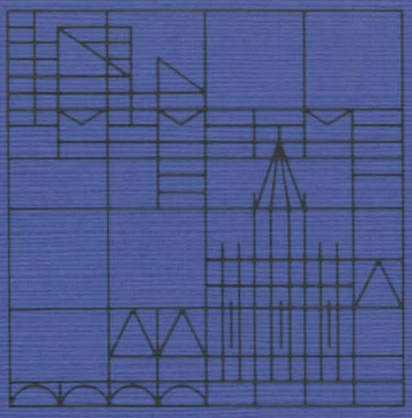


Universität Konstanz
Informationswissenschaft

Pragmatischer Mehrwert von Information.

Sprachspiele mit informationswissenschaftlichen Grundbegriffen

Rainer Kuhlen

Konstanz, Oktober 1989
Bericht 1/89

Inhalt/Content

Vorwort

Pragmatischer Mehrwert von Information.

Sprachspiele mit informationswissenschaftlichen Grundbegriffen 1

The pragmatic added-value of information.

Language games with basic concepts from information science 31

Vorwort

Dieser Text ist weitgehend identisch mit dem Anfang der "Einführung in die Informationswissenschaft", die, obgleich schon mehrfach als entsprechender Kurs im Diplom-Aufbaustudium der Informationswissenschaft gehalten, immer noch nicht als druckfertiges Manuskript vorgelegt werden kann. Dies ist sicher nicht nur mit der Arbeitsbelastung des Autors in einem immer größer werdenden Studiengang zu erklären, die kaum Freiraum für kontinuierliche Arbeit an langfristigen Vorhaben läßt, sondern ist in der systematischen Schwierigkeit angelegt, sozusagen einen Archimedischen Punkt zur Darstellung dieser komplizierten Disziplin der Informationswissenschaft zu finden. Vielleicht ist es vermessen zu behaupten, wir hätten ihn jetzt: Jedoch sieht es so aus, als ob der in diesem Text in den ersten beiden Kapiteln formulierte *pragmatische Primat* (zuerst in Kuhlen 1985) eine konzeptionelle Basis für die weiterhin heterogenen Ansätze und Arbeiten der Informationswissenschaft darstellen könnte. Zudem haben sich die im vierten Kapitel dargelegten analytischen Unterscheidungen im Begriff der Informationsarbeit – *Informationserarbeitung, Informationsaufbereitung, Informationsverarbeitung und Informationsverwaltung* – als tragfähig für eine Ausdifferenzierung einer informationswissenschaftlichen Methodenlehre erwiesen – zumindest im Kontext der Konstanzer Informationswissenschaft. Diese Methodenlehre selber gilt es natürlich in weiteren Teilen vollständiger auszumalen. Wir legen dennoch schon einmal diese Anfangsbemerkungen vor, in erster Linie zur Orientierung und zur Aufmunterung der Studierenden, das zu großen Teilen noch unbearbeitete Feld der Informationswissenschaft durch eigene Beiträge zu erschließen.

Der Text beruht in vielen, im einzelnen nicht mehr rekonstruierbaren Teilen auf mündlichen Beiträgen und Texten, die im Zusammenhang der Erarbeitung eines DFG-Schwerpunktprogramms "Information aus sprachlich repräsentiertem Wissen" formuliert worden sind. Daß dieser Antrag nicht zum Tragen gekommen ist, soll hier nicht weiter bewertet werden, scheint mir jedoch ein weiteres Beispiel für die Schwierigkeit der Informationswissenschaft zu sein, sich, ihre Fragestellungen, Methoden und Lösungsansätze den "etablierten" Wissenschaftlern verständlich zu machen. Unabhängig von dem Gelingen des Schwerpunkt-Programms wurde von den vielen am Planungsprozeß beteiligten Wissenschaftlern (darunter, neben vielen Linguisten, auch "genuine" Informationswissenschaftler wie B.Endres-Niggemeyer, N.Fuhr, D.Henning, G.Knorz, J.Krause, G.Lustig, Th.Seeger, G.Wersig, H.Zimmermann) als Gewinn angesehen, daß über den pragmatischen Primat der Informationswissenschaft grundsätzlich Einigkeit möglich war und daß damit auch das in diesem Text thematisierte Verhältnis von Wissen und Information neu und akzeptabel formuliert werden konnte. Insofern ist dieser Text vermutlich eine Mischung

aus Gedanken des Schwerpunktprogramms und den jahrelangen Bemühungen des Autors mit der "Einführung in die Informationswissenschaft". Für diese "Unterstützung" sei noch einmal allen an diesem Programm Beteiligten herzlich gedankt.

Auch sonst ist der Text an vielen Stellen der laufenden Diskussion in der Informationswissenschaft, vor allem mit Kolleginnen und Kollegen in der BRD, in den USA und dem UK, verpflichtet, ohne daß dies auch nur annähernd vollständig nachgewiesen ist. Ich denke, eine ausführlichere Referenzierung, die vielleicht in einer vollständigeren "Einführung" zu erfolgen hätte, wäre dem eher programmatischen Charakter dieser "Einführung in die Einführung" nicht angemessen.

Auf Anregung des Kollegen Jürgen Krause, Regensburg, sind größere Teile dieses Textes für ein Sonderheft der Zeitschrift "Computer and the Humanities" ins Englische übertragen worden. Es ist das Verdienst meiner Frau, Dr. E. Couper-Kuhlen, wenn sich die Diskussionen um die doch weitgehend sprachabhängigen Sprachspiele im ersten Kapitel in einem englischen Text wiederfinden können. Ich füge diesen englischsprachigen Text bei, um vielleicht auf diese Weise eine etwas größere Resonanz zu erhalten. Für alle Anregungen bin ich offen und dankbar.

Kapitel 1
Sprachspiele mit Information

Die meisten wissenschaftlichen Disziplinen verwenden Begriffe, die zu ihr grundlegend gehören, ohne daß sie in der Lage wären, diese knapp und präzise zu definieren. In der Regel empfindet man dies auch nicht als Mangel, vor allem nicht in naturwissenschaftlichen Disziplinen. Von der heutigen Physik verlangt man in einführenden Darstellungen wohl keine Definition von "Kraft", "Bewegung" oder "Energie", ebensowenig von der Biologie eine Definition von "Leben". Bei den Sozialwissenschaften scheiden sich eventuell die Geister. In der Politikwissenschaft wird man sich möglicherweise doch zu Anfang verständigen wollen, was denn unter "Staat" zu verstehen sei; auch die Soziologen scheinen einer Anfangsdiskussion über "Gesellschaft" nicht abgeneigt zu sein, wohingegen Psychologen, vor allem wenn sie stark experimentell arbeiten, im allgemeinen Diskussionen über "Geist" vermeiden. Vielleicht kann man die Aussage wagen, daß je weiter entwickelt die experimentelle Ausrichtung einer Disziplin ist, desto niedriger das Interesse an nominalistischen terminologischen Diskussionen.

Informationswissenschaft, zumindest im weiteren Umfeld des Information Retrieval, gehört zweifellos zu den experimentellen Wissenschaften[1] und entsprechend entsteht in diesem Kontext kaum ein Bedarf nach einer grundlegenden Diskussion von "Information". Ebenso denken relativ wenige Informatiker (ein neueres Beispiel ist Gitt 1989) über den Basis-Begriff "Information" nach, zumindest nicht öffentlich, sondern konstruieren einfach *Informations*systeme oder entwickeln Methoden dafür. Aber Informationswissenschaft ist nicht nur eine experimentelle Wissenschaft, sondern erhebt auch den Anspruch, Information, ihren Gebrauch, Nutzen oder ihre Wirkungen, in sozialen Umgebungen zu untersuchen. Wegen dieser Position zwischen einem eher technischen und einem eher sozialen Verständnis von "Information" hat es in den letzten 20 Jahren immer wieder intensive Diskussionen zwischen beiden Disziplinen der Informationswissenschaft und Informatik gegeben, ursprünglich sogar in der Absicht, das (exklusive) Recht der jeweiligen Wissenschaft auf die Komponente "Information" im Disziplinentitel zu reklamieren bzw. zu bestreiten. Insofern dieser "Streit" eine (wissenschafts-)politische Angelegenheit gewesen ist, ist er längst, und wohl nicht nur vorübergehend, entschieden: Informatik ist flächendeckend an Hochschulen etabliert, und zwar nicht nur institutionell, sondern, trotz laufender Erweiterungsdiskussionen, z.B. bezüglich Themen der Künstlichen Intelligenz oder einer Neuro-/Bio-Informatik, auch in der wissenschaftlichen Grundkonzeption, wohingegen die Informationswissenschaft zwar nicht mehr länger um ihre

[1] Exemplarisch für diese Richtung das jetzt auch ins Deutsche übersetzte Buch Salton/McGill (1983/87). In Deutschland vertreten eine ähnliche experimentelle Ausrichtung die Darmstädter Arbeitsgruppe unter der Leitung von G.Lustig (vgl. Lustig 1986). Auch die Konstanzer Informationswissenschaft versteht sich in ihrer Forschungsausrichtung als experimentelle Wissenschaft (vgl. Kuhlen 1986).

Existenz kämpfen muß – mit Berlin, Düsseldorf, Konstanz und Saarbrücken gibt es vier informationswissenschaftliche Einrichtungen an wissenschaftlichen Hochschulen –, aber doch recht klein geblieben ist[2] und sich noch immer unsicher über ihren wissenschaftlichen Status ist. Studenten der Informationswissenschaft, und nicht anders geht es den Informationswissenschaftlern, z.B. auf interdisziplinären Konferenzen, sehen sich immer noch dem Problem ausgesetzt, sich ihren Freunden oder Kollegen verständlich machen zu müssen, was denn Informationswissenschaft sei oder worin der Unterschied zur Informatik bestehe.

Da diese Diskussion sich offenbar nicht gänzlich von alleine erledigt, also noch immer Teil der Autonomie- bzw. Identitätsdiskussion der Informationswissenschaft ist, wollen wir uns noch einmal daran machen, das informationswissenschaftliche Verständnis von "Information" zu rekonstruieren. Natürlich gibt es vermutlich nicht *das* informationswissenschaftliche Verständnis, zu ausdifferenziert ist auch diese Disziplin inzwischen; aber da wir keine Definition anstreben (Wissenschaften werden wohl kaum über nominalistische Diskussionen konstituiert), sondern in erster Linie auf die Komplexität von Information hinweisen und multidisziplinäre Ansätze bei der wissenschaftlichen Behandlung von Information ermuntern wollen, dürfte die Chance für einen breiteren Konsens größer sein.

Mit diesem Vorhaben, ein informationswissenschaftliches Verständnis von Information herauszuarbeiten, wollen wir auf keinen Fall eine uns unproduktiv erscheinende Abgrenzungsdiskussion führen. Weltweit scheint uns die Informationswissenschaft dort zu *blühen*, wo sie keine Berührungsängste mit der Informatik hat, wohingegen sie eher ein Schattendasein fristet, wenn sie ihre Unabhängigkeit durch Abwehr technischer oder informatischer Methoden zu behaupten sucht. Aber wir wollen erneut Sensibilität dafür wecken, daß Information, jenseits ihrer informationstechnischen Behandlung, zunächst nach einer *wissenschaftlichen* Behandlung verlangt, welche den sozialen Aspekte von Information Rechnung tragen kann. Wir werden diese die *pragmatischen* Aspekte von Information nennen. Das wollen wir im folgenden versuchen, aber kaum über verbindliche Definitionen, sondern eher spielerisch, indem wir einige *Sprachspiele* (im Sinne L.Wittgensteins) mit Information spielen bzw. *Geschichten* (im Sinne des Phänomenologen W.Schapp) erzählen, in denen Information eine Rolle spielt[3].

a. Ich habe eine Information für dich.
b. Auf diese Information bin ich ganz zufällig gestoßen.
c. Das mag eine Information für x sein, aber nicht für mich.

---

[2] Dies sagt aber wohl weniger etwas über den wissenschaftstheoretischen Status oder die Bedeutung der Informationswissenschaft als über die Art der Verwissenschaftlichung moderner (westlicher) Gesellschaften aus, in denen technisch-naturwissenschaftlich und speziell informations*technisch* ausgerichteten Disziplinen eher gesellschaftliche bzw. wissenschaftspolitische Akzeptanz zukommt, die sich dann bei akademischen Disziplinen auch in Stellen und Ausstattung niederschlägt, als nicht-technischen Wissenschaften.

[3] Die Diskussion weicht gegenüber der englischen Version wegen der Sprachgebundenheit von Sprachspielen leicht ab, führt aber insgesamt zum gleichen Ergebnis.

d. Das ist keine Information, das weiß ich schon.
e. Diese Information bestätigt meinen bisherigen Eindruck.
f. Ohne weitere Information kann ich nichts machen.
g. Diese Information ist zwar ganz interessant, gehört aber nicht hierher.
h. Ich suche nichts Bestimmtes, ich informiere mich bloß.
i. Diese Information paßt mir gar nicht ins Konzept.
j. Durch diese Information wird alles Bisherige auf den Kopf gestellt.
k. Die Informationen von heute sind der Schnee von gestern.
l. Die Information ist mir 1000.- DM wert.
m. "Der Bürger wird sich in der modernen Leistungsgesellschaft dann erfolgreich behaupten können, wenn er über die Informationen verfügt, die er für die Erfüllung seiner beruflichen, persönlichen und sozialen Aufgaben benötigt"[4].

Arbeiten wir einige Eigenschaften von Information an diesen Beispielen heraus. Der Sprecher von

(a) Ich habe eine Information für dich

betrachtet das "Stück [chunk] Wissen", das er besitzt, als einschlägig für den Adressaten, an den er sich wendet, und nennt es deshalb "Information". Er würde kaum sagen: "Ich habe etwas Wissen für dich." Indem er zu verstehen gibt, daß diese Information für jemand Bestimmtes ist, weiß er offenbar einiges um die Situation und das Interesse des Adressaten, d.h. er verfügt über ein bestimmtes *Benutzermodell.* Daraus leiten wir die erste Eigenschaft ab: Information wird in der Regel **adressatenbezogen**, also gerichtet, meistens sogar in der Ausrichtung auf ein bestimmtes Individuum, übermittelt und nicht, wie in der Massenkommunikation üblich, einfach verteilt. Bei anonym verteiltem Wissen spricht die Alltagssprache eher von Nachrichten. In der Massenkommunikation, bei der in der Regel der Adressatenkreis nicht näher bestimmt ist, spricht man daher von Nachrichtensendung. Die Bezeichnung "Informationssendung" ist sicherlich auch nicht unmöglich, aber wohl weniger gebräuchlich. Zuweilen – aber dies wird immer als ungewöhnlich empfunden – werden auch über Massenkommunikationskanäle Informationen adressatenbezogen vermittelt, z.B. wenn im Verkehrsfunk eine gezielte Suchmeldung ausgegeben wird, die sich an eine bestimmte Person richtet.

Sicherlich gibt es auch Information, die nicht adressatenbezogen und gezielt gerichtet übermittelt wird, sondern einfach erworben wird und zwar ohne große

[4] Hans Matthöfer (damals Bundesminister für Forschung und Technologie) im Vorwort zum Programm der Bundesregierung zur Förderung der Information und Dokumentation (IuD-Programm) 1974–1977.

Planung. Möglicherweise ist es sogar Kennzeichen einer gut funktionierenden Informationsumgebung, daß Vorfälle, wie in

(b) Auf diese Information bin ich ganz zufällig gestoßen.

beschrieben, möglich werden. Die gegenwärtig sich herausbildende elektronische Informationskultur begünstigt diese "Zufälle" nur wenig. Information ist zunehmend stärker gerichtet. Die kreativen Nebeneffekte des "Browsing" und von "Serendipity" sind durch elektronische Speichermedien schwieriger zu erzielen. Mit "Browsing" bezeichnet man den Vorgang, daß quasi als Nebeneffekt Information aufgenommen wird, z.B. wenn man an sich gezielt nach einem Eintrag im Wörterbuch sucht und beim "Blättern" weitere interessante Information aufnimmt; wohingegen man von "Serendipity" spricht, wenn man bei der Suche nach einer bestimmten Information auf etwas ganz anderes stößt und es so interessant findet, daß das ursprüngliche Informationsproblem, zumindest vorübergehend, vergessen wird. Entsprechend scheint uns eine besondere informationswissenschaftliche Herausforderung darin zu bestehen, trotz der fortschreitenden Informatisierung[5] von Wissen und Information Verfahren und Systeme zu entwickeln, in denen kreative "Browsing-" und "Serendipity"-Effekte möglich bleiben. Der Überraschungseffekt sollte nicht ohne Not vom Informationsgeschehen entfernt werden, wenn sich dabei die Grenzen zwischen chaotischer und kreativer Informationsaufnahme leicht verwischen.

Im Sinne des Beispiels von

(c) Das mag eine Information für *x* sein, aber nicht für mich

reicht es aber noch nicht aus, daß ein "Stück Wissen" von einem Sender adressatenbezogen ausgeteilt wird. Häufig genug tritt der Fall ein, daß das Benutzermodell des Systems nicht zutreffend war. Erst wenn ein Benutzer sich in der Lage erklärt, für das gelieferte "Stück Wissen" zuständig zu sein, ist eine weitere Bedingung für die Umwandlung von Wissen in Information erfüllt. Information ist adressatenabhängig bzw. rezeptionsbedürftig. Wir wollen diese Eigenschaft von Information in der Zusammenziehung der eben gewählten Charakterisierungen mit **rezeptionsabhängig** bezeichnen. Ein einfach hingeworfenes "Stück Wissen" wird nicht zur Information, wenn es nicht aufgegriffen wird. In unserem Beispiel (c) hat der Rezipient offenbar ein Modell des vermeintlichen Informationsbedürfnisses von *x*, in welches ihm das ursprünglich angebotene "Stück Wissen" besser zu passen scheint.

[5] Unter *Informatisierung* wollen wir die der Tendenz nach vollständige Durchdringung aller gesellschaftlichen Bereiche mit Informations- und Kommunikationstechnologien verstehen; vgl. Abschnitt 4.

Auch wenn Information adressatenbezogen angeboten und offenbar auch vom richtigen Rezipienten aufgenommen wird, ist das noch keine Garantie für eine erfolgreiche Umsetzung von Wissen in Information. In der umgangssprachlichen Formulierung von

(d) Das ist keine Information, das weiß ich schon.

wird die hier anvisierte Unterscheidung von Wissen und Information recht deutlich. Das, was hier als Wissen bezeichnet wird ("das weiß ich schon"), gehört offenbar zum festen und damit verfügbaren Bestand eines Individuums. Von einer Information erwartet man hingegen, daß sie etwas Neues liefert. Der **Neuigkeitswert** ist damit eine wesentliche Eigenschaft von Information. Natürlich ist das nur eine weitere Ausprägung der oben schon erwähnten Eigenschaft der Rezeptionsabhängigkeit. Gleichzeitig wird aber damit die besondere Schwierigkeit von (wirklichen) Informationssystemen deutlich, die bislang in der Regel kaum überwunden werden kann. Zum oben schon kurz angesprochenen Benutzermodell, über das ein erfolgreicher (menschlicher oder maschineller) Informationslieferant verfügen sollte, müßte zentral auch eine Komponente "Benutzergedächtnis" gehören, damit nicht Wissen geliefert wird, das schon bekannt ist. Bei real existierenden Informationssystemen sind jedoch Komponenten wie Benutzermodelle oder Benutzergedächtnisse so gut wie gar nicht vorhanden[6]. Die Fähigkeit, das aus dem Informationssystem herauszubekommen, was für den Benutzer "neu" ist, beruht fast gänzlich auf der Geschicklichkeit des Benutzers oder seines ihn unterstützenden Informationsvermittlers, kaum auf dem "Können" des Systems. Zum Glück (für die Systementwickler) legen aber die meisten Benutzer nicht immer die schärfsten Maßstäbe an die vermittelten Informationen an, so daß man wohl nicht von der Praxis wird abweichen müssen, auch solche Systeme als Informationssysteme zu bezeichnen, die im strengeren Sinne eher als Datenverarbeitungs- oder Wissensverwaltungssysteme aufgefaßt werden sollten.

Wie man aus der Bemerkung in

(e) Diese Information bestätigt meinen bisherigen Eindruck.

sehen kann, ist auch für Information offenbar ein gewisses Ausmaß an Redundanz erträglich, häufig genug zumindest nützlich oder gar erforderlich.

[6] Probleme treten dann auf, wenn bestimmte Informationssysteme, die, wie z.B. bibliographische Online-Systeme, natürlich auf impliziten Benutzermodellen (z.B. über das Informationsverhalten von Wissenschaftlern) beruhen, für ganz andere Benutzerschichten, z.B. Manager in Industrieunternehmen, produktiv gemacht werden sollen. Rudimentäre Ansätze von "Benutzergedächtnis" bzw. von einem Gedächtnis der Dialoghistorie sind in den Möglichkeiten der meisten Online-Systeme zu sehen, z.B. schon einmal gestellte Suchfragen zu speichern oder Interessenprofile zu hinterlegen, aufgrund derer die jeweils neueren Dokumentationseinheiten auf Relevanz durchgesehen werden.

Bei einer bislang unsicheren Einschätzung mag man froh sein, wenn durch bestätigende Wiederholung die bisherige Meinung zu sichererem Wissen wird. Dies wird häufig dann eintreten, wenn etwas nicht als richtig oder falsch bewiesen werden kann, sondern auf *Erfahrungswerten* beruht. Ein zu häufiges Bestätigen des schon Gewußten oder Vermuteten wird dann allerdings als lästige Redundanz empfunden. Dann wird man wohl keinesfalls mehr von Information sprechen wollen. Dies ist sehr häufig bei schlecht konzipierten Mensch-Maschine-Schnittstellen der Fall, wenn anfangs nützliche Systemhinweise (Hilfestellungen, Erklärungen) auch noch in fortgeschrittenen Dialogsituationen wiederholt werden.

Aus dem Beispiel

(f) Ohne weitere Information kann ich nichts machen

wird deutlich, was wir später zusammenfassend den pragmatischen Charakter von Information nennen wollen. Der Begriff "pragmatisch" bezieht sich hier nicht auf die umgangssprachliche Bedeutung von 'ad hoc', sondern betont die zeichentheoretische (semiotische) Funktion der Handlungsrelevanz[7]. Ein "Stück Wissen" muß auch **handlungsrelevant** sein, um zu einer Information zu werden; d.h. es soll einen Benutzer in die Lage versetzen, etwas mit der Information zu machen. Information ist aktives, aktuell gebrauchtes Wissen. Wissen, das ich mir sozusagen auf Vorrat aneigne, z.B. wenn ich Vokabeln einer Fremdsprache auf Verdacht oder als Hausaufgabe lerne, werde ich kaum als Information bezeichnen. Dagegen führt das gezielte Blättern in einem Wörterbuch, dem ich dann das französische Wort "croissant" entnehme, zu einer Information, die ich direkt verwende, um dem Bäcker meinen Wunsch nach eben diesen "croissants" verständlich zu machen. Aus (f) kann man auch entnehmen, daß fehlende Information weiteres Handeln blockieren kann. Man kommt nicht weiter, wenn man nicht ein bestimmtes "Stück Wissen" bekommt, über das man nicht selber verfügt. Solche Situationen treten häufig bei technischen Problemen auf, bei denen man "es einfach wissen muß", um handeln zu können So kann ich keinen *"record"* in einer Programmiersprache schreiben, wenn ich die Syntax des *record*-Typs nicht kenne. Das Handbuch kann mir dann die gewünschte Information liefern. Das Syntaxfragment für sich im Handbuch war noch keine Information, sondern wird erst durch den gezielten Abruf zu einer solchen.

Ein weiteres Merkmal von Information mit Rücksicht auf die Handlungsrelevanz ist mit

(g) Diese Information ist zwar ganz interessant, gehört aber nicht hierher.

[7] Der Begriff des Handelns wird ganz im Sinne der Pragmatik weit gefaßt, d.h. es werden darunter nicht nur physische Akte subsumiert, sondern auch intellektuelle, wie z.B. "eine Entscheidung treffen".

angesprochen. Information ist nicht kontextlos, sondern **kontextabhängig**. Die Bedingung der Kontextstimmigkeit ist dann erfüllt, wenn die Informationen in bestimmte Pläne passen bzw. unter bestimmte Ziele subsumierbar sind. Informationen sind also häufig **zielgerichtet**. In der Regel sollen Informationen einmal gewählte Pläne weiter vorwärts treiben. Sieht man von dem positiven Mitnahmeeffekt des "Browsing" ab (bei dem das Verfolgen eines Plans unterbrochen, gelegentlich, im Falle von "Serendipity", auch abgebrochen wird), so werden ablenkende Informationen meistens als unwillkommen empfunden. Informationen werden in Situationen der Ungewißheit oder Unsicherheit verlangt, die man möglichst schnell durch Verfolgen eines Plans beseitigen will.

Scheinbar im Gegensatz zu dieser Eigenschaft der Zielgerichtetheit (bewußtes Verfolgen eines Plans) steht die hingeworfene Bemerkung aus

(h) Ich suche nichts Bestimmtes, ich informiere mich bloß.

meistens als Antwort auf eine als lästig empfundene Frage "Was suchen oder wünschen Sie?" gegeben. Forschte man allerdings nach, so würde sich ergeben, daß der Sprecher von (h) sehr wohl hinter einem Problem her ist, dies aber selber noch nicht so genau benennen kann oder aber noch keinen anderen an der Problemformulierung oder –lösung (aus welchen Gründen auch immer) beteiligen will. Dies ist für viele Informationssituationen typisch. Ausgangspunkt für die Bereitschaft, Informationen aufzunehmen, ist häufig ein nur sehr allgemeines Gefühl der Unsicherheit; aber weder kann man die Ursache dafür klar ausmachen (Problemidentifikation) noch kann man klar benennen, welche Informationen genau fehlen (Informationsbedarf bzw. Informationsbedürfnis). Man ist eben in einem "anamolous state of knowledge"[8]. Der angesprochene Gegensatz löst sich dann auf, wenn man an die Eigenschaft "zielgerichtet" nicht die unbedingte Anforderung stellt, daß die Zielausrichtung in jedem Moment deutlich bewußt ist. Informationen helfen zuweilen, die Ziele erst genauer herauszuarbeiten bzw. auch diese zu modifizieren.

Die beiden folgenden Aussagen

(i) Diese Information paßt mir gar nicht ins Konzept

(j) Durch diese Information wird alles Bisherige auf den Kopf gestellt

gehören zusammen. Informationen sind – wie wir gesehen haben – kontextabhängig und zielgerichtet. Man neigt häufig dazu – wie man aus bestimmten Theorien der kognitiven Psychologie (kognitive Dissonanz) weiß

[8] Vgl. z.B. Belkin et al. (1982).

–, diejenigen "Stücke von Wissen" als Informationen aufzunehmen, die den bisherigen Kontext stabil belassen bzw. welche die bislang eingeschlagenen Handlungsstrategien zur Erreichung von Zielen unterstützen. Wie man aus (j) entnehmen kann, versucht man häufig, sich gegen neues, nicht ins Konzept passendes Wissen abzuschotten. Information reduziert eben nicht nur Unsicherheit, sondern – und dieses Wissen blockiert häufig weitere Informationsarbeit – führt zuweilen auch zu neuer Verunsicherung. Die wenigstens lieben solche Situationen, wie sie durch (i) und (j) nahegelegt werden, da man auch mit Blick auf informationelle Situationen an gewisser Stabilität interessiert ist. Man erinnere sich nur an die berühmte Szene aus Brechts Galilei, in der sich die Florentiner Gelehrten weigern, durch das Fernrohr zu schauen, weil sie offenbar schon ahnen, daß das bloße Sehen sie mit solcher Information konfrontieren würde, aufgrund derer schwer erträgliche Konsequenzen fast unvermeidbar würden.

Aus diesen Abwehrbestrebungen wird im übrigen u.a. die Notwendigkeit der Werbung für Information abgeleitet. Informationen werden nicht einfach nachgefragt, auch nicht in Situationen zweifelsfreier Ungewißheit, sondern um ihren Gebrauch muß geworben werden. Information ist wohl gleichermaßen eine Hol- und Bringschuld. Allerdings macht das Werben um die Nutzung durchaus Sinn (abgesehen von dem ökonomischen Effekt, daß damit Geld verdient werden kann), denn die beiden Situationen, die Werbung überflüssig oder sinnlos machen, nämlich daß man entweder schon genau weiß, was man will, oder daß man überhaupt nichts Neues wissen will, sind wohl eher Ausnahmen. "Anamolous state of knowledge" und "Neugierde" sind Grundpositionen informationellen Verhaltens.

Die Bereitschaft, sich neuem Wissen zu öffnen, also informationsoffen zu sein, ist wahrscheinlich gleichermaßen von individuellen Dispositionen wie von kulturellen Gepflogenheiten abhängig. Gesellschaften, in denen Neugier und Fortschritt einen hohen Stellenwert haben, können wohl mit größerem Recht als Informationsgesellschaften bezeichnet werden als solche, denen die Wertschätzung dieser Kategorien eher fremd ist. Wir deuten in diesem Zusammenhang nur an, da es offenbar so etwas wie *Informationskulturen* gibt, welche die Übertragung von Modellen informationeller Absicherung, die sich in einem Kulturkreis herausgebildet und vielleicht auch bewährt haben, auf andere Kulturen höchst problematisch erscheinen läßt. Dies wird z.B. an der Diskussion um angepaßte Informationstechnologien für sich entwickelnde Länder deutlich bzw. an der spezielleren Frage, ob Online-Informationsdienste geeignete Instrumente für diese Länder sind.

Die Aussage unter

(k) Die Informationen von heute sind der Schnee von gestern.

ist mit dem bisher Gesagten verträglich. Information — genauer sollten wir jetzt sagen: die Umwandlung von Wissen in Information — ist **kontextabhängig.** Zur Kontextabhängigkeit gehört neben der Kohärenz mit Zielen und Plänen auch die *Zeitabhängigkeit.* Situationen oder Geschichten, in denen freiwillig die Tageszeitungen mit einer gewissen Zeitverzögerung gelesen werden, wirken eher als Karikaturen. Gilt diese Komik schon allgemein für anonym verteilte Nachrichten, so verwandelt sie sich direkt in Schalheit, wenn z.B. eine gestern benötigte Information heute bereitgestellt wird. Der Kontext ist nicht mehr der gleiche. Und damit ist das "Stück Wissen" keine Information mehr. Der Zeitpunkt, zu dem ich eine Information bekomme, entscheidet zudem maßgeblich über ihren Wert. Dies wird u.a. als Begründung dafür verwendet, daß elektronische Informationsdienste nicht zuletzt wegen ihrer Schnelligkeit gegenüber gedruckten oder bibliothekarischen Dienstleistungen einen solchen Mehrwert produzieren, daß eine Abschöpfung über erhobene Gebühren gerechtfertigt ist. Der Mehrwert beruht in diesem Fall also darauf, daß die Information "in time" geliefert werden kann. Dies leitet unmittelbar zur Interpretation der Aussage in (l) über:

(l) Die Information ist mir 1000.- DM wert.

Information ist zumindest in den westlichen Informationsgesellschaften zu einer Ware geworden, die auf einem international organisierten Informationsmarkt gehandelt wird. Der Warencharakter von Information gehört für viele zur Informationsgesellschaft genauso dazu, wie der allgemeine Charakter von Gütern als Ware zur Industriegesellschaft gehört hat. Es scheint allerdings einiges dafür zu sprechen, diese Aussage nicht als gegebenes Naturgesetz hinzunehmen, sondern auch zu problematisieren. Vielleicht wird es in Zukunft eher als eine gegenüber der Industriegesellschaft neue Eigenschaft der Informationsgesellschaft angesehen, daß ihr primäres "Produkt" — die Information — nicht ausschließlich und nicht primär als Ware behandelt werden muß. Der in den sechziger Jahren in der Bundesrepublik thematisierte mündige Bürger, der aber auch heute in politischen Entwürfen zur Informationsgesellschaft (z.B. Späth 1985) angesprochen ist, setzt den informierten Bürger zumindest voraus. Das war zumindest auch in dem Zitat von (m) angesprochen:

(m) "Der Bürger wird sich in der modernen Leistungsgesellschaft dann erfolgreich behaupten können, wenn er über die Informationen verfügt, die er für die Erfüllung seiner beruflichen, persönlichen und sozialen Aufgaben benötigt".

Zusammenfassung: Information – dies hat die Diskussion ergeben – ist offenbar kein frei verfügbares, objektiv definierbares "Stück Wissen", sondern muß unter Berücksichtigung vieler pragmatischer Rahmenbedingungen oder Kontingenzfaktoren, wie Zeit, Geld, soziale Umgebung, individuelle Informationsverarbeitungskapzität oder organisationelle Ziele, jeweils neu erstellt werden. Das ist das, was wir im folgenden "Transformation von Wissen in Information" nennen wollen. Information wird adressatenbezogen und damit gerichtet vermittelt, sie ist in hohem Maße rezeptionsabhängig, d.h. von der Situation und dem Kontext des Informationsnutzer bestimmt. Zur Situation gehört auch die Zeitabhängigkeit. Die Information muß für den Nutzer einen Neuigkeitswert haben und für ihn handlungsrelevant sein. Information wird meistens zielgerichtet verwendet, wobei dem Nutzer das Ziel oder die Nützlichkeit der Information zur Erreichung eines Ziels nicht immer deutlich sein muß. Information kann nicht in beliebigen Mengen aufgenommen werden. Bestimmte Schutzmechanismen blocken manchmal die Zufuhr weiterer Informationen ab, sei es um nicht erwünschte Verunsicherungen zu verhindern oder um die ebenfalls unerwünschte Situation der handlungsunfähig machenden Informationsüberflutung zu vermeiden. Überinformation kann genauso schädlich sein wie Unterinformation. Während letztere zumindest ein Gefühl der Unsicherheit erzeugt, das zu aktivem Informationshandeln führen kann, suggeriert Überinformation oft fälschlicherweise ein Gefühl der Sicherheit, schon alles Relevante zu wissen, obgleich das "Über" nicht immer das "Richtige" heißen muß. Die allgemeine Struktur westlicher Gesellschaften sowie der bei der Informationsarbeit erzielte informationelle Mehrwert legen es nahe, Information unter ökonomischen Aspekten und als Ware zu betrachten, wobei die politische Funktion von Information (als Mittel zur Erreichung von informationeller Rationalität bzw. zum Aufbau einer Informationsgesellschaft im Sinne einer informierten Gesellschaft) den Universalanspruch des Warencharakters von Information einschränken sollte.

Exkurs zum Zusammenhang von politischer und informationeller Rationalität. Wir wollen zum Schluß dieses Kapitels in einem Exkurs noch einmal auf den nur angedeuteten politischen Begriff von Information eingehen, weil auch dadurch einer anderen, nicht nur informationstechnischen Sicht Geltung verschafft wird: Hinter dem oben erwähnten Konzept der informierten Gesellschaft steht ein sehr altes Verständnis von "Information" (vgl. Kuhlen 1987b), der auf den lateinischen Begriff der "informatio" zurückgeht. Vergleichbar der politischen Praxis des freien Diskurs in der griechischen Polis sollte nach dem Cicero'schen Staatsverständnis der als "gebildeter" Bürger gelten, der in der Lage ist, über die ihn angehenden politischen und ökonomischen Geschäfte aus eigenem Wissen zu entscheiden, ohne also auf informationelle Zulieferanten angewiesen zu sein. Nur ein informationell unabhängiger Bürger könne als autonom in Sachen seiner

politischen und ökonomischen Geschäfte gelten. Nun ist klar, daß mit der modernen bürgerlichen Gesellschaft, wie sie sich mit dem Aufkommen neuer Verkehrs- und Produktionsformen im Ausgang der Renaissance entwickelt hat und welche gerade durch die zunehmende Arbeitsteilung und durch die daraus folgende Ausdifferenzierung und Spezialisierung von Wissenschaften die Bedingungen dafür geschaffen hat, daß die umfassende individuelle Aneignung von Wissen schon allein aus quantitativen Gründen unmöglich geworden ist, daß dieses Ideal der informationellen Unabhängigkeit heute nicht mehr zum Kriterium für Bildung und Autonomie gemacht werden kann.

Zu beanspruchen, alles selber zu wissen, ist heute eher Ignoranz als Bildung. Geblieben ist aber die Forderung, seine Angelegenheiten mit guten Gründen vertreten zu können. Dazu muß man nicht selber alles wissen, aber in der Lage sein – methodisch und ökonomisch – , sich dieses Wissen bei Bedarf aneignen zu können. Das Ausmaß, inwieweit eine Gesellschaft die Bedingungen der Möglichkeit dafür bereithält, ihren Bürgern "informationelle Bildung" im Prinzip zu erlauben, kann als Maßstab für eine informierte Gesellschaft angesehen werden. Akzeptiert man diesen Gedanken, so wird deutlich, daß die ausschließliche Rückführung von Information auf ihre ökonomische Funktion bzw. das Beharren auf dem Warencharakter von Information, zu politisch nicht erwünschten Konsequenzen führen muß: "Unter einer Informationsgesellschaft soll eine informierte Gesellschaft verstanden werden, also eine nicht zuletzt durch die Verfügung über Information mündige Gesellschaft, in der die einzelnen Mitglieder nicht durch Zufälle bestimmt oder fremdbestimmt werden, sondern sich im Rahmen ihrer Verantwortung für das Ganze durch gewollte Entscheidungen selbst bestimmen können und in der politische, administrative und ökonomische Handlungen (wissenschaftliche-technische ohnehin) informationell abgesichert sind bzw. sich durch Informationen auf Nachfrage legitimieren müssen und können" (Kuhlen 1987b, 339).

Mit dieser allgemeinen informationspolitischen Aussage wollen wir keinen deterministischen Zusammenhang zwischen der Menge an Information und politischer Rationalität aufstellen. Wir sind nicht der Meinung, daß privates Glück und öffentliche Wohlfahrt zwangsläufig in dem Ausmaß gesteigert werden, wie Informationen zur Verfügung stehen. Politische Entscheidungen sind nicht nur eine Funktion der besseren Informationen, sondern hängen auch von anderen Faktoren ab; aber wir wollen darauf bestehen, daß Handeln, zumal dann, wenn von den Folgen dieses Handelns andere direkt oder indirekt betroffen sind (und das ist bei politischen Handlungen wohl immer der Fall), nur dann als rational akzeptiert werden kann, wenn es informationell in dem Ausmaß abgesichert war, wie es nach dem Stand der Technik möglich und vertretbar gewesen wäre[9].

[9] Ganz zu Recht wird genau mit diesem Argument der Nachweis für die Berechtigung von Investitionen der

Wie erwähnt, der Zusammenhang zwischen mehr Information und besserer Entscheidung wird kein deterministischer sein. Aber das andere Extrem, auf informationelle Absicherung zugunsten des Gefühls, der Intuition, des Charismas oder der Gewohnheit zu verzichten, ist in demokratischen Gesellschaften, in der alle an den Entscheidungsprozessen, die sie betreffen, zumindest nachvollziehend teilhaben sollen, nicht zu akzeptieren. Die *Transparenz* politischen Handelns ist als Wertfaktor der reinen *Effizienz* sicherlich gleichzustellen. Wir wollen also politische Rationalität auf informationelle Absicherung von Seiten der Handelnden und auf Transparenz bzw. Nachvollziehbarkeit von Seiten der Betroffenen abstützen. Politische Rationalität ist weitgehend auch informationelle Rationalität (Kuhlen 1987a; Kuhlen 1989b).

Ende des Exkurses

Kapitel 2
Transformation von Wissen in Information

Im ersten Abschnitt haben wir versucht, durchgehend zwischen "Wissen" und "Information" zu unterscheiden, ohne den Wissensbegriff direkt zu thematisieren. Nach dieser Diskussion können wir das Verhältnis zwischen beiden Begriffen explizit so festlegen, daß Information die Teilmenge von Wissen ist, die von einer bestimmten Person oder einer Gruppe in einer konkreten Situation zur Lösung von Problemen[10] benötigt wird und häufig nicht explizit vorhanden ist. Zur Erläuterung dieser Aussage hilft uns zunächst wieder die Umgangssprache weiter. Die beiden folgenden Aussagen sind umgangssprachlich wohl akzeptabel (mit leichter Präferenz für die erste)

(1) Zur Lösung dieses Problems verfüge ich über einiges Wissen.
(2) Zur Lösung dieses Problems verfüge ich über einige Informationen.

wohingegen in dem Satzpaar

(3) Zur Lösung dieses Problems brauche ich noch weiteres Wissen.
(4) Zur Lösung dieses Problems brauche ich noch weitere Informationen.

eher die Aussage (4) akzeptabel ist (wenn (3) überhaupt toleriert wird). Über die allgemeine Akzeptanz hinaus wird zumindest ein "native speaker" des Deutschen "Wissen" sowohl in (1) als auch in (3) eher auf den internen

öffentlichen Verwaltungen in Informations- und Kommunikationstechnologien erbracht. Dies wurde vor allem bei der Implementierung des Information Resources Management in den Obersten Behörden der Bundesverwaltung der Vereinigten Staaten, aber auch bei der Implementierung des Landessystemkonzepts in Baden-Württemberg deutlich. Der öffentlichen Verwaltung dürfe gegenüber der Wirtschaft kein Informationsnachteil durch schlechtere Ausstattung entstehen (vgl. Marchand/Kresslein 1988).

[10] Der Begriff 'Problem' wird sowohl auf physische Handlungen als auch mentale Prozesse, z.B. Planen, Entscheiden, bezogen.

Zustand eines Subjekts beziehen, während "Informationen" eher als außerhalb dieses Subjekts befindlich angesehen werden. Über Wissen kann man z.B. durch "sich erinnern" verfügen, wobei man lediglich auf sich selber angewiesen ist, wohingegen Informationen durch Interaktion mit etwas außerhalb des Subjekts zu bekommen sind. Information ist also in einen kommunikativen Kontext eingebettet, wobei dies keine traditionelle "face to face communication" sein muß, sondern durchaus — und heute zunehmend mehr — auch in einer "face to file communication" geschehen kann. Informationen holt man nicht aus sich selbst heraus. Wenn man sich informiert, horcht man nicht in sich hinein, sondern sucht nach etwas in externen Quellen[11].

Wissen. Wir wollen die natürliche Sprache nicht überstrapazieren, sondern Festlegungen treffen: Wir wollen unter Wissen den gesicherten Bestand an Modellen über Objekte bzw. Objektbereiche und Sachverhalte verstehen, die partiell in einem Individuum (in Form seines Gedächtnisses), in einer gesellschaftlichen Gruppe, aber auch in einer Organisation, einem ganzen Kulturkreis oder in der Menschheit insgesamt als kognitive Struktur vorhanden sind. Wir haben keine Probleme damit, vom Wissen einer Organisation zu sprechen oder vom abendländischen Wissen (wohl allerdings bei dem Versuch, dieses vollständig aufzulisten)[12]. Wissen ist sozusagen der statische Bestand, die Summe der bisherigen individuellen oder kollektiven Erfahrungen oder Erkenntnisse, die mit guten Gründen in gewissem Ausmaß verallgemeinerbar sind, also nicht nur auf Meinungen beruhen.

Information. Die Eigenschaft der gewissen Verallgemeinerbarkeit kommt Information nicht zwangsläufig zu. Informationen können auch aus Meinungen entstehen. Eine Aussage kann hohen Informationswert dadurch haben, daß man weiß, welche Meinung jemand hat[13]. Um die Diskussion zu vereinfachen, bei der es in erster Linie auf den Transformationsvorgang, nicht um die Einschätzung des Wahrheitswertes der zu transformierenden Aussagen geht, sprechen wir im

---

[11] Ganz in diesem Verständnis hat Fritz Machlup, zusammen mit U.Mansfield, in einer postum herausgegebenen Arbeit Wissen und Information unterschieden: "... Information can be acquired by being told, whereas knowledge can be acquired by thinking ... Information takes at least two persons: one who tells (by speaking, writing, and one who listens, reads, watches" (Machlup/Mansfield 1983, 644).

[12] Unter dem zunehmenden Einfluß der Informationstechnik und vor allem der Künstlichen Intelligenz wird es zunehmend mehr akzeptierter Brauch, nicht nur in metaphorischem oder taktischem Sinne von *Wissen von Rechnern* in deren internen Wissensbasen zu sprechen. Die Aussage (1) von oben könnte also auch wie folgt umformuliert werden: (1') Zur Lösung dieses Problems verfügt der Rechner über einiges Wissen. Nicht ganz sicher ist, ob die Aussagen (2) und (4) ebenso akzeptabel über Rechner gemacht werden können.

[13] Es ist aber nicht zu verkennen – worauf der Konstanzer Philosoph Jürgen Mittelstraß verschiedentlich hingewiesen hat –, daß im Informationsbegriff die unsere Kulturwelt entscheidend bestimmende Unterscheidung zwischen begründbarem Wissen *(episteme)* und privater Meinung *(doxa)* verloren zu gehen droht, wenn nicht entsprechende Referenzierungsmechanismen *(traces)* bei- und offengehalten werden. Beim Entwurf von neuen Typen wissensbasierter Informationssysteme, z.B. Wissensbanken (vgl. Reimer et al. 1989), in die Wissen nur noch in durch Wissensrepräsentationssprachen strukturierter Form eingegeben wird, sollte zwingend die Herkunft der "chunks of knowledge" mitberücksichtigt werden. Dies wird dann umso problematischer, wenn Wissen nicht mehr den "Umweg" über "objektivierte" Texte nimmt, die von entsprechenden Fachleuten vor der Drucklegung kritisch begutachtet werden, sondern direkt in Wissensbanken eingegeben wird, und zwar von all denen, die nicht nur Lese-, sondern auch Schreibberechtigung haben.

folgenden weiter vereinfacht von "Wissen". Wenn man Informationen benötigt, gibt man damit zu verstehen, da man über spezielles Wissen nicht verfügt. Wohl nimmt man an, daß andere darüber verfügen. Was benötigt wird, ist ein Vermittlungsmechanismus zwischen demjenigen, der bestimmtes Wissen hat, und demjenigen, der es als Information benötigt. Dies ist genau das, was wir oben als den kommunikativen Kontext bezeichnet haben.

Damit Wissen zur Information wird, werden mindestens zwei Partner benötigt, wobei der eine Partner — wie wir schon angemerkt haben — durchaus ein "file" (eine maschinelle Datei oder auch ein "normaler" Text) sein kann. Wichtig ist uns hier, daß der eine kommunikative "Partner" unabhängig von dem anderen ist[14]. Diese Unabhängigkeit bringt es mit sich, daß der Umwandlungsprozeß von Wissen in Information auf Kommunikationskanäle angewiesen ist. Wissen muß über Medien codierbar und transportierbar sein. Diese schon von den Kommunikations- und Nachrichtentheoretikern als zentral angesehene Eigenschaft (der Medialität) ist Bedingung dafür, Wissen dergestalt zu bearbeiten, daß in unserem Verständnis daraus Information werden kann. Wissen selber ist durch semantische Eigenschaften bestimmt. Wissen bedeutet etwas und zwar ganz in dem Sinne, wie wir oben Wissen bestimmt haben, als semantisch stimmige und in der Regel auch begründbare Aussagen über Objektbereiche. Auch die berühmten Einhörner sind dann Teile von Objektbereichen, auch wenn diese nur "fiktiv" sind. Semantisch stimmige Aussagen sind auch über diese möglich.

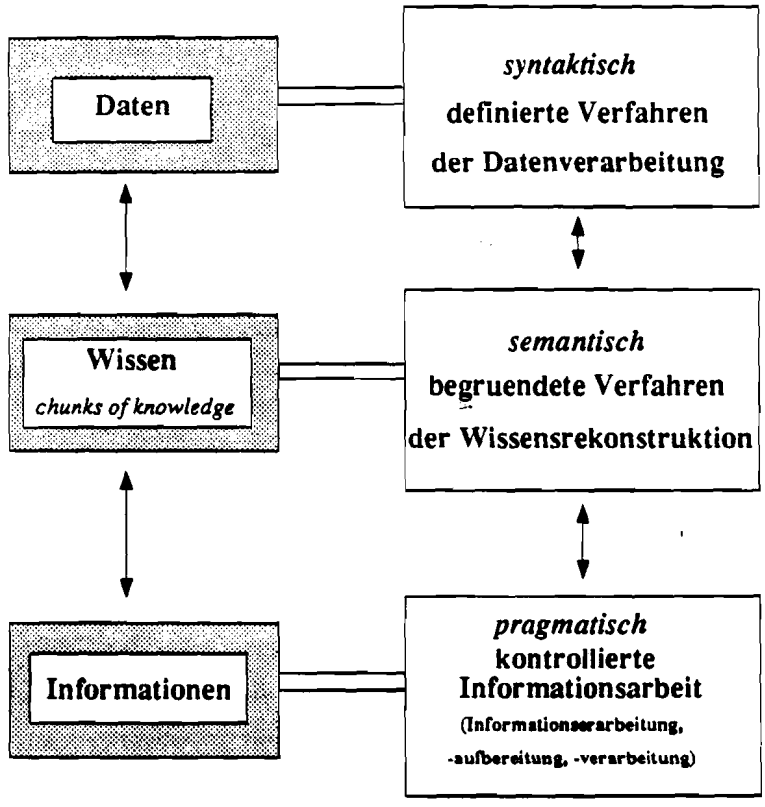
Bedingung für die Einbringung von Wissen in kommunikative Situationen ist, daß Wissen materielle Gestalt annimmt, in der Regel in der Form von Texten. Materialisiertes, nach *syntaktischen* Regeln[15] dargestelltes Wissen kann (und wird von uns) als "Daten" bezeichnet. Wir gebrauchen den Plural, da der Singular "das Datum" in seiner Bedeutung auf die temporale Referenz eingeschränkt ist. Mit gewissem Recht spricht man daher von "Datenverarbeitung" und weniger von "Wissensverarbeitung". Rechner, aber auch Menschen, manipulieren syntaktisch bestimmte Daten. Damit diese Daten (oder Teile von ihnen) nach dem Vermittlungsprozeß als Informationen handlungsrelevant werden können, müssen ihre *semantischen* Eigenschaften reinterpretierbar sein. Wir (oder Rechner)

[14] Wir wollen aus systematischen Grunden Vorgänge wie "sich erinnern" oder "nachdenken" nicht als Ausprägungen kommunikativer Situationen auffassen. Die "Interaktion" mit sich selber wird hier nicht als Kommunikation verstanden.

[15] Die Unterscheidung zwischen syntaktischen, semantischen und pragmatischen Eigenschaften hat für uns eher analytische Funktion; keineswegs soll damit eine Wissenschaftsontologie betrieben werden, nach der etwa die Informatik sich als Datenverarbeitung mit syntaktisch bestimmten Daten, die Künstliche Intelligenz und die Linguistik sich in der Form der Wissensrepräsentation eher mit den semantischen Eigenschaften von Wissen und die Informationswissenschaft sich vor allem mit den pragmatischen Eigenschaften von Wissen, also mit Informationen, zu beschäftigen haben. Ganz im Sinne der von der Semiotik reklamierten Einheit von Zeichen sollten in realen Verarbeitungsprozessen syntaktische, semantische und pragmatische Eigenschaften nicht voneinander getrennt werden. Geschieht das doch oder besteht keine Korrespondenz im Niveau der verschiedenen Ebenen, so entstehen unbefriedigende Situationen, sozusagen "pragmatische mismatches". So macht es z.B. keinen Sinn, ein pragmatisch aufwendig konstruiertes Benutzermodell über ein einfaches Volltext-Informationssystem mit entsprechend armen syntaktischen, aber vor allem primitiven semantischen Eigenschaften zu stülpen oder ein aufwendiges Instrument einer Wissensrepräsentationssprache am Ende in ein einfaches *coordinate indexing* bzw. Boole'sches Retrieval einzubetten.

müssen verstehen können, was die Daten bedeuten, damit wir (oder Rechner) sie in Problemsituationen verwenden können. Sind wir im intellektuellen Prozeß in der Lage, aus den textuellen Daten die uns wichtig erscheinenden Wissensfragmente aufzunehmen, so ist die maschinelle Verarbeitung auf eine Übertragung des Wissens in formale Sprachen, z.B. in Form von semantischen Netzen, Frame-Strukturen oder Mengen von Produktionsregeln, angewiesen.

Da der *Kontext der Handlungssituation* für die Spezifität und Qualität der Information bestimmend ist, hat Information keinen quasi objektiven Charakter, sondern variiert nach den wechselnden Anforderungen und Rahmenbedingungen. Information muß unter Berücksichtigung dieser Rahmenbedingungen erst aus Wissensstrukturen *erarbeitet* werden. Informationen können nicht wie Daten *verarbeitet* werden, sondern entstehen erst durch Anforderungen aufgrund konkreter Bedürfnisse und Problemsituationen. *Information ist* – wenn man es in eine Formel packen wollte – *Wissen in Aktion.*



Semiotischer Zusammenhang von Daten, Wissen, Information

Wir wollen diese Forderung nach Berücksichtigung der Rahmenbedingungen (vor allem die *subjektiven* Benutzerinteressen und die *objektiven* Situationserfordernisse) den *pragmatischen Primat der Informationsarbeit* nennen[16].

In dieser pragmatischen Ausrichtung sind Informationen streng genommen etwas sehr Flüchtiges. Nach ihnen wird in speziellen Situationen verlangt. Danach können sie auch wieder vergessen werden. So geschieht es auch häufig genug, z.B. wenn ich zum wiederholten Male das Rezept zum Erstellen einer delikaten Sauce durchlese, ohne zu versuchen, es mir zu merken, weil ich weiß, daß ich auf das externe Wissen in meinem Kochbuch jederzeit rekurrieren kann. Habe ich aber endlich bei einer Informationsaufnahme das Rezept behalten, so daß ich es bei der nächsten Gelegenheit nicht wieder nachschlagen muß, so habe ich etwas gelernt, habe Wissen erworben. Informationen lernt man nicht. Hat man das, was sie repräsentieren, nämlich bislang das Wissen anderer, gelernt, so sind sie zum eigenen Wissen geworden.

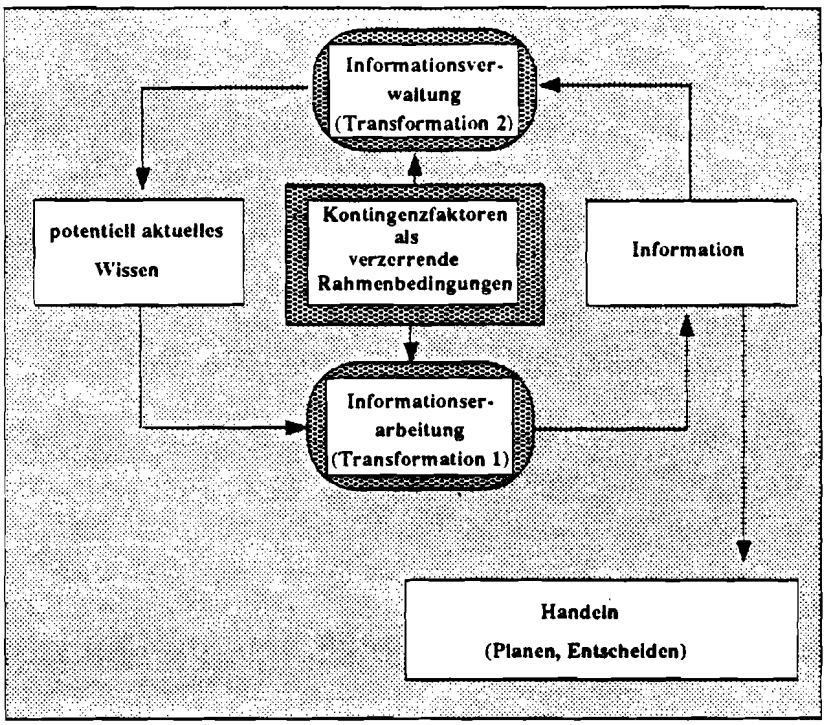
Wir können also von einem doppelten Transformationsprozeß sprechen (vgl. Abb. 2 und Kapitel 4), zum einen beim Übergang von Wissen in Information in Form der *Informationserarbeitung*, zum anderen, falls die Information nicht vergessen wird, als Einbettung der erhaltenen Information in den dauerhafteren Wissensbestand. Wir wollen diesen Übergang *Informationsverwaltung* nennen.

Information — und das ist eine bezüglich seines Warencharakters nicht unwichtige Eigenschaft — ist sozusagen "recycable", entweder durch den gleichen Benutzer in einem anderen Kontext oder auch von neuen Benutzern. Information kann durch ihren Gebrauch die jeweils aktuelle pragmatische Konstellation verlieren; die semantische Referenz bleibt erhalten und kann in einem anderen Kontext oder von einer anderen Person durchaus wieder zu einer Information werden. Auch dies bestätigt den erwähnten *flüchtigen* Charakter von Information. Wird Information als dauerhafter Bestand aufgenommen, so wird man von "Wissen" sprechen.

Die bislang getroffene Unterscheidung von Wissen und Information scheint uns systematisch sinnvoll und analytisch brauchbar zu sein. Wenn man annimmt, daß in einem Informationssystem 'Wissen' (genauer: rekonstruiertes Wissen) eingespeichert ist, dann kann dieses System von unterschiedlichen Individuen aus unterschiedlichen Gründen abgefragt werden. Ist das Ergebnis geeignet, d.h. hilft es dem Suchenden in einem konkreten Handlungszusammenhang, dann wollen wir es Information nennen. Insofern, könnte man argumentieren,

---

16 Wollte man die strengen Maßstäbe des pragmatischen Primats anlegen, also die Berücksichtigung der subjektiven Benutzerinteressen und der objektiven Situationserfordernisse zum Maßstab machen, dann müßten die meisten Systeme, die heute die Bezeichnung *Informationssystem* reklamieren, eher als weitgehend syntaktisch definierte Datenverarbeitungssysteme aufgefaßt werden. Vergleichbar sind auch Arbeiten an wissensbasierten Systemen häufig eher an der Modellierung von Wissensstrukturen interessiert als an einer informationellen Verwertung in praktischen Situationen der Fachkommunikation und sind entsprechend zwar als wichtige, aber nicht als gemeine Informationssystemleistungen einzuschätzen..



Modell der Transformation von Wissen in Information und von Information in Wissen

steht Information nicht als neuer Bestand dem Wissen gegenüber, sondern ist, im Sinne der Datenbank-Theorie, nur eine neue *Sicht.* Wir wollen jedoch den Begriff der Informationsarbeit nicht nur auf die Selektion von Teilen von Wissen aus einer bestimmten Perspektive beschränken, sondern entsprechend dem Taylor'schen Ansatz (Taylor 1986) die Vorgänge der Umwandlung von Wissen in Information als Mehrwert erzeugende Prozesse (*value adding processes*) beschreiben. In der Umwandlung wird *informationeller Mehrwert* erzeugt. Systematisch können wir zwischen indirekten Mehrwertleistungen, also solchen, die eher Hilfsfunktion haben, und direkten, die also direkt Mehrwert erzeugen, unterscheiden. Beispiele für *indirekte Leistungen bzw. Funktionen von Informationssystemen* – hier bezogen auf Themen der automatischen Sprachverarbeitung in der Informationswissenschaft – sind:

- ☐ natürlichsprachiger Zugriff auf eine Datenbank
- ☐ Möglichkeiten der akustischen Eingabe
- ☐ automatische Übersetzungsangebote auf verschiedenen Ebenen, z.B.
  - Wortebene: mehrsprachiger Thesaurus zur Erleichterung der Frageformulierung
  - Satzebene: normierte Syntax zur Erstellung von mehrsprachigen Referaten

- Textebene: Volltextübersetzungen

☐ graphische Aufbereitung der Suchergebnisse.

Beispiele für *direkte informationelle Mehrwertleistungen bzw. Funktionen von Informationssystemen* – ebenfalls aus dem Bereich der automatischen Sprachverarbeitung – sind:

☐ Automatisch erstellte flexible Textkondensate
☐ Inferenz-/Deduktionsleistungen über vorhandene rekonstruierte Wissensstrukturen
☐ Aufbereitung von Wissen nach vorgegebenen oder erschlossenen Benutzermodellen bzw. Situationsanalysen.

Kapitel 3
Sprache und Wissen – Sprache und Information

Nachdem unsere Diskussion der grundlegenden Begriff des Wissens und der Information zu der Formulierung eines pragmatischen Primats der Informationsarbeit geführt hat, wollen wir in diesem Kapitel noch kurz auf die Funktion der natürlichen Sprache für Wissen und Information eingehen, danach den bislang eher intuitiv verwendeten Begriff der Informationsarbeit systematisch diskutieren[17] und zum Schluß die gesamten Überlegungen in den Kontext der allgemeinen Informatisierung der Lebenswelt stellen.

Kann Wissen sprachlich repräsentiert werden?

Ohne Zweifel ist die natürliche Sprache die wesentliche Instanz für die Vermittlung von Wissen. Hierbei sind wir uns in Übereinstimmung mit den Diskussionen in der Künstliche-Intelligenz-Forschung bewußt, daß die natürliche Sprache bzw. genauer die aktuelle Rede oder der aktuelle Text keineswegs vollständig das zu vermittelnde Wissen darstellen kann. Die Interpretation sprachlicher Produkte hängt wesentlich von dem im Hintergrund präsenten Weltwissen ab, und die Kommunikation kann nur funktionieren, wenn dieses Wissen bei den beteiligten Partnern gemeinsam oder zumindest überlappend ist.

Um *Weltwissen* geht es seit den siebziger Jahren der Künstliche-Intelligenz-Forschung – im Gegensatz zu der bis in die sechziger Jahre dominierenden strukturalistischen Linguistik und der danach folgenden primär syntax-orientierten generativen Linguistik. Seitdem scheint es unbestritten zu

[17] Diese beiden Kapitel sind nicht in der englischen Version.

sein, daß bestimmte (wenn nicht alle) sprachlichen Phänomene – bevorzugt wird hier auf Probleme der Referenzauflösung, z.B. von Anaphern, verwiesen – von Maschinen nicht zu bearbeiten seien, wenn nicht das automatische System über Wissen über den zur Debatte stehenden Weltausschnitt verfüge. Intelligente Sprachverarbeitung – sei es für die automatische Übersetzung, für den natürlichsprachigen Zugriff zu Datenbank- oder Retrieval-Systemen, für die Wissensakquisition bei Expertensystemen bzw. Wissensbanken oder sei es für den Zweck des Textverstehens – könne – so die herrschende Meinung – nicht ohne Weltwissen auskommen. Und auch für uns kann nicht abgeschritten werden, daß eine angemessene ("richtige") Rekonstruktion des Wissens aus der sprachlichen Repräsentationsform auch von in entsprechenden formalen Repräsentationssprachen rekonstruiertem Wissen abhängig ist. Wir problematisieren das Ausmaß der Abhängigkeit. Wir halten die Rolle der natürlichen Sprache in der allgemeinen Kommunikation für so zentral, daß sie nicht nur als Vermittlungsmedium, sondern auch als Repräsentant von Wissen direkt begriffen werden sollte. Entsprechend werden Methoden benötigt, mit denen durch *sprach- bzw. textimmanente Verarbeitung* Information erarbeitet und bereitgestellt werden kann.

Durch den Rekurs auf sprachimmanente Prozesse wollen wir nicht auf den Stand der strukturalistischen, oberflächenlinguistischen oder statistischen Sprachverarbeitung der sechziger Jahre zurückfallen, aber vorschlagen, erneut zu überprüfen, inwieweit Wissen aus sprachlichen Produkten selber ableitbar ist. Es sprechen einige Argumente kognitiver Plausibilität dafür, daß neue Konzepte, neue Strukturen oder neue Regeln nicht durchgängig explizit durch Angabe von Definitionen oder Erklärungsmuster gelernt werden, sondern durch Einbettung in andere alte, schon gewußte Strukturen. Wir gehen davon aus, daß es hier erhebliche produktive Anregungen von den Konnektionismus-Modellen der Psychologie (vgl. Schneider 1987) geben wird, zumal diese, wenn auch mit erheblichem Rechneraufwand, auf große Textmengen ohne vorherige Aufbereitung angewendet werden können. In dem konnektionistischen Ansatz werden "Kenntnisse" beschrieben, die man – im Gegensatz zum fakten- und regelorientierten Ansatz – als "intuitiv" oder unbewußt bezeichnen könnte und die auf assoziativen Beziehungen, z.B. zwischen wahrgenommenen Wörtern in Texten, basieren.

Rekonstruktion sprachlich repräsentierten Wissens

Indem wir also die natürliche Sprache als eine der Formen der (materialen) Repräsentation von Wissen begreifen bzw. vorschlagen zu untersuchen, in welchem Umfang sprachliche Systeme diese Repräsentationsleistung tatsächlich übernehmen, weichen wir von dem eingebürgerten Sprachgebrauch und

der Forschungspraxis der Künstlichen Intelligenz ab (vgl. Wettler 1990). Wissensrepräsentation (als Übersetzung des amerikanischen *knowledge representation*) bezieht sich dort in der Regel auf die Darstellung des Wissens in weitgehend formal definierten Sprachen – Wissensrepräsentationssprachen, wie z.B. *frame*-Sprachen, semantische Netze oder Produktionsregelsysteme –, ohne daß zwingend der "Umweg" über die natürliche Sprache gewählt werden muß (wenn auch die Benennungen in diesen Sprachen wiederum Fragmente der natürlichen Sprache sind). Der in der "klassischen" Künstliche Intelligenz verwendete Ansatz beschreibt Wissen als Struktur von logischen Propositionen, wobei angenommen wird, daß wir zu diesem Wissen einen introspektiven Zugang haben und daß wir es weitgehend auf Formulierungen in der Prädikatenlogik zurückführen können. Entsprechend soll Wissen in formalen Modellen repräsentiert werden, das in der natürlichen Sprache immer nur unvollständig wiedergegeben werde. Insofern macht der Umkehrschluß der sprachverarbeitungsorientierten Künstlichen-Intelligenz-Forschung Sinn, daß sprachliche Phänomene (Texte) von Rechnern nur dann "verstanden" werden können, wenn diese über sprachunabhängig repräsentiertes (Welt-)Wissen verfügen.

Wir schlagen demgegenüber vor, die Ableitung von Wissensstrukturen aus sprachlichen Produkten als *Rekonstruktion von Wissen* zu bezeichnen. Wissen, das häufig in sprachlichen Strukturen repräsentiert ist, soll mit Blick auf eine weitere Verarbeitung durch Rechner – und dazu wird wesentlich die Erarbeitung von Information gehören – rekonstruiert werden. Für diese Rekonstruktion sind im Prinzip alle Wissensrepräsentationstechniken der Künstlichen Intelligenz einschlägig. Jedoch ist bei den in der Fachkommunikation anfallenden großen Textmengen der bislang auf kleinere Welten beschränkte Weltwissensansatz als generelles Modell solange nicht in der Informationspraxis adaptierbar[18], wie die Repräsentation weitgehend auf einer intellektuellen Analyse beruht und entsprechend mit den Quantitätsproblemen der Fachinformation nicht zurechtkommt. Aus informationswissenschaftlicher Sicht sind daher alle sprachimmanenten Alternativen der Wissensrekonstruktion, ebenso wie Verfahren maschinellen Lernens, zur Entlastung der kontextuellen Weltwissensmodelle interessant. Allerdings wird aufgrund des postulierten pragmatischen Primats immer außersprachliches Weltwissen, zumindest über den Bedarf bzw. das Vorwissen des Benutzers bzw. den situativen Kontext, verfügbar sein müssen. Insofern sind die in der Künstlichen-Intelligenz-Forschung entwickelten Sprachen

[18] Wir brauchen uns hier auf die grundsätzliche Diskussion nicht einzulassen, inwieweit die bislang entwickelten Wissensrepräsentationsverfahren (semantische Netze, Frame-Sprachen, Inferenz- und Produktionsregeln, Scripts u.a.) (vgl. Reimer 1989) kognitiv plausibel sind, inwieweit Wissen also über Introspektion zugänglich ist und auf formale Strukuren (Fakten, Implikationen, Regeln) abgebildet werden kann. Entsprechend der funktionalen Sicht der Informationswissenschaft zielt die entsprechende Forschung ja nicht auf ein Simulationsmodell menschlicher Intelligenz ab, vielmehr wird durchgängig ein Performanzansatz verfolgt. Die anzustrebende Leistung liegt in der Bereitstellung von Information, nicht in der Modelllierung von Wissen.

als Möglichkeit der Rekonstruktion von Wissen über Benutzer und Situationen für informationswissenschaftliche Forschung auf jeden Fall unverzichtbar.

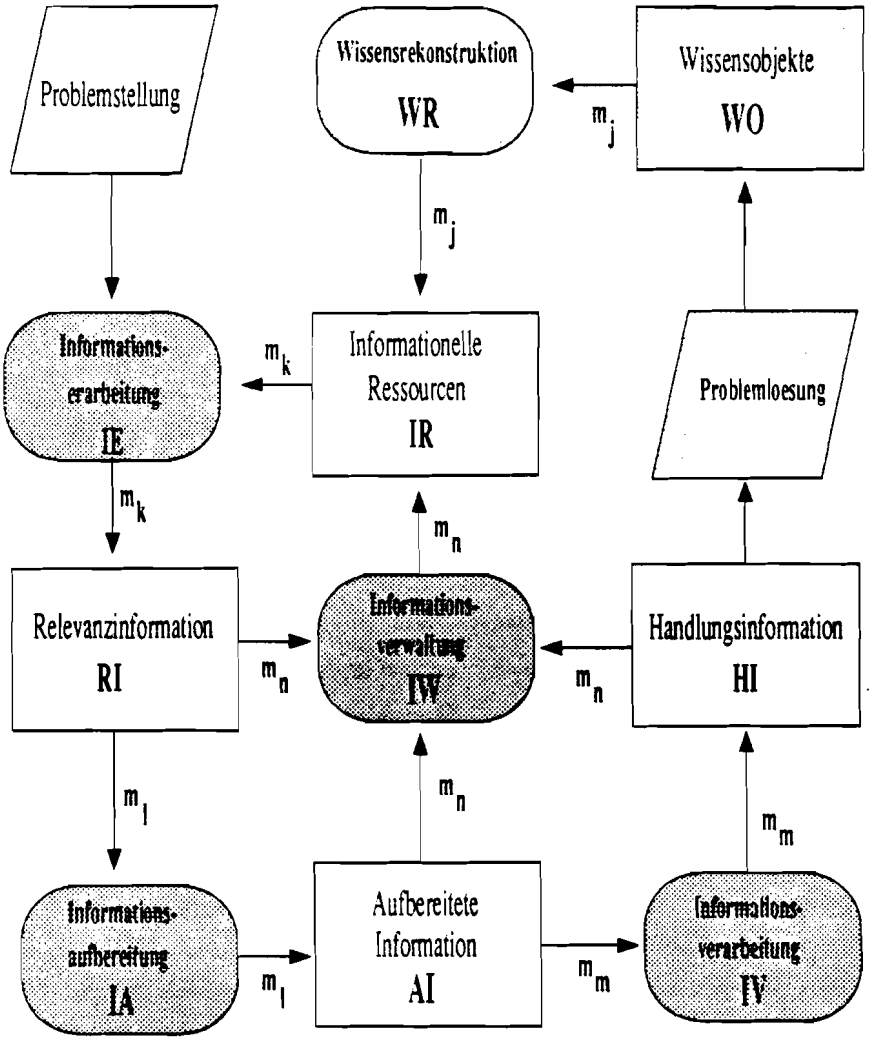
Die Ansprüche informationswissenschaftlicher Sprachverarbeitung sind gegenüber der Künstlichen Intelligenz insofern bescheidener, als es dem Benutzer von Informationen ja allemal obliegt, die in der angebotenen Information rekonstruierten Wissensfragmente in seine eigenen Wissensstrukturen einzubetten. Insofern wird es systematisch akzeptabel sein, wenn die dem Benutzer repräsentierte Information partiell dem Informationssystem "unverstanden" bleibt – wichtig ist, daß der menschliche Benutzer das Wissen aus der übermittelten Information entnehmen kann. Die tatsächliche hermeneutische Verschmelzung des neuen Wissens mit dem alten bleibt die intellektuelle Leistung des Benutzers. Wegen des hohen Eigenanteils der menschlichen Leistung ist es also verständlich, daß nicht das gesamte Weltwissen des jeweiligen Fachgebietes rekonstruiert werden muß, damit Wissen aufgenommen werden kann. Insofern scheint uns der *Rekonstruktionsanspruch* der Informationswissenschaft leichter einlösbar zu sein als der *Wissensrepräsentationsanspruch* in der Künstlichen Intelligenz.

Die nächste Frage, ob Information sprachlich dargestellt werden muß, ist einfacher und eindeutiger zu beantworten als die oben diskutierte Frage des Verhältnisses von Sprache und Wissen. Es besteht überhaupt kein Zweifel daran, daß aus Wissensstrukturen erarbeitete Information in vielfältig medialer Weise repräsentiert werden kann[19]. Unklarer ist schon wieder, welche mediale Gestalt – ob Graphik, ob Text oder bewegte Bilder mit und ohne Ton – in welcher Situation von welchem Individuum am besten rezipiert und verarbeitet wird. Sollen aber Informationssysteme aufgebaut werden, die eben dies können – nämlich die jeweils kritische Situation und die individuelle Rezeptionsfähigkeit einschätzen –, dann muß man mehr darüber wissen. Die kognitive Basis heutiger multimedialer Systeme ist zumindest noch sehr unsicher. Multimedialität ist technisch leichter möglich als rezeptionstheoretisch begründbar.

Kapitel 4
Differenzierungen im Begriff der Informationsarbeit

Wir haben in der bisherigen Diskussion verschiedentlich die Bezeichung "Informationsarbeit" verwendet, ohne deren begriffliche Eigenschaften explizit gemacht zu haben. Wesentliche Facetten davon haben wir schon bei der Darstellung des doppelten Transformationsvorgangs (Informationserarbeitung, Informationsverwaltung) angedeutet (vgl. Abb. 2). Wir wollen in diesem Kapitel genauer zwischen den Zuständen und Verfahren/Methoden der Informationsarbeit

[19] Wir verwenden hier den Ausdruck *repräsentieren* wieder genauso wie oben. Auch Information ist an sich nicht zugänglich, sondern muß in einer materialen Gestalt fixiert/repräsentiert sein, um kommunizierbar zu werden.



unterscheiden (vgl. Kuhlen 1987a), wobei die in Abb. 3 dargestellten Unterscheidungen eher analytische Funktion haben. Wenn wir zwischen "Relevanzinformation", "aufbereiteter Information" und "Handlungsinformation" unterscheiden, dann muß man sich gegenwärtig halten, daß dies alles unter dem umfassenden Begriff der Information zu subsumieren ist. Ansonsten dürfte aus den Unterscheidungen in Abb. 3 lediglich die Facette "Handlungsinformation" als Information angesehen werden, da nach der zentralen Aussage im ersten Teil – Information ist aktives, aktuell gebrauchtes Wissen, also Wissen in Aktion – nur die Handlungsinformation diese Bedingung voll erfüllt.

Bei der folgenden Diskussion gehen wir in Übereinstimmung mit dem bisher Gesagten von den beiden allgemeinen Annahmen aus, daß 1) Informationsarbeit im wesentlichen aus der Transformation von Wissen (materialisiert in Wissensobjekten) in Information besteht und 2) daß Informationsarbeit durch eine konkrete

Problemstellung provoziert wird und (hoffentlich) zu einer Problemlösung führt, aus der dann wieder neues Wissen entstehen kann ...

Bezüglich der verschiedenen *Zustände* unterscheiden wir mit den folgenden Konventionen

- **IR** die Menge der prinzipiell vorhandenen informationellen Ressourcen, in denen das potentiell einschlägige Wissen (W) der verfügbaren Wissensobjekte (WO) in einer (in der Regel für Rechner geeigneten) Form der Repräsentation gespeichert und abrufbar ist.
  IR:= (WO, m_j); $m_j \in$ WR, wobei WO Wissensobjekte sind, und WR die Menge der Verfahren der Inhaltserschließung bzw. der Wissensrekonstruktion ist (WR wird weiter unten formal definiert)[20].
- **RI** die Menge der für eine Problemlösung relevanten Informationen. RI wird aus IR durch Verfahren $m_k \in$ IE erarbeitet (IE wird weiter unten formal definiert):
  RI:= (IR, m_k) ; $m_k \in$ IE
- **AI** die Menge der aus RI aufbereiteten Informationen. AI wird aus RI durch Verfahren ml ϵ IA gewonnen (IA wird weiter unten formal definiert):
  AI:= (RI, m_l) ; $m_l \in$ IA
- **HI** die Menge der für eine Problemlösung tatsächlich verwendeten Information. HI ist also eine Teilmenge von AI und wird durch Verfahren mm ϵ IV gewonnen (IV wird weiter unten formal definiert):
  HI:= (AI, m_m) ; $m_m \in$ IV

Bezüglich der verschiedenen *Verfahren/Methoden* unterscheiden wir mit den folgenden Konventionen:

- **WR** die Menge der Wissensrekonstruktionsverfahren, mit denen die (z.B. in natürlicher Sprache repräsentierten) unterschiedlichen Wissensobjekte (z.B. Texte, Graphiken, Zahlenwerke) in einer Form aufbereitet werden, in der sie für Rechner verarbeitbar sind:
  WR:= $\{m_j \mid m_j$: WO $\rightarrow$ IR$\}$, wobei m_j Verfahren/Methoden der Rekonstruktion von Wissen aus Wissensobjekten durch Abbildungen in (WO, RI) sind.

---

20 Wie schon erwähnt, geht Wissen, einmal in Information umgewandelt, nicht verloren, sondern kann im zweiten Transformationsschritt wieder in Wissen verwandelt werden. Wir können daher im Ressourcen-Begriff zwischen *Original-Ressourcen* (IR-ORG) und *Feedback-Ressourcen* (IR-FB), die im Anschluß an eine aktuelle Nutzung durch Rücktransfomation von Information in Wissen entstehen. Unter Vorgriff der unten gegebenen Konventionen kann ersteres wie folgt definiert werden:

a) I R-ORG:= (W, m_j): $m_j \epsilon$ WR

Ressourcen durch feedback können wie folgt definiert werden:

b) IR-FB:= (RI, m_n) v (AI, m_n) v (HI, m_n); $m_n \epsilon$ IW.

- **IE** die Menge der Informationserarbeitungs-Verfahren, mit denen das in den Ressourcen IR rekonstruierte Wissen mit Rücksicht auf die Problemsituation in Information umgewandelt werden kann:
  IE:= $\{m_k \mid m_k: IR \rightarrow IR\}$, wobei m_k Verfahren der Erzeugung von Relevanzinformation durch Abbildungen in (IR, RI) sind.
- **IA** die Menge der Informationsaufbereitungs-Verfahren, mit denen die erarbeitete Relevanzinformation mit Rücksicht auf die anstehende Problemlösung aufbereitet werden kann:
  IA:= $\{m_l \mid m_l: RI \rightarrow AI\}$, wobei m_l Verfahren der Aufbereitung durch Abbildungen in (RI, AI) sind.
- **IV** die Menge der Informationsverarbeitungs-Verfahren, mit denen aus der aufbereiteten Information die Menge ausgewählt wird, die tatsächlich zur Problemlösung gebraucht wird:
  IV:= $\{m_m \mid m_m: AI \rightarrow HI\}$, wobei m_m Verfahren der Verarbeitung von Information durch Abbildungen (AI, HI) sind.
- **IW** die Menge der Verfahren zur Verwaltung von erarbeiteten, aufbereiteten oder handlungsrelevanten Informationen:
  IW:= $\{m_n \mid m_n: RI \rightarrow AI \vee AI \rightarrow IR \vee HI \rightarrow IR\}$, wobei m_n Verfahren der Speicherung durch Abbildungen in (RI, IR), (AI, IR) und (HI, IR) sind.

Die Gesamtheit der Verfahren m_j, m_k, m_l, m_m, m_n machen den methodischen Kern der Informationswissenschaft aus, wobei vor allem m_j und m_k im wesentlichen Verfahren des erweiterten Information Retrieval im Sinne der Einheit von Inhaltserschließung bzw. Wissensrekonstruktion und Retrieval/Suche) und m_n im wesentlichen Verfahren der Dateiverwaltung und der Datenbanksysteme) methodisch besser beherrscht sind als die Verfahren des m_l und des m_m. Eine Darstellung der Informationswissenschaft unter methodischen Gesichtspunkten müßte sich an dieser Differenzierung orientieren[21].

Entsprechend unserer Diskussion in Kapitel 3 sehen wir Verfahren der *Wissensrekonstruktion (m_j)* (in Umgebungen der Künstlichen Intelligenz weitgehend als Wissensrepräsentation bezeichnet) als grundlegend für alle weitere Informationsarbeit an. Die pragmatisch gesteuerte Informationsarbeit, die also die faktische Situation und die konkreten Informationsbedarfslagen berücksichtigt, hängt weitgehend von der Leistungsfähigkeit der Wissensrekonstruktion ab. In irgendeiner Form, sei es nun auf der schlichten Basis der Invertierung von Volltexten oder auf der aufwendigen Basis der Übertragung des in Wissensobjekten enthaltenen Wissens in mächtige *Wissensrepräsentationssprachen*, muß Wissen aufbereitet und mit Blick auf eine spätere Nutzung in entsprechenden

[21] Entsprechend ist auch die Konstanzer "Einführung in die Informationswissenschaft" gegliedert, für welche dieser Beitrag die systematische Einleitung ist.

informationellen Ressourcen bereitgehalten werden. Damit kein Mißverständnis entsteht, wir rechnen Formen der Wissensrekonstruktion, auch durch Techniken der Künstlichen Intelligenz, ohne Frage zum methodischen Kernbereich der Informationswissenschaft – zur Informationsarbeit im engeren Sinne gehören sie aber wohl eher nicht.

Unter *Ressourcen* wollen wir alle Formen dargestellten und rekonstruierten Wissens verstehen, die sich in irgendeiner Form zu einem Produkt oder Dienstleistung zusammenschließen, nicht nur in elektronischen Formen (Online-Banken, CD-Produkte), sondern auch in tradierten, wie Bibliographien oder Referateorgane. Ressourcen enthalten potentielle Information, nicht schon Information selber. Der Ressourcen-Begriff ist deshalb gut geeignet, weil durch ihn auf zwei wesentliche Momente verwiesen wird: zum einen Wissen bzw. rekonstruiertes Wissen als *Potential* für Handeln, zum andern als wiederverwendbare Quelle (re-source: vgl. Levitan 1982), die sich also in der einmaligen Benutzung nicht erschöpft. Auch diese letzte Eigenschaft macht es schwierig, die Ressource Information (nach unserem Verständnis handelt es sich eher um rekonstruiertes Wissen) allein unter ökonomischen Gesichtspunkten zu behandeln (vgl. Kuhlen/Finke 1988).

Ressourcen enthalten, wie gesagt, potentielle Information. Aus ihnen muß Information erarbeitet werden. Zu den Methoden der *Informationserarbeitung* (m_k) gehören alle Techniken des Information Retrieval, insofern sie sich auf Frageformulierung und Suche beziehen. Allerdings erweist es sich aufgrund des Standes der Technik noch weitgehend als schwierig, den eingangs formulierten pragmatischen Primat einzulösen. Techniken des Information Retrieval produzieren eher Dienstleistungen "von der Stange", weniger individuelle und situationsgerecht erarbeitete Informationsleistungen.

Verfahren der *Informationsaufbereitung* (m_l) beruhen bislang zum großen Teil auf intellektuell zu leistender Filterungs- und Verdichtungsarbeit. Für die Tätigkeit von Informationsvermittlern ist diese Leistung entscheidend, denn hier müssen die pragmatischen Ansprüche eingelöst werden. Für die Informationsaufbereitung werden zunehmend auch Methoden der graphikgestützten Präsentation verwendet. Erst ansatzweise kommen auch automatische Verfahren, z.B. der automatischen Textkondensierung (Kuhlen et al. 1989a; Kuhlen 1990a und b), zum Einsatz.

Die *Informationsverarbeitung* (m_m) beruht auf bislang nur sehr unzulänglich verstandenen kognitiven Prinzipien (wie sie z.B. im Rahmen von Entscheidungstheorien erarbeitet werden). Es ist symptomatisch für den noch unzureichend entwickelten Stand des Wissens in der Informationswissenschaft, daß die Vorgänge, welche die für die Informationsarbeit entscheidende Transformation von Wissen

in Information zu ihrem Ende führen, am wenigstens erforscht sind und deshalb auch bislang kaum beeinflußt werden können.

Kapitel 5
Schluß: Die Herausforderung der Informatisierung der Lebenswelt an die Informationsarbeit

In der Gegenwart wird unser Verständnis von *Wissen* und *Information* entscheidend durch die seit etwa 25 Jahren immer weiter fortschreitende *Informatisierung* bestimmt. Die Prozesse der Informatisierung werden in ihren Auswirkungen ähnlich einschneidend sein, wie in der Vergangenheit die allgemeine Industrialisierung den Lebensweltbereich und die Entwicklung der Buchdruckkunst den intellektuellen Bereich bestimmt haben. Die Informatisierung ist weitgehend Ursache dafür, daß Wissen sich zunehmend auf sehr verschiedene Weise materialisiert bzw. in sehr unterschiedlichen Formen repräsentiert wird. Waren im 18. und 19. Jahrhundert Bücher, Zeitschriften und Referateorgane die einzigen Formen der Wissensdarstellung (sieht man von eher informellen Formen wie Flugblättern, Bekanntmachungen und ähnlichen Gelegenheitsprodukten ab), so ist die heutige Produktpalette kaum mehr überschaubar:

Wissen kann heute auf sehr unterschiedliche Weise mit Hilfe elektronischer Medien verwaltet und als Produkte angeboten werden. Neben die fast schon traditionell zu nennenden Formen der *Online*-Informationsbanken, die entweder bibliographische Informationen mit und ohne *Abstracts*, mit und ohne Volltexte oder faktische Information in weitgehend numerischer Form enthalten, treten neue Formen wie elektronisch verteilte "Zeitschriften", "*image*"-Datenbanken, die Graphik und/oder bewegte Bilder verarbeiten, oder – unter dem Eindruck der neuen *compact-disque*(CD)-Technologie – multimediale Enzyklopädien, die Texte, Bilder und Töne verarbeiten und bereitstellen können. Die weitere Entwicklung deutet sich an: Revolutioniert wird der Umgang mit Wissen dann, wenn es gelingt, in elektronischen Speichern nicht nur die referentiellen Produkte von Wissen (das sind Texte, *Abstracts*, Titel, Graphiken, Tabellen etc.) bereitzuhalten, sondern Wissen(sstrukturen) selber. Dies geschieht auf kleinen Wissensgebieten seit einiger Zeit durch Expertensysteme und wird sich in Form von Wissensbanken in größerem Stil fortsetzen (vgl. Projekt **wit**; Reimer et al. 1989).

Wissen wird zwar immer noch überwiegend *face to face* übermittelt oder mit Hilfe gedruckter Medien dargestellt; zunehmend mehr sind aber – und vor allem in Bereichen, die, wie das gesamte Gebiet der Fachinformation und -kommunikation, auch unter ökonomischen Zwängen stehen – elektronische Medien dazwischengeschaltet.

Die Vorgänge der Informatisierung der Arbeitswelt und der allgemeinen Lebenswelt haben es wohl nötig gemacht, das Verhältnis von grundlegenden Begriffen wie Information und Wissen neu zu bedenken. Da war der Hauptzweck dieser Darlegungen. Informationsarbeit, ohne daß die Bezeichnung dafür gebräuchlich gewesen wäre, hat es immer schon gegeben. Die Informatisierung hat die dabei beteiligten Probleme wohl eher noch verschärft. Wir stehen der an sich paradoxen Situation gegenüber, daß Informations- und Kommunikationstechnologien, an sich dafür konzipiert, den Umgang mit Information zu erleichtern, diesen zunächst komplizieren. Verschiedentlich mag man den Eindruck gewinnen, daß neue Verfahren der Informationstechnik und -methodik nur deshalb entstehen, um Nachteile vorausgegangener Informationstechniken zu kompensieren. Die neueren Diskussionen, z.B. um *Desktop-Publishing* und *Hypertext* (vgl. Kuhlen et al. 1989b; Kuhlen 1989a), scheinen für manche diesen Verdacht, zumindest partiell zu bestätigen. Wir sollten es jedoch nicht bei diesem zweifellos berechtigten kulturkritischen Argument belassen. Vielmehr sehen wir es als eine zentrale Aufgabe der Informationswissenschaft an, die ohne Frage in den neuen Informations- und Kommunikationstechnologien enthaltenen Potentiale kreativ und kontrolliert zu nutzen – und das kann in unserem Kontext nur heißen, den Anspruch des pragmatischen Primats bei der Informationsarbeit schrittweise weiter einzulösen.

Literaturhinweise

Belkin, N.J.; Oddy, R.; Brooks, H. (1982): ASK for information retrieval. *Journal of Documentation* 38, 61–71, 145–164.

Gitt, W. (1989): Der Zweck der Information klärt nur die fünfte Ebene. *PC Woche* 39, 25.9.1989, 17–19.

Kuhlen, R. (1985): Verarbeitung von Daten, Repräsentation von Wissen, Erarbeitung von Information. Primat der Pragmatik bei informationeller Sprachverarbeitung. In: B. Endres-Niggemeyer; J. Krause (Hrsg.): *Sprachverarbeitung in Information und Dokumentation. Proceedings der Jahrestagung der Gesellschaft für Linguistische Datenverarbeitung (GLDV)*. Hannover, 5.-7. März, 1985. Springer IFB 114: Berlin etc., 1–22.

Kuhlen, R. (1986): Informationslinguistik. Theoretische, experimentelle, curriculare und prognostische Aspekte einer informationswissenschaftlichen Teildisziplin. Mit Beiträgen von U.Hahn, R.Kuhlen und U.Reimer. Niemeyer, Sprache und Information 15: Tübingen.

Kuhlen, R. (1987a): Ambivalenz fortgeschrittener informationeller Arbeitsteilung bei komplexen Verwaltungsvorgängen. In: A. Windhoff-Héretier (Hrsg.): *Verwaltung und ihre Umwelt. Festschrift für Thomas Ellwein.* Westdeutscher Verlag: Opladen, 234–257.

Kuhlen, R. (1987b): Information in der informierten Gesellschaft. *Gewerkschaftliche Monatshefte*, 6/87, 337–352.

Kuhlen (1989a): Hypertext – nur ein neues –hyper oder eine realistische Perspektive der Darstellung von Wissen und der Erarbeitung von Information? Erscheint in "Proceedings Deutscher Dokumentartag 1989" in Bremen.

Kuhlen, R. (1989b): Rahmenbedingungen der Akzeptanz für den Einsatz wiussensbasierter Verfahren am Beispiel der Sachbearbeitung einer Kreditabsicherungsbank. SFB 221/B3–4/89: Konstanz.

Kuhlen, R. (1990a): Abstracts – Abstracting – Intellektuelle und maschinelle Verfahren. Erscheint in der 3. Auflage der "Grundlagen der praktischen Information und Dokumentation" [auch als Bericht 5/88 der Informationswissenschaft Konstanz] .

Kuhlen, R. (1990b): Artikel "Information Retrieval – Verfahren des Abstracting". Erscheint in: *Computational Linguistics. Ein internationales Handbuch zur computergestützten Sprachforschung und ihrer Anwendung*. Walter de Gruyter-Verlag: Berlin.H

Kuhlen, R.; Finke, W. (1988): Informationsressourcen-Management. Informations- und Technologiepotentiale professionell für die Organisation verwerten. *Zeitschrift für Organisation* Teil 1: 5, 314–323; Teil 2: 6, 399–403.

Kuhlen, R. et al. (1989a) [R.Hammwöhner; G.Sonnenberger; U.Thiel] : TWRM-TOPOGRAPHIC: Ein wissensbasiertes System zur situationsgerechten Aufbereitung und Präsentation von Textinformation in graphischen Retrievaldialogen. *Informatik. Forschung und Entwicklung* 4, 89–107.

Kuhlen, R. et al. (1989b) [M.Böhlen; M.Diefenbach; W.Reck; H.Weber] : Hypertext – Grundlagen und Funktionen der Entlinearisierung von Text. Teil 1: Modellierung und Realisierung einer Hypertextbasis in einem Ausbildungssystem. *Nachrichten für Dokumentation* Heft 5; Teil 2: System HEIDI: Hypertext-Einführung in die Informationswissenschaft. *Nachrichten für Dokumentation* Heft 6.

Levitan, K.B. (1982): Information resources as "goods" in the life cycle of information production. *Journal of the American Society of Information Science*, Jan., 44–54.

Lustig, G. (1986): Automatische Indexierung zwischen Forschung und Anwendung. Georg Olms Verlag: Hildesheim etc.

Machlup, F. ; Mansfield, U. (1983): Semantic quirks in the study of information. In: Machlup, F.; Mansfield, U. (eds.): *The study of information.* J.Wiley: New York.

Marchand, D.A.; Kresslein, J.C. (1988): Information resources management and the public administrator. Chap. 15 in: Rabin, J.; Jackowski, E.M. (eds.): *Handbook of Information Resource Management.* New York, Basel, 395–455.

Reimer, U. (1989): FRM – Ein Frame-Repräsentationsmodell und seine formale Semantik. Springer IFB 198, Berlin etc.

Reimer et al. (1989) [Sonnenberger, G.; Kuhlen, R.; Mußgnug, K.; Senf, K. (1989): **wit:** Automatischer Aufbau von Wissensbasen durch Wissensakquisition aus Texten. Universität Konstanz, Informationswissenschaft: Konstanz.

Salton, G.; McGill, M.J. (1987): Information Retrieval – Grundlegendes für Informationswissenschaftler. McGraw-Hill: Hamburg etc. [übersetzt von W. von Keitz aus der englischen Ausgabe von 1983: Introduction to Modern Information retrieval].

Schneider, W. (1987): Connectionism: Is it a paradigm shift for psychology? *Behavior Research Methods, Instruments & Computers* 19, 2, 73–83.

Späth, L. (1985): Wende in die Zukunft. Die Bundesrepublik Deutschland auf dem Weg in die Informationsgesellschaft. Spiegel-Buch Rowohlt: Hamburg.

Taylor, R.S. (1986): Value-added processes in information systems. Alex Publishing Corp.: Norwood, NJ.

Wettler, M. (1990): Wissensrepräsentation: Typen und Modelle. Erscheint in: *Computational Linguistics. Ein internationales Handbuch zur computergestützten Sprachforschung und ihrer Anwendung.* Walter de Gruyter-Verlag: Berlin.

University of Constance
Department of Information Science

Box 5560
D-7750 Constance

The pragmatic added-value of information. Language games with basic concepts from information science

Rainer Kuhlen

Constance, October 1989

To appear in: Computer and the Humanities 1990

The pragmatic added-value of information.

Language games with basic concepts from information science

Summary. It is the purpose of this paper to reconstruct an information science understanding of information, in particular to work out the pragmatic aspects of information science by playing "language games" with the central concept of information. We distinguish the concept of information from the concept of knowledge in such a way that information is the subset of knowledge which is needed by but not available to a specific person in a concrete situation in order to solve a problem. Information can be thought of as knowledge in action. We would like to call this postulate the pragmatic primacy of information work. Information work is primarily the transformation of knowledge into information taking into account contingencies such as time, cost, social environment, cognitive ability and organizational goals. In the future we must construct information systems which take the pragmatic primacy of information work seriously. This is the only way to add value to existing knowledge and transform it into active information.

1. Language games with information

Most scientific disciplines either avoid supposedly fundamental generic concepts or use them without being able to provide appropriate definitions. This unconcerned usage is in general not considered a disadvantage, particularly not in disciplines belonging to the (natural) sciences. Contemporary introductions into physics, for example, are not expected to define precisely basic concepts such as "power", "movement", or "energy". The same is true of "life" with respect to biology. These concepts are either not used or are used without any terminological ambitions. The situation may be somewhat different in social sciences or the liberal arts. Political scientists probably need to agree on an understanding of the concept of state and, similarly, sociologists often begin theoretical discussions with an explanation of their specific understanding of "society", whereas psychologists, in particular when their research is highly experimental, tend to avoid a generic discussion about "mind". It may be that the more advanced the experimental sta-

tus of a discipline, the lower the interest in nominalistic terminological discussions.

Information science, at least with respect to the broader field of information retrieval, belongs without doubt to the *experimental sciences*, and accordingly there is no need for a basic discussion of "information". Likewise very few contemporary computer scientists (who are called "Informatiker" in German) reflect on the basic concept of information - they just build information systems. Yet information science is not only an experimental science but also a science that claims to reflect upon the use of *information in social environments*. Perhaps it is this intermediary position between the technical and the social sciences which is responsible for the fact that in the last twenty years, at least in Germany and the English-speaking world, there has been an intensive discussion concerning the relationship between computer science and information science and indeed about the right of each discipline to claim information as its fundamental concept. To the extent that this has been a "political" question of public recognition and status in the realm of academic disciplines, the battle has been won.

Computer science is a well-established and widespread academic discipline in most universities, in general as a department ("Fachbereich" or "Fakultät") of its own or in close contact with mathematics. On the other hand *information science*, although it certainly is no longer struggling for existence - there are currently four information science departments at German universities (*Berlin, Düsseldorf, Konstanz, Saarbrücken*) and in the UK and the USA it is even more widely established - is still unsure of its scientific status. Students of information science, and indeed full-blown information scientists (e.g. at interdisciplinary conferences) are still confronted with the problem of how to explain to their friends or colleagues from other academic disciplines the *difference* between computer science

("Informatik" in German) and information science, as well as how to motivate the need for an independent science of information.

Since this is still an ongoing discussion it may be helpful to *reconstruct the information science understanding of information*. This is what we propose to do in the following. In so doing we do not aim to delimitate information science from computer science (the experimental work of information science can be done as well in computer science and/or artificial intelligence environments), but to make people sensitive to the fact that information is, above and beyond its technical or computational aspects, also a social phenomenon, which needs to be treated by a scientific discipline. A technical understanding of information, even a predominantly technical one, runs the risk of suppressing the social, or as we will call it, the *pragmatic aspects of information*.

We would like to work out these pragmatic aspects of information science by playing some "*language games*" (in the Wittgensteinian sense) or telling some "*stories*" (in the understanding of the phenomenologist Wilhelm Schapp) where information plays the central role. This *ordinary language approach* is not intended to provide a definition for information or information science (sciences cannot be constituted by nominalistic discussions), but to evoke the complexity of information and to encourage a multidisciplinary approach in the scientific treatment of information.

Language games are in general language-dependent. This article was originally written in German and certain nuances of meaning may obtain only for the German language. Therefore the language games which follow and from which we would like to derive some properties of information are also given in the original German version. Nevertheless we hope that the games will be interpretable in spite of the language barriers, in particular as concerns the difference between "information" and "knowledge".

(a) I have got a piece of information for you.

Ich habe eine Information für dich.

The speaker of (a) considers the chunk of knowledge s/he has to be pertinent to the receiver and therefore calls it "information". It is quite unlikely that the speaker would say: I have some knowledge for you. The information is being related explicitly to a specific person ("for you") and this obviously depends on the speaker's knowledge of the addressee's situation and interests. The speaker has a certain model of the person whom the information is being related to, also called a *user model*. This leads us to the first property of information. Information is normally **addressee-related**, mostly directed to a single individual and is not simply 'broadcast' or distributed as is the case in mass communication. Anonymously distributed knowledge tends to be referred to as 'news' in everyday language. Occasionally, mass communication channels are also used for direct transfer of information, for example, when a search message for some single individual is distributed via radio or television, but this usually has a slightly alienating effect. In this case the news will be interpreted as information too.

There is of course also information which is not transferred in a directed and addressee-related fashion but is simply picked up:

(b) I picked up this bit of information by accident.

Auf diese Information bin ich ganz zufällig gestoßen.

For the description of this phenomenon it is useful to distinguish between the notions of *browsing* and *serendipidity*. Everybody knows that while searching for certain information, for example looking up the meaning of a

word in a dictionary, one may incidentally pick up "side-effect information". This is called *browsing*. Sometimes one becomes so fascinated by this side information that the original goal is forgotten; this effect is then called *serendipity*. Both in the browsing and the serendipity case the borderline between creative and chaotic information input is of course fuzzy. But it may be characteristic of a good information environment that browsing and serendipity effects are possible. Lack of them is one of the drawbacks in contemporary electronic information culture, which is in the main directed and does not leave information seeking to chance. It is a challenge for information science research to develop systems or methods which can preserve the creative side-effects of browsing and serendipity in the electronic age.

Not only must a bit of information be addressee-related or directed towards a certain person; in addition, the information must be grasped by that person. Information depends on actual reception by the addressee; and even this may not be enough. In fact, it may happen (as in c) that the speaker's user-model was not appropriate and therefore, although the information was directed, it was received but not interpreted as information.

(c) This is not information, I already know it.

Das ist keine Information, das weiß ich schon.

In the everyday formulation of (c), the distinction between knowledge and information which we have in mind comes to light. What is called knowledge here ("I already know it") is a fixed and retrievable part of the intellectual resources of an individual, whereas we expect information to be "new". **Novelty** is an important feature of information. And novelty, of course, depends upon the recipient's intellectual state; information thus, as we know it, is **recipient-dependent.**

This is one of the main drawbacks of current commercial (on-line) information systems; they have been designed for a more or less anonymous market. Information systems in general are neither provided with specific user-models nor do they have a component which could be called user memory. The lack of user-models is responsible for inappropriate "information". The "information" delivered is not tailored to a special user's interests. This is one of the reasons why on-line data bases originally designed for use in scientific and technical environments are not well received by other user groups, for example in management. Furthermore the lack of user-memories often makes people very angry when they get the same information over and over again. In everyday conversation we simply take it for granted that our interlocutor is keeping in mind what has already been said. Why not in "intelligent" information systems communication?

Current information systems have only very weak pragmatic components at their disposal. They are extremely limited in their capacity to consider the consequences of their actions. The relevance of the system's output depends almost completely on the skill of the human end-user (or of the information transfer specialist who is assisting the end-user). Therefore, based on our discussion so far, current commercial information systems do not deserve the name "information" because they are not able to reflect a user's needs; they don't have a pragmatic component. In the strict sense, they should be called *data or knowledge administration systems*. Of course we only wish to reflect on this problem here, and are not advocating seriously a change in current terminology or a rigorous restriction of the use of "information systems". The term "information system" may be appropriate even for contemporary systems if we include the pragmatic capacity of the end-user in the general system design.

(d) This information confirms my present impression.

Diese Information bestätigt meinen bisherigen Eindruck.

The feature of newness or novelty has been emphasized as important for information. But a certain amount of *redundancy* is sometimes useful and, for the purpose of effective learning, even necessary. Example (d) makes it clear that it is sometimes helpful if an already existing opinion is confirmed by the same information (in particular if it stems from a different source). In social environments, where the truth value of a proposition cannot be proved but depends on *experience or agreement*, the repetition of a piece of information is necessary (and often not sufficient) for the transformation of an opinion into knowledge. For the design of information systems it is extremely difficult to manage the redundancy property of information. For example, systems with poorly designed man-machine interfaces tend to repeat help-messages again and again, which is useful at the beginning but cumbersome for advanced users. Here again, the degree of reduncancy depends on appropriate information about the status of the user.

(e) Without any further information I can do nothing.

Ohne weitere Information kann ich nichts machen.

Example (e) makes explicit what we have called the *pragmatic aspect of information*. A chunk of knowledge needs to be relevant for a specific action, be it a physical or an intellectual action. Information must enable a person to do something. *Information is active knowledge needed at this very moment.* Knowledge which one "buys in stock", for example by trying to memorize a vocabulary list as part of one's daily homework, is hardly called information But looking up the French word "croissant" in a dictionary in order to explain one's wish in a bakery is information which one does not have but needs. Lack of information, as can be seen from (e), may block further action. This often occurs in technical environments, where one simply needs to know in order to do. One cannot write a "record" in a specific programming language if one does not know the syntax of the record-

type. A hand-book may provide the needed information. The corresponding chapter in the hand-book cannot be called information per se, only the current recall of it makes it a piece of information.

(f) This information is clearly interesting but is out of place.

Diese Information ist zwar ganz interessant, gehört aber nicht hierher.

(g) This information does not suit my plans.

Diese Information paßt mir gar nicht ins Konzept.

(h) This information turns everything upside down.

Durch diese Information wird alles Bisherige auf den Kopf gestellt.

In (f) a further feature of information, likewise emphasizing its relevance for action, is manifest. Information is not context-free but **context-dependent**. Context is often defined in terms of plans or *goals*. The condition of context-dependency is thus fulfilled if the information fits into specific plans or is appropriate to certain goals. Information is in general **plan-dependent or goal-related**. Normally information is supposed to promote plans which already exist; diverging information, even if potentially relevant, will often be rejected or found troublesome if it doesn't fit into existing plans (cf. sentence (g)).

Few people can deal with situations in which a sentence like (h) would be appropriate, because as human beings we tend to prefer informational stability. One good example of this propensity is the famous scene in Brecht's *Galilei* where the scholars in Florence refuse to look through

Galilei's telescope because they are afraid that the new information they would gain in doing so will turn their world upside down.

The willingness to be open to new knowledge, to accept it as potentially important information, is undoubtedly just as dependent on cultural custom as on personal disposition. Societies in which concepts like progress and curiosity are high on the scale of acceptance are more likely to be called *information societies* than societies in which these concepts do not carry such positive connotations. We mention this because of its obvious relevance to the problem of transferring models of information-related behaviour from one society to another.

(i) I'm not looking for anything particular, I only want to inform myself.

Ich suche nichts Bestimmtes, ich informiere mich bloß.

Information is demanded in situations of *uncertainty*, where the wish is normally to be rid of this uncertainty as soon as possible. The question whether a piece of information fits into plans or is appropriate to certain goals is one of the most difficult in the theory of human information processing. Before discussing this problem, let us briefly interpret sentence (i), which at first glance appears to be in conflict with the goal-relatedness which has been derived from (f).

Sentence (i) is often used as a means of getting rid of a questioner, perhaps a librarian, who seeing a user walking around wants to be helpful ("what are you looking for?") but is more often than not merely bothersome. If one were to analyze this situation somewhat more carefully, it would probably turn out that the speaker of (i), in spite of claims to the contrary, really does have a problem but either does not know how to explain it, only has a vague feeling as to what it might be, or does not want

somebody else to participate in solving it. The first alternative - that the user does not know what s/he really wants - is typical of many information situations.

The starting point for seeking information or for being willing to receive information is very often a *general and vague feeling of uncertainty.* But neither the cause for this uncertainty (problem identification) nor the means of how to get rid of this uncertainty (concrete demands or information wishes) may be clear. The user is - as some information scientists have put it - in an *anomalous state of knowledge.*

The conflict of (i) with the interpretation of sentence (f) can be resolved if the criterion of goal-relatedness is not applied too rigorously. We don't need to be fully aware of goals at the outset. Information is helpful in modifying existing goals, making them more precise or more concrete, and even formulating them from scratch.

This amounts to an important modification of our interpretation of information so far. Information may not be related only to articulated needs. Observations from cognitive psychology (e.g. the theory of cognitive dissonance) that users tend to refuse information when it does not fit into their plans (cf. sentence (g)) lead to the conclusion that information systems should be designed to strike a balance between helping the user satisfy articulated demands and supplying information which is "objectively" necessary for the situation - whatever "objectively" necessary infomation may mean. We would like to point out that powerful information systems should be *supportive systems* in the sense of being able to compensate for insufficient human information processing behaviour. This insufficiency may lie in slow processing, in limited memory, or in the propensity to ignore relevant information or commit oneself too early to a ostensibly appealing solution.

The difficulty in knowing what kind of information is needed and the human tendency to neglect relevant information or to shun more information, even in critical situations, can be used as a means of justifying the need to advertise information. If one always knows what kind of information is needed, then advertising is unnecessary. If there is no chance of convincing people to open themselves to new information, then advertising makes no sense. The current status with respect to information receptiveness is clearly somewhere in between these two extreme positions.

(j) What is information today is yesterday's old hat.

Die Information von heute ist der Schnee von gestern.

The message in sentence (j) is compatible with our interpretation so far. Information - or as we prefer to say: the transformation of knowledge into information - is **context-dependent.** Context encompasses, in addition to plans or goals, a temporal frame which is dependent on 'now'. Tales of people voluntarily reading their newspapers with a delay of weeks or months give the impression of being caricatures. This is not only true for anonymously distributed mass communication news. Even today's information will seem stale if it is what one urgently needed yesterday. The advantage of modern on-line bibliographic information systems, in comparison to traditional library catalogues, lies mainly in their speed. The ability to deliver information "in time" is the main *value-added effect* of electronic systems and can thus be used as a justification for royalty claims. This brings us to sentence (k).

(k) I am willing to pay $ 500 for this information.

Diese Information ist mir 1000 DM wert.

(1) In a competitive society only those who have access to the information they need for their professional, social or personal commitments will be successful. (Hans Matthöfer, former minister for research and technology, in his preface to the Federal Programme for the Promotion of Information and Documentation, 1974-1977)

"Der Bürger wird sich in der modernen Leistungsgesellschaft dann erfolgreich behaupten können, wenn er über die Information verfügt, die er für die Erfüllung seiner beruflichen, persönlichen und sozialen Aufgaben benötigt."

Information, at least in modern western societies, has become a **commodity** which is distributed on an international information market. The merchandise character of information is considered by many economists to be the primary indicator of a so-called 'information society'. Sentence (1) suggests furthermore that access to and usage of information is the basic requirement for active participation by responsible citizens ("*mündige Bürger*") in all public affairs. Therefore information must be regarded not only as commodity but also as a *public good*.

In sum, information cannot be considered as a freely available, objectively definable piece of knowledge, but must be created or produced with regard to many contingencies including time, cost, social environment, cognitive ability and organizational goals.

Information is without a doubt based on knowledge, but depends in addition on the specific situation in which it is used. Information is addressee-dependent, determined by the situation and the context of the user. Information is judged by its novelty and needs to be relevant for a specific action. Information is in general goal-related or plan-oriented, whereby the user need not be fully aware of the goal.

Information is sometimes also needed to shape plans or goals.

Information cannot be absorbed in unlimited quantity. Over-information is just as harmful as under-information. While under-information creates a feeling of uncertainty, which a person will usually try to overcome by seeking additional information, over-information sometimes deceives a person into thinking they have more than enough information, although the "over" may not always mean the right information. In such a situation a piece of relevant information may not be recognized as pertinent and may be rejected. This can even happen in critical situations, where a person may prefer to stick to a first solution and try to avoid further information seeking effort.

Information is mainly considered an economical factor in modern societies but must also be thought of as a political factor and as an important prerequisite for responsible participation in public life.

2. Transformation of knowledge into information

In the preceding section we have tried to distinguish the concept of information from the concept of knowledge in such a way that *information is the subset of knowledge which is needed by but is not available to a specific person in a concrete situation in order to solve a problem*. In the following, we would like to clarify this still somewhat abstract formulation, by referring to everyday language again. The following two sentences are probably both acceptable - at least in German - , with a slight preference for the first one:

(m) I have some knowledge to solve this problem.

Zur Lösung dieses Problemes verfüge ich über einiges Wissen.

(n) I have some information to solve this problem .

Zur Lösung dieses Problems verfüge ich über einige Informationen.

whereas of the two sentences

(o) To solve this problem I need additional knowledge.

Zur Lösung dieses Problems brauche ich noch weiteres Wissen.

(p) To solve this problem I need additional information.

Zur Lösung dieses Problems brauche ich noch weitere Informationen.

only (p) seems to be acceptable. A native speaker of German would probably refer *knowledge* in (m) and (o) to an *internal mental state* of the subject, whereas *information* (in p) would be considered as being *outside or independent of the subject*. To some (n) may be unacceptable for this reason.

Knowledge, traditionally used to describe a mental state in a human being, is nowadays also increasingly being used as a description of the internal "mental" state of a computer. The knowledge of a computer is the internal set of semantic structures represented in an appropriate (formal) language in combination with procedures to process these structures. For the purpose of our discussion it is not necessary to distinguish between

human and machine knowledge. The consequences are the same in both cases.

Knowledge can be accessed by *memorizing* (or in case of a machine by retrieval). Knowledge can be activated from a personal mental treasure independent of other people, whereas *information* depends on *interaction* with someone or something outside the subject. Information is thus embedded in a *communicative context*, although from our argument so far it should be clear that this communicative context need not entail face-to-face communication but may be a *face-to-file*: traditionally this has meant a person reading printed materials (files), but nowadays it increasingly applies to people accessing machine files.

Information - in contrast to knowledge - is not gathered by introspection. For systematic reasons we would like to exclude the set of communicative situations from procedures such as memorizing, thinking or recalling. People who intend to inform themselves or to gather information do not listen to themselves but look for external resources.

It is not our intent to over-interpret natural language formulations but let us propose the following *convention*:

> With "knowledge" we understand the fixed stock of models about objects and facts. These models are accessible by human beings or machines, useful for individuals, groups, organizations. Knowledge is the summation of individual, group-related, organizational and/or cultural experiences and judgements, based on good reason, and is thus different from opinion.

From a rationalistic point of view one might assume that *information* should also be based on reliable knowledge. It is even part of the professional ethic of information specialists to deliver only information which is well founded. But there is no systematic way to delimitate information from knowledge, and information from *opinion*. Sometimes an opinion is of high information value, for example what *x* believes may be real information to *y* if *x* takes decisions which are relevant for *y*. Thus it should be pointed out that the *pragmatic information concept* (information being relevant to a person in a certain situation) threatens the *distinction between truth and opinion* which is fundamental to our western culture.

If someone says they need information, they are in effect stating that they do not dispose of specific knowledge which they believe someone else has. What is needed in such a situation is a *transfer mechanism* between the person who has a chunk of knowledge needed and the person who needs it as information. When we talk about the communicative context of information it is this transfer situation we have in mind.

The *transformation of knowledge into information* depends on at least two partners, whereby one may be a file or a machine. Following this thought through, it must be admitted that communication between two machines can also be described as a transformation of knowledge into information, if one machine stores knowledge (knowledge structures) and the other machine needs some specific information to perform specific tasks such as counselling or diagnosing.

What is important for our transformation process is that the two partners are independent of one another. This independency means that the transformation only can take place if there is a *communication channel* which is accessible to both partners and if the message to be transmitted via these channels can be understood by both sides.

Knowledge - in order to become information - needs to be *transmittable and codable*. Knowledge in itself is determined by internal mental semantic structures, knowledge as a basis for information needs to be materialized, for example in the form of natural language, or in the machine case, in the form of a knowledge representation language (such as semantic networks, frame languages, predicate calculus, production rules).

Knowledge which is materialized and syntactically defined and structured is called data. Therefore it makes sense to call machine processes *data processing* and not knowledge processing. Machines do not process mental (knowledge) structures; they process data as materialized knowledge structures. The result of the processing needs to be reinterpreted in its semantic structure in order to become information which is relevant for an action.

The context of the situation determines the degree of specificity and the quality of information. Therefore information cannot be objectively defined but varies according to different requirements and *contingency factors*. What information is needs to be worked out according to these factors. Information cannot be processed context-free like data but grows out of concrete needs and problem situations.

> **Information** - to put it in a nutshell - can be thought of as **knowledge in action**. We would like to call this postulate the *pragmatic primacy of information work*. Information work is primarily the *transformation of knowledge into information*.

In this pragmatic sense information is - literally speaking- a very *volatile phenomenon*. It is asked for only in very specific situations; afterwards it can be forgotten. This happens often, for instance when one repeatedly looks up the recipe for a delicious sauce in a cookbook rather than memorizing it

because one knows that the information can be always retrieved again. When one finally retains the information mentally so that the next time the cookbook does not need to be consulted, then one has learned something, has acquired some knowledge.

We do not learn information. If one learns something which so far has been the knowledge of someone else or has been generalized or derived from external experiences, then it becomes one's own knowledge. Therefore we can speak of a transformation process in a double sense: firstly, the transformation of knowledge into information which we would like to call "*Informationserarbeitung*" (not information processing but acquiring information), and secondly, when that information is not forgotten but kept, the transformation of volatile information into lasting knowledge which we would like to call "*Informationsverwaltung*" (administration of information).

Information - and this is an important property with respect to its merchandise character - is *recycable*, reusable either by the same user in a different context or by new users. The pragmatic value of information may get lost in usage, but the semantic reference still remains and can be activated in a different situation or by a different person and thus regain a new pragmatic value. This confirms the volatile character of information. Information transferred into a stable state can be called knowledge.

Value-adding. We believe that the distinction between knowledge and information makes sense systematically and is analytically useful. If one assumes that an information system stores knowledge (more precisely: reconstructed knowledge), then different persons with different interests can pose different questions to this system. If the result helps the user in a concrete problem situation, then the retrieved piece of knowledge is called information.

Therefore one could argue that information is not different from the original subset of knowledge but is only a new persepctive (*a view*) on existing knowledge. However, we prefer not to restrict information work to the selection of parts of knowledge from a certain point of view but would like to describe the process of the transformation of knowledge into information as a value-adding process. The transformation produces *informational surplus value.*

Systematically, it is possible to distinguish between *indirect value-adding* processes and *direct value-adding* processes. Indirect processes have more of a support function, whereas direct processes produce surplus values immediately.

Examples of indirect value-adding taken from the domain of natural language processing are:

- natural language access to a data base
- acoustic input facilities
- automatic translation support on different levels

  word level: multilingual thesauri in order to facilitate a multilingual query formulation for retrieval purposes
  sentence level: controlled syntax for writing multilingual abstracts
  text level: automatic translation of full texts

- graphically supported presentation of retrieval results

Examples of direct value-adding are:

- automatically produced flexible text condensates
- inference procedures from existing knowledge structures

- selection of knowledge according to given or automatically constructed user models or situational analyzers respectively.

Of course the border between indirect and direct value-adding fluctuates. What is important for our discussion is that information work (the transformation process) often changes the structure of given knowledge and thus creates something new, which is information which again can be transformed into knowledge. The acquisition of knowledge - this is our argument - occurs in the main via the "detour" of information.

3. The computerization of our world ("Lebenswelt"): a determining factor in the production of knowledge and the use of information

Our contemporary understanding of knowledge and information is increasingly determined by a process of *computerization*. By this we mean the more or less complete penetration of all areas of life ("*Lebenswelt*") by machine computation and machine communication. Processes of computerization are becoming as influential for the world we live in as industrialization was for production and, earlier, the invention of the printing press was for our intellectual life.

Computerization is also changing the conditions of physical, intellectual and social work. Computerization is the main reason why knowledge is being materialized and represented in increasingly different ways by means of a wide variety of primarily electronic tools. We would like to call this the *diversification* of knowledge production and information services.

In the 18th and 19th centuries books, journals and ultimately abstracting journals were the predominant, if not the sole ways of representing knowledge. But the contemporary range of knowledge and information products is so broad that it is almost impossible to grasp it in its entirety:

knowledge can be administrated with the help of electronic media in many ways and can be distributed in equally as many different ways.

In addition to the usual on-line data bases (bibliographic retrieval systems with or without abstracts, full text systems or fact retrieval systems, which manage primarily numeric structures), there are now new products such as *electronically produced and distributed journals*; *image data banks*, which process graphics and/or moving pictures, or, with the development of storage-intensive compact disk technology, *multimedia encyclopedias*, which process and supply textual, graphical and audio-visual material all at once.

The next development is already on the horizon: electronic storages which provide not only the referential products of knowledge (texts, abstracts, titles etc.) but the knowledge structures themselves promise to revolutionize our way of dealing with knowledge. These are already being used in domain-specific expert systems, but will become available on a large scale in new information products such as *knowledge banks*.

The broad variety of knowledge and information products - and this follows directly from computerization - is made possible by an equally broad *diversification of tools for knowledge/information production and distribution:* no longer do we have only typewriters or printing machines but computers of all sizes, interrelated via telecommunication lines; optical storage techniques; all kinds of software systems for text processing, desk-top-publishing, hypertext, electronic publishing, automatic translation; systems for the construction of data banks, shells for expert systems, knowledge acquisition support systems, etc., etc. ....

Knowledge will of course continue to be transferred in the near future in face-to-face situations via the printed media, but electronic media will be used more and more, particularly in professional environments where knowledge and information are considered economical factors.

Naturally the distinction between knowledge and information which is fundamental for information science is valid in a completely computerized society as well. Future knowledge-based systems, which are becoming more and more powerful as far as automatic knowledge acquisition, knowledge administration, inferencing and access techniques are concerned, will not solve the information problem automatically. Information, as we hope to have demonstrated, depends on social and cognitive circumstances. We must construct information systems which take the pragmatic primacy of information work seriously. This is the only way to add value to existing knowledge and transform it into active information.