

Wissensmanagement und Sprachtechnologie

Peter Bosch
IBM Deutschland Informationssysteme GmbH
Institut für Logik und Linguistik
Global Services
Vangerowstr 18, 69115 Heidelberg
pbosch@de.ibm.com

Zusammenfassung

Wissensmanagement zählt technologisch ebenso wie kulturell zu den spannendsten Herausforderungen der globalen Informationsgesellschaft. Wesentliche technologische Aspekte des Wissensmanagements sind ohne Sprachtechnologie und *linguistic engineering* nicht zu bewältigen, es sei denn, wir wären bereit, textuell repräsentiertes Wissen auszuschließen. Mit aktuellen Möglichkeiten des *Text Mining* kann textuelles Wissen in die Prozesse des Wissensmanagements integriert und seine Rolle für Innovationsprozesse kann erheblich verstärken werden. Zugleich entsteht damit die Möglichkeit, das in den Geisteswissenschaften geschaffene Wissen in den Innovationsprozeß einzubeziehen.

Wissensmanagement

Wenn das Thema „Wissen“ heißt, haben traditionell die Philosophen das Wort und die wirklich wesentliche Frage lautet: *Wie ist gesichertes Wissen möglich?* Geradezu blphemisch muß dann die andere Frage lauten, um die es hier gehen soll: Wie können wir Wissen vernünftigerweise – und das heißt: effizient und effektiv – verwalten? Diese Frage soll das Wissensmanagement beantworten – wobei wir den Begriff des Wissens verwenden, ohne irgendwelche Wahrheitsansprüche mit dem zu verbinden, was da „Wissen“ genannt wird.

Das Gebiet der Wissensverwaltung, unter diesem Namen oder dem eher im Trend liegenden Namen „Wissensmanagement“, ist eine relativ neue Sache.

Es geht im Wissensmanagement um

- den Aufbau,
- die Nutzung,
- die Weitergabe und
- die Aufbewahrung

von Wissen. – Einige Elemente erscheinen vertraut und sind es dennoch nicht ganz; unter anderem deshalb, weil sie im Unternehmenskontext und im Kontext neuer Informationsmedien eine veränderte Rolle spielen. Am ehesten sind noch jene Elemente des Wissensmanagements wiederzuerkennen, die aus dem traditionellen Arbeitsgebiet der Bibliothekare, Archivare und Dokumentare hervorgehen. Aber ist der Inhalt wirklich derselbe?

Solange Wissen in relativ wenigen gedruckten Büchern abgelegt ist, von denen jedes (nicht von irgendwem, sondern vornehmlich von relativ wenigen Gelehrten und ihren Helfershelfern)

geschrieben, gesetzt, korrekturgelesen, gedruckt, gebunden, rezensiert, in Bibliotheken aufgenommen und entsprechend katalogisiert wird, ist das eine Sache, und ein Hegel kann sich gegenüber seinen Studenten – die, wie immer, auch damals schon zu wenig lasen – in der Vorlesung den arroganten Verweis erlauben, man *könne* alles lesen und er, Hegel, *habe* alles gelesen. Auch zu Zeiten meines Studiums erlaubte sich gelegentlich ein Professor – und nicht völlig ironisch, sondern mit einem gewissen Stolz auf die anachronistische Übertreibung – die Wiederholung dieses Hegelzitats.

Doch seit Hegel hat sich die Szene drastisch verändert: nicht nur zählt schon lange nicht mehr nur das, was Gelehrte schreiben, oder was zwischen zwei Buchdeckel gebunden ist. Was im Internet zu lesen ist, könnte auch der anmaßendste Vielleser nicht behaupten, gelesen zu haben. Und sehr zum Pech unseres arroganten gebildeten Viellesers ist ja nicht all diese Information im Internet für jeden unwichtig. Und dann ist das Internet auch nicht das einzige Kommunikationsnetz.

Wenn auch die Kommunikationsnetze zunächst als extremer Fall erscheinen mögen, so haben sie doch noch die hilfreiche Eigenschaft, daß zumindest alle Information in elektronischer Form – und zum großen Teil in einheitlichem Format – vorliegt. Dies ist nicht der Fall in der Situation, die der Industrie Kopfzerbrechen bereitet. Von meinen Kollegen bei Siemens habe ich gelegentlich den Seufzer gehört "Wenn Siemens nur wüßte, was Siemens alles weiß!" Hier geht es auch um das Wissen in den Köpfen der Mitarbeiter, wie in den letzten paar Jahren das teuerste Reengineering-Thema des Jahrhunderts zeigt: die Jahr-2000-Umstellung. Wer gegenwärtig noch das erforderliche Wissen besitzt, um mit in vielerlei Hinsicht veralteten Großrechner-Betriebssystemen und der dazugehörigen Middleware und Anwendungen umzugehen, der kann sich wahrhaftig eine goldene Nase verdienen. Informatiker-Rentner werden quasi aus dem Altersheim geholt und stöbern in FORTRAN- und COBOL-Kode, weil sonst nicht nur ab 01.01.2000 ihre Rente nicht mehr gezahlt würde, sondern weil auch ein paar andere Prozesse in der Wirtschaft global schief laufen würden: Wissensmanagement ist nicht bloß ein Thema der Aufbewahrung und Archivierung, sondern auch ein Personalthema: Ausbildung, Weitergabe von Wissen, Aufbau von Kompetenz, Skill-Management, das Ausfindigmachen der richtigen Mitarbeiter für anstehende Aufgaben, sind Themen der Investitionssicherung und damit der Wettbewerbsfähigkeit. Der Umfang des Problems ist mit dem aus jeder Proportion fallenden Wissenszuwachs nicht mehr nur mit klassischen Ansätzen zu bewältigen. Nur noch in seltenen Glücksfällen hilft es, ein oder zwei Bekannte oder Kollegen anzurufen, die sich mit der zur Diskussion stehenden Frage auskennen und uns sagen, wer gegenwärtig die Ansprechpartner sind und was wir zu dem Thema lesen müssen. Es geht um Informationskultur. Die alte Informationskultur kann die Probleme der Wissensverwaltung im aktuellen Ausmaß nicht mehr mit ihren Mitteln verarbeiten. Eine Kulturrevolution steht nicht ins Haus, aber der Aufbau einer neuen Informationskultur, in der Wissen in einem neuen Sinn als kostbares Gut behandelt wird.

Was hier neu ist, ist natürlich nicht, daß Wissen besser bewacht werden müßte und wir noch kleinlicher mit Intellectual Property umgehen sollten. Diese Seite hat alte Wurzeln in der Wissenschaft, und manch ein Scharlatan wäre schneller entlarvt worden, wenn er seine kleinen Forschungsgeheimnisse nicht so gut gehütet hätte. Der protektionistische Umgang mit Unternehmenswissen und Intellectual Property steht mit gutem Grund in dieser Tradition. Hier ist nicht eine kulturelle Veränderung angesagt, sondern nur die konsequente Übertragung wohl etablierter Grundsätze auf neue Kommunikationsmedien; hier geht es um das traditionelle Thema Informationssicherheit. – Nein, die wesentliche kulturelle Veränderung, die erfolgreiches Wissensmanagement verlangt, besteht darin, daß wir lernen, zwar dem Mitbewerber gegenüber unser Wissen für uns zu behalten, aber dennoch einen effektiven und effizienten Austausch von Wissen innerhalb der eigenen Organisation, und selektiv mit Kunden,

Distributoren und Zulieferern zu ermöglichen. Was kulturell zu erreichen wäre, ist eine Haltung, die Mitarbeiter, Kollegen und Vorgesetzte, und ebenso Geschäftspartner, nicht als potentielle Mitbewerber betrachtet sondern ihnen vorhandenes Wissen freigebig zugänglich macht, ja anträgt. Das neue Verständnis ist, daß der Besitzer von Wissen eine Bringschuld hat.

In diesem Beitrag geht es mir jedoch nicht vornehmlich um Vorschläge zu einer neuen Wissenskultur. Es geht eher, innerhalb der Problematik des Wissensmanagements, um Technologieentwicklung, speziell die Entwicklung der Sprachtechnologie.

Technologische Beiträge zum Wissensmanagement konzentrieren sich bisher auf das Management des Ablaufs von Geschäftsprozessen (Workflow), Dokumentenmanagement oder Archivierung und auf Data Mining. Im folgenden geht es um eine neue Komponente in diesem Ensemble: um Text Mining.

Text Mining

Data Mining ist in zwischen in weiten Kreisen zu einem Begriff geworden. Die Grundüberlegung geht aus von den im Zuge der Ausbreitung der Informationstechnologie allorts entstehenden Datenfriedhöfen und dem Gedanken, daß manches Nützliche darin verborgen sein könnte. Nicht, daß man versehentlich wertvolle Daten weggeworfen hätte statt sie zu archivieren. Es geht um Information, die *implicit* in den Datenmengen vorhanden ist und nur mit mehr oder weniger cleveren Schürftechniken (daher „mining“) extrahiert werden kann. Ein Beispiel, mit dem die Datenschürfer ihre Aufgabe gern illustrieren bezieht sich auf den Datenfriedhof, den die Registriertassen der Supermärkte tagtäglich erzeugen. Mit einfachsten Mitteln läßt sich feststellen, daß z.B. Kunden, die Spargel kaufen, relativ häufig auch gekochten Schinken oder rohen Schinken kaufen. Nicht, daß uns dies überraschen würde (wir wollten ja auch gern ein Beispiel sehen, das sofort einleuchtet), aber es dürfte klar sein, daß wir mit unserer kulinarischen Intuition nur gelegentlich und unsystematisch solche Zusammenhänge im Kaufverhalten des Supermarktkunden aufspüren können. Data Mining bietet die Möglichkeit, solche Zusammenhänge systematisch aufzudecken und bietet sie unserem Supermarktleiter zur Optimierung an: Er kann nun getrost die Kunden mit einem Spargel-Sonderangebot in den Laden locken und zugleich sicher sein, daß er sich über die Preisgestaltung beim Schinken schadlos hält.

Worum es beim Data Mining geht ist die Aufdeckung von zunächst verborgenen jedoch hochsignifikanten Korrelationen zwischen diversen Parametern. Klassifikations- und Clustertechniken bilden den Kern der hierfür erforderlichen Algorithmen.

In welchem Sinn also kann nun neben *Data Mining* auch von *Text Mining* gesprochen werden? Wie und mit welchem Ziel können ähnliche Ansätze wie beim Data Mining auf Texte angewendet werden? Der Gedanke mag zunächst überraschen, weil es von der Anwendung und Zielsetzung her gerade nicht um quantitative Eigenschaften von Texten geht. Wenn wir Wissensmanagement betreiben, interessiert uns z.B. die Tatsache, daß in der *Süddeutschen Zeitung* bestimmte Wörter hochfrequent auftreten, die im *Spiegel* eher selten sind, eigentlich nicht (wobei mir ehrlich gesagt, immer schleierhaft war, warum und zu welchem Zweck sich überhaupt irgend jemand für derartige Ergebnisse interessieren könnte). Wir wollen mit automatischen Verfahren im Text Mining *inhaltliche* Zusammenhänge und Eigenschaften von Texten entdecken.

Die Aufgabe, mit automatischen Verfahren Textinhalte zu erfassen gehört klassisch zu den Aufgaben der Künstlichen Intelligenz, genauer zu dem Untergebiet natürlichsprachliches Verstehen (*natural language understanding*) oder zu dem Arbeitsgebiet der Erstellung von Textzusammenfassungen (*abstracting*). Ein wesentliches Hemmnis für die praktische

Anwendung dieser KI-Ansätze bestand jedoch immer darin, daß ohne eine sehr umfassende Repräsentation des in jedem Fall relevanten Alltagswissens und des für die jeweiligen Texte relevanten Domänenwissens (oder Fachwissens) keine tatsächlich funktionierenden Systeme erstellt werden können. Das für die jeweilige Sprache relevante linguistische Wissen, d.h. die Syntax, Semantik und Morphologie, können diese Aufgabe allein nicht erledigen. Der Aufbau der zusätzlich erforderlichen wirklich sehr umfassenden Wissensrepräsentationen ist jedoch – trotz einiger sehr beeindruckender Versuche – nicht soweit vorangekommen, daß hier ernsthaft von vielversprechendem Fortschritt gesprochen werden könnte. – Ich will nicht sagen, daß der KI-Ansatz gescheitert sei. Er hat mit Sicherheit zu wichtigen neuen Erkenntnissen über Eigenschaften und Struktur der menschlichen Sprache und des menschlichen Wissens geführt – nicht jedoch zu in einem ernst zu nehmenden Sinn zu industriell verwertbaren Anwendungen. Das kann noch kommen und diese Forschungsrichtung muß in jedem Fall mit vollem Einsatz weiter betrieben werden.

Mittlerweile jedoch haben sich – fast, wie es scheinen mag, in Konkurrenz zu KI-Ansätzen – stochastische Ansätze entwickelt, die für viele Zwecke schlicht besser geeignet sind. Besser, weil sie zunächst einmal ohne die schwer lastende Voraussetzung der großen Wissensrepräsentationen auskommen und damit ein großes Hemmnis für ihre Machbarkeit aus dem Wege geräumt ist. Besser auch, weil sie sehr viel billiger entwickelt werden können.

Nehmen wir als Beispielanwendung die Gruppierung von Texten in Ähnlichkeitsklassen, d.h. Clustering. Es ist gewiß nicht absurd, zu vermuten, daß Texte eines Autors sich von Texten anderer Autoren erkennbar unterscheiden und daß z.B. ein Literaturhistoriker kompetente Aussagen über die Autorenschaft eines Textes machen kann. Die Hypothese stochastischer Clustering-Techniken ist nun, daß sich eine Ähnlichkeitsklasse von Texten, also ein Textcluster – wie z.B. die Texte eines Autors im Vergleich zu den Texten anderer Autoren – mit Hilfe quantitativer Parameter charakterisieren läßt. Dies können z.B. Parameter relativer Worthäufigkeiten sein, Parameter relativer Häufigkeit syntaktischer, morphologischer oder auch semantischer Merkmale (vorausgesetzt es stehen Lexika zur Verfügung, die solche Eigenschaften hergeben und vorausgesetzt die Verfügbarkeit automatische Taggingtechniken) und natürlich Kombinationen solcher Parameter und Relationen darüber. Solche Charakterisierungen von Textclustern sind notabene immer *relative* Charakterisierungen, abhängig von den kontrastierenden Texten. Dies bedeutet, daß Clusteringtechniken eine Gruppe von Texten, relativ zu einem bestimmten Textkorpus bestimmen. Da diese Techniken automatisch ausgeführt werden, – d.h. die Algorithmen probieren sozusagen die ihnen verfügbaren Parameter an dem jeweiligen Korpus aus und verwenden am Ende solche Parameter, die zu einem den Vorgaben entsprechenden Clustering führen – und da diese Algorithmen i.a. sehr schnell ausgeführt werden können, ist die Relativität nur von Vorteil: Clustering kann auch rekursiv angewendet werden, um immer feinere Gruppierungen zu erstellen: innerhalb eines Clusters werden weitere Cluster erstellt.

Wenn es darum geht, die Texte eines Autors im Vergleich zu den Texten anderer Autoren mit Hilfe quantitativer Parameter charakterisieren, so ist diese Anwendung zunächst noch eine rein formale und nicht ersichtlicherweise eine inhaltliche Charakterisierung. Die Hypothese im Text Mining ist jedoch, daß mit denselben Mitteln inhaltliches Clustering erreicht werden kann. Dies ist eine empirische Hypothese, die in der Anwendung auf diverse Klassifikationsabsichten getestet werden kann.

Nehmen wir an, wir haben bereits ein Korpus gruppierter Texte vorliegen, etwa in einem Archiv nach Sachgebieten. Dann können wir unsere Algorithmen mit Hilfe von Techniken des maschinellen Lernens die verwendete Klassifizierung „lernen“ lassen und auf weitere bisher nicht klassifizierte Texte anwenden. Auf diese Weise können wir z.B. die Klassifikation eines

bestehenden Archivs von Zeitungsmeldungen auf die offene Klasse neu eingehender Zeitungsmeldungen übertragen. Die Klassifizierung ist inhaltlicher Art und im vorliegenden Fall stellen wir fest, daß die Klassifizierung sehr genau nachgebildet werden kann und zwar ausschließlich mit dem Parameter relativer Worthäufigkeiten.

Durch intelligenten Einsatz solcher rein quantitativer Parameter können Ergebnisse erreicht werden, die auch scheinbar subtile Unterscheidungen erfassen. Wenn wir mit boolescher Stichwortsuche in einer Textdatenbank medizinischer Fachveröffentlichungen Artikel suchen, die sich mit Aspirin als möglicher Ursache von Kopfschmerzen beschäftigen, bekommen wir in jedem Fall Probleme mit der Treffsicherheit (precision), weil uns die Suche auch Artikel liefert, in denen Aspirin als Heilmittel für Kopfschmerzen behandelt wird. Eigentlich würden wir gern mit strukturierten Termen suchen, wie *URSACHE(aspirin, kopfschmerz)* oder *HEILMITTEL(aspirin, kopfschmerz)*. Doch dies sind *semantische* Relationen, die im klassischen Ansatz zunächst einmal eine relativ subtile linguistische Analyse der zu durchsuchenden Texte verlangen und dann eine Abbildung auf eine konzeptuelle Repräsentation. In einem quantitativen Ansatz ist es prinzipiell denkbar, strukturierte Terme durch rekursives Clustering in der Datenbank nachzubilden. Etwa indem wir zunächst ein Cluster bilden, in dem es sowohl um Aspirin als auch um Kopfschmerzen geht (das bedeutet normalerweise nicht, daß wir im Clustering nur diese Terme verwenden, sondern auch Terme, die für unser Kopus oder bereits untersuchte Korpora ähnlicher Art signifikant mit ihnen korrelieren). Ein weiteres Clustering innerhalb der so erzeugten Gruppe könnte Cluster erzeugen, in denen es primär um Ursachen von Beschwerden geht und ein anderes, in dem Effektivität von Heilmitteln behandelt wird, so daß wir Ergebnisse erzielen, die der gewünschten Differenzierung entsprechen.

Ich habe mich soeben bewußt vorsichtig ausgedrückt und Formulierung gewählt *wie es sei denkbar daß, man könne dieses oder jenes versuchen*, etc. Dieses Denken ist charakteristisch für den quantitativen Ansatz: man arbeitet empirisch und experimentell. Solange wir das Experiment nicht ausgeführt haben, können wir über die Resultate nur Vermutungen anstellen, aber Sicherheit erhalten wir nur aus der Empirie. Die verwendeten Algorithmen werden an kleinen Mengen von relativ gut bekanntem Datenmaterial trainiert und dann auf große Mengen von ähnlichem Material angewendet. So funktioniert Clustering, so funktioniert stochastisches Tagging, und so funktioniert Textklassifikation.

Ein Cluster kann beschrieben werden durch die quantitativen Eigenschaften, die es definiert. Im einfachen Fall: eine Menge hochfrequenter Lexeme. Grundlage des semantischen Clusterings ist der Gedanke, daß sich – wie immer relativ zu einer bestimmten Textmenge oder einer bestimmten Textdatenbank – regelmäßige Beziehungen zwischen solchen Mengen hochfrequenter Lexeme und bestimmten Konzepten bestehen.

Mit diesem Gedanken ergibt sich konsequent eine Konzeption robuster semantischer Suche, die weit über thesaurusbasierte Suche hinausgeht. In einem Thesaurus wird die Menge der Suchwörter, vereinfacht gesagt, durch Synonyme erweitert. Eigentlich suchen wir Fundstellen für das Konzept Arbeit, aber der Thesaurus sorgt dafür daß die Suchmaschine alle morphologisch aus „Arbeit“, „Tätigkeit“, „Beruf“, etc. hervorgehenden Zeichenketten sucht. Doch die Grundlage des Thesaurus ist im wesentlichen die Intuition und die Zuhilfenahme von Lexika durch die Autoren des Thesaurus. Im maschinellen Lernen werden die Frequenzkorrelationen in entsprechenden Clustern untersucht. In einer hierauf basierenden Suche in Textdatenbanken werden im wesentlichen mit Hilfe der bereits bekannten Charakterisierung eines Begriffsclusters *Arbeit* weitere Texte für dieses Cluster gesucht. – Ob das Wort „Arbeit“ darin vorkommt ist relativ nebensächlich.

Die sich hiermit ergebende Robustheit kann nun den Anwendungsbereich für das Textmining erheblich erweitern, wie wir das im Institut für Logik und Linguistik in den letzten Monaten

getan haben: wir wenden Textmining nicht nur auf relativ „saubere“ Texte an, sondern auch „unsaubere“ Texte, wie sie z.B. aus der (immer mit einer gewissen Fehlerquote behafteten) Spracherkennung hervorgehen oder aus Systemen der (ebenfalls niemals hundertprozentigen) automatischen Übersetzung. Selbst wenn die Übersetzung nur mäßig gut ist oder die Spracherkennung nur mittelmäßige Ergebnisse liefert, so können die Sprachaufnahmen und die ursprünglichen Texte dennoch – mit Hilfe der robusten Clusteringtechniken – ebenso sauber eingeordnet werden wie „saubere“ Texte. In unseren Anwendungen ging es notabene nicht darum, unsere Textdatenbank mit den aus der Spracherkennung resultierenden Texten oder mit Übersetzungen anzureichern. Die Aufgabe bestand darin, eine Datenbanken, die Sprachaufnahmen bzw. fremdsprachige Texte neben regulären deutschen Texten enthielt, inhaltlich zu ordnen.

Die Techniken des Text Mining stehen noch am Anfang ihrer Entwicklung und wir haben guten Grund von ihrer weiteren Entwicklung noch erhebliche Fortschritte zu erwarten. Daß Text Mining uns interessante Einsichten in die menschliche Kognition oder speziell die menschliche Sprachverarbeitung liefern wird ist eher zu bezweifeln. Aber das ist auch nicht ihr Ziel. Es geht um linguistic engineering: es sollen effiziente und effektive Lösungen für Probleme des Wissensmanagements geliefert werden.

Eine ganz wesentliche Konsequenz aus den aussichtsreichen Entwicklungen des Text Mining möchte ich noch kurz ansprechen. Wenn wir in der Lage sind, sprachliche Texte nach ihren Inhalten automatisch zu verarbeiten, bedeutet dies, daß die natürliche menschliche Sprache ihre Rolle in der Kommunikation und in der Informationsspeicherung behaupten und erweitern kann. Wir können das natürlichste Medium der Informationsspeicherung und Wissensrepräsentation, mit dem jeder Mensch umgehen kann, das ein Maximum an ergonomischem Komfort bietet, tatsächlich in den Kontext der Informationstechnologie herüber retten. Mehrdeutigkeiten und Ungenauigkeiten des sprachlichen Ausdrucks werden dabei in Kauf genommen und stören die Informationsverarbeitung nicht. Texte werden besser zugänglich gemacht und besser verwaltet, bleiben aber letztendlich als sprachliche Texte erhalten, die wie eh und je gesprochen, geschrieben, abgehört und gelesen werden. Der aufwendige Umweg über eine Übersetzung und Abbildung in formale Sprachen – nur um einer zunehmend zum alten Eisen gehörigen und zu nichts anderem fähigen Informationstechnologie einen Gefallen zu tun – bleibt uns erspart.

Wissensmanagement kann sich mit den Methoden des Text Mining mehr als wir in der Vergangenheit zu träumen gewagt haben der ergonomischen Vorteile der sprachlichen Repräsentation zunutze machen. Nicht nur Wissen in der Form und Zahlen und Tabellen, wie es im Data Mining erschlossen wird, sondern auch Wissen in sprachlicher Form kann effizient verwaltet werden.

Wissensmanagement, Innovation und Geisteswissenschaften

Unser Thema in dieser Veranstaltung ist die Frage nach Beiträgen der Geisteswissenschaften zu Innovationen. Ich glaube nicht, daß das Innovationspotential der Geisteswissenschaften als solches zur Diskussion steht. Das im Wesen interpretierende Herangehen der Geisteswissenschaften an Probleme ist ein absolut notwendiges Element in jedem Innovationsprozeß, der von der Gesellschaft getragen wird. Innovationen, die wir Menschen nicht verstehen, tragen wir auch nicht mit, sondern richten uns gegen sie. Innovation braucht eine Kultur, die sie trägt, und in dieser Kultur sind die intellektuell und künstlerisch vermittelnden, verstehenden und interpretierenden Vorgehensweisen der Geisteswissenschaften ein wesentlicher Bestandteil. – Dieses Thema weiter auszuführen ist jedoch vielleicht eher Sache eines Geisteswissenschaftlers. Worum es mir in diesem Beitrag ging ist etwas anderes:

ich glaube, daß wir von der Technologie her noch viel tun müssen, um die *Bedingungen der Möglichkeit* eines Beitrags der Geisteswissenschaften zur Innovation zu schaffen oder zu erhalten. Und das heißt unter anderem, daß wir das in den Geisteswissenschaften in der Form von Texten geschaffene Wissen den technologischen Prozessen des Wissensmanagement zugänglich machen: in Texten vorliegendes Wissen kann nur dann die ihm zukommende Rolle für Innovation spielen, wenn es zunächst einmal auf eine Art zugänglich gemacht wird, die nicht von uns verlangt, daß wir, *alles gelesen haben*, sondern uns das sichere Auffinden und eine selektive Nutzung ermöglicht.