
KLAUS FUCHS-KITTOWSKI & TANKRED SCHEWE

Informationsverarbeitung, - recherche und - erzeugung in den Biowissenschaften

1. Einführung - Wissenschaft und Innovation - Biochemie/ Molekularbiologie und Bioinformatik

Das Thema unserer Tagung: „Wissenschaft und Innovation“ wird wohl gerade in diesen Tagen kaum deutlicher belegt als durch die innovativen Forschungsergebnisse der Biochemie/Molekularbiologie, in Verbindung mit der Informatik und der sich damit verstärkt herausbildenden Bioinformatik.

Die Entschlüsselung des Humangenoms zu Beginn dieses Jahres konnte nur mit Hilfe besonders leistungsfähiger Computer in Angriff genommen werden und die riesige Menge dabei gewonnener Informationen kann nur mit Hilfe komplexer Datenbanken verwaltet und interpretiert werden. Solche Problemstellungen verlangt nach Spezialisten, die über hervorragende Kenntnisse auf dem Gebiet der Biochemie/Molekularbiologie und der Informatik verfügen. Dies führte in jüngster Zeit zu einer verstärkten Entwicklung der Bioinformatik. Mit den biologischen Daten aus der Humangenomschlüsselung alleine ist es bekanntlich noch nicht viel getan. Die Sequenzierung, die Übersetzung der Erbinformation in eine Buchstabenfolge ist nur der erste Schritt. Biowissenschaftliche Forschungseinrichtungen, wie auch Biotech-Unternehmen stellen sich jetzt der außerordentlich schwierigen, möglicherweise langwierigen und sehr kostspieligen Forschungsaufgabe, aus den Daten der Sequenzierung auf Ursachen von Krankheiten zu schließen, um auf dieser Grundlage wiederum Therapien gegen schwere Erkrankungen entwickeln zu können. Auch hierbei kommt der Bioinformatik eine Schlüsselstellung zu. Die Bioinformatik verkörpert praktisch zwei bzw. mehrere innovative Technologien, die auf Ergebnissen der Grundlagen- und Angewandten Forschung beruhen und weitere Forschungen stimulieren. Diese Forschungen werden einerseits staatlich gefördert andererseits sind sie aufgrund der sich abzeichnenden Bedeutung der Forschungsergebnisse für die künftige Wirtschaft, für auf Gewinn orientierte Unternehmen von großem Interesse und werden daher auch von ihnen finanziert. So können wir gerade in dieser Woche im Wirtschaftsteil der Berliner Zeitung¹ lesen, dass der Berliner Pharmakonzern

Schering die Hälfte seiner Biotech-Tochter Metagen an die Beteiligungsfirma APEX verkauft. Der Preis von 42 Millionen Euro kommt dem Unternehmen als Startkapital zu. Metagen wird sich Schering zufolge auf die Entwicklung von Medikamenten für die Krebstherapie konzentrieren. Nach der erfolgreichen Entschlüsselung des menschlichen Genoms stehe nun die Umsetzung des Wissens in vermarktungsfähige Medikamente im Vordergrund. Metagen sei auf Grund seiner weltweit führenden Technologie beim Vergleich von Erbinformationen von Krebs- und Gesunden Gewebe in der Lage, neue Therapieansätze zu entwickeln.

Hiermit wird die von Heinrich Parthey im Einleitungsreferat² zu unserer Tagung vertretene These verdeutlicht, dass naturwissenschaftliche Anwendungsforschung, aber auch Grundlagenforschung, die zu „erfolgreichen, technologischen Innovationen zur Sicherung von Produktion und kulturellen Fortschritt“³, wie Günter Spur es in seinem Buch „Technologie und Management“ formuliert, führt, auch von der Wirtschaft zu refinanzieren ist. Dies soll und darf den Staat jedoch nicht aus seiner Verantwortung für eine entsprechende, großzügige Forschungsförderung entbinden. Das Spannungsfeld von Wissenschaft und Technik, Wirtschaft und Gesellschaft zu verstehen, die sich vollziehenden Wandlungsprozesse zu erkennen und zum Wohle der Menschen zu beeinflussen, ist damit deutlich auch eine unternehmerische Aufgabe, dies ist und bleibt aber eine wichtige Aufgabe der Politik, die für die Entwicklung von Wissenschaft und Technik zur Sicherung der Produktion sowie des sozialen und kulturellen Fortschritts eine besondere Verantwortung trägt.

Im Folgenden soll nur ein spezieller Problembereich aus dem großen Problemfeld Biochemie/Molekularbiologie und Bioinformatik behandelt werden. Es geht speziell um neue Möglichkeiten der Biowissenschaften zur Informationsverarbeitung, -recherche und -erzeugung im Zusammenhang mit einer der entscheidenden technischen Innovationen unserer Tage - der Entwicklung des Internets. Dabei erlangt u. E. ein schon früh im Zuge der Entwicklung des Netzes geäußelter Gedanke - die Notwendigkeit von Informationszentralen - erneut Bedeutung. Denn es kann gezeigt werden, dass die Informationsrecherche eines semantischen feedbacks bedarf, wenn die Entstehung der Semantik im sozialen Prozess berücksichtigt werden muss.

- 1 Schering verkauft Metagen-Anteile, Berliner Zeitung, Nr. 68, Mittwoch, 21. März 2001.
- 2 Parthey, H., Formen von Institutionen der Wissenschaft und ihre Finanzierbarkeit durch Innovation. In diesem Jahrbuch.
- 3 Spur, G., Technologie und Management - Zum Selbstverständnis der Technikwissenschaft. München-Wien: Carl Hanser Verlag 1998.

2. *Entwicklung der modernen Informations- und Kommunikationstechnologie*

Paradoxie der Informationsverarbeitung - Beherrschung großer Datenmengen

„Euch haben sie das Netz ums Haupt geworfen“⁴. Diese Feststellung von Friedrich Schiller im „Wilhelm Tell“ charakterisiert die gegenwärtige Situation in der Informatik, in Wissenschaft und Wirtschaft.⁵ Auch wenn Schiller diesen Satz in einem anderen Zusammenhang sagen lässt.

Kaum eine Innovation der letzten Jahre verdient eine so große Aufmerksamkeit wie die Entwicklung des Internets, geht es doch nicht nur um die Entwicklung der modernen Informations- und Kommunikationstechnologien, sondern damit um die Entwicklung der Infrastruktur einer künftigen Informationsgesellschaft für alle.

Ohne die lokalen, regionalen sowie globalen digitalen Informations- und Kommunikationsnetze geht heute in der Wissenschaft und in der Wirtschaft nichts mehr. Die Bedeutung der digitalen Medien wird in Zukunft noch wesentlich wachsen. Sie wird in dem Maße zunehmen, wie Daten, Information und Wissen zu einem unverzichtbaren Allgemeingut und zu einem immer wichtigeren Wirtschaftsgut werden.

Denn mit dem Internet kann heute global eine Fülle an Daten und Informationen zur Verfügung gestellt werden.

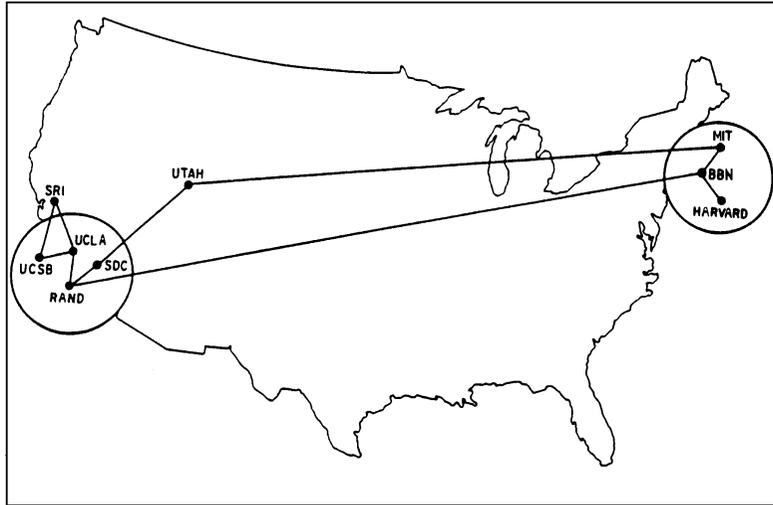
Mit dem Internet hat sich eine Revolution in der Informationsvermittlung vollzogen, die mit der Erfindung des Buchdrucks durch Gutenberg vergleichbar ist. Im ausgehenden 20. Jahrhundert und zu Beginn dieses neuen Jahrtausends sind wir in der Tat Zeugen einer weiteren technologischen Revolution, ohne dass wir schon ihre ganze Dynamik, die sie zur Veränderung des Wirtschafts- und Privatlebens entfaltet, voll überschauen könnten.

Ende der 60er Jahre wurde vom Militär der USA das sog. ARPAnet geschaffen, das die militärischen Computer untereinander vernetzte, vgl. Abbildung 1. Zu Beginn der 80er Jahre wurde dieses Netz immer mehr auch von Forschungseinrichtungen genutzt und der Betrieb des Netzes wurde der National Science Foundation der USA übertragen. Dies ermöglichte auch den schrittweisen Anschluss weiterer Länder an das Netz, vgl. Abbildung 2.

4 Schiller, F., Wilhelm Tell. In: Schillers Werke. Weimar: Volkerverlag 1958. S. 452.

5 Geihs, K., Chancen und Risiken von Kommunikationsnetzen. In: Neue Wege des Informationsmanagements in Banken - Chancen und Risiken von Kommunikationsnetzen - Internes Knowledge-Management, Neuntes Symposium des Informationsrings Kreditwirtschaft e.V., ik report, Band 9. Frankfurt am Main, Zürich 1996. S. 25

Abbildung 1: ARPANET, aus: Hauben, M. / Hauben, R., *Netizens - On The History and Impact of Usenet and the Internet*⁶



Manfred Bonitz publizierte ein ihm von G. M. Dobrov überlassenes Bild einer digitalen Vernetzung zwischen USA, Österreich, VR Polen und der UdSSR aus dem Internationalen Institut für Systemanalyse (IASA), vgl. Abbildung 2.⁷

An der Ausgestaltung und der weiteren gedanklichen Entwicklung war der Autor zu dieser Zeit durch die aktive Teilnahme an IASIA-Konferenzen bzw. Workshops⁸ sowie späteren Beratungen z.B. in Budapest und durch unsere Mitarbeit an der entsprechenden Hauptforschungsrichtung an der Akademie der Wissenschaften der DDR beteiligt.

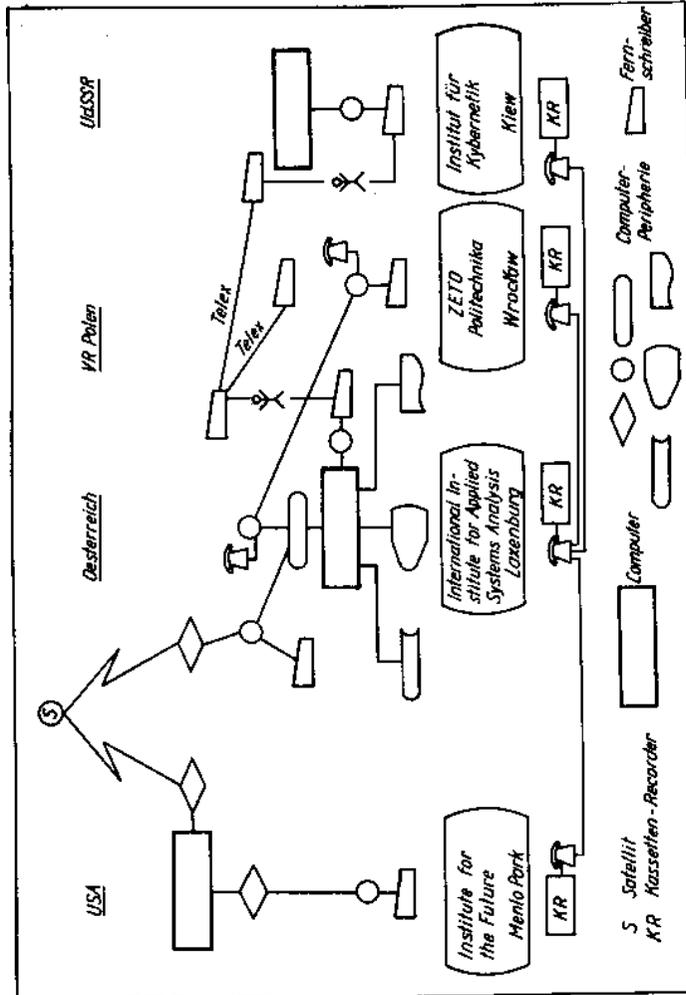
Unser Referat auf dem IASIA-Workshop hatte den Titel: „Mensch-Maschine Kommunikation: Ein Problem der Verbindung von semantischer und syntakti-

6 Hauben, M. / Hauben, R., *Netizens - On the History and Impact of Usenet and the Internet*. Los Alamitos, California, Washington, Brussels, Tokyo: IEEE Computer Society Press 1997, S. 117.

7 Bonitz, M., *Wissenschaftliche Forschung und wissenschaftliche Information*. Berlin: Akademie-Verlag 1981, S. 147; Dobrov, G.M. / Rendolf, R.Ch. / Rauch, V.D., *Informacionnye seti dlja meshdunarodnykh kolektivnykh issledovanij*. Meshdun. forum inform dokum. 3(1978)3, S. 3 - 12.

8 Fuchs-Kittowski, K. / Lemgo, K. / Schuster, U. / Wenzlaff, B., *Man/Computer Communication: A Problem of Linking Semantic and Syntactic Information Processing*. - In: *Workshop on Data Communications*. International Institute for Applied Systems Analysis 2361 Laxenburg, Austria, September 16-19, 1975, CP-76-9. S. 169 - 188.

Abbildung 2: *Neue Kommunikationsformen zur Unterstützung internationaler kollektiver Forschung (nach G.M. Dobrov). Publiziert von Bonitz, M., Wissenschaftliche Forschung und wissenschaftliche Information. Berlin: Akademie-Verlag 1981. S. 147.*



scher Informationsverarbeitung“. Hier wurde verdeutlicht, was heute, mit dem Wissenstransferzyklus generell anerkannt wird, dass mit der Mensch-Maschine Interaktion maschinelle (syntaktische) und menschliche (semantische) zwei qualitativ voneinander unterschiedene Formen der Informationsverarbeitung miteinander kombiniert werden müssen. Es wurde dann weiter verdeutlicht, dass es

verschiedene Paradoxien der Informationsverarbeitung gibt u.a. das Paradoxon des Aufbaus und der Nutzung von Datenbeständen für die Problembearbeitung. Das Paradoxon der Beherrschung großer Datenmengen wurde in Folgendem gesehen:

„Die zunehmende Komplexität und der wachsende Umfang der Verantwortungsbereiche, die aus der gesellschaftlichen Entwicklung der Arbeitsteilung resultieren, erfordern die *inhaltliche Beherrschung* ständig größer werdender Datenmengen durch den einzelnen. Nachgewiesenermaßen versagt die menschliche Kapazität bei der Forderung nach detaillierter *semantischer Beherrschung* von Datenmengen ab einer bestimmten Größenordnung. ... Die Überschaubarkeit und damit Beherrschbarkeit von vielen Millionen Daten ist damit für den einzelnen Nutzer faktisch unmöglich (Paradoxie der Beherrschbarkeit großer Datenmengen)“.⁹

Zur Überwindung dieser u.a. Paradoxien der Informationsverarbeitung wäre ein indirekter Dialog, vermittelt über sog. Informationszentralen erforderlich. Zu dem Gedanken der Einrichtung von Informationszentralen zur Vermittlung zwischen Netz-Nutzer und dem digitalen Netz gab es Diskussion auf dem IIASA-Workshop und auch noch am folgenden Tag, im Shuttle von Wien nach Laxenburg. Einer der Pioniere¹⁰ der Netzentwicklung, Donald W. Davies aus Großbritannien, der den Begriff „packet switching“ eingeführt hatte, sprang plötzlich auf

9 Fuchs-Kittowski, K. / Kaiser, H. / Tschirschwitz, R. / Wenzlaff, B., Informatik und Automatisierung. Berlin: Akademie Verlag 1976, S. 384 ff.

10 Hauben, M. / Hauben, R., Netizen - On the History and Impact of Usenet and the Internet. Los Alamitos, California, Washington, Brussels, Tokyo: IEEE Computer Society Press 1997, S. 117.

Website - Eine andere Bezeichnung für ein Web, eine Webpräsenz, eine Angebot im Internet, einen Webserver.

Websites - Seiten auf einem Server, die aus HTML-Code bestehen und das WWW bilden.

vergl. Kruse, J., Electronic Commerce und Online-Marketing - Chancen, Risiken und Strategien. München, Wien: Carl Hanser Verlag 1999. S. 575.

Der oberflächliche Umgang mit den englischen Fachtermini bei der Übersetzung hat zu einer weit verbreiteten Sprachverwilderung des Deutschen geführt.

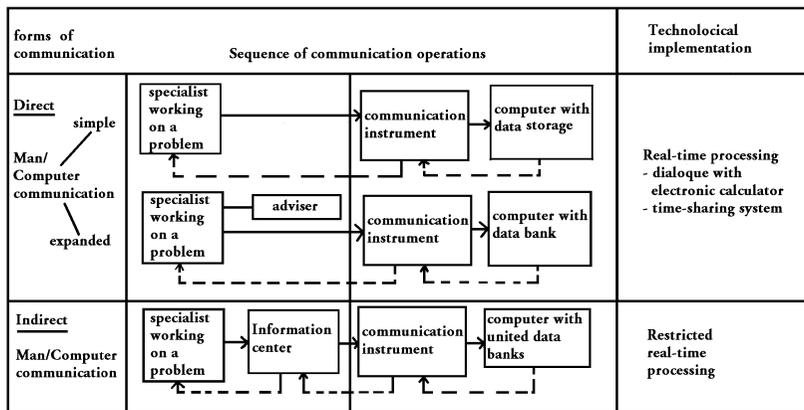
Zur „Web-Site“: Das englische web site bezieht sich darauf, dass sich die Information an einem bestimmten ORT im Informationsnetz befindet, der angeklickt werden muss, um sie zu lesen.

Dieser Prozess ähnelt natürlich im übertragenen Sinne dem Umblättern der Seiten eines Buches. Die Seite eines Buches ist im Englischen aber „page“ (siehe den Ausdruck Homepage, der gern verwendet wird). Solche gibt es durchaus auch im Web, d.h. Dokumente, die aus vielen Seiten bestehen (paginiert sind - eine Frage der Formatierung des Dokuments).

Das Dokument selbst ist aber nur auf einer website lokalisiert. Auch liegt die Besonderheit des Anklickens einer website gerade darin, dass auch Informationen aus einem völlig anderem Datensatz zugänglich werden - im übertragenem Sinne müsste man ein anderes Buch aus dem Regal nehmen und die entsprechende Seite aufschlagen.

und rief: „Sie haben recht, wenn einmal das technische Netz steht, werden die Informationszentralen das eigentliche Netz bilden!“, vgl. Abbildung 3.

Abbildung 3: *Man/Computer communication in (on-line) dialog operation*¹¹



Heute gibt es das Internet und es soll gezeigt werden, dass damit in der Tat die Bildung von Informationszentralen bzw. kompetenten Informationsgemeinschaften zur Bewertung von Informationen/Wissen bzw. der Websites verbunden sein muss.

Es sei noch ein Hinweis auf Aktivitäten in der damaligen Zeit erlaubt.

G. M. Dobrov, mit dem wir zusammen am Orgware-Konzept^{12, 13} arbeiteten und insbesondere der Projektleiter - Netzentwicklung im IIASA, A. Butrimenko, hatten mich darum gebeten, dass ich mich auch an der Ausarbeitung einer Studie zum „grenzüberschreitenden Datenaustausch“ beteilige. Grenzüberschreitender Datenaustausch ist heute in einer enger zusammengewachsenen Welt fast selbstverständlich. Damals war der grenzüberschreitende Datenaustausch jedoch auch

- 11 Fuchs-Kittowski, K. / Lemgo, K. / Schuster, U. / Wenzlaff, B., Man/Computer Communication: A Problem of Linking Semantic and Syntactic Information Processing. - In: Workshop on Data Communications. International Institute for Applied Systems Analysis 2361 Laxenburg, Austria, September 16-19, 1975, CP-76-9. S. 169 - 188.
- 12 Dobrov, G. M., Organisationstechnologie als Gegenstand der Systemanalyse. Teil 1: Grundsatzfragen und Teil 2: Aspekte und Ebenen der Organisationstechnologie. - In: Wissenschaftliche Zeitschrift der Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Reihe, (Berlin). 5 (1979), S. 613 - 622 und 675 - 684.
- 13 Fuchs-Kittowski, K., Wechselbeziehungen zwischen Automat und Gesellschaft - zu Strategien des Einsatzes der automatisierten Informationsverarbeitung als Rationalisierungs- und Erkenntnismittel. - In: Wissenschaftliche Zeitschrift der Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Reihe (Berlin). 5 (1979), S. 707 - 718.

für die westlichen Länder, z.B. aufgrund unterschiedlicher Preise, unterschiedlicher Hard- und Software, noch ein Problem. Vor allem galt es aber für uns, den Eisernen Vorhang zu überwinden. Die politische Zerrissenheit der damaligen Welt taucht in der Skizze von Dobrov nicht auf, aber die Überwindung des Eisernen Vorhangs vermittels dieses Wissenschaftsnetzes war deutlich ein wichtiges Ziel. Man bezog mich in die Studie ein, da ich als Beteiligter am Projekt „Modeling of Health Care Systems“¹⁴ über medizinische Daten aus der DDR verfügte. Dabei ging man von der Annahme aus, das solche Daten „neutral“ seien, vielleicht doch hin und her geschickt werden könnten. Beim ersten Anlauf stieg der polnische Vertreter auch dabei aus, denn man könne doch aus den medizinischen Daten auf die militärische Stärke des Landes schließen. So blieben für die ersten Versuche eines den „Eisernen Vorhang“ überschreitenden Datenflusses, nur die Literaturrecherche übrig. Woran sich die DDR, wie aus der Abbildung 2 ersichtlich, zu diesem Zeitpunkt noch nicht beteiligte. Weitere Auseinandersetzungen um die Beteiligung der DDR an dieser Netzentwicklung und -nutzung habe ich schon an anderer Stelle erwähnt.¹⁵ Das Netz unterstand damals eben doch noch überwiegend dem Militär, so dass alle Versuche eines den Eisernen Vorhang überschreitenden Datenflusses von beiden Seiten zurückgepfiffen wurden.

Was heute unter dem Begriff Internet, Information Highway oder Cyberspace als Infrastruktur einer globalen Informationsgesellschaft für alle diskutiert wird, existierte zu Beginn nur in einer einfachen Form, als ARPA-Net zu militärischen Zwecken und erst seit kurzer Zeit zum Zwecke des Wissenstransfer in der modernen Wissenschaft. Dies sei hier gesagt, damit man das Erreichte auch wirklich wertschätzen kann.

3. Biochemisch-molekularbiologische Forschung und die Entwicklung von Hochleistungscomputern zur Unterstützung neuer Kommunikationsformen internationaler, kollektiver Forschung

Mit dem Internet ist eine neue Qualität der unmittelbaren, technisch vermittelten Kommunikation realisiert worden, wie sie der Buchdruck nicht erreichen konnte.

14 Fuchs-Kittowski, K. / Gudermuth, P., Providing Data for the Management and Planning of Public Health. - In: Systems Modeling in Health Care - Proceedings of an IIASA Conference November 22-24, 1977. Hrsg. v. E.N. Shigan, Laxenburg, Austria: International Institute for Applied Systems Analysis 1978, S. 300 - 313.

15 Fuchs-Kittowski, K., Digitale Medien und die Zukunft der Kultur wissenschaftlicher Tätigkeit. - In: Wissenschaft und Digitale Bibliothek: Wissenschaftsforschung, Jahrbuch 1998. Hrsg. v. K. Fuchs-Kittowski / H. Laitko / H. Parthey / W. Umstätter. Berlin: Gesellschaft für Wissenschaftsforschung 2000 S. 9 - 66.

Zugleich wird deutlich, dass wissenschaftliche Kommunikation notwendigerweise das Bedürfnis nach wissenschaftlicher Information hervorruft. Dieses Bedürfnis zu befriedigen, gilt als Lösung des Problems der wissenschaftlichen Information.

So schreibt M. Bonitz:

„Dieses Problem ist vergleichbar mit den globalen Problemen der Gegenwart, wie der Deckung des Energie- und Rohstoffbedarfs oder der Erhaltung einer sauberen Umwelt für die rapide wachsende Erdbevölkerung. Für solche Probleme ist charakteristisch, dass wesentliche Parameter ein beschleunigtes Wachstum aufweisen: Energie- und Rohstoffbedarf, Umweltgefährdung usw.

Beim Problem der wissenschaftlichen Information sind es die Informationsmassive sowie die Anzahl der Wissenschaftler und der kommunikativen Verbindungen.“

Manfred Bonitz schreibt weiter:

„Wird bedacht, daß die großen Existenzprobleme der Menschheit zu ihrer Lösung aller Potenzen der Wissenschaft bedürfen, die Wissenschaft ihre Produktivkraft nur ganz entfalten kann, wenn ihr ein adäquates System der Wissenschaftlichen Kommunikation zur Verfügung steht, so folgt hieraus die Unaufschiebbarkeit der Anstrengungen zur Bewältigung des Problems der wissenschaftlichen Information.“

Und weiter heißt es:

„Für solche Probleme gilt weiterhin, dass sie an den Ländergrenzen nicht halt machen und dass Empfehlungen, welche die Wissenschaftler ausarbeiten, erst dann voll wirksam werden können, wenn ihre Realisierung im internationalen Maßstab erfolgt¹⁶.

Das Internet ist aus der Kultur wissenschaftlicher Tätigkeit an den Universitäten und Hochschulen und anderen Forschungseinrichtungen heute nicht mehr wegzudenken. Es ermöglicht den raschen Austausch dokumentierten Wissens zwischen den Wissenschaftlern sowie den Zugriff auf ein fast unendlich erscheinendes, globales Angebot an Informationen. Die Zugriffszeiten liegen dabei im Sekundenbereich. Es verschwinden in der Tat die Grenzen zwischen Ländern und Kontinenten. Dies ist eine revolutionäre Veränderung, deren Bedeutung man nur richtig würdigen kann, wenn man die jetzige Situation mit der in früheren Zeiten vergleicht. Technisch gesehen benötigte man für die Übermittlung von Informationen Tage und Wochen. Sozial gesehen, kommt es offensichtlich zu einem solchen Zusammenwachsen der Welt, bei dem ein „Eiserner Vorhang“ keinen Bestand mehr haben kann, ob sich dabei aber auch die Menschen wirklich näher kommen, verlangt wesentlich mehr, als nur die Bereitstellung einer technologischen Infrastruktur.

16 Bonitz, M., Wissenschaftliche Forschung und wissenschaftliche Information. Berlin: Akademie Verlag 1981. S. 143 - 144.

Trotz weltweiter Kommunikation kann es zu einer Verarmung menschlicher Beziehungen, zu einer Einschränkung des persönlichen Gesprächs zwischen den Menschen kommen, wenn es nicht gelingt, zeitgleich mit der Verbesserung der technologischen Infrastruktur und der damit gegebenen massenhaften Vermittlung dokumentierten Wissens auch die zwischenmenschliche Kommunikation und die Bedingungen für Kreativität zur Erzeugung neuen Wissens zu stärken. Telekooperationssysteme (wie Konferenzsysteme) sollen die kooperative Arbeit zwischen weit entfernten Wissenschaftlergruppen unterstützen, aber keinesfalls zu einer Einschränkung der unmittelbaren Kommunikation zwischen den Wissenschaftlern auf Konferenzen usw. führen.

3.1. Besonders leistungsfähige Informations- und Kommunikationstechnologie zur Entschlüsselung des Humangenoms

An der Entschlüsselung des Humangenoms waren leistungsstarke Gruppen in den USA, Großbritannien, Frankreich, Japan und Deutschland beteiligt.

Der wissenschaftliche Wert des Humangenom-Projekts wird gegenwärtig sehr kontrovers diskutiert. Die Einschätzungen bewegen sich vom Vergleich mit der Landung des ersten Menschen auf dem Mond bis zur Betrachtung als reines Prestigeobjekt, dessen finanzieller und personeller Aufwand nicht im Verhältnis zum erzielten Ergebnis stünde. Sicherlich hat jede der beiden Seiten in gewisser Hinsicht recht. Natürlich ist die *Entschlüsselung des menschlichen Genoms* ein bedeutender Meilenstein in der medizinischen Forschung, *der ohne die Revolution der Informationstechnologie nicht möglich gewesen wäre.*

Andererseits darf die Bedeutung dieser Leistung nicht überbewertet werden. Wie so oft in der wissenschaftlichen Forschung werden mit der Lösung eines Problems gleichzeitig neue Probleme aufgedeckt und Grenzen erkannt. *Der Umgang mit der nunmehr bekannten Struktur des Humangenoms birgt eine Reihe von Hindernissen in sich.* So besteht eine nicht zu unterschätzende individuelle Variabilität der Basensequenz, von der nicht immer leicht erkennbar ist, ob sie pathologische Konsequenzen hat oder nicht. Eine weitere Schwierigkeit ist die Tatsache, dass die Gene sowohl aminosäurekodierende als auch nicht kodierende Bereiche (Promotoren, Intron und andere nicht translatierte Regionen) enthalten und nicht in allen Fällen sicher zwischen kodierenden und nicht kodierenden Regionen unterschieden werden kann.

Die Optimisten sagen dagegen: Zwischen kodierenden und nicht kodierenden Sequenzen kann leicht unterschieden werden. Daher weiß man mit Sicherheit, dass mehr als 90% oder sogar 95% der 3 Mrd. Bausteine bzw. der Gesamtsequenz für kein Protein kodiert und auch kein Signal für Transkription und Trans-

lation enthält und möglicherweise nur die Bedeutung eines Vorratsspeichers für Nukleotide besitzt.

Hier wird also die Divergenz zwischen Genverfügbarkeit und Genexpression weniger problematisch gesehen. Genetische Information ist jedoch auf jeden Fall mehr, als ihre syntaktische Struktur - die DNA.¹⁷

Die Analyse der Gensequenz bei gesunden Menschen und bei Personen bzw. Embryonen, die an genetisch bedingten Krankheiten leiden oder später leiden werden, hat dazu geführt und wird weiterhin dazu führen, dass man z.B. mit pränataler DNA-Diagnostik einschätzen kann, ob das Kind hinsichtlich dieser Krankheit krank oder gesund sein wird. Zukünftig werden über den Rahmen von Erbkrankheiten hinaus Sequenzvergleiche verschiedene Aussagen über die Lebenserwartung oder das Risiko für bestimmte Krankheiten ermöglichen. Ein anderes Beispiel der Nützlichkeit ist in der Rechtsprechung der Vergleich von DNA-Sequenzen von kriminalistisch gesicherten Spuren und Verdächtigen.

So wichtig die Kenntnis der Struktur eines Gens auch ist, das angestrebte Ziel des Molekularbiologen ist es, aus der Struktur des Gens die *Struktur des Genprodukts*, des Proteins, vorherzusagen. Aus der Basensequenz ist lediglich die Primärstruktur, d.h. die Aminosäuresequenz, ableitbar. Die funktionellen Eigenschaften eines Proteins werden aber durch seine dreidimensionale Struktur bestimmt. Erst deren Kenntnis gestattet uns Einblicke in seine Wirkungsweise. Aufgrund thermodynamischer Triebkräfte wird die Raumstruktur eines Proteins durch seine Primärstruktur bestimmt; im Prozess der kotranslationalen Faltung der Peptidkette stellen sich die energetisch günstigsten Zustände ein. Diese jedoch richtig vorherzusagen, wird seit ca. 30 Jahren intensiv versucht – mit mäßigem Erfolg. Auch hier kann nur die Revolution der Informationstechnologie perspektivisch zum Durchbruch verhelfen. Gegenwärtig wird an der Entwicklung eines Hochleistungscomputers (*„Blue Gene“*) für die *Vorhersage der Proteinfaltung* gearbeitet, der voraussichtlich in fünf Jahren verfügbar sein soll. *Blue Gene* wird hinsichtlich seiner Leistungsfähigkeit alles bisherige an Computertechnik in den Schatten stellen. Er wird 500 mal schneller arbeiten als die derzeit leistungsstärksten Computer und 1000 mal schneller als *Deep Blue*, der 1997 den Schachweltmeister Gary Kasparov besiegte.

17 Rosenthal, H.A. / Fuchs-Kittowski, K., Genetische Information ist mehr als ihre syntaktische Struktur - die DNA. In: Ethik und Sozialwissenschaften - Streitforum für Erziehungskultur, Licus Verlag, Jg. 12/ 2001, Heft 1, S. 43 - 46.

3.2. Entschlüsselung des Humangenoms und computerisierte Speicherung und Auswertung der biologischer Informationen aus der Genomanalyse

Heute ist jede biochemische Forschung undenkbar, die sich nicht auch molekularbiologischer Techniken bedient. Andererseits ist mit der Identifizierung und Strukturaufklärung von Genen zugleich auch die Wirkungsweise der *Genprodukte*, d.h. der entsprechenden Eiweiße, aufzuklären. Letzteres ist aber Gegenstand der traditionellen Biochemie. Aus diesen Überlegungen heraus erscheint es heute sinnvoll, von einer gemeinsamen Fachdisziplin der Biochemie *und* Molekularbiologie zu sprechen.

Mit der wachsenden gegenseitigen Durchdringung von Biochemie und Molekularbiologie hat die Bedeutung von Datenbanken in der biochemischen Forschung zugenommen.

Mit welcher riesigen Informationsmasse dabei gearbeitet werden muss, wird am Beispiel des gegenwärtig intensiv bearbeiteten und viel diskutierten *humanen Genoms* deutlich. Die neuesten Daten (Nature) zeigen, dass der Mensch kaum mehr als 30.000 Gene hat.

Tabelle 1: *Erhöhung der Computerleistung für die Sequenzierung der DNA*

| | | | | |
|--|------------------|------------------|-------------------|------|
| 1. Genetische Informationsmenge des Menschen 3×10^9 bp (bp = Basenpaare) | | | | |
| 2. Menge der insgesamt in allen Datenbanken der Welt gespeicherte Informationsmenge für Viren, Bakterien, Pilze, Tiere und Menschen in bp: | | | | |
| 1987 | 1989 | 1992 | 2000 | 2001 |
| 10×10^6 | 30×10^6 | 70×10^6 | 100×10^6 | |
| 3. Sequenzierungsgeschwindigkeit pro Arbeitsplatz (Maschine) und Tag in bp | | | | |
| | rohe Daten | | fertige Sequenzen | |
| | 20×10^3 | | 3×10^2 | |
| 4. Informationszuwachsrate ab 1995/2000 (wurde auf der Basis zu erwartender wissenschaftlich-technischer Entwicklung geschätzt) | | | | |
| | 10^6 bp / Std. | | | |
| 5. Zeitdauer der vollständigen Sequenzierung der DNA eines Probanden mit 3×10^9 bp (rohe Daten): in Tagen | | | | |
| | 1992 | 2000 | | |
| | 10^5 | 10^2 | | |

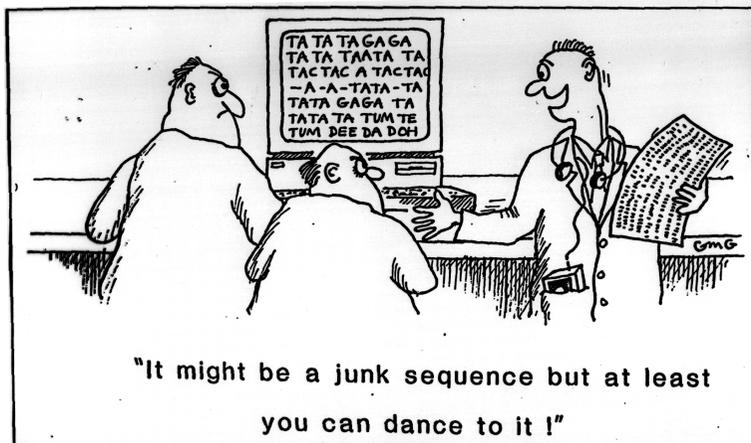
Die Analyse des menschlichen Genoms ist in den letzten Jahren enorm vorange-
trieben worden.

Die Informationsmengenwachstumsrate wurde noch mehr gesteigert, als in der Voraussicht vorhergesehen war, so dass nun das Projekt praktisch abgeschlossen werden konnte, vgl. Tabelle 1.

Durch ständige Verbesserung der Sequenzierungstechniken konnte innerhalb von 3 Jahren der Preis von 2 \$ pro Base auf 0,3 \$ pro Base gesenkt werden. Der größte Teil des Genoms - 90% - wurde inzwischen entziffert, für 10% gibt es noch eine Unsicherheit. Diese hochgenaue und fast vollständige Gesamtsequenz wurde bis 2003 erwartet, wurde aber aufgrund des hohen Drucks denen die Wissenschaftler ausgesetzt waren, schon jetzt erreicht.

Eine Tücke dieser gewaltigen Informationsmasse besteht darin, dass ein sehr großer Teil des Genoms anscheinend nicht zur Expression gelangt und wir deshalb die Funktionen dieser Abschnitte nicht kennen.

Abbildung 4: *Mit dem Scherz soll vermittelt werden, dass besonders große Datenmengen bewältigt werden müssen, um eine codierte Sequenz zu finden. (Nature)*



Unter den zur Expression gelangenden Genen sind für die medizinische Forschung in erster Linie diejenigen von Interesse, deren Ausfall oder Veränderung zu Erkrankungen führt. Mit den Methoden der modernen Gentechnologie können nämlich bislang als unheilbar geltende Erbkrankheiten in der Perspektive erfolgreich behandelt werden. Bisher sind lediglich ca. 500 bis 1000 solcher „Krankheitskeime“ beim Menschen bekannt. Daraus folgt, dass der medizinorientierte Biochemiker mit einem gewaltigen Überschuss an Information konfrontiert wird, aus dem er die für ihn relevante Information selektiv herausfil-

tern muss. Dies ist ohne die heutigen Mittel der Informationstechnologie kaum möglich. Dazu gehören Computer, das Internet, diverse Datenbanken sowie Softwarepakete, die die Daten miteinander verknüpfen.

Durch den Vergleich der Basen- bzw. Aminosäuresequenzen von Genen, Genabschnitten, Transkriptionsprodukten oder Polypeptiden mit den in Datenbanken gespeicherten Sequenzen bekannter Gene oder Genprodukte ist es möglich festzustellen, ob diese bereits bekannt sind, um sie dann entsprechend zuzuordnen. Sollte sich herausstellen, dass es sich um ein bisher unbekanntes Gen oder Genprodukt handelt, so ist für den Biochemiker von Interesse, mit welchen bereits bekannten Genen gewisse Ähnlichkeiten bestehen. Daraus können wichtige Schlussfolgerungen hinsichtlich der Evolution der Gene und ihrer Verwandtschaft gezogen werden.

In der Abbildung 5 wird speziell die Bedeutung des Computereinsatzes für die Sequenzanalyse sowie die Nutzung von Datenbanken für den Vergleich der Basen bzw. Aminosäuresequenzen mit den in Datenbanken gespeicherten Sequenzen bekannter Gene oder Genprodukte hingewiesen, (vgl. Kasten Experimentauswertung, Vergleich und Datenbanken). Damit wird die besondere Bedeutung der modernen Informations- und Kommunikationstechnologien für den biochemischen/molekularbiologischen Forschungsprozess verdeutlicht und in den Gesamtprozess der Kultur wissenschaftlicher Tätigkeit auf diesen Gebieten eingeordnet.

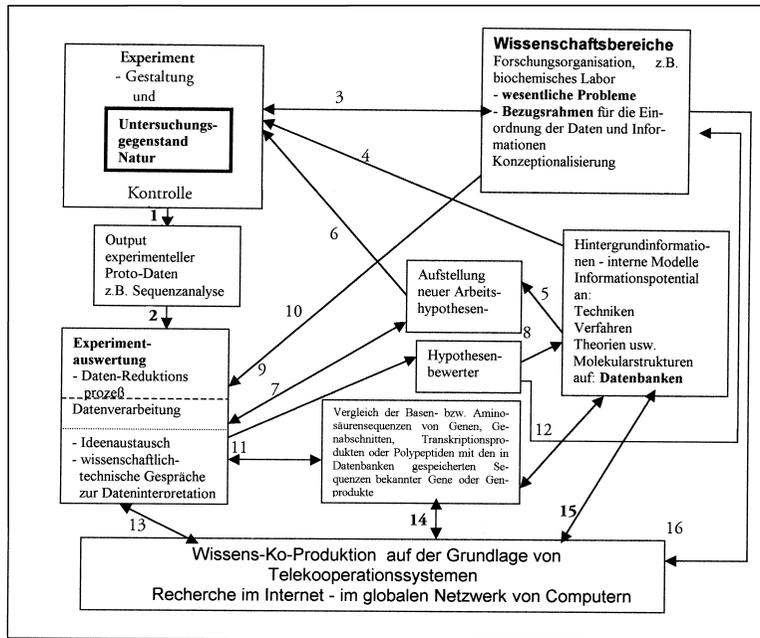
3.3. Der wissenschaftliche Forschungsprozess als Informations- und Kommunikationsprozess in der Wissenschaftsorganisation

Speziell unter dem Einfluss des Einsatzes moderner Informations- und Kommunikationstechnologien zur Unterstützung des wissenschaftlichen Forschungsprozess, wird dieser selbst als ein Informations- und Kommunikationsprozess in Organisationen der Wissenschaft verstanden. Unter dem Einfluss bestimmter erkenntnistheoretischer Positionen kann er dabei entweder auf ein reines Datenverarbeitungssystem reduziert werden, z.B. aus der Sicht des logischen Positivismen, oder er wird als ein Informationssystem verstanden, wobei jedoch immer noch von der sozialen Organisation abstrahiert wird, in dem er mit dem Informationssystem identifiziert wird.

Wie die Abbildung 5 zeigen soll, vollzieht sich der wissenschaftliche Forschungsprozess, mit seiner intensiven Unterstützung durch moderne Forschungsgeräte, leistungsfähige Computer und globale sowie lokale digitale Netze, in der sozialen Organisation der Wissenschaft. Dies bei der Frage nach der „Objektivität der Wissenschaft“¹⁸ zu berücksichtigen, bedeutet die Überwindung eines naiv-

realistischen Standpunktes, demzufolge eine Erfahrbarkeit der Realität „an sich“, unabhängig von sensorischen oder kognitiv bedingten Einflüssen des erkennenden Subjekts, möglich sei. Es muss aber nicht bedeuten, dass man in einen post-modernen Relativismus oder gar dem Solypzismus verfällt.

Abbildung 5: Struktur des experimentellen Forschungsprozesses



Legende zur Abbildung 5

Die Abbildung 5 zeigt den experimentellen Forschungsprozess in den Naturwissenschaften. Es zeigt die Beziehungen zwischen dem Bereich (der Welt) der bei einer experimentellen Untersuchung gewonnenen Daten (Tatsachen) und der wissenschaftlichen Beschreibung der Welt.

Die Pfeile geben die Richtung der Determinierung an.

Wäre das Wissenschaftssystem mit Pfeil 1 und 2 erschöpfend dargestellt, wäre die Beschreibung der Welt, im Sinne einer einfachen Widerspiegelung eindeutig vom Untersuchungsgegenstand allein bestimmt.

Mit der Überwindung des naiven Realismus, der letztlich unterstellt, dass man die Welt so erfasst wie sie ist, wurde klar, dass man die Aktivität des erkennenden Subjekts berücksichtigen muss. Dies bedeutet die Einsicht, dass wissenschaftliche Ergebnisse ihre wissenschaftliche Erklärung in dem sie hervorbringenden Prozess finden.

18 Mühlhölzer, F. Die Debatte um die Objektivität der Wissenschaft. - In: Deutsche Zeitschrift für Philosophie, Berlin. 49 (2001)1, S. 151 - 166.

Daten können nicht unproblematisiert zu Hypothesenbildung- und -bestätigung oder Falsifizierung herangezogen werden. Denn bevor sie als Daten interpretierbar und reduzierbar sind und als empirischer Beweis gelten können, wurden Praktiken der Experimentgestaltung und -kontrolle, der Replikation, der wissenschaftlichen Argumentation und Kritik, als wesentliche Teile der wissenschaftlichen Methode angewendet (Pfeil 1).

Nun wird jedoch noch ein weiterer Schritt getan. Es wird mit dem Schema verdeutlicht, dass wissenschaftliche Erkenntnis nicht nur Ergebnis individueller Analysen und Einsichten in komplexe Zusammenhänge darstellt, sondern auch zu beachten ist, wie die wissenschaftliche Erkenntnis erzeugt wird.

Die Wissensproduktion bringt Wissensprodukte hervor. Diese Wissensprodukte sind nicht mehr nur in Bezug auf den Untersuchungsgegenstand „Natur“ zu sehen. Die Wissensprodukte müssen auch in Bezug zu ihren Erzeugungsprozess gesehen werden. So erfährt die wissenschaftliche Tätigkeit, das Forschungshandeln, größeres Interesse. Der Untersuchungsgegenstand (das Objekt) aus der Natur wird im Labor mit bestimmten Organisationsbedingungen konfrontiert (Pfeil 3)¹⁹. Es ist ein wichtiger Unterschied ob z.B. Feldforschung im Sinne der Agrarwissenschaft durchgeführt oder ob diese abgelöst wird von der Laborarbeit bei der Biotechnologie. Dies führt zu einer Erhöhung der Macht des Forschers gegenüber dem Untersuchungsgegenstand. Hierbei spielen die Faktoren, wie die Verfügbarkeit, die schnellere Zugänglichkeit (Bereitstellung) von Wissen und Gerät, schnellere Durchführung der Untersuchungsprozesse sowie Wiederholbarkeit und Kontrollmöglichkeiten eine wichtige Rolle (Pfeil 4)²⁰. Die schnelle Bereitstellung von Techniken, Verfahren, Theorien und Hintergrundwissen über online Datenbanken wird für den experimentell arbeitenden Forscher zu einer wichtigen Stütze seiner Arbeit. Insbesondere kann das bereitgestellte Wissenspotential wesentlich die Aufstellung neuer Arbeitshypothesen unterstützen (Pfeil 5), die dann die Voraussetzung für neue experimentelle Anfragen an die Natur bilden (Pfeil 6)²¹. Grundlage für die Hypothesenbewertung und die Aufstellung neuer Hypothesen ist die Experimentauswertung, die speziell im wissenschaftlichen Gespräch im Laboratorium bzw. wissenschaftlichen Einrichtung erfolgt (Pfeile 7). Positiv bewertete Hypothesen werden ins Informationspotential aufgenommen und bei Bedarf zur Verfügung gestellt (Pfeil 8). Im wissenschaftlichen Dialog werden die Daten interpretiert, erfolgt die Hypothesen- und Theorienbildung (Pfeil 9). Die eigentliche Wissenserzeugung ist also eine Wissens-Co-Konstruktion. In einem experimentell arbeitenden Labor entwickelt sich eine entsprechende Gesprächskultur der wissenschaftlichen Arbeit.

Wir haben also auch hier die Kombination von syntaktischer und semantischer Informationsverarbeitung, wie sie in jedem Wissenstransferzyklus auftritt. Das Zusammenspiel zwischen den instrumentellen Formen syntaktischer Information und Kommunikation mit den menschlichen Formen des Forschungshandeln muss speziell untersucht und herausgearbeitet werden. Die Einbettung des instrumentellen Informations- und Kommunikationssystem in die soziale Organisation der Wissenschaft wird nochmals darin deutlich, dass nur die soziale Organisation, die wissenschaftlichen Gespräche im Labor, die kompetente wissenschaftliche Gemeinschaft, den Bezugsrahmen für die Einordnung der

- 19 Knorr-Cetina, K., *The Manufacture of Knowledge*. Oxford: Pergamon Press 1981; dt., *Die Fabrikation von Erkenntnis*. Frankfurt am Main: Suhrkamp 1984.
- 20 Fuchs-Kittowski, K. / Parthey H., *Veränderung der Forschungssituation durch Entwicklung der Informationstechnologie*. - In: *Arbeitstagung Forschungstechnologie '87 Informationstechnologien als Teil der Forschungstechnologie in den experimentellen Wissenschaften*, Tagungsmaterial, Akademie der Wissenschaften der DDR, 1987.
- 21 Prigogine, I. / Stengers, I., *Dialog mit der Natur. Neue Wege naturwissenschaftlichen Denkens*. 2. Auflage. München, Zürich: R. Piper & Verlag 1981.

Daten, für die inhaltlichen Bestimmung der Daten und Informationen zur Erzeugung von Wissen liefern kann (Pfeil 10).

Moderne biochemische und molekularbiologische Forschung ist ohne den Einsatz moderner, sehr leistungsfähiger Informations- und Kommunikationstechnologie undenkbar. Wie das Beispiel der Entschlüsselung des Humangenom zeigt, kommen gerade auch aus dieser Forschung weitere Anforderungen die Leistungsfähigkeit der Rechner und Kommunikationsnetze noch zu steigern.

Benötigt werden Hochleistungscomputer wie Deep Blue oder künftig "Blue Gene".

Es ging und geht hier insbesondere um die Senkung der Zeitdauer für die Sequenzierung der DNA eines Probanden, es geht um die Möglichkeit des Vergleichs von Molekularstrukturen, es geht um einen schnellen und sicheren Zugriff auf die neuesten Forschungsinformationen (vgl. Pfeil 11 u. 12.)

Der Einsatz moderner Informations- und Kommunikationstechnologien im Forschungsprozess steht heute in einem engen Zusammenhang mit dem Einsatz von Telekooperationssystemen zur Unterstützung kooperativer wissenschaftliche Arbeit (Wissens-Ko-Produktion) relativ unabhängig von Raum und Zeit sowie mit der Nutzung des Internets zur Bereitstellung wissenschaftlicher Informationen. (vgl. Pfeil 14 u. 15).

Eine naiv-realistische Erkenntnisposition wird in der heutigen Wissenschaftsphilosophie kaum noch vertreten. Sie gilt als überwunden. Moderne Erkenntnispositionen, wie der hypothetische Realismus der evolutionären Erkenntnistheorie oder eines konstruktiven Realismus, wie er hier vertreten werden soll, erkennen die aktive, konstruktive Leistung des erkennenden Subjekts an. Damit muss jedoch nicht bestimmten Übertreibungen in der Wissenschaftssoziologie gefolgt werden, nach denen das wissenschaftliche Ergebnis in der sozialen Organisation der Wissenschaft ausgehandelt würde, wie in der Politik. Wichtig und richtig ist jedoch u. E., dass die Bedeutungen unserer Begriffe mit denen wir an den Untersuchungsgegenstand herangehen, mit denen wir nach Literatur in online Datenbanken suchen, in einem sozialen Prozess, in der wissenschaftlichen Gemeinschaft, entstehen.

Mit Pfeil 16 soll daher nochmals, jetzt in Bezug auf die Recherche im Internet, hervorgehoben werden, dass die soziale Organisation, die kompetente wissenschaftliche Gemeinschaft, den Bezugsrahmen für die Einordnung der Daten liefert. Kein Informationssystem ist für sich selbst in der Lage diesen Bezugsrahmen zu erstellen.

In der gewöhnlichen Auffassung von der Wissenschaft wird die Welt der Natur als etwas von uns unberührtes (objektives) angesehen und die Wissenschaft als ein Versuch eine wörtliche Beschreibung dieser Welt (durch Protokollsätze) zu gewinnen. Diese Konzeption von der Wissenschaft entspricht den Vorstellungen eines naiven Realismus, der von einer reflektiven Beziehung zwischen der Welt und ihrer wissenschaftlichen Beschreibung ausgeht.

Eine solche Position impliziert, dass die Forschungstechnologie sowie andere intellektuelle Gegebenheiten bestimmte Bedingungen der Forschung darstellen, die der Gewinnung von Wissen förderlich sein können, aber nicht die Wissensinhalte beeinflussen. In Wirklichkeit gibt es keine Erkenntnis, keine wissenschaftlichen Inhalte, die nicht wesentlich von der sozialen Form des Forschungshandelns mit bestimmt sind.

Das was wir als „wirklich“ und „wahr“ bezeichnen, ist nicht allein vom Forschungsgegenstand determiniert, sondern auch von den theoretischen Konzepten, den Untersuchungsmethoden. Unser Wissen ist eine Konstruktion über die Welt. Erkenntnis ist also eine kulturell erzeugte Welt. Damit wird die Existenz ei-

ner vom Menschen unabhängig existierenden Realität nicht gelegnet. Es wird aber deutlich gemacht, dass wir zu ihr keinen Zugang außerhalb wissenschaftlicher Praktiken der Wissenserzeugung haben.

4. *Informationsrecherche im Internet*

Zu Beginn des neuen Jahrhunderts und Jahrtausends stehen wir auch an der Schwelle von der Industrie- zur Informations- und Wissensgesellschaft. Moderne Kommunikationsnetze mit Multimedia-Endgeräten stellen die technische Infrastruktur für eine solche Entwicklung dar. Für diese Infrastruktur wurden Begriffe wie, Information Highway, Cyber-Space oder Hyper-Space geprägt. Vielfältig und kontrovers sind die Erwartungen und Befürchtungen, die angesichts der Ambivalenz der Wirkungen moderner Informations- und Kommunikationstechnologien mit dieser technologischen Entwicklung verbunden werden.

Die neuen Anwendungen sind auf die Erhöhung der Lebensqualität für den Einzelnen und auf die Steigerung der Effektivität und Effizienz der Geschäftsprozesse gerichtet. Teleworking, Teleshopping, Online-Datenbanken sowie interaktive Teleunterhaltung sind dabei unser Arbeits- und Lebensweise wesentlich zu verändern.

Ein Problem, welches unzweifelhaft auf uns zukommen wird, ja uns heute schon sehr belastet, ist die *Informationsüberflutung*. Auf der einen Seite leiden wir unter dem Übermaß an angebotener Information, fehlt es an erforderlichen Mechanismen zur Filterung der Informationen, auf der anderen Seite sind wir oft nicht in der Lage, gezielt eine bestimmte Information wieder aufzufinden. Das Problem des Wiederauffindens von Informationen - des Information Retrieval - wird uns in Zukunft noch weit mehr beschäftigen. Wir brauchen leistungsfähige Werkzeuge zur Navigation, um in der Datenflut effektiv surfen zu können. Die Entwicklung des World Wide Web war ein wichtiger Schritt aber nicht die Lösung sondern erst der eigentliche Beginn des Problems. Insbesondere sind folgende technischen Probleme zu lösen:

- es bedarf schneller Kommunikationsnetze, um die großen Datenmengen zu übertragen,
- es bedarf leistungsfähiger Rechner, um die Daten verarbeiten und verteilen zu können,
- es bedarf leistungsfähiger Software, um die Komplexität der verteilten Umgebung zu bewältigen und dem Internet-Nutzer den Zugang zu erleichtern,

- es bedarf insbesondere neuer Methoden der Informationsrecherche im Internet. Hierbei auch Möglichkeiten der Website-Evaluation - der Bewertung der Informationsinhalte durch den Nutzer.

Das World Wide Web (WWW) wurde 1991 von dem Informatiker Tim Berners-Lee am Genfer Hochenergieforschungszentrum CERN zum Informationsaustausch zwischen Wissenschaftlern entwickelt.

Das World Wide Web im Internet ermöglicht den Zugriff auf multimediale Dokumente. Die Kommunikation ist hierbei also nicht auf die Übermittlung von Texten beschränkt, sondern es werden auch Sprach- und Videoinformationen übermittelt. Es ermöglicht eine einfache Textformatierung und die Einbindung von Grafik sowie den Verweis (Hyperlink) auf Dokumente, welche auf einem beliebigen Server im Internet liegen.

Es entstand die Auszeichnungssprache HTML (Hypertext Markup Language) und das neue Internet-Protokoll HTTP (Hypertext Transfer Protocol).

Aufgrund des Hypertext-Charakters wurde das ganze Projekt World Wide Web (weltweites Netz) genannt.

4.1. Die Ausmaße des Webs

Seither wächst das Web exponentiell.

Die Ausmaße des Webs sind heute etwa:

- Juli 2000: 2,1 Milliarden Websites,
- täglich kommen etwa 7 Millionen hinzu und
- Januar 2001 sollen es schon über 3,5 Milliarden Sites sein, wurde in einer Studie von Cyveillance ermittelt.

Das globale Wissen wäre nun nicht weiter entfernt als das nächste Computerterminal - wenn man es wirklich finden könnte.

Aufgrund dieses rapiden Wachstums hat das World Wide Web kaum eine Organisation noch eine Struktur.

Aufgrund der Tatsache, dass jedermann, aus jedem Land, mit beliebigen Hintergrund, in beliebiger Sprache eine Website schreiben kann, wird das WWW zu einem ständig wachsenden weltweiten „informationellen Müllhaufen“.

Den wachsenden „Informationsmüll“ zu filtern, wird daher immer wichtiger.

Die Informationsüberflutung zwingt dazu, relevante Information aus der Überfülle an Information zu selektieren. Das Informationsmanagement zur Gewinnung von Wissen aus den In-Beziehung gesetzten Informationen und speziell die Informationsrecherche ist eine grundlegende Anforderung an den Nutzer des Internets.

Um so erstaunlicher ist es, dass die bisherigen Suchdienste offensichtlich nicht in der Lage sind die Situation zu meistern:

Es ist keine vollständige Katalogisierung möglich, da:

- es keinen Zwang zur Registrierung gibt, um ein Dokument ins Web zu stellen,
- es kein einheitliches Klassifikationsschema gibt,
- jeden Tag neue Dokumente hinzu kommen,
- jeden Tag wieder Dokumente herausgenommen werden,
- jeden Tag Dokumente verändert werden und
- einige der Dokumente von dynamischer Natur sind.

Das Internet ist um ein Vielfaches größer als eine Suchmaschine bisher abdecken kann.

Selbst Google, der derzeit größte Internet-Index, hat gerade einmal, mit 1,25 Milliarden Seiten, die Hälfte der verfügbaren Seiten erfasst. Die beiden nächst größeren Indizes haben demnach gerade einmal 1/4 der verfügbaren Seiten (Ink-tomi und Webtop, über 500 Millionen) indiziert. Das bedeutet nicht, dass damit Google nicht auch jetzt den Wissenschaftler bei der Suche nach Informationen, zur Vorbereitung auf Konferenzen, bei der Suche nach für sie wichtiger Literatur wertvolle Hilfe leistet.

Dies haben auch die Betreiber von Suchmaschinen erkannt und deshalb arbeiten heute verschiedene Suchdienste zusammen z.B. Alleskar und Fireball.

Die riesige Größe des Internets macht deutlich, dass die Qualität der Seiten von den Suchmaschinen stärker bewertet werden muss. Spezielle Verzeichnisse (Portale) und kleine, spezialisierte Suchmaschinen werden daher an Bedeutung gewinnen.

In der Studie von Cyveillance wurden auch bestimmte Durchschnittswerte für Websites ermittelt:

- Durchschnittlich Größe einer Website: 10.060 Byte
- Durchschnittliche interne Links pro Seite: 23
- Durchschnittliche externe Links pro Seite: 5,6
- Durchschnittliche Anzahl von Bildern pro Seite: 14,4
- Anteil US-Websites: 84,7%
- Internationale Websites: 15,3%

Heute gibt es schon eine Dominanz des kommerziellen Web. Die kommerziellen Seiten machen 83% aus. Wissenschaft und Bildung dagegen nur 6%.

4.2. Die Webmap

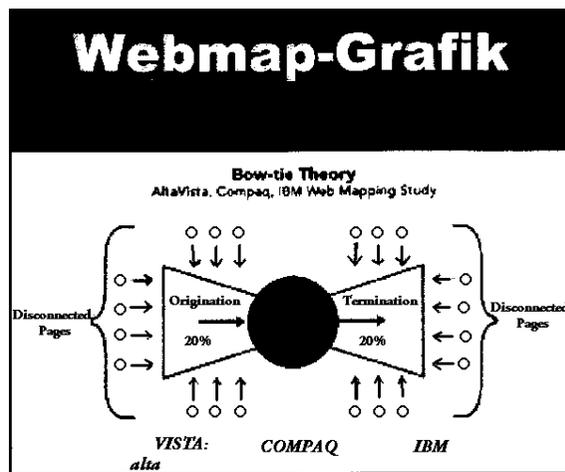
Eine Studie von Compaq, IBM und Alta Vista vom Mai 2000 zeigt den Betreibern von großen Suchmaschinen, dass es immer wichtiger für sie wird, das gesamte Web zu erfassen, denn nur so könnten sie ihre Kompetenz darstellen, vgl. Abbildung 6. Nur wer das ganze Web indiziert habe, könne im Wettbewerb bestehen.

Der gewonnene komplette Plan des WWW ergibt viele „tote Ecken“.

Es gibt vier Regionen mit verschiedenen Stufen der Verlinkung:

- Stark verknüpfter Kern - 39%
- Ursprung - 21%
- Abschluss - 19%
- Distanz - 14%

Abbildung 6: *Die Webmap - Ergebnis einer Studie von Alta Vista, Compaq und IBM*



Der neue Plan des WWW gibt der Forschung die Möglichkeit effektivere Suchtechniken zu entwickeln, die Pfade der Websites zu verfolgen und diese nicht abzuschneiden. Mit dem durch die Studie gewonnenen Verständnis der Webgeographie können die Websites-Betreiber die Position ihrer Suchmaschine stabilisieren oder auch erhöhen. Mit mehr richtigen Startpunkten können die Crawler mehr Seiten des gesamten Web's erfassen.

4.3. Suchdienste

Sucht man im World Wide Web nach Informationen, so geschieht dies unter zu Hilfenahme eines Suchdienstes.

Zur Suche im Web stehen folgende Möglichkeiten zur Verfügung:

1. Suchmaschinen
2. Verzeichnisse
3. Metasuchmaschinen
4. Online-Datenbanken
5. Agenten
6. Spezialsuchmaschinen

Fast alle Suchdienste benutzen heute eine Kombination aus Verzeichnissen (Katalog) und Suchmaschine.

Was ist eine Suchmaschine?

Eine Suchmaschine ist ein interaktives Werkzeug zur Unterstützung des Menschen beim Auffinden von Informationen, die durch das World Wide Web (WWW) verfügbar sind.

Die Web Suchmaschinen sind Datenbasen, die Referenzen zu Tausenden von Ressourcen haben. Die Nutzer sind in der Lage mit den Datenbasen zu interagieren, indem sie Fragen stellen, wodurch die Datenbasen danach befragt werden, ob sie Ressourcen enthalten, die den entsprechenden Kriterien genügen.

Der Index wird automatisch durch Robots - die eigentlichen Datensammler erzeugt, auch Spider, Crawler oder Worm genannt, die durch das WWW „kriechen“ und die Daten zur Schaffung des Index zusammensuchen.

Diese Begriffe sind irreführend, da die Robots nicht durch das Internet hindurchkriechen, sondern wie ein normaler Browser die Seiten vom Webserver anfordern und anschließend automatisch auswerten.

Das Interface liefert dem Nutzer eine Formblatt, in das er einen „Suchbegriff“ - Zeichenkette - eingeben kann. Dies kann ein Wort oder eine Phrase oder ein Datum sein. Die Suchmaschine verarbeitet den Suchbegriff gegen die Datenbasis/Index und liefert eine Liste der Ressourcen die den Kriterien entsprechen und zeigt dem Internet-Nutzer das Ergebnis.

Es gibt zwei Möglichkeiten, wie Suchmaschinen zu indizierenden Websites gelangen:

- Die Website wird vom Website-Betreiber bei den Suchmaschinen angemeldet.
- Die Robots folgen den Links angemeldeter bzw. bereits indizierter Websites.

85% der Nutzer benutzen Suchmaschinen, 90% folgen Links aus Linklisten, jedoch auch aus Katalogen.

Das Internet ist also eine weltweite Informationsquelle, doch 44% der Internet-Nutzer sind mit der gegebenen Webnavigation und den vorhandenen Suchmaschinen unzufrieden. Die Tatsache, dass beinahe die Hälfte aller Nutzer unzufrieden ist, *verweist auf die Größe des Problems*.

Die Arten, in denen die verschiedenen Suchmaschinen die Daten auswerten und indizieren, unterscheiden sich wesentlich.

Jede Website kann im nicht sichtbaren Bereich einige Meta-Tags zur Beschreibung der Seite enthalten. Manche Suchmaschinen, wie Lycos, beachten diese Meta-Tags nicht, während die Angaben der Meta-Tags für andere Suchmaschinen besonders wichtig sind. Zum anderen unterscheiden sich die Suchmaschinen in der Anzahl der Worte, die ausgewertet werden. Manche indizieren die gesamte Website, während andere nur die ersten Worte oder Sätze auswerten.

Aufgrund dieser unterschiedlichen Vorgehensweise werden Suchmaschinen in folgende Grundkategorien eingeteilt:

- Volltextsuchmaschinen
- Speichern von Meta-Daten als Verschlagwortung
- Speichern von Wort-Statistiken - Stichwortsuche.

Für die Wortstatistik wird jedes Schlüsselwort nur einmal abgelegt, zusammen mit der Information, wie oft es in der Website an bestimmten Stellen (Überschrift, Meta-Tags, Text) enthalten ist.

Verzeichnisse

Web-Verzeichnisse (Kataloge) erhalten ihre Informationen, indem die angemeldeten Seiten zumindest bei der Aufnahme in den Katalog, vom Menschen angesehen wird.

Dadurch ist der Umfang der indizierten Websites wesentlich kleiner als bei den Suchmaschinen. F. J. Davadason geht in seinem Vortrag: „On Searching the Web“²² davon aus, dass Verzeichnisse nur 1-2% der verfügbaren Websites indiziert haben.

22 Davadason, F.J., Präsentation: „On Searching the Web“, Siemens, DAAD, AIT Summer School, Conversion of Computing, Communication and Content, at the Asian Institute of Technology (AIT), Bangkok, 2000 (Lehrmaterial des Center for Library & Information Resources).

Meta-Suchmaschinen

Meta-Suchmaschine im WWW machen die Suche effizienter. Die Anfrage wird zentral eingegeben und das Ergebnis erhält man von multiplen Suchmaschinen. Die Meta-Suchmaschinen unterscheiden sich von anderen Suchmaschinen und Verzeichnissen:

- Die einzelne Suchmaschine und Verzeichnisse stellen eine Sammlung oder Datenbasis an Ressourcen zur Verfügung nach denen gefragt wird.
- Meta-Suchmaschinen besitzen keine eigene Datenbasis. Meta-Suchmaschinen geben eine an sie gestellte Anfrage an verschiedene Suchmaschinen und Verzeichnisse weiter, werten die Ergebnisse der einzelnen Suchmaschinen aus und fassen sie zu einer Ergebnisliste zusammen. Meta-Suchmaschinen realisieren also eine Dienstleistung durch die eine einzelne Anfrage an multiple Datenbasen geschickt werden. Damit steht ein größerer Datenbestand zur Verfügung, der mit einer Anfrage durchsucht werden kann.

Ein Meta-Sucher ist also eine Dienstleistung zur Auffindung Internetressourcen. Bekannte Suchmaschinen dieser Art sind u.a.: Apollo 7, Metacrawler, Metager.

Online-Datenbanken

Für die Informationssuche im Internet sind Online-Datenbanken in Form von Lexika und Nachschlagewerke, für die Wissenschaft auch insbesondere „Handbücher“ von besonderem Interesse, da man dort deutlicher die Spreu vom Weizen trennen kann und somit sicherer qualitativ hochwertige Informationen findet.

Der Inhalt von Online-Datenbanken und daraus generierte dynamische Webseiten werden jedoch von den Suchmaschinen nicht indiziert. Die wesentlichen Gründe dafür liegen in der schnellen Veränderung der Inhalte dieser Datenbanken und im Umfang dieser Datenbestände.

Zu solchen Online-Datenbanken gehören:

- Enzyklopädia Britannica
- Meyers Lexikon
- Weltchronic
- Wissen.de

Es gibt über 110 gebührenfreie Datenbanken im WWW.

Agenten

Unter Agenten im WWW versteht man eine neue Art von Suchdiensten. Wie der Name Agent schon zum Ausdruck bringt, soll die vom Anwender auf den Agen-

ten delegierte Suchfrage autonom und zufriedenstellend erfüllt werden. Solche Suchagenten sollen in der Lage sein, den Auftrag anzunehmen und die im WWW verfügbaren Daten zu interpretieren, um sie dem Auftrag entsprechend bereitzustellen. Sie sollen darüber hinaus auch lernfähig sein, so dass durch die gesammelten Erfahrungen immer treffsicherere Ergebnisse geliefert werden.

Es wird zwischen allgemeinen, spezialisierten und Meta-Suchagenten unterschieden.

Welche der Suchmaschinen, Meta-Suchmaschinen, Agenten oder Meta-Suchagenten unterschiedlicher Leistungsfähigkeit auch immer zum Einsatz kommen werden, wird man bestimmte erkenntnistheoretische und anthropologische Gesichtspunkte und insbesondere Erkenntnisse aus der Kybernetik 2. Ordnung, der hier entwickelten Biologie der Erkenntnis und Wissenskonstruktion berücksichtigen müssen.

5. *Informationsrecherche auf der Grundlage spezifischer Datenbanken (im Internet) und die Biowissenschaften*

Biochemie und Molekularbiologie sind zunehmend miteinander verflochten und bilden eine gemeinsame Fachdisziplin.²³ Durch die Einbeziehung von Datenbanken und anderen Mitteln der Informationstechnologie hat der interdisziplinäre und integrale Charakter der biochemisch-molekularbiologischen Forschung zugenommen. Auf diese Weise sind grundlegende Entdeckungen nicht mehr allein den Spezialisten vorbehalten. Dies wird am Beispiel der kürzlichen Entdeckung einer neuen Gruppe von entzündungshemmenden Arzneimitteln belegt.

Dieses konkrete Beispiel aus einem Arbeitsgebiet der Biochemie - der Eikosanoidforschung - zeigt, wie durch die Einbeziehung von modernen Datenbanken neue Erkenntnisse erzielt wurden, die mit der traditionellen Herangehensweise in der Forschung nicht möglich gewesen wären. Dieses Beispiel ist die Entdeckung der Cyklooxygenase-2 als Basis für die Entwicklung einer neuen Gruppe von Arzneimitteln.

Sprechen wir von Informationsrecherchen auf der Grundlage spezifischer Datenbanken, dann soll hier das medizinische Literatursuchsystem *PubMed* besonders hervorgehoben werden. Mit seiner Hilfe ist es für den medizinischen Biochemiker möglich geworden, mit der gewaltigen und ständig wachsenden Flut an

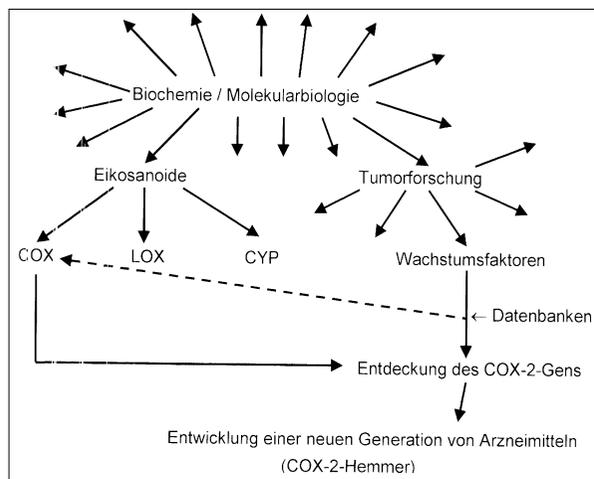
23 Schewe, T., Neue Horizonte der biochemisch-molekularbiologischen Forschung im Lichte der Revolution der Informationstechnologie zur Jahrtausendwende. In: Stufen zur Informationsgesellschaft für alle. Festschrift zum 65. Geburtstag von Klaus Fuchs-Kittowski. Hrsg. v. Chr. Floyd / W. Hofkirchner / C. Fuchs. Frankfurt a.M.: Peterlang-Verlag 2002, im Druck.

Literatur sinnvoll umzugehen. Durch geschickte Wahl geeigneter Schlüsselwörter können mit geringem Aufwand die relevanten Quellen schnell und weitgehend vollständig herausgefiltert werden. Damit ist es für den Wissenschaftler auch erheblich leichter, sich in neue Gebiete, mit denen er sich zuvor überhaupt nicht beschäftigt hat, schnell einzulesen.

5.1. Die Entdeckungsgeschichte der *Cyklooxygenase-2*

Während vor der gegenwärtigen Revolution der Informationstechnologie neue Entdeckungen den Spezialisten vorbehalten blieben, die sich Jahre und nicht selten Jahrzehnte lang mit ein und demselben Gebiet beschäftigten, ist es heute möglich geworden, dass fundamentale Entdeckungen von Wissenschaftlern gemacht werden, die ursprünglich auf einem völlig anderen Gebiet tätig waren. In Anbetracht des ständig wachsenden Wissens und der fortschreitenden Differenzierung der biochemisch-molekularbiologischen Forschung in immer mehr Spezialgebiete, erscheint diese Entwicklung auf dem ersten Blick paradox. Die Erklärung für diesen Widerspruch liegt in der Vernetzung des Wissens über Datenbanken. Diese Situation ist in der Abbildung 7 schematisch dargestellt.

Abbildung 7: *Spezialrichtungen der Biochemie / Molekularbiologie und Wechselwirkung zweier dieser Richtungen über Datenbanken bei einer Arzneimittelentwicklung*



Die Pfeile sollen die Vielzahl von Einzelgebieten der Biochemie/Molekularbiologie veranschaulichen, auf die sich einzelne Gruppen von Wissenschaftlern spezia-

lisiert haben und denen sich zum Teil spezielle Fachzeitschriften widmen. Zwei dieser Gebiete sollen beispielhaft herausgegriffen werden, die Eikosanoide und die Wachstumsfaktoren, da sie für die Entwicklung einer neuen Gruppe von entzündungshemmenden Arzneimitteln, den sog. COX-2-Hemmern, in Verbindung mit der Informationstechnologie eine entscheidende Rolle spielt.

Eikosanoide sind Oxygenierungsprodukte mehrfach ungesättigter Fettsäuren mit 20 Kohlenstoffatomen, insbesondere der Arachidonsäure. Am längsten bekannt sind die Prostaglandine, die aber nicht nur in der Samenflüssigkeit vorkommen, wo sie zuerst nachgewiesen wurden, sondern im Organismus weit verbreitet sind und dort in niedrigen Konzentrationen vielfältige biologisch-regulatorische Funktionen ausüben. Außer den Prostaglandinen sind heute eine Reihe weiterer Gruppen von biologisch aktiven Eikosanoiden bekannt, die sich hinsichtlich ihres biochemischen Bildungswegs unterscheiden. Entsprechend der Natur des Enzyms der Fettsäureoxygenierung unterscheiden wir drei Hauptgruppen von Eikosanoiden, die *Cyklooxygenaseprodukte* (COX), die *Lipoxygenaseprodukte* (LOX) und die *Cytochrom P-450-Produkte* (CYP). Die Eikosanoide haben ein besonderes Interesse der medizinischen Forschung erlangt, da sie neben ihren physiologischen Funktionen auch für die pathologischen Vorgänge bei Entzündungen, Asthma und anderen Erkrankungen eine wichtige Vermittlerrolle spielen. So stellte sich heraus, dass das seit mehr als 100 Jahren bekannte *Aspirin* (Acetylsalicylsäure) ein Hemmstoff der Cyklooxygenase und damit der Prostaglandinsynthese ist. Im Zuge der Eikosanoidforschung wurden weitere COX-hemmende Arzneimittel entwickelt (z.B. Indomethacin, Diclofenac, Naproxen, Ibuprofen, Phenylbutazon u.a.), für die der Sammelbegriff NSAID (non-steroidal antiinflammatory drugs) eingeführt wurde, da sie für den Arzt eine Alternative zu den entzündungshemmenden Steroidhormonen (Prednisolon, Dexamethazon u.a.) bildeten, die bei den Patienten unerwünschte Nebenwirkungen verursachen können. Leider stellte sich aber heraus, dass die NSAID ebenfalls Nebenwirkungen ausüben, allerdings andere als die Steroide. So führt der fortwährende Gebrauch von Aspirin nicht selten zu Magengeschwüren und zu Schädigungen der Nieren. Dies erklärt sich aus der fehlenden Selektivität der „klassischen“ COX-Hemmer, indem sie nicht nur die Entzündungsphänomene wie den Schmerz unterdrücken, sondern auch die physiologisch wichtigen regulatorischen Funktionen der Prostaglandine, zu denen u.a. der Schutz der Magenschleimhaut und die Beteiligung an der Nierenfunktion gehören. Die Eikosanoidspezialisten suchten deshalb gezielt nach den „*NSAID zweiter Generation*“, die diese Nebenwirkungen nicht ausüben – lange Zeit allerdings ohne Erfolg. Der entscheidende Anstoß für den Durchbruch kam kurioserweise aus einem völlig anderen Forschungsgebiet.

Wachstumsfaktoren sind spezifische Proteine (Polypeptide), die wie die Eikosa-noide wichtige Zellfunktionen regulieren, sich aber von den aus Fettsäuren stam-menden Eikosanoiden strukturell grundlegend unterscheiden. Neben essentiellen physiologischen Funktionen spielen bestimmte Wachstumsfaktoren auch eine Rolle beim Tumorwachstum. Daher erweckten sie verständlicherweise das beson-dere Interesse der *Krebsforscher*. Wie andere Gebiete der Biochemie bedient sich auch die Forschung über Wachstumshormone zunehmend molekularbiologischer Techniken und Vorgehensweisen. So lag es nahe, den Einfluss von Wachstums-faktoren auf die Genexpression in Tumorzellen zu untersuchen. Solche Forschun-gen wurden auch von der Arbeitsgruppe um *Harvey Herschman* von der University of California in Los Angeles betrieben. Um 1990 stellte diese Gruppe fest, dass die Behandlung bestimmter Tumorzellen mit einem solchen Wachs-tumsfaktor oder Verbindungen, die dessen Wirkung simulieren, zur Expression von ca. 10 zusätzlichen Genen führte, die in unbehandelten Zellen nicht auftrat. Mit den üblichen Methoden der Molekularbiologie wurde eines dieser Transkrip-tionsprodukte isoliert und dessen Basensequenz bestimmt. Diese wurde mit den in Datenbanken gespeicherten Sequenzen bekannter Gene verglichen. Es stellte sich heraus, dass das durch dieses Gen kodierte Protein bis dahin unbekannt war, aber eine verblüffend hohe Ähnlichkeit mit einem anderen schon bekannten Pro-tein aufwies, der Cyklooxygenase (COX). Der Arbeitsgruppenleiter gestand in ein-tem 1997 in Berlin-Steglitz gehaltenen Vortrag, dass er nach Erhalt dieser Infor-mation aus der Datenbank in ein Lehrbuch der Biochemie schauen musste, was die Cyklooxygenase eigentlich ist und welche Funktionen sie hat. Die nachfolgen-den Untersuchungen seines Labors ergaben, dass es sich bei dem neuen Protein tatsächlich um eine Cyklooxygenase handelte, die auch durch die damals bekann-ten NSAID gehemmt wird. Nach Erscheinen der ersten Publikationen über die neuen Beobachtungen (Kujubu/Fletcher/Varum/Lim/Herschman 1991²⁴; Flet-cher/Kujubu/Perrin/Herschman 1992²⁵) setzte ein wahrer Run der Eikosano-idspezialisten auf das neue Enzym ein. Quasi über Nacht war Herschman selbst ein Vorreiter der Eikosanoidforschung geworden. Man fand schnell heraus, dass diese COX-2 im Gegensatz zur länger bekannten COX-1 durch verschiedene entzündungsauslösende Substanzen induziert wird, wohingegen die entzündungs-

- 24 Kujubu, D.A. / Fletcher, B.S. / Varum, B.C. / Lim, R.W. / Herschman, H.R., TIS10, a phorbol ester tumor promotor-inducible mRNA from Swiss 3T3 cells, encodes a novel prostagladin syn-thetase/cyclooxygenase homologue. - In: Journal of Biological Chemistry 266 (1991), Heft 20. S. 12866 - 12872.
- 25 Fletcher, B.S. / Kujubu, D.A. / Perrin, D.M. / Herschman, H.R., Structure of the mitogen-inducible TIS10 gene and demonstration that the TIS10-encodet protein is a functional prosta-gladin G/H synthase. - In: Journal of Biological Chemistry 267 (1992) Heft 7, S. 4338 - 4344.

hemmenden Steroidhormone die Genexpression der COX-2 unterdrücken. Wir wissen heute, dass COX-1 und COX-2 zwar identische Reaktionen katalysieren, dabei jedoch sehr unterschiedliche biologische Funktionen erfüllen. Die COX-1 ist ein konstitutives („house-keeping“) Enzym, das in der Zelle immer vorhanden ist und für die physiologischen Funktionen der Prostaglandine zuständig ist, während die COX-2 erst bei Auslösung von Entzündungsvorgängen experimentiert wird und für die dabei ablaufenden durch Prostaglandine vermittelten *pathologischen* Vorgänge wie Schmerz, Schwellung, Rötung, Austritt von Gewebsflüssigkeit usw. verantwortlich ist. Jetzt ließen sich auch die unerwünschten Nebenwirkungen des Aspirins und anderer NSAID erklären. Aspirin hemmt nicht nur die COX-2 (das ist das eigentliche Ziel der NSAID-Therapie, ohne dass man das bis vor kurzem wusste), sondern auch die COX-1, die man lange Zeit fälschlich als das pharmakologische Zielobjekt ansah. Dadurch wird der Schutz der Magenschleimhaut durch das Eikosanoid Prostaglandin E₂ unterbunden und die Bildung von Magengeschwüren begünstigt. Trotz der großen strukturellen Ähnlichkeiten zwischen COX-1 und COX-2 gelang es pharmakologischen Forschern in relativ kurzer Zeit, selektive COX-2-Hemmer zu entwickeln, die dementsprechend frei von diesen Nebenwirkungen sind. Solche Arzneimittel werden heute von den Ärzten unter anderem zur Behandlung von Gelenkentzündungen erfolgreich eingesetzt.

5.2. Informationsentstehung - Die Problematik der unterschiedlichen Sprachspiele

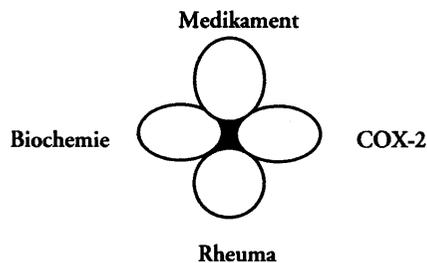
Das dargestellte Beispiel aus der jüngsten Eikosanoidforschung belegt, dass durch die Vernetzung der ständig wachsenden Informationsfülle mit Hilfe der Informationstechnologie die Wissenschaftler trotz der fortschreitenden notwendigen Subspezialisierungen enger zusammenrücken. Dieses Zusammenrücken ist sowohl regional als auch fachgebietsbezogen zu verstehen. Damit trägt die Informationstechnologie maßgeblich zur Steigerung der Interdisziplinarität des biochemisch-molekularbiologischen Forschungsprozesses bei.

Mit dem Beispiel lässt sich jedoch auch die „Zielgruppenproblematik“ verdeutlichen, die sich aus den unterschiedlichen „Sprachspielen“ im Sinne von Ludwig v. Wittgenstein ergibt.

Sucht beispielsweise ein Biochemiker nach Literatur über die therapeutische Nutzung des COX-2 (Cyclooxygenase-2), so kann er sie, wie geschildert, über verschiedene Deskriptoren finden. Die Bedeutung der Indexe und der Suchbegriffe kann allerdings in unterschiedlichen „Sprachspielen“ entstanden sein. Das soziale Umfeld in dem die Bedeutung der Begriffe entstand, die der Autor beim Schreiben des Originalartikels benutzte, wird sich in sehr vielen Fällen von dem

sozialen Umfeld unterscheiden, das die Bedeutungen der Begriffe bestimmt, die beim Indexieren unterstellt werden. Die Begriffsbedeutungen des die Literatur Suchenden werden wiederum andere sein, denn dieser lebt und arbeitet meist nochmals in einem anderen sozialen Umfeld. Erkennt man einmal an, dass die Bedeutungen der sozialen Information im sozialen Prozess gebildet werden, so wird klar, dass sie nicht für jeden gleich und auch nicht eindeutig auf Syntaxbereiche fixierbar sind.

Abbildung 8: *Venn-Diagramm zur Suche nach Dokumenten, charakterisiert durch vier unterschiedliche Indexierbegriffe*



Legende zur Abbildung 8:

In modernen bibliographischen Informationsrecherchesystemen ist es nicht ungewöhnlich, dass man viele Millionen Dokumente in einer Datenbank hat, die weltweit zur Verfügung stehen. Jede Datenbasis spezialisiert sich gewöhnlich auf einen speziellen Wissensbereich, so z.B. auf medical abstracts. Die Dokumente werden professionell erarbeitet indexiert und gespeichert um wieder von professionellen Bibliothekaren und Informationswissenschaftlern aufgefunden zu werden.

Wenn wir nun aber davon ausgehen, dass wir ein solches System nutzen wollen, ohne dass wir Biochemiker/Molekularbiologe sind oder ein Bibliothekswissenschaftler und wollen etwas über die therapeutische Nutzung des Cox-2 (Cyclooxygenase-2) erfahren, kann man die Literatur über verschiedene Deskriptoren auffinden.

Die Bedeutung der Worte, die einmal der Autor, zum anderen vom Indizierer und dann wiederum vom Suchenden bzw. vom Datenbanknutzer verwendet werden, unterscheiden sich.

Dies weist auf die Bedeutung von semantischen Rückkopplungen hin, wie sie in einem erweiterten bzw. vervielfachten Wissenstransferzyklus möglich werden.^{26, 27}

26 Fuchs-Kittowski, F / Vogel, E., Kooperative Online-Beratung im Electronic Commerce: Der COCo-Ansatz zur kooperativen Wissenserzeugung. - In: Mensch & Computer. Hrsg. v. H. Oberquelle / R. Oppermann / J. Krause. Stuttgart, Leipzig, Wiesbaden: B.G. Teubner 2001.

Dies weist darüber hinaus auf eine prinzipielle Schwierigkeit automatisierter Suchsystemen hin, die nur auf der Grundlage syntaktischer Strukturen die Wiederauffindung der Literatur realisieren können.

Es wird also wichtig, nicht nur die Informationen bzw. ihre Träger anzubieten, sondern auch zu beschreiben, für wen diese gedacht sind.

Auf der anderen Seite ist es aber auch notwendig, dass die Benutzer der Suchdienste bei einer Suchanfrage ihr persönliches Benutzerprofil mit angeben können. Dies ist aber bei den meisten Suchdiensten im Internet zur Zeit nicht der Fall. Die personenbezogenen Kriterien (z.B. zu welcher Interessengruppe gehört der Nutzer) verändern sich nur langsam und könnten daher vom Suchdienst mit berücksichtigt werden.²⁸

Sprechen wir von Informationsrecherchen auf der Grundlage spezifischer Datenbanken, wie hier von dem medizinischen Literatursuchsystem *PubMed*, so kann man vielleicht zunächst unterstellen, dass alle drei: Autor, Indexierer und Suchender, Spezialisten auf dem gleichen Fachgebiet sind, so dass das Problem der unterschiedlichen Semantik, aufgrund unterschiedlicher „Sprachspiele“, sich nicht unbedingt stellt. Gehen wir aber davon aus, dass, wie unser Beispiel zeigt, die Nutzung der modernen Informations- und Kommunikationstechnologien maßgeblich zur Erhöhung der Interdisziplinarität des biochemisch-molekularbiologischen Forschungsprozesses beiträgt, dann muss auch hier die Tatsache unterschiedlicher Sprachspiele beachtet werden.

Dies bedeutet insbesondere die Beachtung der Erkenntnis, dass der Informationsverarbeitungsansatz, wie er sich aus den Kognitionswissenschaften, insbesondere aus der bisherigen Forschung über Künstliche Intelligenz entwickelt hat, der Entstehung von Bedeutungen nicht Rechnung tragen kann, da er Information immer schon als gegeben voraussetzt²⁹. Erst mit der Theorie der Selbstorganisation, verbunden mit dem Konzept der Informationsentstehung bzw. Bildung von Bedeutungen, wie sie zuerst in der Theorie der Biologie³⁰, speziell mit der

27 Fuchs-Kittowski, K., Wissens-Ko-Produktion - Verarbeitung, Verteilung und Entstehung von Informationen in kreativ-lernenden Organisationen. - In: Stufen zur Informationsgesellschaft für alle. Festschrift zum 65. Geburtstag von Klaus Fuchs-Kittowski, Hrsg. v. Chr. Floyd / W. Hofkirchner / C. Fuchs, Frankfurt: Peter Lang Verlag 2002, im Druck.

28 Nawrocki, K., Bewertung der Informationsrecherche im Internet und Vorstellung eines neuen Ansatzes. Diplomarbeit, Berlin: Fachhochschule für Technik und Wirtschaft Berlin 2001.

29 Fuchs-Kittowski, K., Information - Neither Matter nor Mind: On the Essence and on the Evolutionary Stage Conception of Information. - In: The Quest For A Unified Theory of Information. Hrsg. v. W. Hofkirchner. World Futures General Evolution Studies. Australia, Canada, Amsterdam, Volume 13, Gordon and Breach Publishers 1994, S. 331 - 350.

Theorie der Lebensentstehung von M. Eigen entwickelt wurde, fand auch dieser Gesichtspunkt Beachtung.³¹

Sprechen wir hier von den Möglichkeiten und Grenzen der Literaturrecherche auf der Grundlage spezifischer Datenbanken, so sollte man aus der Theorie der Selbstorganisation des Lebendigen lernen, dass die hier relevanten Bedeutungen der Begriffe auch nicht vorgegeben sind, sondern in einem sozialen Prozess der Selbstorganisation entstehen.³² Dies wird dazu zwingen, dass der Mensch in den Such- und Bewertungsprozess einbezogen bleibt.

6. *Semantische Rückkopplungen als Grundlage der Selbstorganisation Dokumente verbreitender Informationssysteme*

Die Informationsrecherche ist eine immer wiederkehrende Anforderung für jeden Internetnutzer. Es ist daher erstaunlich, dass man bisher gezwungen ist mit Suchdiensten zu arbeiten, die in ihrer Konzeption in sofern veraltet sind, dass sie eine Nutzerbewertung der Dokumentenqualität nicht ermöglichen. Die Nutzer haben sich mit den bisherigen Suchdiensten abgefunden, aber kaum einer ist wirklich zufrieden mit den aufwendigen Recherchen.

Bei den gegebenen technischen Möglichkeiten des Internets, dass z.B. der einzelne Nutzer direkt mit dem Informations- bzw. Dokumentenanbieter (Mediator) in Kontakt treten können, dass sie selbst in den Leistungserstellungsprozess z.B. bei der Softwareentwicklung, bei der Erarbeitung von Dokumenten, eingreifen können und so zu Co-Produzenten ihres Wissens werden, ist es nicht hinnehmbar, dass ein Internetnutzer bisher nicht von den Erfahrungen anderer Internetnutzer lernen kann.

Wir nehmen daher ein ursprünglich von S. Brier erarbeitetes Schema, das zur Verdeutlichung der erforderlichen semantischen Rückkopplungsprozesse erarbei-

30 Fuchs-Kittowski / K. Rosenthal, H.A., Selbstorganisation, Information und Evolution - Zur Kreativität der belebten Natur. - In: Information und Selbstorganisation - Annäherung an eine vereinheitlichte Theorie der Information. Hrsg. v. N. Fenzl / W. Hofkirchner / G. Stockinger. Wien: Studien-Verlag 1998, S. 141 - 188.

31 Brier, S., What is a Possible Ontological and Epistemological Framework for a True Universal 'Information Science?' The Suggestion of a Cybersemiotics. - In: The Quest For A Unified Theory of Information. Hrsg. v. W. Hofkirchner. World Futures General Evolution Studies. Australia, Canada, Amsterdam, Volume 13, Gordon and Breach Publishers, 1994, S. 79 - 99.

32 Fuchs-Kittowski, K. / Heinrich, L.J. / Rolf, A., Information entsteht in Organisationen - in kreativen Unternehmen: Wissenschaftstheoretische und methodologische Konsequenzen für die Wirtschaftsinformatik. In: Wirtschaftsinformatik und Wissenschaftstheorie - Bestandsaufnahme und Perspektiven. Hrsg. v. J. Becker / W. König / R. Schütte / O. Wendt / S. Zelewski, Wiesbaden: Betriebswirtschaftlicher Verlag Dr. Th. Gabler GmbH, 1999, S. 329 - 361.

- von Sonderdrucken geschieht. Bei Wissens-Co-Konstruktion werden auch Produzenten-
gruppen in der scientific community gebildet.
2. verweist auf die Möglichkeit des direkten Zugriffs eines Bibliothekars oder Informations-
Broker auf eine Dokumentensammlung. Im Falle einer entsprechenden Spezialisierung z.B.
als Biochemiker/Molekularbiologe, der auch ständig die laufenden Fachtagungen besucht,
passt er in das Sprachspiel des Biochemiker/Molekularbiologen als Produzent und Nutzer.
 3. verweist auf den direkten Zugriff des Endnutzers über ein Online-System.
- Hier treten aber die beschriebenen Schwierigkeiten der meist ungleichen Verwendung der
Begriffsbedeutungen bei der Wissensorganisation auf.
 4. verweist auf die Möglichkeiten des Menschen, des Bibliothekars bzw. Informations-Brokers
als Mediator der Sammlung oder auch als Systemgestalter durch Verbesserung der Systeme
die elektronische Recherche zu unterstützen. Hier kann die globale Suche durch Suchma-
schinen: Recherche-, Selektions- und Präsentationssysteme insbesondere durch die gezielte
Spezialisierung der Daten/Wissensbasen verbessert werden.
 5. Mit dem Bewertungsabgabe-Programm führen wir eine weitere (von S. Brier in 3 und 4
wohl implizierte, aber doch ungenügend berücksichtigte), u. E. aber besonders wichtige,
semantische und syntaktische Rückkopplung ein, die Bewertung durch den kompetenten
Nutzer oder Nutzergruppen (Informationsgemeinschaft). Dies ist besonders wichtig bei der
Suche nach wissenschaftlicher Literatur. Bei der Nachfrage nach medizinischen Dokumen-
ten. Gewinnt aber angesichts des Mülls im Internet auch allgemeine Bedeutung für den nor-
malen Nutzer.

Auch die übrigen eingezeichneten gestrichelten Linien, die semantische Rückkopplungen darstel-
len, werden natürlich zur Bewertung der Qualität der Dokumente genutzt. Hier können die
Bedeutungsveränderungen im sozialen Prozess, wenn man bewusst darauf orientiert, auch
Berücksichtigung finden.

Wir wollen hier jedoch einen speziellen Kanal ausweisen, um auf eine relativ neue Entwicklung
Aufmerksam zu machen. Die Bewertung der Informationsrecherche im Internet erhält zuneh-
mend größere Aufmerksamkeit - „Web-Site-Evaluation“.

Die Interaktionsmöglichkeiten des Internets, die verstärkte Personalisierung von
Angeboten sowie die Bildung großer Communities z.B. scientific communities
zur Lösung eines speziellen Problems führt dazu, dass auch Informations-Com-
munities gebildet werden. Dies ermöglicht die Herausbildung von Diensten, die
im Unterschied zu den herkömmlichen Suchdiensten, die Informationsgewin-
nung und insbesondere die Bewertung ihrer Qualität konsequent auf den kom-
petenten Endnutzer bzw. Nutzergruppen verlagern. Im Vergleich zu den her-
kömmlichen Katalogen bekommt hier jeder Internetnutzer die Möglichkeit, die
gefundene Website zu bewerten und selbst einzelne Websites im System zu pfle-
gen. Dies erweist sich als ein wichtiger Weg zur Verbesserung der Informations-
qualität von Suchdiensten im Inter- und Intranet.

Es sei hier verwiesen z.B. auf Online-Beratung im Electronic Commerce^{33, 34}. Im Zusammenhang mit der Kundenberatung über das Internet wird immer häufiger auch gleichzeitig ermöglicht, dass Unternehmen von den Kunden lernen. Um das Potential des World Wide Web auszuschöpfen gilt es die Möglichkeiten beiderseitiger Kommunikation zu nutzen.

Meinungen der Kunden werden eingeholt, Buchrezensionen (z.B. Amazon.de). Die Kunden können die Bewertung lesen und sich so ein eigenes Urteil bilden.

Ciao, als größtes Verbraucherportal der Welt sammelt Erfahrungsberichte über Produkte und Services.

- Numerische Durchschnittsbewertung (z.B. eBaay.de)
- Durchschnittliche Kundenbewertung (Amazon, Moviefinder, eBay Feedback Profile).

Ähnlich wie bei den Textkommentare bewerten hier andere Kunden dargestellte Produkte in Form von numerischen Wertungsskalen (meist 1-5). Diese Meinungen werden zu Durchschnittswerten zusammengefasst.³⁵

Die bevorzugte Bewertungsmethode ist zur Zeit die manuelle Bewertung durch Experten, wobei die Qualifikation dieser Experten natürlich zum Hauptproblem wird. Das andere Problem ist die Unmenge an Dokumenten im Internet, die auch durch viele Experten nicht vollständig bewertet werden kann. Es muss also die zu bewertende Menge an Dokumenten von vornherein eingeschränkt werden.

Die manuelle Bewertung durch den kompetenten Benutzer oder Benutzergruppen bedeutet, dass der Benutzer die Möglichkeit hat, ein Dokument subjektiv zu bewerten. Um die subjektive Sicht eines Benutzers nicht vorherrschend werden zu lassen, muss über die Bewertung aller Benutzer ein Mittel gebildet werden. Wie von K. Nawrocki herausgearbeitet wird³⁶ besteht der Vorteil dieser Vorgehensweise darin, dass die Belastung einiger weniger Experten auf Tausende von Endnutzern verteilt wird. Ein möglicher Nachteil besteht in der Vertrauenswürdigkeit gegenüber einer Bewertung durch einen Experten. Das Problem der

33 Fuchs-Kittowski, F., Online-Beratung im Electronic Commerce - Kundenberatung über das Internet. Fraunhofer Institut für Software- und Systemtechnik ISST, <http://www.isst.fhg.de/~ffuchs/>, 2. November 2000.

34 Fuchs-Kittowski, F., Kooperative Wissenserzeugung und -nutzung in wissensintensiven Geschäftsprozessen. - In: Professionelles Wissensmanagement - Erfahrungen und Visionen. Hrsg. v. H-P. Schnurr / S. Staab / R. Studer / G. Stumme / Y. Sure. Aachen.: Shaker Verlag 2001.

35 Nawrocki, K., Bewertung der Informationsrecherche im Internet und Vorstellung eines neuen Ansatzes. Diplomarbeit, Berlin: Fachhochschule für Technik und Wirtschaft Berlin 2001.

36 ebenda

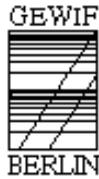
Vertrauenswürdigkeit kann durch stärkere Benutzertransparenz entschärft werden. Im Bereich der Wissenschaft, insbesondere z.B. in der Medizin werden sich Spezialistengruppen - Informationsgemeinschaften - bilden. Ist das technische Netz gegeben, wird dieses Netz der Informationszentralen bzw. Informationsgemeinschaften, das eigentliche Netz sein, welches die Grundlage einer Informationsgesellschaft für alle bilden kann. Auftretende Konflikte müssen auch hier, wie generell im wissenschaftlichen Leben, durch Erhöhung der Transparenz, verbesserte Kollegialität und erhöhte Internationalität gelöst werden.

7. Ausblick

Der Buchdruck hatte unmittelbare Auswirkungen auf die gesellschaftliche Entwicklung. Er hat zur Aufklärung beigetragen und damit die Neuzeit eingeleitet.³⁷ Die Verfügbarkeit an entsprechender Literatur für Viele war offensichtlich eine entscheidende Triebkraft. Ob aber der globale Informationszuwachs vermittels des Internets eine ähnliche Rolle spielen kann - eine neue Aufklärung einleitet, wird vor allem davon abhängig sein, ob es wirklich gelingt eine qualitative, nach humanistischen und demokratischen Werten erfolgende Bewertung der Information durchzusetzen und speziell hinsichtlich der wissenschaftlichen Information die Spaltung der Welt in „information rich“ and „information poor“ zu überwinden. Nur so kann der Weg für eine neue Aufklärung, die nun auch global erfolgen muss, vom Internet mit bereitet werden.

37 Kolditz, L., Informationsverarbeitung in der Wissenschaft - Buchdruck und Internet. - In: Sitzungsberichte Der Leibniz - Sozietät (Berlin). Jahrgang 1999, Heft 6, S. 71 - 90.

Gesellschaft für
Wissenschaftsforschung



Heinrich Parthey,
Günter Spur (Hrsg.)

**Wissenschaft
und
Innovation**

Wissenschaftsforschung
Jahrbuch 2001

Sonderdruck

Mit Beiträgen von:

*Wolfgang Biedermann • Manfred Bonitz •
Werner Ebeling • Klaus Fuchs-Kittowski •
Siegfried Greif • Christoph Grenzmann •
Horst Kant • Mathias Köbel •
Rüdiger Marquardt • Heinrich Parthey •
Andrea Scharnhorst • Tankred Schewe •
Günter Spur • Walther Umstätter*

Wissenschaftsforschung
Jahrbuch **2001**

Wissenschaft und Innovation:

Wissenschaftsforschung Jahrbuch 2001 / Heinrich Parthey; Günter Spur (Hrsg.). Mit Beiträgen von Wolfgang Biedermann ... - Berlin: Gesellschaft für Wissenschaftsforschung 2002.

Das Werk ist in allen seinen Teilen urheberrechtlich geschützt.

Jede kommerzielle Verwertung ohne schriftliche Genehmigung des Verlages ist unzulässig. Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in Systeme(n) der elektronischen Datenverarbeitung.

© Gesellschaft für Wissenschaftsforschung,
1. Auflage 2002
Alle Rechte vorbehalten.

Verlag:
Gesellschaft für Wissenschaftsforschung
c/o Prof. Dr. Walther Umstätter, Institut für
Bibliothekswissenschaft der Humboldt-Universität zu
Berlin, Dorotheenstr. 26, D-10099 Berlin

Druck: BOOKS on DEMAND GmbH,
Gutenbergring, D-22848 Norderstedt

ISBN 3-934682-35-9

Preis 15,80 €

Jahrbücher Wissenschaftsforschung

Wissenschaftsforschung: Jahrbuch 1994/95.

Hrsg. v. Hubert Laitko, Heinrich Parthey u. Jutta Petersdorf. Mit Beiträgen von Siegfried Greif, Günter Hartung, Frank Havemann, Horst Kant, Hubert Laitko, Karlheinz Lüdtke, Renate Müller, Heinrich Parthey u. Manfred Wölfling. Marburg: BdWi – Verlag 1996. 306 Seiten (ISBN 3-924684-49-6) 39,80 DM

Wissenschaftsforschung: Jahrbuch 1996/97.

Hrsg. v. Siegfried Greif, Hubert Laitko u. Heinrich Parthey. Mit Beiträgen von Siegfried Greif, Christoph Grenzmann, Claudia Hermann, Gunter Kayser, Karlheinz Lüdtke, Werner Meske, Heinrich Parthey, Roland Wagner-Döbler, Manfred Wölfling u. Regine Zott. Marburg: BdWi – Verlag 1998. 254 Seiten (ISBN 3-924684-85-5) 38,00 DM

Wissenschaft und Digitale Bibliothek: Wissenschaftsforschung Jahrbuch 1998.

Hrsg. v. Klaus Fuchs-Kittowski, Hubert Laitko, Heinrich Parthey u. Walther Umstätter. Mit Beiträgen von Manfred Bonitz, Klaus Fuchs-Kittowski, Siegfried Greif, Frank Havemann, Horst Kant, Hubert Laitko, Karlheinz Lüdtke, Heinrich Parthey, Wolfgang Stock, Walther Umstätter, Roland Wagner-Döbler, Petra Werner u. Regine Zott. Berlin: GeWif 2000. 368 Seiten. (ISBN 3-934682-30-8) 38,00 DM

Wissenschaft und Innovation: Wissenschaftsforschung Jahrbuch 1999.

Hrsg. v. Siegfried Greif u. Manfred Wölfling. Mit Beiträgen von Siegfried Greif, Christoph Grenzmann, Hans-Eduard Hauser, Frank Havemann, Gunter Kayser, Andrea Scharnhorst, Roland Wagner-Döbler, Manfred Wölfling u. Janos Wolf. Berlin: GeWif 2001. 227 Seiten. (ISBN 3-934682-33-2) 13,00 €

Organisationsinformatik und Digitale Bibliothek in der Wissenschaft: Wissenschaftsforschung Jahrbuch 2000.

Hrsg. v. Klaus Fuchs-Kittowski, Heinrich Parthey, Walther Umstätter u. Roland Wagner-Döbler. Mit Beiträgen von Manfred Bonitz, Christian Dahme, Klaus Fuchs-Kittowski, Frank Havemann, Heinrich Parthey, Andrea Scharnhorst, Walther Umstätter u. Roland Wagner-Döbler. Berlin: GeWif 2001. 239 Seiten. (ISBN 3-934682-34-0) 14,00 €

Inhaltsverzeichnisse der Jahrbücher Wissenschaftsforschung im Internet:
www.wissenschaftsforschung.de