V. (1977): Meta-analysis of psachotherapy outcome studies. American Psychologist 1977, 752-60

Soergel D. (1971): Dokumentation und Organisation des Wissens. Berlin 1975

Sparck, J.K. (1979): Search term relevance weighting. Journal of Documentation 35, 30-48

Spencer C. C. (1967): Subject searching with Science Citation Index: American Documentation 18/2, 87ff.

Spiegel-Rösing, I. S. (1973): Wissenschaftsentwicklung und Wissenschaftssteuerung. Frankfurt

Spiegel-Rösing, I. S. (1977): Science studies: Bibliometric and content analyses. Social Studies of Science 7, 97-113

Spinner, H. F.(1984): Zur Soziologie des Rezensionswesens. Soziologie 1/1984, 49-78

Stahn, P. (1977): Datenbanken. Nachrichten für Dokumentation 28

Stegemann H. (1980): Informationsmarketing und Benutzerforschung. Berlin

Stegmüller, W. (1975): Die Evolution des Wissens: Nichtkumulativer Wissenschaftsfortschritt und Theoriendynamik. Stegmüller, Hauptströmungen der Gegenwartsphilosophie Bd. II. Stuttgart, 483-534, 550-552

Steinbuch, K. (1978): Die Enteignung unseres Denkens. München-Berlin

Stern J. (1971): An analysis of the review literature. Proceedings of the American Society Information 8, 91-96

Stock, K. F. (1982?): Einführung in die elektronische Datenverarbeitung für Archivare. München

Stock, W. G. (1981): Die Wichtigkeit der wissenschaftlichen Dokumente. Nachrichten für Dokumentation 32 (4/5), 164f.

Strasser, H./Knorr, H. D. (Hg. , 1976): Wissenschaftssteuerung. Frankfurt

Streicher, H. (1984): Die Angst des Vorstands vor dem Computer. Office Management 32 (11), 1074-76, Baden-Baden

Study committee "research on the theoretical basis of information". FID 527, Moskau 1976

Stukenbröcker, B. (1979): IuD Informationsvermittlung aus Online-Systemen im Unternehmen. Nachrichten für Dokumentation 30 (1), 13-16

Textor, R. B. (1967): Computer summarization of the coded cross-cultural literature. Social Science Information VI, 137-146

Thesaurus of Psychological Index. Washington 1977

Thomas, B. J. (1996): The Internet for Scientists and Engineers. Online Tools and Resources. Oxford etc. (2nd Ed.)

Thomas, K. (1978): Ein Soziologe blickt auf die medizinische Soziologie.

Thun, v. Schul F. et al. (1974): Überprüfung einer Theorie der Verständlichkeit von Kurzzusammenfassungen wissenschaftlicher Veröffentlichungen. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 192-206

Toulmin, S. (1966): Is there a limit to scientific growth? Science Journal 2

Trümpener, H. J./Weingart P. (1978): Empirische Forschungsprozeßanalyse. Bielefeld

Tze-Ching Li (1980): Social Science Reference Sources: A Practical Guide, Westport

Voß, H. (1984): Nutzungsmöglichkeiten von Bildschirmtext an wissenschaftlichen Hochschulen.

Ullitzka, N. (1979): Interaktion von Politiksystem unter Legitimationsaspekten. Bielefeld

Urban, D. (1978): Wissenschaftlermobilität und Wissenschaftstransfer - eine empirische Studie zum Verhältnis von Stellenmarkt. Bielefeld

Velten, D. (1981): ICONOS: une nouvelle base de donnees sur la photographie. In Documentaliste, 18, 106ff.

Vernimb, C. O. (1977): The european network for scientific, technical, economic and social information. Nachrichten für Dokumentation 18

Verzeichnis deutscher Informations-Dokumentationsstellen. GID-IZ, Aufgabe 3-1978/79 München

Vickery, B. C. (1971): Dokumentationssysteme. München-Pullach-Berlin

Vinogradov V. A. (1977): Problems of international cooperation in social science information and documentation. International forum on information and documentation. Den Haag

Wagner, W. (1977): Uni-Angst und Uni-Bluff. Wie studieren und sich nicht verlieren. Berlin

Walberg, H. J./Haertel, E. H. (1980): Research integration: Introduction and overview. Evaluation in education: An international review series 4, 5-10

Walford A.J. (Ed., 1975): Guide to reference material. 3. Aufl. London

Warnken, K. (1981): The information brokers. N. Y.

Weinberg, A. M. (Hg., 1970): Probleme der Großforschung. Ffm.

Weingart, P. (Hg., 1972): Wissenschaftssoziologie. Bd. I: Wissenschaftliche Entwicklung als sozialer Prozeß. Ffm.

Weingart, P. (Hg., 1974): Wissenschaftssoziologie. Bd. II: Determinanten

wissenschaftlicherEntwicklung. Ffm.

Wersig, G. (1971): Information-Dokumentation-Kommunikation. Pullach-Berlin

Wersig, G. (1973): Informationssoziologie. Ffm.

Wersig, G. (Hg., 1983): Informatisierung und Gesellschaft. München

Wesemann, A. (1997): Das ganze Internet ist ein einziger Kunstfehler. kursiv 4 (1), 30-32

Wetzle, E.L. (1980): Kostenmodell für eine Informationsvermittlungsstelle. Nachrichten für Dokumentation 31, 77-82

Wiesner, H. (1978): Bildschirmarbeitsplätze - Problemfelder und Aufgaben der Arbeitsgestaltung. afa-informationen 28 (6), 14-20

Williams, M. E. (1978): Online retrieval - today and tomorrow. Inline Review 2 (4), 353-366

Wilson, T. D. (1980): On Information Science and the Social Sciences. Social Science Information Studies 1, 3-13

Wingert, F. (Hg., 1978): Klartextverarbeitung. Berlin 1979

Whitley R. (1974): Social processes of scientific development. London-Boston

Witte, E. (Hg., 1980): Telekommunikation für den Menschen. Berlin

Woodward A. M. (1977): Roles of reviews in information transfer. Journal of the American Society for Information Science 28 (3), 175-180

Woodward, A.M. (1976): Citations to review serials. Journal of Documentation 32 (4), 290 ff.

7. Anhang

[Übersicht 1](#): Nachschlagewerke (Beispiele)

[Übersicht 2](#): Datenbasen-/CD-ROM-Kataloge

[Übersicht 3](#): Inhalte von Datenbasen

[Übersicht 4](#): Sozialwissenschaftlich relevante Hosts

[Übersicht 5](#): Datenbasen für Soziologen

Übersicht 6: Hosts mit der Kommandosprache CCL

Übersicht 7: Elementare Suchfunktionen am Beispiel CCL

Übersicht 8: Zugriffsmöglichkeiten über Datenfelder

Übersicht 1: Nachschlagewerke (Beispiele)

a) Periodisch erscheinende Nachschlagewerke

Informationszentrum Sozialwissenschaften: *Forschungsarbeiten in den Sozialwissenschaften 1969ff.* Bonn

Jährliche erscheinende Dokumentation sozialwissenschaftlicher Forschungsarbeiten, sowohl theoretischer als auch empirischer Art, auf der Basis der Erhebung des Informationszentrums Sozialwissenschaften, Bonn. Die Forschungsarbeiten sind nach Ländern, innerhalb eines Landes nach Orten (Sitze der Institutionen) und nach dem Namen der Institutionen alphabetisch geordnet. Personen, Sach- und geographisches Register.

Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit/Bundesministerium für Arbeit und Sozialordnung: *Forschungsdokumentation zur Arbeitsmarkt- und Berufsforschung.* Nürnberg/Bonn

Erscheint dreimal jährlich, dazu einmal jährlich mit einem Sonderband über abgeschlossene Forschungsarbeiten. Auf der Basis der Erhebung des Informationszentrums Sozialwissenschaften, Bonn sowie der Tätigkeit der Informationsstelle für Forschungen auf dem Gebiet der Beschäftigung beim Bundesministerium für Arbeit und Sozialordnung und zusätzlichen Recherchen des Instituts für Arbeitsmarkt- und Berufsforschung (IAB). Textteil: Die Forschungsprojekte sind nach der systematischen Klassifikation des IAB geordnet. Der Registerteil enthält folgende Register: (Namens-)Register der forschenden Institutionen, Ortsregister der forschenden Institutionen, Forscher-Register, Auftraggeber-Register, Register der finanzierenden Institutionen, Schlagwort-Register, Geographisches Register, Systematisches Register. Im Anhang: Systematische Klassifikation des IAB.

Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit: *Literaturdokumentation zur Arbeitsmarkt- und Berufsforschung.* Nürnberg

Neben der allgemeinen Literaturdokumentation gibt das IAB auch zahlreiche themenzentrierte Sonderhefte (z. B. "Berufliche Sozialisation", "Schattenwirtschaft, Alternativökonomie") heraus, entweder in alleiniger Verantwortung oder gemeinsam mit Kooperationspartnern. Diese Sonderhefte enthalten in der Regel eine inhaltliche Einführung sowie Autoren- und Forscher-Register, Titel-, Schlagwort-, z. T. Experten-Register.

Zentralarchiv für empirische Sozialforschung an der Universität zu Köln in Zusammenarbeit mit Informationszentrum Sozialwissenschaften: *Empirische Sozialforschung 1982 ff.* Frankfurt/Main, New York: Campus Verlag, 1983ff.

Jährlich erscheinende Dokumentation von empirischen Untersuchungen auf der Basis der Erhebung des Informationszentrums Sozialwissenschaften, Bonn. Aufnahmekriterium: Die Daten der dokumentierten

Forschungsarbeiten müssen einer quantitativen Sekundäranalyse zugänglich sein (Teilmenge der o.g. Dokumentation "Forschungsarbeiten..", detailliertere Darstellung).

b) Nichtperiodisch erscheinende Nachschlagewerke

Umfragen aus der empirischen Sozialforschung 1945-1982. Datenbestandskatalog des Zentralarchivs für empirische Sozialforschung. Frankfurt/Main, New York 1984

Dokumentiert weit über 1000 öffentlich zugängliche Studien aus dem Bereich der empirischen Sozialforschung. Die Datensätze stammen aus privatwirtschaftlicher sowie universitärer Forschung. Schwerpunkt der Sammlung sind Repräsentativstudien für die Bundesrepublik, darunter auch Zeitreihen- und Panelstudien.

Bette, K.-H. et al. (Hg.): *Deutsche Soziologie - German Sociology*. Auswahlbibliographie 1945-1977. Bonn/Göttingen

Der Band ist in insgesamt 65 Einzelkapitel (26 Hauptkapitel) gegliedert und durch Namens-, Institutionen- und Sachregister erschlossen.

Übersicht 2: Datenbasen-/CD-ROM-Kataloge/-Zeitschriften (Beispiele)

1) Online-Basen-Kataloge

Directory of Online Databases. Cuadra/Elsevier, New York

Scientific Consulting Schulte-Hillen (1992) : Handbuch der Wirtschaftsdatenbanken 1992. Hoppenstedt V., Darmstadt

Scientific Consulting Schulte-Hillen (1992): Handbuch der Datenbanken für Naturwissenschaft, Technik, Patente 1992. Hoppenstedt V., Darmstadt

Marcaccio, K. Y./DeMaggio, J. A. (Eds., 1989): Computer-readable Databases. A Directory and Data Sourcebook. Gale Research Inc. Detroit, Michigan

Staud, J. L. (1987): Online Wirtschaftsdatenbanken 1987/Online Business Databases 1987. Bern etc., 2 Bände

2) CD-ROM-Kataloge

Desmarais, N. (1991): CD-ROMS in Print. Meckler

Mitchell, J. (Ed. 1990): The CD-ROM Directory 1991. London

3) CD-ROM-/Online-Zeitschriften, Newsletters, Informationsdienste

ECHO facts for user (p.r. für den Host ECHO: European Commission Host Organisation), Luxemburg. mail: echo@echo.lu, URL <http://www.echo.lu/> (v.a. I*M-Europe)

Cronolog. Knight-Ridder Information, European Edition, London. mail: chronolog@dm.krinfo.ch, URL <http://www.krinfo.com> (p.r. für die Hosts DIALOG und data star)

Information World Review. News on Internet, intranet, online and CD-ROM content. Oxford, UK (Learned Information Europe Ltd). mail: subscriptions@learned.co.at, URL: <http://www.learned.co.uk>

Online & CD-ROM Review. The International Journal of Online, Optical and Networked Information. Oxford, UK (Learned Information Europe Ltd). mail: subscriptions@learned.co.at, URL: <http://www.learned.co.uk>

database. Wilton

review of information science. <http://www.inf-wiss.uni-konstanz.de/RIS/>

Schriften zur Informationswissenschaft. Universitätsverlag Konstanz

Übersicht 3: Inhalte von Datenbasen

a) Referenzinformationen:

Datenbasennachweise: Enthalten Informationen über Datenbasen, z. B. I*MGUIDE von ECHO (European Commission Host Organization), Luxemburg

Nachweise von Forschungsprojekten:

Projekthinweise beruhen in der Regel auf freiwilligen Angaben der beteiligten Forscher zu abgeschlossenen, laufenden und geplanten Forschungsprojekten (z. B. FORIS, Informationszentrum Sozialwissenschaften, Bonn).

Literaturnachweise:

wissenschaftliche Zeitschriftenaufsätze (in der Regel neben den üblichen bibliographischen Angaben mit Autorenadresse, Abstract und Schlagwörtern versehen) (z. B. SOLIS)

Beiträge zu Sammelbänden (z. B. SOLIS)

Bücher: z. B. Übersetzungen, *Index Translationum*,
<http://www.unesco.org/culture/xtrans/html_eng/index4.shtml>

Graue Literatur: z. B. Europ. Datenbase SIGLE, SOLIS

Bibliothekarische Standortnachweise:

Verbund Österreich <http://www.bibvb.ac.at/verbund-opac.htm> Aleph-Linz
<http://aleph.edvz.uni-linz.ac.at/ALEPH>

Tages-, Wochen-, Fachpresse: z. B. FRANCE ACTUALITE, NYTIS (= New York Times Information Service)

b) Primärinformationen:

Presseartikel (Volltexte): z. B. Die Presse, Wien, Der Spiegel, TAZ

Meldungen der Nachrichtenagenturen: z. B. UPI NEWS; DPA, APA

Terminologien: z. B. Mesh, medizinisches Vokabular

numerische/statistische Daten: z. B. STATIS-BUND (Statistisches Informationssystem des Bundes)

Fakten z. B. JURIS (Juristisches Informationssystem: Bundesgesetze und -verordnungen, Verwaltungsvorschriften, Rechtssprechung) in der BRD, RDB (Rechtsdatenbank) in Österreich

Electronic Publishing:

Volltexte von wissenschaftlicher Grauer Literatur oder von Fachzeitschriften, z. B. VDI-Nachrichten (Verein Deutscher Ingenieure), Elektronische Zeitschriften im WWW, z. B. RIS = Review of Information Science : A Peer-reviewed Electronic Journal Promoting Academic Research in Information Science. ISSN 1431-5971, <http://www.inf-wiss.uni-konstanz.de/RIS/>:

TRANS, Zeitschrift für Kulturwissenschaften <<http://www.inst.at/trans>> ,
FQS: FORUM QUALITATIVE SOZIALFORSCHUNG
<<http://www.qualitative-research.net/fqs/fqs.htm>>

e-journal: "kommunikation@gesellschaft" <<http://www.kommunikation-gesellschaft.de/>>

Künstliche Intelligenz:

Expertenwissen (Expertensysteme)

Übersicht 4: Wissenschaftlich relevante Hosts

(Auswahl)

DATA STAR

primär wirtschaftlich/technischer Host, jedoch auch psychologische und soziologische Datenbasen

DIALOG

Datenbasen zu sämtlichen wissenschaftlichen Disziplinen, auch sozial- und geisteswissenschaftliche. International, insbesondere angelsächsischer Sprachraum, WWW-Adresse: <<http://dialog.krinfo.com/>>, s.a. Web Science (WWW-Zugang zu Dialog-Datenbasen)

DIMDI (Deutsches Institut für Medizinische Dokumentation und Information)

Medizinische und biowissenschaftliche, soziologische und psychologische Datenbasen. Kommandosprache: CCL, WWW-Adresse: <<http://www.dimdi.de>> (SOLIS und FORIS werden leider mit 31.12.2002 auf diesem Host eingestellt, verbleiben aber auf STN und GBI)

ECHO (European Comission Host Organisation)

Kommandosprache: WWW-Adresse: <<http://www.echo.lu>>

EDV

Sozialwissenschaftliche und Rechtsdatenbanken, Mediendokumentation

FIZ TECHNIK

technisch-naturwissenschaftliche Datenbasen, INFODAT (Informations-wissenschaften)

GBI

Gesellschaft für Betriebswirtschaftliche Information: Wirtschaftsdatenbasen
<<http://www.gbi.de/>>

INIST

Datenbasen zu sämtlichen wissenschaftlichen Disziplinen, auch allen sozial- und geisteswissenschaftlichen. International, insbesondere romanischer Sprachraum, Schwerpunkt: Frankreich

STN (Scientific Technical Network)

Computerverbund BRD-USA-Japan; primär naturwissenschaftlich-technische Datenbasen, jedoch auch SOLIS, FORIS

Übersicht 5: Datenbasen für SoziologInnen (Auswahl)

a) Referenzdatenbasen (Nachweis von Projekten, Literatur)

SOWIDOK (eher arbeitnehmerorientierte sozialwissenschaftliche und Verbandsliteratur, Arbeiterkammer Wien); Host: EDV

FORIS (sozialwissenschaftliche Forschungsprojekte, Informationszentrum Sozialwissenschaften Bonn); Hosts: DIMDI, STN

SOLIS (Soziologische Literatur, IZ Bonn); Hosts: DIMDI, STN

FoDokAb (Forschungsprojekte Arbeitsmarkt- und Berufsforschung); nicht online zugänglich

LitDokAb (Literatur zu Arbeitsmarkt- und Berufsforschung); nicht online zugänglich

PHILIS (Philosophie Informationsdienst), Düsseldorf: Wissenschaftstheorie, Sozialphilosophie, Soziologische Theorie (bes. Frankfurter Schule), DDR-Literatur; <http://philmain.phil-fak.uni-duesseldorf.de/>

Sociological Abstracts (USA); Hosts u.a.: DIMDI, data star, DIALOG

Social Sci Search (USA, Zitationsdatenbase: Suche nach zitierten Autoren, Werken möglich); Hosts u.a.: DIMDI, data star, DIALOG

Francis Sociologie (Datenbasenversion des Nachschlagewerkes "Bulletin Signalétique - Sociologie", Frankreich; international, ihr Vorteil liegt insbesondere in den zahlreich erfaßten Publikationen in *romanischen* Sprachen, die von den angelsächsischen Datenbasen wenig abgedeckt werden; enthält allerdings nur sehr kurze Abstracts); Host: INIST

b) Datenarchive, Faktendatenbasen

Zentralarchiv für empirische Sozialforschung Köln: Datensätze zur Sekundäranalyse

ISIS, Statistisches Zentralamt, Wien

STATIS BUND: Statistiken BRD

c) Allgemeine, auch für Sozialwissenschaftler relevante Daten

LoC: Buchbestände d. Library of Congress, Washington (sehr umfassend, international; Host: u.a. BRZN, oder über das WWW

d) Datenbanken zu verwandten Gebieten

Psyndex (deutschsprachige psycholog. Literatur, Datenbasenversion des periodisch erscheinenden Nachschlagewerkes "Psychologischer Index"); Host: DIMDI

PsycInfo (internationale psycholog. Literatur, USA; s.a. "Psychological Abstracts. Produzent: American Psychological Association); Hosts u.a.: DIMDI, data star, DIALOG

francis sämtliche humanwiss. Datenbasen, z. B. ethnologie; Host: INIST

Übersicht 6: Hosts mit der Kommandosprache CCL/DIRS/ GRIPS**(Beispiele)**

DIMDI (Deutsches Institut für medizinische Dokumentation und Information, Köln)

Weißhausstr. 27, 5000 Köln 41, WWW-Adresse: <http://www.dimdi.de>

Offeriert neben zahlreichen medizinischen und biowissenschaftlichen Datenbasen, die z.T. durchaus auch von sozialwissenschaftlicher Relevanz sein können (etwa in den Bereichen Sozialmedizin, Sportwissenschaften) eine Reihe von soziologischen und psychologischen Datenbasen:

FORIS (*Forschungsinformationssystem Sozialwissenschaften*) nur bis 31. 12. 2002

PsycInfo (psychologische Literatur aus Zeitschriften und Sammelbänden, primär englischsprachige)

Psyndex (Datenbankversion des Nachschlagewerkes "Psychologischer Index"; sehr interessant, aber relativ kostenintensiv)

Psyndextrain (Trainingsdatenbank, Auszug aus *Psyndex*, lizenzgebührenfreies Üben möglich)

Social Sci search (Zitationsdatenbase, international)

Sociological Abstracts (soziologische Literatur, primär englischsprachige,

Datenbasenversion des gleichnamigen Nachschlagewerkes)

SOLIS (Soziologisches Literaturinformationssystem) nur bis 31. 12. 2002

ECHO (Host der Europäischen Gemeinschaft, Luxemburg)

PO Box 2373, L-1023 Luxembourg, <http://www.echo.lu>

kostenlos, u.a.

Trainingsdatenbank für CCL (Common Command Language),

Datenbasen der Datenbasen und Informationsbroker (*I'MGUIDE*),

THESAURI (Wörterbücher kontrollierter Terminologie aus EG-Mitgliedsstaaten, USA und Kanada)

Terminologiedatenbasen (z. B. *EURO-DICAUTOM*),

zahlreiche Datenbasen zu Forschungs- und Entwicklungsprojekten der Europäischen Gemeinschaft, zum Europarecht,

sowie eine Datenbase der Vereinten Nationen (*DUNDIS*).

Der ECHO-help desk kann kostenfrei telephonisch erreicht werden;

e-mail: echo@echo.lu

Auch viele über telnet recherchierbare Rechner im WWW verwenden die Kommandosprache CCL oder ähnliche Dialekte.

Übersicht 7: Elementare Suchfunktionen am Beispiel CCL -

wichtigste Befehle von A bis Z

Die Befehle müssen nicht unbedingt ausgeschrieben werden (Ausnahme: "Stop" und "info"), es genügt die Eingabe der unterstrichenen Zeichen.

Suchfunktionen

base (gefolgt von Datenbasenbezeichnung): Auswahl einer Datenbase,

base (ohne Datenbasenbezeichnung): Auflistung verfügbarer Datenbasen

display Wörterbuchfunktion: Ausgabe alphabetisch nachfolgender suchbarer Wörter

find Suchbefehl

print Druckbefehl (am Wirtsrechner, postalische Zusendung)

show Online-Anzeige der Treffer

stop Beendigung des Recherchedialogs mit einem Host

tab Anzeige des Rechercheprofils der verwendeten (erfolgreichen) Suchstrategie mit Trefferzahlen, Verknüpfungen etc. (Suchprotokoll)

Fortsetzung einer unterbrochenen Befehlsabarbeitung

continue Fortsetzung eines unterbrochenen Kommandos

more Fortsetzung einer unterbrochenen Ausgabe

Unterstützende Funktionen

help (kurz: **?**) Dialogunterstützung (Erläuterung der Kommandos)

info allgemeine Information

Übersicht 8: Zugriffsmöglichkeiten über Datenfelder

A) *Freitextrecherche (FT)* in sämtlichen Datenfeldern

"FT" muß nicht eingegeben werden, erscheint jedoch als "Datenfeld" in den Systemantworten bzw. Profiltabellen

B) über einzelne *Datenfelder* (in CCL: **FIELD**),

z. B.

- über *Deskriptoren* (kontrollierte Terminologie) Deskriptoren sind in Thesauri (Wörterbücher kontrollierter Terminologie) verzeichnet, vgl. z. B. den Thesaurus der American Psychological Association (APA) oder des Educational Resources Information Center (ERIC)

In CCL: **CT** (controlled terms)

- über Autorennamen. In CCL meist: **AU**, seltener: **PER**

- über alle *signifikanten* (nicht auf einer Stop-Wort-Liste stehenden) *Wörter* aus den *Publikationstiteln*. In CCL: **TI**

- über alle *signifikanten* (nicht auf einer Stop-Wort-Liste stehenden) *Wörter* aus den *Abstracts*. In CCL: **AB**

- über *zitierte Publikationen/Autoren* (z. B. Social Sci Search). In CCL: **RF** = reference, **RJ** = referenced journal, **RP** = referenced patent No.

Insbesondere zur *Eingrenzung der Recherche*, d. h. zur Reduzierung der Trefferzahl, werden folgende Datenfelder verwendet:

- *Publikationsjahr (PY)*

- *Sprache der Publikation (LA)*

z. B. **AND LA=DE** oder **NOT LA=FR**

- *Dokumenttyp (DT)*: z. B. Buch, Zeitschriftenaufsatz, Arbeitspapier

8. Aktualisierungen

"zwecks automatischer uebermittlung von aktualisierungen ersuchen wir alle, die sich die datei herunterladen, im eigenen interesse um uebermittlung ihrer email adresse.

selbstverstaendlich sind wir auch an kritik, anregungen usw. interessiert"

Gerhard.Froehlich@jku.at

8.1 ECHO

From: "DGD" <dgd@darmstadt.gmd.de>

To: <dgd-news@gmd.de>

Subject: ECHO Datenbasen

Date: Wed, 28 Oct 1998 14:48:03 +0100

Sehr geehrte Damen und Herren,

ECHO (European Commission Host) wurde am 1. Oktober 1998 eingestellt.

Ueber folgende URLs haben Sie von diesem Zeitpunkt ab Zugang zu folgenden vier Datenbanken:

TED : the new URL is <http://ted.eur-op.eu.int/>

I*M-GUIDE : the new URL is <http://www.echo.lu/imguide/>

I*M-FORUM : the new URL is <http://www.echo.lu/imforum/>

EURODICAUTOM : the new URL is <http://www.echo.lu/edic>

Mit freundlichen Gruessen

Hans Nerlich

STARTSEITE
