



# In Richtung Summarizing für Diskurse in K3

Rainer Kuhlen

## Zusammenfassung

Der Bedarf nach *Summarizing*-Leistungen, in Situationen der Fachinformation, aber auch in kommunikativen Umgebungen (Diskursen) wird aufgezeigt. *Summarizing* wird dazu in den Kontext des bisherigen (auch automatischen) *Abstracting/Extracting* gestellt. Der aktuelle Forschungsstand, vor allem mit Blick auf *Multi-Document-Summarizing*, wird dargestellt. *Summarizing* ist eine wichtige Funktion in komplex und umfangreich werdenden Diskussionen in elektronischen Foren. Dies wird am Beispiel des *e-Learning*-Systems K3 aufgezeigt. Rudimentäre *Summarizing*-Funktionen von K3 und des zugeordneten K3VIS-Systems werden dargestellt. Der Rahmen für ein elaborierteres, Template-orientiertes *Summarizing* unter Verwendung der vielfältigen Auszeichnungsfunktionen von K3 (Rollen, Diskurstypen, Inhaltstypen etc.) wird aufgespannt.

## 1 Einleitung

*Summarizing* ist eine grundlegende Informationsleistung und beruht auf methodisch geschulter Informationskompetenz. Zweck des *Summarizing* ist es, komplexe Wissensobjekte so zusammenzufassen, dass andere Menschen, die keine Zeit haben, alle möglicherweise einschlägigen Wissensobjekte vollständig durchzusehen, doch den Überblick über ein Wissensgebiet oder über die Produktion eines aktuellen Projektes behalten können.

Der Bedarf nach solchen Zusammenfassungen entstand Anfang des 19. Jahrhunderts angesichts der steigenden Publikationszahlen in den wissenschaftlichen Disziplinen, und er ist heute größer denn je. Die Zahlen der „großen drei“ naturwissenschaftlichen Fächer sprechen für sich. Pro Jahr entstehen in der Chemie fast 750.000 Artikel (fast 85% auf Englisch), in der Physik ca. 300.000 und in der Biologie 550.000 – so gut wie alle werden mit Zusammenfassungen (im Fachinformationsgebiet *Abstracts* genannt) versehen – oft genug als Leistung der Autoren selber, auch wenn Einiges dafür spricht, dass die Qualität von Autorenreferaten geringer ist als die der von Informationsspezialisten erzeugten.

Aber nicht nur auf dem Gebiet der traditionellen Fachinformation besteht *Summarizing*-Bedarf. Die Verwendung von Textverarbeitungssoftware in allen (privaten und professionellen) Lebensbereichen, die Ausweitung der vielfältigen Formen elektronischer Kommunikation und die flächendeckende Darstellung von Wissen jeder Art in unzähligen Websites hat eine Explosion von Wissensobjekten jedweder medialen Form bewirkt. Einher geht dies fast zwangsläufig mit einem Orientierungsproblem. Suchmaschinen sind mit ihrer Referenzfunktion *eine* umfassende Antwort auf die Schwierigkeit, aus der Fülle des produzierten Wissens diejenigen relevanten Objekte nachgewiesen zu bekommen, die aktuell benötigt werden.

*Summarizing* ist ebenfalls eine (zuweilen substituierende, in der Regel aber referenzierende) Orientierungsleistung.

Ein besonderes Orientierungsproblem ist gegenwärtig bei Formen elektronischer Kommunikation auszumachen. Wir meinen damit vor allem die Unübersichtlichkeit in elektronischen Kommunikationsforen, in denen bei interessanten Themen sehr schnell sehr umfangreiche und durch die Vernetzung sehr komplexe Diskurse entstehen, für die erst ansatzweise Orientierungsformen existieren. Dazu gehört vor allem die Darstellung von Diskursen als *Threads*, also die in der Regel hierarchisch geordnete Auflistung der die Beiträge charakterisierenden Themen (vgl. Abschnitt 3.1 und Abb. 7). Ansatzweise werden auch Mittel der Visualisierung zur Darstellung von Kommunikationsstrukturen verwendet (vgl. Abschnitt 4.1 und Abb. 9).

Bislang ist aber kaum versucht worden, *Summaries* von Diskursen in elektronischer Kommunikation automatisch zu erstellen. Als Vorarbeit für ein geplantes Diskurs-*Summarizing* in den Systemen K3 und eAgora (vgl. Abschnitt 4) wollen wir zunächst das Diskurs-*Summarizing* in den Kontext des bisherigen *Abstracting/Summarizing* stellen und anschließend ein Anwendungsszenario für ein mögliches Diskurs-*Summarizing* entwickeln, wie es für die Orientierung in virtuellen Kommunikationsforen in der Hochschulausbildung zum Einsatz kommen soll (K3), aber auch im Bereich von eScience (kollaboratives Wissensmanagement in virtuell und verteilt arbeitenden Forschungsgruppen – eAgora) (vgl. Abschnitt 3).

## 2 Informationsmethodisches *Summarizing*

### 2.1 Formen des *Abstracting*

Im informationswissenschaftlichen Zusammenhang wird *Summarizing* in erster Linie als *Abstracting* verstanden (Kuhlen 2004a). Gemeint ist damit die Zusammenfassung von in der Regel wissenschaftlichen Texten, die als indikative *Abstracts* (hier synonym mit *Referate* verwendet) der Orientierung dienen, und als informative *Abstracts* auch zusammenfassend direkt über die wesentlichen Inhalte, Methoden und Ergebnisse informieren sollen, und zwar in der Regel in textueller Form. *Abstracts* sollen nach herrschender Meinung keine Wertungen enthalten, sondern das wiedergeben, was im Originaltext vorzufinden ist, nach Möglichkeit unter Berücksichtigung der Formulierungen in den Ausgangstexten. Letztere Anforderung dient auch als Rechtfertigung für das (automatisch betriebene) *Extracting*. Es findet sich jedoch auch die Form des kritischen Referats, das in seinem normativen Charakter, wenn auch wohl kaum im Umfang, schon der Textform des *Reviews* nahe kommt. Eine attraktive Ausprägung ist das sogenannte *Slanted abstract*, das auch als adaptives *Abstract* bezeichnet werden kann, weil es mit Blick auf eine eng definierte Zielgruppe oder sogar auf eine einzige Person hin angefertigt wird.

Neben den textuellen Ausprägungen von *Abstracts* sind auch *Abstracts* als grafische Visualisierung und als strukturierte Positionsreferate entwickelt worden, die, mit vorgegebenen *Templates* arbeitend, als semantisch typisierte Wissensobjekte aufgefasst werden können [Hammwöhner et al. 1989]. Die Zusammenfassung von mehreren Texten (oder Wissensobjekten allgemein) wurde in der Dokumentationspraxis bislang nicht Gegenstand des *Abstracting*. Diese Leistung – meistens auch in kritischer, nicht nur synthetischer Absicht – wird als Fortschrittsbericht (im Englischen treffender als *State-of-the-art-report*) bezeichnet. Dies ist eine in der Wissenschaft außerordentlich geschätzte Leistung, die als selbständige Arbeit pub-

liziert werden kann, meistens aber auch als unverzichtbarer Bestandteil einer jeden wissenschaftlichen Qualifikationsarbeit (Dissertation, Habilitationsschrift) angesehen wird.

## 2.2 Automatisierung des *Abstracting*

Angesichts der eingangs angedeuteten quantitativen Herausforderung, die immer umfänglicher werdende Produktion von Wissensobjekten überschaubar zu halten, ist es verständlich, dass die Automatisierung des *Abstracting* bzw. des Summarizing als Lösung des Problems angesehen wird. *Abstracting* gehört als *Extracting* zu den ersten Anwendungsgegenständen automatischer Sprachverarbeitung. Vorbild bis heute ist, wenn auch ständig modifiziert, der Luhn-Algorithmus [Luhn 1958], der Sätze mit signifikant hohem Vorkommen signifikanter Wörter extrahierte<sup>1</sup>. Die in den 80er und 90er Jahren entwickelten wissensbasierten und textlinguistischen Verfahren, z.B. automatische Diskursanalyse mittels der *Rhetorical Structure Theory*, waren für die Künstliche-Intelligenz-Forschung ein attraktives Forschungsfeld zur Simulation menschlicher intelligenter Performanz. Auch in der Informationswissenschaft wurde damit experimentiert [Endres-Niggemeyer 1999]. Exemplarisch für ein wissensbasiertes und linguistisch fundiertes System sei hier nur das in Konstanz entwickelte Topic/Topographic-System genannt [Hahn 1990; Reimer 1990; Sonnenberger 1988; Hammwöhner et al. 1989; später: Hahn/Reimer 1999]. Allerdings ist nicht zu übersehen, dass Interesse und Zuversicht in die Leitungsfähigkeit bzw. Generalisierbarkeit des wissensbasierten und textlinguistischen Paradigmas in den letzten Jahren stark abgenommen haben (nicht zuletzt wegen der weiterhin kaum zu überwindenden semantischen Komplexitäts-, aber auch der schlichten Quantitätsprobleme in realistischen Anwendungssituationen).

## 2.3 *Abstracting/Extracting* im praktischen Einsatz

Gegenwärtig finden praktische Zusammenfassungsleistungen, meistens auf einfacher statistischer Grundlage, aber zuweilen auch mit elaborierteren Verfahren, immer mehr Eingang in die Textverarbeitung, z.B. bei Word/Office, aber auch als Leistung eines allgemeinen Textmanagements<sup>2</sup>. Zum Teil elaborierter in den Verfahren und häufig auch für mehrere Sprachen werden *Summarizer* mit Blick auf Web-Dienste angeboten<sup>3</sup>. *Summarizer* werden im Medienbereich auch als adaptive *Harvester* eingesetzt, indem nach einem vorgegebenen Interessenprofil neue Inhalte aus Webseiten zusammengefasst und fortlaufend als „Zeitung“ präsentiert werden.

---

<sup>1</sup> In [Kuhlen 1989; 2004b] werden diese frühen Ansätzen dargestellt, die u.a. durch einfache linguistische Verfahren weiterentwickelt wurden, wie Extrahieren von Topic-Sätzen (Sätze zu Beginn oder Ende eines Absatzes) oder textsyntaktisches Verknüpfen zum Bilden von Kohäsion von ansonsten isolierten extrahierten Sätzen.

<sup>2</sup> Z.B. *Brevity Document Summarizer* von Lextek - <http://www.lextek.com/brevity/>), *Automatic Text Summarizer* - <http://search.iiit.net/~jags/summarizer/index.cgi>; *Extractor* von DBI (<http://www.extractor.com/>) oder *Metafer Summarizer* (<http://www.metafer.de/de/test/onlinetest.html>).

<sup>3</sup> z.B. der *X-Summarizer* von Arexera (<http://www.arexera.de>), der *Copernic Summarizer* (<http://www.copernic.com>), die *DIGEST-Software* (<http://www.digestsoftware.de/info.htm>), der *Inxight Summarizer* (<http://www.inxight.com/products/sdks/sum/>), der *NetOwl-Summarizer* mit adaptiven Eigenschaften, *Pertinence Summarizer* (<http://www.pertinence.netf>), der *DBI Text Summarizer* (Version 7.1) (<http://www.download.com/3000-2070-10361551.html>) oder *Sinope Summarizer*, integriert in den Internet Explorer (<http://www.sinope.info/en/index.php>)

## 2.4 Multidocument Summarizing (MDS)

Erst in der letzten Zeit ist textübergreifendes *Abstracting*, als *Multidocument Summarizing* (MDS) Gegenstand der computerlinguistischen Forschung geworden, wobei allerdings nur noch selten textgrammatische Verfahren zum Einsatz kommen, sondern eher elaboriertere statistische Verfahren oder auch neuronale Netzwerke, die aber natürlich auch zu den computerlinguistischen Textverarbeitungsverfahren zu rechnen sind. Hier hat die *Document Understanding Conference* (DUC) von 2001 ein brauchbares Testbett für MDS-Anwendungen entwickelt. Mit MDS wird derzeit intensiv experimentiert, z.B. über die Verwendung von Clustertechniken relevanter Sätze aus verschiedenen Texten [Siddharthan et al. 2004], von „lead sentences“ und Häufigkeiten informationsintensiver Konzepte anstelle von Textwörtern [Schiffman/Nenkova/McKeown 2002], einer Kombination von Zufallssuchtreffern und der Extraktion benachbarter Sätzen [White/Cardie 2002]. MDS-Grundlagen können aber auch Retrievalergebnisse [Mori 2002], thematisch aufeinander bezogene Web-Sites [Amitay/Paris 2000; Zhang et al. 2002] bzw. allgemeine Textcorpora unterschiedlichster Art sein. Der Erfolg von MDS hängt stark von der Kohärenz der zusammenfassenden Dokumente ab (vgl. die Referenzen in [Schiffman/Nenkova/McKeown 2002]), aber auch, wie eine *Summarizing*-Kohäsion und -Kohärenz aus den aus verschiedenen Texten extrahierten und somit isolierten Sätzen erstellt werden kann [Barzilay/Elhadad/McKeown 2002].

Auch das *Cross-Lingual Summarizing/Extracting* gewinnt in zunehmend internationaler und damit mehrsprachiger Wissenschaft an Bedeutung [Sudo et al. 2004] oder das *Summarizing* von elektronischer Kommunikation, z.B. eMail-Threads [Lam et al. 2002; Wan/McKeown 2004] oder Chat Logs [Zhou/Hovy 2005], und von Interaktions-/Gesprächssituationen [Zechner 2001]. Methodisch vielversprechend, auch mit Blick auf das hier speziell interessierende *Diskurs-Summarizing*, sind Ansätze des *Summarizing* auf der Grundlage von Hypertexttechnologie (z.B. das *HyperGen Summarization Tool*), wobei einzelne *Summaries* hypertextuell über ausgezeichnete/typisierte Links vernetzt werden<sup>4</sup>.

Insgesamt ist also weltweit und nicht nur in der Informationswissenschaft angesichts fortschreitender internationaler Kommunikation ein verstärktes Interesse an *Summarizing*-Verfahren auszumachen (vgl. den Sammelband [Mani/Maybury 1999], [Endres-Niggemeyer 1999; 2004] und [Endres-Niggemeyer et al. 1993]), die sich zudem keineswegs mehr auf das klassische Eintext-*Abstracting* beschränken. Dem trägt z.B. auch die aktuelle Informationspolitik in Deutschland Rechnung: In der Bekanntmachung des BMBF zur Förderung von „eScience und vernetztes Wissensmanagement“ wird angesprochen, dass Verfahren der Integration, Extraktion, Komprimierung und Abstrahierung für die „medial unterstützte Nutzernavigation in komplexen, auch mehrdimensionalen Informationsbeständen“ für den Erfolg des vernetzten Wissensmanagements wichtig sind<sup>5</sup>.

## 3 *Summarizing* im Kontext des kollaborativen Wissensmanagement in der Lehre

Zusammenfassungen bereitzustellen und somit Produktion und Nutzung von Wissen zu begünstigen, kann auch als Leistung des Wissensmanagement angesehen werden. Wissensma-

---

<sup>4</sup> Vgl. die Folien von Kavi Mahesh in:  
[http://crl.nmsu.edu/Research/Projects/minds/core\\_summarizer/talk/](http://crl.nmsu.edu/Research/Projects/minds/core_summarizer/talk/)

<sup>5</sup> <http://www.bmbf.de/foerderungen/3179.php>

nagement, bislang überwiegend verstanden als Koordination des z.B. in einer Forschergruppe oder einer Organisation(seinheit) vorhandenen Wissens zur Unterstützung der Produktion neuen Wissens, bezieht sich zunehmend auf den Prozess der Erzeugung von Wissen selber, vor allem in kollaborativ arbeitenden Gruppen [Kuhlen 2004b]. Dieser Paradigmenwechsel trägt dem Rechnung, dass die Produktion neuen Wissens immer mehr auf dem Diskurs, auf der Kommunikation mit heterogenen, oft genug in getrennten Räumen arbeitenden Partnern beruht – begünstigt durch die Entwicklung der elektronischen Kommunikationsnetze.

Dieses neue Paradigma hat ebenfalls Konsequenzen für das traditionelle Gebiet der Fachinformation, das sich bislang dem dokumentarischen Auftrag verpflichtet fühlt(e). Dessen Aufgabe, Wissenschaftlern die Information über weltweit publiziertes Wissen zu jeder Zeit und an jedem Ort bereitzustellen, bleibt gewiss erhalten. Die neue Herausforderung besteht allerdings darin, dass sich Informationsarbeit (also alle Vorgänge des Sammelns, Aufbereitens, Ordners, Bereitstellens, Bewertens, Austauschens und Nutzens) in die kollaborativen Prozesse des Wissensaustauschs und der Wissensproduktion direkt einbindet. Informationelle Absicherung und kommunikativer Austausch werden stärker denn je integrale Bestandteile der kreativen wissenschaftlichen Arbeit.

In diesem Zusammenhang bekommt auch das *Summarizing* eine neue Funktion. Zum einen sollen sich *Summaries* ganz im Sinne des angesprochenen MDS (Abschnitt 2.4) auf vielfältige, durchaus auch multimodale Wissensobjekte beziehen, die in diesen kollaborativen Umgebungen entstehen. Zum andern besteht ein steigender Bedarf nach Orientierung in den zunehmend elektronisch ablaufenden Diskursen wissenschaftlicher Kommunikation. Diskurs-*Summaries* sollen hier diese Orientierungsleistung erbringen. Beides ist Gegenstand des Forschungsprojektes eAgora [Kuhlen et al. 2005b], wobei sich das MDS bei eAgora vor allem auf die Metadaten der verschiedenen Wissensobjekte beziehen wird. Gemeint sind damit Titel, vorhandene *Abstracts* und term-orientierte Indexate, einschließlich der formalen Beschreibungselemente, die in erster Linie auf die Verfasser, deren Institutionen und die verwendeten Publikationsmedien referenzieren und damit wichtige Orientierungshilfe leisten.

Wir beziehen Diskurs-*Summaries* hier auf Situationen des kollaborativen Wissensmanagements in der Hochschulausbildung, inwieweit also Summaries dazu beitragen können, sehr schnell komplex werdende Diskurse, z.B. in elektronischen Kommunikationsforen, für die Studierenden und für die die Diskurse bewertenden Dozenten übersichtlich und nachvollziehbar zu halten.

### **3.1 Information und Kommunikation im K3-System**

Kollaboratives Wissensmanagement wird in Konstanz in der Lehre seit einigen Semestern über das im Rahmen eines BMBF-geförderten Projektes entwickelte K3-System organisiert. Abb. 1 stellt die bislang mit K3 durchgeführten Kurse zusammen:

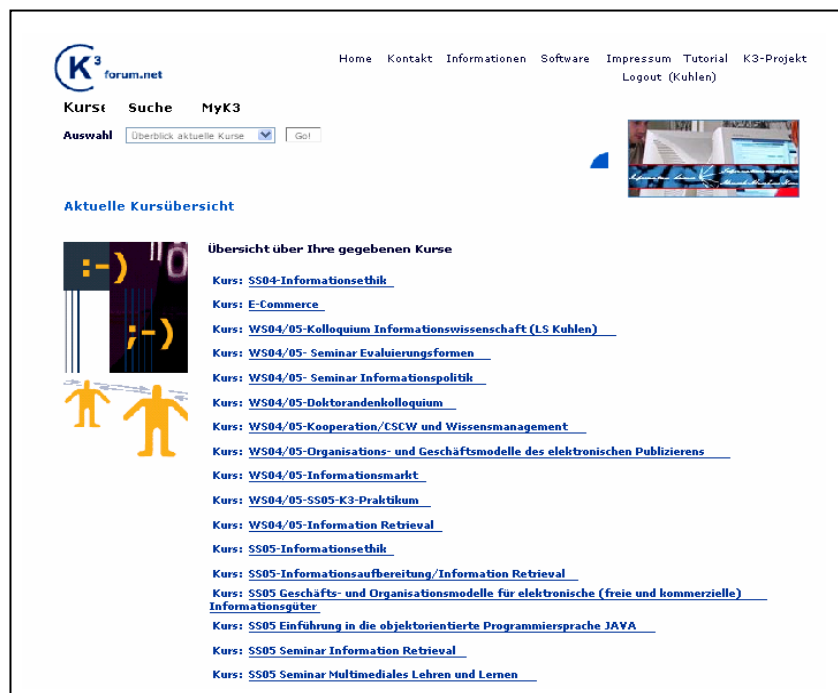


Abb. 1 Kurse mit K3

K<sup>3</sup> steht für Kollaboration, Kommunikation, Kompetenz<sup>6</sup>. Erläutern wir kurz diese drei zentralen K3-Konzepte.

- Kollaboration: Studierende arbeiten (kollaborativ und zeitlich begrenzt) in virtuell organisierten Gruppen an Themen des jeweiligen Kurse (z.B. „Wem gehören Wissen und Information“ in einem Kurs Informationsethik – vgl. Abb. 2). Die Gruppen bekommen dazu Arbeitsaufträge (vgl. links in Abb. 2) mit anfänglichen Textreferenzen (vgl. rechts in Abb. 2), die durch verschiedene Aufgaben weiter spezifiziert werden (vgl. links in Abb. 3).

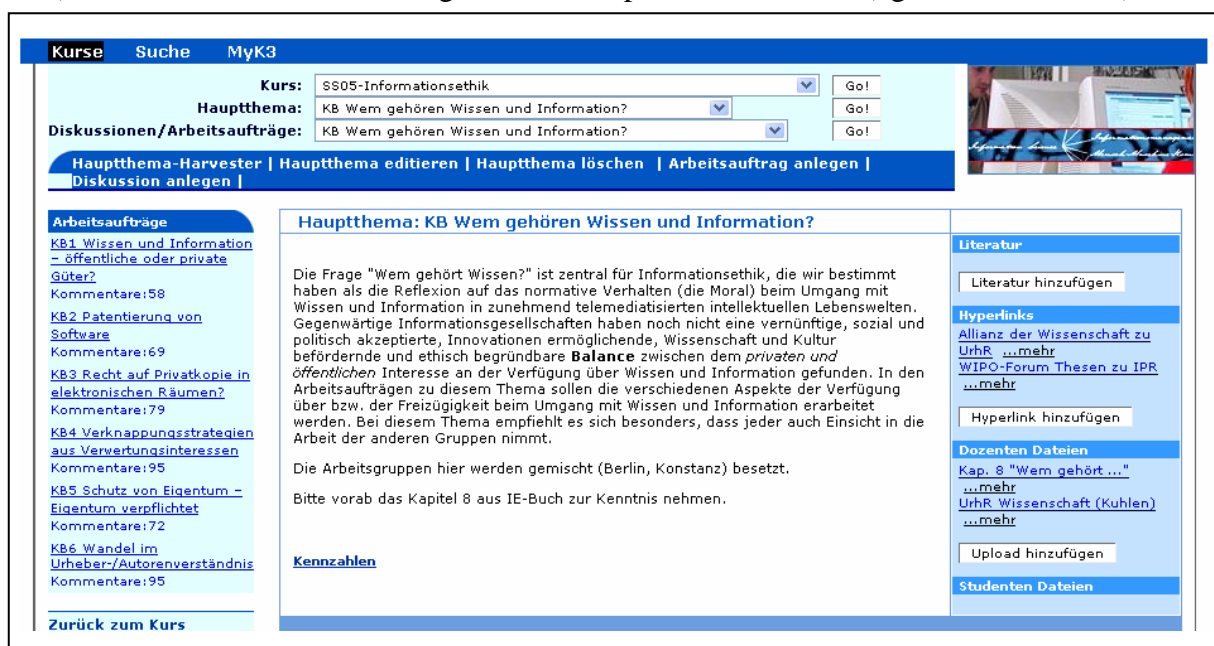


Abb. 2 Hauptthema aus Kurs Informationsethik mit Arbeitsaufträgen und Referenzobjekten

<sup>6</sup> Für einen Überblick über K3 vgl. [Kuhlen 2002 ; Semar et al. 2004 ; Kuhlen et al. 2005a]

Arbeitsaufgaben	User: (Anzahl)	Diskurstyp: (Anzahl)	Rolle: (Anzahl)	Datum:	Freitext:
<a href="#">KB4.0 Rollenorganisation</a> Kommentare:24 <a href="#">KB4.1 Schutzmechanismen</a> Kommentare:9 <a href="#">KB4.2 Pauschalierung</a> Kommentare:14 <a href="#">KB4.3 DRM vs Pauschalierung</a> Kommentare:38	bitte auswählen	bitte auswählen	bitte auswählen	bitte auswählen	
<a href="#">Zurück zum Hauptthema</a>  <b>NEU</b> Neue Beiträge /max.5 <a href="#">das "s" ist weg</a> Yan - 06.07.2005 06:51:25 <a href="#">Gute Arbeit!</a> DanielK - 05.07.2005 11:04:07 <a href="#">Pers. Meinung DanielK</a> <a href="#">Konstanz</a> DanielK - 05.07.2005 10:59:44 <a href="#">gute Präsentation</a> htwin - 05.07.2005 09:54:45	<b>Arbeitsauftrag KB4</b> Kuhlén 15.06.2005 07:30:37 Aufgaben:4 Kommentare:95 <b>Verknappungsstrategien aus Verwertungsinteressen</b>  <b>Startdatum:</b> 15.06.2005 00:00:00 <b>Enddatum:</b> 04.07.2005 00:00:00  <b>Moderator:</b> ally ; <b>Präsentator:</b> Yan ; <b>Rechercheur:</b> DanielK ; <b>Summarizer:</b> htwin ; <b>Arbeitsgruppe:</b> KB4 Verknappungsstrategien aus Verwerterinteressen: ally, Yan, htwin, DanielK,  <b>Bewertung:</b> Typ: Dozent Note: 1.3 Bemerkung: Teilnahmegrad: 0.93 Interaktionsgrad: 0.94 Studentbeiträge: 124 Absicherungsgrad: 0.33 Qualität: 1.7  Informationsgüter, die in kommerziellen Umgebungen erzeugt werden, werden in der Regel als Privateigentum angesehen. Daraus wird das Recht abgeleitet, diese Güter über Verknappungsstrategien, Schutzmechanismen zu sichern. Dieses Recht wird auch zunehmend durch rechtliche Maßnahmen (seit den WIPO-Gesetzen bis zur Gegenwart des deutschen UrhRs) geschützt. Welchen Einfluss haben Lizenzierungsformen und Verknappungsstrategien auf den Umgang mit Wissen und Information? Was passiert, wenn Wissens- und Informationsobjekte kaum noch gekauft, sondern durchgehend lizenziert werden?				

Abb. 3 Arbeitsauftrag mit spezifizierenden Aufgaben

Kollaboratives Arbeiten wird in der Ausbildung unter der Annahme gefördert, dass durch kollaboratives Arbeiten nicht nur mehr Wissen produziert wird, als es durch die Summe der individuell erzeugten Wissensbeiträge geschieht, sondern dass dadurch auch die Qualität der Ergebnisse gesteigert werden kann. Die Hypothese der Überlegenheit des kollaborativen Ansatzes des Wissensmanagements in der Ausbildung (gegenüber dem individualistischen Lernansatz) beruht auch darauf, dass die in vernetzten virtuellen Gruppen zwangsläufig bestehenden Wissensasymmetrien – Asymmetrien sowohl bezüglich des jeweils eigenen Wissens als auch der Kompetenz, über externe Ressourcen verfügen zu können – durch Austausch und Teilen produktiv ausgenutzt werden können [Paechter 2003]. Zudem fördert die Kommunikation fordernde Umgebung (in K3 sind es vor allem asynchrone elektronische Foren) laufendes *Feedback*, sowohl von Seiten des/der Kursleiter als auch von den anderen Mitgliedern in der Gruppe oder sogar von Studierenden des Kurses insgesamt. *Feedback* ist ein wichtiges Mittel, um Motivation und Leistungsbereitschaft der Studierenden zu entwickeln und hoch zu halten [Griesbaum 2004; Griesbaum/Rittberger 2005]. Auch die abschließende Beurteilung der jeweiligen Gruppen und der Rollen (vgl. unten) wird in K3 angezeigt (vgl. Abb. 4 und Abb. 5).

<b>Arbeitsauftrag KB4</b> <b>Verknappungsstrategien aus Verwertungsinteressen</b>  <b>Startdatum:</b> 15.06.2005 00:00:00 <b>Enddatum:</b> 04.07.2005 00:00:00  <b>Moderator:</b> ally ; <b>Präsentator:</b> Yan ; <b>Rechercheur:</b> DanielK ; <b>Summarizer:</b> htwin ; <b>Arbeitsgruppe:</b> KB4 Verknappungsstrategien aus Verwerterinteressen: ally, Yan, htwin, DanielK,  <b>Bewertung:</b> Typ: Dozent Note: 1.3 Bemerkung: Teilnahmegrad: 0.93 Interaktionsgrad: 0.94 Studentbeiträge: 124 Absicherungsgrad: 0.33 Qualität: 1.7	Kuhlén 15.06.2005 07:30:37 Aufgaben:4 Kommentare:95
---	---

Abb.4 Bewertung kollaborativer Gruppenarbeit

Rollen bewerten:				
Rolle	UserName	Startdatum	Enddatum	bereits bewertet?
Moderator	ally	18.06.2005		Typ: Dozent Punkte: 35 Note: 1.3
Präsentator	Yan	18.06.2005		Typ: Dozent Punkte: 30 Note: 1.7
Rechercheur	DanielK	18.06.2005		Typ: Dozent Punkte: 40 Note: 2.0
Rechercheur	Yan	15.06.2005	17.06.2005	Typ: Dozent Note: 0.0
Zusammenfasser	htwin	18.06.2005		Typ: Dozent Punkte: 37 Note: 1.3

Abb. 5 Bewertung der Rollenfunktionen

- **Kommunikation:** Für K3 wurde ein elektronisches Kommunikationsforum entwickelt, das über reichhaltige Funktionen die erwünschten kollaborativen Prozesse unterstützt. Kommunikation ist in K3 asynchron organisiert. Wenn auch für gewisse Aufgaben, vor allem zur Lösung von aktuellen Organisationsproblemen bzw. für die Darstellung institutioneller Sprechakte [Schulmeister 2005], synchrone Kommunikation (z.B. über *Chat* oder *Video-Conferencing*) besser geeignet sein mag, ist die asynchrone Organisation von Wissen auch für eine gemeinsame Produktion von Wissen besser geeignet. Diskurse beruhen auf fundierten (erarbeiteten, nicht spontan erstellten) Beiträgen (in K3 sogenannten *Diskursobjekte*), die zudem noch durch (in K3 sogenannte) *Referenzobjekte* (Referenzen auf einschlägige Publikationen oder Links auf entsprechende Websites; vgl. Abb. 10) abgesichert sein sollen.

Die Kommunikation wird auch dadurch gefördert und gesteuert, dass die Gruppenteilnehmer Rollen übernehmen müssen. In K3 sind bislang im Einsatz: Moderator, Rechercheur, *Summarizer* und Präsentator (vgl. Abb. 4 und Abb. 5). Uns wird hier natürlich vor allem der *Summarizer* interessieren.

Weiterhin wird Kommunikation durch die diskurs-syntaktische und -semantische Kennzeichnung der einzelnen Beiträge, z.B. Kennzeichnung als These, Frage bzw. Kennzeichnung des semantischen Bezugs (vgl. Abb. 6 und Abb. 7), strukturiert. Diese Kennzeichnungen können entsprechend für ein strukturiertes, template-orientiertes *Summarizing* eingesetzt werden. Wir gehen darauf in Abschnitt 4.2.2 näher ein.



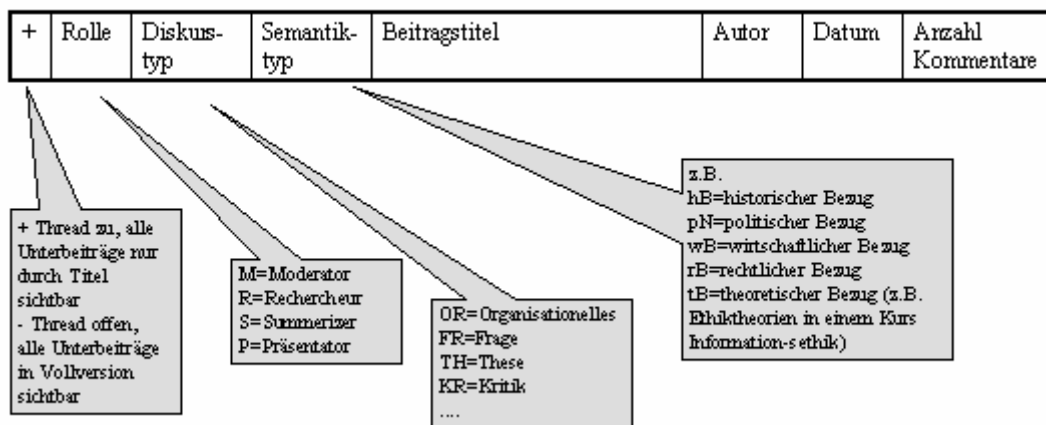


Abb. 6 Auszeichnung/Typisierung von Diskursobjekten in K2

In einer realen Kommunikationssituation sieht das z.B. so aus:

+	Aufgabe	<b>P2.1 Privacy - Datenspuren - Grundsätze der Sicherung</b>		Kuhlen	04.04.2005 18:02:32	Kommentare: 38
+	IM	ORGANISATIONELLES	<b>Lassen wir es angehen...Grundgerüst zum Diskurs</b>	Ponzo	11.05.2005 02:59:49	Kommentare:0
+	IM	NEUES THEMA	<b>Grundprinzipien von Garstka bzw. Kuhlen</b>	Ponzo	13.05.2005 02:03:55	Kommentare:8
+	ME	FRAGE	<b>Grundsatz der informationellen Symmetrie</b>	Ponzo	13.05.2005 02:20:21	Kommentare:7
+		ERGÄNZUNG	<b>informationelle Symmetrie überhaupt immer sinnvoll?</b>	Jens	16.05.2005 17:58:57	Kommentare:6
+		ERGÄNZUNG	<b>interessantes Interview</b>	Jens	16.05.2005 18:05:21	Kommentare:5
+	RE	ERGÄNZUNG	<b>Konkretes Beispiel</b>	JakobH	16.05.2005 18:52:54	Kommentare:4
+	RE	ERGÄNZUNG	<b>wichtigste Treffer bei Heise</b>	JakobH	16.05.2005 19:03:20	Kommentare:0
+	IM	ERGÄNZUNG	<b>Also Erfassung von Distributionsdaten...</b>	Ponzo	16.05.2005 20:40:34	Kommentare:2
+	KR	KRITIK	<b>Kritik an kontextbezogener Werbung</b>	Jens	17.05.2005 22:41:07	Kommentare:0

Abb. 7 K3-Thread (Ausschnitt) mit Kennzeichnung/Typisierung der Diskursobjekte

- Kompetenz: K3 unterstützt als informationswissenschaftliches Vorhaben auch die Bildung von Informations- und Kommunikationskompetenz – beides Ziele, die, wie z.B. in der sogenannten stefi-Studie belegt [Klatt et al. 2001], in der Hochschulausbildung insgesamt stark vernachlässigt werden. Was wir unter Kommunikationskompetenz in K3 verstehen, sollte durch die beiden Abschnitte zu Kollaboration und Kommunikation ausreichend angedeutet sein. Informationskompetenz [Kuhlen 2003] in K3 wird zum einen dadurch gefördert, dass die Studierenden angehalten werden (und dies auch explizit als Leistung bewertet wird), ihre Beiträge durch Referenzobjekte (s. oben) abzusichern (vgl. Abb. 8). Explizit zuständig dafür ist der jeweilige Inhaber der Rollenfunktion Rechercheur, aber auch alle anderen sollen sich um das Einbringen von Referenzobjekten kümmern. Ein Kriterium für die Leistung der Gruppe insgesamt ist somit der Absicherungsgrad, definiert als  $n(\text{studentische Referenzobjekte})/m(\text{studentische Diskursobjekte insgesamt})$ .

ERGÄNZUNG JakobH 16.05.2005 19:03:20 Kommentare:0

**wichtigste**  
**Treffer bei**  
**Heise**  
 Recherche  
 Diskurstyp: Ergänzung

Ich möchte hier noch einmal die interessantesten Treffer des Heise-Newsletters wiedergeben. Das ist sicherlich besser, so dass ihr nicht alles lesen mässt.

Kommentar anlegen    Kommentar bewerten

**Hyperlinks**  
 Datenschützer: Googles  
 Mail-Service in Deutschland  
 unzulässig ...mehr  
 Datenschützer erhöhen Druck  
 auf Google ...mehr  
 CFP: Heftige Kritik an Googles  
 Gmail ...mehr

Abb. 8 Referenzobjekte in K3 (einzufügen nur im Kontext eines Diskursobjekts, nicht in diesem selber; über eine Harvesting-Funktion können so alle Referenzobjekte in der Wissensbasis des K3-Kurses zusammengestellt werden)

Informationskompetenz in K3 bedeutet aber auch die Fähigkeit, einen in K3 sehr schnell komplex werdenden Diskurs so zusammenzufassen, dass es den einzelnen Gruppenmitgliedern leicht fällt, den Überblick über den aktuellen bzw. dann finalen Stand des erarbeiteten Wissens zu behalten. *Summaries* werden in K3 (bislang) von den Studierenden selber erbracht. *Summaries* spielen als spezielle Leistung der Studierenden zum Nachweis von Informationskompetenz im Konstanzer K3-System eine wichtige Rolle. Gehen wir darauf etwas näher ein.

### 3.2 *Summarizing* in K3

Regelwerke für die Zusammenfassung von Diskursen existieren bislang nicht als Teil der Informationsmethodik. Da zudem nicht davon ausgegangen werden kann, dass die Teilnehmer an K3-Kursen elementares Wissen über *Abstracts* und *Abstracting* haben [Kuhlen 2004a], wird in K3 eine Anleitung für die Rollenfunktion des *Summarizers* bereitgestellt:

#### **Aufgabe**

Die Aufgabe des *Summarizers* (Zusammenfassers) im K3-System besteht darin, den Diskussionsverlauf und das Ergebnis eines Arbeitsauftrags zusammengefasst wiederzugeben. Diese „Zusammenfassung“ muss an einer geeigneten Stelle in K3 als Kommentar, der mit dem Beitragstyp „Zusammenfassung“ und, falls der Autor die Rolle des *Summarizers* inne hat, mit dem Rollenvermerk „*Summarizer*“ gekennzeichnet und eingegeben werden. Ziel dieser „Rollenfunktion“ ist die Beförderung von informationsmethodischer Kompetenz. Das Ergebnis der Arbeit des *Summarizers* ist somit eine informative Zusammenstellung der Ergebnisse des kollaborativen Arbeitens.

#### **Was ist eine Zusammenfassung?**

Eine Zusammenfassung ist die Darstellung der wesentlichen Ergebnisse und Schlussfolgerungen einer Diskussion oder von Teilen einer Diskussion. Sie wird im allgemeinen zum Ende einer Diskussion erstellt. Sie kann aber auch bereits im Laufe einer Diskussion als Zusammenfassung des bisherigen Diskussionsverlaufs, quasi als Zwischenergebnis, erstellt werden.

Zusammenfassungen sollen kurz und informativ sein. Die Leistung besteht a) in der Synthese der Argumente (was ist erreicht worden?), b) in der Konzentration auf das Wesentliche und c) in der sprachlichen Aufbereitung. Die Zusammenfassung soll sich eigener Bewertung enthalten, kann aber - vor allem bei den Zwischenzusammenfassungen - auch auf offen gebliebene Probleme hinweisen. Die Zusammenfassung soll sich nicht auf organisatorische Fragen oder Metadiskussionen beziehen, sondern im Sinne des Wissensmanagements nur auf die erarbeiteten Inhalte.

## Was ist zu tun?

Zusammenfassen (darstellen) der Ergebnisse der Diskussion zu einem Arbeitsauftrag in K3. Der Umfang sollte eine Seite (bzw. 2000 Zeichen) nicht überschreiten.

Wöchentlich ist eine Zwischenzusammenfassung zu erstellen. Sie soll eine informative Zusammenstellung der Ergebnisse des kollaborativen Arbeitens in der jeweiligen Woche darstellen.

Zum Ende des Arbeitsauftrags ist eine Endzusammenfassung (Schluss-, Ergebniszusammenfassung) zu erstellen, sie soll sich auf den gesamten Arbeitsauftrag beziehen. Sie dient damit auch als Inputmaterial für den Präsentator.

Eine Zusammenfassung kann (soll) auch "laufend" erstellt werden - wenn der Summarizer das Gefühl hat, dass eine Diskussion zu einem Teilthema zu einem gewissen Ende gekommen ist.

Die Zusammenfassung ist zu strukturieren, damit der Leser sich schnell einen Überblick verschaffen kann.

Die Arbeit des *Summarizers* ist mit dem Moderator und Präsentator zu koordinieren.

Die Zwischenzusammenfassung und die Endzusammenfassung sind jeweils mit einem „sprechenden“ Titel zu versehen.

Die Zusammenfassung muss als Kommentar mit dem Beitragstyp „Resultat“ versehen werden. Ebenso muss der *Summarizer* dieses Kommentars angeben, ob er die Rolle des (eentlichen) *Summarizers* zur Zeit inne hat, dies geschieht durch das Anklicken der entsprechenden „Klickbox“ beim Eingeben der Zusammenfassung. Dies ist u.a. deshalb notwendig, damit eine automatische Punktevergabe möglich ist.

Die Zusammenfassung kann zusätzlich mit einem externen Tool erstellt und als File "*uploaded*" und damit zu einem Referenzobjekt werden.

Die in K3 eingestellte Zusammenfassung kann (und sollte) im weiteren Verlauf des Kurses von allen Teilnehmern des Kurses (also nicht nur von den Mitgliedern der jeweiligen Arbeitsgruppe) kommentiert werden.

### Kriterien zur Bewertung der Rolle *Summarizer*

S1 Abdeckungsgrad der Gesamtzusammenfassung – inwieweit sind die zentralen Themen angesprochen; inwieweit bezieht sich die Zusammenfassung wirklich auf die Diskussion?

S2 Qualität der Zusammenfassung(en) - Validität der Argumentation in der/den Zusammenfassung/en

S3 Zwischenzusammenfassungen - Vorhandensein von Zwischenzusammenfassungen (z.B. in Abhängigkeit von den Aufgaben oder von der zeitlichen Entwicklung)

S4 Form der Zusammenfassung(en). (Strukturiertheit, Fehlerfreiheit, Lesbarkeit/Verständlichkeit)

**Gewichtung:** S1 5; S2 4; S3 2 und S4 3

### 3.3 *Summarizing*-Leistungen der Studierenden

Trotz der relativ detaillierten Vorgaben für die zu erstellenden *Summaries* ist es für Studierende offenbar außerordentlich schwierig, *Summaries* von Diskursen zu erstellen. Nach den Erfahrungen im Kurs Informationsethik (SS 2005)<sup>7</sup> gibt es eine große Spannbreite in Intensi-

---

<sup>7</sup> Im Kurs Informationsethik mussten in 3 K3-Phasen über jeweils 3 Wochen zu insgesamt 5 Hauptthemen (*Digital Divides*, *Privacy*, Menschen-/Kommunikationsrechte, Wissensökologie und „Wem gehört Wissen?“) insgesamt 23 Arbeitsaufträge in virtuellen Gruppen (jeder Studierende musste an drei Gruppen, also 3 Arbeitsaufträgen teilnehmen) bearbeitet werden. Jeder Arbeitsauftrag wurde durch durchschnittlich 4 Aufgaben spezifiziert – insgesamt also fast 100. Dieses starke instruktionalistische Moment im ansonsten konstruktivistischen Lernansatz verdeutlicht den immensen Aufwand, den ein K3-Dozent vorab erbringen muss (ganz zu schweigen von dem Aufwand für Feedback und Bewertung). Keine Rede also davon, dass innovative (kollaborative) *e-Learning*-Verfahren als Rationalisierungsinstrumente mit Blick auf Lehrende angesehen werden können.

tät und Extension der erstellten *Summaries*. Das reicht von einer fast vollständigen, hintereinander aufgelisteten Darstellung der Diskursobjekte, über tabellenartige Strukturversuche bis zu knappen, wenige Zeilen umfassenden Zusammenfassungen der für wesentlich gehaltenen Beiträge bzw. zur stichwortartigen Auflistung der für zentral eingeschätzten Themen.

Die jeweiligen *Summarizer* haben so gut wie gar nicht die in K3 möglichen Strukturierungsmöglichkeiten für ihre *Summaries* genutzt (vgl. Abschnitt 4.2.2). Allenfalls wurden verschiedentlich die in den Aufgaben enthaltenen, vom Dozenten formulierten Fragen als Gliederung für die *Summaries* verwendet. Die *Summary*-Leistungen beruhten auf den persönlichen Einschätzungen der *Summarizer*. Es ist in manchen *Summaries* unverkennbar (und als Bias nachvollziehbar), dass den eigenen Beiträgen – auch der jeweilige *Summarizer* war angehalten, sich an den Diskursen aktiv zu beteiligen – ein höherer Stellenwert eingeräumt wurde als den Beiträgen der anderen (zuweilen wurde das auch von den anderen Teilnehmern kritisiert).

Wenn in K3 an intellektuell erstellten *Summaries* festgehalten wird – entsprechend dem Lernziel des Erwerbs an Informationskompetenz –, ist vorgesehen, *Summaries* über eine Wiki-Funktion von der Gruppe insgesamt erstellen zu lassen. Dies käme dem kollaborativen Ansatz entgegen.

## 4 *Summarizing*-Möglichkeiten aus K3

Wie zu Beginn erwähnt, sind *Abstracting*- bzw. *Summarizing*-Leistungen weder ausschließlich auf textuelle Referenz-/Wissensobjekte bezogen noch müssen sie als sprachliche Texte realisiert werden. Es ist also zweckmäßig, wie auch in der Textlinguistik üblich, von einem weiteren Textbegriff und damit auch von einem weiteren *Summary*-Begriff auszugehen, der sich auf die kohärente Darstellung eines Wissenszusammenhangs bzw. der Gesamtmenge der in einer definierten kommunikativen Situation vorkommenden Diskursobjekte bezieht, wobei die Darstellung unabhängig von der medialen Realisierung ist. Entsprechend kann auch eine Hypertextstruktur als Visualisierung von vernetzten Wissensobjekten oder eine Tabelle als Text bzw. als *Summary* angesprochen werden.

### 4.1 *Summarizing* über K3Vis-Hypertextstrukturen

K3 bietet, wie in Abb. 9 angedeutet, die Möglichkeit, sich die Diskursstrukturen über das zugeordnete System K3Vis (Visualisierung von Kommunikationsstrukturen in K3) anzeigen zu lassen. Davon haben die Studierenden allerdings wenig Gebrauch gemacht<sup>8</sup>. Tatsächlich sind

---

Die *Summaries* sollten sich auf die Diskurse zu den (23) Arbeitsaufträgen beziehen. Da die Diskurse sich aber in der Regel auf der Ebene der Aufgaben, nicht des gesamten Arbeitsauftrags, bewegten, haben es manche *Summarizer* vorgezogen, entgegen der Arbeitsanweisung, die *Summaries* auf der Ebene der Aufgaben zu belassen. Es liegen also insgesamt mehr *Summaries* als Anzahl Arbeitsaufträge vor. Von Zwischenzusammenfassungen wurden nur spärlich Gebrauch gemacht. Das wurde nach Rückfrage als zu zeitaufwändig angesehen. Auch hat es – wie an sich vorgesehen – kaum eine Diskussion über die dann in K3 eingestellten *Summaries* gegeben. Allenfalls wurde Zustimmung signalisiert bzw. – selten – eine Kritik an der Einseitigkeit der *Summaries* geäußert. Angesichts der knapp bemessenen Zeit für die Diskurse und der unmittelbar nach Ende der Diskurszeit in einer Präsenzveranstaltung vorgesehenen Präsentation der Ergebnisse, die sich auch auf die Zusammenfassung bezog, sind diese Mängel verständlich.

<sup>8</sup> K3Vis basiert auf *Scalable Vector Graphics* (SVG) - dem Grafik-Standard von W3C für vektorbasierte Grafikanwendungen im Internet. Es gab bislang noch keine (produktiven) Web-Browser, die

aber durch diese grafisch repräsentierten Hypertextstrukturen, in denen frei navigiert werden kann und die in vorgesehenen Erweiterungen flexibel komprimiert bzw. aufgezoomt werden können, intuitiv einleuchtende Übersichten und damit Orientierungshilfen möglich. Weiterhin sind durch K3Vis Kommunikations-Cluster (also momentan besonders aktive Diskurse) oder Kommunikations-Deadlocks (unbeantwortete Fragen, nicht bearbeitete Hypothesen etc.) erkennbar, die z.B. von den Moderatoren zur Steuerung benutzt, aber auch von den Summarizer als Hinweise auf besonderes intensive Diskussionen bzw. defizitäre Diskussion benutzt werden können. Abb. 9 zeigt einen besonders intensiven Diskursausschnitt.

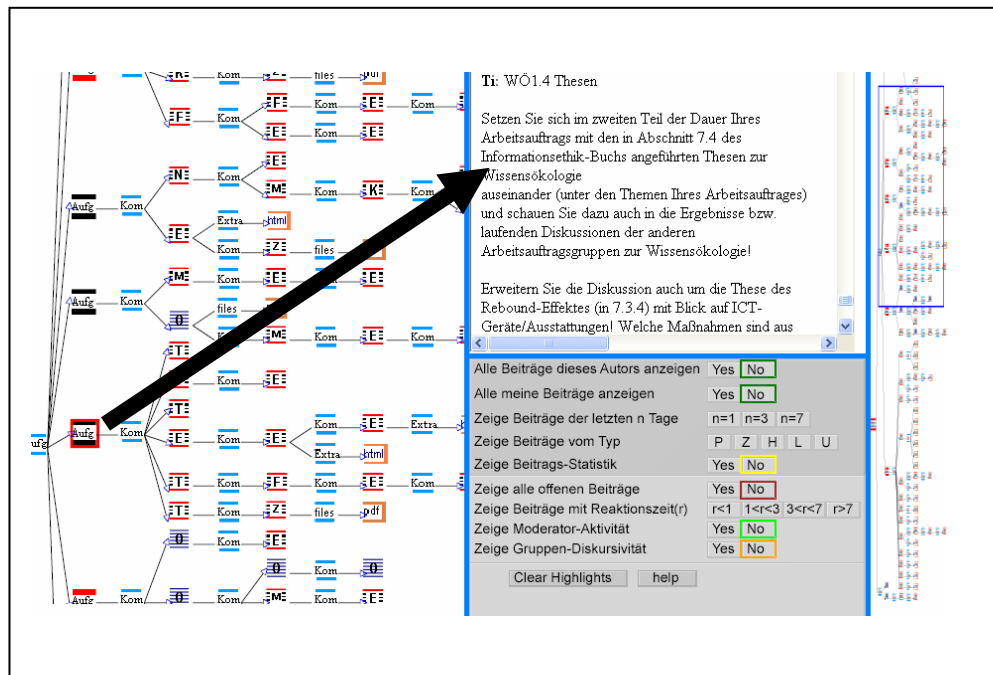


Abb. 9 Ausschnitt aus einer Visualisierung von Kommunikationsstrukturen (aus K3) – rechts ist der Ausschnitt aus einem größeren Thread markiert, der links aufgefaltet wird. In der Mitte wird der rechts markierte Beitrag als Text angezeigt.

## 4.2 Summaries auf der Grundlage von K3-Funktionen

### 4.2.1 Harvesting

K3 bietet die Möglichkeiten, auf den oberen K3-Ebenen (Gesamtkurs, Hauptthema, Arbeitsauftrag) sich die während des Diskurses eingetragenen und den jeweils Diskursobjekten zugeordneten Referenzobjekte (Literaturhinweise, Links auf Websites, in K3 hereingeladene Dateien) insgesamt anzeigen zu lassen<sup>9</sup>. Der entsprechende Kurs-/Hauptthema-

---

SVG native unterstützen. Deshalb war für die Anzeige von SVG ein *plugin* notwendig, das bislang nur auf dem Internet Explorer realisiert werden konnte. Das war für die meisten Studierenden offenbar zu restriktiv.

<sup>9</sup> Eine ähnliche Leistung wird in K3 durch die Report-Funktion erbracht, durch die die individuellen Beiträge eines Kursteilnehmers in einer Datei (PDF) zusammengestellt werden. Weiterhin gibt es in K3 eine differenzierte Filterfunktion, durch die auf der Ebene der Arbeitsaufträge a) die Beiträge des jeweiligen Arbeitsgruppenmitglieds, b) die Beiträge entsprechend dem Diskurstyp (vgl. Abschnitt 4.2.2), c) die Beiträge der jeweiligen Rolle, d) die Beiträge, spezifiziert nach Datum und e) die Beiträge spezifiziert nach einzelnen Stichwörtern zusammengestellt und dann entsprechend ausgewertet werden können.

/Arbeitsauftrags-*Harvester* sammelt also die entsprechenden Referenzobjekte zusammen und stellt sie in einem gesonderten Eintrag zusammen (vgl. Abb. 10).

Diese *Harvester*-Leistungen gehen in K3 bislang nicht über die tabellarische Auflistung der jeweiligen Referenzobjekte hinaus. Sie sind jedoch z.B. bei weiterer Auswertung der den Referenzobjekten zugeordneten Metadaten für *Summmarizing*-Leistungen vielfältig verwendbar.

Harvester zum Kurs: **SS05-Informationsethik**  
 / Hauptthema: **K2 Digital divide(s)**  
 / zum Arbeitsauftrag (Diskussion): **KN2 DD2 Ausprägungen der digital divide(s) aus der Sicht der Länder des Nordens/Westens**



Alle Literaturangaben:			
<b>Alle Hyperlinks:</b>			
<a href="#">Wissenskluft</a>	...mehr Infos	Stand/Eintrag am: 04.06.2005	referenziert von: dzweri4
<a href="#">Does the Internet open up opportunities for disabled people?</a>	...mehr Infos	Stand/Eintrag am: 07.06.2005	referenziert von: dzweri4
<a href="#">How People with Disabilities Use the Web</a>	...mehr Infos	Stand/Eintrag am: 07.06.2005	referenziert von: dzweri4
<a href="#">Minitel - The resistible rise of French videotex</a>	...mehr Infos	Stand/Eintrag am: 10.06.2005	referenziert von: Jetter
<a href="#">"www.fortschritt.fr"</a>	...mehr Infos	Stand/Eintrag am: 10.06.2005	referenziert von: Jetter
<a href="#">(N)onliner-Atlanten</a>	...mehr Infos	Stand/Eintrag am: 01.06.2005	referenziert von: Ponzo
<a href="#">Digitale Teilung</a>	...mehr Infos	Stand/Eintrag am: 01.06.2005	referenziert von: Ponzo
<a href="#">Aktionsbündnis für barrierefreie Informationstechnik</a>	...mehr Infos	Stand/Eintrag am: 08.06.2005	referenziert von: Ponzo
<a href="#">Barrierekompass</a>	...mehr Infos	Stand/Eintrag am: 08.06.2005	referenziert von: Ponzo
<b>Alle Uploads:</b>			
Präsentation <a href="#">Ergebnispräsentation</a>	von: Franz Wanner	Stand/Eintrag am: 14.06.2005	referenziert von: Ponzo
Präsentation <a href="#">Fertige Präsentation</a>	von: Franz Wanner	Stand/Eintrag am: 14.06.2005	referenziert von: Ponzo
Externe Texte <a href="#">disability and digital divide</a>	von: institute of social and ethical accountability	Stand/Eintrag am: 08.06.2005	referenziert von: dzweri4
Zusammenfassung <a href="#">Zusammenfassung Teil 1</a>	von: Jetter	Stand/Eintrag am: 12.06.2005	referenziert von: Jetter
Zusammenfassung <a href="#">Zusammenfassung Komplett</a>	von: Jetter	Stand/Eintrag am: 14.06.2005	referenziert von: Jetter
Externe Texte <a href="#">Landes-Behindertengleichstellungsgesetz</a>	von: Land Baden-Württemberg	Stand/Eintrag am: 08.06.2005	referenziert von: Ponzo
Externe Texte <a href="#">Gesetzesbegründung</a>	von: Land Baden-Württemberg	Stand/Eintrag am: 08.06.2005	referenziert von: Ponzo

Abb. 10 *Harvester*-Leistung auf Arbeitsauftragsebene (Ausschnitt für Links und *Uploads* zum Arbeitsauftrag 2 (*Digital Divides* aus der Sicht der Länder des Nordens/Westens) des Hauptthemas *Digital Divides*)

#### 4.2.2 Verwendung der Auszeichnungen der K3-Diskursobjekte

In K3 müssen, wie erwähnt, die Beiträge im Diskurs (also die Diskursobjekte) entsprechend ihrer Funktion im Diskurs ausgezeichnet werden<sup>10</sup>. Nicht gekennzeichnete Beiträge werden von K3 nicht akzeptiert.

In K3 wird bislang zwischen 4 Klassen von Diskurstypen unterschieden, die eine quasi text-syntaktische Funktion haben<sup>11</sup>:

<sup>10</sup> Kennzeichnungen der Beiträge sind in Kommunikationsforen auf vielfältige Weise möglich. Als Minimalform werden häufig para-sprachliche Kennzeichnungen über Smileys oder Emoticons verwendet. In elektronischen Diskursen können in Anlehnung an Searle's Sprechakttheorie illokutive Kennzeichnungen wie *representatives*, *directives*, *commissives*, *expressives*, *declaratives* verwendet werden, da sie auf Grund ihrer pragmatischen (illokutiven) Funktion ergänzende Kommentare stimulieren können [Searle 1969]. [Schulmeister 2005] schlägt Sprechaktkennzeichnungen vor wie: *Kommunikativa* (sagen, sprechen, reden, fragen, antworten, widersprechen, zustimmen, zugeben), *Konstative* (behaupten, beschreiben, mitteilen, erklären, bemerken, versichern), *Repräsentativa* (denken, meinen, hoffen, fürchten, lieben, hassen, mögen, wünschen, verbergen), *Regulativa* (auffordern, befehlen, bitten, verlangen, ermahnen, verbieten, erlauben), *Operativa* (definieren, ableiten, schließen, begründen, klassifizieren, bezeichnen, identifizieren) und *Institutionelle Sprechakte* (begrüßen, danken, gratulieren, wetten, verkündigen, verurteilen, bezeugen), ohne allerdings auszuführen, wozu diese Kennzeichnungen verwendet werden sollen.

<sup>11</sup> In K3 sind zur Kennzeichnung von Diskursobjekten auch semantische Inhaltstypen vorgesehen, die, anders als die quasi neutralen textsyntaktischen Auszeichnungen, kurs-spezifisch definiert werden. Für den Kurs Informationsethik können z.B. vorgesehen werden: *ethiktheoretischer* Bezug (in dem entsprechenden Beitrag wird Bezug genommen auf allgemeine Ethiktheorien oder andere theore-

## 1. Organisation des Diskurses

### **Diskurstyp Organisationelles (OR)**

Beiträge des Typs OR enthalten Metainformationen. Sie beziehen sich z.B. darauf, wie die Arbeit organisiert werden soll, können auch allgemeine soziale, emotionsstiftende Funktion haben. Sie sind für den Erfolg des Diskurses sicherlich nicht unwesentlich, treiben ihn aber nicht direkt weiter. Sie sollten daher nicht unbedingt als „off topics“ bezeichnet werden – eine typische Kennzeichnung, vor allem in Mailing-Listen, mit der man sich quasi entschuldigt, dass man nicht genau zum Thema der Liste redet.

## 2. Initialisierung des Diskurses

### **Diskurstyp Frage (FR)**

Eine Frage (FR) kann einen Diskussionsstrang eröffnen, kann aber natürlich auch jederzeit innerhalb eines laufenden Diskursstranges verwendet werden.

### **Diskurstyp These (TH)**

Eine These (TH) soll die Position des Beitragenden deutlich machen. Thesen (wie Hypothesen) können durchaus tentativ formuliert werden (und nähern sich damit zuweilen der Frage an). Die These sollte eindeutig formuliert und in der Regel durch (ein) Referenzobjekt/e belegt werden. Die Auszeichnung TH kann auch für Hypothesen verwendet werden. Formal gilt hier das Gleiche wie für Thesen. Auch Thesen und Hypothesen sollen in erster Linie einen Diskussionsstrang eröffnen, können aber natürlich auch jederzeit innerhalb eines laufenden Diskursstranges verwendet werden.

### **Diskurstyp Neues Thema (NE)**

Bei dieser Kennzeichnung (NE) sollte beachtet werden, dass hier wirklich ein neuer Sachverhalt innerhalb eines laufenden Diskurses, z.B. bei der Bearbeitung einer Aufgabe, eingebracht werden soll. Die Kennzeichnung NE hat also hohe diskursstimulierende Funktion.

## 3. Erweiterung des Diskurses

### **Diskurstyp Ergänzung (ER)**

Mit der Kennzeichnung Ergänzung (ER) wird direkt an einen vorausgegangenen Beitrag angeknüpft, und zwar in der Regel prinzipiell und durch neue Argumente unterstützend. Durch ER werden also auch Diskurstypen wie Modifikation, Unterstützung, neues Argument, ... abgedeckt.

---

tische Begründungen, wie Utilitarismus, kategorischer Imperativ, Naturrechtslehre, religiöse Überzeugungen, Menschenrechte,...), *historischer* Bezug (in dem entsprechenden Beitrag wird Bezug genommen auf Vorgänge, Argumentationen, Moralvorstellungen etc., die in der Vergangenheit, in vorelektronischer Zeit, entstanden sind. Durch den historischen Bezug können gegenwärtige Phänomene auch relativiert bzw. in einen größeren Kontext gestellt werden), *ökonomischer* Bezug (In dem entsprechenden Beitrag wird Bezug genommen auf ökonomische Argumente und Positionen (national, EU, international, WTO etc.), auch, welche Auswirkungen ethische Argumentationen speziell auf die Informationswirtschaft (aber auch die Wirtschaft insgesamt) haben können.), *politischer* Bezug (in dem entsprechenden Beitrag wird Bezug genommen auf politische Argumente und Positionen (national, EU, international), auch, welche Auswirkungen ethische Argumentationen auf die politische Realität haben können), *rechtlicher* Bezug (in dem entsprechenden Beitrag wird Bezug genommen auf rechtliche Regelungen (national, EU, international), wie Urheberrecht, Patentrecht, Datenschutz, bzw. es werden Argumentationen aus dem Rechtsbereich eingebracht).

### **Diskurstyp Kritik (K)**

Mit der Kennzeichnung Kritik (KR) wird direkt an einen vorausgegangenen Beitrag angeknüpft, und zwar in der Regel eher ablehnend, sei es destruktiv oder durch Hinzufügung neuer Argumente, die die Berechtigung der Kritik belegen/nahe legen und/oder die Diskussion im Sinne eines „thematischer Sprungs“ weiterführt. Kritik kann aber durchaus auch das bestehende Thema (eben mit neuen kritischen Rhema) weiterführen.

## 4. Ergebnis des Diskurses

### **Diskurstyp Resultat (RE)**

Mit der Kennzeichnung Resultat (RE) wird ein Diskussionsstrang in einem laufenden Diskurs zu einem (vorläufigen) Ende gebracht. Mit dieser Kennzeichnung wird also das Ergebnis der Diskussion gekennzeichnet. Der solchermaßen ausgezeichnete Beitrag kann dann natürlich auch wieder Anlass für eine weiterführende Diskussion geben. Der Diskurstyp RE ist insofern zu unterscheiden von der Leistung des *Summaries*, die in Wahrnehmung der Rolle des *Summarizers* entsteht, als der Diskurstyp RE von allen Mitgliedern der aktuellen Gruppe verwendet werden kann, wenn sie der Meinung sind, dass ein Diskussionsstrang zusammengefasst zu einem Ende gebracht werden kann/sollte. Der *Summarizer* wird aber in der Regel diesen Diskurstyp wahrnehmen, da er ja nicht nur die Zusammenfassung der Gesamtdiskussion eines Arbeitsauftrags leisten soll, sondern laufend versuchen soll, den Stand der Diskussion zusammenzufassen.

Die Beiträge sind entsprechend durch die Diskurstypen gekennzeichnet (vgl. z.B. in der Thread-Darstellung in Abb. 7 oder in der K3Vis-Darstellung in Abb. 9), so dass durch die Kennzeichnung eine Orientierung und pragmatische Anzeige (Aufforderungen zum Handeln) möglich ist. Wie erwähnt, können sich die Studierenden über die Filterfunktion Beiträge entsprechend dem Diskurstyp anzeigen lassen, z.B. alle Fragen oder alle neuen Themen, so dass sie spezifisch reagieren können<sup>12</sup>.

Es leuchtet unmittelbar ein, dass die Diskurstypen für ein strukturiertes *Summarizing* ausgenutzt werden können. Soll das *Summary* ausschließlich ergebnisorientiert sein, kann dafür der Diskurstyp *Resultat* verwendet werden. Die Gesamtheit der Diskurstypen (sinnvollerweise unter Ausklammerung des Diskurstyps *Organisationelles*) können ein *Summary-Template* bilden, dessen *Slots* entweder mit allen Beiträgen zu dem jeweiligen Diskurstyp oder mit entsprechend automatisch erstellten Zusammenfassungen aller Diskursobjekte des jeweiligen Typs gefüllt werden.

Um ein Beispiel zu geben: Im Arbeitsauftrag 1 zur Wissensökologie (Zum Konzept der Nachhaltigkeit mit Blick auf Wissen) wurden 11 Diskursobjekte mit dem Diskurstyp *These* ausgezeichnet. Über die Gesamtheit dieser 11 Beiträge (ein Text von ca. 5000 Zeichen) wurde die *Summarizing*-Funktion des in Anm. 2 erwähnten *Metafer Summarizer* aktiviert (mit einer Reduktion von 25%), die folgendes Ergebnis erbrachte:

Es stellt sich also eine zweite Frage: Möglicherweise wenden wir das Prinzip der schwachen Nachhaltigkeit an (Ressource A wird durch A+ ersetzt), doch bleibt zu klären, ob es unter den gegebenen Argumenten sinnvoll ist? Wie Du so schön sagst: "wir wären heute oft froh, wenn wir

---

<sup>12</sup> Eine zusätzliche Orientierung und pragmatische Leistung (Aufforderung zur Beteiligung) ist durch die Anzeige NEU gegeben, die nach dem Einloggen in K3 die aktuell neu hinzugekommenen Beiträge speziell auszeichnet.



das Wissen das unsere Vorfahren hatten und das wir durch die Technologieentwicklung aufgegeben haben, heute noch besitzen würden.

Möchte zur Veranschaulichung der These, dass "Ökologie" auf Wissen anwendbar ist, noch folgendes Bild einführen: bekannt sind doch ökologische Kreisläufe (z. B. Wasser fällt auf die Erde, verdunstet, fällt wieder hernieder, ermöglicht dadurch Wachstum usw.). Bei Störungen des Kreislaufes (z.B. durch starke Verschmutzung des Wassers, Verdichtung des Bodens ...) kommt es zu weitreichenden Störungen dieses und auch anderer, davon abhängiger Kreislaufsysteme.

Dieses Bild lässt sich meiner Meinung nach auch gut auf Wissen anwenden: Wissen generiert in einem Kreislauf neues Wissen, aber nur wenn es ungehindert "fließen" kann, wenn der Kreislauf nicht gestört wird, das heißt, wenn seine Verbreitung nicht eingeschränkt wird ...Bedarf es einer Kontrolle der Ressource Wissen, um eine auf langfristige Ziele orientierte Entwicklung zu gewährleisten? Ich denke schon, doch wer sollte diese Kontrolle ausüben? Ressourcen können nur genutzt werden, wenn ihr Vorkommen bzw. das Wissen, ohne sie handeln zu können, garantiert werden.

## 5 Ausblick

Wie zu Beginn erwähnt, wird in diesem Beitrag noch kein entwickeltes *Summarizing*-Verfahren für Diskurse vorgestellt. Die Arbeiten dazu in den Projekten/Systemen K3 und eAgora sind erst ganz am Anfang. Der Bedarf nach *Summarizing*-Leistungen in komplexen Kommunikationssituationen (Diskursen) sollte jedoch deutlich geworden sein, ebenfalls die Möglichkeit der Verwendung der aktuell in der Entwicklung befindlichen vielfältigen und vielversprechenden MDS-Verfahren (vgl. Abschnitt 2.4). Beim Diskurs-*Summarizing* handelt es sich um ein spezielles MDS mit in der Regel nicht sehr umfangreichen Texten (im K3-Fall sind es Diskursobjekte). Unter Ausnutzung der vielfältigen K3-Spezifikationen (Rollen, Diskurstypen, Inhaltstypen) sollten sich *Summarizing*-Leistungen ableiten lassen, die nicht unbedingt als Texte repräsentiert werden müssen, sondern, unter Weiterführung der K3Vis-Ansätze, auch zu flexiblen und adaptiven Visualisierungs- und hypertextuellen Navigationsformen führen können. Allerdings ist hierfür noch erheblicher Forschungs- und Entwicklungsaufwand erforderlich.

## 6 Referenzen

[Amitay/Paris 2000] Amitay, Einat; Paris, Cécile : Automatically summarising web sites – Is there a way around it? - <http://www.ics.mq.edu.au/~einat/publications/cikm2000.pdf>

[Barzilay/Elhadad/McKeown 2002] Barzilay, Regina; Elhadad, Noemie; McKeown, Kathleen: Inferring strategies for sentence ordering in multidocument news summarization. *Journal of Artificial Intelligence Research* 17, 2002, 35-55

[Endres-Niggemeyer 1999] Endres-Niggemeyer, Brigitte: Summarizing information. Springer-Verlag: Heidelberg etc. 1999

[Endres-Niggemeyer 2004] Endres-Niggemeyer, Brigitte: Automatisches Textzusammenfassen. In: Lobin, H.; Lemnitzer, L. (Hrsg.). *Texttechnologie*. Stauffenburg: Tübingen 2004, 407-432

[Endres Niggemeyer et al. 1993] Endres-Niggemeyer, B.; Hobbs, J.; Sparck Jones, K.: Summarizing text for intelligent communication. Technical Report Dagstuhl Seminar Report 79, 13.12-19.12.1993 -

[Griesbaum 2004] Griesbaum, Joachim: Curriculare Vermittlung von Informationskompetenz: Konzepte, Ziele, Erfahrungen eines experimentellen Retrievalkurses (Projekt K3), in: Bekavac, Bernard; Herget, Josef; Rittberger, Marc (Hg.): *Information zwischen Kultur und Marktwirtschaft*. Proceedings des 9. Internationalen Symposiums für Informationswissenschaft (ISI 2004), Chur, 6.-8. Oktober 2004. Konstanz: UVK Verlagsgesellschaft mbH, 2004, 23-37

[Griesbaum/Rittberger 2005] Griesbaum, Joachim; Rittberger, Marc: A collaborative lecture in information retrieval for students at universities in Germany and Switzerland. In: Proceedings of the World Library and Information Congress: 71st IFLA General Conference and Council. *Libraries - A voyage of discovery*. August 14th - 18th 2005, Oslo, Norway

[Hahn 1990] Hahn, Udo: Lexikalisch verteiltes Text-Parsing. Eine objektorientierte Spezifikation eines Wortexpertensystems auf der Grundlage des Aktorenmodells: Springer (Informatik-Fachberichte, 243 - Subreihe Künstliche Intelligenz): Berlin 1990

[Hahn/Reimer 1999] Hahn, Udo; Reimer, Ulrich: Knowledge-based text summarization: Saliency and generalization operators for knowledge base abstraction. In: Mani, Inderjeet; Maybury, Mark T. (eds.): *Advances in Automatic Text Summarization*. MIT Press 1999, 215-232

[Hammwöhner et al. 1989] Hammwöhner, Rainer; Kuhlen, Rainer; Sonnenberger, Gabriele; Thiel, Ulrich: TWRM-TOPOGRAPHIC. Ein wissensbasiertes System zur situationsgerechten Aufbereitung und Präsentation von Textinformation in graphischen Retrievaldialogen. *Informatik. Forschung und Entwicklung* 4, 1989, 89-107

[Klatt et al. 2001] Klatt, R.; Gavriilidis, K.; Kleinsimlinghaus, K.; Feldmann, M. u.a.: Nutzung elektronischer wissenschaftlicher Information in der Hochschulausbildung Barrieren und Potenziale der innovativen Mediennutzung im Lernalltag der Hochschulen. Endbericht Dortmund 2001. (<http://www.stefi.de/download/bericht2.pdf>)

[Kuhlen 1989] Kuhlen, Rainer: Information Retrieval: Verfahren des Abstracting. In: Batori, I. S.; Lenders, W.; Putschke (eds.): *Computational linguistics. An international handbook on computer oriented language research and applications*. Berlin etc.: de Gruyter

[Kuhlen 2002] Kuhlen, Rainer: K3 – Wissensmanagement über kooperative verteilte Formen der Produktion und der Aneignung von Wissen zur Bildung von konzeptueller Informationskompetenz durch Nutzung heterogener Informationsressourcen. Vorhabensbeschreibung Konstanz 2002 - <http://www.inf-wiss.uni-konstanz.de/FG/Forschungsprojekte/k3/vorhabensbeschreibung.pdf>

[Kuhlen 2003] Kuhlen, Rainer: Informationskompetenz und Vertrauen als Grundlage informationeller Autonomie und Bildung. Was bedeutet die fortschreitende Delegation von Informationsarbeit an Informationsassistenten? In: T. Christaller; J. Wehner (Hrsg.): *Autonome Maschinen*. Westdeutscher Verlag: Wiesbaden 2003, 186-206

[Kuhlen 2004a] Kuhlen, Rainer: Abstracting. In: R. Kuhlen; T. Seeger; D. Strauch (Hrsg.): *Grundlagen der praktischen Information und Dokumentation*. Saur-Verlag: München etc. 2004, 189-206

[Kuhlen 2004b] Kuhlen, Rainer: Change of paradigm in knowledge management - Framework for the collaborative production and exchange of knowledge. Plenary Session, 03 August 2003, of the World Library and Information Congress: 69th IFLA General Conference and Council, Berlin 2003. In: H.-C. Hobohm (ed.): *Knowledge Management. Libraries and Librarians Taking Up the Challenge*. IFLA Publications 108. K.G. Saur: München 2004, 21-38

[Kuhlen et al. 2005a] Kuhlen, Rainer et al.: K3 – an e-learning forum with elaborated discourse functions for collaborative knowledge management. Proceedings E-LEarn05 - World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, Vancouver Oktober 2005

[Kuhlen et al. 2005b] Kuhlen, Rainer et al.: eAgora – Modellierung von Wissensräumen (Antrag an das BMBF - [http://www.inf-wiss.uni-konstanz.de/FG/IV/eAGORA/antrag\\_eAgora.pdf](http://www.inf-wiss.uni-konstanz.de/FG/IV/eAGORA/antrag_eAgora.pdf))

[Lam et al. 2002] Lam, Derek et al.: Exploiting e-mail structure to improve summarization. - <http://www.dcs.shef.ac.uk/~saggion/ReadingGroup/Papers/TR2002-02.pdf>

[Luhn 1958] Luhn, H.P.: The automatic creation of literature abstracts. *IBM Journal* April 1958, 159-163

[Mani/Maybury 1999] Mani, I.; Maybury, M.T. (eds): *Advances in automatic text summarization*. MIT Press 1999

[Mori 2002] Mori, Tatsunori: Information gain ratio as term weight: The case of summarization of IR results. In: *Proceedings of the 19th International Conference on Computational Linguistics*. Taipei 2002

[Paechter 2003] Paechter, M.: *Wissenskommunikation, Kooperation und Lernen in virtuellen Gruppen*. Lengerich et al.: Pabst Science Publishers

[Reimer 1990] Reimer, U.: Automatic acquisition of terminological knowledge from texts. In: *Proc. 9th European Conf. on Artificial Intelligence*, 1990, 547-549

[Schiffman/Nenkova/McKeown 2002] Schiffman, Barry; Nenkova, Ani; McKeown, Kathleen: Experiments in multidocument summarization. *HLT '02 San Diego 2002* - <http://www1.cs.columbia.edu/nlp/newsblaster/papers/hlt02-dems.pdf>

[Semar et al. 2004] Semar, Wolfgang; Griesbaum, Joachim; König-Mistic, Jagoda; Lenich, Andreas; Schütz, Thomas: K3 – Wissensmanagement über kooperative verteilte Formen der Produktion und der Aneignung von Wissen zur Bildung von konzeptueller Informationskompetenz durch Nutzung heterogener Informationsressourcen - Stand und Aussichten. In: Hammwöhner, Rainer; Rittberger, Marc; Semar, Wolfgang (Hg.): *Wissen in Aktion - Der Primat der Pragmatik als Motto der Konstanzer Informationswissenschaft*. Konstanz: UVK Verlagsgesellschaft mbH, 2004. S. 329 – 347

[Searle 1969] Searle, J. R.: *Speech acts*. Cambridge University Press: Boston, MA 1969

[Siddharthan et al. 2004] Siddharthan, Advaith; Nekova, Ani; McKeown, Kathleen: Syntactic simplification for improving content selection in multi-document summarization. In: *Proceedings of the 20th International Conference on Computational Linguistics*. Genf 2004

[Schulmeister 2005] Schulmeister, Rolf – *Der Diskurs*. Powerpoint für LearnTec 2005

[Sonnenberger 1988] Sonnenberger, Gabriele: Flexible Generierung von natürlichsprachigen Abstracts aus Textrepräsentationsstrukturen. In: *Wiener Workshop Wissensbasierte Sprachverarbeitung*. 4. Österreichische Artificial Intelligence Tagung. Wien 1988. Springer Informatik Fachberichte 176: Heidelberg etc. 1988, 72-82

[Sudo et al. 2004] Sudo, Kiyoshi; Sekine, Satoshi; Grishman, Ralph: Crosslingual information extraction system evaluation. In: *Proceedings of the 20th International Conference on Computational Linguistics*. Genf 2004

[Wan/MacKeown 2004] Wan, Stephen; MacKeown, Kathleen: Generating overview summaries of ongoing email thread discussions. In: *Proceedings of the 20th International conference on Computational Linguistics*. Genf 2004

[White/Cardie 2002] White, Michael; Cardie, Claire: Selecting sentences for multidocument summaries using randomized local search. In: *Proceedings of the Workshop on Automatic Summarization*. Association for Computational Linguistics. Philadelphia, July 2002, 9-18

[Zhang et al. 2002] Zhang, Y.; Zincir-Heywood, N.; Milios, E.: World Wide Web site summarization. Technical Report CS-2002-8, Faculty of Computer Science, Dalhousie University, October 2002.

[Zechner 2001] Zechner, Klaus: *Automatic Summarization of spoken dialogues in unrestricted domains*. Dissertation, Carnegie Mellon University, School of Computer Science. Pittsburgh 2001

[Zhou/Hovy 2005] Zhou, Liang; Hovy, Eduard: Fine-grained clustering for summarizing chat logs - [http://www.isi.edu/~liangz/PAPERS/chi\\_thread\\_2005.pdf](http://www.isi.edu/~liangz/PAPERS/chi_thread_2005.pdf)