



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

Inhalterschließung betrachtet aus
allgemeinbildender Perspektive

verfasst von

Günter Lautner, Bakk. rer. soc. oec.

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer. nat.)

Wien, 2013

Studienkennzahl lt. Studienblatt:	A 190 884 423
Studienrichtung lt. Studienblatt:	190 Lehramtsstudium UniStG 884 UF Informatik und Informatikmanagement UniStG 423 UF Chemie UniStG
Betreuer:	Univ.-Prof. Dr. Wilfried Grossmann

Die Natur.

*Jedem erscheint sie in einer eigenen Gestalt. Sie verbirgt sich in
tausend Namen und Termen und ist immer dieselbe.*

(Johann Wolfgang von Goethe)

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Diplomarbeit selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe.

Die aus den fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde im In- noch im Ausland als Prüfungsarbeit vorgelegt.

Wien, am 16.4.2013

(Günter Lautner)

Abstract

Die traditionelle Betrachtung des Fachgebietes der Inhaltserschließung respektive Sacherschließung attribuiert ihm einen professionellen Charakter und die Assoziation dedizierter Fachkreise. Moderne digitale Technologien und durch sie geförderte Arbeitsstile führen zur Diagnose der Erweiterung betroffener Kreise bis hin zu Anwendern von Informationssystemen und des Internets. Eine dadurch effiziente Fragestellung behandelt die vorliegende Arbeit: Wie viel und welches Wissen der Spezialisten kann und muss in der breiten Anwendermasse verfügbar sein, und wie lässt es sich didaktisch im Unterrichtsfach Informatik auf allgemeinbildendem Niveau integrieren? Die Beantwortung zieht einerseits einschlägige Literatur zur Hilfe und andererseits fließen praxisnahe Erfahrungswerte aus einem Unternehmen ein. Die Erkenntnisse führen zur Empfehlung, das Stoffgebiet, idealiter im fächerübergreifenden Modus im Unterricht – aber auch in betrieblichen Weiterbildungen –, zu verorten. Dem Lehrstoff ist das grundlegende Thema Sprache immanent, welche nicht nur wesentliche Aufmerksamkeit aufgrund ihrer Phänomene wie Synonymie und Homonymie in der Inhaltserschließung erhält, sondern ein großes Potential in der Übertragung von Erlerntem in andere Bereiche – allen voran die Kommunikation – aufweist. Die Schule bietet durch ihr Wesen selbst, ihre Instrumente wie Lernmanagementsysteme, der Aufgabenstellung der vorwissenschaftlichen Arbeit, Förderung von Lese- und Schreibkompetenzen sowie der Behandlung der Berufsorientierung passende Ankerpunkte, um das Vorwissen und Alltagserfahrungen von Lehrenden und Lernenden zu berücksichtigen. Durch Kriterien fundamentaler Ideen und den positiven Konsequenzen ökonomischer Auswirkungen von effizientem Umgang mit Wissen ist die Platzierung des Themengebietes im Lehrstoff begründbar. Die alltäglichen Berührungspunkte und die Empfehlung der Integration im Lehrstoff werden durch Vorschläge für Unterrichtsmaterialien ergänzt.

Vorwort

Schüler lernen fürs Leben. Diese Vorstellung vertrete ich nicht nur aufgrund von Beobachtungen, Erfahrungen und meiner Lernbiografie. Vielmehr richte ich als zukünftiger Lehrer der Unterrichtsfächer Informatik und Chemie meine Ambitionen nach dieser Maxime. Ein großer Teil des Lebens spielt sich im Beruf ab. Ergo, Schüler lernen auch für ihre(n) zukünftigen Beruf(e). Für die Erstellung dieser Arbeit lagen in meinen Augen außerordentlich vorteilhafte und fruchtbare Rahmenbedingungen diesbezüglich vor: Im Zuge meines zur Diplomarbeit zeitgleichen Angestelltenverhältnisses in der Informatikabteilung des Unternehmens Julius Blum GmbH erhielt ich durch die Involvierung im Dokumentenmanagement-Team viele Inputs aus der realen Berufswelt sowie dem praxisnahen Dokumentenmanagement. Vor allem das Spannungsfeld zwischen Machbarkeiten im technischen Sinne und den Herausforderungen im organisatorischen Sinne stellt eine große Aufgabe in Unternehmen dar. Im hier bearbeiteten Thema Inhaltserschließung zählen zu diesen Aufgaben auch die Einschätzung und die Förderung von Kompetenzen der Mitarbeiter eines Unternehmens hinsichtlich des Umgangs mit Wissen, das in Dokumenten gespeichert ist. Welches Basiswissen wird notwendig, wenn Mitarbeiter sowie auch Internetnutzer anhand vorgegebener oder freier Schlagworte Dokumente inhaltlich erschließen (müssen)? Dieses Problemfeld bildete die primäre Triebkraft meiner Arbeit, die das Ziel hat, obige Maxime im Blickwinkel der Schulbildung zu unterstützen. Wenn Schüler fürs Leben lernen sollen, müssen Verantwortliche die Selektion des Lehrstoffs hinsichtlich geforderter Basiskenntnisse fürs Leben entsprechend berücksichtigen – und, nicht weniger wichtig, adäquat didaktisch aufbereiten.

Ich hoffe, mit dieser Arbeit die Brücke zwischen Beruf und Schule bezüglich des behandelten Themas geschlagen zu haben. Neben der Beantwortung der Forschungsfrage stellt das Artefakt eine Einstiegslektüre für Verantwortliche in Unternehmen und Schulen, die sich mit der inhaltlichen Erschließung von Dokumenten beschäftigen.

Für Feedback aus der Anwendung der Materialien und Anregungen für eine Weiterentwicklung hinsichtlich der Thematik Inhaltserschließung auf allgemeinbildendem Niveau stehe ich offen gegenüber.

Um eine leichtere Lesbarkeit des vorliegenden Dokumentes zu ermöglichen, verzichte ich auf die laufende Nennung der weiblichen und männlichen Form sowie auf Binnenmajuskeln. Durch die Angabe der männlichen Form schließe ich intentional stets auch die weibliche Form ein.

Inhaltsverzeichnis

Wichtige Hinweise zur Verwendung der Arbeit	1
1 Einleitung	2
1.1 Forschungsfrage.....	6
1.2 Aufbau der Arbeit.....	7
2 Einführung in die Inhaltserschließung.....	9
2.1 Definition, Einordnung und Abgrenzung.....	9
2.2 Was wird inhaltlich erschlossen? – Das Dokument.....	11
2.3 Aufgabe, Ziel und Kosten der Inhaltserschließung	14
2.4 Für wen wird erschlossen?	18
2.5 Kennzahlen des Nutzbarmachens	21
2.6 Teilprozesse der inhaltlichen Erschließung	27
3 Sprache als Medium.....	30
3.1 Sprachliches Zeichen und semiotisches Dreieck.....	30
3.2 Der Begriff.....	36
3.3 Konsequenzen sprachlicher Phänomene	40
4 Methoden der Inhaltserschließung.....	46
4.1 Automatische Methoden.....	47
4.2 Das Inhaltsverzeichnis	50
4.3 Das Abstract	51
4.4 Das Indexieren	53
4.5 Das Register	57
4.6 Der Thesaurus	60
4.7 Die Klassifikation	61
4.8 Das Tagging.....	64
5 Argumente für Schule und Allgemeinbildung.....	69
5.1 Wissensmanagement in der Schule.....	69

5.2	Metadaten im eLearning Bereich	70
5.3	Berufsorientierung	73
5.4	Vorwissenschaftliche Arbeit	74
5.5	Fundamentale Ideen	76
5.6	Lesen, Schreiben und Lernen lernen	78
5.7	Querschnittscharakter des sprachlichen Aspekts	81
5.8	Inhaltserschließung im „2.0“ – Kontext	83
5.9	Programmierung	85
6	Unterrichtsmaterialien.....	88
6.1	Material: Disambiguierung im Wikipedia.....	88
6.2	Material: Kennzahlen der Dokumentation.....	89
6.3	Material: Gegenstand – ein Ausschnitt der Wirklichkeit.....	90
6.4	Material: Mehrdeutigkeit und Unschärfe.....	91
6.5	Material: Inhaltsverzeichnis erstellen	91
6.6	Material: Register erstellen.....	93
6.7	Material: Tagging	93
6.8	Material: Abstract, Zusammenfassung.....	94
6.9	Material: Begriffsbildung.....	95
7	Schluss	96
I	Literaturverzeichnis	100
II	Anhang	105
II.1	Abkürzungsverzeichnis	105
II.2	Beispiel eines alten Kartenkatalogs	106
II.3	Beispiele für Produkte inhaltlicher Erschließung	107
II.4	Relationenkürzel im Thesaurus	109
II.5	Unternehmen Julius Blum GmbH.....	110
	Lebenslauf.....	111

Abbildungen

Abbildung 1: Einordnung der Inhaltserschließung	11
Abbildung 2: Informationsverlust und Ballast	25
Abbildung 3: Kompromiss Recall und Precision (Stock, 2006, S. 64)	27
Abbildung 4: Sprachliches Zeichen nach de Saussure (2001, S. 78)	31
Abbildung 5: Signifikat als kulturelle Einheit semantischer Systeme	34
Abbildung 6: Triadische Relation nach Eco	35
Abbildung 7: 'Computer' anhand des semiotischen Dreiecks illustriert	35
Abbildung 8: Beispiele für Merkmale Teller und Vase	36
Abbildung 9: Abstraktions- und Bestandshierarchie (Fugmann, 1999, S. 21)	38
Abbildung 10: Grundstruktur eines Informations- systems (Salton & MacGill, 1987, S. 11)	41
Abbildung 11: Beispiele für Synonymie (Bertram, 2005, S. 40)	42
Abbildung 12: Beispiel für Schreibweisenvielfalt (Bertram, 2005, S. 42)	42
Abbildung 13: Überblick über die Funktionen eines Informations- systems (Salton & MacGill, 1987, S. 12)	44
Abbildung 14: Auslegungen von "The man saw the pyramid on the hill with the telescope" (Vickery & Vickery, 2004, S. 201)	48
Abbildung 15: Beispiel für den abstrakten Aufbau eines Registereintrags (Kiel & Rost, 2002, S. 61)	59
Abbildung 16: konkretes Register-Beispiel (Kiel & Rost, 2002, S. 62)	59
Abbildung 17: Deskriptorensatz für 'Hochzeit' (Bertram, 2005, S. 214)	61
Abbildung 18: Narrow und Broad Folksonomy inkl. Häufigkeitsverteilung (Vander Wal, 2005a)	65
Abbildung 19: Bibsonomy.org Tag Cloud zum Artikel von Mathes (2004)	66
Abbildung 20: Die beliebtesten Tags in BibSonomy.org	66
Abbildung 21: Hashtags im Mikroblogging Dienst Twitter	67
Abbildung 22: Hashtags im Sozialen Netzwerk Google Plus	68
Abbildung 23: Implizites und explizites Wissen (Capaul & Seitz, 2011, S. 530)	69
Abbildung 24: Struktur LOM Metadatenstandard (e-teaching.org, 2012)	72
Abbildung 25: Suche in Inhaltsverzeichnissen der Vorarlberger Landesbibliothek	92
Abbildung 26: Karte eines Zettelkataloges	106
Abbildung 27: Beispielregister einer Website	107
Abbildung 28: Beispielregister mit Absätzen als DBE	108
Abbildung 29: Relationenkürzel eines Thesaurus (DIN 1463-1, 1987, S. 11)	109

Wichtige Hinweise zur Verwendung der Arbeit

Die vorliegende Arbeit beschäftigt sich im Kern mit dem Thema Inhaltser-schließung. Im Zuge der Beantwortung der Forschungsfrage wird ein primärer Bezug zum Unterrichtsfach Informatik hergestellt. Das Thema impliziert jedoch die Behandlung eines fundamentalen Aspekts: Die Natur der Sprache. Aus diesem Grund ist es durchaus legitim, den Blick von der Informatik weg, hin zu einem anderen Unterrichtsfach zu wenden. Legt eine Schule besonderen Wert auf die Recherche in Bibliotheken im Unterrichtsfach Deutsch, so vermag die dortige Behandlung der Inhaltser-schließung, aus einem weniger informatischen Blickwinkel, das Verständnis über die Entwicklung von Suchstrategien durchaus zu verbessern. Als fruchtbarste Verankerung sieht der Autor dieser Arbeit die fächerübergreifende Integration im Lehrstoff.

Da die Inhaltser-schließung aufgrund hoher Kosten durch Informationsverluste und Doppelarbeiten in Unternehmen ebenso eine die Allgemeinbildung betreffende Rolle spielt, eignet sich die Aufnahme der Inhalte dieser Arbeit in innerbetriebliche Schulungen. Hier existieren wiederum Zugänge möglicher Integration: In Schulungen der jeweiligen Informationssysteme (wie Dokumenten- und Contentmanagementsysteme), in persönlichkeitsbildende Schulungen, die kommunikative Kompetenzen oder die Sensibilisierung über den Wert des Unternehmenswissens fördern.

Ferner elaboriert die vorliegende Arbeit Vorschläge für Unterrichtsmaterialien. Es ist davon abzuraten, diese ohne vorherige Beschäftigung mit dem Thema Inhaltser-schließung anzuwenden, da das implizite Thema Sprache (sprachliches Zeichen, semiotisches Dreieck) großes Potential zu Transferleistungen im Lernenden in andere Bereiche (Kommunikation, ...) aufweist.

1 Einleitung

Unter dem Begriff Inhaltserschließung werden alle Methoden und Hilfsmittel zur inhaltlichen Beschreibung von Dokumenten zusammengefasst (vgl. Bertram, 2005, S. 18). Sie ist Teil der Dokumentation (engl. Information Retrieval), deren Zweck es ist, Dokumente und Informationen gezielt nutzbar und wiederauffindbar zu machen (vgl. Gaus, 2005, S. 11). Kiel und Rost (2002, S. 16ff.) nennen erste Versuche zur Erfüllung der Bedürfnisse Wissen wiederzubeleben, wiederzuentdecken und zu organisieren im Übergang vom Mittelalter zur Neuzeit. In der genannten Zeit beherrschte im Durchschnitt bei weitem nicht die Majorität der Menschen den Umgang mit Schriftsprache (vgl. Glück, 1987, S. 174ff.), in Deutschland des ausgehenden Mittelalters nur ca. fünf Prozent (vgl. Genuneit, 2001, S. 51). Nicht die Allgemeinheit war mit Fragestellungen, wie beispielsweise wo und wie Wissen in Schriften zur Behandlung einer aktuellen Aufgaben-, Frage- oder Problemstellung zu finden ist, konfrontiert. Die Gegenwart spiegelt ein teils konträres Bild wider. Der Schriftgebrauch entwickelte sich zum Konstituens der modernen Zivilisation. Das Wissen – und somit auch die Anzahl der Dokumente – wuchsen enorm an. *„Alle 20 Jahre wird so viel publiziert wie in allen Jahrhunderten davor.“* (Umstätter, 2004, S. 237). Indes blieb die Art der Fragestellung bezüglich des Erschließens konstant: *„[...] wie man sich gezielt und schnell einen Zugang zu Dokumentinhalten und eine Orientierung über sie verschaffen kann.“* (Bertram, 2005, S. 13). Es entwickelte sich das Fachgebiet Inhaltserschließung, die auch Sacherschließung genannt wird, im Laufe der Zeit zum Thema dedizierter Bereiche. Zu diesen zählen die einschlägigen Wissenschaften wie Informations-, Dokumentations- und Bibliothekswissenschaft und Künstliche Intelligenz (vgl. Kiel & Rost, 2002, S. 39) sowie themenassoziierte Berufsbilder wie das des Dokumentars, Bibliothekars oder Archivars (Gaus, 2002). Sie wird in der Bibliotheksarbeit *„[...] als primäre und spezifische Aufgabe des höheren Bibliotheksdienstes angesehen.“* (Lorenz, 1998, S. 9).

Erfordert die Gegenwart eine Blickwendung aus Sicht der Beurteilung, ob die Thematik ausschließlich als Spezialgebiet genannter Kreise verbleibt? Die Beantwortung dieser Fragestellung wird von den technologiegetriebenen

Veränderungen der gegenwärtigen Zeit genährt. Die digital vernetzte Welt mit ihrem breiten Repertoire an Kommunikationsmedien, die teils durch technologische Errungenschaften und hochgradige Vernetzung erst ermöglicht wurden, involviert viel umfassender Gruppen von Menschen als je zuvor. *„Insbesondere das World Wide Web hat Information Retrieval zu einem Thema gemacht, das (plötzlich) eine breite Öffentlichkeit betrifft und teilweise auch interessiert.“* (Ferber, 2003, S. 31). Im Speziellen trägt die Philosophie des Web 2.0 Paradigmas, welches die partizipative Einstellung in Web 2.0 Diensten betont (vgl. O'Reilly, 2005), bei, potentiell jeden Internetnutzer dazu zu befähigen, neben herkömmlichen Konsumierens nicht nur *„[...] Inhalte zu produzieren, sondern auch [...] inhaltlich zu erschließen.“* (Peters & Stock, 2008, S. 77).

Ein unterhalb dieses oberen Endes des Nutzerkreisspektrums Internetcommunity angesiedeltes Anwendungsfeld von Inhaltserschließungskompetenzen erscheint nicht weniger signifikant, denn ein Unternehmen kommt heute ohne Dokumentation seines Unternehmenswissens kaum aus. Betriebliches Informationsmanagement bedarf meist des Einsatzes von Dokumentenmanagementsystemen, in denen *„semantische Metadaten“* (Schütz, 2004, S. 340), zu denen *„[...] die klassischen inhaltserschließenden Attribute wie Volltextindizierungswerte, Schlagworte, Abstracts oder klassifizierende Attribute [gehören]“* (ebd.), Dokumente inhaltlich beschreiben. Digitale Dokumentationen speichern Informationen, die von Mitarbeitern zur Aufgaben- und Problemlösung genutzt und verwertet werden – eine *„[...] Sensibilisierung für die Notwendigkeit eines bewussten Umgangs mit Informationen [...]“* wird erforderlich, schreibt Herget (2004, S. 251), um letztendlich auch Wettbewerbsfähigkeit eines Betriebes zu gewährleisten (ebd., S. 253). Um die Qualität innerbetrieblicher Dokumentation aus Sicht der Erschließung sicher zu stellen, gewinnen für Unternehmen respektive deren Mitarbeiter Grundkenntnisse im Wissen über inhaltliche Erschließung fortlaufend mehr an Bedeutung – nicht nur um die Einsicht und Akzeptanz beim Ausfüllen obligatorischer inhaltsbeschreibender Eingabefelder im Mitarbeiter zu fördern, sondern um einen Perspektivenwechsel in das Informationen suchende Sub-

jekt (vgl. Bertram, 2005, S. 19), meist ein anderer Mitarbeiter, vollziehen zu können.

Umfassendes und detailliertes Wissen über inhaltliche Erschließung, ihre wissenschaftlichen und anwendungsbezogenen Zusammenhänge und ihre Methoden, verbleibt stets beim Spezialisten. Da die gegenwärtige Involvement, wie oben diagnostiziert, weitreichend und die Repräsentationsform inhaltlicher Erschließung das Medium Sprache ist, kristallisiert sich „[...] *das Bewußtsein für sprachliche Phänomene und ihre problematischen Folgen für das Retrieval*“ (Bertram, 2005, S. 46) unter anderem zum Gegenstand der Betrachtung, welcher hier als Vorschlag in den Geltungsraum des Allgemeinwissens ragen kann, heraus. Aus didaktischer Sicht strahlt eine Förderung des Verständnisses über Sprache und ihrer Natur, weil fundamental, und der Umgang mit Wissen in die meisten, wenn nicht alle Bereiche des Lebens. Kiel und Rost (vgl. 2002, S. 21) konstatieren die Notwendigkeit, Grundkenntnisse der Wissensorganisation als Kompetenz (engl. information literacy) in der Allgemeinbildung zu verankern, damit sie in der breiten Masse verstanden und angewendet werden können.

Die einem Unternehmen und seinen Mitarbeitern förderlichen Grundkompetenzen über Inhaltserschließung können durchaus in jeweiligen Schulungen der genutzten Anwendungen aufgebaut werden. Ökonomischer erscheint die Situation, wenn das Basiswissen und das Problembewusstsein über die Notwendigkeit inhaltlicher Erschließung und sprachlicher Phänomene bereits vom Mitarbeiter mitgebracht werden. Da Schüler von heute Mitarbeiter von morgen sind, kommt die Platzierung des Themas auf vorberuflichem Niveau der Wirtschaftlichkeit entgegen. Die innerbetrieblichen system- und anwendungsspezifischen Schulungsmaterialien können dann didaktisch an das Vorwissen der Mitarbeiter anknüpfen.

Obige Begründungen, warum ausgewählte Lehrstoffgebiete der Inhaltserschließung zur Allgemeinbildung zählend gesehen werden können, provozieren die Frage, wo der Lehrstoff effektiv ankern soll. Da digitale Technologien im betreffenden Themengebiet traditionelle Mittel wie Karteikarten und

Zettelkataloge (siehe Anhang II.2) antiquieren, bietet sich in der Bildung der Informatikunterricht an.

Eine Quellenrecherche zum Zeitpunkt des Verfassens dieser Arbeit ergab, dass wenige Unterrichtsmaterialien zum Thema zu finden sind. Ausgenommen sind im Ergebnis Materialien, die ihren Ursprung in dafür vorgesehenen Institutionen wie einschlägige universitäre Fakultäten oder Fachhochschulen haben. Diese vermitteln naturgemäß Wissen über ein allgemeinbildendes Niveau hinaus an.

Eine Sichtung betreffender Lehrpläne führt zur Erkenntnis, dass die Thematik nicht explizit angeführt wird. Die offen formulierten Anforderungen an den Unterricht am Beispiel des vom Bundesministerium für Unterricht, Kunst und Kultur (BMUKK, 2004, S. 1) herausgegebenen Lehrplans für den Informatikunterricht der AHS Oberstufe lassen den Interpretationsraum offen, die vorliegende Thematik unter folgendem Aspekt zu subsumieren:

„Schülerinnen und Schüler sollen Informationstechnologien soweit beherrschen, dass sie damit Aufgaben selbst mit großen Datenmengen bewältigen können. Sie sollen sich kooperative und kommunikative Arbeitsweisen unter Einsatz von Kommunikationstechnologien aneignen.“

Das Themenfeld der Inhaltserschließung existiert genuin in Gegenwart großer Datenmengen wie Bibliotheksbeständen, Wissensdatenbanken, Dokumentationen in Unternehmen und Organisationen oder Inhalte von Webanwendungen. Als alltägliche Aufgaben lassen sich oberflächlich das Ablegen und die Suche von Informationen identifizieren. Zudem fordert der genannte Lehrplan die „[...] Entwicklung einer wissenschaftlichen Arbeits- und Dokumentationsweise [...]“ sowie „[...] den sicheren Umgang mit Standardsoftware [...] zur Dokumentation, zur Publikation von Arbeiten [...]“ (BMUKK, 2004, S. 2). Eine Dokumentationsweise endet nicht mit der Erstellung eines Artefaktes durch ein geeignetes Werkzeug, sondern führt in einem höheren kooperativen und kommunikativen Kontext zu konsekutiven Überlegungen, wie es wiederauffindbar abgelegt beziehungsweise publiziert werden kann, respektive, wie es inhaltlich und formal zu erschließen ist. Artefakte der Inhaltserschließung finden sich in vielen alltäglichen Anwendungen -

von Registern in Büchern über Klassifikationen in Bibliotheken bis zur Möglichkeit des Taggings in Unternehmens- und Webanwendungen. Umso mehr dürfen sie Eingang in den Informatikunterricht finden, besonders unter der Prämisse, dass *„[...] der Informatikunterricht [...] beispielhaft für den sinnvollen Einsatz verfügbarer Technologien sein [muss]“*, wie im analysierten Lehrplan (ebd.) verankert.

Ferner liefert der Bezug des Verfassers dieser Arbeit zum Thema aus beruflicher Sicht einen weiteren Beweggrund, das Themenfeld Inhaltserschließung fundiert zu bearbeiten. Im Konkreten laufen aktuelle innerbetriebliche Fragestellungen in der Suchbarkeit von hunderttausenden Dokumenten in verschiedenen Systemen des weltweit agierenden Unternehmens Julius Blum GmbH in Höchst/Vorarlberg (siehe Anhang II.5) ein. Aufgrund der Erkenntnisse dieser Arbeit folgen im genannten Unternehmen Maßnahmen zur Verbesserung der aktuellen Situation durch Umsetzung von Erschließungsmethoden und Vorschläge der Integration abgestimmter Lerneinheiten in intern gehaltenen Schulungen für Mitarbeiter.

1.1 Forschungsfrage

Die vorgeschlagene Verortung der Thematik Inhaltserschließung im Informatikunterricht löst eine Folgefrage aus, die in diese Arbeit als leitende Forschungsfrage eingeht: Wie lässt sich eine Berücksichtigung der Thematik Inhaltserschließung im Lehrstoff eines allgemeinbildenden Niveaus im Detail begründen und welche Herausforderungen und Möglichkeiten induziert sein didaktisches Design für Unterricht, Schule und Betrieb?

Die Bearbeitung der Forschungsfrage erfolgt einerseits literaturgestützt und andererseits durch empirische Werte aus den praxisnahen Anforderungen des Wissens- und Dokumentenmanagements im oben genannten Unternehmen, welches durch seine Größe und den existierenden Anforderungen in den Fragestellungen als repräsentativ eingestuft werden kann. Das resultierende Ergebnis, ein die Forschungsfrage beantwortendes Dokument mit einer ausreichenden Einführung in die Thematik für Lehrende zum Aufbau eines notwendigen Wissensfundaments für Unterricht und Schulung sowie einer Sammlung von Unterrichtsmaterialien, adressiert folgende Zielgruppen:

Die Rezipienten der vorliegenden Arbeit bewegen sich im Bereich der Sekundarstufe I und II Schulen, Bildungsinstitutionen, welche informatische Grundkenntnisse lehren, sowie unternehmensinterne Schulungsleiter, die der Aufgabe nachgehen, Mitarbeitern im Zuge von Schulungen in betreffenden Systemen zum Thema Inhaltserschließung Wissen zu vermitteln. Zudem helfen die Kenntnisse auch den Verantwortlichen des Schulmanagements bei der Wahl und Umsetzung von Erschließungsmethoden in schulinternen Dokumentationen.

1.2 Aufbau der Arbeit

In seiner Gliederung folgt das vorliegende Dokument einem aufbauenden Charakter, indem zur Beantwortung der Forschungsfrage folgende Kapitel notwendig sind:

Neben den in der Einleitung allgemeinen Definition sowie Diagnose der Involvierung und Verortung des behandelten Themas unterstützt Kapitel zwei die Beantwortung der Forschungsfrage, indem a) ein notwendiges Basiswissen über die inhaltliche Erschließung erarbeitet wird, b) Zusammenhänge und Konsequenzen aufgezeigt werden und c) auf die allgemeinen didaktischen Herausforderungen eingegangen wird.

Kapitel drei vertieft das der Inhaltserschließung zugrundeliegende Kernthema Sprache. Aus ihrer aufgezeigten Natur werden erstens die Konsequenzen erörtert und die Notwendigkeit eines Problembewusstseins über sie behandelt. Zweitens liefert es hinsichtlich der gestellten Forschungsfrage, da Sprache ein Konstituens in unserer Gesellschaft bildet, Gründe zur Berücksichtigung im Lehrstoff.

Das vierte Kapitel offeriert einen Überblick über die hier relevanten Methoden der Inhaltserschließung, um a) den Aufbau einer Wissensplattform über Methoden der Inhaltserschließung für Lehrende zu ermöglichen und b) den Umfang der Unterrichtsmaterialien hinsichtlich der Fragestellung abzugrenzen.

Argumentationen für die Berücksichtigung des Themas in der Schule respektive im Unterricht elaboriert Kapitel fünf.

Als ein Resultat dieser Arbeit enthält Kapitel sechs erarbeitete Vorschläge für Unterrichtsmaterialien, die in einer Unterrichts- oder Schulungsplanung gezielt Eingang finden können.

Die pädagogischen und didaktischen Zugänge zu den einzelnen Aspekten der Inhaltserschließung werden nicht in einem dedizierten Kapitel behandelt, sondern in den einzelnen Kapiteln themenbezogen erarbeitet. Hier wird die wissenschaftliche Erkenntnis, dass sich das Aktivieren des Vorwissens der Lernenden als bedeutenden Vorteil für den Lernerfolg verhält, berücksichtigt (vgl. Krause & Stark, 2006, S. 38ff.; Mietzel, 2003, S. 305ff.). Dies fließt einerseits in die didaktischen Argumentationen in den einzelnen Kapiteln ein und liefert andererseits eine Grundlage für die Konstruktion der Unterrichtsmaterialien.

Um betont dem Ziel des allgemeinbildenden Aspekts Rechnung zu tragen, bleiben Betrachtungen wie Systemarchitekturen, Datenbankmodelle und Speicherformate in dieser Arbeit weitgehend unberücksichtigt – ausschließlich Kapitel „5.9 Programmierung“ richtet den Fokus auf technische Momente.

2 Einführung in die Inhaltserschließung

Das einführende Kapitel grenzt das Thema terminologisch und inhaltlich ab. Der Begriff Inhaltserschließung wird eingeordnet und ihr wesentlicher Gegenstand – das Dokument – in einem Diskurs bestimmt. Aufgaben und Ziele inhaltlicher Erschließung werden, auch unter ökonomischer Betrachtung, definiert. Wer die Akteure und Zielgruppen der Tätigkeit sein können, sowie grundlegende Kennzahlen und Teilprozesse der Inhaltserschließung sind erläutert. Dem Ziel, die Thematik der inhaltlichen Erschließung aus Perspektive des didaktischen Designs in Hinblick auf die in der Einleitung formulierte Forschungsfrage wird in den einzelnen Unterkapiteln Rechnung getragen. Da die Natur der Sprache der Inhaltserschließung größte Berücksichtigung beschert, ist sie dem dedizierten Kapitel „3 Sprache als Medium“ gewidmet.

2.1 Definition, Einordnung und Abgrenzung

Das Resultat der inhaltlichen Erschließung zeichnet sich durch den Teilprozess aus, bei dem Dokumente mit inhaltsbeschreibenden Informationen assoziiert werden. Die Beschreibung der Dokumente besteht dabei aus „*einzelnen Wörtern und/oder ganzen Sätzen*“ (Bertram, 2005, S. 18). Einzelne Worte kommen beim Indexieren (siehe Kapitel 4.4), ganze Sätze beim Abstracting (siehe Kapitel 4.3) zur Anwendung (vgl. ebd., S. 24). Im Bibliothekswesen, in dem auch der Begriff Sacherschließung neben Inhaltserschließung synonymisch Anwendung findet, erklärt Lorenz (1998, S. 10): „*Die Sacherschließung bedeutet [...] die Kennzeichnung des Inhalts bzw. seiner wesentlichen Elemente (in komprimierter Form) durch Zuteilen und Auswahl von Begriffen [...]*“. Gaus (2005, S. 13) identifiziert das Erschließen, neben dem Beschaffen und Erfassen, Speichern sowie Recherchieren, als Teilgebiet der Dokumentationsarbeit und nennt es indexieren, in dem es um „*[...] das Feststellen und Kennzeichnen des Inhalts einer Dokumentations-einheit mithilfe so genannter Deskriptoren [...]*“ geht, die im einfachsten Fall Schlagwörter sind. Die zum Gebiet der Inhaltserschließung gehörenden Artefakte wie Register, Schlagwortlisten oder Klassifikationssysteme finden sich aus Kiels und Rosts (2002, S. 38) Perspektive der Wissensorganisation im

Prozess des Ordners als „*Mittel der Wissensorganisation*“ (ebd., S. 57ff.) wieder.

Die obigen Beschreibungen aus der Literatur zeigen eine inkonsistente Begrifflichkeit in der Thematik und diverse Einordnung ihrer selbst auf. Wird das inhaltliche Erschließen bei Gaus (2002, S. 13) dem Indexieren gleichgesetzt, so differenzieren andere Autoren und zählen das Indexieren zu den Methoden der inhaltlichen Erschließung (vgl. Bertram, 2005, S. 67; vgl. Riggert, 2009, S. 6) beziehungsweise widmen ein ganzes Werk, namentlich – „*Inhaltserschließung durch Indexieren: Prinzipien und Praxis*“ (Fugmann, 1999) - dieser Methode. Die vorliegende Arbeit sieht das Indexieren als Methode der Inhaltserschließung und der Begriff Sacherschließung ist dem der Inhaltserschließung gleichgesetzt. Kapitel 4 beschäftigt sich mit den hier relevanten Erschließungsmethoden.

Von der inhaltlichen Erschließung unterscheidet sich die Formalerschließung, in der sich die Beschreibung von Dokumenten auf äußerliche formale Kriterien beschränkt, wie zum Beispiel der Verfassername, der Titel und die ISBN-Nummer (vgl. Strauch & Rehm, 2012, S. 181). Kiel und Rost sprechen auch von einer „*passiven Inhaltserschließung*“ (2002, S. 87), Bertram von der „*Formalbeschreibung*“ (2005, S. 52). Der wesentliche Unterschied dieser Erschließungstypen resultiert im Aufwand der Tätigkeit. Für das Bestimmen der Werte formaler Erschließungskriterien, wie zum Beispiel Autorennamen(n), muss der Inhalt des betreffenden Dokumentes nicht analysiert werden. Sie sind durch Angaben am Dokument üblicherweise vorgegeben und einfach zu übernehmen. Somit bedarf die formale Erschließung in der Regel keiner intellektuellen Leistung. Gemeinsam mit der inhaltlichen Erschließung liefern die Prozesse für ein Dokument Metadaten, also Daten über es. Gödert, Lepsky und Nagelschmidt (2011, S. 59) fassen unter dem Begriff Informationserschließung die Teilaufgaben der formalen und inhaltlichen Erschließung zusammen.

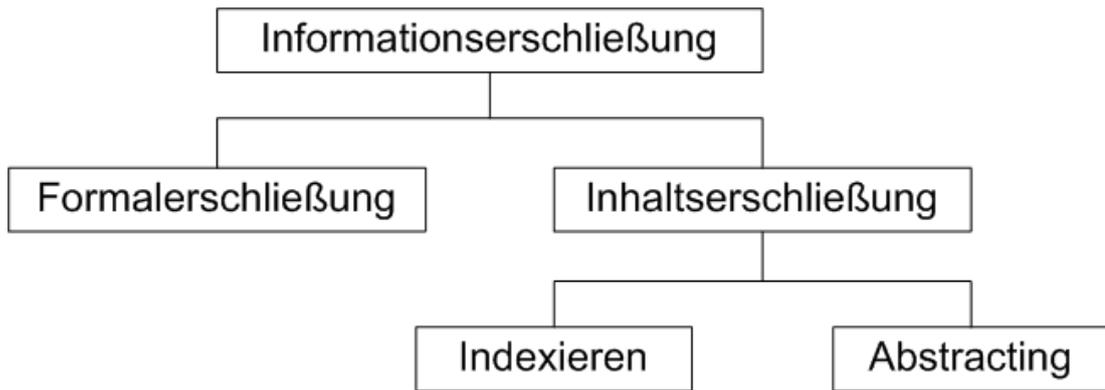


Abbildung 1: Einordnung der Inhaltserschließung

2.2 Was wird inhaltlich erschlossen? – Das Dokument

Der Begriff Inhaltserschließung wirft die Frage auf, wessen Inhalt der zu erschließende ist, oder anders gefragt: Was denn beinhaltet? Die obigen Definitionen aus der Literatur in Kapitel 2.1 nennen dafür Begriffe wie Dokument und Dokumentationseinheit. Jedoch: *„Was ein Dokument ist, diese Frage an einen Notar gerichtet wird eine andere Antwort hervorrufen als einem Historiker gestellt.“* (Riggert, 2009, S. 52).

Riggert verdeutlicht die Denotations- sowie Konnotationsvielfalt des Begriffs Dokument aus Unternehmenssicht, indem er es im Konventionellen *„[...] ein Schriftstück [nennt], in dem die Gedanken von Menschen, mit Hilfe von herkömmlichen Zeichen und Symbolen, dargestellt werden“* und damit die Assoziation zum papierbasierten Artefakt feststellt, welches jedoch auch elektronisch repräsentierbar ist. Er weitet dann den Umfang des Dokumentbegriffs aus und verweist auf die Möglichkeit, in elektronischen Dokumenten multimediale Inhalte wie Animationen, Video und Sound aufnehmen zu können, eine Vorstellung, die die Papierform des Dokumentes übertrifft. (Vgl. Riggert, 2009, S. 52f.).

Wird der Gedanke fortgesetzt, so muss für den Dokumentbegriff kein beinhaltendes elektronisches Schriftdokument für multimediale Inhalte existieren, sondern es lässt sich ein multimedialer Inhalt ebenfalls als Dokument erkennen, wobei diese zum Beispiel Bild- oder Tondokumente sind. Letztendlich führt der Dokumentbegriff in seinem Umfang in die ISO 5963 (1985)

Norm, in der das Dokument als *„Any item, printed or otherwise, which is amenable to cataloguing or indexing“* definiert ist, mit der Anmerkung *„This definition refers not only to written and printed materials in paper or microform versions (for example books, journals, diagrams, maps) but also non-printed media (for example machine-readable records, films, sound recordings), and three-dimensional objects or realia used as specimens.“* Laut dieser Definition können alle realen Objekte, aber auch Entitäten ohne räumliche Ausprägung als Dokument, welches inhaltlich beschreibbar ist, dienen.

Gaus (2005, S. 11f.) formuliert den Dokumentbegriff allgemeiner als Dokumentationseinheit und listet zum Verständnis des Begriffs unter anderem folgende Beispiele auf:

- ein Buch,
- ein Zeitschriftenband,
- ein Aufsatz in einer Zeitschrift,
- eine einzelne Buchseite,
- ein einzelner Satz (z.B. Bibelvers, Gesetzestext),
- eine einzelne Angabe, ein einzelner Messwert, ein einzelnes Datum oder Faktum,
- das Datenblatt eines Werkstoffs,
- ein Bild einer Gemäldesammlung,
- eine Personalakte,
- eine Krankenakte.

Die Fachsprache weist ferner den Begriff „dokumentarische Bezugseinheit“ (DBE) auf, welcher sich im Dokumentationsprozess auf ein Dokument oder gar einen Teil eines Dokumentes, wie zum Beispiel die Artikel einer Zeitung (vgl. Bertram, 2005, S. 21), oder auf die Fahrzeuge eines Fuhrparks oder Maschinen einer Fabrik (vgl. Gaus, 2005, S. 12) bezieht, wobei die Zeitungsartikel beziehungsweise die Maschinen als DBE und die Zeitung bzw. die Betriebsbücher der Maschinen als Dokument angesehen werden können. Beide Autoren vereinfachen die Terminologie in ihren Werken und vernach-

lässigen den Begriff DBE zugunsten von Dokument (Bertram) bzw. Dokumentationseinheit (Gaus). In der vorliegenden Arbeit wird die DBE ebenfalls einem Dokument gleichgesetzt, die Dokumentationseinheit soll jedoch wie nach DIN 31631-1 (1984, S. 2) die Entität sein, welche im Dokumentationsprozess stellvertretend für eine dokumentarische Bezugseinheit (Dokument) steht. Sie entspricht somit dem das Dokument repräsentierenden Datensatz in Informationssystemen oder der Karteikarte im Traditionellen (vgl. Bertram, 2005, S. 21). Dokumente erhalten meistens eine eindeutige Kennzeichnung – Signatur genannt. Im Bibliothekswesen signalisiert sie zudem den Standort des Dokumentes (vgl. Gaus, 2005, S. 15). Liegen Dokumente elektronisch vor, navigieren Nutzer üblicherweise über einen Verweis (engl. Link) zum Dokument.

Die Anwendung des Begriffs Dokument für beliebige Entitäten in Bezug auf die inhaltliche Erschließung stellt den Lehrenden vor eine fachdidaktische Herausforderung. Im Vorwissen der Schüler, welches den Begriff mit hoher Wahrscheinlichkeit noch nicht thematisiert und in einem alltäglichen Bezug konstruiert, bleibt der Zugang „any item“ vorerst unverfügbar oder er wirkt irritierend. Am ehesten entspricht im Schüler die Vorstellung eines Dokumentes die der Urkunde, zum Beispiel die Geburtsurkunde in seiner Dokumentenmappe, die des Reisedokumentes wie dem Pass, oder die des Textverarbeitungsdokumentes, welches in der populärsten Anwendungen mit dem Menüpunkt ‚Datei/Neu/Leeres Dokument‘ erstellt und im Dateisystem als ‚.doc(x)-Datei‘ vorliegt. Es werden im schulalltäglichen Sprachgebrauch weder das Email, das Zeugnis, das ausgeteilte Arbeitsblatt im Unterricht, das Schulbuch, das Klassenbuch, der Testbogen, die an die Wand projizierte Overheadfolie, der vorgeführte Dokumentarfilm, ... als Dokumente bezeichnet noch als solche vom Schüler erkannt.

Auch im erwachsenen Lernenden darf nicht bedingungslos vorausgesetzt werden, dass der Dokumentbegriff aus Sicht der Dokumentation gewusst wird oder verstanden ist. Denn im Allgemeinen wird mit dem Begriff Dokument „[...] eine hohe inhaltliche Qualität und eine rechtliche Bedeutung verbunden“ (Riggert, 2009, S. 52) oder in ihm „schwingt [...] mit, dass es sich

um eine wertvolle Sache und nicht um einen Gebrauchsgegenstand handelt.“ (Gaus, 2005, S. 12), wie auch der Duden Online¹ unter anderem die Bedeutung „*Beweisstück, Zeugnis*“ des Wortes Dokument mit den Beispielen „*der Film ist ein erschütterndes Dokument des Krieges*“ und „*etwas als historisches Dokument aufbewahren*“ anführt.

Somit lässt sich für die Lehre über Inhaltserschließung eine unabdingbare Voraussetzung als Ziel im Unterricht identifizieren, welches das Schaffen eines gemeinsamen Verständnisses und Vorwissens einfordert. Ein Verständnis, das dem Lernenden erlaubt, alles „*[...] was als Träger inhaltlicher Daten in Betracht kommt*“ (Bertram, 2005, S. 21), als Dokument anzusehen.

2.3 Aufgabe, Ziel und Kosten der Inhaltserschließung

Das Bestücken von Dokumenten mit beschreibenden Informationen über sie führt bei isolierter Betrachtung der Tätigkeit ausschließlich zu Aufwänden. Der Nutzen erwächst erst durch die Kenntnis des übergeordneten Prozesses - der Dokumentation. Diese wird im Englischen als Information Retrieval (IR) bezeichnet und verfolgt das Ziel, Informationen nutzbar und zurückgewinnbar zu machen. Die vier Teilgebiete der Dokumentation umfassen das Beschaffen und Erfassen, das Erschließen, das Speichern und das Recherchieren von Informationen. Die Tätigkeiten der Inhaltserschließung ordnen sich der Dokumentation respektive dem Information Retrieval als Teilprozess unter. Im Gesamtprozess wird erkennbar, dass Dokumentation weit über das reine Sammeln, Ordnen und Ablegen von Informationen oder Dokumenten hinausgeht. Das Erschließen ist als Vorarbeit zur Möglichkeit des Recherchierens zu sehen, die das Suchen und Finden in einem befriedigenden Ausmaß erst ermöglicht. (Vgl. Gaus, 2005, S. 11ff.).

Aus dem Blickwinkel des Ordnungsprinzips Gleiches-zu-Gleichem definieren Gödert, Lepsky und Nagelschmidt (2011, S. 70) als Ziel der Inhaltserschließung, dass „*[...] ein Dokument bei inhaltlich gleichen – oder zumindest ähnlichen – Dokumenten eingeordnet werden kann*“, unabhängig davon, ob es sich um eine physische Ordnung, wie Bücher im Bücherregal, oder virtuelle, wie Datensätze in einer Datenbank, handelt.

¹ URL: <http://www.duden.de/rechtschreibung/Dokument> [09.10.1012].

Fugmann (1999, S. 11) bezeichnet die Aufgabe der inhaltlichen Erschließung „[...] die Wiederverwertung von vorliegender Information dadurch zu erleichtern, daß sie diesbezügliche Aufzeichnungen wiederauffindbar macht, und zwar ohne daß der Mensch hierbei bezüglich seines Suchvermögens überfordert wird und ohne daß hierbei untragbar große Kosten entstehen“, wobei er unter dem menschlichen Suchvermögen die begrenzten menschlichen Eigenschaften – Geduld, Gedächtnisleistung, Aufmerksamkeit – des Suchenden versteht.

Gaus (2005, S. 14) konstatiert, dass der Arbeitsaufwand für die Inhaltserschließung durch Indexieren erheblich ist. Indes definiert Fugmann die inhaltliche Erschließung als eine das Wiederfinden erleichternde Maßnahme, „[...] ohne daß hierbei untragbar große Kosten entstehen.“ (1999, S. 11). Der Einsatz einer Erschließungsmethode ist somit ein Kompromiss, der den Nutzen und die Kosten der eingesetzten Methode im Kalkül berücksichtigt. Als Beispiel einer untragbaren Relation führt Fugmann (1999, S. 12) treffend die Analogie eines ständig in sein Lager produzierenden Unternehmens, das jedoch keinen Zugriff auf dieses hat, an - jede Nachfrage führe dann zu einer erneuten Produktion und somit zu Doppelarbeiten. In dem Vergleich entsprechen die Produkte den Informationen und das unzugängliche Lager dem unerschlossenen Unternehmenswissen.

Monetäre Konsequenzen für unauffindbare Informationen und Doppelarbeiten für Unternehmen sind schwer einzuschätzen. Eine Studie der IDC zeigt auf, dass ein typisches Unternehmen mit 1000 Wissensarbeitern (engl. knowledge workers), die 30% ihrer Arbeitszeit (2,5 h/Tag) mit der Suche nach Informationen verbringen und ein Jahresgehalt von 80.000 Dollar beziehen, ca. 2,5 Mio Dollar pro Jahr wegen unproduktiver Informationssuche verschwendet. Zu diesem Betrag kumulieren sich die Aufwände der zu leistenden Doppelarbeit aufgrund nicht auffindbarer Informationen in der Höhe von ca. 5 Mio Dollar pro Jahr und schwer quantifizierbare Opportunitätskosten von ca. 15 Millionen Dollar pro Jahr. (Vgl. Feldman & Sherman, 2001, S. 7f.).

Ein bekanntes Beispiel insuffizienter Informationserschließung und gleichzeitiges Antezedens zur Verbesserung der Bemühungen im Information Retrieval wurde durch den ersten künstlichen Erdsatelliten Sputnik, welcher für die USA überraschenderweise von der UdSSR 1957 ins All geschossen wurde, bewirkt (vgl. Stock, 2006, S. 40). Sputnik sendete ein verschlüsseltes Signal zur Erde, welches es von den USA zu entschlüsseln galt; Diese waren ein halbes Jahr lang mit dem Dechiffrieren des Signals zugange, obwohl die verwendeten Signale und ihre Verschlüsselung lange vor dem Start des Satelliten von amerikanischen Stellen übersetzt und in heimischen Bibliotheken verfügbar waren (vgl. Rauch, 1988, S. 8f., zit. nach: Stock, 2006, S. 40).

Eine zu besprechende Analogie, die an die Erfahrung des Schüleralltags anknüpft, illustriert folgende Situation: Ein Schüler kommt von der Schule nach Hause und löst die Mathematikhausaufgabe schriftlich. Am Abend wollen seine Erziehungsberechtigten die Aufgaben gemeinsam durchgehen, der Schüler findet jedoch die Ausarbeitung unter seinen Mitschriften und Schulbüchern nicht und muss die Aufgaben erneut schriftlich lösen. Wäre eine inhaltserschließende Maßnahme notwendig, um die Doppelarbeit zu vermeiden?

Eine bedingungslose Notwendigkeit für Aufwände inhaltserschließender Verfahren liegt beim Umgang mit Dokumenten nicht a priori vor. Gaus fragt hier nach der „*Dokumentationswürdigkeit*“ (2005, S. 13) eines Dokumentes, bevor es in eine Dokumentation aufgenommen und somit inhaltlich erschlossen wird. Die Leitfrage hierbei ist nicht, ob ein Dokument subjektiv für gut oder schlecht empfunden wird, sondern „*[...] ob es zum bearbeiteten Sachgebiet gehört, ob es einschlägig ist und ob es für den einen oder anderen Benutzer der Dokumentation interessant sein könnte.*“ (ebd.), oder ob es spezifizierten Kriterien zeitlicher, formaler oder inhaltlicher Natur sowie Kriterien des Nutzerbedarfs (für Betriebsangehörige oder die Öffentlichkeit) entspricht (vgl. Stock, 2006, S. 54). Auch Kiel und Rost (2002, S. 87f.) warnen vor un gerechtfertigten Aufwänden: „*Stopp! Bevor Sie also munter an die Erarbeitung eines Thesaurus [ein Instrument der Inhaltserschließung, siehe Kapitel 4.6, d. Verf.] gehen, überlegen Sie erst einmal, ob dieser Aufwand notwendig ist*

und ob nicht eine ‚passive Inhaltserschließung‘ ausreicht.“ Hier ist festzuhalten, dass eine Formalerschließung nicht die einzige Alternative zu einem Thesaurus ist. Es existiert eine Reihe von Methoden, die jeweils Vor- und Nachteile bieten und sich somit für eine bestimmte Anforderung angemessener in den anfallenden Kosten als der Thesaurus verhalten. Kriterien, ob sich grundsätzlich eine Erschließungsvariante gegenüber einer formalen Erschließung rechnet, sind:

- Der Dokumentenumfang,
- die Geschwindigkeit des Wachstums und
- der Nutzerkreis der Dokumentation (vgl. ebd.).

Bezogen auf das obige Hausaufgabenbeispiel stellt der Einsatz einer Dokumentation eine übertriebene Maßnahme dar. Eine zu stärkende Disziplin in der Haltung einer physischen Ordnung der überschaubaren Menge an Unterlagen nach dem Prinzip Gleiches-zu-Gleichem genügt, um die gesuchten Informationen für die abendlichen Hausaufgabenbesprechungen wiederzufinden.

Würden hingegen Berechnungen zu einer Evaluierungsphase eines Bauunternehmens zu einem Brückenbauprojekt zählen, auf die alle Projektmitglieder und externen Statiker jederzeit Zugriff haben müssen, dann sind die Berechnungsblätter einer innerbetrieblichen Dokumentation würdig. Sie zählen zum Wissen des Unternehmens und sind im erschlossenen Zustand für einen späteren Bedarfsfall, womöglich einer folgenden Projektphase und gegebenenfalls von einem anderen Personenkreis, nutzbar und wiederauffindbar. Werden die recherchierten Informationen nicht gefunden, muss unter anderem teure Doppelarbeit geleistet werden.

Ähnlich gelagert kann auch die Dokumentation des Schulmanagements, weil eine Art innerbetriebliche Dokumentation, gesehen werden. Als Illustration eines dokumentationswürdigen Dokuments sei hier ein 15seitiges Protokoll einer dreistündigen Abschlusskonferenz genannt, welches am Anfang des darauffolgenden Schuljahres bezüglich der Nachvollziehbarkeit einer Abstimmung des konferierenden Lehrkörpers als Basis für eine Entscheidung

vorgelegt werden muss. Sind das originale Dokument und das eingescannte nicht auffindbar, würden – beispielhaft – Kosten für rechtlichen Beistand in einer Sache aufgrund unnachweisbarer Entscheidungen verursacht. Nicht nur das Schulmanagement in administrativer Hinsicht, sondern auch das Wissensmanagement der Schule profitiert von erschlossenen Dokumenten (siehe Kapitel 5.1).

Übermäßig anfallende Kosten ergeben sich nach obigen Schilderungen einerseits durch Ignoranz der Notwendigkeit einer einzuführenden Erschließungsmethode und durch die Wahl eines inadäquaten Erschließungsverfahrens entstehen ebenso Kosten, die vermeidbar und zudem schwierig in Zahlen auszudrücken sind.

Andererseits darf die Anwenderseite, die in einem System recherchiert, in der Kostenfrage und somit in ökonomischer Hinsicht nicht unberücksichtigt bleiben. Ein Dokument wird in der Regel nur einmal oder wenige Male inhaltlich erschlossen und wird darauf folgend Gegenstand einer beliebig oft durchgeführten Recherche. Weist das suchende Subjekt mangelnde Kenntnisse über den Suchprozess und dessen zugrundeliegenden Erschließungsprinzipien auf, so entstehen unkalkulierbare Kosten durch Konsequenzen überdurchschnittlich hoher Suchzeiten oder unauffindbarer Dokumente. Gaus (2005, S. 54) schreibt, dass die Freitextsuche in inhaltlich nicht erschlossenen Dokumenten viel Wissen voraussetzt, welches für eine gebundene Indexierung (siehe Kapitel 4.4) notwendig ist. Daraus lässt sich schließen, dass eine effiziente Recherche – innerhalb jeweils akzeptierter Toleranzgrenzen bezüglich Ballast und Informationsverlust (siehe Kapitel 2.5) berücksichtigend (vgl. Knorz, 2004, S. 182) – ohne ausreichendes Wissen über sie schwierig ist. Diesem Faktor entgegen Kenntnisse über die Anwendung von Erschließungsmethoden, damit einhergehende Grundlagen über die Natur der Sprache (siehe Kapitel 3) und Konzepte bezüglich Suchstrategien.

2.4 Für wen wird erschlossen?

Die Klärung, was Inhaltserschließung ist und die Zielsetzung ihrer Anwendung mit der Wiederauffindbarkeit und des Nutzbarmachens führen zu der wesentlichen Frage, für wen Inhalte erschlossen werden? Wiederauffindbar-

keit und Nutzbarmachung setzen potentielle in der Zukunft liegende Bedarfsfälle voraus, da ansonsten jegliche Bemühung inhaltlicher Erschließung paradox wäre.

Nachweisbar ist die Praxis, Wissen für den Bedarfsfall nutzbar zu machen, mindestens ab der Antike. Diese Epoche brachte die Tradition der *Ars Memoria* (Gedächtniskunst) – die heute Gegenstand des psychologischen Forschungsfeldes der mnemonischen Lernstrategien wie *Loci-Methode* darstellen (vgl. Stangl, 2006, S. 91f.) – hervor, in der man sich zu merkende Informationen Gegenständen in Räumen eines imaginierten Gebäudes zuordnet. Durch das geistige Durchschreiten der Räume wird Wissen wiederauffindbar und nutzbar. Diese Art der individuellen Erschließung von Wissen für sich selbst nutzt bereits grundlegende heute noch angewendete Prinzipien. (Vgl. Kiel & Rost, 2002, S. 19).

So hilfreich diese Art der Erinnerungsmethode im Persönlichen auch scheint, umso unmöglicher erweist sie sich, wenn der Nutzungshorizont des Individuums überschritten wird. Das imaginäre Gebäude ist nur in einem Subjekt, dem Erbauer, verfügbar und infolgedessen für kein anderes weder zugänglich noch nutzbar.

Die *Ars Memoria* ist, wie auch alle anderen Methoden der inhaltlichen Erschließung, „[...] niemals Selbstzweck, sondern erfährt ihre Bestimmung dadurch, den Nutzern Zugang zu und Orientierung über Dokumenteninhalte zu verschaffen.“ (vgl. Bertram, 2005, S. 19). Dass ‚den Nutzern‘ auch im Singular stehen kann, wie die *Ars Memoria* illustriert, zählt eher zur Ausnahme, obgleich dieser Modus durch aktuelle Web 2.0 Anwendungen potentieller wird, da durch die Option Dokumentationseinheiten wie Lesezeichen (engl. bookmark) inklusive der ihnen zugeteilten Tags als privat zu markieren, eine nur persönliche Unterstützung im Wiederauffinden und im Nutzen fördert.

Um das Verständnis des Sonderfalls, in dem der Nutzer nur für sich selbst erschließt, zu erleichtern, soll obiges Hausaufgabenbeispiel eines Schülers, in dem die Dokumente nicht dokumentationswürdig waren, fortgesetzt werden.

In Anlehnung an das von Gaus (2005, S. 18) gegebene Beispiel einer Dokumentation, sei der Schüler nun zum Student graduiert, der in einer Vorlesung eine Mitschrift führt und vor seinem schriftlichen Kolloquium eine Dokumentation, die er dort verwenden darf, erstellt: Er nimmt alle 250 Seiten seiner Mitschrift, nummeriert sie fortlaufend durch und vergibt ihnen Schlagworte, die den Inhalt der Seiten kennzeichnen. Auf einer zusätzlichen Seite führt er die Schlagworte alphabetisch an und fügt jedem Schlagwort die Seitennummer(n) hinzu, auf der oder denen das zugehörige Thema zu finden ist. So macht er sein Wissen in Form eines Registers (siehe Kapitel 4.5) in seiner Mitschrift nutzbar und findet es während des Lernens und der Prüfung wieder. In diesem Beispiel erschließt der Schüler nur für sich selbst Inhalt, mit dem aktuellen Bedarfsfall ‚Lernen und Prüfung‘ – und einem möglichen zukünftigen, in dem er beispielsweise in seiner Mitschrift nach Informationen zu einer beruflichen Fragestellung suchen wird.

Die Existenz der *Ars Memoria* zeigt auf, dass der Mensch ein limitiertes Erinnerungsvermögen und „*Suchvermögen*“ (Fugmann, 1999, S. 11) hat und genau deshalb zur Kompensation dieser sich Hilfsmitteln und Methoden bedient. Durch Erweiterung des Nutzerkreises vom Einzelnen, der nur für sich selbst erschließt, zu beliebig vielen Nutzern wird die Methode der Gedächtniskunst unbrauchbar und eine alternative Erschließung notwendig.

Die Nutzbarmachung für einen beliebigen Nutzerkreis geht mit den Konsequenzen einher, dass ein geeignetes Medium für die Repräsentation notwendig wird (vgl. Bertram, 2005, S. 22), eine unbestimmte zeitliche Distanz zwischen Erschließungs- und Nutzungszeitpunkt vorliegt und sich der Nutzertyp ebenso beliebig verhält.

Die Klärung der Repräsentation im Medium Sprache, in dem das Artefakt der Erschließung fixiert wird, behandelt Kapitel 3, in dem auch die Herausforderung der zeitlichen Differenz von Erschließung und Nutzung zur Sprache kommt.

In ihrem Gegenstandsbereich enthält diese Arbeit ausschließlich den Nutzertyp Mensch als Anwender erschlossener Dokumente und grenzt somit The-

men, die die Maschine als suchendes Objekt einschließen, ab. Dazu zählen das semantische Web und Teile des Wissensmanagements. Die automatische Indexierung ohne intellektuelle Leistung wird als Methode der inhaltlichen Erschließung in Kapitel 4.1 konzis erörtert.

2.5 Kennzahlen des Nutzbarmachens

Die mittels Suchanfrage relevanten – und möglichst nur solche – Dokumente oder Dokumentationseinheiten aus einer (umfangreichen) Dokumentation zu selektieren, um einen Informationsbedarf, welcher der Grund einer Suche ist, zu befriedigen, ruft erst erschließende Maßnahmen als notwendige Voraussetzung in den Dokumentationsprozess.

Vom Information Retrieval, um eine fundamentale Abgrenzung bezüglich Informationsbedarfs- und Fragetypen zu unternehmen, mit seinen inhaltlichen Erschließungsmethoden unterscheidet sich das Data Retrieval durch Exaktheit und Determinismus im Suchprozess (vgl. van Rijsbergen, 1979, S. 1). Suchergebnisse sind hierbei, sofern die Datenbasis korrekt ist, weder unscharf noch unsicher. Abfragekriterien lassen sich durch Abfragesprachen wie SQL in strukturierten Datenbanken formulieren. Als Beispiel sei hier die Abfrage eines Buches über das formale Kriterium ISBN in einer Bücherdatenbank gegeben: `SELECT Titel, Autor FROM Buecher WHERE isbn='0408709294'`. Vorausgesetzt, dieses Medium existiert im Bestand, so bringt das Abfrageergebnis exakt dieses, mit dem Suchergebnis „Information Retrieval, C.J. van Rijsbergen“. Das Data Retrieval erfährt in der vorliegenden Arbeit keine weitere Behandlung.

Die über eine Formalerschließung hinausgehende Bemühung zielt idealiter auf eine erwünschte Recherchequalität ab, welche Fragen nach inhaltlichen Aspekten in einem Dokumentenbestand ermöglicht.

Welche Fragearten können identifiziert werden? Stock zeigt anhand der Erkenntnisse von Frants, Shapiro und Voiskunskii (vgl. 1997, S. 34ff.) zwei dominante Typen auf, die vom nach Informationen Suchenden zum Ausdruck gebracht werden. Zum einen existiert ein Fragetypus, der die Beseitigung eines Faktendefizits im Wissen des Fragestellers bewirken soll. Fragen

dieses Charakters lauten beispielhaft innerbetrieblich „Wie lautet das Gründungsjahr unseres Unternehmens?“, im Schulbetrieb „Wer bekleidete erstmalig das Amt des Direktors unserer Schule?“ oder aus Schülerperspektive im Zuge einer Projektarbeit im Unterrichtsfach Geographie und Wirtschaftskunde „Wie lautet die Hauptstadt von Neuseeland?“. Durch ein Suchergebnis als Folge einer Suchanfrage mit exakten Suchkriterien – meist durch Anwendung von Individualbegriffen (siehe Kapitel 3.2) –, wird das Defizit durch retournierte Fakten in der Antwortmenge eliminiert und die Recherche abgeschlossen. Zum anderen lässt sich eine Art des Fragens nennen, welche die Faktenlieferung im Suchergebnis nicht zu beantworten vermag. Exemplarische Fragen zu diesem Typ sind innerbetrieblich „Wie entwickelte sich unsere Produktinnovationskraft seit der Einführung computervermittelter Kommunikationskanäle im Gegensatz zu vorher?“, im Schulbetrieb „Wie wirken sich Weiterbildungsprogramme pädagogischer Natur auf unser Schulklima aus?“ oder in der Recherche eines Unterrichtsprojekts „Auswirkungen klimatischer Veränderungen auf die Länder Europas seit dem 19. Jahrhundert?“. Eine erhaltene Dokumentenmenge auf variationsreiche Suchformulierungsmöglichkeiten enthält Informationen, die nicht alle Aspekte der gestellten Frage beinhalten. Dies führt meist zur Modifikation der Fragestellung und Fortführung der Recherchetätigkeit, um ausreichend zu relevanten Informationen zu gelangen. (Vgl. Stock, 2006, S. 51f.).

Besonders die oben geschilderte zweite Art eines Informationsbedarfs (thematische Suche) bedarf einer genaueren Betrachtung mittels der im Informationswesen üblichen Kennzahlen, die die Relevanz eines Suchergebnisses beschreiben. Maßnahmen, um relevante Dokumente zu finden, erfordern Kompetenzen auf der Inputseite (Erschließung) und der Outputseite (Recherche) (vgl. Bertram, 2005, S. 19). Zu den Kompetenzen zählt auch die Kenntnis der Bewertung eines Ergebnisses. Hierzu werden die Begriffe der Relevanz, Pertinenz, Nützlichkeit, Recall und Präzision im Folgenden behandelt.

Stock heißt den Begriff der Relevanz fundamental und problematisch zugleich, da Anbieter von Internetsuchmaschinen – die auch zum Fachgebiet

Information Retrieval zählen – ihn im Zusammenhang ihrer Relevance Ranking Algorithmen für die Ordnung in Trefferlisten verwenden. Im Informationswesen bezieht sich die Relevanz auf eine nutzerunabhängige, also objektive Betrachtung. Wissen, welches in Dokumenten fixiert ist, wird bezüglich der Befriedigung eines Informationsbedarfs dann als relevant eingestuft, wenn es objektiv in Entscheidungsprozessen dienlich ist oder Wissensdefizite zu kompensieren imstande ist. (Vgl. Stock, 2006, S. 68).

Die Unabhängigkeit des Relevanzbegriffs vom individuellen Informationsbedürfnis des einzelnen Nutzers beschreiben Kiel und Rost mit: *„Relevanz bezeichnet die Verknüpfung von Suchbegriffen [...] mit Dokumenten, die Wissen zum Thema enthalten. Bei der Relevanz geht es also um die allgemeine Einschlägigkeit der einem Dokument zugeordneten Suchbegriffe.“* (Kiel & Rost, 2002, S. 50). In einer Trefferliste zeichnen sich relevante Dokumente aus, indem die vom Nutzer formulierten Suchbegriffe mit denen des Dokumentes übereinstimmen (vgl. Bertram, 2005, S. 20).

Die zu einem Thema relevanten Treffer einer Trefferliste stellen für den Nutzer, der eine Suchanfrage getätigt hat, nicht a priori „relevante“ Informationen dar. Um von der „objektiven Relevanz“ zu unterscheiden, wird die subjektive Relevanz, die sogenannte Pertinenz, eingeführt.

Pertinenz bezieht sich auf die individuelle Situation des suchenden Nutzers, berücksichtigt somit sein kognitives Modell (vgl. Stock, 2006, S. 68). Kraft dieses bringt er seine Suchanfrage zum Ausdruck, *„[...] gemäß seinen Vorkenntnissen, seinem Modell der Außenwelt, seinem sozio-ökonomischen Hintergrund, seiner Sprachkenntnisse [...]“* (ebd., S. 57). Suchen zwei Nutzer nach Informationen in einem Informationssystem mit identen Suchformulierungen, so erhalten sie eine Trefferliste gleicher Relevanz, indes können die pertinenten Ergebnisse divers ausfallen (vgl. Kiel & Rost, 2002, S. 49f.). Insbesondere kann ein Treffer *„[...] nur dann pertinent sein, wenn der Nutzer über die Fähigkeit verfügt, das entsprechende Wissen gemäß seinem kognitiven Modell zu erfassen und zu verstehen.“* (Stock, 2006, S. 69).

In dieser Aussage begegnen sich die Fachgebiete des Informationswesens und der Didaktik. Das hier verwendete kognitive Modell und das in der Didaktik anzuknüpfende Vorwissen im lernenden Subjekt bedingen den Erfolg des jeweiligen Vorgangs Suchen beziehungsweise Lernen. Werden im Unterricht vom Lehrer Recherchequellen wie CD-ROMs oder Internetquellen wie Onlineenzyklopädien und –lexika oder einschlägige Datenbanken, zur Erfüllung einer Aufgabenstellung eingesetzt, müssen diese nicht nur sprachlich, sondern inhaltlich adäquat zum Vorwissen der Schüler verhalten. So zählen zum Einsatz von (öffentlichen) partizipativen Wissenssammlungen, wie die Wikipedia, nicht nur die Aufklärung über qualitative Belange (siehe Unterrichtsmaterial 6.1), sondern auch das Eruiieren des Vorwissens der Schüler, da eine Suchanfrage zwar durchaus relevante Treffer bringen kann, diese jedoch für den Schüler kognitiv noch unzugänglich sind, also nicht pertinent sind.

Befähigen pertinente Suchergebnisse, damit ist das Wissen in den pertinenten Dokumenten gemeint, den Nutzer auf Basis seines kognitiven Modells neues handlungsrelevantes Wissen zu generieren beziehungsweise es praktisch einzusetzen, kann es als nützlich gesehen werden. Je nach Zugang kann es sinnvoll sein, die Nützlichkeit der Pertinenz undifferenziert gleichzusetzen, sowie statt Pertinenz verallgemeinert von Relevanz zu sprechen. (Vgl. Stock, 2006, S. 69).

So wird unter Nutzern, die keinen wissenschaftlichen Zugang zu den Konzepten und der Terminologie der Informationswissenschaft haben, vereinfacht von relevanten Treffern gesprochen.

Die binäre Betrachtung, aus der heraus ein Dokument entweder als relevant oder nicht relevant bewertet wird, ermöglicht zwei weitere mathematisch formulierbare Kennzahlen im Informationswesen, die Precision (Genauigkeit) und den Recall (Vollständigkeit) (vgl. Stock, 2006, S. 72).

Zur Erläuterung sei ein fiktives Informationssystem im Umfang von 10.000 Dokumenten eines Industrieunternehmens gegeben. Das System decke die inhaltlich erschlossene Dokumentation der Forschungsabteilung ab. Ein in

einem Entwicklungsprojekt entstandener Informationsbedarf über erprobte Konzepte von Steckverbindungen zwischen Produktkomponenten bei gegebenen Rahmenbedingungen resultiert in einer Trefferliste, die **100** Dokumente listet. Unter der Prämisse, die Dokumentation speichert **200** relevante Dokumente, von denen die Suche **80** gefunden habe, zeigt sich folgende Aufstellung zur Bewertung:

	relevant	nicht relevant	gesamt
Insgesamt gespeichert	200	9800	10000
In der Suche gefunden	80	20	100
In der Suche nicht gefunden	120	9780	9900

Die in der Trefferliste aufscheinenden nicht relevanten Dokumente – im Beispiel **20** an der Zahl – werden als Ballast bezeichnet, indes wird der Umfang der nicht gefundenen jedoch relevanten Dokumente – hier **120** – Informationsverlust genannt (vgl. Stock, 2006, S. 63).

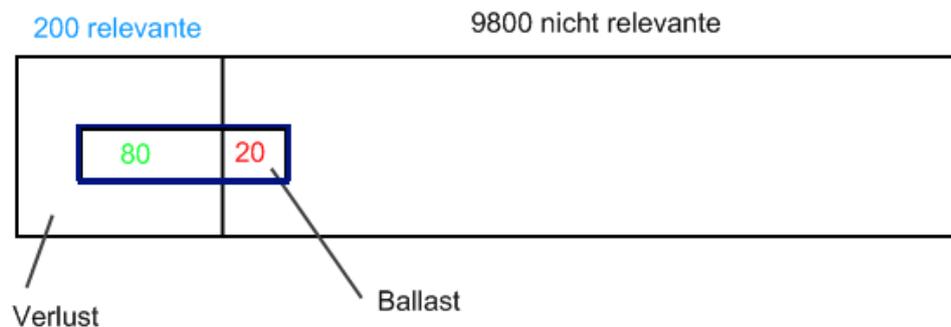


Abbildung 2: Informationsverlust und Ballast

„Die Precision, die Fähigkeit eines Systems, nur relevante Dokumente nachzuweisen [...]“ (Salton & MacGill, 1987, S. 172), gibt Auskunft über die Genauigkeitsquote. Sie wird durch das Verhältnis der in der Suche nachgewiesenen relevanten Dokumente zur Gesamtzahl der in der Suche nachgewiese-

nen Dokumente ausgedrückt (vgl. ebd., S. 175). Im obigen Beispiel ergibt sich:

$$Precision = \frac{80}{80 + 20} = \frac{80}{100} = 0,8$$

Die zweite Kennzahl, Recall genannt, „[...] die Fähigkeit eines Systems, alle relevanten Dokumente nachzuweisen [...]“ (Salton & MacGill, 1987, S. 172), wird als Verhältniszahl zwischen der in der Suche nachgewiesenen relevanten Dokumentenzahl durch die Anzahl der relevanten Dokumente im System angegeben (vgl. ebd., S. 175) und errechnet sich im aktuellen Beispiel wie folgt:

$$Recall = \frac{80}{80 + 120} = \frac{80}{200} = 0,4$$

Die Berechnung des Recalls berücksichtigt einen Wert, welcher keiner direkten Bestimmung zugänglich ist: Der Nutzer weiß nicht, welche und wie viele relevanten Dokumente ihm im Ergebnis unverfügbar bleiben – im Beispiel konkret **120** Dokumente.

Trotzdem eignen sich beide Kennzahlen um modellhaft einen Sachverhalt zu illustrieren. In thematischen Fragen – also nicht in Faktenfragen (siehe oben) – hängen diese Kennzahlen zusammen: Wird durch eine allgemeinere Formulierung der Suchanfrage der Recall erhöht, erniedrigt sich dadurch gleichzeitig die Precision. Diese Art der Modifikation der Suchformulierung reichert die Trefferliste mit Ballast an. Umgekehrt sinkt mittels engerer Formulierung der Suche der Recall zugunsten einer höheren Precision. (Vgl. Stock, 2006, S. 63f.).

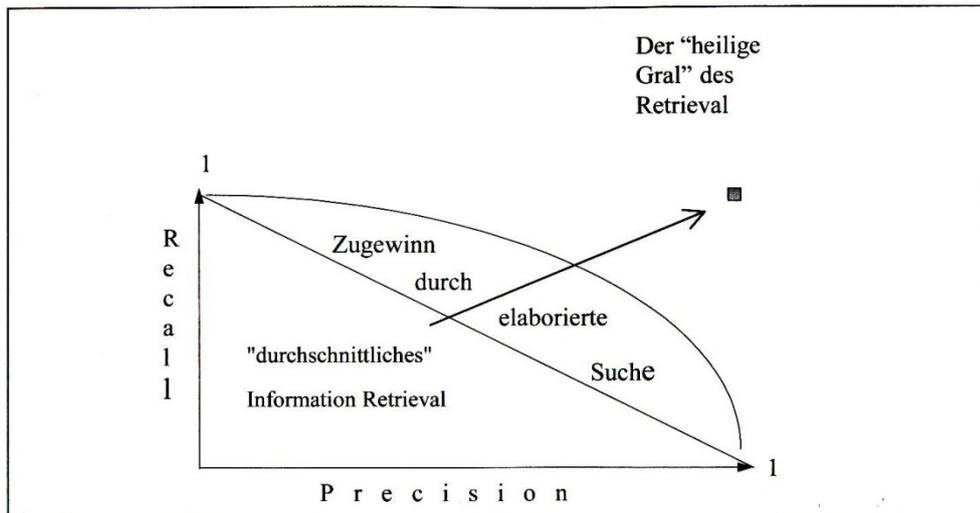


Abbildung 3: Kompromiss Recall und Precision (Stock, 2006, S. 64)

Abbildung 3 zeigt die Korrelation zwischen Recall und Precision. Ein Ziel fortgeschrittener Methoden des Information Retrievals – dazu zählt neben der Suchkompetenz auch die adäquate Anwendung und Durchführung der Tätigkeiten der Inhaltserschließung – ist es, den Kompromiss zwischen Recall und Precision derart zu beeinflussen, sodass beide Kennzahlen möglichst hoch ausfallen und somit, wie Stock schreibt, sich dem „*heiligen Gral*“ (Evans, 1994) des Information Retrieval anzunähern versuchen (vgl. Stock, 2006, S. 64f.).

Durch Manipulation einer Suchformulierung verändern sich auch die Trefferliste und ihre Qualität. Das Unterrichtsmaterial 6.2 behandelt diesen Zusammenhang und begleitet den Lernenden im Verständnis über die Bewertung einer Trefferliste und Vorstellung von Recall (Vollständigkeit) und Precision (Genauigkeit). Um das obige Denkmodell im didaktischen Design angemessen zu berücksichtigen, ist die Klärung über das Wesen eines Denkmodells sinnvoll, um das Vorwissen der Lernenden zu aktivieren und vorbereitend zu ergänzen.

2.6 Teilprozesse der inhaltlichen Erschließung

In ihrem Ablauf erfordert die inhaltliche Erschließung eines Dokumentes zwei Teilprozesse, die Inhaltsanalyse und die Inhaltsdarstellung. Vom Analyseprozess, welcher in der (intellektuellen) Erschließung wesentlichster Teil

ist, hängt ihre Qualität stark ab. Der Prozess stellt die Aboutness eines Inhaltes fest, ein Qualitätskriterium der inhaltlichen Erschließung. Sie klärt die Frage, um was es im betreffenden Dokument geht. Bertram schlägt zur Übersetzung des englischen Begriffs die ‚Begriffliche Abdeckung‘ vor, welche Ausprägungen in zwei Dimensionen vorweist: Die Abdeckung in die Tiefe, damit ist spezifisch und erschöpfend gemeint, und in die Breite, welches ein Indexieren mit wenigen Allgemeinbegriffen bedeutet. Zur weiteren Herausforderung der Inhaltsanalyse zählt die prospektive Identifikation der Zielgruppe, für die die Dokumente nutzbar gemacht werden soll. Die Schwierigkeit besteht darin, nicht nur mögliche Blickwinkel der gegenwärtigen, sondern auch der zukünftigen Nutzer, deren Informationsbedarf schwer abzuschätzen ist, zu berücksichtigen. Der Prozess der inhaltlichen Analyse gliedert sich in die Schritte Orientieren, Strukturieren und Auswählen. Orientierungen liefern Artefakte wie Inhaltsverzeichnisse (Überschriften im Dokument) (siehe Kapitel 4.2) oder Zusammenfassungen – es basiert also auf dem genauen oder oberflächlichen Rezipieren des zu erschließenden Dokumentes oder existenter Verdichtungen über es. Das Strukturieren orientiert sich anhand von Leitfragen, die im folgenden Absatz angeführt sind. Das Auswählen identifiziert schlussendlich die zu übernehmenden Elemente. (Vgl. Bertram, 2005, S. 20 und 22f.).

Dem kategorialen Begriff (siehe Kapitel 3.2) wird fundamentalen Wert in Bezug auf Indexsprachen (siehe Kapitel 3.3) zugeschrieben (vgl. Fugmann, 1999, S. 49f.). Bertram (2005, S. 23) skizziert die mögliche Anwendung von Kategorien in der Inhaltsanalyse und schreibt: *„Beim Strukturieren können Kategorien als Hilfen dienen, anhand derer sich Leitfragen formulieren lassen, z.B.:*

- *Was ist das hauptsächliche Thema?*
- *Wer ist die Zielgruppe?*
- *Welche Theorien oder Gesetzmäßigkeiten werden aufgeführt?*
- *Um welche Prozesse geht es?*
- *Welche Anwendungsgebiete werden thematisiert?*

- Welche Ereignisse, Schlussfolgerungen, Prognosen werden dargestellt?
- Gibt es Zeit- oder Ortsangaben?
- Werden Personen oder Institutionen bzw. Organisationen erwähnt?“

Der wichtige Teilprozess Inhaltsanalyse in der Inhaltserschließung eröffnet der schulischen Unterrichtsdidaktik im Lehrstoff Optionen der Integration fachübergreifender Natur. Die Tätigkeit der Inhaltanalyse erfordert Lese- und Schreibkompetenz und impliziert das Thema Lernstrategien (siehe Kapitel 5.6). Diese Kompetenzen und Strategien sind ebenso bedeutend für Berufstätige, die mit einem hohen Aufkommen von Informationen täglich konfrontiert sind.

Der Inhaltsanalyse folgt die Inhaltsdarstellung. Je nach anzuwendender Methode wird ein Dokument mit einem Indexat oder Abstract angereichert (siehe Kapitel 4).

3 Sprache als Medium

Grundlage der Inhaltserschließung ist üblicherweise die Anwendung der Schriftsprache (vgl. Bertram, 2005, S. 30), die das Fixieren einer inhaltlichen Repräsentation eines Dokumentes in einem – für potentiell Viele – zugänglichen Medium ermöglicht. Die alltägliche Sprache gilt als Konstituens unserer Kommunikation, und kann so als vorausgesetzte Kompetenz für alle Beteiligten des Dokumentationsprozesses gesehen werden. Eine möglichst ungestörte Kommunikation bedarf auch Kenntnisse über die Natur der Sprache. Das Potential einhergehender Missverständnisse in der Anwendung von Sprache, sei es in schriftlicher oder mündlicher Form, wird erst als solches in einem Kommunikationsvorgang explizit. Im Alltag kann sich dies in der Feststellung „wir reden aneinander vorbei, aber meinen das Gleiche“, in einer Erklärung folgenden Frage „weißt du, was ich meine?“ oder Ähnlichem äußern – im Dokumentationsprozess hingegen durch Informationsverlust und Ballast in Trefferlisten. Wodurch wird Unschärfe und Mehrdeutigkeit in natürlichen Sprachen begründet? Dieses Kapitel behandelt Eigenschaften der Sprache, indem es auf das sprachliche Zeichen nach de Saussure, das semiotische Dreieck und die Begriffsbildung eingeht. Abschließend wird Bezug auf didaktische Möglichkeiten genommen.

3.1 Sprachliches Zeichen und semiotisches Dreieck

Der Mensch redet nicht in Gegenständen, Sachen oder Ereignissen, sondern verwendet sprachliche Zeichen, die an Stelle von Gegenständen treten. Ferdinand de Saussure (2001, S. 76), der als wesentlicher Begründer der Semiotik, die Lehre der Zeichen, gilt, kritisiert: *„Für manche Leute ist die Sprache im Grunde eine Nomenklatur, d. h. eine Liste von Ausdrücken, die ebensovielen Sachen entsprechen. [...] Diese Ansicht [...] setzt fertige Vorstellungen voraus, die schon vor den Worten vorhanden waren [...]“*. Diese Kritik alltäglicher Anschauung erinnert analog naturwissenschaftlicher Fachdidaktik den dort erkannten nicht-normativen Konzepten und Erklärungsmustern (Krist, 1999) oder auch Präkonzepte (Clement, 1982) genannt. Denn eine Sache respektive ein Gegenstand ist definiert als *„[...] ein Ausschnitt aus der sinnlich wahrnehmbaren, der gedachten oder denkbaren Wirklichkeit“*

(Felber & Budin, 1989, S. 22), über den zumindest eine Aussage gemacht werden kann (vgl. Fugmann, 1999, S. 18). Die Grenzziehung eines Ausschnitts der Wirklichkeit schlägt sich in der „*Vorstellung*“, wie sie de Saussure nennt, nieder, welche einem Gedankengebilde entspricht, „[...] *die von Mensch zu Mensch etwas verschieden sind.*“ (Felber & Budin, 1989, S. 23). Das bedeutet, die Vorstellungen über Gegenstände sind nicht genuin vorgegeben, sondern in jedem Subjekt konstruiert und (geringfügig) divers. Dies führt zu einer Unschärfe zeichenvermittelter Kommunikation, die besonders im Alltag stärker zum Tragen kommt, als in durch Terminologearbeit begleiteten Bereichen wie den Wissenschaften (vgl. Kiel & Rost, 2002, S. 29).

Der Mensch redet auch nicht in Vorstellungen, sondern bringt diese über etwas zum Ausdruck, das de Saussure als Lautbild bezeichnet. Damit meint er nicht den tatsächlichen physikalischen Laut einer Vorstellung, sondern den psychischen Eindruck, welcher mit der Vorstellung eines Gegenstandes untrennbar in Verbindung steht. Diesen zweiten Teil – der des Lautbildes – bildet gemeinsam mit der Vorstellung – dem Gedankengebilde – ein sprachliches Zeichen. Es besitzt doppelseitigen Charakter, dessen Teile im Gehirn assoziiert sind. De Saussure schlägt die Terminologie betreffend für *Vorstellung* Bezeichnetes und für *Lautbild* Bezeichnendes oder Bezeichnung vor. (Vgl. de Saussure, 2001, S. 77ff.).

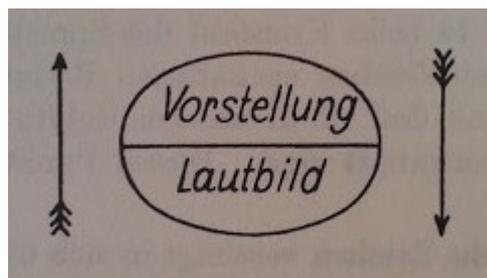


Abbildung 4: Sprachliches Zeichen nach de Saussure (2001, S. 78)

Formuliert ein Redender eine Aussage über einen Berg, so wird die gedachte Vorstellung – das Gedankengebilde *BERG* – in sein Bewusstsein gerufen und löst das assoziierte Lautbild aus, welches über die Sprechorgane zum physikalischen Laut umgesetzt wird. Der Rezipient hört den über die Luft übertra-

genen Laut und identifiziert das entsprechende Lautbild. Das Lautbild als Teil eines sprachlichen Zeichens ruft wiederum die assoziierte Vorstellung *BERG* des Zuhörers – nicht die des Redners – in sein Bewusstsein. (Vgl. de Saussure, 2001, S. 13ff.).

Lautbilder, als Teil des sprachlichen Zeichens, werden nicht irrelevant, wenn statt mündlich schriftlich kommuniziert wird, da die Sprache, die auf Zeichen basiert, so de Saussure, „[...] *vergleichbar mit einem Blatt Papier* [ist]: *das Denken ist die Vorderseite und der Laut die Rückseite; man kann die Vorderseite nicht zerschneiden, ohne zugleich die Rückseite zu zerschneiden* [...]“ (de Saussure, 2001, S. 134). Ohne den oralen Sprechakt, zum Beispiel bei im Geist gesprochenen Gedichten, verwenden der Mensch trotzdem Lautbilder, weil „[...] *die Worte der Sprache für uns Lautbilder sind* [...]“ (ebd., S. 77).

Das heißt, dass auch beim Lesen und Schreiben sprachliche Zeichen gegenüber der Redekonstellation unverändert zur Anwendung kommen. Für die Inhaltserschließung, die im Wesentlichen auf Schriftsprache basiert, bedeutet dies unvermeidbare Konsequenzen hinsichtlich Unschärfe und Mehrdeutigkeit im Dokumentationsprozess, die es in nuce zu berücksichtigen gilt.

Die zweite Quelle möglicher Unschärfe und Ambiguität in der Sprache ist die Art der Beziehung im Zeichen, die zwischen Lautbild und Vorstellung herrscht: Sie ist ebenfalls nicht vorgegeben, sondern arbiträr. Beliebig nicht in dem Sinne, dass eine Vorstellung einem zufälligen Lautbild während der Anwendung, zum Beispiel beim Sprechen, zugeordnet wird, sondern, dass es keinen natürlichen Grund gibt, warum ein bestimmtes Lautbild einer bestimmten Vorstellung zugeordnet ist. De Saussure begründet dies schon dadurch, dass eine Vorstellung in jeder Sprache lautbildlich anders repräsentiert werden kann. So wird im Deutschen die Vorstellung *DREI* mit dem Lautbild *drei* und im Englischen mit *three* assoziiert. (Vgl. de Saussure, 2001, S. 79).

Die im Zeichen lautbildliche Entsprechung *drei* für die Vorstellung *DREI* ist völlig unmotiviert. Der Grad der Motivierung kann je nach Zeichen variieren.

Das Zeichen *dreizehn* lässt einen höheren Grad erkennen, da es die Glieder *drei* und *zehn* erkennen lässt. (Vgl. de Saussure, 2001, S. 156f.).

Ferner bleibt ein Zeichen nicht unbedingt hinsichtlich sprachlicher Entwicklung über die Zeit hinweg konstant. Eine Umgestaltung des Zeichens, „[...] eine Verschiebung des Verhältnisses zwischen dem Bezeichneten und der Bezeichnung [...]“ (de Saussure, 2001, S. 88), wurzelt in der Veränderlichkeit der Vorstellung, des Lautbildes oder beider. De Saussure (ebd.) führt hier unter anderem das Beispiel *dritteil* aus dem ursprünglichen Deutsch, welches heute mit *Drittel* bezeichnet wird. Hier blieb die Vorstellung konstant, das Lautbild hat eine Änderung erfahren. Das Wort *billig* illustriert in den verschiedenen Stadien seiner Entwicklung den Wandel der Vorstellung bei konstantem Lautbild: Zu Lebzeiten von Johann Wolfgang von Goethe noch in der Bedeutung *angemessen* und *fair*, wandelte es sich im Laufe der Zeit zu *niedrig im Preis* bis hin zu *minderwertig in der Qualität*, ausgenommen in der Rechtssprache, in der die ursprüngliche Bedeutung noch gängig ist (vgl. Keller & Kirschbaum, 2003, S. 86f.). Die Phänomene der Verschiebung sind allgegenwärtig, denn: „Keine Sprache kann sich der Einflüsse erwehren, welche auf Schritt und Tritt das Verhältnis von Bezeichnetem und Bezeichnendem verrücken. Das ist eine Folge der Beliebigkeit des Zeichens.“ (de Saussure, 2001, S. 89).

Ein unterschiedliches Ausschneiden der Wirklichkeit in Bezug auf einen Gegenstand kann zudem kultureller Ursache sein. Ein sprachliches Zeichen existiert niemals isoliert, sondern nur durch Abgrenzung und im Wettbewerb zu anderen, und erhält dadurch eine Wertigkeit, die sich von anderen Zeichen unterscheidet (vgl. de Saussure, 2001, S. 138ff.). Etwas Bezeichnetes als kulturelle Einheit kann sich somit beispielsweise in verschiedenen Sprachen oder Kulturen divers abgrenzen. Eco arbeitet dies anhand eines Beispiels von Hjelmslev (1957, S. 104, zit. nach Eco, 2002, S. 86) heraus, das die Möglichkeit unterschiedlicher Bedeutungskreise illustriert:

FRANZÖSISCH	DEUTSCH	DÄNISCH	ITALIENISCH
arbre	Baum	trae	albero
bois	Holz	skov	legno
	Wald		bosco
forêt			

Abbildung 5: Signifikat als kulturelle Einheit semantischer Systeme

Eco (2002, S. 86) erklärt zur schematischen Abbildung 5, dass zu erkennen ist,

„wie das französische Wort /arbre/ denselben Bedeutungskreis deckt wie das deutsche Wort ‚Baum‘, während das Wort /bois/ sowohl dafür gebraucht wird, das zu bezeichnen, was die Italiener ‚legno‘ nennen, als auch das, was sie ‚bosco‘ nennen, während /forêt/ eine größere und dichtere Ansammlung von Bäumen bezeichnet. Das deutsche Wort /Holz/ dagegen bezeichnet ‚legno‘, nicht aber ‚bosco‘, und belässt die italienischen Signifikate ‚bosco‘ und ‚foresta‘ unter der allgemeinen Benennung /Wald/.“

Das Beispiel zeigt, dass sich neben Subjektivität auch kulturelle Aspekte in den Konsequenzen der Anwendung von Sprache als Medium der inhaltlichen Erschließung wiederfinden können. Diese werden im Dokumentationsprozess beispielsweise schlagend, wenn Abstracts lexikalisch übersetzt werden.

Das zweidimensionale Zeichen vermisst die Anordnung des durch es vertretenen Gegenstandes, welcher jedoch durch den psychischen Charakter beider Seiten und ihrer losen Beziehung deutbar wird. Eine Darstellung, die den Gegenstand, auf das das Zeichen referenziert, explizit berücksichtigt, illustriert eine triadische, welche als semiotisches Dreieck bezeichnet wird. Eco (1977, S. 28) hat die triadische Relation zwischen Signifikant, Signifikat und Referent wie folgt skizziert:

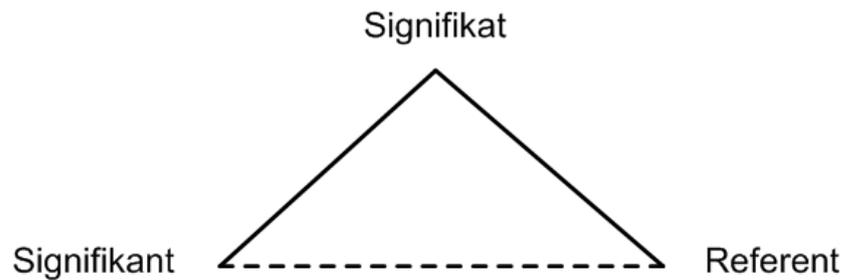


Abbildung 6: Triadische Relation nach Eco

Die Abbildung 6 verdeutlicht anhand einer durchbrochenen Linie, dass zwischen einem Referenten (Gegenstand) und seinem Signifikanten (Bezeichnendes) keine direkte Beziehung herrscht. Sie wird über ein Signifikat (Bezeichnetes) vermittelt. Dies zeigt auch auf, dass ein Signifikant anwendbar ist, ohne Anwesenheit eines Referenten und auch ohne realiter zu existieren, wie zum Beispiel mythologische Figuren. (Vgl. Eco, 1977, S. 29).

Der ‚Computer‘ wird in Abbildung 7 in den drei Ecken der triadischen Relation nach Eco exemplarisch dargestellt.

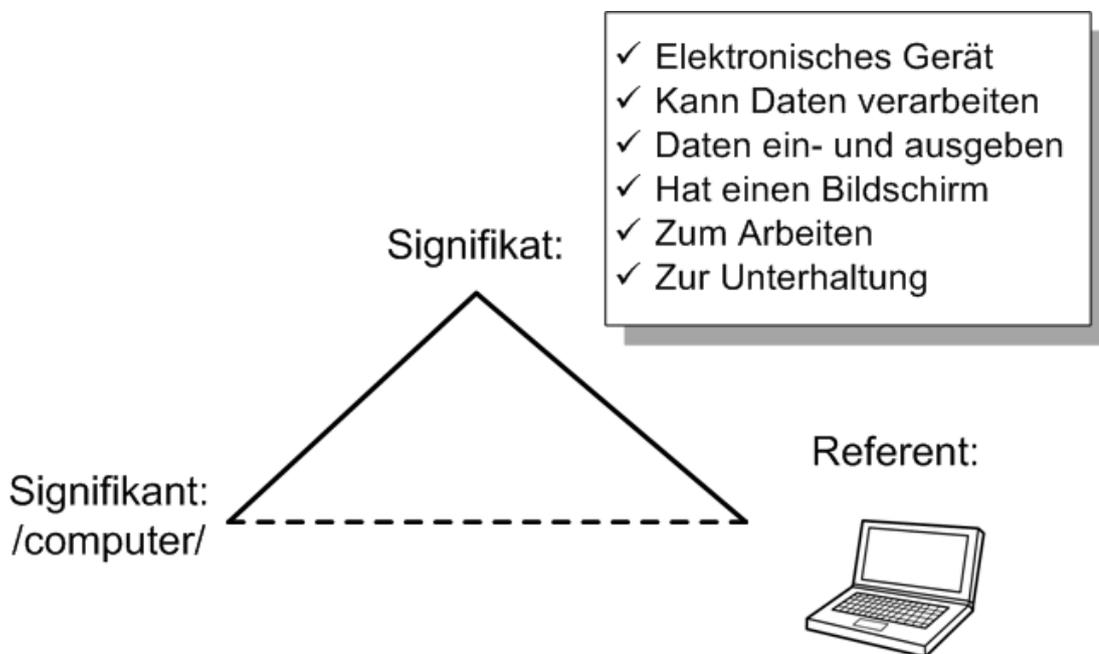


Abbildung 7: 'Computer' anhand des semiotischen Dreiecks illustriert

Im Unterricht scheint der terminologische Umfang, der durch die vielen teils unterschiedlichen wissenschaftlichen Betrachtungen (vgl. Eco, 1977, S. 30; vgl. Posner, 1988, S. 185ff.) der triadischen Relation, dem Verständnis über das Thema selbst im Wege zu sein. Die Anknüpfung an einfache und schon bekannte Worte im Vorwissen der Lernenden ermöglicht die Erarbeitung auf einer bereits gemeinsamen Basis. Zum einfachen und annehmbaren Wortschatz gehören – wahrscheinlich bis dahin noch nicht thematisierte Worte – Begriff und Bezeichnung, welche für de Saussures Vorstellung (Gedankengebilde) und Lautbild stehen.

3.2 Der Begriff

Um ein Verständnis hinsichtlich des Begriffs Begriff zu ermöglichen, ist die Behandlung der Begriffsbildung unabdingbar. Wie oben besprochen, definiert sich ein Gegenstand als Ausschnitt der Wirklichkeit. Die Grenzen des Ausschnitts werden über Merkmale gezogen: *„Wenn wir etwa von einer Reihe von Bezugsobjekten denken, sie seien Teller und von einer Reihe anderer Bezugsgegenständen, sie seien Vasen, dann fassen wir Gegenstände in unserer Umgebung gedanklich durch die Zuordnung von Merkmalen zu Einheiten zusammen.“* (Kiel & Rost, 2002, S. 28). Durch Merkmale werden diejenigen Eigenschaften von Gegenständen wiedergegeben, die ihn ausmachen und zur Bildung eines Begriffes relevant sind (vgl. Bertram, 2005, S. 33).

Die im Zuge der Erklärung eines sprachlichen Zeichens markierte Unschärfe im obigen Diskurs wird über folgendes Beispiel, der Unterscheidung der Gegenstände Teller und Vase deutlich. Im Beispiel von Kiel und Rost (2002, S. 28) könnten die Merkmalsmengen der beiden Gegenstände so definiert sein:

	Teller	Vase
Merkmale	<ul style="list-style-type: none"> • eher flach • gewöhnlich rund • dient dazu, Essen zu servieren oder von ihm zu essen 	<ul style="list-style-type: none"> • Hohlgefäß • im Allgemeinen mit einer Öffnung • dient dazu, geschnittene Blumen aufzunehmen

Abbildung 8: Beispiele für Merkmale Teller und Vase

Ein Indiz für Unschärfe zeigt sich in den Ausdrücken *eher*, *gewöhnlich* oder *im Allgemeinen* in den Begriffsmerkmalen, die im Alltag üblich sind. (Vgl. Kiel & Rost, 2002, S. 28).

In einer Reihe von 100 Beispielobjekten vom typischen Teller in feinen Nuancen abgestuft bis zur typischen Vase mag es etliche Varianten geben, die nicht von jedem Menschen genau einer bestimmten Merkmalsmenge zugeordnet werden. Über ein derartiges Experiment kann Unschärfe visualisiert werden (siehe Unterrichtsmaterial 6.3). Das heißt, ein Geschirrproduzent würde mit hoher Wahrscheinlichkeit die Merkmale für die Gegenstände Teller und Vasen anders definieren als ein Informatiker. Selbst unter Geschirrproduzenten verhalten sich die Merkmalsmengen möglicherweise nicht kongruent und führen zu diversen Begriffsbildungen. „*Diese Subjektivität ist unvermeidbar und tief im Mechanismus der menschlichen Erkenntnis verwurzelt*“ (Fugmann, 1999, S. 19) und führt vor, wie jegliches Bewusstsein über ihre Existenz in kommunikativer Hinsicht, so auch in der Inhaltserschließung, förderlich ist.

Um eine hohe Qualität für Erschließungen und Suchstrategien zu gewährleisten, wird der Begriff zum Gegenstand weiterer Differenzierung in die Begriffsarten Allgemeinbegriff und Individualbegriff, da „[...] *die Unterschiedlichkeit [...] gravierend sich auch in der Art und Weise auswirkt, wie die Begriffe in einem Text sprachlich ausgedrückt werden.*“ (Fugmann, 1999, S. 33).

Begriffe werden durch Merkmalsmengen spezifiziert. Im Zuge der Begriffsbildung verhält sich die Menge variierbar. Kann einem Begriff ein weiteres sinnvolles Merkmal hinzugefügt werden, so führt dies zur Bildung eines Unterbegriffs, wobei der Ausgangsbegriff als Allgemeinbegriff bezeichnet wird. Somit definiert ein Allgemeinbegriff eine Klasse von miteinander verwandten Gegenständen, die eine Merkmalsmenge gemein haben. Ausgehend vom Allgemeinbegriff durch Hinzufügen eines Merkmals und Bildung eines Unterbegriffs, schmälert dieser Vorgang die Klasse der Gegenstände in ihrem Umfang, auf die er referenziert. Eine wiederholte Anwendung auf den erzeugten Unterbegriff macht ihn zu einem Allgemeinbegriff und bildet wiederum

einen Unterbegriff. Dieser Vorgang entwickelt eine Abstraktionshierarchie, die schlussendlich im Individualbegriff endet. Diesem kann im betrachteten Fachgebiet dann kein sinnvolles Merkmal mehr hinzugefügt werden. Die Gegenrichtung, also eine Verallgemeinerung eines Begriffes, d.h. der Bildung eines Oberbegriffs mittels der Eliminierung von Merkmalen, führt zu einem Punkt, an dem in einem betrachteten Fachgebiet keine Fortsetzung des Vorgangs sinnvoll erscheint. Die Wurzel einer Hierarchie bildet eine sogenannte Kategorie. Sie stellen die obersten Allgemeinbegriffe der entwickelten Abstraktionshierarchien dar, und, so Fugmann (1999, S. 24), inkludieren insbesondere „[...] all diejenigen Begriffe, nach denen man in einem Informationssystem unter allen Umständen zuverlässig suchen können möchte.“ Neben einem abstraktiven Hierarchietyp, wie oben beschrieben, welche die Operation „is a“ respektive „ist eine Art von“ bei der Begriffsbildung appliziert, existiert ein partitiver Typ, Bestandshierarchie genannt. Hierarchien dieses Typs konstruieren sich über die „has a“ beziehungsweise „besitzt ein“ Relation. (Vgl. Fugmann, 1999, S. 20ff.).

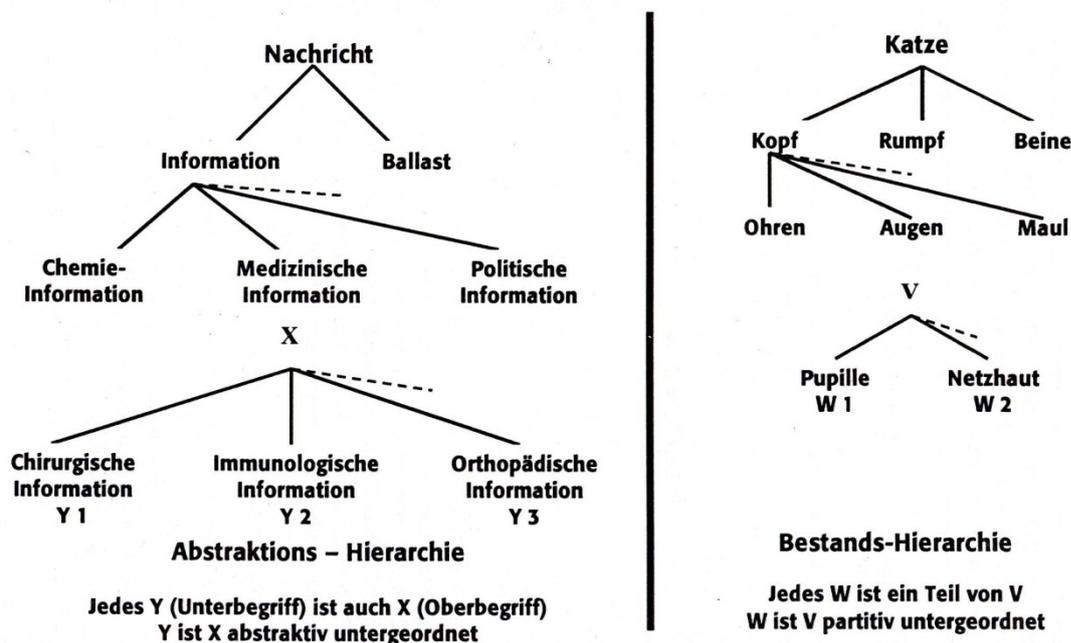


Abbildung 9: Abstraktions- und Bestandshierarchie (Fugmann, 1999, S. 21)

Zur didaktischen Aufbereitung der Begriffsarten Allgemein- und Individualbegriff für Unterrichts- und Schulungseinheiten eignet sich die Feststellung von Wüster (1959/60, S. 187, zit. nach Felber & Budin, 1989, S. 32), dass das Zeigen auf Gegenstände dessen Individualbegriff, nicht Allgemeinbegriff, ins Bewusstsein ruft:

„Der Individualbegriff kann einem Gesprächspartner entweder durch Hinweisen auf den betreffenden Gegenstand oder dadurch ins Bewußtsein gerufen werden, daß ihm ein Ersatzgegenstand vorgeführt wird, von dem durch Vereinbarung oder Gewohnheit feststeht, daß er für Verständigungszwecke den Originalgegenstand vertreten soll. Ein solcher Ersatzgegenstand ist ein ‚Zeichen‘ für den Originalgegenstand. Allgemeinbegriffe können überhaupt nur durch Zeichen identifiziert werden. Denn auf Begriffe kann man nicht hinweisen. Zeigt man auf einen individuellen Vertreter eines Allgemeinbegriffes, so wird dadurch nur der Individualbegriff dieses speziellen Vertreters vermittelt.“

Warum erweist sich die Betrachtung der beiden Begriffsarten als nützlich bezüglich inhaltlicher Erschließung und Suchstrategien? Fugmann konstatiert, dass Individualbegriffe fast immer lexikalisch, indes Allgemeinbegriffe nicht-lexikalisch ausgedrückt werden. Mit lexikalisch meint er die Ausdrucksweise, die der Ausdrucksseite eines sprachlichen Zeichens entspricht. Nichtlexikalisch bedeutet, dass ein Begriff – die inhaltliche Seite eines sprachlichen Zeichens – oft definitorisch oder paraphrasierend umschrieben wird und er dadurch einer nahezu unbegrenzten Mannigfaltigkeit im Ausdruck unterliegt. (Vgl. Fugmann, 1999, S. 29ff.).

Eine Bezeichnung eines Begriffs respektive wie oben elaboriert, der Signifikant – das Bezeichnende – oder die lexikalische Ausdrucksweise eines sprachlichen Zeichens und seine Unterscheidung vom Begriff gehen fundamental in die Thematik der Didaktik ein. Dies gilt nicht nur für das überfachliche Verständnis der Sachverhalte über die Natur der Sprache, sondern auch, wie Fugmann (1999, S. 18) bemerkt, um einer Quelle von Missverständnissen bezüglich Inhaltserschließung in Informationssystemen, welche zu Informationsverlusten und Ballast im Dokumentationsprozess führen, zu entgegnen.

Im Informationswesen, somit in der inhaltlichen Erschließung, führen vor allem die Unterlassung einer Differenzierung zwischen Wort und Begriff und ein linguistischer Zugang zur Verwirrung. In der Linguistik werde der Begriff als Bedeutung eines Wortes gesehen, indes kann ein Wort viele Begriffe bedeuten (bezeichnen) und ein Begriff durch viele Worte oder Wortgruppen bezeichnet werden. Diese Differenzierung in Begriff und Bezeichnung ermöglicht erst das Überwinden der Fehlvorstellung, es seien ausschließlich die Wörter des Verfassers eines Dokumentes oder die Wörter des Suchenden nach Informationen, die zu einem befriedigenden Suchresultat führen. Vielmehr lagert der übliche Fall dort, wo das Suchen sich an einem Begriff orientiert, unabhängig davon, wie dieser in einem Dokument fixiert, lexikalisch oder nicht-lexikalisch, wurde. Die Suche nach Informationen, in denen Begriffe ausschließlich lexikalisch in einem Dokument zum Ausdruck gebracht werden, existiert zum Beispiel bei der Suche nach Personen mittels Name oder Personalnummer, Suche nach eindeutigen Individualbegriffen wie Produktnamen oder terminologischen Studien, in denen die Bedeutungen eines Terminus in Erfahrung gebracht wird. (Vgl. Fugmann, 1999, S. 33ff.).

3.3 Konsequenzen sprachlicher Phänomene

Die Konsequenzen sprachlicher Phänomene werden im Dokumentationsprozess genau dann autoritativ, wenn in einer Suchanfrage die Entscheidung getroffen werden muss, ob ein Dokument entsprechend der Anfrage Relevanz besitzt.

Diese Entscheidung fällt letztendlich ein im System implementierter Vergleichsoperator **ÄHNLICH**, eine Retrievalfunktion. Sie berechnet die Ähnlichkeit der Suchanfrage mit den gespeicherten Dokumenten. Ist eine Ähnlichkeit zwischen Suchanfrage und zu vergleichendem Dokument gegeben, erscheint sie in den nachgewiesenen Dokumenten, der Trefferliste. Die nachfolgende Skizze verdeutlicht die theoretische Beziehung zwischen Suchanfragen und Dokumenten. (Vgl. Salton & MacGill, 1987, S. 11f.).

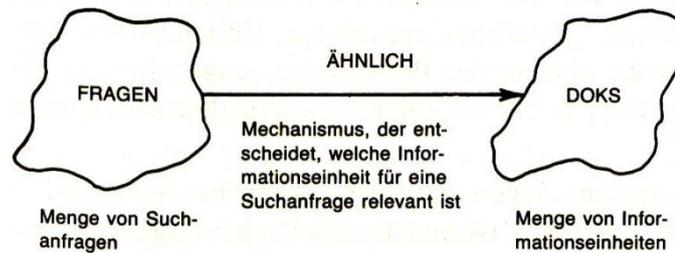


Abbildung 10: Grundstruktur eines Informationssystems
(Salton & MacGill, 1987, S. 11)

Die Skizze von Salton und MacGill zeigt einen unmittelbaren Vergleich von Suchanfragen und Dokumenten. Eine Retrievalfunktion dieser Art, so lässt sich die Darstellung interpretieren, liegt vor, wenn eine inhaltliche Erschließung über Indexierung unterbleibt und die Suche direkt im Volltext oder invertierten Volltext stattfindet. Dann wären alle oder nahezu alle Worte eines Dokumententextes zugleich suchbare Worte. Genau dann werden Notwendigkeit, Ziele und Tätigkeiten der Disziplin Inhaberschließung, mit ihren Teilprozessen Inhaltsanalyse und Inhaltsdarstellung, die Ballast und Informationsverlust in der Recherche reduzieren, augenscheinlich. Folgende Beispiele sollen dies illustrieren (vgl. Bertram, 2005, S. 37ff.):

- Homonymie: Die Benennung „Schloß“ steht für die Begriffe „Gebäude“ und „Schließvorrichtung“, wobei zwischen den Begriffen keine etymologische oder semantische Verwandtschaft herrscht.
- Homographie: Die verschiedene jedoch gleich geschriebene Benennung „Baumast“ steht für die Begriffe „Baumast“ (gesprochen Baumast) und „Baumast“ (gesprochen Baum-Ast)
- Polysemie: Wie Homonymie, jedoch liegt zwischen den Begriffen etymologische oder semantische Verwandtschaft vor, zum Beispiel „Flügel“, „Virus“ und „Zelle“.

- Synonymie: Liegt vor, wenn Benennungen denselben Begriff bezeichnen. Abbildung 11 zeigt Beispiele von Synonymen:

Unterschiedliche Rechtschreibung:	<i>Chicorée – Schikoree</i>
Kurzform – Langform:	<i>IuD – Information und Dokumentation</i>
Offizielle Sprache – Umgangssprache:	<i>Gefängnis – Knast</i>
Eingedeutschtes Wort – Fremdwort:	<i>Meinungsforschung – Demoskopie</i>
Unterschiedliche Sprachen:	<i>Liebe – Love – Amour</i>
Regionalismen:	<i>Fasching – Karneval – Fastnacht</i>
Unterschiedliche historische Epochen:	<i>Leningrad – Petrograd – St. Petersburg</i>
Politisch-ideologisch bedingte Unterschiede	<i>Reichskristallnacht – Pogromnacht</i>
‘Ostsprache’ – ‘Westsprache’	<i>RGW – COMECON</i>

Abbildung 11: Beispiele für Synonymie (Bertram, 2005, S. 40)

- Quasi-Synonymie: Benennungen überlappen sich – zum Beispiel über verschiedene Wortformen („Spielen“ / „Spiel“) oder Konnotationen („Frau“, „Weib“ und „Dame“).
- Paraphrase: Umschreibung eines Begriffes, der entweder lexikalisch vorhanden („Fahrstuhl“ umschrieben mit „Käfigartige Einrichtung zur automatischen Beförderung von Personen innerhalb von Gebäuden“) oder auch inexistent (Nomen für „Schwimmen“) ist.
- Schreibweisenvielfalt: Aufgrund unterschiedlicher Rechtschreibung, vor allem bei Namen ursächlich. Bertram zeigt ein Beispiel auf:

Schreibweisen des Eiskunstläufers Евгений Плющенко	
Vorname	Nachname
Jevgenij	Plushchenko
Evgenij	Plushenko
Evgeni	Pluschenko
Evgeny	Pluschtschenko
Jewgeni	
Ewgeny	

Abbildung 12: Beispiel für Schreibweisenvielfalt (Bertram, 2005, S. 42)

- Wortformenvielfalt: Die Einhaltung grammatikalischer Regeln erfordert die Verwendung verschiedener Wortformen, wie zum Beispiel Singular- und Pluralformen.
- Ausdrucksvielfalt: Besonders abseits von Fachinformation können blumige Sprache, Umgangssprache oder Metaphern zu Problemen führen. Dokumenttitel unterliegen oft einer gewissen Ausdrucksfreiheit, die nicht direkt auf den Inhalt von Dokumenten schließen lässt.
- Begriffskombinationen: „Pferd“ kommt in Kombination mit „Pferdestall“, „Reitpferdsport“ und „Arbeitspferd“ vor.
- Implizite Inhalte: Ein Artikel über die Sesamstraße muss nicht das Wort Kindersendung beinhalten. Über eine adäquate Inhaltserschließung soll es jedoch über eine Suche mit „Kindersendung“ auffindbar sein.
- Mangelnde Transparenz von Begriffsbeziehungen: Aus den Benennungen „Schmetterling“, „Fliege“ und „Heuschrecke“ ist nicht ersichtlich, dass es sich um Insekten handelt. Andere Benennungen täuschen verwandtschaftliche Beziehungen vor.
- Irrelevanz: Die Existenz eines Wortes in einem Volltext erlaubt keinen Rückschluss darüber, ob es den Text inhaltlich beschreibt.

Um nicht komplette Textinhalte zum Suchbaren zu ernennen, und damit vor allem, die Möglichkeit zu schaffen, im Durchsuchbaren Wichtiges von Unwichtigem zu trennen, wird eine intermediäre Indexsprache, oder auch Dokumentationsprache genannt, notwendig. Mit ebenso hoher Wichtigkeit kompensiert der Einsatz einer Indexsprache im Gegensatz zur Natursprache *„die Mannigfaltigkeit“* und *„einen Mangel an Voraussehbarkeit ihrer Ausdrucksweise“*, so Fugmann (1999, S. 21 und 22). Der Vergleich vollzieht sich dann auf einer repräsentativen Ebene, wie in oft zitierter Abbildung von Salton und MacGill skizziert:

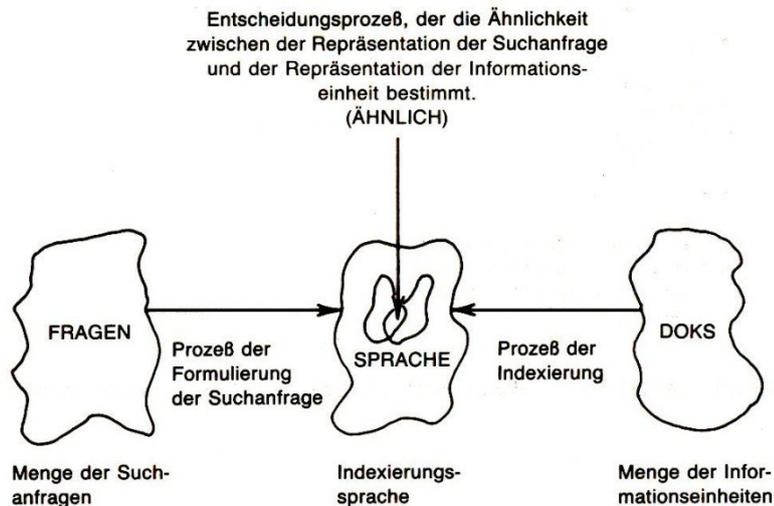


Abbildung 13: Überblick über die Funktionen eines Informationssystems
(Salton & MacGill, 1987, S. 12)

Bertram (2005, S. 46) fasst Konsequenzen natursprachlicher Phänomene für dokumentarische Tätigkeiten zur Reduktion negativer Auswirkungen auf den Dokumentationsprozess mit den Punkten

- „Schreibfehler vermeiden bzw. korrigieren;
- Wortgut normieren;
- Homonyme und Polyseme eindeutig machen;
- Synonyme zusammenführen;
- Paraphrasen lexikalisieren;
- Umgangssprache und Metaphern in formale Sprache übertragen;
- [...]
- Implizite Inhalte explizit machen;
- Begriffsbeziehungen transparent machen;
- Wichtiges von Unwichtigem trennen und darstellen“

zusammen.

Natursprachliche Phänomene rufen nicht nur in der Inhaltserschließung Konsequenzen hervor. Durch das Erkennen der Problematik natürlicher Sprachen in der Erschließung von Dokumenten, sei es im schulischen Unter-

richt (siehe Kapitel 5.7) oder in Mitarbeiterschulungen, ermöglicht die „Übertragung“ (vgl. Schröder, 2002, S. 120f.; vgl. Schwill, 1994, S. 7f.) der Erkenntnisse in andere Bereiche oder Situationen.

4 Methoden der Inhaltserschließung

Das vorliegende Kapitel behandelt gängige und vor allem im aktuellen Bezug relevante Methoden der inhaltlichen Erschließung, die einer didaktischen Berücksichtigung und der Beantwortung der Forschungsfrage dienen. Die einzelnen Methoden können im Zuge dieser Arbeit nur konzis und summarisch vorgestellt werden, da jede einzelne den Umfang eines einzigen Buches sprengen würde. Detaillierte Informationen sind in der gegebenen Literatur zu finden. Dieses Kapitel assoziiert nicht mit didaktischen Themen und Möglichkeiten, wie aufgrund methodenunabhängiger und interdisziplinärer Verbindungen in den vorangegangenen Kapiteln praktiziert. Die diesem Kapitel folgenden werden dann auf die Methoden rekurrieren, um sie didaktisch anknüpfend jeweils zur Anwendung zu bringen. Im aktuellen Kontext stufen sich hinsichtlich Inhaltsanalyse und Inhaltsdarstellung automatische Erschließungsmethoden, in denen die Maschine die Aufgaben der Indexierung übernimmt, für den didaktischen Diskurs als wenig relevant ein. Völlig vernachlässigt dürfen sie hier jedoch nicht bleiben, da die Qualität der Suche von der Erschließung abhängt, ergo auch in der Anwendung maschineller Methoden das Bewusstsein des Suchenden über die Natur der Sprache und Art der Erschließung sich als vorteilhaft erweisen.

Auf eine an dieser Stelle induzierbare Frage über denkbare Qualitätsdifferenzen automatischer und manueller respektive intellektueller Methoden sei hier, ausgenommen mit diesem Absatz, nicht weiter eingegangen. Die Fragestellung lässt sich keiner einfachen Beantwortung zuführen, da sie stark von vorherrschenden Rahmenbedingungen und Überzeugungen abhängt. Fugmann (1999, S. 122) stellt die Frage, ob „[...] die Indexierung zum Programmierbaren oder zum Unprogrammierbaren [gehört]?“. Er untermauert seine Meinung, die eine Unprogrammierbarkeit des Indexierungsprozesses ausdrückt, durch eine Inferenz, die sich aus den Prämissen, dass sich indeterminierte Prozesse jeglicher Programmierung widersetzen und der Charakter des Indexierungsprozesses aufgrund seiner immanenten Unvorhersehbarkeit indeterminiert ist, herleitet (vgl. ebd., S. 123). Andere Autoren sehen den Zugang pragmatischer, wie Salton (1987, S. 64), dem die Bewertung der

Indexierungsverfahren über dessen Retrievalresultate erfolgen sollte, denn automatische Verfahren sind nicht nur schneller und kostengünstiger, sondern kommen auch an Recall- und Precision-Werte der manuellen Methoden heran. Bertram (2005, S. 112f.) relativiert und nennt das Aufgabengebiet automatischer Verfahren dort, wo große, digitale oder digitalisierte sowie möglichst homogene Dokumentenbestände, wie sie etwa im Pressebereich existieren, vorliegen, deren sprachliche Form zum Erfolg automatischer Analyse für die Erschließung führt, und nicht ein Verstehen der Dokumente.

Die Erklärung der vorgestellten Methoden beginnt mit einem kurzen Überblick der automatischen, gefolgt von den einzelnen manuellen bzw. intellektuellen Methoden.

4.1 Automatische Methoden

Ihre treibende Kraft erfährt die Entwicklung automatischer Indexierverfahren in erster Linie aufgrund des rasanten Wachstums der zu erschließenden Dokumentenbestände. Dieses hohe Aufkommen neuer Dokumente, die zum Beispiel durch tägliche Arbeit, wissenschaftliches Forschen oder literarische Tätigkeit erstellt werden, würde durch ausschließlich intellektuelles Erschließen viel Zeitinvestition mit einhergehenden hohen Kosten verursachen.

„Unter automatischer Inhaltserschließung faßt man Verfahren zusammen, die die Prozesse der Inhaltsanalyse und –darstellung auf maschinellem Wege leisten. Sie schließen das automatische Abstracting und das automatische Indexieren ein.“ (Bertram, 2005, S. 97).

Bertram schließt, da oft synonymisch verwendet, die Volltextsuche über invertierte Inhalte, die sich auf Zeichenketten und deren alphabetische Sortierung konzentriert, aus, da das Verfahren nichts weglässt und nichts hinzufügt. Da das Invertieren unspezifische Rahmenbedingungen erfordert – und deswegen auch einfach und vielseitig einzusetzen ist, weil auch in vielen Standardprodukten integriert –, wird durch das Entfallen der im Kapitel 3.3 beschriebenen Konsequenzen eine vermeintliche Einsparung beim Indexieren seitens der Suchenden als Aufwand wieder hervorgerufen (siehe Kapitel 2.3). Automatische Methoden zur inhaltlichen Erschließung stellen sich unter

Zuhilfenahme von computerlinguistischen, statistischen sowie begriffsorientierten Verfahren denselben Herausforderungen wie manuelle Methoden. Allen automatischen Verfahren ist ein wesentliches Defizit immanent: Sie verstehen die Bedeutung von Dokumenteninhalten nicht. Daher stellen Paraphrasen, Umgangssprache, Metaphern, Neologismen, Ellipsen und Ambiguitäten für diese Verfahren große Hürden in der Gewährleistung qualitativer Indexierung dar. So kann im maschinellen Erschließungsprozess ein Pressebericht über ein Fußballspiel bei Schlechtwetter irrtümlicherweise mit ‚Wetterbericht‘ indexiert werden. (Vgl. Bertram, 2005, S. 98ff.).

Ein Beispiel (siehe auch Unterrichtsmaterial 6.4), um die Schwierigkeiten der Identifikation von Interpretationsmöglichkeiten und Auswahl einer wahrscheinlichen, liefert der Satz „*The man saw the pyramid on the hill with the telescope*“ (Vickery & Vickery, 2004, S. 201), dessen Auslegungen Abbildung 14 piktografisch dargestellt.

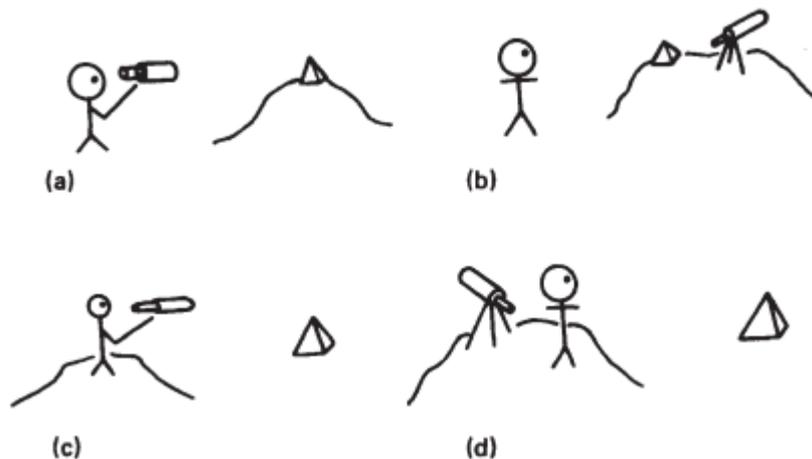


Abbildung 14: Auslegungen von "The man saw the pyramid on the hill with the telescope" (Vickery & Vickery, 2004, S. 201)

Analysiert der Autor des Dokumentes, welches den Satz enthält, für eine Indexierung, weiß er genuin, wie der Satz zu interpretieren ist. Ein Rezipient des Dokumentes wählt aus dem inhaltlichen Zusammenhang heraus. Problematisch erweist sich die Aufgabe für ein automatisches Verfahren, da ein Kontext nur über ein Textverständnis hergestellt werden kann. „Dies er-

schöpfend maschinell abzubilden, dürfte den Retrievalsystemen noch einige Zeit verwehrt bleiben.“ (Stock, 2006, S. 96).

Um den Anforderungen der inhaltlichen Erschließung auf maschinellem Wege gerecht zu werden, unternehmen Hersteller automatischer Verfahren Formalisierungen der Prozesse. Bertram (2005, S. 101ff.) erläutert Aufgaben der automatischen Verfahren wie folgt:

- Stopwort-Eliminierung: Worte wie *,der, die, das, und, oder, zu, in, ein‘* heißen Stopwörter. Sie dienen der Textkohärenz und sind in der Textanalyse nicht von Bedeutung.
- Wortformenreduktion: Durch die Rückführung der Worte auf ihre Stammform werden verschiedene Wortformen auf genau einen Indexterm eingeschränkt. Beispielsweise ist die Stammform von Auswirkungen *wirk*, und von Wissenschaftlerinnen *Wiss*.
- Dekomposition: Besonders in der deutschen Sprache existieren viele Komposita. Diese sollen vom System sinnvoll in begriffliche Bestandteile zerlegt werden.
- Wortgruppenerkennung: Mehrworttermini wie zum Beispiel *„öffentliche und private Einrichtungen‘* oder Paraphrasen müssen vom System erkannt und gegebenenfalls übersetzt oder lexikalisiert werden.
- Namenserkennung: Die Erkennung von Namen stellt eine wichtige und zugleich sehr anspruchsvolle Aufgabe dar. In *„Essen ist angesagt‘* könnte das System *Essen* als Ort oder Mahlzeit identifizieren werden (, und angesagt als modern, oder angekündigt).
- Wortbindestrichtilgung: Formulierungen wie *Ein- und Ausgang* müssen erkannt und in *Eingang und Ausgang* umgewandelt werden.
- Wortableitung: Dies betrifft beispielsweise das Wort *bibliothekarisch* in der Form des Adjektivs in sein Nomen *Bibliothek* zu überführen.

4.2 Das Inhaltsverzeichnis

„Das Inhaltsverzeichnis gibt einen allgemeinen Überblick über die Gliederung des Textes und hilft – wenn die Überschriften prägnant gewählt sind – dadurch bei der Suche nach Informationen und Aussagen oder nach speziellen Kapiteln oder Aufsatztiteln.“ (Kiel & Rost, 2002, S. 57). Es ist ein *„[...] verkürztes Abbild eines Dokuments.“* (Gaus, 2005, S. 51).

Ein Inhaltsverzeichnis stellt vorderhand kein explizit erzeugtes Indexat dar. Das Artefakt resultiert aus einer bereits vorgegebenen Einteilung des betreffenden Dokumentes, es kann also nur dann (automatisch) expliziert werden, wenn das Dokument eine Gliederung aufweist (siehe Unterrichtsmaterial 6.5).

Die Gliederungstiefe eines Inhaltsverzeichnisses muss nicht der vollen Tiefe der Gliederung des Dokumentes entsprechen, Zwischenüberschriften ab der beispielsweise dritten Ebene könnten im Inhaltsverzeichnis unberücksichtigt bleiben. Eine Textgliederung – und somit auch das resultierende Inhaltsverzeichnis – reflektiert die Ordnung der Inhalte im Dokument, welche inhaltlicher und/oder argumentativer Natur ist. Inhaltsverzeichnisse helfen dem Suchenden in Form eines Ordnungs- und Verweissystems mit Überschriften benannte Inhalte und Textabschnitte über Seitenzahlen zu finden und sich zu orientieren. (Vgl. Kiel & Rost, 2002, S. 57ff.)

Der Verweischarakter lässt sich besonders in elektronischen Dokumenten erkennen, in denen oftmals Einträge im Inhaltsverzeichnis visuell als Links präsentiert werden und per Klick direkt zur verwiesenen Stelle navigiert werden kann.

Im Zuge der Inhaltsanalyse eines intellektuell inhaltlich zu erschließenden Werkes – zum Beispiel für die Klassifizierung oder Indexierung – treten Inhaltsverzeichnisse im Schritt der Orientierung als Hilfsmittel auf, da diese, analog zu Zusammenfassungen, bereits inhaltsverdichtet sind (vgl. Bertram, 2005, S. 23). Diese Eigenschaft verhilft auch automatischen Verfahren durch digital vorliegende oder digitalisierten Inhaltsverzeichnissen von Werken wie Büchern oder Zeitschriften zu indexieren (vgl. Hauer, 2003, S. 193ff.).

4.3 Das Abstract

„Unter Abstracting versteht man eine Methode der inhaltlichen Erschließung, bei der der Inhalt einer DBE komprimiert und im Kontext wiedergegeben wird.“ (Bertram, 2005, S. 49).

Laut DIN 1426 (1988, S. 2) fällt es, wie auch Inhaltsverzeichnis (siehe Kapitel 4.2), unter Inhaltsangaben von Dokumenten, da es sich um eine verkürzte Darstellung von Inhalt handelt. Hier interessieren vor allem die Zusammenfassung und das Kurzreferat (Abstract).

Die Norm definiert die Zusammenfassung als „eine Darstellung der wesentlichen Ergebnisse und Schlußfolgerungen eines Dokuments oder von Teilen eines Dokuments.“ (DIN 1426, 1988, S. 2). Die Zusammenfassung zählt nicht primär zu den Artefakten der Inhaltserschließung zum Zweck einer Orientierung bietenden Methode, da sie Bestandteil des Dokumentes selbst und ohne es im Allgemeinen nicht verständlich ist. Daher markiert die Norm Kurzreferat und Zusammenfassung nicht synonym zu verwenden. (Vgl. ebd.).

Ohne Originalvorlage – dem Dokument – hingegen verhilft das Abstract dem Rechercheur bei der Selektion von inhaltlich erschlossenen Dokumenten (vgl. Kuhlen, 2004, S. 190), indem beim Abstracting – auch Referieren genannt – nach DIN 1426 (1988, S. 2) die Dokumentinhalte „kurz und klar“, „informativ“ und „ohne Interpretation und Wertung“ verdichtend wiedergegeben werden.

Neben der Erfüllung einer Orientierungsfunktion, eignen sich Abstracts, falls sie über eine Volltextsuche durchsuchbar sind, für den erleichterten Zugang zu Dokumenten, ähnlich einem Indexat (siehe Kapitel 4.4). Um die Zugangsfunktion zu verbessern, empfiehlt Bertram (2005, S. 52f.) folgende Maßnahmen:

- Zentrale Begriffe sollten mit verschiedenen Benennungen belegt werden: Zu *Dritte Welt* auch *Entwicklungsländer* einfließen lassen.
- Negative Formulierungen vermeiden, da Suchanfragen meist in positiver Form beschrieben sind: ...*nicht frei indexieren*... durch ...*gebunden indexieren*... ausdrücken.

- Paraphrasen lexikalisieren: anstatt *Bundesland zwischen Rhön und Alpen* besser *Bayern* schreiben.
- Wortbindestrichtilgungen vermeiden: *Berufsbilder und –aussichten* werden zu *Berufsbilder und Berufsaussichten*.

DIN 1426 (1988, S. 2f.) definiert folgende Kriterien zur Erstellung von Kurzreferaten:

- Vollständigkeit: Ein Abstract expliziert alle wesentlichen Sachverhalte wie Hypothesen, Zielsetzung, Gegenstand, Verfahren und Methode, Ergebnis, Schlussfolgerung, Anwendung, Zeitraum und geographischer Raum derartig, sodass ein Rückgriff auf das Originaldokument nicht erforderlich wird.
- Genauigkeit: Das Kurzreferat erfordert die genaue Wiedergabe des Inhaltes, ohne diesem etwas hinzuzufügen.
- Objektivität: Wertungen dürfen nicht in den Abstract eingehen. Da es zur verdichteten Repräsentation eine Auswahl und Darstellung der inhaltlichen Sachverhalte bedarf, soll sich der Aufbau des Abstracts am Original orientieren.
- Kürze: Abstracts sind so kurz wie möglich zu verfassen.
- Verständlichkeit: Dieses Kriterium wird erfüllt, wenn möglichst weit verbreitete Fachausdrücke, international und national eingeführte Nomenklaturen, Maßangaben, Formelzeichen, Symbole und Abkürzungen verwendet und ungebräuchliche Ausdrücke vermieden werden – jedoch vom Verfasser neu eingeführte Ausdrücken sollen erklärt und erläutert werden.

Aus inhaltlicher Perspektive differenziert die DIN 1426 (1988, S. 3) folgende Typen des Kurzreferates:

- Informatives Kurzreferat: Ein Abstract dieses Typs ist der übliche Fall und gibt so viel wie möglich an wichtigen Informationen des Originals wieder, unter Berücksichtigung von Zielsetzungen, Methoden, Ergebnisse und Schlussfolgerungen sowie Daten und Fakten.

- Indikatives Kurzreferat: Der indikative Typ deutet an, wovon das Original handelt, indem es nur auf die Sachverhalte und die Art ihrer Behandlung indiziert, ohne über genaue Resultate zu referieren.
- Informativ-indikatives Kurzreferat: Dieser Typ stellt eine Mischform dar und erlaubt die ausgesuchten Sachverhalte informativer darzustellen und andere wiederum nur in erwähnender Art und Weise.

4.4 Das Indexieren

Im Gegensatz zur erschließenden Methode des Abstractings, in der Dokumentinhalte über ganze Sätze und im Kontext repräsentiert werden, reichert das Indexieren eine Dokumentationseinheit mit einzelnen Worten oder Bezeichnungen, auch Indexterme genannt, an. Das Indexieren kommt in der inhaltlichen Erschließung von Dokumentenbeständen, wie sie in Datenbanken, Informationssystemen oder Bibliotheken existieren, sowie bei der Erstellung von Registern (siehe Kapitel 4.5) zur Anwendung und erleichtert somit in erster Line den Zugang zu Dokumenten. Das Ergebnis der Indexierung heißt im ersten Fall Indexat, im zweiten Fall Register (engl. Index). Stichwort, Schlagwort, Deskriptor und Notation sind Arten der Indexterme. (Vgl. Bertram, 2005, S. 67f.).

Indexterme werden wie folgt differenziert:

- Stichwort: *„Ein Stichwort ist ein aus einem gegebenen Dokumenttext extrahiertes und diesen kennzeichnendes Wort.“* (Bertram, 2005, S. 68). Stammt das extrahierte Wort aus dem Dokumenttitel, heißt es Titel-Stichwort, ist der Dokumenttext die Quelle, dann Text-Stichwort (vgl. Gaus, 2005, S. 53).
- Schlagwort: Ein Schlagwort kann, muss aber nicht aus dem Dokument entnommen sein, um es inhaltlich zu beschreiben (vgl. ebd.) und ist somit additiv oder durch Extraktion ermittelbar.
- Deskriptor: Wird ein Indexterm einem kontrollierten Vokabular entnommen, dem ein Thesaurus zugrunde liegt, heißt er Deskriptor (vgl. Knorz, 2004, S. 184).

- Notation: Erfolgt die Ermittlung über eine Klassifikation, handelt es sich um eine Notation (vgl. ebd.).

Stichworte werden durch die Extraktionsmethode, Deskriptoren und Notationen durch die sogenannte Additionsmethode und Schlagworte über beide Methoden ermittelt. Dies hat zur Konsequenz, dass eine Extraktion nur auf verbale Dokumente anwendbar und so für Bilddokumente unbrauchbar ist. Das Indexieren wird je nach Art der Verbindlichkeit in freies und gebundenes Indexieren eingeteilt. Das freie Indexieren erlaubt dem Indexer eine beliebige Zuteilung sinnvoller Indexterme, ohne diese aus einem vorgegebenen Vokabular, wie einer Indexsprache oder dem Vokabular des betreffenden Dokumentes, entnehmen zu müssen. Diese Art der Indexierung erfordert ein hohes Bewusstsein über die Natur der Sprache. Beim gebundenen Indexieren muss sich der Indexer an ein vorgegebenes Vokabular halten, dadurch wird eine höhere Konsistenz beim Indexieren erreicht. Andererseits erfordert das gebundene Indexieren die Entwicklung eines Vokabulars, welches zudem, um einer hohen Aktualität und Spezifität gerecht zu werden, einer fortlaufenden Pflege bedarf. Das freie und das gebundene Indexieren werden in der Praxis oft parallel eingesetzt, um Vorteile beider Varianten im Dokumentationsprozess zu einfließen zu lassen. (Vgl. Bertram, 2005, S. 80ff.).

Gaus vergleicht das gebundene Indexieren mit dem Einordnen eines Dokumentes in ein vorgegebenes Ordnungssystem, das dem Indexer und dem Rechercheur bekannt und vertraut sein muss (vgl. Gaus, 2005, S. 53f.).

Für die Zuteilung von Indextermen im Allgemeinen, jedoch insbesondere in der Variante des freien Indexierens, gelten für Indexterme laut DIN 31623-2 (1988, S. 4) folgende Eigenschaften:

- Wiedergabetreue: Die zuzuteilenden Indexterme sind in Anlehnung an das zu erschließende Dokument zu ermitteln. Damit diese eindeutig, verlässlich und prägnant sind sowie dem Wesen einer Kategorie (siehe Kapitel 3.2) entsprechen, wird gegebenenfalls eine Änderung oder Ersatz der Terme notwendig.

- **Eindeutigkeit:** Indexterme sollen auch losgelöst vom Kontext des Dokumentes Aussagekraft und Informationswert besitzen, insbesondere auch in Unabhängigkeit weiterer zugeteilter Terme. Bezeichnungen sind zu prüfen, präzisieren und disambiguieren.
- **Verlässlichkeit:** Aus dem Aspekt der Konsistenz der Indexierung gelten Indexterme als verlässlich, je spezifischer und genauer sie den Inhalt eines Dokumentes beschreiben.
- **Prägnanz:** „*Der Term sollte kurz, knapp, treffend, genau, einprägsam und unkompliziert formuliert sein*“ (Bertram, 2005, S. 73).
- **Gebräuchlichkeit:** Indexterme sollten sich am jeweiligen Sprachgebrauch des Fachgebietes orientieren, und keineswegs Ausdrücke einer Werkstattsprache (Slang, Jargon) einbeziehen. Ebenfalls kritisch wirken sich Individualprägungen der Dokumentenautoren, Neologismen und Akronyme aus.

Als allgemeine Maxime gilt: „*Je komplizierter die Wortform, desto schlechter die Vorhersehbarkeit für die Nutzer!*“ (Bertram, 2005, S. 74). Wortformen finden sich nicht immer als einfache Sachverhalte vor und müssen somit über zusammengesetzte Benennungen, wie zum Beispiel Komposita oder Mehrworttermini, formuliert werden. Dies ruft drei mögliche Vorgehensweisen hervor (vgl. ebd., S. 74ff.):

- **Präkombination:** Bei präkombinierten Indextermen werden seine Teile der Komposita oder Mehrworttermini bereits als zu einer Einheit kombinierten Benennung dem kontrollierten Vokabular entnommen beziehungsweise beim freien Indexieren im Indexiervorgang zu einem Indexterm verknüpft. Beispiele hierfür sind: *Straßenverkehrsrecht, Arbeitsmarktentwicklung, Information und Dokumentation*.
- **Präkoordination:** Beim syntaktischen Indexieren kommt das Prinzip der Präkoordination zur Anwendung. Es verfolgt die Abbildung eines komplexen Sachverhaltes über das Zusammensetzen der Einzelteile beim Indexieren, beispielsweise in der Syntax *Straßenverkehr/Recht, Arbeitsmarkt:Entwicklung*. Die mit diesen Termen indexierten Do-

kumente werden ausschließlich über genau diese präkoordinierten Benennungen in der Suche nachweisbar.

- Postkoordination: Im Gegensatz zur präkoordinierten Ermittlung von Indextermen ermöglicht das Verfahren der Postkoordination die Recherche nach den einzelnen Teilen eines komplexen Sachverhaltes sowie als Ganzes durch boolesche Operatoren. Das durch die Komponenten des Sachverhaltes indexierte Dokument mit den Indextermen *Arbeitsmarkt # Entwicklung* erscheint in der Trefferliste durch die Suchformulierung *Arbeitsmarkt AND Entwicklung*.

Bertram unternimmt die Problematik der Postkoordination anhand des Beispielindexates

Zuckerrübe # Mais # Weizen # Anbau # Import # Export

zu verdeutlichen. Durch diesen gleichgeordneten Indexterm lässt sich nicht unterscheiden, ob das indexierte Dokument von Zuckerrübenimport und Zuckerrübenexport oder Zuckerrübenanbau, Weizenexport oder Weizenanbau und so weiter handelt. Eine Möglichkeit den authentischen Sachverhalt zu explizieren, bietet die Bildung von folgenden dedizierten Schlagwortketten

Zuckerrübe # Import # Export,

Mais # Import # Export und

Weizen # Anbau,

welche im Retrievalmechanismus separat verglichen werden. Hierdurch wird unter anderem das Thema Maisanbau ausgeschlossen. Neben der Schlagwortkettenbildung kommen in Praxis Rollenindikatoren

Zuckerrübe=Import # Zuckerrübe=Export # Mais=Import #

Mais=Export # Weizen=Anbau

und Kopplungsindikatoren

Zuckerrübe (1) # Mais (2) # Weizen (3) # Import (1,2) #

Export (1,2) # Anbau (3)

zum Einsatz, um Abhängigkeiten im Indexat zum Ausdruck zu bringen. Via Präkombination ist obige Abhängigkeit in der Form

Zuckerrübenexport # Zuckerrübenimport # Maisimport #

Maisexport # Weizenanbau

im Indexat zu fixieren. (Vgl. Bertram, 2005, S. 78f.).

4.5 Das Register

„Unter einem Register versteht man eine Suchhilfe zur ergänzenden Erschließung von Dokumenten mit der Funktion, einen alternativen Zugang auf Dokumenteninhalte zu ermöglichen.“ (Bertram, 2005, S. 116).

Die ISO 999 (1996) definiert das Register (engl. index) als „*alphabetically or otherwise ordered arrangement of entries, different from the order of the document or collection indexed, designed to enable users to locate information in a document or specific documents in a collection*“.

Das Erstellen eines Registers als Artefakt inhaltlicher Erschließung verfolgt unter anderem die folgenden Zwecke:

- Es ist ein Hilfsmittel, um einen raschen inhaltlichen Zugang zu bekannten Dokumenten zu ermöglichen und
- über erinnerte Begriffsbenennungen gewünschte Inhalte aufzufinden sowie
- die Fragestellung von Suchenden zu klären, ob gewünschte Informationen in einem unbekanntem Dokument vorliegen (vgl. Bertram, 2005, S. 106);
- Zudem soll es Wichtiges von Beiläufigem unterscheiden und dabei letzteres ausschließen,
- Informationen, die im Dokument verstreut sind, zusammenführen,
- verwandtschaftliche Beziehungen zwischen Registereinträgen, den Schlagwörtern, wiedergeben und
- dadurch die Suchenden auch über eine anfängliche Benennung des aufzusuchenden Begriffes zu diesem unterstützend hinleiten (vgl. Fugmann, 1999, S. 143).

Ein bekanntes Beispiel eines Registers, so wie es in Fachbüchern üblich hinzugefügt wird, ist das Sachwortregister, indem über alphabetisch sortierte Schlagworte rasch auf gewünschte Inhalte zugegriffen werden kann (siehe Unterrichtsmaterial 6.6). Im Sachwortregister entspricht der Registereintrag einem Deskriptor, der mit einer oder mehreren Seitenzahlen auf die jeweilige Dokumentationseinheit verweist. (Vgl. Gaus, 2005, S. 106).

Auf den ersten Blick scheint das Erstellen eines Registers kein ‚Indexieren‘ im üblichen Sinn zu sein. Dennoch ist die angewandte Methode zur Registererstellung das Indexieren – liegt ein Buch zur Erstellung eines Registers vor, wird es Buchindexieren genannt (vgl. Fugmann, 1999, S. 143ff.). Irritierend wirkt gegebenenfalls das Ziel: Werden sonst ganze Dokumentenbestände wie Bibliotheken oder Dokumentendatenbanken inhaltlich erschlossen, so bezieht sich beim Register der erschließende Vorgang auf ein Dokument oder eine Sammlung von Dokumenten. Die dokumentarische Bezugseinheit (DBE) reduziert sich im Umfang von beispielsweise einem Buch oder einem Worddokument auf Teile dieser, wie auf eine Seite, eine Folge von Seiten, einen Aufsatz, einen Artikel oder auf einen Absatz (zur Absatznummerierung siehe Abbildung 28 im Anhang II.2). Register werden nicht nur auf Dokumente wie Bücher, sondern auch auf „[...] Patentschriften, Filme, Tonaufnahmen, graphische Materialien, Landkarten und Objekte [...]“ sowie bei der Erschließung „thematisch komplexer Websites“ angewendet, führt Bertram (2005, S. 116) an. Als Beispiel für eine mit einem Register versehene Website sei hier der A-Z Index des OECD Auftritts angeführt (siehe Abbildung 27 im Anhang II.2).

Im Unterschied zum Inhaltsverzeichnis erlaubt ein Register die Anwendung eines diversen Ordnungsprinzips gegenüber dem Dokumenteninhalt. Die zugeteilten Deskriptoren lassen sich alphabetisch, systematisch oder chronologisch anordnen. Somit kann ein Dokument oder eine Sammlung von Dokumenten mit mehreren, verschieden sortierten Registern angereichert werden. Nicht nur verschieden geordnete Registereinträge, sondern das Erzeugen von Registern aus unterschiedlichen Perspektiven ermöglicht dem Suchenden mehrere Zugänge zum Dokumenteninhalt. Ein Dokument kann ein Sachwortregister und ein Namensregister aufweisen. (Vgl. Gaus, 2005, S. 106ff.).

Im Aufbau besteht ein Register aus Registereinträgen, welche folgende Elemente aufweisen, wie Kiel und Rost (vgl. 2002, S. 61) beschreiben:

- Registereingang (Stich- oder Schlagwort, Eigenname, Notation, Nummer,...)

- Ggf. eine Zusatzinformation zum Eingang (z.B. Genauere Spezifikation bei Homonymen wie ‚Bank [Geldinstitut]‘)
- Registerausgang mit Angabe der Fundstelle wie Seitennummer, Lexikonartikeltitel, Dokumentnummer,...)
- Und/oder Verweise innerhalb des Registers wie Siehe- bzw. Siehe auch-Verweise)

Die folgenden Abbildungen verdeutlichen die Elemente eines Registereintrags:

<i>Register-Eingang</i>	<i>evtl. Zusatz-information</i>	<i>Registerausgang (auch Registerinformation genannt)</i>	<i>evtl. (Quer-)Verweise auf andere Registereinträge</i>
-------------------------	---------------------------------	---	--

Abbildung 15: Beispiel für den abstrakten Aufbau eines Registereintrags (Kiel & Rost, 2002, S. 61)

Gliederung	(von Dokumenten)	S. 23, 34f., 39ff., 45-49, 60 [statt Seitenzahlen könnten hier bei Bibliographien auch Dokumentnummern oder Notationen stehen]	s.a. alphanumerisches Gliederungssystem; s.a. numerisches Gliederungssystem
------------	------------------	---	---

Abbildung 16: konkretes Register-Beispiel (Kiel & Rost, 2002, S. 62)

Fugmann schreibt, dass eine Vorbemerkung zum Register Angaben über die Reihenfolge der Einträge und Typ des Registers enthalten soll. Zudem sind die bei der Registererstellung unberücksichtigten Teile des Dokumentes wie etwa Anhang, Fußnoten, Vorwort, Literaturverzeichnis, Widmung oder Inhaltsverzeichnis anzuführen. Registereinträge selbst sollten nicht auf Passagen verweisen, die wiederum auf andere Stellen verweisen. (Vgl. Fugmann, 1999, S. 143f.).

4.6 Der Thesaurus

„Ein Thesaurus im Bereich der Information und Dokumentation ist eine geordnete Zusammenstellung von Begriffen und ihren (vorwiegend natürlichsprachigen) Bezeichnungen, die in einem Dokumentationsgebiet zum Indexieren, Speichern und Wiederauffinden dient.“ (DIN 1463-1, 1987, S. 2).

Thesauri unterstützen im Dokumentationsprozess die Möglichkeit der Feinerschließung von Dokumenten über terminologische Kontrolle, indem Synonyme vollständig erfasst, Homonyme und Polyseme als solche gekennzeichnet und Begriffe durch eine determinierte Benennung (siehe Signifikat und Signifikant in Kapitel 3.1) eindeutig vertreten werden. Ist in einer thesaurusgestützten Dokumentationsprache beziehungsweise Indexsprache der Sachverhalt des Heiratens abzubilden, so wird dieser in einem Deskriptorensatz modelliert. Nach dem Sammeln von Synonymen zu „Hochzeit“ wie „Eheschließung“, „Vermählung“, „Trauung“ und „Heirat“ wird für die vorzuziehende Benennung des Sachverhaltes optiert. Ist diese „Hochzeit“, dann besitzt der Deskriptorensatz den aktiven Deskriptor respektive das Gebrauchsvokabel „Hochzeit“, der für das Indexieren und das Retrieval verfügbar ist. Seine Synonyme werden Nicht-Deskriptoren genannt, dienen als Zugangsvokabular (passive Deskriptoren) und gehen nicht als Indexterm in ein Indexat ein. Obige Vorgehensweise modelliert eine Äquivalenzklasse, wie in Abbildung skizziert. Neben diesem Relationentyp, der über das Relationenkürzel BF (Benutzt für) vermerkt werden, vermag ein Thesaurus auch hierarchische und assoziative Relationen abzubilden. Die Hochzeit lässt sich unter dem Oberbegriff (OB) „Familienfeier“ subsumieren und als Unterbegriff (UB) ist die „Goldene Hochzeit“ bekannt. Im Sachverhalt zeigen sich der „Ringtausch“ als Teilbegriff (TP) und die „Verlobung“ als verwandter Begriff (VB). Ferner kann eine Begriffskombination (KB) wie „Hochzeitsreise“ in das Vokabular Eingang finden. Der Sachverhalt ist in Abbildung dargestellt. (Vgl. Bertram, 2005, S. 209ff.).

Hochzeit
BF <i>Eheschließung</i>
BF <i>Heirat</i>
BF <i>Trauung</i>
BF <i>Vermählung</i>
OB Familienfeier
UB Goldene Hochzeit
TP Ringtausch
VB Verlobung
KB <i>Hochzeitsreise</i>

Abbildung 17: Deskriptorensatz für 'Hochzeit' (Bertram, 2005, S. 214)

Eine Liste aller Relationenkürzel befindet sich im Anhang II.4 auf Seite 109.

„Thesauri weisen also eine komplexe Beziehungsstruktur auf, die sich als Folge terminologischer und begrifflicher Kontrolle ergibt. Dabei stellen Äquivalenzrelationen weitgehend das Ergebnis terminologischer Kontrolle dar. Hierarchie- und Assoziationsrelationen sind das Ergebnis begrifflicher Kontrolle.“ (Bertram, 2005, S. 214).

4.7 Die Klassifikation

„Klassifikationssysteme sind Hilfsmittel zur Ordnung von Gegenständen oder Wissen über Gegenstände.“ (DIN 32705, 1987, S. 2). Der Begriff Gegenstand ist im Kapitel 3.1. definiert. DIN 32705 (ebd.) versteht unter Ordnung die Aufgaben in

- technisch-praktischer Perspektive, wie das An-, Bei-, Neben-, Zu-, Zusammen- und Einordnen sowie das Einteilen und Gruppieren,
- wissenschaftlicher Perspektive, die unter Verwendung logischer Hilfsmittel zur adäquaten Darstellung von Wissen führt, sowie
- erkenntnisvermittelnder Perspektive, die das Aufhellen von Zusammenhängen anhand geordneten Wissens ermöglicht,

und subsumiert dadurch Klassifikationssysteme unter den Mitteln der Handhabung von Wissen sowie auch der Förderung der Wissenserkenntnis.

Im Gegensatz zu Thesauri dienen Klassifikationen der inhaltlichen Groberschließung von Dokumenten (vgl. Bertram, 2005, S. 150), zudem eignen sie sich für „[...] Vollständigkeits- und Überblicksrecherchen, also für Recherchen auf einem hohen Abstraktionsniveau [...]“ (ebd., S. 152). Kiel und Rost illustrieren das Konzept anhand der Analogie, Kleidungsstücke nach bestimmten Kriterien in die Fächer des Kleiderschranks einzusortieren, unter dem Grundsatz, dass reale Gegenstände an nur genau einem Ort einsortierbar sind (vgl. Kiel & Rost, 2002, S. 65f.). Heutige Technologien ermöglichen jedoch Dokumente mehreren Klassen zuzuordnen (vgl. Bertram, 2005, S. 152).

Ist vom Begriff der Klassifikation die Rede, so umfasst dieser laut Manecke (2004, S. 127) Folgendes:

- Den Prozess der Klassifikationserarbeitung, das heißt der schrittweisen Strukturierung des Wissens respektive dessen Begriffe in einem Klassenbildungsprozess und Zuteilung der erhaltenen Klassen an einen bestimmten Platz im System,
- das Ergebnis der Klassifikationserarbeitung, das Klassifikationssystem, welches eine Aufstellung von Begriffen und deren Beziehungen darstellt, und
- der Prozess des Klassierens beziehungsweise des Klassifizierens, in dem Dokumente den Klassen des Klassifikationssystems zugeordnet werden.

Klassieren und Klassifizieren sind hier synonym verwendet, oftmals wird die Klassenbildung als Klassifizieren und das Zuordnen von Objekten zu Klassen Klassieren genannt (vgl. Bertram, 2005, S. 150).

Das zentrale Element der Modellierung einer Klassifikation stellt die Klasse dar. Sie entstehen durch die Anwendung eines klassifikatorischen Merkmals, auch Klassen genannt, auf eine Menge von Begriffen und fassen diese zu Gruppen zusammen. (Vgl. Manecke, 2004, S. 127).

Werden zum Beispiel Dokumente eines Unternehmens nach dem Klassen Produktionsstandort unterteilt, so ergeben sich unter anderem die Klassen

Deutschland, China und Australien. Ist jedoch als klassifikatorisches Merkmal die Marktregion gewählt, führt dies gegebenenfalls zu den Klassen Asien, Ozeanien und Europa. Bei der Fragestellung, wonach und wie zu klassifizieren ist, existieren in der Literatur gängige Prinzipien, wie das der Einschlägigkeit oder der Beständigkeit, die von Bertram zusammengefasst sind (vgl. Bertram, 2005, S. 153f.).

Wird im obigen Beispiel der Begriff Europa weiter unterteilt, entsteht eine hierarchische Beziehung, in der wiederum die Klassen Westeuropa, Osteuropa, Südeuropa und Nordeuropa vorkommen. Hierarchien dieser Art werden Monohierarchien genannt. Sie definieren, dass Unterbegriffe genau einen Oberbegriff haben. Zudem lassen sich auch Polyhierarchien bilden, in denen Unterbegriffe mehreren Oberbegriffen zugeordnet sind. (Vgl. Manecke, 2004, S. 128)

Ein aus einem Klassifikationssystem gewählter Indexterm ist eine Notation, welche anstatt mit einer natürlichsprachigen eine künstliche Bezeichnung in der Form einer nach bestimmten Regeln gebildeten Zeichenfolge eine Klasse repräsentiert (vgl. Manecke, 2004, S. 127). Notationen bleiben in modernen Informationssystemen dem Benutzer verborgen (vgl. Bertram, 2005, S. 152).

In der Recherche erfährt der Nutzer erleichternde Möglichkeiten der Suchformulierung, da durch die Angabe eines Oberbegriffes auch alle Dokumente, die mit den jeweiligen Unterbegriffen klassiert wurden, nachgewiesen werden. Ferner erleichtern Notationen die Bedeutungserklärung von Fachausdrücken und die Umschreibung von Begriffen, wenn keine lexikalische Benennung verfügbar ist. Zum Beispiel repräsentiert in der Dezimalklassifikation die Notation „663.432“ das „Waschen und Weichen von Gerste zur Malzbereitung in der Brauerei. Weichwasserbehandlung“ als Paraphrase. Da Klassifikationssysteme nur grob inhaltlich erschließen, ergänzen sie oftmals in einer hybriden Erschließung die natürlichsprachigen Indexierungen. (Vgl. Manecke, 2004, S. 132).

Beispiele für bekannte Klassifikationssysteme sind die Internationale Dezimalklassifikation (DK), die Internationale Patentklassifikation (IPC bzw. Int.

CI), die Colon-Klassifikation (CC) sowie die RIS Klassifikation (vgl. Manecke, 2004, S. 132ff.).

4.8 Das Tagging

Aus dem Englischen übersetzt ist ein Tag ein Etikett oder ein Schildchen. Ein Tag ist dazu da, um es an einen Gegenstand anzubringen, damit er gekennzeichnet wird. Genau dieses Prinzip wird auch in einer Art der Indexierung verwendet. Potenziell jeder Nutzer – vorwiegend in Web 2.0- und Enterprise 2.0-Diensten – wird in die Lage versetzt, durch seine Sicht der Dinge und frei von Regeln mittels Tags Dokumente inhaltlich zu erschließen. *„Diese Art der freien Schlagwortvergabe durch jedermann bezeichnen wir als ‚Folksonomy‘, wobei die freien Schlagworte hier ‚Tags‘ genannt werden.“* (Peters & Stock, 2008, S. 78). Tags sind von der Art des Indexterms Stichworte oder Schlagworte (siehe Kapitel 4.4) und stammen aus dem Vokabular der Nutzer – es ist nicht vorgegeben und somit unkontrolliert.

Trotz einer arbiträren Vergabe von Tags können nach Golder und Huberman Tags typisiert werden – neben dem in der Inhaltsanalyse ausschließlich verfolgten Aspekts, die Aboutness (siehe Kapitel 2.6), zeigen Analysen von Folksonomys unter anderem die Nutzerperspektiven wie *„identifying what it is“*, *„identifying who owns it“* und *„identifying qualities or characteristics“* (Golder & Huberman, 2006, S. 203) auf.

Vander Wal kreierte das heute geläufige Kunstwort Folksonomy (vgl. Peters & Stock, 2008, S. 78) – mit der Bedeutung ‚Taxonomie des Volkes‘ – und markiert den Vorteil in der resultierenden Multiperspektivität: *„The value in this external tagging is derived from people using their own vocabulary and adding explicit meaning, which may come from inferred understanding of the information/object as well as. The people [...] provide their meaning in their own understanding.“* (Vander Wal, 2005b).

Vander Wal (2005a) differenziert in seinem Beitrag zwischen zwei Arten von Folksonomys:

- **Narrow Folksonomy:** In engen Folksonomys erlaubt das Taggingsystem einen Tag maximal einmal pro Objekt zu vergeben, Tags *„are sin-*

gular in nature“ (ebd.). Oft nimmt der Autor die Rolle des ersten Taggers ein – je nach System ist auch anderen Nutzern das Recht, Tags zu vergeben, eingeräumt. Abbildung 18 illustriert das Konzept der Narrow Folksonomy.

- **Broad Folksonomy:** Weite Folksonomys spiegeln wider, „*how a broad range are tagging one object.*“ (ebd.). Das angestrebte Ziel ist es, dass sich viele Nutzer beim Tagging beteiligen. Taggingssysteme dieses Typs unterstützen die mehrfache Vergabe desselben Tags an ein Dokument. In Abbildung 18 wurde der Tag ,2‘ 13 Mal und der Tag ,1‘ zehn Mal dem Objekt respektive Dokument zugewiesen. Dadurch kann die Folksonomy ausgewertet werden. Eine Häufigkeitsverteilung folgt dem Power Law. Tags mit den höchsten Häufigkeiten zeigen die populären Ansichten der Vielen. Für Ersteller von Taxonomien eignen sich Broad Folksonomys in der Terminologiearbeit durch Extraktion der häufigsten Tags.

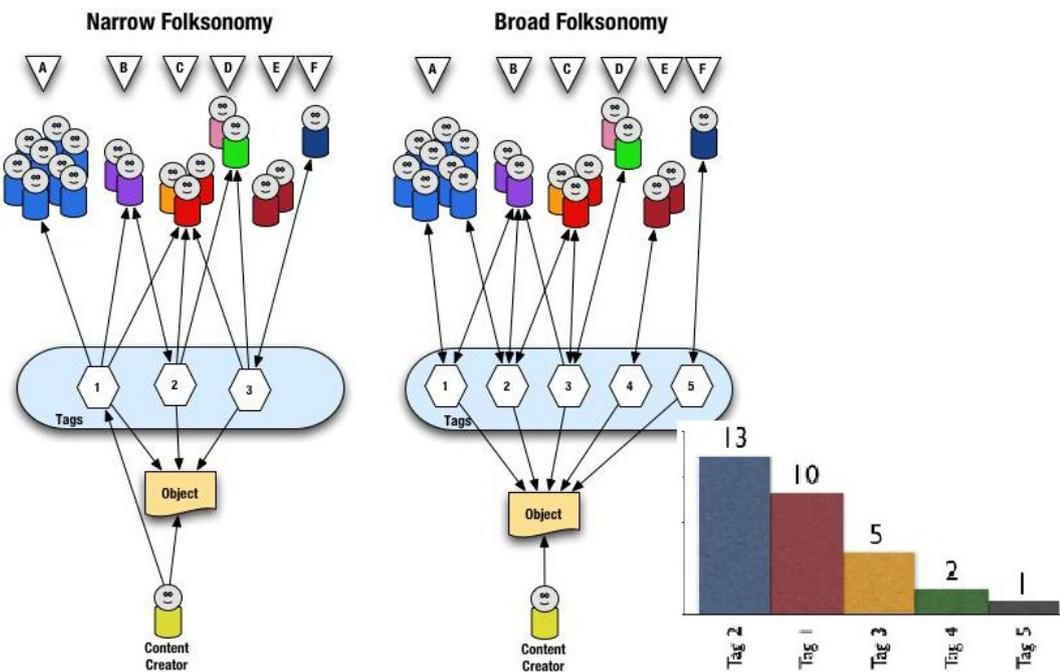


Abbildung 18: Narrow und Broad Folksonomy inkl. Häufigkeitsverteilung (Vander Wal, 2005a)

Häufigkeiten fließen in eine populäre Art der Visualisierung von Folksonomys ein. Tagging-Systeme präsentieren dem Nutzer in sogenannten Tag Clouds eine alphabetische Auflistung des eigenen oder des kollektiven unkontrollierten Vokabulars, in der häufige Schlagworte typografisch hervorgehoben sind (vgl. Gaiser & Panke, 2008, S. 24).



Abbildung 19: Bibsonomy.org Tag Cloud zum Artikel von Mathes (2004) [2013-02-23]



Abbildung 20: Die beliebtesten Tags in BibSonomy.org [2013-02-22]

Die letzten beiden Abbildungen illustrieren Tag Clouds, die sich in der Datenquelle der visualisierten Tags unterscheiden. Abbildung 19 fasst alle Tags zu genau einer Ressource (Dokument) zusammen, indes selektiert Abbildung 20 die beliebtesten Tags aller getaggten Ressourcen im Social Bookmarking Dienst BibSonomy.org. Tag Clouds unterstützen eine alternative Art des Suchens. Mathes (2004) subsumiert sie unter dem Phänomen der Serendipität, durch welches der Nutzer mittels „*Browsing*“ durch die Tags unerwartet auf etwas stoßen kann, das er anfangs nicht als explizite Ursache seiner Suche hatte.

Ein Tag, der beispielsweise wie in Abbildung 19 und Abbildung 20 in einer Tag Cloud dargestellt wird, kann als Hyperlink interpretiert werden. Nach dessen Aktivierung wird jedoch nicht auf eine bestimmte Ressource verwiesen, sondern dynamisch auf eine Menge von Ressourcen (vgl. Gaiser & Panke, 2008, S. 24).

Den Hyperlink-Charakter verkörpern sogenannte Hashtags augenscheinlich. Sie werden nicht in einem dedizierten Tagging-Eingabefeld vom Nutzer vergeben, sondern direkt im Fließtext über ein #-Zeichen (Hash-Zeichen) einleitend expliziert.² Zur Veranschaulichung zeigen die folgenden Screenshots Hashtags eines Twitter Tweets (Abbildung 21, Hashtags #asteroid und #aida) und Google Plus Beitrags (Abbildung 22, Hashtags #Asteroid, #Didymos und #AIDA) der European Space Agency (ESA). Durch Klicken auf den Hashtag werden in den jeweiligen Diensten Inhalte angezeigt, die mit dem entsprechenden Hashtag gekennzeichnet sind.



Abbildung 21: Hashtags im Mikroblogging Dienst Twitter

² URL: <http://de.wikipedia.org/wiki/Hashtag> [2013-02-24]

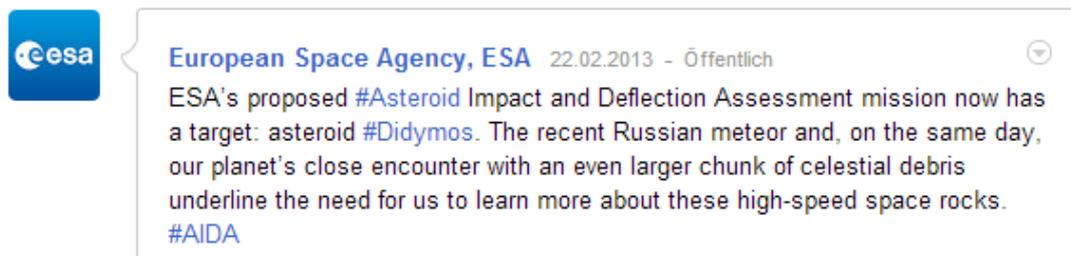


Abbildung 22: Hashtags im Sozialen Netzwerk Google Plus

5 Argumente für Schule und Allgemeinbildung

Dieses Kapitel untersucht Argumente, welche die Integration in den Lehrstoff des Informatikunterrichts sowie die Sensibilisierung der Thematik Inhaltser-schließung in der Schule und in Unternehmen stützen.

5.1 Wissensmanagement in der Schule

Im Kapitel 2.3 wurde ein Beweggrund zur inhaltlichen Erschließung eines Protokolls im Licht des administrativen Schulmanagements gegeben. Eine noch größere Wirkung in der Anwendung von inhaltserschließenden Methoden erwächst im Management des Wissens einer Schule.

Wissen kann in explizites und implizites Wissen klassifiziert werden. Explizites Wissen wurde bereits in einem Medium kodifiziert. Implizites Wissen setzt sich aus noch nicht expliziertem Wissen und aus tacit knowledge, welches nicht kodifizierbar ist, zusammen. (Vgl. Capaul & Seitz, 2011, S. 530).

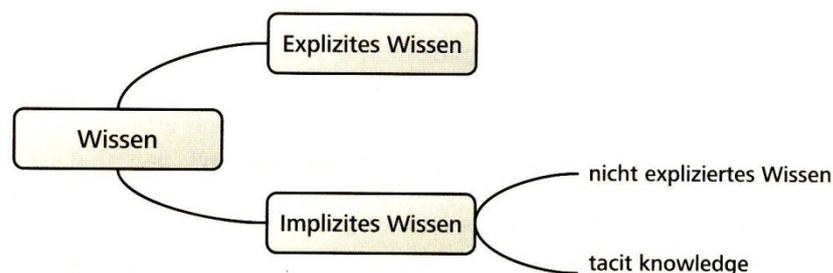


Abbildung 23: Implizites und explizites Wissen (Capaul & Seitz, 2011, S. 530)

Durch Integration von Wissen wird das Wissen des Einzelnen in Form von Handbüchern, Leitfäden, Checklisten und Prozessbeschreibungen für Kollegen zugänglich (vgl. Capaul & Seitz, 2011, S. 533). Dieses dann kodifizierte Wissen liegt in Form von Dokumenten vor. Zugänglich werden diese einerseits durch einen offenen Umgang mit Wissen und andererseits durch Dokumentation, die als Teilprozess das formale und inhaltliche Erschließen vorsieht (siehe Kapitel 2).

Neben den eben genannten Dokumenten liegen in der Schule bezüglich des Wissensbereichs Unterrichtsvorbereitungen, -erfahrungen und -materialien große Nutzenpotenziale vor, welche in dieser Domäne genuin begründet

vorzufinden sind. Zu diesen Prämissen zählen die Tatsachen, dass Wissen für eine erfolgreiche Lehrtätigkeit vorausgesetzt wird und es zudem Gegenstand der schulischen Tätigkeit ist - steigt die Bereitschaft, Wissen in der Schule zu teilen, so entstehen Wissensbestände, bestehend aus Dokumenten wie Fachartikel, Zusammenfassungen, Bücher, Unterrichtsmaterialien, Zeitschriften und anderen Artefakten (vgl. Capaul & Seitz, 2011, S. 534ff.).

Capaul und Seitz sehen technische Instrumente im schulischen Management sekundär, neben Emailversand und Internetzugang genüge die Möglichkeit des Dokumentenup- und downloads (vgl. Capaul & Seitz, 2011, S. 540).

Unter diesen Voraussetzungen, wie sie Capaul und Seitz für das Schulmanagement schildern, läuft eine Schule, welche den Lehrkörper zum Teilen ihrer Dokumente bezüglich Fachwissen ihres Unterrichtsfaches und der Pädagogik erfolgreich ermutigt, Gefahr, eine adäquate Nutzbarmachung der Bestände zu übersehen. Um auf Fragestellungen – beispielsweise über Potentiale interdisziplinärer Unterrichtsgestaltung aufgrund vorhandener Materialien und Literatur – befriedigende Resultate der Suchanfragen in der schulischen Dokumentation zu erhalten oder auch Doppelarbeiten zu vermeiden, bedarf es inhaltserschließender Maßnahmen, wie sie auch in betrieblichen und wissenschaftlichen Bereichen angewendet werden. Capaul und Seitz (2011, S. 535) konstatieren Konsequenzen bezüglich Doppelarbeiten in der Schule über das Monetäre hinaus: *„Darüber hinaus bindet die fortwährende Neuerfindung des Rads auch intellektuelle und zeitliche Ressourcen und beeinträchtigt damit ebenso die Fähigkeit der Schule als Ganzer, neues Wissen zu generieren.“*

5.2 Metadaten im eLearning Bereich

Die in Kapitel 5.1 angesprochenen Unterrichtsmaterialien liegen entweder in analoger oder digitaler Form in Schulen vor. Durch den Einzug von Lernmanagementsystemen (LMS) im schulischen Bildungsbereich, wie auch in der betrieblichen Weiterbildung, werden Unterrichtsmaterialien in Form von einzelnen Lernobjekten oder gar ganzen Kursen in LM-Systemen erstellt.

Am Markt existiert ein großes Angebot an Lernmanagementsystemen, die in der Schule installiert und gewartet werden können. In Österreich wurde zudem vom BMUKK das Projekt edumoodle (www.edumoodle.at) ins Leben gerufen, um Schulen das Betreiben eines LMS ohne eigener Hard- und Software für Moodle-Server zu ermöglichen.³

Um Lernobjekte und –module zwischen Institutionen und Lehrkräften austauschen und wiederverwenden zu können, werden standardisierte Metadatenpezifikationen notwendig. Es wurden zahlreiche Standards entwickelt, die Lernobjekte aus unterschiedlichen Aspekten, wie Lerninhalte, -materialien, -aktivitäten, -strategien, -modelle und –methoden sowie Aggregation zu Lernpaketen, beschreiben. Zu den populären Standards zählen zum Beispiel CanCore, Dublin Core, LOM, IMS, LORN und SCORM. (Vgl. e-teaching.org, 2012).

Um Lernobjekte präzise zu beschreiben bedarf es umfangreichen Angaben über sie. Durch Metadaten werden „[...] *Lerninhalte vergleich-, austausch- und wieder verwertbar [...]*“ (e-teaching.org, 2012). Ein wichtiger Aspekt stellt die Auffindbarkeit in Datenbanken im Zuge einer Suche nach Lernmaterialien für LMS dar. Zur inhaltlichen Erschließung stellt der IEEE LOM Standard eine Metadatenkategorie „*Classification*“ zur Verfügung, mit dem Ziel: „*The Classification category describes this learning object in relation to a particular classification system.*“ (IEEE 1484.12.1-2002, 2002, S. 7). Zudem bietet die „*General*“-Kategorie das „*Description*“-Feld, um es in Form eines kurzen Textes zu beschreiben, und das „*Keyword*“-Feld verfolgt die Methode des freien Indexierens (siehe Kapitel 4.4), um ein Lernobjekt mit inhaltsbeschreibenden Worten oder Phrasen anzureichern. Der LOM Standard bietet zahlreiche Kategorien, um ein Lernobjekt formal und inhaltlich zu erschließen:

³ URL: http://www.edumoodle.at/moodle/pluginfile.php/16148/mod_page/content/10/Produktblatt_edumoodle-at.pdf [2013-03-03]

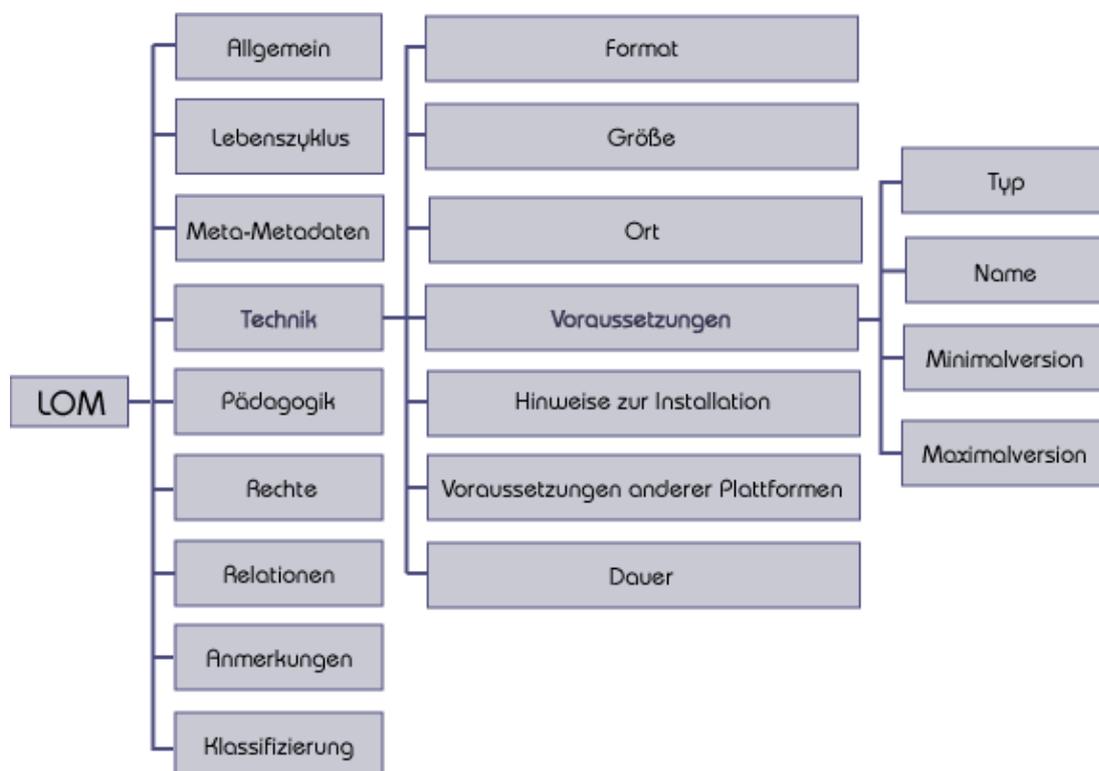


Abbildung 24: Struktur LOM Metadatenstandard (e-teaching.org, 2012)

In Österreich existieren zahlreiche Initiativen, die sich mit der Entwicklung von digitalen Materialien für den Unterricht beschäftigen. Ein Erlass des Bundesministeriums für Unterricht, Kunst und Kultur (BMUKK, 2003) legt eine Rahmenregelung für „[...] die Entwicklung von elektronischen Content für den lehrplanmäßigen Unterricht an weiterführenden Schulen [...]“ fest. Digitale Lernmaterialien werden u.a. von „[...] engagierten Lehrenden, oft auch mit Schülerbeteiligung [...]“ (ebd.) erstellt. Dies zeigt, dass im Umfeld eLearning Lehrende, sowie auch Lernende mit der Thematik der Inhaltser-schließung konfrontiert sind und zudem sinnvolle Anknüpfungspunkte in der Didaktik induzieren: Einerseits erfordert die eContent-Bereitstellung die Einhaltung der Metadaten-spezifikation, die auf dem LOM-Standard basiert, beispielsweise hinsichtlich einer gegebenen Schulbuchklassifikation (vgl. ebd.). Andererseits soll es in einer Datenbank vor der Erstellung eines Lernobjektes überprüfbar werden (Recherche), ob sich bereits idente oder ähnliche Objekte finden lassen, um Doppelarbeiten zu vermeiden.

5.3 Berufsorientierung

Neben den grundlegend involvierten didaktischen Themenfelder Sprache und Lernen bietet das Thema Inhaltserschließung auch in den Berufsorientierungsmaßnahmen in der Schule Ansatzpunkte.

Das Thema Inhaltserschließung ist ein wesentlicher Teilprozess in der Dokumentation (siehe Kapitel 2). Die Dokumentation respektive das Information Retrieval sind Schwerpunkte in den Berufen des Informationswesens (vgl. Gaus, 2002, S. 13). Zu diesen zählen neben den Berufen der Datenverarbeitung und Informatik unter anderem die Fachbereiche Archiv, Bibliothek, Buchwissenschaft, Information- und Dokumentation, Medizinische Dokumentation, Computerlinguistik und Museum (vgl. ebd., S. 7). Das Bundesministerium für Unterricht, Kunst und Kultur (BMUKK, 2012a) verortet als Unterstützung zur Berufswahl im Unterricht Maßnahmen der Berufsorientierung im Ausmaß von einer Wochenstunde: *„Im österreichischen Schulwesen ist der Berufsorientierungsunterricht in der 7. und 8. Schulstufe in allen Schularten [...] als ‚verbindliche Übung‘ verpflichtend im Ausmaß von je 32 Unterrichtsstunden [...] verankert.“*

Im Lehrstoff des Informatikunterrichts, aber auch im fächerübergreifenden Modus, vermag die Behandlung der Thematik Inhaltserschließung die geforderten Bildungs- und Lehraufgaben unterstützend zu erfüllen. Zu diesen zählt zum Beispiel im Bereich Natur und Technik die *„Auswirkung neuer Technologien in den verschiedenen Berufs- und Lebensbereichen“* (BMUKK, 2000, S. 1). Die Schnelllebigkeit der Technik bringt zwar Fortschritte in der automatischen inhaltlichen Erschließung hervor, jedoch bleiben die grundlegenden Rahmenbedingungen – wie die Natur der Sprache – stets erhalten und rufen Konsequenzen in der Recherche hervor. Der Umgang mit umfangreichem Wissen bedarf einer Methodik, die in vielen Berufsfeldern nicht nur von Nutzen ist, sondern notwendig wird. Ein Einblick über das Thema Inhaltserschließung, welches genug Anknüpfungspunkte zum Schüleralltag (Bibliothek, Internet) vorweist, ermöglicht das Wecken von Interessen in den Disziplinen des Informationswesens. Zum Überblick über entsprechende Berufe existieren Wegweiser, wie das Werk von Gaus (2002), welches sich

insbesondere an „[...] Schüler der Hauptschulen, Realschulen und Gymnasien, die kurz vor dem Abschluss ihrer Schulausbildung und damit vor der Berufswahl stehen“ (ebd. Gaus, 2002, S. 7), wendet.

Ferner bietet das Kompetenzparadoxon Anlass zur Diskussion im Informatikunterricht, es offenbart: *„IT ermöglichen den Zugang zu immer mehr Informationen und Wissensbeständen, gleichzeitig wachsen die Anforderungen an die Qualifikation der Nutzer/innen.“*⁴ Ein Informatikunterricht, welchem das Thema Inhaltserschließung nicht fremd bleibt, vermag Antworten im lernenden Subjekt herbeizuleiten und das Paradoxon aufzulösen. Zudem bietet er einen Anknüpfungspunkt zum Thema Berufsorientierung.

5.4 Vorwissenschaftliche Arbeit

Mit der Kundmachung der Novellierung des österreichischen Schulunterrichtsgesetzes im BGBl. I Nr. 52/2010 wurde die rechtliche Grundlage für die vorwissenschaftliche Arbeit geschaffen: § 34 SchUG legt fest, dass ein Teil der abschließenden Prüfung aus *„[...] einer abschließenden Arbeit [...], die selbständig [...] zu erstellen ist“* besteht, wobei sie *„[...] in höheren Schulen auf vorwissenschaftlichem Niveau; mit Abschluss- oder Diplomcharakter [...]“* angesiedelt ist.

Im Zuge der geforderten Arbeit kommen Schüler oft das erste Mal mit der Methodik der Wissenschaft in engeren Kontakt. *„Zum Verfassen der Vorwissenschaftlichen Arbeit ist eine Reihe von Kompetenzen notwendig, auf deren Erwerb der Unterricht der Oberstufe ausgerichtet sein muss und die unter dem Begriff ‚Informationskompetenz‘ subsumiert werden [...].“* (BMUKK, 2012b, S. 10). Zur Informationskompetenz zählen das Erkennen des Informationsbedarfs und die Befriedigung dieses durch Finden und Beschaffen der notwendigen Informationen (vgl. Franke, Klein, & Schüller-Zwierlein, 2010, S. 3). Dies rückt die Tätigkeit der Recherche in den Fokus des vorliegenden Dokumentes, die mit ca. 20 % des Gesamtaufwandes in der Erstellung einer vorwissenschaftlichen Arbeit einzuschätzen ist (vgl. BMUKK, 2012b, S. 10f.).

⁴ URL: http://www.bmukk.gv.at/medienpool/17660/berufsorientierung_7.pdf

Als Informationsquellen in der Recherche dienen unter anderem Dokumente, die einem Dokumentationsprozess unterliegen – Bücher in Bibliotheken, wissenschaftliche Zeitschriften, Fachartikel und Internetseiten. Das Lokalisieren relevanter Dokumente für das Verfassen einer vorwissenschaftlichen Arbeit erfordert Kenntnisse über Suchstrategien (vgl. ebd.), die wiederum ein Problembewusstsein über sprachliche Phänomene wie zum Beispiel Synonymie und Homonymie (siehe Kapitel 3.3) voraussetzen.

Findet das Stoffgebiet der Inhaltserschließung im Unterricht der Oberstufe Eingang, dann erwerben Schüler die notwendigen Basiskompetenzen automatisch. Die Aufnahme des Lehrstoffs über Inhaltserschließung lässt sich dadurch didaktisch sinnvoll rechtfertigen. Auf ihm aufbauend können Unterrichtseinheiten oder Beiträge im Zuge der Betreuung vorwissenschaftlicher Arbeiten durch Lehrkräfte stattfinden, um spezielle Anforderungen und Problemstellungen in der Recherche (individuell) zu behandeln.

Neben Informationskompetenzen führt das Verfassen einer vorwissenschaftlichen Arbeit genuin zur Thematik der Inhaltserschließung, da in § 8 Abs. 5 der Prüfungsordnung AHS (in BGBl. II Nr. 174/2012) ein inhaltserschließendes Artefakt eingefordert wird: *„Im Rahmen der schriftlichen Arbeit ist ein Abstract im Umfang von zirka 1 000 bis 1 500 Zeichen, inklusive Leerzeichen, zu erstellen, in welchem das Thema, die Fragestellung, die Problemformulierung und die wesentlichen Ergebnisse schlüssig darzulegen sind.“*

Eine vorwissenschaftliche Arbeit ist vom Schüler mit einem informativen Abstract (siehe Kapitel 4.3) anzureichern. Durch Kenntnis der beiden Funktionen – Zugangs- und Orientierungsfunktion – von Abstracts begründet sich im Verfasser der vorwissenschaftlichen Arbeit eher die Forderung nach einem Abstract. Ferner fördert durch das Erschließen des Sinns eines Abstracts im Schüler die Qualität des Artefaktes.

Durch das Gesetz im Umfang nicht gefordert, jedoch als Übung für den Lehrstoff der Inhaltserschließung stringent platzierbar, kann zudem die Fragestellung nach adäquater Indexierung des Werkes durch den Schüler selbst didaktisch sinnvoll verankert werden (siehe Unterrichtsmaterial 6.8).

5.5 Fundamentale Ideen

Computer sind heute aus dem Alltags- und Berufsleben kaum mehr wegzudenken. Umso umfangreicher wird das Wissen in der Informatik und ebenso herausfordernd gestaltet sich die Selektion der Inhalte in den Unterrichtsfächern. Ein Argument gegen die Behandlung der Inhaltserschließung im Lehrstoff äußert sich im Umfang: *„Die Kritik an den Schulen richtete sich meist gegen die Stofffülle und die damit verbundene Überforderung der Schüler.“* (Schröder, 2002, S. 70). Schwill (1994, S. 12) propagiert daher aus fachdidaktischer Perspektive sich auf langlebige Konzepte zu konzentrieren, sogenannte fundamentale Ideen, die er wie folgt definiert:

„Eine fundamentale Idee bezgl. eines Gegenstandsbereichs (Wissenschaft, Teilgebiet) ist ein Denk-, Handlungs-, Beschreibungs- oder Erklärungsschema, das

*(1) in verschiedenen Gebieten des Bereichs vielfältig anwendbar oder erkennbar ist (**Horizontalkriterium**),*

*(2) auf jedem intellektuellen Niveau aufgezeigt und vermittelt werden kann (**Vertikalkriterium**),*

*(3) zur Annäherung an eine gewisse idealisierte Zielvorstellung dient, die jedoch faktisch möglicherweise unerreichbar ist (**Zielkriterium**),*

*(4) in der historischen Entwicklung des Bereichs deutlich wahrnehmbar ist und längerfristig relevant bleibt (**Zeitkriterium**),*

*(5) einen Bezug zu Sprache und Denken des Alltags und der Lebenswelt besitzt (**Sinnkriterium**).“*

Das Stoffgebiet der Inhaltserschließung lässt sich primär nicht als fundamentale Idee identifizieren. Es beinhaltet jedoch Momente, welche fundamentalen Charakter im Alltags- und Berufsleben aufweisen, wie zum Beispiel die Natur der Sprache. Die Inhaltserschließung eignet sich daher, um das grundlegende Problembewusstsein über natürliche Sprachen zu behandeln und der Sensibilisierung ihrer Konsequenzen in vielen anderen Bereichen. Daher empfehlen sich die eben zitierten Kriterien zur Begründung der Verortung des Stoffgebietes im allgemeinbildenden Informatikunterricht:

- **Horizontalkriterium:** Die Inhaltserschließung zählt nicht mehr ausschließlich zu den Aufgaben von Spezialisten, sondern besitzt vielseitige

ge Berührungspunkte breiter Massen. Dies involviert Menschen die Dokumente verfassen, zum Beispiel:

- Schüler und Studenten, die ein Abstracts zu ihrer (vor)wissenschaftlichen Arbeit verfassen
- Blogger, die ihre Blogartikel über Tags erschließen
- Anwender sozialer Netzwerke wie Twitter und Google Plus, die über Hashtags ihre Posts erschließen
- Anwender von Webdiensten, in denen Fotos und Videos veröffentlicht werden können, machen diese durch Tags auffindbar
- Angestellte, die betriebsrelevante Dokumente in Dokumentenmanagementsystemen ablegen und über semantische Metadaten erschließen
- Lehrer, die Unterrichtsmaterialien im Kollegium oder über Institutionen hinweg teilen und inhaltlich beschreiben (siehe Kapitel 5.2)
- Ersteller von Webseiten, die über den Dublin Core Standard Suchmaschinen im Internet ermöglichen, ihre Seiten adäquat zu indexieren.

Ferner ist das Wissen über Inhaltserschließung von wesentlichem Vorteil bezüglich der Entwicklung von Suchstrategien in der Recherche nach Informationen – eine Aufgabe, die in unserer Wissensgesellschaft alltäglichen Charakter besitzt.

- **Vertikalkriterium:** Mit zunehmender Schulstufe kann detaillierter in die Thematik eingedrungen werden. Tagging als unkomplizierteste Methode für Anwendungsfälle von Volksschülern, über das tiefere Verständnis des sprachlichen Zeichens und semantischen Dreiecks auf AHS Niveau bis hin zu computerlinguistischen Algorithmen auf Hochschulebene.
- **Zielkriterium:** Als Zielvorstellung lässt sich die Situation anstreben, in der es Informationsbedürfnisse mit möglichst geringem Aufwand zu befriedigen gilt. Dies erfordert das Erkennen des kausalen Zusammenhanges zur Inhaltserschließung und ihren grundlegenden Herausforderungen.

- **Zeitkriterium:** Das Nutzbarmachen von Wissen hat eine lange Vergangenheit. Selbst durch moderne Technik bleibt die intellektuelle Erschließung eine wichtige Methode neben automatischen Verfahren.
- **Sinnkriterium:** Für Schüler und Erwachsene existieren ausreichend Bezugspunkte, die das Stoffgebiet der Inhaltserschließung direkt (Tagging, Abstracting, Erschließen in Web 2.0 oder Dokumentenmanagementsystemen) sowie als Konsequenz (Recherche in Bibliothekskatalogen, Suchmaschinen im Web, Informationssystemen in Unternehmen) berühren. Der natürlichsprachige Bezug prädestiniert den Lehrstoff für den fächerübergreifenden Unterricht, wie mit Psychologie (siehe Kapitel 5.7), den Sprachfächern (siehe Kapitel 2.6) und der Informatik.

Schwill (1994, S. 18) erkennt Sprache als fundamentale Idee: *„Mit der Idee der Sprache verbindet sich die Zielvorstellung, daß sich alle informatikrelevanten Sachverhalte durch gewisse Zeichenfolgen beschreiben lassen, die nach einfachen Bildungsgesetzen aufgebaut sind und die untereinander auf effiziente algorithmische Weise semantikerhaltend transformierbar sind.“* Diese Art der Sprache, die nach Regeln funktioniert, offeriert den Unterschied zur natürlichen Sprache und den Gebrauch von Index- oder Dokumentationssprachen zu behandeln. Ein Computerprogramm kann nicht wie ein Text in beliebige Worte gekleidet werden. Ferner bedient sich die inhaltserschließende Methode Klassifikation (siehe Kapitel 4.7) der fundamentalen Idee *„Strukturierte Zerlegung“*, wie sie von Schwill (ebd., S. 19) genannt wird.

5.6 Lesen, Schreiben und Lernen lernen

Da für den Teilprozess der Inhaltsanalyse (siehe Kapitel 2.6) beim inhaltlichen Erschließen mit den Schritten Orientieren, Strukturieren und Auswählen der Sinn eines Dokumentes vom Bearbeiter erschlossen werden muss – *„Aufgabe ist das Verstehen des Dokumenteninhaltes“* (Lorenz, 1998, S. 9) – rückt didaktisch das dafür notwendige Thema der Lesekompetenz in das Querschnittsfeld. Die Messung der Lesekompetenz, neben der mathematischen und naturwissenschaftlichen, als Teil der regelmäßig seit 2000 durch-

geführten PISA-Studien der OECD⁵ – die jeweils zu Zeiten der Veröffentlichung ihrer Untersuchungsergebnisse hohe mediale Resonanz bewirken – betont die Wichtigkeit dieser Kompetenz in unserer Gesellschaft zum einen. Zum anderen untermauern Autoren die Relevanz von Schreib- und Lesekompetenzen bezüglich ihrer Anwendung im Alltag auf hohem Niveau über den funktionalen Analphabetismus: *„Auch wer schriftsprachliche Anforderungssituationen vermeidet, verhält sich als Analphabet/in, selbst wenn er/sie Lese- und Schreibkompetenzen besitzt.“* (Döbert & Hubertus, 2000, S. 21). In der Arbeitswelt ist das Verschlagworten, also das inhaltliche Erschließen, von selbsterstellten oder auch erhaltenen Dokumenten keine ungewöhnliche Situation: Seien hier der Empfang eines Besprechungsprotokolls eines externen Meetings und seine Ablage im betrieblichen Dokumentenmanagementsystem oder das Bereitstellen einer Anleitung über ein gemeinschaftliches Lesezeichen in einer Social Business Plattform beispielhaft genannt – beide Vorgänge bedürfen des Ausfüllens von inhaltserschließenden Feldern wie beispielsweise Schlagwörter eines kontrollierten Vokabulars, Tags oder Beschreibungsfelder, die Zusammenfassungen oder Abstracts aufnehmen.

Da ein sinnerfassendes Verständnis eines Textes auf Lesekompetenz basiert, eignet sich das informatische Thema Inhaltserschließung dafür, interdisziplinäre Aktivitäten im schulischen Unterricht zu forcieren und die Förderung der betreffenden Kompetenzen zu unterstützen. Im fächerübergreifenden Modus mit Sprachfächern ergeben sich hier mögliche Aufgabenteilungen, in denen beispielsweise im Deutschunterricht die Fertigkeiten des Orientierens und Strukturierens während der Inhaltsanalyse geübt und anschließend die Inhaltsdarstellungen im Informatikunterricht in eine Übungsdatenbank eingearbeitet werden. Den Lernenden wird hierdurch ein realitätsnaher Anwendungsfall durch exemplarisches Lehren (vgl. Schröder, 2002, S. 63f.) vorgeführt, der sich durch Erkennen des Gesamtprozesses und der Anreicherung mit für den Sprachunterricht ungewöhnlichen Werkzeugen – zum Beispiel Dokumentendatenbanken – abwechslungsreich gestaltet und sich dadurch motivational positiv auswirken kann und somit Übertragungseffekte

⁵ URL: <http://www.oecd.org/pisa/aboutpisa/> [11.12.2012]

ermöglicht. Ein Vorschlag für ein derartiges Projekt zeigt Unterrichtsmaterial 6.8.

Die Inhaltsdarstellung richtet sich nach den gewählten Erschließungsmethoden, welche Kapitel 4 erläutert. Ein weiterer fruchtbarer Anknüpfungspunkt zur schulischen Didaktik liefert die Methode des Abstractings (siehe Kapitel 4.3). Beim inhaltlichen Erschließen über das Erstellen von Abstracts werden *„[...] die Inhalte von Dokumenten im Kontext wiedergegeben [...]“* (Bertram, 2005, S. 24), also in kohärenten Sätzen und zusammenfassender Weise.

Ist das Zusammenfassen von Inhalt im Blickwinkel der Inhaltserschließung einer Orientierungs- und Zugangsfunktion in Form von Abstracts respektive Kurzreferaten förderlich (vgl. Bertram, 2005, S. 50), so verfolgt es im Kontext des Lernens eine Organisationsstrategie, um neues Wissen im lernenden Subjekt zu organisieren und zu strukturieren (vgl. Friedrich & Mandl, 2006, S. 4).

Beide Perspektiven haben gemein, sich im Prozess des Verdichtens ähnlicher Operationen zu bedienen. Dazu zählt insbesondere die Fertigkeit der Identifikation wichtiger Aussagen, wie Ballstaed (2006, S. 118) lernstrategisch konstatiert: *„Zentrale Bedingung für jede Art von Zusammenfassung ist das Erkennen der Wichtigkeit von Textaussagen (Importanz, main ideas) und damit gleichzeitig der Unwichtigkeit anderer.“* Bertram (2005, S. 46) listet diesen Aspekt als Aufgabe bezüglich inhaltlicher Erschließung.

Ballstaedt (2006, S. 117) heißt im Kontext Lernen das Zusammenfassen von Textinformation eine mentale Schlüsselkompetenz, es *„[...] ist eine multifunktionale Fertigkeit, die Kompetenzen des Lesens und Schreibens erfordert. Diese zentrale Schlüsselqualifikation in der Wissensgesellschaft spielt in der [...] Bildung in vielen Berufsfeldern eine wesentliche Rolle. Sie dient nicht nur der kognitiven Bewältigung der Informationsflut, sondern auch dem tieferen Verstehen und dem langfristigen Behalten.“*

Neben den Möglichkeiten der Schreib- und Lesekompetenzförderung bieten die verwandten Stoffgebiete Abstracting und Textzusammenfassen somit Optionen, lernstrategische Aspekte zu behandeln. Diese werden umso wichti-

ger, weil Lehrpläne durch das rasche Wachstum von Wissen dieses nicht vollständig abzudecken imstande sind, und daher Strategien für ein lebenslanges Lernen notwendig werden (vgl. Schröder, 2002, S. 28).

5.7 Querschnittscharakter des sprachlichen Aspekts

Das Thema natürliche Sprache respektive die Eigenschaften zeichenvermittelter Kommunikation siedelt sich nicht ausschließlich im Informatikunterricht an, sondern besitzt fundamentalen und somit überfachlichen Charakter. Um über die Vorstellungen einer alltäglichen Sichtweise auf natürliche Sprachen hinaus zu kommen, wird ein wissenschaftlicher Zugang notwendig, wie in Kapitel 3 mittels de Saussure und Eco verdeutlicht. Dieser bringt sich auch als Bestandteil der schulischen Bildung im Unterrichtsfach Psychologie und Philosophie ein. Der AHS-Lehrplan⁶ als Beispiel trägt diesbezüglich mit Stichworten wie „[...] *begriffliche Differenzierungsfähigkeit* [...]“, „[...] *sprachliche und kommunikative Prozesse zu fördern* [...]“ und „[...] *begriffliche Genauigkeit anwenden* [...]“ bei. Er akzentuiert unter den didaktischen Grundsätzen die „[...] *fächerübergreifenden Aspekte zu allen geistes- und naturwissenschaftlichen Fächern durch Querverbindungen aufzuzeigen*“ – hier als Vorschlag einer Querverbindung die Inhaltserschließung – und listet den Lehrstoff „*Sprache und Denken*“.

Beispielhaft als Lehrstoff sei hier das Kapitel „2.2 Begriffsbildung“ im Schulbuch *Arbeitsbuch Psychologie* angeführt, welches den Begriffsbegriff wie folgt mit – nur – einem Satz behandelt: „*Ein ‚Begriff‘ ist dann gegeben, wenn eine Anzahl von Objekten oder Ereignissen auf Grund gewisser übereinstimmender Merkmale mit einem gemeinsamen Namen belegt wird, oder allgemeiner: wenn darauf in übereinstimmender Weise reagiert wird.*“ (Illichmann, 2000, S. 218). Die Existenz des Themas der Begriffsbildung im Lehrstoff des Unterrichtsfaches Psychologie und Philosophie lädt ein, es interdisziplinär mit einem Anwendungsfeld weitreichender Konsequenzen – zum Beispiel Inhaltserschließung oder Recherche im Internet und in Bibliotheken – komplementär nach dem Prinzip der Veranschaulichung (vgl. Schröder, 2002, S. 167ff.) zu erweitern. „*Veranschaulichung im Unterricht*

⁶ URL: http://www.bmukk.gv.at/medienpool/11865/lp_neu_ahs_13.pdf [25.10.2012]

heißt, den Unterrichtsstoff so darzubieten, dass die Schüler ihn mit Hilfe ihrer Sinnesorgane und entsprechend ihrer Auffassungsfähigkeit umfassend und zutreffend erkennen können.“ (ebd., S. 167).

In Bezug auf linguistische Fächer können Unterschiede in der Betrachtung im Unterricht herausgearbeitet und aufgezeigt werden. Selbst die Literatur liefert durch widersprüchliche Feststellungen Diskussionspunkte: So stellt Fugmann (1999, S. 34) in der Differenzierung zwischen Wort und Begriff fest, dass der Ausdruck „englischer Begriff“ Irritation stiftet und damit nicht Begriffe, sondern deren Ausdrucksweise gemeint seien. Dem gegenüber steht zum Beispiel Ecos (2002, S. 86) Interpretation von Signifikaten als kulturelle Einheiten semantischer Systeme (siehe Abbildung 5), die den „englischen Begriff“ wiederum ermöglicht.

Die Behandlung der Sprache im Unterricht als Querschnittsaufgabe offeriert die fruchtbare Möglichkeit einer Abstimmung zwischen Unterrichtsfächern, wie oben geschildert mit dem Fach Psychologie und Philosophie, aber auch den Sprachfächern, um den Lernenden die Reichweite eines angemessenen Verständnisses zu illustrieren und im speziellen Fall hier, die Konsequenzen natürlicher Sprache bezüglich Nutzbarmachung von Inhalten – letztendlich dem in Dokumenten fixierten Wissen der Menschheit – aufzuzeigen.

In der Schule sowie auch in betrieblichen Schulungen begrenzt ein vorgegebenes Zeitkontingent den Umfang der Ausführung und didaktischen Aufbereitung. Die hier herangezogene Betrachtung der Begriffsbildung eignet sich für eine Unterrichtseinheit (siehe Unterrichtsmaterial 6.9) in zeitlicher Hinsicht, aber auch in lehrstofflicher, um potentielle nicht-normative Konzepte über Sprache (siehe Kapitel 3.1) aufzubrechen und zu lockern versuchen und, der Thematik hier bezüglich, Konsequenzen in der inhaltlichen Erschließung und Suche von Dokumenten zu vergegenwärtigen. Ein weiterer Zugang zur Begriffsbildung, in dem *„[...] nicht nur eine Konjunktion von Merkmalen, sondern auch mitunter eine Disjunktion von Merkmalen einen Begriff bestimmen kann“* und wie es überhaupt zu Begriffen kommt, elaboriert Stock (2009, S. 404) – eine Behandlung in diesem Detailgrad in Unterrichtssequenzen übersteigt den Aufbau eines Basiswissens.

Eine Methode, wie in Informationssystemen mit Mehrdeutigkeit umgegangen werden kann, findet im Unterrichtsmaterial 6.1 Eingang.

5.8 Inhalterschließung im „2.0“ – Kontext

In der Einleitung dieser Arbeit wurde als markanter Grund für den allgemeinbildenden Charakter des Erschließungsthemas die Involvierung breiter Massen durch das Internet festgestellt. Im Besonderen lässt sich das Web 2.0 als treibende Kraft für notwendige Kompetenzen erkennen. Web 2.0 Dienste unterstützen eine „*architecture of participation*“ (O'Reilly, 2005) und erlauben neben dem passiven Konsumieren auch aktive Beteiligung durch Kontribuieren eigener Dokumente wie Blog- und Forenbeiträge, Posts in Sozialen Netzwerken, Fotos, Videos, Wikiartikel oder Lesezeichen. Die bedeutsame Methode der Erschließung im partizipativen Web ist das Tagging (siehe Kapitel 4.8). Neben dem Tagging erlauben manche Dienste auch das Verfassen einer kohärenten Beschreibung zu einem Inhalt, hier kommt das Wissen über Abstracts (siehe Kapitel 4.3) dem Ersteller und schlussendlich dem (unvorhersehbaren) Nutzerkreis der Inhalte des Dienstes zugute.

Ferner halten Konzepte des Web 2.0 und die Entwicklung von Enterprise Social Software, welche Kommunikation, Vernetzung und Kollaboration innerbetrieblich, aber auch über Unternehmensgrenzen hinweg fördern, Einzug im betrieblichen Kontext. Unter dem Schlagwort „*Enterprise 2.0*“ (McAfee, 2006) gilt es für viele Softwarehersteller in ihren Produkten das erfolgreiche Web 2.0 Paradigma auf Unternehmen übertragbar zu machen.

Social Software wird durch McAfee (2006) anhand des Akronyms SLATES charakterisiert: Es steht für **S**earch, **L**inks, **A**uthoring, **T**ags, **E**xtension, **S**ignals. Im Kontext der vorliegenden Arbeit stechen die Schlagworte Search und Tags ins Auge und provozieren die Frage, ob für das Tagging und die Suche in Folksonomys kompetenzfördernde Maßnahmen notwendig sind?

Aus Anwendungsperspektive konstatiert Müller-Prove (2008, S. 21) hinsichtlich Social Tagging: „*Auf Seiten der Anwender bedeutet das [Social Tagging, d. Verf.], dass sie eine Medienkompetenz im Umgang mit diesen neuen semantischen Ordnung- und Visualisierungsmethoden entwickeln müssen.*“

Selbst ein Hersteller von Social Business Software sagt über Social Tagging aus, Mitarbeiter müssen dazu gebracht werden, „[...] dass sie sich das Tagging angewöhnen. Das geht nicht von heute auf morgen und das geht nicht ohne Übung, das geht nicht ohne Schulung.“⁷ Ein anderer Hersteller einer Collaboration Software deutet auf das Thema Informationssicherheit hin: „Educate users about which aspects of their social tagging activity are public and which are private. Train users to mark tags as private when they do not want other users to see that they have applied a tag to a Web page.“⁸

Die geforderten Kompetenzen betreffen qualitative Belange, in der Mitarbeiter Verantwortung über die Qualität von Folksonomys übernehmen, kulturelle Belange, in der Mitarbeiter über die Methode des Taggings aufgeklärt und für sie motiviert werden, und sicherheitsrelevante Aspekte, die über Sicherheitsrisiken des gemeinschaftlichen Taggings aufmerksam machen.

Alle drei Aspekte können durch die Integration der Inhaltserschließung und im Speziellen der Methode Tagging im Unterricht oder in Schulungen behandelt und im Lernenden das notwendige Basiswissen konstruiert werden.

Social Tagging Anwendungen existieren inzwischen viele und haben das Potenzial, an Alltagserfahrungen von Schülern anzuknüpfen. Im Internet erfreuen sich Social Bookmarking Dienste – wie Delicious.com, falkd.com oder BibSonomy.org – großer Beliebtheit. Sie ermöglichen das Teilen von Lesezeichen im Internet. Die Lesezeichen werden über Tags indexiert und sind unter anderem über Tag Clouds durchsuchbar (siehe Abbildung 19 und Abbildung 20, Seite 66). Social Bookmarking Lösungen werden auch in Betrieben von Mitarbeitern angewendet, um Dokumente innerhalb der Unternehmensgrenzen zu teilen und erschließen.

Dass Tagging ein höheres Potenzial besitzt als nur Lösungen für (Social) Bookmarking Anwendungen zu unterstützen, zeigt die Integration der Tagging-Funktionalität in Bibliothekssuchmaschinen zur Realisierung hybrider Indexierung durch Involvierung des partizipativen Nutzers. Als angemeldeter

⁷ URL: http://www.youtube.com/watch?feature=player_detailpage&v=XH5vBaXPACk , von 6m13s bis 6m22s [2013-03-10]

⁸ URL: [http://technet.microsoft.com/en-us/library/ff608006\(v=office.14\).aspx](http://technet.microsoft.com/en-us/library/ff608006(v=office.14).aspx) [2013-03-10]

Nutzer der Suchmaschine `u:search`⁹ der Universität Wien und `CatalogPlus`¹⁰ der Technischen Universität Wien ist eine Tagging Funktion verfügbar. Sie erlaubt, nachgewiesene Titel mit Tags zu verschlagworten und ermöglicht die Suche nach Titeln über Tag-Wolken oder Tag-Listen. Andere Angebote, wie das der Staats- und Universitätsbibliothek Hamburg Carl von Ossietzky¹¹ ermöglichen das Speichern von Titeln in eigenen Listen, in denen die Titel über persönliche Tags erschlossen werden können. Beispiele dieser Art können im Unterricht, idealiter im konkreten Anwendungsfall wie in Unterrichtseinheiten über Inhaltserschließung oder der Recherche für eine vorwissenschaftliche Arbeit, einfließen.

5.9 Programmierung

Aus technischer Perspektive offeriert das Unterrichtsfach Informatik interessante Anknüpfungspunkte zum Thema Inhaltserschließung.

Lernen Schüler im Unterricht das Design und die Entwicklung von Webseiten, dann führt die Fragestellung nach der Suchbarkeit der erstellten Seiten in den Fachbereich der inhaltlichen Erschließung. Suchmaschinen implementieren zahlreiche Algorithmen zur Indexierung des Webs. Der Autor einer Webseite wird durch sogenannte HTML Meta-Tags sowie Meta-Angaben nach Dublin Core befähigt, sie im HTML-Dateikopf formal und inhaltlich zu beschreiben und dadurch zu ihrer Auffindbarkeit beitragen: *„Es ist also sinnvoll, diese Meta-Angaben in allen inhaltlich relevanten HTML-Dateien zu notieren und dabei redaktionell genauso sorgfältig zu sein wie bei dem Text, der im Browser-Fenster angezeigt wird.“*¹² In Bezug auf die Inhaltserschließung sind die HTML-Meta-Tags *keywords* und *description* relevant, die wie folgt beispielhaft kodiert werden:

```
<meta name="description" content="Diese Dokument handelt von...">  
<meta name="keywords" content="Auto Automobil PKW">
```

⁹ URL: <http://usearch.univie.ac.at> [2013-03-11]

¹⁰ URL: <http://catalogplus.tuwien.ac.at> [2013-03-11]

¹¹ URL: <http://beluga.sub.uni-hamburg.de> [2013-03-11]

¹² URL: <http://de.selfhtml.org/html/kopfdaten/meta.htm> [2013-03-11]

Aus dem Dublin Core Standard¹³ eignen sich zu inhaltlichen Beschreibung insbesondere die Terme *abstract*, welcher „*a summary of the resource*“ zu definieren erlaubt, und *subject* mit dem Zweck: „*Typically, the subject will be represented using keywords, key phrases, or classification codes. Recommended best practice is to use a controlled vocabulary.*“ In einem HTML-Dokument werden Metadaten nach dem Dublin Core Standard folgendermaßen fixiert:

```
<meta name="DCTERMS.abstract" content="Dieses Dokument ..." />
<meta name="DCTERMS.subject" content="Auto" />
<meta name="DCTERMS.subject" content="Automobil" />
<meta name="DCTERMS.subject" content="PKW" />
```

In der Programmierung rückt ein weiterer Ansatzpunkt in den Fokus. Die in Kapitel 3.3 erwähnte Retrievalfunktion „ähnlich“ kann – in starker Vereinfachung – zur Veranschaulichung verschiedener Vergleichsmöglichkeiten herangezogen werden. Angenommen, eine Tabelle DOKUMENTE einer Microsoft Access Datenbank sei gegeben. Die Tabelle enthält das Feld ABSTRACT, welches Abstracts – also kohärenten Text – zu den Dokumenten speichert. Sollen in einer Abfrage alle Dokumente selektiert werden, die zum Suchkriterium „Wirtschaft“ relevant sind, so führt folgende Strategie in der Erstellung einer SQL Anweisung nicht zum gewünschten Ziel:

```
SELECT DOKUMENTE.Titel, DOKUMENTE.Autor
FROM DOKUMENTE
WHERE (((DOKUMENTE.ABSTRACT)=" Wirtschaft "));
```

Vielmehr stellt folgendes SQL-Statement eine Lösung dar:

```
SELECT DOKUMENTE.Titel, DOKUMENTE.Autor
FROM DOKUMENTE
WHERE ((InStr([DOKUMENTE].[ABSTRACT]," Wirtschaft ")>0);
```

¹³ URL: <http://dublincore.org/documents/dcmi-terms> [2013-03-11]

Diese Aufgabenstellung gibt dem Lernenden die Möglichkeit anhand eines Beispiels den Unterschied der Vergleichsmöglichkeiten von Daten zu verstehen. Zudem können Boolesche Operatoren AND sowie OR an dieser Stelle im Unterricht Eingang finden oder gegebenenfalls rekapituliert werden.

Ferner kann die Problematik Mehrdeutigkeit natürlicher Sprachen bezüglich Umsetzung in formale Sprachen wie Programmiersprachen zum Gegenstand der Betrachtung werden, um eine Übertragung der Erkenntnisse aus dem Stoffgebiet Inhaltserschließung zu erlangen. Folgendes Beispiel illustriert die mögliche Interpretation der in Java umzusetzenden Spezifikation: „Wenn sich Max im Urlaub in Kärnten verliebt, dann informiere Hans“.

Interpretation 1:

```
If (person.Name == "Max" &&
    urlaub.Ort == "Kärnten" &&
    person.IstVerliebtIn != "") Mail.Send("Hans", "Max hat sich in
Kärnten in jemanden verliebt!");
```

Interpretation 2:

```
If (person.Name == "Max" &&
    urlaub.Ort != "" &&
    person.IstVerliebtIn == "Kärnten") Mail.Send("Hans", "Max hat
sich in das Bundesland Kärnten verliebt ");
```

Neben der Programmierung vermag das Thema Modellierung die Integration der Inhaltserschließung durch sinnvolle Ansatzpunkte im fortgeschrittenen Unterricht zu rechtfertigen. Die Abbildung einer Narrow oder Broad Folksonomy bedarf eines Normalisierens in der Entwicklung eines Datenmodells. Ein Tag in einer Narrow Folksonomy wird durch die Relation Dokument und Tagname, eine Broad Folksonomy durch die komplexere Relation Dokument, Tagname und Nutzer beschrieben.

6 Unterrichtsmaterialien

Im Folgenden werden Vorschläge für Unterrichtsmaterialien für den gezielten Einbau in Lehreinheiten angeführt. Die fachlichen Grundlagen für die Materialien sind in den jeweiligen referenzierten Kapiteln zu finden.

6.1 Material: Disambiguierung im Wikipedia

Wikipedia¹⁴ ist eine Online-Enzyklopädie, zu der jeder Internetnutzer beitragen kann. Die Aufgabenstellung der Schüler ist es, Wikipedia zu untersuchen. Im Vordergrund steht der Sachverhalt der Mehrdeutigkeit. Für den Fall, dass ein Wort mehrere Begriffe bezeichnet, schlägt Wikipedia Lösungen zur Begriffsklärung¹⁵ vor. Diese sollen vom Lernenden selbstständig oder in Gruppen erkannt und in eigenen Worten anhand eines Beispiels erklärt werden.

Das Ergebnis der Untersuchung kann beinhalten:

- Pro Wikipedia-Seite wird ein Begriff behandelt
- Ein Wort kann in der Wikipedia mehrere Begriffe bezeichnen
- Zur Auflösung der Mehrdeutigkeit werden Begriffsklärungsseiten erstellt
- In Begriffserklärungsseiten wird gegebenenfalls nach Fachgebieten unterteilt und
- ein mehrdeutiger Begriff wird dann beispielsweise als „Wort (Begriff)“ aufgelöst
- Sucht man nach einem Wort, wird man zu einem anderen Artikel weitergeleitet, der ein Synonym zum Suchkriterium ist

Pro Gruppe wird ein unterschiedliches Untersuchungsergebnis der Klasse präsentiert – anhand eines Beispiels, idealiter am Beamer.

Diese Aufgabenstellung kann einerseits als Einstieg in das Stoffgebiet mit nachfolgender theoretischer Auseinandersetzung (Wort, Begriff, Synonymie/Homonymie) oder auch als praxisnaher Anwendungsfall, der nach der Behandlung der allgemeinen Theorie verortet ist, durchgeführt werden.

¹⁴ URL: <http://www.wikipedia.at> [2013-03-18]

¹⁵ URL: <http://de.wikipedia.org/wiki/Wikipedia:Begriffskl%C3%A4rung> [2013-03-18]

Die Behandlung der Disambiguierung in der Wikipedia bietet sich an, über die Natur der Online-Enzyklopädie zu diskutieren und auf mögliche Gefahren urheberrechtliche Aspekte hinzuweisen.

6.2 Material: Kennzahlen der Dokumentation

Kapitel 2.5 beschäftigt sich mit den Kennzahlen der Dokumentation, die den kausalen Zusammenhang zwischen Inhaltserschließung und Trefferliste aufzeigen. Die Thematik eignet sich aufgrund populärer Suchmöglichkeiten im Internet über Suchmaschinen didaktisch aufzubereiten. Wird jedoch eine Web-Suchmaschine wie Google herangezogen, ist unbedingt zu betonen, dass die Relevanzberechnung viele dem Anwender verborgene Algorithmen verwendet. Vorteilhafter für den Lernenden ist die Anwendung eines Bibliothekskatalogs, um hier irritierende Effekte zu vermeiden.

Nachdem den Lernenden die theoretischen Ansätze von Relevanz, Precision sowie Ballast und Verlust (siehe Abbildung 2) erklärt wurden, kann eine praktische Übung das Verständnis fördern. Als Vorbereitung sucht sich die Lehrkraft einen Onlinebibliothekskatalog (OPAC) aus, welcher für die Übung eingesetzt wird. Als Beispiel für die Lernenden bereitet die Lehrkraft eine thematische Fragestellung inklusive Suchformulierung vor, die eine überschaubare Trefferzahl im Ergebnis bringt, zum Beispiel das Thema „Förderungen von erneuerbarer Energien in Österreich“ mit den Suchkriterien „erneuerbare Energie Österreich Förderung“. Die Aufgaben für die Lernenden lauten:

- Änderung der Suchformulierung, damit sich der Recall erhöht
- Änderung der Suchformulierung, damit sich die Precision erhöht
- Feststellen von Ballast in der Trefferliste
- Erklären von Informationsverlust anhand der ersten Abfrage

Ferner kann eine Rechercheaufgabe den Lernenden gestellt werden, in der sie selbstständig die Problematik des Relevanzbegriffs im Sinne der Suchanbieter im Internet eruieren. Im Web existieren viele Hinweise darauf, dass verschiedenste Aspekte in die Berechnung einfließen, wie Präferenzen des

Nutzers oder bezahlte Links. Vom Lernenden sind das Rechercheergebnis sowie die Suchstrategie zu dokumentieren.

6.3 Material: Gegenstand – ein Ausschnitt der Wirklichkeit

Das folgende Experiment dient zur Veranschaulichung sprachlicher Unschärfe hinsichtlich des subjektiven Ausschneidens der Wirklichkeit, die in Kapitel 3 besprochen wurde. Das Experiment eignet sich für den Einstieg in die Thematik Sprache, um von einem demonstrierten Fall induktiv zum allgemeinen Wissen zu gelangen, wie den Erklärungsmodellen „Sprachliches Zeichen“ und „semiotisches Dreieck“.

Das Experiment läuft wie folgt ab: Die Klasse wird durch Durchzählen in zwei Gruppen geteilt (zum Beispiel: Durchzählen – gerade Nummern und ungerade Nummern bilden je eine Gruppe – dies ist wichtig, um nachher keinen vereinfachten Rückschluss zu ermöglichen). Die Schüler erhalten ein leeres A4 Blatt, mit der Aufgabe, Gruppe A skizziere einen Teller und Gruppe B möge eine Vase skizzieren. Auf die Blattrückseite soll dann noch das Wort (Teller oder Vase) geschrieben werden, um zu fixieren, um was es sich in der Skizze handelt. Dann werden die Artefakte eingesammelt und ausgewertet. Die Lehrkraft zeigt je eine Skizze und Schüler identifizieren den Gegenstand.

Die Auswertung kann auch statistisch erfolgen, es genügt jedoch die alleinige Demonstration, dass zum Beispiel eine gemeinte Vase nicht von allen Schülern als solche zu erkennen ist.

Dies bildet einen hilfreichen Ausgangspunkt, um Schülern mit Erklärungen über Eigenschaften der Sprache – sprachliches Zeichen, semiotisches Dreieck, Begriff, Bezeichnung, Benennung, Wort – zu begegnen.

Erst dann soll der Bezug zu Problematik sprachlicher Phänomene in der Inhaltserschließung hergestellt werden. Zur Veranschaulichung eignet sich dort beispielsweise eine Tag Cloud zu einer Ressource aus dem Internet, die viele Tags, also viele Sichtweisen enthält.

Besonders wichtig ist die Möglichkeit an dieser Stelle, dass Schüler die Problematik der Unschärfe auf andere Bereiche übertragen können, wie zum

Beispiel die Gefahr von Missverständnissen in der alltäglichen Kommunikation.

6.4 Material: Mehrdeutigkeit und Unschärfe

In Kapitel 4.1 wurde erläutert, dass automatische Erschließungsmethoden im Prozess der Inhaltanalyse den Nachteil besitzen, dass sie die Bedeutung von Dokumenten nicht eindeutig errechnen können. Die Ursache liegt in der Unschärfe und Mehrdeutigkeit der natürlichen Sprache. Sachverhalte, die in kohärenten Sätzen niedergeschrieben sind, können in ihrer Bedeutung oft mehrfach ausgelegt werden. Dieses Unterrichtsmaterial soll Lernende für diesen Umstand sensibilisieren und insbesondere die Tragweite dieses sprachlichen Phänomens verdeutlichen.

Aufgabe der Schüler ist es zuerst, anhand vorgegebener Beispiele Interpretationsmöglichkeiten zu identifizieren und zu erklären. Im zweiten Schritt suchen Schüler selbstständig nach Beispielen, indem sie im Internet Recherchieren oder eigenhändig mehrdeutige Sätze kreieren.

Als Beispiele können folgende Sätze herangezogen werden:

- Beispielsatz aus Kapitel 4.1: „*The man saw the pyramid on the hill with the telescope*“ (Vickery & Vickery, 2004, S. 201).
- Wie haben sie Essen gefunden?
- Essen ist angesagt!
- Wir müssen die Maschine einstellen.
- Jim saw the tree.

Die Lernenden beschäftigen sich mit den Sätzen, indem sie auch Wörterbücher zur Hand nehmen, und beispielsweise „einstellen“ oder „finden“ nachschlagen.

6.5 Material: Inhaltsverzeichnis erstellen

Kapitel 4.2 behandelt das Inhaltsverzeichnis als inhaltsverdichtende Methode. Inhaltsverzeichnisse werden auch in Bibliothekskatalogen als durchsuchbare Artefakte integriert.

Das Erstellen von Inhaltsverzeichnissen hängt stark mit der Gliederung eines Dokumentes zusammen. Der Lehrstoff hier kann auch mit der Lehre der Funktionen von Textverarbeitungsprogrammen (Gliederung, Formatvorlagen) kombiniert werden.

Als Vorschlag einer Aufgabenstellung, die die inhaltserschließende Seite betrifft, können von der Lehrkraft vorgelegte (umfangreichere) Dokumente von Schülern gegliedert werden, um in einer Textverarbeitungssoftware automatisch ein Inhaltsverzeichnis zu generieren. Dabei soll Wert auf aussagekräftige Überschriften gelegt werden.

Von der Rechercheite kann vorgeführt werden, wie in Bibliothekskatalogen Inhaltsverzeichnisse als durchsuchbare Texte zu Büchern verwendet werden können. Hier bietet sich beispielsweise die Vorarlberger Landesbibliothek mit der URL <http://vlb.vorarlberg.at/home.html> an. Auf dieser Website kann in der Suchmaske das Feld „Inhaltsverzeichnis“ ausgewählt werden, wie folgende Abbildung zeigt.



Vorarlberger Landesbibliothek

Indersuche | Sitzung be

Suche in In

Einfache Suche

Geben Sie ein Wort oder eine Wortfolge ein

Feld

Inhaltsverzeichnis

OK Verwerfen

Abbildung 25: Suche in Inhaltsverzeichnissen der Vorarlberger Landesbibliothek

Lernende formulieren Suchanfragen und erkennen somit einen weiteren Zweck für die Erstellung von Inhaltsverzeichnissen.

6.6 Material: Register erstellen

Anknüpfend an das Unterrichtsmaterial 6.5 lässt sich das Stoffgebiet Inhaltserschließung vertiefen. Schüler sollen a) lernen, mit Registern umzugehen und b) ein kleines Register selbst zu erstellen.

Wie in Kapitel 4.5 erklärt, können Register aus verschiedenen Perspektiven erstellt werden. Eine einfache Form des Registers ist ein Sachwortregister, in dem die Indexterme direkt aus einem vorgelegten Dokument entnommen und im Register alphabetisch angeordnet werden. Dies kann wiederum im Zuge des Lernens einer Textverarbeitung durchgeführt werden.

Diese Übung kann auch als Gruppenarbeit angedacht werden: In einem Online-Textverarbeitungsprogramm wie Google Drive oder ZoHo kann je eine Gruppe von ca. 5 Schülern gemeinschaftlich ein Register erstellen, indem Schüler 1 die Seiten 1 bis 2, Schüler 2 die Seiten 3-4 und so weiter indexieren.

Neben dem klassischen Buchregister, oder Register in einem Textverarbeitungsdokument, empfiehlt sich, Schüler auch Register von Internetauftritten (siehe Anhang II.3) untersuchen zu lassen. Schüler sollen neben der klassischen Suchfunktion in Webseiten auch lernen, in einem Register einen Überblick über das Dokument zu erhalten und gegebenenfalls zu einem bestimmten Inhalt navigieren. Neben klassischen Webauftritten weist zum Beispiel auch die Wikipedia ein Register auf („Von A bis Z“¹⁶).

6.7 Material: Tagging

Die Erschließungsmethode Tagging (siehe Kapitel 4.8) wurde vor allem in Web 2.0 Diensten (siehe Kapitel 5.8) populär. Anhand der Methode lässt sich das inhaltliche Erschließen leicht üben und veranschaulichen. Öffentliche Dienste, wie Social Bookmarking, würden sich nicht erfreuen, wenn Lehrkräfte diese für Übungseinheiten missbrauchen. Daher empfiehlt sich, eine Übungsumgebung einzurichten, die nach Abschluss der Lerneinheiten wieder bereinigt werden kann. Das Lernmanagementsystem Moodle bietet die Möglichkeit, Tagging zu aktivieren.

¹⁶ URL: http://de.wikipedia.org/wiki/Spezial:Alle_Seiten [2013-03-18]

Unter http://docs.moodle.org/23/en/Using_tags werden die Funktionen erklärt.

In einer Moodle-Instanz wird ein Kurs für den Lehrstoff der Inhaltserschließung angelegt und das Tagging aktiviert. Jeder Schüler hat dann die Aufgabe Blogbeiträge zu verfassen und diese über Tags inhaltlich zu erschließen. In Blogbeiträge können auch Fotos ohne zusätzlichen Text aufgenommen werden – dann wird die Funktion des Taggings augenscheinlich: Ohne jegliche inhaltliche Beschreibung zu einem Foto kann es kaum gezielt wiedergefunden werden.

Ist eine gewisse Menge an Blogbeiträgen von den Schülern verfasst worden (Lehrer können den Kurs mit Beispielbeiträgen vorab füllen – damit sich Schüler daran orientieren können), so bietet Moodle dann die Möglichkeit der Visualisierung von Folksonomys über Tag Clouds. Die Lernenden haben nun die Aufgabe, mit Tag Clouds bestimmte Blogbeiträge zu finden, indem sie a) sich an selbsterstellte Beiträge erinnern und b) sich von einem fremd-erstellten Eintrag erzählen lassen und diese dann über die Tag Cloud suchen.

Zudem soll die Untersuchung von Tag Clouds zur adäquaten Interpretation der verschieden groß und fett geschriebenen Tags in der Tag Cloud führen.

6.8 Material: Abstract, Zusammenfassung

Kapitel 5.6 arbeitet unter anderem heraus, dass sich das Zusammenfassen von Text und die Inhaltserschließung ähnlicher Operationen bedienen. Dies eröffnet die Option, fächerübergreifend eine Prozessteilung didaktisch zu berücksichtigen. Im Sprachunterricht kann die Inhaltsanalyse von ausgewählten Texten erfolgen, um anschließend im Informatikunterricht die Inhaltsdarstellungen in Form von Indextermen oder eines Abstracts zu realisieren.

Aus Kapitel 5.4 ergibt sich die sinnvolle Verankerung des Lehrstoffes über Abstracts im Zuge der Erstellung einer vorwissenschaftlichen Arbeit.

6.9 Material: Begriffsbildung

Diesem Material muss das sprachliche Zeichen und semiotische Dreieck (siehe Kapitel 3.1) vorausgegangen sein. Auf diesem Vorwissen aufbauend erfolgt die Erläuterung des Begriffs Begriff, welcher über eine Merkmalsliste kreiert wird, wie in Kapitel 3.2 beschrieben. Die Behandlung des Begriffs erlaubt eine fächerübergreifende Didaktik (siehe Kapitel 5.7).

Anhand eines Allgemeinbegriffs wird durch Hinzufügen und Entfernen von Merkmalen illustriert, wie sich die Menge der Referenten verändert und je nach Richtung zur Kategorie oder Individualbegriff entwickelt.

Als Beispiel kann der Allgemeinbegriff *Kraftfahrzeug* mit den Merkmalen

- hat Räder
- hat einen Antrieb
- kann etwas befördern

herangezogen werden.

Wird dem Allgemeinbegriff *Kraftfahrzeug* das Merkmal „befördert Personen“ hinzugefügt, so wird er zum Beispiel zum Individualbegriff *Personenkraftwagen*. Dadurch wird die Menge der Gegenstände, die durch den Begriff referenziert werden, geringer.

Der kategoriale Begriff *Fahrzeug* entsteht durch das Entfernen des Merkmals „hat einen Antrieb“.

Neben einer Abstraktionshierarchie *Fahrzeug* – *Kraftfahrzeug* – *Personenkraftwagen* kann auch eine Bestandshierarchie (siehe Abbildung 9) demonstriert werden.

7 Schluss

Die Disziplin der Inhaltserschließung, oder auch Sacherschließung genannt, erwächst aus Fachbereichen, die durch dedizierte Berufsbilder wie Dokumentar und Bibliothekar praktiziert werden. In der Einleitung wurde diagnostiziert, dass durch den modernen Stand der Technik heute weit mehr Menschen mit dem Thema in Berührung kommen. Es betrifft viele Internetnutzer, die ihre selbst produzierten Inhalte – aber auch fremde Inhalte – über die aktuell populäre Methode Tagging erschließen. Insbesondere verlangen Web 2.0 Dienste im Internet sowie Enterprise 2.0 Dienste in Unternehmen entsprechende Kompetenzen ihrer Nutzer, um Qualität beim Taggen zu gewährleisten. Ferner wenden Berufstätige in Unternehmen und Organisationen Dokumentenmanagementsysteme an, um Dokumente inklusive semantischer Metadaten zu speichern. Da die Berührungspunkte zu inhaltserschließenden Momenten sich in der breiten Masse identifizieren lassen, ist aus Sicht dieser Arbeit eine Verortung grundlegender Kenntnisse der Thematik im allgemeinbildenden Informatikunterricht zu empfehlen.

Die Inhaltserschließung stellt neben Suchstrategien den kausalen Zusammenhang zur Trefferlistenqualität dar. Ballast und Informationsverlust verursachen hohe wirtschaftliche Kosten, die durch ausreichendes Problembewusstsein über die Herausforderungen der inhaltlichen Erschließung reduziert werden können. Die Technik alleine vermag heute (noch) nicht die Methoden der intellektuellen Erschließung zu antiquieren. Sind im suchenden Subjekt Kenntnisse über grundlegende Phänomene der natürlichen Sprache präsent, werden Suchstrategien verbessert. Diese Kenntnisse werden durch das Verstehen der Grundzüge der Inhaltserschließung automatisch erlernt, da sie im Kern die betreffenden Themen behandelt. Zu diesen Themen zählt primär der Umgang mit der natürlichen Sprache und ihren Eigenschaften und Konsequenzen hinsichtlich dem Wiederfinden und Nutzbarmachen von Informationen. Wissen über Eigenschaften der Sprache, wie es mit dem sprachlichen Zeichen und dem semiotischen Dreieck zu verdeutlichen ist, kommt dem Lernenden in vielen Bereichen des Lebens zugute. Zudem lässt sich durch die Kriterien fundamentaler Ideen der Lehrstoff über die Natur

der Sprache anhand des Anwendungsfalles Inhaltserschließung gerechtfertigt integrieren. Die vorgeschlagene Interdisziplinarität des Lehrstoffes „Natur der Sprache“ mit anderen Fächern bewirkt im Lernenden eine Verdeutlichung der Signifikanz im Alltag und Beruf. Anknüpfungspunkte zu Sprachfächern – Lese- und Schreibkompetenz – sowie Psychologie – Sprache und Denken – eröffnen im fächerübergreifenden Modus Optionen fruchtbarer Zusammenarbeit unter Lehrenden, positive Effekte auf motivationaler Ebene im Lernenden sowie die Möglichkeit der Übertragung von Erkenntnissen in andere Gebiete. Lernstrategien, welche für das lebenslange Lernen vordergründig zu sehen sind, lassen sich mit dem Lehrstoff der Inhaltserschließung explizit ansprechen und üben.

Informationen werden in Dokumenten fixiert. Um das Verständnis des gesamten Dokumentationsprozesses zu erleichtern, schlägt diese Arbeit aus didaktischer Sicht vor, ein Konnotations- oder auch Denotationsdefizit hinsichtlich des Begriffs Dokument im Vorwissen des Lernenden zu lösen. Ebenfalls grundlegend ist die Einsicht, dass in der Regel für einen (anonymen) Nutzerkreis und nicht nur für sich selbst inhaltlich erschlossen wird. Dieser Umstand soll sich im didaktischen Design niederschlagen, indem auf den Begriff Gegenstand als subjektiv gefärbter Ausschnitt der Wirklichkeit eingegangen wird. Begriffe wie Relevanz, Pertinenz, Recall, Precision, Ballast und Informationsverlust eignen sich in Unterrichtsmaterialien zur Erklärung und Illustration sprachlicher Unschärfe und subjektiver Betrachtung der Wirklichkeit. Den Schülern bekannte Anwendungsfälle, wie die Suche im Internet oder in Bibliothekskatalogen sowie Tagging im Internet, verknüpfen den Lehrstoff sinnvoll mit Alltagserfahrungen der Lernenden.

Relevante Methoden der intellektuellen Inhaltserschließung für den Lehrstoff lassen sich auf solche reduzieren, die Schülern – aber auch Mitarbeitern in Unternehmensschulungen – größtenteils vertraut sind. Inhaltsverzeichnisse, Abstracts und Register sind aus Büchern bekannt, Klassifikationen mit hoher Wahrscheinlichkeit aus Sicht von Bibliotheksnutzern und Tagging durch den Anwendungsbereich Internet. Tagging als populäre Methode kann als freies Indexieren durch den Nutzer gesehen werden. Um qualitative Resultate

durch Tagging zu erhalten, bedarf es Kompetenzen der Erschließung und Interpretation der Taggingartefakte, die durch einen Beitrag in Unterricht und Schulung durch Behandeln der Inhaltsererschließung exemplarisch zu erwerben sind.

Der Abstract, den die Schüler als Teil ihrer Abschlussprüfung verfassen müssen, begegnet ihnen in der vorwissenschaftlichen Arbeit. Durch Kenntnis der Zielsetzung und des Aufbaus eines Abstracts, welche Gegenstand von Lektionen über inhaltliche Erschließungsmethoden ist, erschließt sich dem Schüler der Zweck des Artefaktes sowie die Vorgehensweise in der Erstellung.

Der Schule respektive dem Schulmanagement bieten inhaltlich erschlossene Dokumente, wie administrative, organisatorische und vor allem Unterrichtsmaterialien, effizienten Umgang mit Wissen. Besonders Metadatenstandards (wie LOM) in Lernmanagementsystemen erfordern auch im lehrenden Subjekt Einsicht über die Notwendigkeit einer adäquaten Erschließung von Lernobjekten, um das Teilen ihrer zu vereinfachen und Doppelarbeiten im Lehrkörper innerschulisch oder über Institutionen hinweg zu vermeiden.

Die Berücksichtigung didaktisch aufbereiteter Lektionen zum Thema Inhaltsererschließung im Unterrichtsfach Informatik – idealiter fachübergreifend – besitzt aus Sicht dieser Arbeit großes Potential die in einer Wissensgesellschaft geforderten Kompetenzen bezüglich Umgang mit Wissen zu adressieren. Die elaborierten Gründe sprechen für eine Integration in den Lehrstoff. Besonders die im Lehrplan geforderte Anknüpfung an Vorkenntnisse und Alltagsbezüge kommt dem Thema entgegen. Hinreichendes Basiswissen über inhaltliche Erschließung impliziert ebenso Wissen über die Natur der Sprache – mit diesem Rüstzeug finden Schüler von heute als Mitarbeiter von morgen einfacher Informationen – und ermöglichen es Mitmenschen durch bewusste Anwendung inhaltserschließender Methoden im (Internet)Alltag sowie im Beruf, Dokumente wiederzufinden und Informationen rückzugewinnen.

„Der Unterrichtsgegenstand Informatik leistet durch die drei Dimensionen Wissensentwicklung, Unterrichtsmanagement und Wissensdarstellung einen unverzichtbaren Beitrag zur Allgemeinbildung.“ (BMUKK, 2004, S. 2). Durch Integration des in der vorliegenden Arbeit beschriebenen Lehrstoffs Inhaltserschließung im Unterricht fände die Ergänzung einer vierten Dimension Wissensrepräsentation statt.

I Literaturverzeichnis

- Ballstaedt, S.-P. (2006). Zusammenfassen von Textinformationen. In H. F. Friedrich, H. Mandl (Hrsg.), *Handbuch Lernstrategien* (S. 117-126). Göttingen, Wien [u.a.]: Hogrefe.
- Bertram, J. (2005). *Einführung in die inhaltliche Erschließung : Grundlagen - Methoden - Instrumente*. Würzburg: Würzburg : Ergon.
- BMUKK (2000). Verbindliche Übung Berufsorientierung. URL: <http://www.bmukk.gv.at/medienpool/794/ahs19.pdf> [2013-03-01]
- BMUKK (2003). *Elektronischer Content für Österreichs Schulen. Entwicklung, Verbreitung und Pflege*. (Zl. 16.700/164-II/8/03). URL: http://www.bmukk.gv.at/medienpool/10108/eContent_Schulen.pdf.
- BMUKK (2004). Informatik. URL: http://www.bmukk.gv.at/medienpool/11866/lp_neu_ahs_14.pdf [2012-09-09]
- BMUKK (2012a). Berufsorientierung (BO). URL: <http://www.bmukk.gv.at/schulen/unterricht/ba/berufsorientierung.xml> [2013-03-01]
- BMUKK (2012b). Vorwissenschaftliche Arbeit. Eine Handreichung. Standardisierte, kompetenzorientierte Reifeprüfung an AHS. URL: http://www.bmukk.gv.at/medienpool/22700/reifepruefung_ahs_lfvwa.pdf
- Capaul, R., Seitz, H. (2011). *Schulführung und Schulentwicklung. Theoretische Grundlagen und Empfehlungen für die Praxis* (3., erw. u. aktual. Aufl.). Bern (u.a.): Haupt.
- Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics*, 50, 66-71.
- de Saussure, F. (2001). *Grundfragen der Allgemeinen Sprachwissenschaft*. (Hrsg. von Charles Bally und Albert Sechehaye. Unter Mitw. von Albert Riedlinger. Übers. von Herman Lommel. - 3. Aufl. / Mit einem Nachw. von Peter Ernst. De Gruyter). Berlin.
- DIN 1426 (1988). Inhaltsangaben von Dokumenten. Kurzreferate, Literaturberichte (Okt. 1988)
- DIN 1463-1 (1987). Erstellung und Weiterentwicklung von Thesauri. Einsprachige Thesauri (Nov. 1987)
- DIN 31623-2 (1988). Indexierung zur inhaltlichen Erschließung von Dokumenten. Gleichordnende Indexierung mit Deskriptoren (Sep. 1988)
- DIN 31631-1 (1984). Kategorienkatalog für Dokumente. Begriffe und Gestaltung (Jän. 1984)
- DIN 32705 (1987). Klassifikationssysteme. Erstellung und Weiterentwicklung von Klassifikationssystemen (Jän. 1987)

- Döbert, M., Hubertus, P. (2000). *Ihr Kreuz ist die Schrift. Analphabetismus und Alphabetisierung in Deutschland*. Münster, Stuttgart: Ernst Klett Verlag.
- e-teaching.org (2012). Standards & Metadaten. URL: <http://www.e-teaching.org/projekt/nachhaltigkeit/metadaten/> [2013-03-03]
- Eco, U. (1977). *Zeichen Einführung in einen Begriff und seine Geschichte / Aus d. Ital. übers. von Günter Memmert*. Frankfurt/Main: Suhrkamp.
- Eco, U. (2002). *Einführung in die Semiotik* (Autoris. dt. Ausg. / von Jürgen Trabant [9., unveränd. Aufl.]). München: Fink.
- Evans, R. (1994). Beyond Boolean: Relevance Ranking, Natural Language and the New Search Paradigm. *Proceedings of the 15th National Online Meeting, New York NY* (S. 121-128).
- Felber, H., Budin, G. (1989). *Terminologie in Theorie und Praxis*. Tübingen: Narr.
- Feldman, S., Sherman, C. (2001). The High Cost of Not Finding Information. An IDC White Paper.
- Ferber, R. (2003). *Information Retrieval. Suchmodelle und Data-Mining-Verfahren für Textsammlungen und das Web* (1. Aufl.). Heidelberg: dpunkt-Verl.
- Franke, F., Klein, A., Schüller-Zwierlein, A. (2010). *Schlüsselkompetenzen: Literatur recherchieren in Bibliotheken und Internet* (1. Aufl.). Stuttgart (u.a.): Metzler.
- Frants, V. I., Shapiro, J., Voiskunskii, V. G. (1997). *Automated Information Retrieval. Theory and Methods*. San Diego: Academic Press.
- Friedrich, H. F., Mandl, H. (Hrsg.). (2006). *Handbuch Lernstrategien*. Göttingen, Wien [u.a.]: Hogrefe.
- Fugmann, R. (1999). *Inhaltserschließung durch Indexieren : Prinzipien und Praxis*. Frankfurt am Main: DGD.
- Gaiser, B., Panke, S. (2008). "With my head up in the clouds" - Social Tagging aus Nutzersicht. In B. Gaiser, T. Hampel, S. Panke (Hrsg.), *Good tags - bad tags. Social Tagging in der Wissensorganisation* (S. 23-35). Münster: Waxmann.
- Gaus, W. (2002). *Berufe im Informationswesen. Ein Wegweiser zur Ausbildung* (5., überarb. Aufl.). Berlin: Springer.
- Gaus, W. (2005). *Dokumentations- und Ordnungslehre : Theorie und Praxis des Information-Retrieval*. Berlin Heidelberg: Springer.
- Genuneit, J. (2001). Wie alles anfang... - Zur Qualitätsdiskussion über Lesen- und Schreibenlernen im 16./17. Jahrhundert. In W. Stark (Hrsg.), *Qualifiziertes Alphabetisieren in Schule und Erwachsenenbildung.: Eine Fachtagung, Evangelische Akademie Bad Boll* (S. 51-76). Stuttgart, Düsseldorf, Leipzig: Ernst Klett Verlag.

- Glück, H. (1987). *Schrift und Schriftlichkeit. Eine sprach- und kulturwissenschaftliche Studie*. Stuttgart: Metzler.
- Gödert, W., Lepsky, K., Nagelschmidt, M. (2011). *Informationserschließung und Automatisches Indexieren: Ein Lehr- und Arbeitsbuch*: Springer.
- Golder, S. A., Huberman, B. A. (2006). Usage patterns of collaborative tagging systems *Journal of Information Science*, 32(2), 198-208.
- Hauer, M. (2003). Inhaltserschließung : Wissensressourcen zutage fördern. Digitalisierung von Aufsätzen und anderen Texten mit maschineller Inhaltserschließung am Beispiel der Vorarlberger Landesbibliothek. *BuB Forum für Bibliothek und Information*, 55(3), 192-196.
- Herget, J. (2004). Informationsmanagement. In R. Kuhlen, T. Seeger, D. Strauch (Hrsg.), *Grundlagen der praktischen Information und Dokumentation* (5., völlig neu gefasste Ausg., S. 245-255). München: K G Saur.
- Hjelmslev, L. (1957). "Pour une sémantique structurale" (cf. Hjelmslev, 1959). zit. nach: Eco, U. (2002). Einführung in die Semiotik (Autoris. dt. Ausg. / von Jürgen Trabant [9., unveränd. Aufl.] Vol. 105). München: Fink.
- IEEE 1484.12.1-2002 (2002). Draft Standard for Learning Object Metadata. URL: http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf
- Illichmann, A. (2000). *Arbeitsbuch Psychologie* (2. Aufl., überarb. Nachdr.). Wien: ÖBV, Pädagog. Verl.
- ISO 999 (1996). Information and documentation -- Guidelines for the content, organization and presentation of indexes
- ISO 5963 (1985). Documentation - Methods for examining documents, determining their subjects and selecting indexing terms
- Keller, R., Kirschbaum, I. (2003). *Bedeutungswandel. Eine Einführung*. Berlin [u.a.]: de Gruyter.
- Kiel, E., Rost, F. (2002). *Einführung in die Wissensorganisation : Grundlegende Probleme und Begriffe*. Würzburg: Ergon Verlag.
- Knorz, G. (2004). Informationsaufbereitung II: Indexieren. In R. Kuhlen, T. Seeger, D. Strauch (Hrsg.), *Grundlagen der praktischen Information und Dokumentation* (5., völlig neu gefasste Ausg., S. 179-188). München: K G Saur.
- Krause, U.-M., Stark, R. (2006). Vorwissen aktivieren. In H. F. Friedrich, H. Mandl (Hrsg.), *Handbuch Lernstrategien* (S. 38-49). Göttingen ; Wien [u.a.]: Hogrefe.
- Krist, H. (1999). Die Integration intuitiven Wissens beim schulischen Lernen. *Zeitschrift für Pädagogische Psychologie*, 13(4), 191-206.
- Kuhlen, R. (2004). Informationsaufbereitung III: Referieren (Abstracts - Abstracting - Grundlagen). In R. Kuhlen, T. Seeger, D. Strauch (Hrsg.), *Grundlagen der praktischen Information und Dokumentation* (5., völlig neu gefasste Ausg., S. 189-206). München: K G Saur.

- Lorenz, B. (1998). *Klassifikatorische Sacherschließung : Eine Einführung*. Wiesbaden: Harrassowitz Verlag.
- Manecke, H.-J. (2004). Klassifikation, Klassieren. In R. Kuhlen, T. Seeger, D. Strauch (Hrsg.), *Grundlagen der praktischen Information und Dokumentation* (5., völlig neu gefasste Ausg., S. 127-140). München: K G Saur.
- Mathes, A. (2004). Folksonomies - Cooperative Classification and Communication Through Shared Metadata. URL: <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html> [2013-02-22]
- McAfee, A. P. (2006). Enterprise 2.0: The Dawn of Emergent Collaboration. *MIT Sloan Management Review*, 47(3), 20-28.
- Mietzel, G. (2003). *Pädagogische Psychologie des Lernens und Lehrens* (7., korrigierte Aufl.). Göttingen [u.a.]: Hogrefe, Verl. für Psychologie.
- Müller-Prove, M. (2008). Modell und Anwendungsperspektive des Social Tagging. In B. Gaiser, T. Hampel, S. Panke (Hrsg.), *Good tags - bad tags. Social Tagging in der Wissensorganisation* (S. 15-22). Münster: Waxmann.
- O'Reilly, T. (2005). What Is Web 2.0. Design Patterns and Business Models for the Next Generation of Software. URL: <http://oreilly.com/pub/a/web2/archive/what-is-web-20.html> [2013-02-12]
- Peters, I., Stock, W. G. (2008). Folksonomies in Wissensrepräsentation und Information Retrieval. *Information - Wissenschaft und Praxis*, 59(2), 77--90.
- Posner, R. (1988). Metamorphosen des semiotischen Dreiecks. *Zeitschrift für Semiotik*, 10(3), 185-187.
- Rauch, W. (1988). *Was ist Informationswissenschaft?* : Verlag Jos. A. Kienreich, zit. nach: Stock, W. G. (2006). Information Retrieval Informationen suchen und finden. München: Oldenbourg.
- Riggert, W. (2009). *ECM - Enterprise Content Management: Konzepte und Techniken rund um Dokumente*: Vieweg+Teubner Verlag.
- Salton, G., MacGill, M. J. (1987). *Information retrieval - Grundlegendes für Informationswissenschaftler*. Hamburg [u.a.]: McGraw-Hill.
- Schröder, H. (2002). *Lernen - Lehren - Unterricht. Lernpsychologische und didaktische Grundlagen*. München ; Wien: Oldenbourg.
- Schütz, T. (2004). Dokumentenmanagement. In R. Kuhlen, T. Seeger, D. Strauch (Hrsg.), *Grundlagen der praktischen Information und Dokumentation* (5., völlig neu gefasste Ausg., S. 339-349). München: K G Saur.
- Schwill, A. (1994). Fundamentale Ideen: Eine Studie zur Methodologie der Informatik. URL: <http://ddi.cs.uni-potsdam.de/Forschung/Schriften/Methodologie1994.pdf>
- Stangl, W. (2006). Mnemotechnik. In H. F. Friedrich, H. Mandl (Hrsg.), *Handbuch Lernstrategien* (S. 89-100). Göttingen ; Wien [u.a.]: Hogrefe.

- Stock, W. G. (2006). *Information Retrieval Informationen suchen und finden*. München: Oldenbourg.
- Stock, W. G. (2009). Begriffe und semantische Relationen in der Wissensrepräsentation. *Information - Wissenschaft und Praxis*, 60(8), 403-420.
- Strauch, D., Rehm, M. (2012). *Lexikon Buch- Bibliothek- Neue Medien*: K. G. Saur Verlag GmbH.
- Umstätter, W. (2004). Szientometrische Verfahren. In R. Kuhlen, T. Seeger, D. Strauch (Hrsg.), *Grundlagen der praktischen Information und Dokumentation* (5., völlig neu gefasste Ausg., S. 237-243). München: K G Saur.
- van Rijsbergen, C. J. (1979). *Information Retrieval*. London: Butterworth; hier verwendet: <http://www.dcs.gla.ac.uk/Keith/pdf/Chapter1.pdf> [14.01.2013].
- Vander Wal, T. (2005a). Explaining and Showing Broad and Narrow Folksonomies. URL: <http://www.vanderwal.net/random/entrysel.php?blog=1635> [2013-02-20]
- Vander Wal, T. (2005b). Folksonomy Definition and Wikipedia. URL: <http://www.vanderwal.net/random/entrysel.php?blog=1750> [2013-02-20]
- Vickery, B. C., Vickery, A. (2004). *Information Science in Theory and Practice*. München: Saur.
- Wüster, E. (1959/60). Das Worten der Welt, schaubildlich und terminologisch dargestellt. *Sprachforum* 3, 3/4, 183-204. zit. nach: Felber, H., Budin, G. (1989). *Terminologie in Theorie und Praxis* (Vol. 1989). Tübingen: Narr.

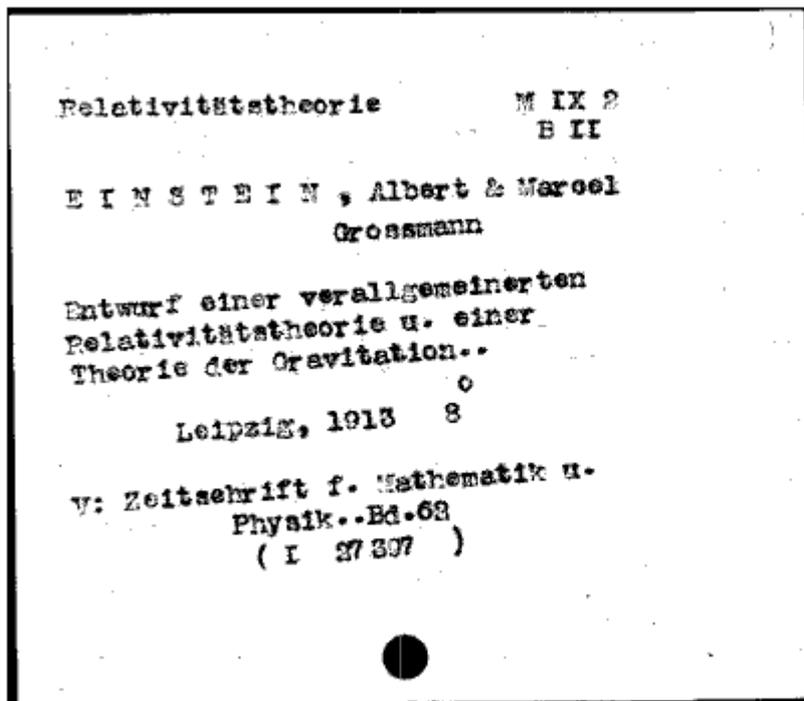
II Anhang

II.1 Abkürzungsverzeichnis

AHS	Allgemeinbildende höhere Schule
BMUKK	Bundesministerium für Unterricht, Kunst und Kultur
CC	Colon Classification
DBE	Dokumentarische Bezugseinheit
DE	Dokumentationseinheit
DIN	Deutsche Institut für Normung
DK	Internationale Dezimalklassifikation
Ebd.	ebenda
IEEE	Institute of Electrical and Electronics Engineers
IPC	Internationale Patentklassifikation
IR	Information Retrieval
ISO	International Organization for Standardization
LMS	Learning Management System
LOM	Learning Objects Metadata
OECD	Organisation for Economic Co-operation and Development
OPAC	Online Public Access Catalogue
RIS	Rechtsinformationssystem
SCORM	Sharable Content Object Reference Model

II.2 Beispiel eines alten Kartenkatalogs

Im kleinen Lesesaal des Hauptgebäudes der Universität Wien „befindet sich aus technischen Gründen nach wie vor der alte Schlagwortkatalog bis 1931. Sie können alle alten Kartenkataloge auch als Image-Katalog auf unserer Homepage aufrufen.“¹⁷ Abbildung 26 zeigt exemplarisch eine Karte, welche ein Werk von Albert Einstein und Marcel Grossmann formal und inhaltlich erschließt.



Zettel: 752776

Abbildung 26: Karte eines Zettelkataloges¹⁸

Links oben druckt die DE das Schlagwort „Relativitätstheorie“ an, womit es inhaltlich erschlossen wird. Das Schlagwort dient zugleich als Ordnungskriterium im zugrundeliegenden Katalog. Zu den formalen Kriterien zählen die Autorennamen „E I N S T E I N, Albert & Marcel Grossmann“, der Titel „Entwurf einer verallgemeinerten Relativitätstheorie u. einer Theorie der Gravitation“, Ort „Leipzig“ und das Jahr „1913“ der Veröffentlichung in der Zeitschrift „Zeitschrift f. Mathematik u. Physik..Bd.63“. „I 27397“ ist die Signatur, unter der das Werk in der Aufstellung zu finden ist.

¹⁷ URL: http://bibliothek.univie.ac.at/hauptbibliothek/kleiner_lesesaal.html [2013-02-05]

¹⁸ URL: <http://www.univie.ac.at/ubwdb/cgi-bin/katzoom.cgi?katalog=1> [2013-02-05]

II.3 Beispiele für Produkte inhaltlicher Erschließung

Register einer Website anhand OECD Auftritt¹⁹:

The screenshot shows the OECD website's 'A to Z Index (A-B-C)' page. At the top, there is the OECD logo with the tagline 'BETTER POLICIES FOR BETTER LIVES'. To the right, there are social media links and a search bar. The main navigation menu includes 'OECD Home', 'About', 'Countries', 'Topics', 'Statistics', and 'Newsroom'. Below the navigation, the page title is 'About the OECD' and the sub-page title is 'A to Z Index (A-B-C)'. A sidebar on the left lists various categories like 'Who does what', 'What we do and how', etc. The main content area displays a grid of links for each letter from A to Z. The 'A' section is expanded, showing a list of links such as 'About OECD', 'Acronyms and Abbreviations', 'Administrative Tribunal', 'Africa', 'Africa Investment Initiative (NEPAD-OECD)', 'Africa Partnership Forum', 'African Economic Outlook', 'Agreements', 'Agricultural Certification and Standardisation', 'Agricultural Outlook', 'Agricultural Policies in OECD Countries: At a Glance', 'Agricultural Policy Monitoring and Evaluation', 'Agriculture', 'Aid for Trade', 'Aid for Trade at a Glance', 'Aid statistics', 'Alerts by e-mails', 'Ambassadors', 'Annual Budget', 'Annual Report: OECD Secretary-General's Report to Ministers (.pdf of 3.38 MB)', 'Anti-Bribery Convention', 'Architecture for schools and universities', 'Archives', 'At a Glance Publications', 'Audio Material', 'Auditor-General', 'Australia', and 'Austria'. The 'B' section is also visible, listing links like 'Badoes (how to obtain)', 'Balkans', 'Belgium', 'Berlin Centre', 'Bestsellers', 'Better Life Index', 'Better Life Initiative', and 'Bioeconomy'.

Abbildung 27: Beispielregister einer Website

¹⁹ URL: <http://www.oecd.org/about/atozindexa-b-c.htm> [2013-04-15]

Beispiel für einen Registereintrag:

Abbildung 28 illustriert einen Registereintrag, dessen Registerausgang auf einen Absatz im Buch als Dokumentationseinheit verweist. Die Seite 102 und der exemplarische Registereintrag aus Fugmann's Werk wurden eingescannt und zur Veranschaulichung übereinander gelegt.

291. Daher ist es eine wichtige

Aufgabe der Indexierung,
die Ausdrucksweisen für Begriffe
aus den Dokumenten in eine
die zumindest *besser voraus-*
ungeregelt natursprachige A

292. Nur dort, wo schon die natursprachig
es angenähert bei den Individualbegriffen
eine Übersetzung in eine Indexsprache ve

293. Aus dieser Sicht ist für uns im folgen

"Indexieren"
erstens das Erkennen der (wie
Essenz eines Textes und
zweitens das Wiedergeben die
spezifischen und gut vorausse-
d.h. „indexsprachigen“ Form.

294. Bietet die Indexsprache nicht genüge
Indexierung und Anfrage nur ungenau bes
man aber nichts anderes erwarten als ents

295. Das *Klassifizieren* von Dokumenten is
ante vom Indexieren, indem hier ausschlie
onssystem Verwendung finden, welche do

Hat man diese Aufgabe einer Indexsprach
Reihe von Regeln für die Anwendung derselben. Sie werden in den nächsten Kapiteln
abgehandelt.

**Voraussehbarkeit der Ausdrucks-
weise**

*siehe auch natursprachiger Aus-
druck, Mannigfaltigkeit der
Ausdrucksweise, Informations-
verlust, Programmierbarkeit,
Volltexte*

mangelhaft bei Allgemein-

begriffen 76 86B 89 318

nicht ersetzbar durch Gedächtnis

113

Wegfall beim freien Indexieren

354

beim kontrollierten Indexieren

345

durch Indexsprachenwortschatz

112 125 172 290 291 293 340

350

gut bei Individualbegriffen 73B

86B 99 292 293 319

entbehrlich beim Kleinsystem

115

mangelhaft bei Natursprache 89

290 318 322 342

Wegfall bei Paraphrasen 318

Erzielung nicht programierbar 90

Erfordernis bei Recherche 84 90

288 289

bei Schlagwörtern 85

durch Thesaurus 172

gering bei Fragethesaurus 332

333 336

Wegfall bei Untereinträgen 472

Abbildung 28: Beispielregister mit Absätzen als DBE (Fugmann, 1998, S. 102 und
Registereintrag ‚Voraussehbarkeit der Ausdrucksweise‘)

II.4 Relationenkürzel im Thesaurus

Kurzzzeichen zu deutschen Benennungen			Kurzzzeichen zu englischen Benennungen	
Kurzzzeichen		Benennung	Kurzzzeichen	Benennung
zwei-stellig	ein-stellig			
Wenn Abstraktions- und Bestandsrelation nicht unterschieden werden:				
TT		Top Term (Kopfbegriff einer Hierarchie)	TT	Top term
OB		übergeordneter Begriff	BT	Broader term
UB		untergeordneter Begriff	NT	Narrower term
VB	V	Verwandter Begriff (Assoziationsrelation)	RT	Related term
Bei Differenzierung der Hierarchierelationen verwendet man:				
Entweder:				
OA	O	Oberbegriff (Abstraktionsrelation)	BTG	Broader term generic
UA	U	Unterbegriff (Abstraktionsrelation)	NTG	Narrower term generic
oder:				
SP	S	Verbandsbegriff (Bestandsrelation)	BTP	Broader term partitive
TP	T	Teilbegriff (Bestandsrelation)	NTP	Narrower term partitive
Kurzzzeichen für Verweisungen:				
BS	B*)	Benutze Synonym oder Quasi-Synonym	USE	Use
BF	B*)	Benutzt für Synonym oder Quasi-Synonym	UF	Used for
BK		Benutze Kombination von Deskriptoren	USE	Use
KB		Benutzt in Kombination	UFC	Used for combination
BO		Benutze Oberbegriff	USE	Use
FU		Benutzt für Unterbegriff	UF	Used for
*) Wenn das Kurzzzeichen „B“ benutzt wird, müssen Deskriptor und Nicht-Deskriptor durch unterschiedliche Darstellung (z. B. Typographie) gekennzeichnet werden.				

Abbildung 29: Relationenkürzel eines Thesaurus (DIN 1463-1, 1987, S. 11)

II.5 Unternehmen Julius Blum GmbH

Die Julius Blum GmbH mit Hauptsitz in Höchst/Vbg ist ein international tätiges Unternehmen, welches Möbelfunktionsbeschläge wie Scharnier-, Klappen- und Auszugsysteme sowie zugehörige Verarbeitungshilfen herstellt. Das Unternehmen beschäftigt ca. 5700 Mitarbeiter weltweit. Im Jahre 1996 wurde zur Unterstützung von Kommunikation und Kollaboration die Groupware Lotus Notes/Domino gruppenweit eingeführt. Lotus Domino deckt einerseits die Anforderungen der E-mailkommunikation und des Kalender- und Ressourcenmanagements ab, andererseits dient es als Dokumentenmanagementsystem. Mehr als 3700 Benutzer arbeiten systemweit in knapp 3000 Lotus Notes Datenbanken auf 52 Dominoservern in 24 Ländern. Das Dokumentenaufkommen pro Datenbank bewegt sich zwischen einigen Wenigen bis zu 250.000. Alle Dokumente nehmen eine Speicherkapazität von rund 3 Terabyte in Anspruch. In Lotus Notes Datenbanken werden verschiedenste Typen von Dokumenten gespeichert, welche ein breites Spektrum von Anforderungen abdecken. Dieses reicht von der Ablage von Informationen, über Dokumente, welche in Workflows involviert sind, bis hin zur sicheren Verwaltung von personenbezogenen, projektbezogenen und unternehmenskritischen Dokumenten. Durch den integrierten Replikationsmechanismus von Lotus Domino wird der Anforderung eines weltweit verteilten Systems Rechnung getragen. Das rasante Wachstum der Unternehmensinformationen generiert ein erhebliches Datenaufkommen in Form von Dokumenten in Lotus Notes Datenbanken. Um Verbesserungen in der Suchbarkeit zu fördern, sollen Möglichkeiten einer fortschrittlichen Inhaltserschließung herausgearbeitet und ausgewählte Lösungswege umgesetzt werden. Dies erfolgt durch eine IST-Zustandsanalyse, Erhebungen in Form von Abstimmungen mit Fachabteilungen, Interviews mit Spezialisten der Lotus Notes Abteilung und Dokumentenmanagementgruppe sowie der Recherche in wissenschaftlichen Publikationen.

Lebenslauf

Name	Günter Lautner
Geburtsjahr	1974
Staatsbürgerschaft	Österreich
Schulabschluss	1995: HTL für EDV und Organisation, Spengergasse, Wien V
Universität	2003 – 2004: Ernährungswissenschaften, Universität Wien 2004 – 2013: Lehramtsstudium UF Informatik und UF Chemie, Universität Wien und TU Wien 2007 – 2009: Bakkalaureatsstudium Informatikmanagement, Universität Wien und TU Wien
Beruf	Informatiker
Zertifikat	Principal CLP (Certified Lotus Professional)
Stipendien	Selbsterhalterstipendium der Stipendienstelle Wien 2 x Leistungsstipendium der Universität Wien
Auslandsaufenthalt	November 2002 bis Mai 2003 in Neuseeland Dezember 2009 bis Februar 2010 in Neuseeland
Fremdsprachen	Englisch