

Diplomarbeit

**Terminologiarbeit
im Bereich
Wissensorganisation**

von

Stefan Hauser

betreut von

Dipl.Soⁱⁿ./Wiss.Dok.ⁱⁿ Jutta Bertram

Fachhochschul-Studiengang Angewandtes Wissensmanagement
Eisenstadt 2007

Ehrenwörtliche Erklärung

Ich habe diese Diplomarbeit selbständig verfasst, alle meine Quellen und Hilfsmittel angegeben, keine unerlaubten Hilfen eingesetzt und die Arbeit bisher in keiner Form als Prüfungsarbeit vorgelegt.

Ort und Datum

Unterschrift

Abstract

The aim of this thesis is to provide a concise survey of current developments within German-speaking terminology work in the domain of knowledge organisation and an analysis of and comparison between two recently re-edited printed publications and the Wikipedia portal devoted to this subject.

In the first part, key concepts are defined and located within the subject matter and a theoretical framework is provided. In the second part, the field ‘thesaurus’ is singled out in order to exemplify terminology work. A catalogue of criteria in both form and content is developed in order to analyse and compare entries related to this field, permitting conclusions in terms of the different treatment of the field in the abovementioned publications.

In doing so, various aspects of terminology work are taken into account: Firstly, the methodical level dealing with models and processes applied, secondly, the organisational level scrutinising the ways terminology is produced and thirdly, the functional level dealing with the outcomes. In the second part, special emphasis is put on this latter aspect, by analysing and comparing entries pertinent to the field in terms of macrostructure, microstructure, and mediostructure. Furthermore, consistency of definitions is looked into.

The comparative part of the thesis demonstrates differences between the publications by applying the catalogue of criteria to the treatment of ‘thesaurus’. These differences are ascribed to constitutive factors of each publication, such as methods of production (non-collaborative vs. collaborative) or the anticipated user (expert vs. non-expert).

Keywords: knowledge organisation, terminology, terminology work, thesaurus

Stichwörter: Wissensorganisation, Terminologie, Terminologearbeit, Thesaurus

Executive Summary / Abriss

Die vorliegende Arbeit setzt sich zum Ziel, eine Untersuchung von Terminologiearbeit im Bereich Wissensorganisation vorzunehmen. Um Aktualität zu gewährleisten, wurden zwei Publikationen ausgewählt, die innerhalb der letzten Jahre in neuen Auflagen herausgebracht worden sind, nämlich die *Grundlagen der praktischen Information und Dokumentation* (2004, im Folgenden abgekürzt als GpID) und die *Terminologie der Information und Dokumentation* (2006, abgekürzt TID). Um einen Vergleich über das Medium Buch hinaus zu ermöglichen, wird außerdem als neues, elektronisches Medium Wikipedia (Stand: Oktober 2006, abgekürzt Wp) herangezogen. Da eine Untersuchung zu Terminologiearbeit als solcher den Rahmen dieser Arbeit sprengen würde, wird ein Themenkomplex herausgegriffen und Terminologiearbeit und ihre Produkte mit seiner Hilfe exemplifiziert. Die Wahl fiel auf den Komplex Thesaurus, aufgrund der zentralen Rolle, die dieses Instrument in der Wissensorganisation seit einigen Jahrzehnten einnimmt. Publikationen, die sich den Methoden der Wissensorganisation widmen, kommen um eine ausführliche Behandlung dieses Instruments daher nicht herum.

Die leitenden Fragen der Arbeit können so formuliert werden:

- Wie wird Terminologiearbeit in den genannten Publikationen realisiert?
- Wie sehen demgemäß die Produkte dieser Terminologiearbeit aus?
- Welche Unterschiede sind dabei feststellbar und auf welche Faktoren können sie zurückgeführt werden?

Bevor diese Fragen behandelt werden können, ist es angebracht, Klarheit über die Voraussetzungen und die grundlegenden Begriffe zu schaffen. Häufig wird „Wissensorganisation“ sehr weit verstanden als jede Form des Ordners oder In-Beziehung-zueinander-Setzens von Wissen oder von Begriffen. Ein sehr enges Verständnis meint dagegen damit nur die Bearbeitung bibliografischer Daten. Eine wieder andere Interpretation setzt sie mit Wissensmanagement gleich (so in GpID, S. 131). Auf die verschiedenen Formen von Wissen wird in der Arbeit nicht eingegangen. Die Rolle von Begriffen und ihren sprachlichen Repräsentati-

onen, den Bezeichnungen, sowie die Beziehung zwischen diesen beiden Aspekten sehe ich jedoch als wesentlich an. Erst aufgrund der nicht eindeutigen Zuordnung von Bezeichnungen zu Begriffen werden explizite Festlegungen erforderlich. Dazu gehören Thesauri ebenso wie Termini, deren Zusammenstellung in Terminologien und ihre Erstellung im Rahmen von Terminologearbeit. Die Stufen, die für eine systematische Durchführung von Terminologearbeit durchlaufen werden müssen, werden ebenfalls dargestellt. Schließlich werden im Theoretischen Teil noch Überlegungen über die Verschiedenartigkeit der untersuchten Publikationen aufgrund der medialen Form (gedruckt vs. digital) und der Methoden in der Erstellung (einzelner Autor vs. kollaborative Verfasserschaft) angestellt und über die Folgen, die sich für die Terminologearbeit daraus ergeben.

Der Praktische Teil eröffnet mit einer kurzen Charakterisierung der drei zu untersuchenden Publikationen und gibt einen Einblick in ihren Aufbau. Dies ist eine wesentliche Voraussetzung zur anschließenden Typologisierung der Publikationen, um ihnen ihren spezifischen Platz im Bereich Fachliteratur zuzuweisen.

Wichtig dabei ist, die textsortenspezifischen Unterschiede zwischen den Publikationen im Auge zu behalten, auch wenn die untersuchten Aspekte jeweils die gleichen sind. Glossar, Register mit Definitionsteil und Enzyklopädie sind je verschiedene Textsorten, in denen die Behandlung des gleichen Themas in je anderer Weise erfolgen muss, um den jeweiligen Zielsetzungen angemessen zu sein.

Zunächst werden, ausgehend vom Eintrag ‚Thesaurus‘ und von dort den Verweisen folgend, die einschlägigen Lemmata aufgelistet. Bei allen zu erwartenden Gemeinsamkeiten zeigen sich hier bereits einige Unterschiede, wie sie dann bei der mikrostrukturellen Untersuchung der Definitionen deutlicher zu Tage treten. Als Vergleichsfolie werden zudem die in DIN 1463 *Erstellung und Weiterentwicklung von Thesauri* verwendeten Termini herangezogen.

Um konkrete Unterschiede herauszuarbeiten, erweist sich Roelckes Ansatz als sehr hilfreich: Im Anschluss daran wird die Anordnung der Artikel in den Publikationen untersucht (Makrostruktur), die Anordnung der Angaben innerhalb der Artikel (Mikrostruktur) und die Vernetzung der Artikel miteinander (vgl. Roelcke 2005, S. 132ff.).

Inhalt

Abstract	I
Executive Summary / Abriss	II
Verzeichnis der verwendeten Abkürzungen	V
1 Einleitung	1
2 Theoretischer Teil	2
2.1 Begriffsklärungen	2
2.1.1 Wissensorganisation	2
2.1.2 Terminologie.....	9
2.1.2 Terminologiearbeit	10
2.3 Methoden und Modelle der Terminologiearbeit	12
2.3.1 Vorstufen systematischer Terminologiearbeit	13
2.3.2 Systematische Terminologiearbeit	14
2.4 Überlegungen zum Vergleich von „konventionellen“ mit „neuen“ Publikationen... 16	
2.5 Folgerungen	23
3 Praktischer Teil	25
3.1 Darstellung der drei zu untersuchenden Publikationen.....	25
3.2 Typologisierung der Publikationen und Analyseschema.....	29
3.3 Auflistung der Lemmata.....	31
3.4 Makrostruktur	34
3.5 Mikrostruktur	35
3.5.1 Lemma.....	36
3.5.2 Kopf.....	38
3.5.3 Einträge	41
3.6 Mediostruktur	47
4 Zusammenfassung und Ausblick	53
5 Quellen	57
6 Anhang	60
7 Lebenslauf	66

Verzeichnis der verwendeten Abkürzungen

BID	Bibliothek, Information, Dokumentation
DIN	Deutsches Institut für Normung bzw. Deutsche Industrie-Norm
EB	Encyclopædia Britannica
GpID	Kuhlen/Seeger/Strauch (2004): Grundlagen der praktischen Information und Dokumentation
IuD	Information und Dokumentation
TID	Beling/Port/Strohl-Groebel (2006): Terminologie der Information und Dokumentation
Wp	Wikipedia

1 Einleitung

Die vorliegende Untersuchung befasst sich mit Terminologearbeit im Bereich Wissensorganisation (wie diese Ausdrücke hier verstanden werden sollen, wird im Abschnitt 2.1 *Begriffsklärungen* dargelegt). Nach einer Phase einer gewissen Vernachlässigung ist dieses Thema in den vergangenen Jahren wieder stärker in den Vordergrund gerückt, was sich auch in einer Reihe von Publikationen niederschlägt. Drei davon werden hier herausgegriffen und zu einem Überblick und Vergleich über aktuelle Ansätze deutschsprachiger Terminologearbeit im genannten Feld herangezogen: Kuhlen/Seeger/Strauch: *Grundlagen der praktischen Information und Dokumentation* (2004, kurz: GpID), Beling/Port/Strohl-Goebel: *Terminologie der Information und Dokumentation* (2006, kurz: TID) und das Wikipedia-Portal *Bibliothek, Information, Dokumentation* (Stand: Oktober 2007, kurz: Wp). Exemplarisch wird dabei der Themenkomplex Thesaurus ausgewählt, da er ein zentrales Instrument der Wissensorganisation darstellt, und anhand seiner beschrieben, zu welchen verschiedenen Ergebnissen Terminologearbeit kommt und welche Qualitätskriterien dabei erfüllt werden.

Im Mittelpunkt der Betrachtung stehen dabei Fragestellungen unter folgenden Aspekten:

- Fachlich-inhaltliche Aspekte, mit einer ergebnis- bzw. produktorientierten Fokussierung und Fragen wie: Wie widersprüchlich oder konsistent fallen Begriffsklärungen und die Definition von Begriffsbeziehungen aus? Auf welche Grundlagen (z.B. Normen) wird dabei recurriert?
- Methodische Aspekte, welche die Grundlage für das Obige bietet: Welche unterschiedlichen Methoden von Terminologearbeit existieren und welche Methoden werden im jeweiligen Fall angewandt?
- Aspekte der Arbeitsorganisation: Wie hängen Unterschiede bei den Ergebnissen mit Unterschieden in der Art des Erstellungsprozesses zusammen? Im konkreten Fall: Sind Unterschiede zwischen den „klassischen“ gedruckten Publikationen und der digitalen und kollaborativen Methode auszumachen?

Um die Vergleichbarkeit zwischen den drei „Publikationen“ (unter diesen Ausdruck sollen sie im Folgenden subsumiert werden) herzustellen, werden Analyse-kriterien ausgearbeitet.

Als Grundlage dazu werden zunächst jedoch einige theoretischen Voraussetzungen sowie die wesentlichen Begriffe geklärt, um einerseits den Untersuchungsgegenstand der vorliegenden Arbeit einzugrenzen (Terminologiarbeit) und den Bereich herauszuarbeiten, in dem diese Terminologiarbeit geleistet wird (Wissensorganisation).

2 Theoretischer Teil

2.1 Begriffsklärungen

2.1.1 Wissensorganisation

Um das Thema „Terminologiarbeit im Bereich Wissensorganisation“ behandeln zu können, ist es zunächst erforderlich, den Bereich „Wissensorganisation“ ein- und von benachbarten Feldern abzugrenzen, um so das Feld dieser Untersuchung eindeutig zu bestimmen. Dafür eignen sich zum einen die grundlegende, einführende Literatur in diesem Bereich, zum anderen die Selbstbeschreibung von Institutionen, die sich gebildet haben, um in diesem Bereich Erfahrungsaustausch und Kommunikation zu ermöglichen.

Eine erste Annäherung mag der Artikel *Wissensorganisation* in Wikipedia bieten, wo lakonisch festgehalten wird:

Unter Wissensorganisation fasst man verschiedene Methoden und Systeme zur Organisation von Informationen bzw. dem mit ihnen ausgedrückten Wissen zusammen. Dabei spielen vor allem die Dokumentation – d.h. das Sammeln, Ordnen und Verfügbarmachen – sowie das Informations- und Wissensmanagement eine Rolle.¹

Und, etwas konkretisierend, wenngleich nicht sehr präzise, im Artikel zur *International Society for Knowledge Organization*:

Die Wissensorganisation betrachtet Methoden und Systeme zum Ordnen von Information und Wissen, z.B. Datenbanken, Bibliotheken, Wörterbücher, Enzyklopädien, Taxonomien, Klassifikationen, Wikis. Aufgrund des

¹ <http://de.wikipedia.org/wiki/Wissensorganisation> [zuletzt aufgerufen am 19.10.2007]

*interdisziplinären Gegenstandes der Wissensorganisation stammen die Mitglieder aus unterschiedlichen Wissenschaftsdisziplinen wie Informationswissenschaft, Informatik, Linguistik, Philosophie u.a.*²

Wissensorganisation wird dabei von Wissensrepräsentation abgegrenzt:

*Während bei der Wissensorganisation der Schwerpunkt auf der Ordnung bestehender Wissensbestände mit Hilfe von Metadaten liegt, versucht man im Rahmen der Wissensrepräsentation die Informationen direkt abzubilden. Die Grenzen sind jedoch fließend, da die meisten Systeme zur Wissensorganisation bereits gewisse Strukturen repräsentieren und jede Wissensrepräsentation [eine] gewisse Strukturierung voraussetzt. Auch Glossare und Enzyklopädien enthalten sowohl geordnete Begriffe (Organisation) als auch Definitionen (Repräsentation). [...] In der Regel geschieht die Wissensorganisation mit Hilfe von verschiedenen Arten von Begriffssystemen, die unter anderem helfen, eine einheitliche Terminologie festzulegen.*³

In ihrem Aufsatz *Wissensorganisation – eine neue Wissenschaft?* zeichnet Inge-
traut Dahlberg die Begriffsgeschichte nach (Dahlberg 1994, S. 225f.). Der Aus-
druck „Wissensorganisation“ bzw. „Organization of Knowledge“ kann bis zum
Beginn der 1930er Jahre zurückverfolgt werden, als er im Titel zweier Bücher von
Henry Evelyn Bliss aufschien.⁴ Da nach Dahlbergs Angaben bei der Gründung
der *Internationalen Gesellschaft für Wissensorganisation (ISKO)* die Benennung
„Wissensordnung“ alternativ vorgeschlagen wurde, der bei der Gründung der Ge-
sellschaft für Klassifikation als Äquivalent zu Klassifikation galt, würde daraus
folgen, dass Wissensordnung das Gleiche wäre wie Klassifikation, worauf auch
die Namensänderung der Zeitschrift von „International Classification“ zu „Know-
ledge Organization“ hindeuten könnte. Wenn Dahlberg allerdings im weiteren
Verlauf „Knowledge Organization“ als den „zusammenfassenden Begriff“ be-
zeichnet, wird er in einem weiteren Sinne verwendet als „Klassifikation“ (Dahl-
berg 1994, 225f.). „Knowledge Organization“ bezeichne

*die Objekte und Tätigkeiten der Begriffstheorie, der Klassifikation und In-
dexierung und der Wissensrepräsentation.* (Dahlberg 1994, S. 226)

² http://de.wikipedia.org/wiki/International_Society_for_Knowledge_Organization [19.10.2007]

³ <http://de.wikipedia.org/wiki/Wissensorganisation> [zuletzt aufgerufen am 19.10.2007]

⁴ „The Organization of Knowledge and the System of the Sciences“ (1929) und „The Organization of Knowledge in Libraries“ (1933)

Unter Wissensrepräsentation fasst sie dabei „nicht nur die logische Struktur begrifflicher Darstellungen“, sondern auch „alle Probleme der Begriffsformulierung durch geeignete Benennungen“, d.h. die Terminologie (ebd.).

In einem nächsten Schritt sucht Dahlberg zu klären, mit welchem Gegenstandsbereich es die Wissensorganisation zu tun hat und bestimmt diesen – abweichend von der oben zitierten Definition – aus der in der Benennung enthaltenen Begriffskombination heraus und analysiert (verkürzt, aber sinngemäß): Wissensorganisation organisiert Wissen. „Wissen im Sinne von Gewußtem“ wird etwas klarer, dabei aber „völlig subjektbezogen“ definiert als

subjektiv oder objektiv zureichend begründete Gewißheit vom Bestehen einer Tatsache oder eines Sachverhalts. Wissen ist nicht übertragbar, es kann nur selbst erarbeitet werden durch Nachdenken.“ (Dahlberg 1994, S. 226)

Diese Definition ist in der Tat völlig subjektorientiert. Sie kursiert im Wissensmanagement-Alltag in der popularisierten Variante „Wissen ist etwas, was seinen Ort, salopp formuliert, zwischen zwei Ohren hat und nicht zwischen zwei Modems“ (Malik 2004, S. 63). Ohne sich auf die philosophisch-erkenntnistheoretische Frage nach der Zureichendheit von Gewissheit oder Plausibilität als Wahrheitskriterium einlassen zu wollen (bzw. zu können), sei immerhin darauf hingewiesen, dass eine solch subjektivistische Wissensdefinition dem Gebrauch des Ausdrucks bei Dahlberg selbst zuwiderläuft, es sei denn, man versteht „Wissensorganisation“ so eng, dass nur mentale Prozesse des jeweiligen Subjekts darunterfallen. Tatsächlich zielt Wissensorganisation jedoch nicht auf die Etablierung von begrifflichen Systemen im Bewusstsein eines Subjekts (was durchaus auch damit verbunden sein mag), sondern eben auf die intersubjektiv nachvollziehbare „Tätigkeit der Begriffstheorie“, wie Dahlberg es in ihrer ersten Definition ausgeführt hat. Mit anderen Worten: Es sind keine „Gewissheiten“, die organisiert werden, sondern Begriffe (welchen ontologischen Status man diesen auch immer zuschreiben und welches Kommunikationsmodell man in weiterer Folge vertreten mag)⁵. Übrigens wird die oft gebrauchte Rede vom Wissen als „zwischen den Ohren“ in der Literatur kaum je durchgehalten und bald stellt sich heraus, dass

⁵ Vgl. etwa Roelckes Ausführungen zum Status von Fachwörtern (Roelcke 2005, S. 53ff.)

nun doch Informationen gemeint sind (als Beispiel dafür mag die Beschreibung eines Aus- und Weiterbildungsangebots in Wissensmanagement von Auer Consulting & Partner dienen)⁶.

Es ist übrigens aufschlussreich, dass Dahlberg diese ihre stark subjektivistische Definition von Wissen in der zwölf Jahre später auf Englisch erschienenen und überarbeiteten Version des Artikels zwar zunächst wiederholt, dann aber stark abgeschwächt hat. Sie erklärt zwar weiterhin, dass Wissen immer das Wissen von jemandem sei, aber:

we do know – have the experience and certainty about this fact – that it can yet be shared via our ability to deal with the ,things of the world’, and by using our linguistic abilities to express our experience and insights.
(Dahlberg 2006, S. 12)

Ausdrücklicher als früher spricht sie von der Vermittelbarkeit jedweder Art von Wissen. Das vornehmste Werkzeug des Menschen dazu ist die Sprache.

Dass der Begriff Wissensorganisation über rein subjektive, d.h. intrapersonale, Prozesse hinausgeht, darauf verweist auch die alternativ dazu erwogene Bezeichnung „Wissensordnung“, ohne dass man das schon als *ordo rerum* (im traditionellen Sinne) verstehen müsste. Solche Versuche einer begrifflichen Struktur, welche die der realen Welt inhärente Ordnung abbildet, finden sich bis in unsere Zeit:

In unserer Information Coding Classification haben wir – von den Objekten der Wirklichkeit ausgehend – neun ontische Bereiche angenommen, die im Sinne der Integrationsstufentheorie nach J.K. Feibleman und Nicolai Hartmann aufeinander aufbauen und sich jeweils von Aspektbereichen spezifizieren lassen. (Dahlberg 1994, S. 232)

Wie bereits angedeutet, erscheint bei Dahlberg der Wissensbegriff unzureichend geklärt. Exemplarisch zeigt sich dies an dem Widerspruch zwischen der Definition „[...] Wissen ist nicht übertragbar, es kann nur selbst erarbeitet werden durch Nachdenken“ und der zugeschriebenen Eigenschaft in der nur zwei Absätze später gebrauchten Wendung „tradierbare[s] Wissen“ (Dahlberg 1994, S. 226). Die anfangs scharfe Unterscheidung wird wieder verwischt. Entscheidend ist hier aber der Schritt, von der notwendigen „Repräsentation“ des Wissens zu sprechen. Diese ermöglicht die Sprache. Damit rückt nun die Sprache in den Mittelpunkt der

⁶ http://www.hrm-auer.ch/downloads/NDK_Fly.pdf [19.10.2007]

Betrachtung. Ohne Sprache ist Wissensorganisation nicht zu leisten; sie ist Instrument dieser organisierenden Tätigkeit und auch Objekt, insofern eine Herausbildung von Begriffen ohne Sprache kaum denkbar ist.⁷

Dahlberg zieht einen Vergleich zum Aufbau der Materie und sieht die Wissens-elemente (= Begriffsmerkmale, man könnte auch sagen „die Seme“) als Atome, die Wissenseinheiten (= Begriffe, Bündel von Semen) als Moleküle und die Wissenskomplexe (= Begriffsverknüpfungen, d.h. Aussagen) als Molekülverbände. Von *Wissenselementen*, *Wissenseinheiten* und *Wissenskomplexen* zu sprechen halte ich für zumindest missverständlich; denn nach der Definition von Wissen als einer „Gewissheit vom Bestehen einer Tatsache oder Sachverhaltes“ (Dahlberg 1994, S. 226) sollte man besser von *Bedeutungselementen* etc. sprechen, da es hier eben nicht um einen Bezug zur außersprachlichen Welt geht.

Für die Wissensorganisation ist Wissen immer explizites Wissen, und das heißt: sprachlich darstellbares und mitteilbares Wissen. Die Sprache ist im Sinne des Bühlerschen Organon-Modells das Zeichensystem, das die Codierung und Übermittlung von Wissen ermöglicht.

Zwar ist es richtig zu sagen, Wissensorganisation befasse sich mit der Organisation von Begriffen (so indirekt Dahlberg 1994, S. 225), und diese seien als abstrakte Entitäten vorsprachlich. Diese Betrachtungsweise erweitert jedoch unnötig die Kluft zwischen Begriffen und den ihnen zugeordneten Benennungen. Da Wissen in kulturell relevanter Form immer kodiert und kommuniziert werden muss, dies aber (so wie Wissen oben definiert wurde) immer in sprachlicher Form geschehen muss, so ist die Sprache und ihre Untersuchung letztlich auch eine unverzichtbare Voraussetzung von Wissensorganisation. Zumal Begriffe und Begriffssysteme in ihrer Genese natürlich ohnehin aufs engste mit der Sprache verknüpft sind.

Häufig ist davon die Rede, dass Begriffe *Wissen* enthalten würden. Betrachten wir den Begriff Junggeselle. Er wird konstituiert bzw. enthält die Seme ‚Mensch‘, ‚männlich‘, ‚unverheiratet‘, ‚im heiratsfähigen Alter‘. Statt „Seme“ würde Dahlberg sagen „Wissenseinheiten“; diese wären das „in Begriffen enthaltene Wissen“

⁷ Auffällig ist, in welchem Ausmaß die Bedeutung der Sprache in manchen Bereichen des Wissensmanagements ignoriert oder sogar geringgeschätzt wird, etwa durch die derzeit vorherrschende Überbetonung der Kommunikationsleistung von Visualisierungen.

(z.B. Dahlberg 1994, S. 227). Ist es aber sinnvoll zu sagen, der Begriff Junggeselle enthalte das „Wissen“ ‚männlich‘, ‚unverheiratet‘ usw.? Das scheint mir dem Sprachgebrauch zuwider zu laufen. „Ein Junggeselle ist ein unverheirateter Mann“, hier handelt es sich um eine analytische Aussage, die kein Wissen, jedenfalls keinen Wissenszuwachs mit sich bringt. Es erscheint mir im Sinne eines nicht-inflationären, klar abgegrenzten Gebrauchs von „Wissen“ sinnvoll, den Begriff auf sprachlich explizit darstellbare, nicht-analytische Urteile einzuschränken.

Die Frage „Wissensorganisation – eine neue Wissenschaft?“ (so der Titel des Aufsatzes von Dahlberg, 1994, S. 229ff.) soll hier nicht behandelt, allerdings Zweifel daran angemeldet werden, ob Inhalt und Arbeitsweise der Wissensorganisation den Kriterien entsprechen, die für eine Wissenschaft notwendig sind. Wissenschaftlich ist sicher die Untersuchung der Struktur von Begriffen und ihre Hierarchisierung – was allerdings die sprachwissenschaftliche Teildisziplin der Semantik leistet, deren Ergebnisse die Wissensorganisation übernimmt.

Indirekt, aber auf eine sehr griffige Weise definiert Dahlberg den Begriff Wissensorganisation, indem sie die Aufgaben von damit befassten Institutionen (etwa der *ISKO*) umreißt:

[...] die notwendigen begrifflichen Arbeiten forschend und alle Wissenschaften durchforstend weiterzubringen [...]. Diese Arbeiten können nicht den Terminologen und Wissensorganisatoren allein überlassen bleiben, es müssen die Wissenschaftler selbst Anteil nehmen, da sie ja auch daran interessiert sein sollten, daß ihre ‚Produkte‘, d.h., das durch sie gefundene neue Wissen, in einer reproduzierbaren Ordnung wiedergefunden werden kann. Da es heute also mehr denn je darum gehen muß, neues Wissen auf bestehendes Wissen zu beziehen durch entsprechende begriffliche Arbeiten, weil wir für unser Wissen die notwendige Ordnung schaffen, bzw. wiederherstellen wollen, ist die Arbeit an den Begriffen ein Anliegen, das unseres Erachtens höchste Aufmerksamkeit verdient. Es geht dabei um eine Analyse bestehender und neuer Begriffe, deren reproduzierbare Zuordnung, sinnvolle Einordnung und ihre wiederauffindbare Anordnung, kurz also, um eine objekt- und methodengerechte Wissensorganisation.
(Dahlberg 1994, S. 235)

Birger Hjørland zitiert in seinem Artikel *What is Knowledge Organization (KO)?*

J.D. Andersons Definition von Wissensorganisation:

The description of documents, their contents, features and purposes, and the organization of these descriptions so as to make these documents and

their parts accessible to persons seeking them or the messages that they contain. Knowledge organization encompasses every type and method of indexing, abstracting, cataloguing, classification, records management, bibliography and the creation of textual or bibliographical databases for information retrieval. (Andersen 1996, zit. in Hjørland⁸)

Hjørland weist darauf hin, dass Wissensorganisation sowohl den Vorgang als auch das Ergebnis der organisierenden Tätigkeit umfasst, d.h. den Prozess wie auch das Produkt. In weiterer Folge differenziert er den Begriff Wissensorganisation je nach seinem Umfang:

1. Im engeren Sinne befasst sich Wissensorganisation mit bibliografischen Daten („bibliographical records“)
 - Bibliografische Daten in Bibliothekskatalogen
 - Etwas weiter gefasst: bibliografische Daten in (elektronisch zugänglichen) Datenbanken, z.B. unter Einsatz von Thesauri
 - Noch etwas weiter: Wissensorganisationssysteme auch in anderen Institutionen des ‚kulturellen Gedächtnisses‘, z.B. Archive und Museen, deren Prinzipien der Wissensorganisation jedoch von den gleichen epistemologischen Grundsätzen abgeleitet seien.
 - Wiederum weiter: die Erschließung von Volltext- und Internet-Dokumenten. Um diesen Bereich in Wissensorganisation einzuschließen, wäre deren Definition zu modifizieren.
2. Im weiteren Sinne beschäftigt sich das Fach damit, wie etwa in wissenschaftlichen Disziplinen, Enzyklopädien, Wörterbüchern, Sprachen, Genres, Theorien und im Bewusstsein Wissen geordnet ist (man beachte den hier offenbar auch zugrundeliegenden weiteren Wissensbegriff). Dabei seien die Ordnungsprinzipien einerseits Abbild der Struktur der (jeweiligen) Wissenschaft und ihrer Begriffe, andererseits immer auch kulturell bedingt.

Im Rahmen der vorliegenden Arbeit soll „Wissensorganisation“ zunächst im hier zuletzt beschriebenen weiteren Sinne verstanden werden, um im zweiten Teil anhand eines Anwendungsfalles im engeren Sinne verwendet und betrachtet zu werden.

Im letzten Abschnitt ihres Artikels geht Dahlberg auf „Knowledge Organization and the Tasks of ISKO“ ein (Dahlberg 2006, S. 17f.) und kommt dabei ganz all-

⁸ http://www.db.dk/bh/Lifeboat_KO/CONCEPTS/knowledge_organization.htm [19.10.2007]

gemein auf die mit zunehmender Verwissenschaftlichung und Vertechnologisierung der modernen Welt gewachsene Rolle der Wissensorganisation zu sprechen:

It seems of primary importance today, to relate new knowledge to existing knowledge by corresponding conceptual work, as we need to recreate the necessary order for our knowledge. And this work deserves, in my eyes, much more attention in order to keep track of scientific development and make it transparent by analyzing existing and new concepts, defining and possibly assigning them to their pertinent position in a systematic order and to render them retrievable from there for any application. (Dahlberg 2006, S. 18)

In ähnlicher Weise bestimmt die Deutsche Sektion der ISKO auf ihrer Homepage selbst ihre Aufgaben, indem sie festhält, ihr Feld Wissensorganisation

klärt Grundlagen und Prinzipien der Prozesse und Formen, unter denen Wissen produziert, gestaltet, kommuniziert, dokumentiert und recherchiert wird. [...] Bei der Arbeit an Begriffordnungen, Terminologien, Ontologien, der Aufbereitung und Synthese von Wissen oder dem Management von Wissen stellt sich die Frage nach der jeweils geeigneten Wissensstrukturierung immer wieder aufs neue.⁹

2.1.2 Terminologie

Im Folgenden soll der Begriff Terminologie umrissen werden, allerdings in aller gebotenen Kürze, was insofern ausreichend scheint, als es hierbei in der Literatur große Übereinstimmung gibt – es hat sich für den praktischen Gebrauch eine Begriffsbestimmung durchgesetzt, wie sie etwa die Arbeitsgruppe *Terminologie und Dokumentation der Konferenz der Übersetzungsdienste europäischer Staaten (KÜDES)* bietet:

Terminologie – [im engeren Sinn] Gesamtbestand des spezifischen Wortschatzes eines Fachgebiets, im weiteren Sinn auch die Terminologearbeit, -lehre und -forschung (KÜDES 2002, S. 95)

Zwar ergeben sich hier durchaus weitere Fragestellungen, aus sprachwissenschaftlicher Sicht etwa diejenige, was genau unter einem Wortschatz zu verstehen sei, aus Sicht des jeweiligen Fachgebiets Fragen nach der Abgrenzung zu benachbarten Fachgebieten. Dennoch hat sich obige, sehr praktikable Definition durchgesetzt. So bieten Arntz et al. keine explizite Definition von Terminologie, setzen sie aber in der Einleitung zu ihrem Buch gleich mit den Begriffen Fachwörter und

⁹ So die Definition der Deutschen Sektion der Internationalen Gesellschaft für Wissensorganisation (ISKO): <http://www.bonn.iz-soz.de/wiss-org/Programmatis.htm> [19.10.2007]

Fachwortbestände (Arntz et al. 2004, S. 1f.)¹⁰. All diese Festlegungen spiegeln die Definition von Terminologie (auch: Fachwortschatz) durch DIN 2342 als „Gesamtbestand der Begriffe und ihrer Benennungen in einem Fachgebiet“ (DIN 2342, S. 3).

Etwas weiter geht Budin, der zwar auf die Uneindeutigkeit des zugrundeliegenden Begriffs Terminologie hinweist, sich aber ebenfalls an Definitionen wie die oben angeführten anschließt und sie durch das Element der Strukturiertheit erweitert:

Terminologie = strukturierte Gesamtheit der Begriffe und der diesen zugeordneten Repräsentationen eines Fachgebietes. (Budin 1996, S. 16)

Für die Zwecke der vorliegenden Arbeit erscheint diese Definition ausreichend.

2.1.2 Terminologiearbeit

Die weiter oben bereits erwähnte Arbeitsgruppe *Terminologie und Dokumentation* der *Konferenz der Übersetzungsdienste europäischer Staaten (KÜDES)* definiert Terminologiearbeit als die Arbeit von Terminologen folgendermaßen:

Terminologen sammeln und prüfen die Terminologie eines bestimmten Fachgebiets in einer, oft aber auch in mehreren Sprachen (übersetzungsorientierte Terminologiearbeit). Sie zeichnen die Fachwortschätze auf, legen bei Bedarf Fachwörter fest (terminologische Festlegung) oder prägen neue, und sie stellen Terminologien in Terminologiesammlungen zusammen. Die Ergebnisse dieser Arbeit, insbesondere die der systematischen Terminologiearbeit, werden den Benutzern in Fachwortlisten, Glossaren oder Fachwörterbüchern zur Verfügung gestellt (terminologische Lexikographie, Terminographie) oder sind bei Terminologie-Datenbanken abrufbar. (KÜDES 2002, S. 13)

Terminologiearbeit bedarf als theoretischer Grundlage der Terminologielehre als der „Wissenschaft von den Fachwortschätzen“ (ebd.). Ein besonderes Naheverhältnis besteht zur Sprachwissenschaft, deren Untersuchungsgegenstand und Untersuchungsmethoden sich teilweise mit den ihren decken, im Unterschied zu welcher sie sich allerdings (meist) auf die Betrachtung einer Fachsprache zu einem bestimmten Zeitpunkt beschränkt, deren Ergebnisse zudem in eine Sprachnormie-

¹⁰ Wichtig erscheint mir ihr Hinweis auf die sprachpolitische Bedeutung von Terminologien bzw. von Terminologiearbeit, da besonders kleinere Sprachen umso eher ins Hintertreffen geraten, je mehr Fachbereiche es gibt, für die sie über keinen Wortschatz verfügt (Ausbau einer Sprache) (Arntz et al. 2004, S. 2).

rung münden – in der Ausdrucksweise des linguistischen Strukturalismus: Synchronie statt Diachronie, oftmals Präskription statt reiner Deskription.¹¹

Terminologearbeit, nach einer anderen Definition „die Erarbeitung, Bearbeitung oder Verarbeitung von Terminologie“ (Arntz et al. 2004, S. 3), ist eine wesentliche Voraussetzung für eine störungsfreie Kommunikation vor allem in Bereichen, in denen komplexes Wissen zur Anwendung kommt, also etwa der technischen Produktion oder der Logistik. Aufgrund des Aufwandes an intellektueller Arbeit und zunehmend auch von elektronischen Systemen ist sie arbeits- und kostenintensiv (z.B. Kosten für die Erfassung und Analyse von Dokumenten, Datenbankabfragen und -erstellung, Fachliteratur, Personalkosten allgemein) und macht daher „ein methodisch korrektes Arbeiten zwingend erforderlich“ (ebd.).

Eine Differenzierung zwischen Terminologearbeit und Terminologiemanagement nimmt Budin vor, indem er Letzteres zwar weit definiert „als jegliche Art von zielgerichtetem Umgang mit terminologischer Information“ (Budin 2006, S. 453), dabei aber in Analogie zum englischen Ausdruck *information management* stärker auf moderne Formen dieses „zielgerichteten Umgangs“ fokussiert, durch deren Einsatz die Diversität und die zunehmende Komplexität terminologischer Systeme beherrscht werden sollen, indem etwa soziale Interaktion zu Etablierung und Gebrauch von Termini digital stattfindet (Budin 2006, S. 457). Zur Unterstützung der kommunikativen, wissensorganisatorischen oder informationstechnologischen Funktionen setzt Terminologiemanagement auf Instrumentarien wie Terminologiedatenbanken, elektronische Thesauri und Ontologien sowie webbasierte Systeme zur Nutzung elektronischer Daten. Besonders Ontologien mit ihrer Möglichkeit zu differenzierterer Darstellung von Begriffsbeziehungen haben dadurch einen Aufschwung erfahren. In ähnlicher Weise gebrauchen Arntz et al. den Ausdruck „rechnergestützte Terminologearbeit“ (Arntz et al., Kap. 9, S. 228ff.).

Budin bietet auch ein von den oben durch Arntz beschriebenen Arbeitsschritten etwas abweichendes „Prozessmodell des Terminologiemanagements“, das auf

¹¹ Zumindest grob gesprochen. Genauer unterscheiden Wright/Budin zwischen „Descriptive Terminology Management“ und „Prescriptive Terminology Management“ (Wright/Budin 1997, S. 329).

ISO 15188 *Project management guidelines for terminology standardization* be-
ruht (Budin 2006, S. 461f.):

*Projektplanung und Designphase (mit den Faktoren Personal-, Zeit-, Fi-
nanzressourcen sowie Abgrenzung des Fachgebiets, Wahl eines Terminolo-
giemanagementsystems und Festlegung seiner Eintragsstruktur)*

*Zusammenstellung und Analyse von Sprachressourcen (eventuell Kombi-
nation von automatisierter mit intellektueller Terminologieextraktion)*

*Terminologieerfassung (mithilfe auch sonst üblicher Datenbanken oder
spezieller Terminologieverwaltungssysteme, dabei auch Aufzeichnung le-
xikalischer Ambiguitäten)*

*Revision, Aufbereitung, Verbreitung (mit regelmäßiger Qualitätssiche-
rung, Darstellung gemäß der Benutzererwartungen)*

*Nutzung, Feedback (z.B. durch Internetzugänge und Integration in Wis-
sensmanagementsysteme, laufende Vervollständigung und Aktualisierung)*

Schließlich sei noch auf die sehr praktisch orientierte Beschreibung des *Deutschen Terminologieportals* hingewiesen. Darin werden die Aspekte „Konstruktion, Produktion, Einkauf, Marketing und Verkauf, Vertragswesen, technische Dokumentation und Übersetzung“ genannt, sowie der Zusammenhang mit Normung, gesetzlichen Auflagen und verbrauchergerechter Sprache, als Bereiche, die wesentlich von den Ergebnissen der Terminologiearbeit abhängen.¹²

Im Rahmen der vorliegenden Arbeit soll Terminologiearbeit als der weitere Begriff gebraucht und der Ausdruck „Terminologiemanagement“ nur dann verwendet werden, wenn explizit (rein) elektronisch gestützte Terminologiearbeit gemeint ist.

2.3 Methoden und Modelle der Terminologiearbeit

Das Ziel von Terminologiearbeit besteht in der Erarbeitung von systematisch aufgebauten, vollständigen und widerspruchsfreien Terminologien, in denen im Idealfall eindeutige Beziehungen zwischen Begriffen und Benennungen bestehen. Dieser Zustand ist in hochformalisierten Kunstsprachen erreicht, in vielen naturwissenschaftlichen Bereichen annäherungsweise verwirklicht, wird in Technik und Wirtschaft verstärkt angestrebt und ist in vielen Bereichen der Human- und

¹² Deutsches Terminologie-Portal: „Terminologiearbeit im Unternehmen“ <http://www.iim.fh-koeln.de/dtp/terminologieDT.html> [19.10.2007]

Geisteswissenschaften in weiter Ferne oder kaum erreichbar (man denke etwa an Soziologie oder Philosophie).

Grundsätzlich können mit von Buol (von Buol 2000, S. 29) drei Typen der Terminologiearbeit unterschieden werden: *Neuerstellung*, *Modifikation* und *Pflege*.

- *Neuerstellung*: Eine Terminologie (Fachwörterbuch) wird gänzlich neu erstellt
- *Modifikation*: ein bereits bestehendes Wörterbuch wird bearbeitet (etwa die Erweiterung einzelner Teile oder im Rahmen mehrsprachiger Terminologiearbeit eine Übersetzung)
- *Pflege und Wartung*: ein bestehendes Wörterbuch wird laufend aktualisiert.

Die vorliegende Arbeit beschäftigt sich mit Terminologiearbeit in der Ausprägung ‚Modifikation‘, wie sie in den im praktischen Teil zu untersuchenden Publikationen ihren Niederschlag findet.

Neben diesen drei Typen können weitere Stufen von Terminologiearbeit unterschieden werden: *deskriptiv* und *normend*, wobei letztere erstere immer voraussetzt. Hinsichtlich ihrer Reichweite kann die ‚genormte‘ Terminologie *intern* in Industrie- und Produktionsunternehmen Gültigkeit besitzen (als hauseigene Terminologie) oder darüber hinaus *extern*, z.B. gesamt- oder überstaatlich, vor allem wenn sie von Normungsgremien festgelegt wird (Terminologienormung) (Arntz et al. 2004, S. 227). Die praktischen Vorteile einer solchen terminologischen Normung liegen auf der Hand und zeigen sich mit der wachsenden ökonomischen, politischen, sozialen und kulturellen Verflechtung umso deutlicher.

2.3.1 Vorstufen systematischer Terminologiearbeit

- In der Alltagspraxis müssen – etwa bei einer erforderlichen Übersetzung – oft einzelne Begriffe punktuell behandelt werden, wobei dann höchstens noch die unmittelbar benachbarten Begriffe im jeweiligen Feld mit einbezogen werden können (dazu und zum Folgenden: Arntz et al. 2004, S. 217f.). Die Ergebnisse werden elektronisch gespeichert, um bei erneutem Bedarf bzw. für eine systematischere Bearbeitung zur Verfügung zu stehen.
- Das beschriebene Vorgehen kann mit der Zeit zu Kompilationen von Fachwörtern führen, indem die Ergebnisse punktueller Problembearbeitung einfach

alphabetisch gesammelt werden. Da dies allerdings oft ohne einheitliche Leitlinien geschieht, sind solche Sammlungen häufig unvollständig, willkürlich in der Auswahl und formal uneinheitlich oder mangelhaft.

- Um die beschriebene Willkür bei der Auswahl der zu behandelnden Begriffe wenigstens etwas einzudämmen, kann ein Fachgebiet zunächst in Untereinheiten gegliedert und diesen die jeweils zu bearbeitenden Begriffe zugeordnet werden. Eine tiefgehende Strukturierung des Begriffsfeldes unterbleibt allerdings, weswegen auch hier eine vollständige Abdeckung des Gebiets schwer erzielbar bzw. kontrollierbar ist.

2.3.2 Systematische Terminologiearbeit

Idealerweise jedoch wird Terminologiearbeit in systematischer Form geleistet, d.h. die Terminologie eines Fachbereichs wird von einem Terminologen als Ganzes und im Zusammenhang erschlossen, statt punktuell wird also flächendeckend über ein gesamtes Fachgebiet gearbeitet. Die Arbeitsschritte, die dazu erforderlich sind, stellen Arntz et al. dar (Arntz et al. 2004, S. 219ff.; *kursiv* die dort aufgelisteten Arbeitsschritte):

- *Organisatorische Vorüberlegungen:* Diese umfassen etwa Fragen nach dem bisherigen Bearbeitungsstand des jeweiligen Fachgebiets, der Zielsetzung und der Zielgruppe, für die die Arbeit geleistet werden soll; zudem eine Vorklärung über die Form der Veröffentlichung der Ergebnisse (z.B. Wörterbuch/Glossar oder terminologische Datenbank), weil davon die Art der Information abhängt, die im Zusammenhang mit den Begriffen gesammelt werden soll.
- *Abgrenzung des Fachgebietes:* Wohl in den seltensten Fällen sind Fachgebiete isoliert zu betrachten, weil sie keine Überlappungen mit benachbarten Gebieten aufweisen. Aus arbeitsökonomischen und systematischen Gründen ist jedoch eine genaue Abgrenzung erforderlich. Nützlich dafür kann der Einsatz von Klassifikationen sein.
- *Aufteilung des Fachgebietes in kleinere Einheiten:* Die Abgrenzung des Fachgebiets ist auch eine Voraussetzung für die Aufteilung in überschaubare Untereinheiten.

- *Beschaffung und Analyse des Dokumentationsmaterials:* Die Ergebnisse von Terminologiarbeit hängen wesentlich vom zu analysierenden Material, man könnte sagen: den Belegstellen für die Verwendung von Benennungen, ab. Um eine gute Qualität des Materials zu gewährleisten, sollte auf Kriterien wie Muttersprachlichkeit des Autors, dessen fachliche Kompetenz und die Aktualität der Quellen Bedacht genommen werden. Bei der Beschaffung von derartigem Dokumentationsmaterial (etwa einschlägige Normen, Fachlexika und fachspezifische Lehrbücher) erweist sich die Kooperation mit BID als sinnvoll.
- *Sammlung und vorläufige Zuordnung der gefundenen Benennungen und Begriffe:* Die aus dem Material extrahierten Informationen werden zu terminologischen Datensätzen zusammengefasst, die standardisierte Elemente (Datenfelder) enthalten.
- *Erarbeitung der Begriffssysteme:* Aus den im Ausgangsmaterial verwendeten Begriffen wird ein möglichst vollständiges Begriffssystem des Fachgebiets entwickelt und dessen Gültigkeit mit Experten abgeklärt.
- *Bearbeitung des Materials im Systemzusammenhang:* Nunmehr können für die Begriffe Notationen vergeben und dadurch auch Synonyme zusammengeführt werden; zudem werden endgültige Definitionen formuliert und die Zusatzinformationen ausgewählt, die zur Verfügung gestellt werden sollen. Somit liegt der vollständige terminologische Datensatz vor.
- *Bereitstellung für den Benutzer:* Je nach intendiertem Nutzerkreis und gewähltem Medium werden die Ergebnisse zugänglich gemacht.

Nach den hier dargestellten Stufen handelt es sich auch bei den drei im Praktischen Teil zu untersuchenden Publikationen zweifellos um Fälle systematischer Terminologiarbeit, wenngleich in unterschiedlichem Ausmaß. Während TID eine möglichst vollständige Abdeckung des behandelten Fachgebiets anstrebt, nimmt GpID eine Auswahl der Lemmata nach „einer gewissen Repräsentativität“ vor (GpID, Einleitung). Bei Wp kann vor allem die Bündelung der Einträge mittels thematischem Portal als Versuch gewertet werden, die Systematik zu erhöhen.

Grundsätzlich unterscheidet von Buol zwei „eigentlich konträre Vorgehensweisen“ in der praktischen Terminologiearbeit, nämlich „einerseits nach einer kooperativen Durchführung, andererseits nach einer formal fundierten, nach Konsistenz ausgerichteten Begriffsbildung“ (von Buol 2000, S. 28). Interessant – und auch im Hinblick auf die Themenstellung dieser Arbeit relevant – ist die hohe Bedeutung, die von Buol kooperativen Formen der Terminologiearbeit zumisst, um zu qualitativ hochwertigen Ergebnissen zu gelangen. Der normenden Terminologiearbeit sei die kooperative Vorgehensweise inhärent, da hier Experten eines Fachgebietes (oder sich überlagernder Fachgebiete) dessen (oder deren) Terminologie standardisieren, was „möglichst hohe Objektivität und Akzeptanz“ erfordert (von Buol 2000, S. 31f.). Auch über die reine Normung hinaus ist jedoch Kooperation und Arbeitsteilung aufgrund steigender Komplexität von Terminologiearbeit unabdingbar.

2.4 Überlegungen zum Vergleich von „konventionellen“ mit „neuen“ Publikationen

Im Dezember 2005 veröffentlichte Jim Giles in *Nature* einen Beitrag über einen Qualitätsvergleich der beiden Nachschlagewerke Encyclopædia Britannica (EB) und Wikipedia.¹³ Dabei waren für 42 naturwissenschaftliche Themen in beiden Werken die je gleichen Artikel untersucht und verglichen worden, und zwar erstmals anhand von *peer reviews*, also von Experten des jeweiligen Fachgebiets im Rahmen eines „Blindversuchs“, d.h. ohne zu wissen, welcher Artikel aus welchem Werk stammt. Das Ergebnis: Es hätten sich in jedem EB-Artikel durchschnittlich drei Mängel (Fehler, Ungenauigkeiten, Auslassungen, irreführende Angaben) gefunden, in jedem Wikipedia-Artikel vier Mängel. Auch wenn das immerhin um ein Drittel mehr Mängel bedeutet, so erscheint das doch relativ wenig, wenn man die Erstellungsmethode bei beiden Werken bedenkt. An gravierenden Fehlern (z.B. falsche Erklärungen von Grundbegriffen) seien insgesamt sogar nur acht Stück gefunden worden, vier pro Werk. Alles in allem also, möchte man meinen,

¹³ Giles (2005): Internet encyclopaedias go head to head.
http://www.nature.com/news/2005/051212/pf/438900a_pf.html [02.07.2007]

ein durchaus passables bis sogar sehr erfreuliches Ergebnis für Wikipedia. Es soll allerdings nicht verabsäumt werden, auf die im März 2006 erfolgte Gegendarstellung vonseiten der EB hinzuweisen,¹⁴ in welcher die Methoden von *Nature* als ungeeignet und die Ergebnisse als ungültig dargestellt werden, und die im selben Monat postwendend erfolgende Antwort von *Nature*.¹⁵ Diese Debatte drehte sich allerdings m.E. zu sehr darum, wie viele Fehler nun tatsächlich vorhanden waren, und verlor darüber andere Aspekte aus den Augen, die ebenfalls genannt und die gerade auch vom pädagogischen Standpunkt aus für ein Nachschlagewerk wichtig sind, nämlich die schlechte *readability* vieler Wikipedia-Artikel („poorly structured and confusing“, Giles 2005) und der unverhältnismäßige Umfang, der umstrittenen oder gar wissenschaftlich abwegigen Theorien eingeräumt wird.¹⁶ Zur Behebung solcher Unzulänglichkeiten und der generellen Verbesserung der Qualität beabsichtigt Jimmy Wales, Mitbegründer von Wikipedia, einerseits Experten stärker zur Erstellung von Artikeln zu motivieren und andererseits für jeden Artikel eine ‚stabile‘ Version festzulegen, die ein einmal erreichtes Qualitätsniveau festschreibt und von der aus dann erst weitere Veränderungen vorgenommen werden können (dazu Giles 2005). So kündigt Giles im *New Scientist Magazine* im September 2007 die bevorstehende Einführung eines neuen Systems bei Wikipedia an, das Vandalismus und Missbrauch erschweren soll, indem als „vertrauenswürdig“ eingestufte Versionen von Artikeln ausgezeichnet werden.¹⁷

Überlegt wird ein zweistufiges Verfahren: Sogenannte *gesichtete Versionen* sind als frei von Vandalismus eingestuft worden,¹⁸ während *geprüfte Versionen* nach

¹⁴ Encyclopaedia Britannica (2006): *Fatally Flawed – Refuting the recent study on encyclopedic accuracy by the journal Nature*; http://corporate.britannica.com/britannica_nature_response.pdf [19.10.2007]

¹⁵ Autor dieser Zeilen fände es – im Anschluss an die im obigen *Nature*-Artikel zitierte Bemerkung von Michael Twidale – eher „shocking to see how many errors there are in Britannica“, findet aber einige der Punkte in der Gegendarstellung durch die EB durchaus nachvollziehbar.

¹⁶ Die englischsprachige Wikipedia bietet eine Liste von Fehlern der EB (zumindest nach Auffassung der Wikipedia). Ein großer Teil davon betrifft falsche Daten oder Fakten, seltener wird auf die Art der Darstellung selbst eingegangen (Aufbau von Artikeln, Informationsmenge etc.): http://en.wikipedia.org/wiki/Wikipedia:Errors_in_the_Encyclop%C3%A6dia_Britannica_that_have_been_corrected_in_Wikipedia [19.10.2007]

¹⁷ *New Scientist Magazine*, Ausgabe 2622 vom 20. September 2007, S. 28f.; <http://technology.newscientist.com/article/mg19526226.200-wikipedia-20--now-with-added-trust.html> [19.10.2007]

¹⁸ Vgl. Wikipedia http://de.wikipedia.org/wiki/Wikipedia:Gesichtete_Versionen [19.10.2007]

Meinung des Prüfers keine falschen Aussagen und groben Lücken enthalten.¹⁹ Bereits eingesetzte Maßnahmen zur Steigerung der Qualität umfassen ein Auszeichnungssystem („exzellenter Artikel“ und „lesenswerter Artikel“), die Einführung einer Artikelwerkstatt zur gezielten Bearbeitung unzureichender Einträge, die Durchführung von Schreibwettbewerben und das vermehrte Augenmerk, das auf die oft vernachlässigten Überblicksthemen gelegt wird. Aktuell sollen einerseits eine Spezialisierung in der Arbeitsteilung zwischen administrativ tätigen Mitarbeitern und Autoren, andererseits eine Dezentralisierung in Fachgruppen, die sich aus Mitgliedern zusammensetzt, die im betreffenden Fach besonders kompetent sind, zur Qualitätssicherung beitragen (eine Beschreibung dieser Maßnahmen in Schulenburg et al. 2007, S. 227ff.). Letztlich führt Wikipedia spätestens mit den zuletzt genannten Schritten ein Redaktionssystem²⁰ ein, was es vom ursprünglichen Gedanken der „freien“ Enzyklopädie wegführt hin zu autorisierten Versionen. Überspitzt könnte man sagen: Die Freiheit ist damit verloren, die Autorität anerkannter Experten (wie bei der EB) aber noch nicht gewonnen.

In der Überschrift zu diesem Abschnitt sind mit dem Hilfsausdruck „neu“ alle Publikationen gemeint, die das konventionelle enzyklopädische Modell insofern hinter sich lassen, als sie einen anderen Ansatz verfolgen: Informationssammlung und Aufbereitung (und in unserem Fall dann: Terminologiarbeit) wird nicht mehr von einem Zirkel eingeweihter Fachleute geleistet, indem Beiträge zu Lemmata in einem langen und aufwändigen Prozess ausgewählt, aufeinander abgestimmt und schließlich veröffentlicht werden, sondern es wird ein Medium bzw. ein Kanal zur Verfügung gestellt, mithilfe dessen – zumindest potenziell – alle Menschen, die Zugriff darauf haben, sich an der Erstellung beteiligen können. An dieser Stelle sind die seit einiger Zeit kursierenden Schlagwörter „Web 2.0“ und „Social Software“ zu nennen, womit die Tendenz gemeint ist, das Internet (technisch) so zu gestalten, dass Inhalte nicht mehr ausschließlich oder auch primär von definierten Sendern vorgegeben und von Empfängern konsumiert werden, sondern dass die

¹⁹ Vg. Wikipedia, http://de.wikipedia.org/wiki/Wikipedia:Gepr%C3%BCfte_Versionen [19.10.2007]

²⁰ Tatsächlich auch so genannt: „Viele dieser [Fach-]Gruppen, die sich heute in der Regel ‚Redaktionen‘ nennen, [...]“ (Schulenburg et al. 2007, S. 229)

Rollen von Sender und Empfänger wechseln können. Beide Schlagwörter sind unscharfe Begriffe, die nicht klar voneinander abgegrenzt sind. Tendenziell könnte man eher ‚Web 2.0 als den theoretischen Ansatz verstehen und ‚soziale Software‘ als eine Bezeichnung für konkrete Anwendungen, die dem Web 2.0-Gedanken verpflichtet sind“ (Bargmann 2007, S. 4). Klassisches Beispiel dafür ist jedenfalls die Online-Enzyklopädie Wikipedia, die sich „von anderen Enzyklopädien dadurch unterscheidet, dass [sie] einen vergleichsweise offenen Zugang bietet. Jeder Mensch mit Internetzugang kann Beiträge erstellen, ändern und erweitern.“ (Bargmann 2007, S. 32).

Die anderen zu behandelnden Publikationen sollen als „konventionell“ in dem Sinne bezeichnet werden, dass bei ihnen Experten eines Fachgebietes für Fachleute und Laien einschlägige Artikel verfassen, die von Anfang an auf hohes Niveau und umfassende Darstellung angelegt sind. Mit den Bezeichnungen „neu“ und „konventionell“ soll ausdrücklich keinerlei Wertung verbunden sein.

Angeregt durch die oben angeführte *Nature*-Studie veröffentlichte das deutsche Computermagazin *c't* in der Ausgabe 6/2007 einen ähnlichen Vergleich zwischen den Online-Enzyklopädien von Bertelsmann, Brockhaus und Microsoft Encarta einerseits und der deutschen Wikipedia andererseits (Wiegand 2007). Offenbar wurden die Analysen und Bewertungen der Artikel zwar „nur“ vom Redaktionsteam vorgenommen,²¹ doch ist die Betrachtung von 56 ausgewählten Begriffen unter den Aspekten Korrektheit, Vollständigkeit, Verständlichkeit, Aktualität und Multimedia gut nachvollziehbar. In unserem Zusammenhang interessiert hier nur die Einschätzung von Wikipedia durch *c't*. Bezüglich Korrektheit der ausgewählten Artikel erkennt *c't* keinen Unterschied zwischen Wikipedia und den Online-Produkten anderer Enzyklopädisten. Bei Vielfalt und Art der Multimedia-Angebote habe sich Wikipedia in der letzten Zeit stark verbessert. Unschlagbar ist Wikipedia in Hinblick auf die Aktualität und die Schnelligkeit der Erstellung und Anpassung von Artikeln – was allerdings auch zu einer Vielzahl über Gebühr de-

²¹ Wodurch dann solche Fehler passieren wie die Behauptung, der Ausdruck *Apoptose* („programmierter Zelltod“) leite sich nicht, wie von der Bertelsmann-Enzyklopädie und von Wikipedia – richtig! – festgestellt, „von dem griechischen Wort für Abfallen her“, sondern tatsächlich ginge er „auf *apo* = *ab*, *weg* und *stosis* = *Senkung* zurück“ (Wiegand 2007, S. 142).

tailreicher Einträge etwa zu Themen der Popkultur führt, für die sich „die Wikipedianer möglicherweise auch stärker zuständig als Brockhaus oder Bertelsmann“ fühlten (Wiegand 2007, S. 138). Dieses Phänomen ist auch auf die relativ schwache Stellung einer zwischengeschalteten redaktionellen Instanz bei der Beitrags-erstellung zurückzuführen.²² Wichtig erscheinen hier vor allem Wiegands Hinweise auf die Vollständigkeit von Wikipedia; so habe es zu fast allen Testbegriffen auch tatsächlich einen Artikel gegeben. Dies führe zu einer unglaublich Fülle von Artikeln, oft über nebensächliche, ephemere Phänomene sowie zu übermäßig langen Artikeln mit geringerer Dichte an relevanter Information. Redundanz, nebensächliche Details und fehlende innere Gliederung könnten die Suche nach Information mühsam werden lassen. Durch die reine Kumulation sprengten Wikipedia-Beiträge oft die Intentionen und Grenzen einer Enzyklopädie, weisen aber dennoch Lücken auf. Vor allem lasse das sprachliche Niveau vieler Beiträge zu wünschen übrig, wie auch die Sachlichkeit der Darstellung, in die zu oft persönliche Sichtweisen der Autoren einflössen. Wichtig auch der Hinweis, die internen Diskussionen und Unterschiede zu früheren Versionen zu beachten (vgl. zu allen genannten Punkten Wiegand 2007, S. 144f.). Die Conclusio für *c't*:

Die ideale Enzyklopädie sollte aus unserer Sicht so leicht bedienbar und so verständlich sein wie die Encarta, sich so konsequent wie der Brockhaus um sprachlich einwandfreie, neutrale Texte bemühen und so viel Wissen enthalten wie die Wikipedia. (Wiegand 2007, S. 145)

Die hier aufgeworfene Frage nach der Qualität beider Produktgruppen stellt sich mit zunehmendem Umfang und Verbreitung von neuen, elektronischen Publikationen, die konventionellen Produkten Konkurrenz machen, verschärft. Der entscheidende Faktor – auch im Hinblick auf den Fokus dieser Arbeit – liegt dabei m.E. weniger im elektronischen Charakter des Mediums selbst (auch wenn dieser die Voraussetzung für einige seiner Besonderheiten ist); denn auch die Enzyklopädien konventioneller Anbieter können und werden ja tatsächlich – auch – auf elektronischem Wege angeboten. Entscheidend ist weniger das Medium als vielmehr zwei miteinander verknüpfte Faktoren des Produktionsprozesses:

(1) die Urheberschaft von Einträgen mit der Polarität Laie vs. Experte

²² Die allerdings gestärkt werden soll; vgl. Schulenburg et al. 2007, S. 228f.

- (2) Arbeitsform Erstellung mit der Polarität einzelner, bestellter Autor vs. offenes Kollektiv mit wechselnder Autorenrolle

Ad (1): Der deutsche Kommunikationstheoretiker Norbert Bolz spricht davon, mit Wikipedia entstehe „ein weltweites Laienwissen, das in Konkurrenz zum Expertenwissen tritt“ (zitiert in Wehn/Welker 2006)²³. Zwar sind Vorteile einer Laienverfasserschaft durchaus vorstellbar; Bolz zählt einige auf: etwa Aktualität und Anwendungsbreite, sowie die oft bessere Verständlichkeit aufgrund eines geringeren Abstraktionsniveaus. Wieso allerdings die Erstellung durch Laien nachgerade *per se* besser sein soll, bleibt unklar. *Per definitionem* sind Experten Menschen, die sich auf einem Fachgebiet hervorragend auskennen. Durchaus verständlich also und wünschenswert, dass klassische Enzyklopädien immer auf solche zurückgegriffen und auch damit geworben haben.

Während Beiträge in klassischen Enzyklopädien also immer von Experten ihres Fachgebiets oder zumindest von professionellen Redakteuren verfasst sind, kann das in Wikipedia zwar auch der Fall sein, muss es aber durchaus nicht. Und vor allem: als Konsument des Beitrags weiß ich im Normalfall nicht, ob der Urheber Laie oder Experte ist. Damit ist die Verlässlichkeit als Quelle entscheidend gemindert – und im Falle wissenschaftlicher Arbeiten nicht mehr gegeben.²⁴ Zudem sind die Verfasser von Einträgen meist selbst User bzw. Konsumenten von Wikipedia, die Überschneidungen zwischen beiden Gruppen also viel größer als bei den konventionellen Publikationen. Dies führt zu einer Verzerrung bei der Themenauswahl und -behandlung, in der Interessen von eher jungen, gegenüber neuen Kommunikationstechnologien offenen Menschen über-, viele schwierige oder „trockene“ Wissensgebiete aber unterrepräsentiert sind (vgl. dazu Wehn/Welker 2006 und neuerdings Schulenburg et al. 2007, S. 227).

Ad (2): Eng damit verbunden ist die Art der Erstellung von Beiträgen: als Werk einzelner Urheber oder eines Kollektivs. Zwar können auch in klassischen Enzyklopädien mehrere Autoren einen Beitrag erstellen. Bei Wikipedia ist dies

²³ Wehn/Welker (2006, 01.September): Weisheit der Massen.
<http://www.telepolis.de/r4/artikel/23/23435/1.html> [19.10.2007]

²⁴ „For God [sic!] sake, you're in college; don't cite the encyclopedia [=Wikipedia]“ hielt angeblich auch Jimmy Wales, deren Gründer, Studenten entgegen (Wehn/Welker 2006).

jedoch der ausdrücklich erwünschte Regelfall. Letztlich geht es um die Nachvollziehbarkeit der Urheberschaft für den Konsumenten: Wer zeichnet für einen Beitrag verantwortlich? Bei Wikipedia ist die Urheberschaft über die Versionskontrolle zwar grundsätzlich, wenngleich mühsam, rekonstruierbar, stößt aber an Grenzen, wenn Autoren nicht namentlich registriert sind, sondern – was häufig vorkommt – nur mit einem Nickname oder gar nur die IP-Adresse des Rechners. Gerade bei eher kontroversiellen Themen bleibt offen, ob die Beiträge tatsächlich *sine ira et studio*, d.h. ohne (implizite) Eigeninteressen des Urhebers, verfasst wurden. So sind etwa Versuche bekannt geworden, Einträge in Wikipedia zu lancieren, und sie so in den Dienst des Marketings großer Firmen zu stellen.²⁵ Der in den Leitlinien genannte Grundsatz des „Neutralen Standpunkts“ (Voß/Danowski 2004, S. 457) findet oft genug keine Beachtung. Zwar stellt die Community, die sich rund um einzelne Themenkomplexe herausgebildet hat, ein gewisses Korrektiv dar; hier aber grundsätzlich auf die „Selbstheilungskräfte des Wiki-Prinzips“ zu vertrauen (Voß/Danowski 2004, S. 461) scheint doch eine allzu optimistische Hypostasierung: bis „ein anderer Fachkundiger den Artikel findet“ (ebd.) und ihn verbessert, kann schon einiges an Schaden angerichtet worden sein.

Die kollektive Erstellung führt auch zu einem für Wikipedia typischen Problem: die reine Kumulation von Informationen. Oft wird die Ansicht vertreten, der Anspruch einer Enzyklopädie bestehe darin, „das gesamte Wissen der Menschheit in einem System zur Verfügung zu stellen“ (so Wehn/Welker 2006)²⁶.

Voß/Danowski meinen, „die selbstgestellte Aufgabe, eine Enzyklopädie zu erstellen, misst sich [...] am Anspruch, das gesamte Wissen der Menschheit zu sammeln“ (Voß/Danowski 2004, S. 458). Auch wenn man diesen Anspruch nur als unerreichbares Leitideal interpretiert, so bleibt er doch erstaunlich hehr – und nachgerade naiv, wie ein kurzer Blick auf die Entstehungsgeschichte der Enzyklopädien zeigt: Zwar versuchten die französischen Enzyklopädisten noch, das gesamte Wissen ihrer Zeit wiederzugeben, doch musste dieser Anspruch bald aufgegeben werden zugunsten einer gut strukturierten, einführenden Behandlung der

²⁵ Vgl. etwa einen Bericht in der *Süddeutschen Zeitung* vom 26.01.2007, <http://www.sueddeutsche.de/kultur/artikel/578/99479/article.html>

²⁶ Von der „einführenden, überblicksartigen und eher allgemein gehaltenen Anlage der Texte“ spricht dann aber – im Widerspruch zum auch dort genannten Anspruch auf Vollständigkeit – der genannte Artikel. Zudem bleibt unklar, was genau hier mit „System“ gemeint ist.

Themen. Bei Wikipedia führt das Festhalten am hehren Ziel zur Unübersichtlichkeit zahlreicher Einträge.

2.5 Folgerungen

Welche Folgerungen ergeben sich aus den obigen Ausführungen für das Kernthema der vorliegenden Arbeit? Inwieweit sind sie für den Vergleich von Terminologiarbeit in „konventionellen“ und „neuen“ Medien bedeutsam? Zieht man die für vorliegende Untersuchung relevanten Einträge in Wikipedia in Betracht, so sind m.E. einige der oben dargestellten Faktoren bedeutsam. Diese Faktoren und ihre Konsequenzen werden hier in eher hypothetischer Weise formuliert (A-D).

(A) Zunächst die Art des Themas, zu dem Einträge erstellt werden. Beim Themenkomplex Thesaurus handelt es sich kaum um ein kontroversielles Thema. Einen „neutralen Standpunkt“ zu finden dürfte unproblematisch sein, emotionale Auseinandersetzungen, Vandalismus, gar *edit wars* sind eher nicht zu erwarten.²⁷ Beim Thesaurus selbst handelt es sich um ein etabliertes Werkzeug im Bereich Bibliothek, Information, Dokumentation, für das zwar neue Einsatzmöglichkeiten erschlossen werden, das aber als Werkzeug eine stabile Entwicklungsstufe erreicht hat.

(B) Mit der Art des Themas eng verbunden ist die Gruppe der potenziellen Verfasser bzw. Autoren der einschlägigen Beiträge in Wikipedia. Als solche kommen nur Menschen in Betracht, die ein gewisses Expertentum erworben haben und deren Interesse ausbildungs- oder berufsbedingt in diesem Bereich liegt. Beiträge von tatsächlichen Laien oder gar „Weisheiten der Massen“ sind nicht zu erwarten. Dies zeigt sich auch an der Anzahl der Bearbeiter laut Versionsgeschichte für das Wikipedia-Portal BID: Eine eher geringe Zahl von Autoren, die jedoch sehr aktiv sind. Der Nutzer oder Leser kann daher berechtigterweise von einer höheren Verlässlichkeit der einschlägigen Artikel im Vergleich zu vielen anderen Themengebieten ausgehen.

(C) Aus ähnlichen Gründen vergleichsweise klein dürfte auch die Gruppe der interessierten Nutzer einschlägiger Artikel sein. Damit sinkt einerseits die Ge-

²⁷ Als Beleg dafür mag die Versionsgeschichte zum Eintrag ‚Thesaurus‘ dienen. Im Jahr 2007 wurden übrigens beinahe alle Beiträge oder Änderungen von namentlich oder zumindest per Nickname genannten Autoren erstellt.

<http://de.wikipedia.org/w/index.php?title=Thesaurus&action=history> [19.10.2007]

fahr von gezieltem Missbrauch von Verfasserseite (etwa zu Werbezwecken), da die Breitenwirkung begrenzt ist. Andererseits dürfte sich auf die Qualität positiv auswirken, dass beide Gruppen themenaffin sind.

(D) Als eine Methode der Qualitätssteigerung von Wikipedia-Inhalten nennen Voß/Danowski auch die Portale, von denen inzwischen ca. 400 verfügbar sind (Oktober 2007). Sie machen viele (nicht notwendigerweise alle) Einträge zu einem Themenkomplex von einer Wikipedia-Seite aus zugänglich und dienen damit der Übersichtlichkeit und Einheitlichkeit der Darstellung. Zudem werden Mängel und Lücken im Zusammenhang deutlich (Voß/Danowski 2004, S. 461).

Unter diesen Gesichtspunkten interessant wäre ein Qualitätsvergleich zwischen Wikipedia und anderen enzyklopädischen Projekten oder Sachlexika, der nicht auf einer zufälligen Auswahl von paarigen Artikeln beruht, sondern ganze Themengebiete (wie in Portalen) einander gegenüberstellt. Ansatzweise hat das Klaus Graf mit seinem Vergleich des Historischen Lexikons Bayerns mit einschlägigen Wikipedia-Artikeln unternommen,²⁸ wo er zu dem Ergebnis kommt, dass viele Artikel in Wikipedia gleichwertig, wenn nicht gar überlegen seien.²⁹ Der wissenschaftlichen Autorität und damit Zitierbarkeit hie stehe da eine Fülle von Einzelnachweisen, eine nachvollziehbare Versionsgeschichte, Aktualität und Aktualisierbarkeit sowie die freie Lizenz gegenüber. Zwar sei die Qualität vieler Artikel noch bei weitem nicht ausreichend, doch würden immer mehr Artikel durchaus das Niveau des Sachlexikons erreichen.

Zusammenfassend lässt sich sagen, dass sich Wikipedia zu einem durchaus beachtlichen Nachschlagewerk entwickelt hat. Die Nachteile bei der Nutzung bestehen vor allem in den großen Unterschieden bei der Qualität der Artikel und in der mangelnden Nachvollziehbarkeit bei der Urheberschaft. Es bleibt abzuwarten, inwieweit diese Mängel in Zukunft behoben werden können.

²⁸ <http://www.historisches-lexikon-bayerns.de> mit (zum Zeitpunkt der Rezension Dezember 2006) über 340 Artikeln. [19.10.2007]

²⁹ <http://www.ub.uni-dortmund.de/listen/inetbib/msg32061.html> [19.10.2007]

3 Praktischer Teil

3.1 Darstellung der drei zu untersuchenden Publikationen

Im Folgenden wird ein Überblick über die drei Publikationen geboten, die in vorliegender Arbeit herangezogen werden, um Terminologiearbeit bzw. deren Ergebnisse zu untersuchen. Es handelt sich dabei um zwei Bücher und ein Internetportal.

- A) Kuhlen, Rainer / Seeger, Thomas / Strauch, Dietmar (Hrsg.) (2004): Grundlagen der praktischen Information und Dokumentation. Begründet von Klaus Laisiepen, Ernst Lutterbeck und Karl-Heinrich Meyer-Uhlenried, 5. völlig neu gefasste Auflage; München (K. G. Saur);
 Band 1: Handbuch zur Einführung in die Informationswissenschaft und -praxis
 Band 2: Glossar (148 Seiten)

Wie oben zitiert, gibt das Handbuch auf dem Einband die „Gründungs herausgeber“ an, nach denen es akronymisch bald als LaiLuMU bezeichnet und unter dieser Bezeichnung im IuD-Bereich bekannt wurde.

Zur nunmehr aktuellen 5. Auflage hat Otto Oberhauser eine Rezension verfasst (Oberhauser 2004, S. 86), in der er grundsätzlich die Verbesserungen dieser 5. Auflage gegenüber der vorangegangenen hervorhebt, da vor allem im 1. Band Beiträge aktualisiert, wenngleich inhaltlich zu wenig aufeinander abgestimmt worden seien, sodass sich darin auch „die historisch gewachsene Kluft zwischen den doch so verwandten Bereichen Bibliotheken und Dokumentation in Deutschland“ zeige (Oberhauser 2004, S. 88). Bezüglich des für vorliegende Arbeit hauptsächlich interessierenden Glossarbandes fällt seine Beurteilung noch zwiespältiger aus. Er beanstandet vor allem die Aufnahme von Einträgen von nur noch historischem Interesse (z.B. ‚Lochstreifen‘), bemängelt die „unsäglichen englischen Übersetzungen“ (ebd., S. 89) für *Dokumentationseinheit* und *Dokumentarische Bezugseinheit* (diesem Hinweis nachzugehen würde allerdings bereits in die Gefilde mehrsprachiger Terminologiearbeit führen) und moniert insgesamt eine gewisse Willkürlichkeit der Glossarbeiträge, die „wohl noch ein reichhaltiges Aktionsfeld für künftige Auflagen“ biete (ebd.). Dennoch bestätigt Oberhauser den Wert des Handbuches, um sich einen Ein- und Überblick über Informationswissenschaft und

-praxis zu verschaffen.

Angemerkt sei noch, dass die Herausgeber des Glossarbandes für die Auswahl der Lemmata das Kriterium „einer gewissen Repräsentativität“ angeben (GpID Einleitung, o.S.) – ohne allerdings konkrete Kriterien dafür zu nennen. Aufgenommen worden seien „zusätzlich Definitionen von Sachverhalten, Techniken oder Methoden, die nicht in Band 1 behandelt werden, aber von historischen [sic!] Interesse sind.“ (ebd.)

- B)** Beling, Gerd † / Port, Peter / Strohl-Goebel, Hildeburg (Redaktion)(2006): Terminologie der Dokumentation und Information. Komitee Terminologie und Sprachfragen (KTS) der Deutschen Gesellschaft für Informationswissenschaft und Informationspraxis (DGI); 2., neu bearbeitete Ausgabe, Frankfurt am Main (Deutsche Gesellschaft für Informationswissenschaft und Informationspraxis), 124 Seiten

Dieses Werk ist über den Webauftritt des Surendorf-Teams

<http://www.surendorf-team.de/> auch online als PDF-Datei verfügbar:

<http://www.surendorf-team.de/TID-II2.pdf> [19.10.2007]. Allerdings unterscheiden sich beide Versionen voneinander. Beide enthalten ein – allerdings nicht gänzlich identisches – Vorwort des mittlerweile verstorbenen Gerd Beling, die als 2. Ausgabe vorliegende Broschur-Ausgabe darüber hinaus ein Inhaltsverzeichnis und ein „Vorwort 2006“ (TID, S. 3 bzw. 5). In beiden folgt „1. Systematisches Register“ und „2. Definitionsteil“ mit einführenden Erläuterungen und „3. Alphabetisches, permutiertes Register“.

Weiterreichende Unterschiede zwischen den beiden Versionen ergeben sich erst im zweiten, dem Definitionsteil (Broschur ab Seite 35, PDF-Datei ab S. 21). In letzterer finden sich englische und französische Äquivalente zu den Benennungen, die in der Broschur-Version grundsätzlich nicht angeführt werden. Durch die Beifügung eines *Deutschen Registers* (ab S. 122), eines *English Index* (ab S. 158) und eines *Index français* (ab S. 187) wird die Online-Version auch sprachübergreifend nutzbar. Die Broschur-Version enthält

nur ein (deutsches) alphabetisches, permutiertes Register (ab S. 105). Wie bereits angemerkt, sind beide Versionen auch sonst nicht immer wort- (geschweige denn satz-)identisch. Abgesehen von den oben genannten sind die Unterschiede aber meist formaler Natur (etwa die Verwendung von Zahlwörtern oder Ziffern im Definitionsteil, Layout). Die Online-Version ist die neuere, überarbeitete Version, was sich etwa daraus schließen lässt, dass die in der Broschur-Version auftretende Inkonsistenz zwischen dem systematischen Register, 51-03-2 ‚digitale Daten‘, und dem Definitionsteil 51-03-2 ‚diskrete Daten [...]‘ in der Online-Version beseitigt wurde und in beiden Teilen einheitlich unter 51-03-2 ‚diskrete Daten‘ aufgeführt ist.³⁰

Kuriosum am Rande: Am Ende seines Vorworts merkt Beling einigermaßen süffisant an, das Komitee habe sich nicht entschließen können, „irgendeine der Rechtschreibreformen der letzten Jahre zu berücksichtigen“ (TID, S. 8), welche Anmerkung mit dem begründenden Hinweis auf die Altersstruktur des Komitees auch in der Online-Version steht. Allerdings setzt letztere die (gültige) Rechtschreibreform zumindest teilweise um, indem etwa die neuen Regeln zur s-Schreibung zur Anwendung kommen. Die Einträge im Registerteil bleiben in der alten Rechtschreibung, die Definitionen im Definitionsteil folgen teilweise wiederum der neuen („Gewissheit“, aber: „Numerierung“). Eine Inkonsistenz, die nach dem endgültigen Inkrafttreten der nunmehr recht unumstrittenen oder zumindest nicht mehr in öffentlicher Diskussion stehenden Regeln im Jahr 2006 bei nächster Gelegenheit beseitigt werden sollte.

C) Das Wikipedia-Portal Bibliothek, Information, Dokumentation

http://de.wikipedia.org/wiki/Portal:Bibliothek%2C_Information%2C_Dokumentation

Laut Selbstdarstellung auf der Übersichtsseite der auf Wikipedia vorhandenen Portale sind Portale Einstiegsseiten in die Enzyklopädie, die von den Nutzern selbst redaktionell gepflegt werden. Sie präsentieren eine Übersicht der wichtigsten Artikel zu einem Themengebiet und zeigen, welche Artikel neu ge-

³⁰ Jeder Eintrag in Register und Definitionsteil erhält eine Notation, welche die Haupt- und Untergruppe angibt, sowie eine laufende Nummer. Um die Auffindung in TID zu erleichtern, wird im Folgenden neben dem Lemma fallweise die Notation angegeben.

schrieben wurden bzw. welche sich noch im Bearbeitungsstadium befinden.³¹ Damit sind diese – auch Themenportale genannten – „einheitlichen Einstiegspunkt[e]“ tatsächlich „Benutzerschnittstelle[n]“, auch wenn sie keine „*Personalisierungsfunktion*, die letztlich das wesentliche definitorische Merkmal ausmacht“, aufweisen (alle Zitate Bertram 2005, S. 286f.), demnach also keine Portale im engeren Sinne darstellen.³² Laut Hansen/Neumann – die Portale allerdings aus dem Blickwinkel des E-Commerce betrachten – können Internet-Portale nach „Struktur, Inhalt und Präsentation“ für alle Benutzer gleich aussehen oder aber für Benutzer (adaptive Portale) bzw. von Benutzern (adaptierbare Portale) angepasst werden (Hansen/Neumann 2005, S. 642f.). Das Wikipedia-Portal – oder besser: Themenportal – wäre demnach ein vertikales (d.h. auf einen engen Fach-/Themenbereich bezogenes), nicht personalisierbares Portal.

Jedenfalls handelt es sich hierbei um eine gänzlich andere Art der Wissens-erstellung und -verteilung als bei den weiter oben vorgestellten beiden anderen Publikationen. Vor allem sind hier Fragestellungen relevant, wie sie allgemein im Zusammenhang mit öffentlich erstelltem und zugänglichem Content diskutiert werden, exemplifizierbar an vergleichenden Untersuchungen der Qualität von Beiträgen in Wikipedia und der Encyclopædia Britannica und für deutschsprachige Enzyklopädien zwischen Brockhaus, Bertelsmann, Encarta und Wikipedia (s. Abschnitt 2.4). Inwieweit sich diese Ergebnisse auch auf den Untersuchungsbereich dieser Arbeit übertragen lassen, ob also die Qualität von Artikeln zum Themenkomplex Thesaurus sich bei den beiden Publikationsmethoden unterscheidet, ist eine wesentliche Frage, auf die diese Arbeit eingeht. Das Portal Bibliothek, Information, Dokumentation wurde am 01.04.2004 von Jakob Voss erstellt, der nach wie vor zu den Hauptbeiträgern gehört. Übrigens existierten mit Stichtag 19.10.2007 vergleichbare Portale auch in einer englischen – *The Library and Information Science Portal*³³ – einer franzö-

³¹ http://de.wikipedia.org/wiki/Portal:Wikipedia_nach_Themen [19.10.2007]

³² Allerdings weist Bertram auf die von Rösch ausgemachten unterschiedlichen Verwendungen des Ausdrucks im anglo-amerikanischen und deutschen Sprachraum hin, wobei in letzterem darunter „lediglich der einheitliche Zugang zu bestimmten Ressourcen verstanden wird“ (Bertram 2005, S. 287, Fußnote 24).

³³ http://en.wikipedia.org/wiki/Portal:Library_and_information_science [19.10.2007]

sischen – *Portail des sciences de l'information et des bibliothèques*³⁴ – einer polnischen – *Bibliologia w polskiej Wikipedii*³⁵ – und einer arabischen – *تاہول عمل و تابت كمل امل ع:قب اوب*³⁶ – Version. Hier wäre es interessant festzustellen, inwieweit dies den weltweiten Aktivitätsgrad in diesem Fachbereich widerspiegelt. Zumindest wird immer wieder betont, dass der deutsche Sprachraum im Verhältnis zu seiner Sprecherzahl zu den aktivsten gehört, was den Themenbereich BID anbelangt.³⁷

3.2 Typologisierung der Publikationen und Analyseschema

Die vorliegende Arbeit untersucht die lexikografische Beschreibung einer Fachsprache – oder vielmehr: eines Teilbereichs einer Fachsprache – und gehört insofern zur Metalexikografie. Unter Fachlexikografie wird ganz allgemein die „Hervorbringung von Nachschlagewerken innerhalb von spezialisierten menschlichen Tätigkeitsbereichen“ verstanden (Roelcke 2005, S. 126). Die Bezeichnung „Nachschlagewerke“ ist dabei weit genug, um strukturell und funktional heterogene Werke von Glossar bis elektronische Enzyklopädie zu umfassen, wohingegen der Ausdruck „Fachwörterbuch“ zu sehr auf die Beschreibung sprachlicher Aspekte (etwa Angaben zu Morphologie, Phraseologie oder sprachlichem Register) hin ausgerichtet erscheint, wenngleich auch er nach der Definition: „Ein Fachwörterbuch ist eine geordnete Sammlung von Benennungen der Begriffe eines Fachgebiets“ (DIN 2333, S. 1) auf die drei untersuchten Publikationen angewendet werden könnte. Der Einteilung von Roelcke in drei verschiedene Fachwörterbuchtypen – mit sprachbezogenen, sachbezogenen oder beiderlei lexikografischen Angaben – können alle drei Publikationen dem Typus „fachliches Sachwörterbuch“ zugerechnet werden, wobei Wp eine gewisse Tendenz zum Typus „fachliches

³⁴ Französisches Portal:

http://fr.wikipedia.org/wiki/Portail:Sciences_de_l%27information_et_des_biblioth%C3%A8ques [19.10.2007]

³⁵ Polnisches Portal: <http://pl.wikipedia.org/wiki/Portal:Bibliologia> [19.10.2007]

³⁶ Arabisches Portal:

<http://ar.wikipedia.org/wiki/%D8%A8%D9%88%D8%A7%D8%A8%D8%A9:%D8%B9%D9%84%D9%85%D8%A7%D9%84%D9%85%D9%83%D8%AA%D8%A8%D8%A7%D8%AA%D9%88%D8%A7%D9%84%D9%85%D8%B9%D9%84%D9%88%D9%85%D8%A7%D8%AA> [19.10.2007]

³⁷ „From the point of view of terminology the German speaking regions are the most active in the world.“ <http://www.infoterm.info/> [19.10.2007]

Allbuch“ (Roelcke 2005, S. 126) aufweist, da hier in den einzelnen Artikeln neben die Sach- zum Teil auch Sprachangaben treten.

Folgt man der „Gliederung terminologischer Disziplinen“ nach Felber und Budin (Felber/Budin 1989, S. 296), so befasst sich vorliegende Arbeit mit dem Gegenstandsbereich einsprachige Terminologiarbeit und darin mit dem Tätigkeitsbereich „Feststellung des Sprachgebrauchs (Ist-Norm)“, was für GpID und Wp uneingeschränkt, für TID allerdings nur mit Einschränkung gilt, da hier – wie auch in den beiden Vorworten angesprochen (TID, S. 6f.) – durchaus Übergänge zu „Festlegung der Begriffssysteme, Definitionen und Benennungen (Soll-Norm)“ festzustellen sind (Felber/Budin 1989, S. 296).

Im Rückblick auf die Ausführungen über die Spezifika der drei Publikationen im Theoretischen Teil können auch Aspekte der Ziel- bzw. Benutzergruppe zu einer Typologisierung herangezogen werden (vgl. dazu Roelcke 2005, S. 129). GpID, TID und Wp wenden sich als einsprachige Publikationen an einen muttersprachlichen, zumindest aber deutschsprachigen Nutzerkreis. TID wendet sich an Experten („einsprachiges Fachwörterbuch für Experten mit präskriptiver Funktion“, ebd.), GpID an Benutzer, die Expertentum anstreben (dito, mit Einschränkungen, vor allem wenn „Laie“ und „Experte“ als Pole eines Kontinuums und nicht als Dichotomie gesehen werden), während Wp für Laien erstellt wird („einsprachiges Fachwörterbuch für Laien mit deskriptiver Funktion“, ebd.). Damit ergeben sich vor allem für Wp besondere kommunikative Anforderungen, etwa was die Deutlichkeit von Definitionen oder die Nützlichkeit von weiteren Bezügen und Beispielen betrifft.

Im Folgenden soll im Rahmen eines Analyseschemas ein Katalog von Kriterien aufgestellt werden, nach denen die einschlägigen Einträge zum Themenkomplex Thesaurus in den drei zu behandelnden Publikationen untersucht, verglichen und bewertet werden. Unter „Eintrag“ werden dabei alle Artikel, Definitionen etc. in diesen drei Publikationen verstanden. Als brauchbares Modell für die Analyse können dabei Aspekte der Strukturen von Fachwörterbüchern in analoger Weise herangezogen werden, nämlich die *Makrostruktur* als Anordnung der einzelnen Artikel, die *Mikrostruktur* als Anordnung der Angaben innerhalb der Artikel und

schließlich die *Mediostruktur* als Vernetzung der Artikel miteinander (vgl. dazu Roelcke 2005, S. 132ff.). Das Hauptaugenmerk soll allerdings auf Medio- und Mikrostruktur liegen. Zu beachten ist dabei indessen, dass die drei untersuchten Publikationen – wie oben dargestellt – heterogene Textsorten repräsentieren, so dass zu ihrer Beurteilung Kriterien, wie sie etwa Roelcke mit seiner „vereinfachte[n] idealisierte[n] Mikrostruktur von terminographischen Wörterbüchern [...] sowie fachlexikographischen Wörterbüchern“ aufstellt (Roelcke 2005, S. 136), nur bedingt bzw. in adaptierter Form herangezogen werden können. Die untersuchten Publikationen – Glossar, Register mit Definitionsteil und Internetenzyklopädie (so die jeweilige Selbstbeschreibung) – unterliegen zwar jede für sich textsortenspezifischen und kommunikativ-pragmatischen Voraussetzungen, diese sollen im Rahmen der vorliegenden Arbeit jedoch nur insofern berücksichtigt werden, als sie für die inhaltliche Kohärenz und die Funktionalität der spezifischen Produkte von Terminologearbeit von Bedeutung sind.

3.3 Auflistung der Lemmata

Nachfolgende Tabelle bietet einen Überblick über die in den *Grundlagen der praktischen Information und Dokumentation* (GpID), in der *Terminologie der Information und Dokumentation* (TID) und in *Wikipedia* (Wp) zum Themenkomplex Thesaurus aufgeführten Lemmata, wobei für die beiden erstgenannten Publikationen diejenigen wiedergegeben werden, auf die im Eintrag Thesaurus verwiesen wird bzw. – in einem zweiten Kreis – auf die von diesen Einträgen aus wieder verwiesen wird, d.h. dass Verweise vom Artikel Thesaurus aus direkt bzw. über höchstens eine Zwischenstufe zum jeweiligen Lemma führen. Zum Beispiel führt im GpID-Eintrag ‚Thesaurus‘ (Kern) ein Verweis auf den Eintrag ‚Dokumentationssprache‘ (erster Kreis), von diesem wiederum ein Verweis auf den Eintrag ‚Klassifikation‘ (zweiter Kreis). Für Wikipedia jedoch werden aufgrund der großen Zahl der Verweise (d.h. Hyperlinks) nur die im Artikel ‚Thesaurus‘ direkt mit Verweis angegebenen Lemmata angeführt, zudem nur solche, die Thesaurus im Sinne von BID betreffen (also nicht Thesaurus „im Sinne der Sprachwissenschaft (z.B. Synonymenwörterbücher u.ä.)“ (DIN 1463, S. 2). In der Auflistung bleiben

Unterschiede im Numerus der Substantiva unberücksichtigt, diese werden im Singular angegeben.

Gegenübergestellt werden dieser Auflistung die in der DIN 1463 – Teil 1 *Erstellung und Weiterentwicklung von Thesauri* verwendeten einschlägigen Begriffe, wobei die dort im 2. Abschnitt explizit definierten Begriffe mit ‚+‘ gekennzeichnet sind (elf an der Zahl), wesentliche in diesen Definitionen zusätzlich herangezogene einschlägige Begriffe mit ‚~‘.

Lemma	GpID	TID	Wp	DIN	Anmerkungen
Abkürzung	-	-	+		
Abstraktionsrelation	~	-	-	+	
Äquivalenzklasse	~	-	~	~	Wp: als Terminus der mathemat. Mengentheorie
Äquivalenzrelation	+	~	+	+	TID: unter ‚R‘
Assoziationsrelation	+	~	+	+	TID: unter ‚R‘ und ‚A‘
Attribut	-	-	+		TID: ‚attributive Relation‘
Begriff	+	+	+	~	
Begriffliche Kontrolle	+	-	-		
Benennung	~	+	+		GpID: ‚s. Bezeichnung‘
Bestandsrelation	~	-	-	+	
Bezeichnung	+	+	~	~	
Datenbank	+	~	~		Wp: ‚Datenbanksystem‘
Deskriptor	~	+	+	+	DIN auch: ‚Vorzugsbenennung‘, ‚Vorzugsbezeichnung‘
Dezimalklassifikation (DS)	+	~	~		
Dokumentationssprache	+	+	~	+	
Dokumentationssystem	-	+	~		
Dokumentationswissenschaft	-	-	+		
Gedächtnis	+	-	~		
Gesprochene Sprache	-	+	~		
Griechisch	-	-	+		Wp: Link zu ‚Altgriechische Sprache‘
Hierarchie	+	~	~		
Hierarchische Relation	+	~	-	+	TID: unter ‚R‘ und ‚H‘; Wp: Link zu ‚Hierarchie‘
Homonym	+	-	~	~	TID: nur ‚Homonymkontrolle‘
Indexierung	+	+	+	~	DIN auch: ‚Indexieren‘, ‚Inhaltskennzeichnung‘
Information-Retrieval-Modell	+	-	-		
Inhaltserschließung	+	~	~		TID: ‚Erschließung, inhaltliche‘ und ‚inhaltliche Erschließung‘; Wp: Link zu ‚Sacherschließung‘
Klassifikation	+	~	~	~	

Kommunikation	~	+	~		
Kontrolliertes Vokabular	-	-	+		TID: ‚kontrollierte Schlagwortliste‘
Künstliche Sprache	-	+	~		Wp: ‚konstruierte Sprache‘
Latinisiert	-	-	+		Wp: Link zu ‚Latein‘
Luhn, Peter	-	-	+		Link bei Wp blind
Medical Subject Headings	~	-	+		
Merkmal	-	~	~	~	Wp: vorhanden, aber nicht über BID-Portal verlinkt
Natürliche Sprache	-	+	~		
Nicht-Deskriptor	-	~	~	+	Wp: Link auf ‚Verweisungsform‘, nicht über BID-Portal verlinkt
Notation	+	~	~	~	
Oberbegriff	-	-	+	~	TID: ‚Überordnung‘
Polyhierarchisch (Polyhierarchie)	~	~	+		Wp: Link auf ‚Polyhierarchie‘
Polylem	+	-	~	~	TID: verwendet in Def. ‚Homonymkontrolle‘
Recherche	~	~	+		GpID: ‚s. Retrieval‘
Relation	~	+	~	+	DIN auch: ‚Beziehung‘
Retrieval	+	+	+	~	Wp: ‚information retrieval‘
Sacherschließung	-	-	+		
Schlagwort	+	~	~	~	Wp: ‚=Inhalterschließung‘, DIN: ‚Schlagwortsysteme‘
Semantisches Netz	~	-	+		
Speichern	~	-	~	~	GpID: ‚Speicherung‘
Stichwort	+	~	~		
Synonym	~	~	+	~	GpDI: ‚s. Äquivalenzrelation‘, TID: ‚Synonymie‘
Teilbegriff	-	-	-	~	
Terminologische Kontrolle	+	+	~	~	Wp: weitergeleitet von ‚kontrolliertes Vokabular‘
Thesaurus	+	+	+	+	
Thesauruswort	-	+	-		
Unterbegriff	-	~	+	~	TID: ‚Unterordnung‘
Verbandsbegriff	-	-	-	~	
Vokabular einer Dokumentationssprache	-	+	-	+	
Vokabularkontrolle	~	+	-		
Wort	-	+	~		

Legende:

- + Lemma scheint auf als direktes oder indirektes (höchstens 2. Kreis) Referenzlemma vom Eintrag ‚Thesaurus‘ aus. In der Spalte ‚DIN‘: in der DIN 1463 definierte Begriffe.
- ~ Lemma scheint auf, jedoch nicht mit direktem oder indirektem Verweis vom Eintrag ‚Thesaurus‘ aus. In der Spalte ‚DIN‘: in den Definitionen verwendete Begriffe.
- Lemma scheint überhaupt nicht auf.

In der Spalte ‚DIN‘ bleiben Zellen für diejenigen Lemmata (Begriffe) leer, die in dieser DIN nicht behandelt werden, wohl aber möglicherweise in anderen DIN-Festlegungen (etwa ‚Abkürzung‘).

Die 15 fettgedruckten Bezeichnungen begegnen sowohl in allen drei Publikationen als auch in der DIN 1463.

Wp weist auf die Polysemie des Ausdrucks „Thesaurus“ hin, jedoch nicht – wie in ähnlichen Fällen häufig – mittels einer eigenen Seite zur ‚Begriffsklärung‘, sondern im ersten Absatz des Artikels.

3.4 Makrostruktur

Hinsichtlich der Makrostruktur der untersuchten Publikationen sollen hier einige Beobachtungen genügen, die indes schon ausreichend erscheinen, um daraus Schlüsse etwa auf den intendierten Benutzerkreis ziehen zu können.

GpID: Die Anordnung der Lemmata erfolgt alphabetisch. Zusammengesetzte Lemmata (meist aus attributivem Adjektiv und Substantiv) werden gemäß ihrer natürlichen Wortfolge behandelt, wobei Ausnahmen auftreten (s.u. Abschnitt 3.5.1). Dieser rein der sprachlichen Ausdrucksseite (Signifikant) folgende Ansatz empfiehlt sich bei einem Werk, das den Gebrauch grundlegender Termini deskriptiv angibt,³⁸ „Grundlagen“ liefern will und dabei die Bedürfnisse einer breitgefächerten Zielgruppe berücksichtigen muss, zu denen „Wissenschaftler, Praktiker, Dozenten und Studierende“ gehören³⁹.

TID: Die Ausrichtung auf Fachleute, verbunden mit einer Mischung aus deskriptiven, aber durchaus auch präskriptiven Elementen und die Absicht, den Fachbereich hinsichtlich seiner Terminologie einigermaßen erschöpfend zu behandeln (vgl. die beiden Vorworte, TID, S. 5ff.), bedingt eine gänzlich andere makrostrukturelle Lösung, die den Fachbereich semantisch strukturiert, indem sie ihn thematisch in acht Hauptgruppen und weitere Untergruppen aufteilt. Es wird also von der begrifflichen Inhaltsseite (Signifikat) ausgegangen, um so zu einem systemati-

³⁸ Vgl. etwa die Hinweise zum „heutigen Sprachgebrauch“ im Eintrag ‚Retrieval‘ (GpID, S. 107).

³⁹ So die Beschreibung der Nutzergruppe im Verlagsprospekt; <http://www.saur.de/download/prospekte/deutsch/0000011342.pdf> [19.10.2007]

schen Register zu gelangen, dem der Definitionsteil in seinem Aufbau parallel folgt. Den Zugang über die Ausdrucksseite ermöglicht das beigegebene alphabetische Register.

Wp: Das elektronische Medium ermöglicht hier eine Verbindung beider Ansätze, indem im Portal Bibliothek, Information, Dokumentation einerseits ein alphabetisches Register zur Verfügung gestellt wird, dort aber auch ein inhaltlicher Zugang über Kategorien und Unterkategorien möglich ist. So listet etwa die Kategorie ‚Thesaurus‘ derzeit neun Einträge auf.⁴⁰ Die einzelnen Artikel können vom Nutzer natürlich auch ohne Umweg über das Portal aufgerufen werden. Die Enzyklopädie ermöglicht somit unterschiedliche Arten des Zugangs.

3.5 Mikrostruktur

Hinsichtlich des Aufbaus einzelner Einträge in Fachwörterbüchern setzt Roelcke (Roelcke 2005, S. 134) sechs allgemeine Klassen von Angaben an, die zur Untersuchung der Einträge in den drei Publikationen in analoger Weise angewendet werden können. Von diesen sind vor allem die „Lemmazeichengestaltangabe“, also die Art der Anführung des jeweiligen Stichworts, und die „Bedeutungsangabe“, also die präskriptive oder deskriptive Definition, für unsere Zwecke relevant, eingeschränkt auch die „Formangabe“, also grammatische Informationen zum Lemma. Diese Elemente der Mikrostruktur – Angaben beim Lemma und Art der terminografischen Bedeutungsbeschreibung – werden im Folgenden genauer betrachtet.

Nochmals sei festgehalten, dass es sich bei den vorliegenden Publikationen nicht um Fachwörterbücher handelt, es in dieser Arbeit also um die Prüfung von Einträgen in einem fachsprachlichen *Glossar* (GpID) bzw. in einem *Register* mit *Definitionsteil* (TID) bzw. von *Artikeln* eines *Portals* (Wp) geht. Dennoch scheinen Roelckes Kategorien zur Analyse auch dieser Textsorten gut geeignet.

⁴⁰ <http://de.wikipedia.org/wiki/Kategorie:Thesaurus> [19.10.2007]

3.5.1 Lemma

In der Lexikografie bedeutet Lemma „Eintrag, Stichwort in einem Lexikon, Repräsentant eines Lexems (als Explikandum), zu dem Informationen gegeben werden“ (Glück 2005, S. 376). Obzwar damit streng genommen offen bleibt, ob unter Lemma die Grundform des Lexems oder der (ganze?) Eintrag in einem Wörterbuch zu verstehen ist, soll im Rahmen dieser Arbeit ersteres gemeint sein.⁴¹ Im Sinne der von Glück angegebenen Synonyme „Stichwort, Wörterbuchwort“.

DIN 1463 gibt Hinweise für die Ansetzung der in einen Thesaurus einzutragenden Wörter (also des Thesaurusvokabulars). Demnach ist die am meisten verbreitete Schreibweise zu wählen (natürlich sofern sie orthografisch korrekt ist). Bevorzugt sind Substantive zu verwenden, nur für Tätigkeiten oder zur Vermeidung von Mehrdeutigkeit (etwa zwischen Vorgang und Ergebnis des Vorgangs bei vielen Substantivierungen mit dem Suffix *-ung*) können Verben herangezogen werden. Adjektive sind nicht isoliert, sondern attributiv zu Substantiven anzugeben. Das Lemma ist im Nominativ Singular anzusetzen (der Plural nur bei Pluraliatantum oder semantischer Differenzierung zwischen Singular und Plural) (DIN 1463, Teil 1, 1987, S. 3f.).

Analog ist bei Einträgen in Glossaren oder Registern zu verfahren (DIN 2333, S. 3f.).

GpID:

Grundsätzlich folgt die Ansetzung der Lemmata im Glossar den oben angeführten Regeln, doch zeigt bereits ein cursorischer Überblick über die Lemmata unter Buchstabe A, dass nicht gänzlich konsequent verfahren wird.⁴² So existieren etwa die pluralischen Lemmata ‚Abstands-Operatoren‘ (mit Verweis auf den Eintrag ‚Kontext-Operatoren‘), ‚Anhängeszahlen‘ (mit Verweis auf ‚Dezimalklassifikation‘) und ‚Avatare‘. Letzteres besonders irritierend, da der Ausdruck, wenn er ei-

⁴¹ Der Ausdruck „Eintrag“ ist ambig (Stichwort oder ganzer Artikel?), zudem muss nicht jedes Wörterbuchstichwort ein Lemma im Sinne von Grundform eines Wortes sein, da viele Wörterbücher auch abgeleitete Formen als Verweislemmata enthalten.

⁴² Hier sei für einen Augenblick der Themenkomplex Thesaurus verlassen und der Buchstabe A mit allen seinen Lemmata ins Auge gefasst.

nem Muttersprachler des Deutschen unbekannt ist, nicht unbedingt als Pluralform erkannt wird, zumal als Zusatzinformation „*engl.: avatar*“ (also Singular) angegeben wird (GpID, S. 7). Allerdings kann aus dem Plural des in der Definition angeführten *genus proximum* „Kunstfiguren“ auf den Numerus des Lemmas geschlossen werden.

Ähnlich uneinheitlich das Bild bei Lemmata, die aus Adjektiv und Substantiv zusammengesetzt sind. Analog wiederum zur Regel für Thesauri, welche die Ansetzung von zusammengesetzten Deskriptoren in ihrer natürlichen Wortfolge vorschreibt (DIN 1463, S. 2), scheinen Einträge unter den Lemmata ‚Hierarchische Relation‘, ‚Natürliche Sprache‘ und ‚Terminologische Kontrolle‘ auf. Daneben gibt es vereinzelt, doch offenbar recht willkürlich, Ansetzungen von invertierten Lemmata, die auf nicht invertierte Referenzlemmata verweisen (z.B. ‚Anwendungssystem, betriebliches‘ → ‚Betriebliches Anwendungssystem‘, GpID, S. 3). Aber auch nicht-invertierte Lemmata, die auf invertierte Referenzlemmata verweisen, kommen vor (z.B. ‚Arithmetische Operation‘ → ‚Operation, arithmetische‘, GpID, S. 5). Insgesamt also eine wenig konsequente Vorgehensweise, die auf ein Fehlen von Leitlinien während des Arbeitsprozesses schließen lässt. Zumindest existiert kein Einleitungsteil, in dem diese Vorgehensweise erklärt würde, sowenig wie es Hinweise zum Gebrauch des Glossars gibt (die ja auch bei einheitlicherem Vorgehen eher entbehrlich wären).

TID:

Hier stellt sich die Situation etwas anders dar, aufgrund des Aufbaus aus den drei Teilen *Systematisches Register*, *Definitionsteil* und *Alphabetisches, permutiertes Register*. Die Registeringänge sind durchwegs im Singular gehalten, die Überschriften zu den acht Hauptgruppen und die Zwischenüberschriften zu den Untergruppen ganz überwiegend im Plural. In gleicher Weise wird im Definitionsteil verfahren, wo der definierte Terminus in einem Satz eingebettet ist, gemäß dessen syntaktischen Anforderungen der Numerus anderer, hier als *definiens* verwendeter Termini wechseln kann. Beispiel: „Ein Thesaurushauptteil ist die Darstellung aller Deskriptoren (43-17) und Nichtdeskriptoren (43-18) mit ihren zugehörigen Rela-

tionen (41-01) einschließlich zusätzlicher Angaben.“ (TID, S. 65)

Sehr konsequent werden im dritten Teil die Permutationen durchgeführt, sodass zusammengesetzte Lemmata sowohl in nicht-invertierter als auch in invertierter Form als Registereingänge aufscheinen (‚hierarchische Relation‘ und ‚Relation, hierarchische‘), genauso wie nicht-permutierte und permutierte Formen (‚Äquivalenzrelation‘ und ‚Relation, Äquivalenz=‘).

Wp:

Die deutschsprachige Wikipedia bietet im Autorenportal auf einer Seite unter dem etwas irreführenden Titel *Namenskonventionen* Hinweise zur Gestaltung der Lemmata.⁴³ Die darin angeführte „Singularregel“ besagt, dass „Artikelnamen“ (auf der selben Seite auch als „Artikeltitel“ bezeichnet) wenn möglich im Singular anzugeben sind. Als Ausnahmen werden vor allem angeführt: Pluraliatantum, Volksgruppen, Kategorien der Biologie und zusammenhängende Ereignisse der Geschichte. Für die Lemmata des Portals *Bibliothek, Information, Dokumentation* wird diese Grundregel größtenteils eingehalten,⁴⁴ wie bei Wikipedia üblich und auch erwünscht wird oft von der Pluralform auf das singularische Lemma verlinkt (‚re-direct‘), etwa von ‚Thesauri‘ auf ‚Thesaurus‘.

Keinen Hinweis gibt es auf der genannten Seite dazu, wie mit mehrteiligen Lemmata zu verfahren ist. Allerdings gibt etwa das Register im Portal *BID* diese in ihrer natürlichen (nicht invertierten, nicht permutierten) Form, zum Beispiel ‚Kontrolliertes Vokabular‘.

3.5.2 Kopf

Über das bloße Lemma hinaus soll hier auch kurz der Kopf von Einträgen in den drei Publikationen betrachtet werden, womit in der Lexikografie der Teil vor dem Beginn des Eintrags im engeren Sinne gemeint ist. Neben dem Lemma kann der Kopf verschiedene andere Informationen enthalten. Welche Art von Informatio-

⁴³ <http://de.wikipedia.org/wiki/Wikipedia:Namenskonventionen> [19.10.2007]

⁴⁴ Ausnahme etwa ‚Bibliothekarische Regelwerke‘, <http://de.wikipedia.org/wiki/Portal:PID/Register#B> [19.10.2007]

nen gegeben wird, hängt primär vom antizipierten Nutzerkreis ab. Im Falle von GpID handelt es sich dabei neben Experten auch um Fachleute verwandter Bereiche und um Studierende (was auch im Titelwort „Grundlagen“ zum Ausdruck kommt), im Falle von TID um den engeren Kreis der „Fachwelt [von Information und Dokumentation]“ (TID, S. 5). Genauer könnte man jeweils sagen: für die deutschsprachige (Fach-)Welt, woraus sich die Weglassung aller grammatikalischen Angaben (Genus, Pluralbildung) erklärt.

GpID

GpID bietet unterhalb des Lemmas das englische Äquivalent, soweit vorhanden, es sei denn es handelt sich um ein Verweislemma oder das Lemma ist selbst bereits englisch.⁴⁵ Als Lemma angeführte englische Akronyme werden auf diese Weise aufgelöst, deutsche Akronyme dagegen meist erst im Artikel selbst. Für als Lemma aufgeführte aus dem Englischen übernommene Termini werden im Deutschen übliche fachsprachliche Äquivalente erst im Artikel angeführt (z.B. „Retrieval (auch Recherche oder Information Retrieval genannt) [...]“ (GpID, S. 107). Um das Auffinden der entsprechenden Artikel auch dann zu ermöglichen, wenn dem Nutzer nur der englische Ausdruck, nicht aber das deutsche Äquivalent bekannt ist, verfügt der Band am Ende über ein *Englisches Register zum Glossar*. Zu jedem englischen Terminus werden hier das deutsche Äquivalent und (gleichzeitig) das Referenzlemma angegeben. Aus diesem Verfahren lässt sich die zunehmend beherrschende Stellung des Englischen in der Theorie von BID ablesen. Bei polysemen Lemmata werden die verschiedenen Bedeutungen im Artikel angeführt (vgl. etwa den Artikel ‚Hierarchie‘, GpID, S. 49).

TID:

Das systematische Register bietet neben einer Nummerierung in Haupt- und Untergruppen ausschließlich das bloße Lemma, nur in seltenen Fällen einen numerischen Index zur Unterscheidung von Polysemen. Gleiches gilt für das alphabeti-

⁴⁵ Allerdings z.B. „[Lemma:] **Topic maps** [Plural!] / *engl.: topic maps*“, GpID S. 121

sche, permutierte Register, z.B. „Abbildung (1)“, „Abbildung (2)“ (TID, S. 105). Im Definitionsteil wird grammatikalische Information nur indirekt gegeben, sofern das Lemma als *definiendum* syntagmatisch in eine Definition eingebunden ist. Aus diesen Umständen kann auf einen engeren Benutzerkreis von Experten geschlossen werden, wie dies auch im Vorwort angeführt ist.

Wp:

Interessant – gerade auch vor dem Hintergrund der Ausführungen zur Frage von Erstellung und Weitergabe von Wissen durch Experten vs. durch Laien, wie sie im Theoretischen Teil der Arbeit dargelegt werden (vgl. Abschnitt 2.4) – ist die Gestaltung des Kopfes (bzw. die Angabe von zusätzlicher Information zum Lemma) bei Wikipedia. Auf der Seite ‚Was Wikipedia nicht ist‘⁴⁶ wird zwar dezidiert festgehalten:

*Wikipedia ist **kein Wörterbuch** (im Sinne von Sprachwörterbuch). In der Wikipedia sollen in erster Linie Begriffe erläutert und keine gängigen deutschen Wörter erklärt werden, wie dies ein Wörterbuch tut.* (Hervorhebungen im Original)

Tatsächlich bietet etwa der Artikel ‚Thesaurus‘ keine Angaben, wie sie etwa in einem Sprachwörterbuch zu erwarten wären. Allerdings wird in der unmittelbar unterhalb des Lemmas angeführten Kurzdefinition auf geschickte Weise der Plural dargeboten, wie überhaupt oft im Artikeltext die Pluralform expliziert wird (meist als Angabe in Klammern) oder auch Hinweise zur Etymologie enthalten sind (wie im genannten Artikel auf die ursprüngliche Bedeutung des griechischen Wortes). Dies geht bereits über das oben formulierte Grundprinzip, kein Wörterbuch zu sein, hinaus, zumal als ‚Schwesterprojekt‘ zu Wikipedia ein eigenes ebenso freies und kooperatives Wörterbuch existiert, nämlich *Wiktionary* (auch Wikiwörterbuch), das jedoch im Vergleich zur deutschsprachigen Wikipedia klein ist und derzeit (Oktober 2007) weniger als 60 000 Einträge umfasst. Am Ende des Wp-Eintrags ‚Thesaurus‘ gibt es denn einen Link zum gleichen Eintrag in Wiktionary, wo sich Angaben zur Grammatik finden – andererseits bei den Bedeu-

⁴⁶ http://de.wikipedia.org/wiki/Wikipedia:Was_Wikipedia_nicht_ist [19.10.2007]

tungsangaben gerade die BID-spezifische Bedeutung fehlt. Im Gegensatz dazu – und auch entgegen dem oben genannten Grundsatz, kein Wörterbuch zu sein und somit auch keine sprachlichen Angaben zu liefern – heißt es in Wp selbst etwa unter ‚Deskriptor‘:

*Als **Deskriptor** (Plural Deskriptoren) oder **Schlagwort** (Plural Schlagwörter) bezeichnet man meist vorgegebene Bezeichnungen, die zur inhaltlichen Beschreibung eines Objektes ausgewählt werden können.*⁴⁷

In Wiktionary dagegen existiert zu ‚Deskriptor‘ kein Eintrag.

3.5.3 Einträge

In diesem Abschnitt soll anhand einiger Kernbegriffe zum Themenkomplex Thesaurus die Frage beleuchtet werden, wie widersprüchlich oder konsistent Begriffsklärungen bzw. Definitionen ausfallen. Dazu wird aus den in obiger Liste angeführten 15 Lemmata, die sowohl in den drei Publikationen als auch in DIN behandelt werden (siehe Abschnitt 3.3), ein Drittel ausgewählt und näher betrachtet. Die Auswahl erfolgt unter dem Gesichtspunkt ihrer zentralen Bedeutung für den Themenkomplex Thesaurus und ihrer Repräsentativität hinsichtlich der Behandlung dieser Lemmata in den jeweiligen Publikationen. Ausgegangen wird dabei von DIN 2330 *Begriffe und Benennungen*, wo auch Grundsätzliches zum Begriff Definition niedergelegt ist. Die Einträge zu den behandelten Lemmata sind im Anhang nachzulesen.

Deskriptor

Bei allen vier Definitionen dieses für den Themenkomplex Thesaurus zentralen Begriffs handelt es sich um aristotelische Definitionen, welche der Formel des Boethius folgen: *Definitio fit per genus proximum et differentiam specificam*. Also um Inhaltsdefinitionen in ihrer „am weitesten verbreitete[n] Form“, die darin besteht, „den bekannten Oberbegriff und die einschränkenden Merkmale anzugeben, die den zu definierenden Begriff kennzeichnen und ihn von den anderen Begriffen derselben Reihe unterscheiden“ (DIN 2330, S. 8). Oder, wie TID in den Vorbe-

⁴⁷ <http://de.wikipedia.org/wiki/Deskriptor> [19.10.2007]

merkungen zum Definitionsteil ausführt, es solle „jede [verwendete] Definition einen vollständigen Satz, möglichst nach dem Muster ‚Ein a ist ein b mit dem Merkmal c‘ bilden“ (TID, S. 35).

Als *genus proximum* (Oberbegriff) fungiert dabei zweimal „Bezeichnung“ (Wp und DIN), einmal „Thesauruswort“ (TID) und einmal das allgemeine, allerdings attributiv spezifizierte „Wort [innerhalb eines Thesaurus]“; das Lemma ‚Thesauruswort‘ kommt bei GpID nicht vor. In der Begriffsleiter ist letzteres jedoch als der unmittelbar übergeordnete Begriff zu Deskriptor anzusehen, wohingegen mit dem Begriff Bezeichnung auf eine höhere Stufe der Begriffsleiter gesprungen wird. Insofern ist obige Bemerkung dahingehend zu präzisieren, dass es sich zwar um übergeordnete Begriffe, nicht aber um das *genus proximum* handelt.

Entgegen dem ersten Anschein sind alle Definitionen kurz, sie umfassen nur einen Satz, was auch auf GpID zutrifft, wo der erste Satz den Begriff Deskriptor definiert, der folgende, längere, den Begriff Nicht-Deskriptor. Nur der vergleichsweise lange Text von Wp grenzt Deskriptoren (auch: Schlagwörter) von Stichwörtern ab und wendet sich auch anderen, verwandten Begriffen zu (Indexierung, inhaltliche Erschließung). Nicht sehr glücklich verfährt DIN, da hier dem Oberbegriff „Bezeichnung“ in Klammern ein Synonym, jedoch nicht dazu, sondern zum Lemma, beigelegt ist, nämlich „Vorzugsbezeichnung“. Ein weiteres gebräuchliches deutsches Äquivalent, nämlich „Vorzugsbenennung“ (vgl. Bertram 2005, S. 210), fehlt.

Als *differentia specifica* begegnen Varianten der DIN-Formulierung „zur Inhaltskennzeichnung zugelassen“: „...für die Indexierung zugelassen“ (GpID), „...zur Indexierung und zum Retrieval verwandt“ (TID), „...zur inhaltlichen Beschreibung eines Objekts ausgewählt werden können“ (Wp).

Bei allen Unterschieden in der Formulierung liegen in diesem Fall jedoch in der Substanz weitgehend übereinstimmende Definitionen vor.

Assoziationsrelation

Naheliegenderweise setzen sowohl GpID als auch TID als Oberbegriff Relation an (TID einengend: symmetrische Relation) und folgen darin dem Beispiel der

DIN, darüber hinaus auch darin, dass die artspezifischen Merkmale im *definiens* primär negativ angegeben werden, nämlich als Relation, „die weder eindeutig hierarchischer Natur ist, noch als äquivalent angesehen werden kann“. (DIN 1463, wörtlich übernommen in GpID). Es ist m.E. ungünstig, dass GpID von den in DIN angeführten Beispielen einzig „DIESELMOTOR – OTTOMOTOR“ übernimmt, also Begriffe, die auf *einer* Begriffsreihe liegen. Immerhin wird in beiden Werken – über die negative Bestimmung hinaus – eine (explizit als offen bezeichnete) exemplarische Definition geboten, indem verschiedene Arten von Assoziationsrelationen aufgezählt werden. Zudem verweist GpID auf den Nutzen dieser Relation, nämlich jenseits der hierarchischen Struktur „freie“ Querbeziehungen und damit zusätzliche Deskriptoren zu ermöglichen. Am wenigsten konkret bleibt hier TID, das andererseits auf den Unterschied zwischen Assoziationsrelation und assoziativer Relation hinweist, wobei letztere „Assoziation“ im Sinne der Psychologie verwendet, erstere nur bedeutet, dass andere als die in Thesauri sonst anwendbaren (hierarchischen) Relationen vorliegen. Nach dem Gebrauch von TID wäre wohl assoziative Relation als Oberbegriff zu Assoziationsrelation anzusehen.

Hierdurch wird vielleicht der Umstand erklärlich, dass Wp im Artikel zu Thesaurus zwar einen Link ‚Assoziationsrelation‘ aufweist, dieser jedoch auf den Artikel ‚Assoziation (Psychologie)‘ führt, wo der Begriff im Zusammenhang mit Thesauri nicht mehr dargelegt wird (s. Anhang). Hier klafft bei Wp also eine Lücke.

Retrieval

Zunächst fällt bei diesem Begriff die sehr unterschiedliche Länge der Einträge ins Auge. Die mit Thesauri befasste DIN 1463 definiert den Begriff nicht explizit, sondern nur im Rahmen der Definition von Dokumentationssprache als „gezielte Wiederauffindung [von Dokumenten]“ (DIN 1463, S. 2). Ähnlich lakonisch bleibt TID: „Retrieval ist der Vorgang, gespeicherte Daten in Auswahl wieder aufzufinden“ (55-07, S. 76). Diese weitere Definition – bezogen auf Daten im Allgemeinen statt eingeschränkt auf Dokumente – erklärt sich durch die bei TID vorgenommene Differenzierung in die Retrievalarten: Dokumentretrieval (bei Dokumenten), Nachweisretrieval (bei Dokumentationseinheiten) und Datenretrieval

(bei Daten; Synonym: Faktenretrieval), eine Unterscheidung, wie sie weder in GpID noch in Wp getroffen wird. Ferner fällt auf, dass TID Retrieval nicht mit Recherche⁴⁸ gleichsetzt, was aufgrund der oft leidvollen Erfahrung der Diskrepanz zwischen *Suche in* und *Auffinden von* Datenbeständen durchaus nachvollziehbar erscheint. GpID setzt (Information) Retrieval synonym mit Recherche.⁴⁹ Wertvoll – weil für eine Zielgruppe, die an einführenden Grundlagen der praktischen Information und Dokumentation interessiert ist (also noch nicht notwendigerweise Expertenstatus erreicht hat) – erscheint hier der Hinweis auf die heute oftmals erfolgende Gleichsetzung von Recherche mit Online-Retrieval (GpID).

Wp bietet zu ‚Retrieval‘ eine Begriffsklärung, von der ein Link auf ‚Information Retrieval‘ führt. Der Ausdruck wird dort im Sinne des von GpID konstatierten heute häufigen Sprachgebrauchs mit „computergestütztem inhaltsorientiertem Suchen“ wiedergegeben. Insgesamt ergibt sich daher folgendes Bild: TID definiert (darin DIN folgend) allgemein; GpID ebenso, erwähnt aber explizit den speziellen Gebrauch, den allein Wp heranzieht. Etwas ärgerlich übrigens – und im Widerspruch zu den Richtlinien zum Aufbau eines Artikels, die verlangen: „Der erste Satz kennzeichnet den Gegenstand des Artikels (Lemma) treffend“⁵⁰, womit, wie auch die angeführten Beispiele zeigen, die Definition gemeint ist – ärgerlich also, dass nach dieser einleitenden Definition ein Abschnitt folgt, der mit „Definition“ überschrieben ist, tatsächlich aber die in den Richtlinien geforderte „kurze Einleitung mit einer Zusammenfassung der wichtigsten Aspekte des Artikelinhalts“ (ebd., Wikipedia Richtlinien) darstellt. Hier liegt somit ein Beispiel dafür vor, wie in der freien Enzyklopädie die Strukturierung von Einträgen schwieriger einheitlich zu regeln bzw. die Einhaltung kaum gänzlich zu kontrollieren ist, eine der Folgen des Prinzips der Kollaborativität (vgl. dazu im Theoretischen Teil der vorliegenden Arbeit Abschnitt 2.4).

⁴⁸ „Eine Recherche ist die methodische Suche in Datenbeständen.“ (TID 55-03, S. 76)

⁴⁹ Inkonsistent hier Bertram, die dem Registerzugang ‚Recherche‘ in Klammern ‚Retrieval‘ nachstellt und beide somit als Synonyme ausweist, gleichzeitig jedoch beim Registerzugang ‚Information Retrieval (Retrieval) [...] s.a. Recherche‘ auf den „assozierten Registerzugang“ Recherche verweist. Vgl. Bertram 2005, S. 303 [Hinweise zum Gebrauch des Registers]

⁵⁰ http://de.wikipedia.org/wiki/Wikipedia:Wie_schreibe_ich_gute_Artikel [19.10.2007]

Dokumentationssprache

Bei diesem Lemma liegt der Fall einer weitgehend übereinstimmenden Definition in allen drei Publikationen vor. Dokumentationssprache wird jeweils definiert als eine künstliche Sprache⁵¹ und dann mittels Angabe ihrer Funktion näher bestimmt. Dagegen setzt DIN den Oberbegriff weiter an („eine Menge sprachlicher Ausdrücke“), das Verhältnis zu künstlichen Sprachen bleibt zunächst unklar. In den direkt anschließenden Ausführungen zu ‚Thesaurus‘ heißt es dann allerdings: „Begriffe und ihre Bezeichnungen werden eindeutig aufeinander bezogen („terminologische Kontrolle““ (DIN 1463, S. 2). Zusammen mit den Ausführungen in DIN 2330, dass

künstliche Sprachen [...] auf Widerspruchsfreiheit und Eindeutigkeit hin konstruierte Zeichensysteme sind. Dazu gehören die formalen (formalisierten) Sprachen, z.B. der Logik oder Mathematik, die aber der Vermittlung durch eine natürliche Sprache bedürfen (DIN 2330, S. 3),

ergibt sich, dass Dokumentationssprachen künstliche Sprachen darstellen. Interessanterweise erscheint hier gerade die DIN als problematisch und nicht widerspruchsfrei, wenn man nämlich die Aufzählung der Subsysteme natürlicher Sprachen betrachtet, wozu auch Fachsprachen gezählt werden, welche „versuchen, die Mehrdeutigkeit der natürlichen Sprache durch genaue Bedeutungsfestlegung, durch Bilden spezifischer Bezeichnungen u.ä. einzuschränken“ (DIN 2330, S. 3). Wenn Widerspruchsfreiheit und Eindeutigkeit Merkmale künstlicher Sprachen sind, dann wären damit auch Fachsprachen künstliche Sprachen.⁵² Auf die Unterschiede zwischen natürlich-sprachlich basierten und nicht-natürlich-sprachlich basierten Dokumentationssprachen weist der Wp-Artikel hin, mit indes sachlich höchst ungenauen Erläuterungen.⁵³ Ungeklärt bleibt bei Wp auch das Verhältnis zwischen ‚Dokumentationssprache‘ und ‚kontrolliertes Vokabular‘.

⁵¹ Bei Wp wird bei der Suche nach dem Lemma ‚Künstliche Sprache‘ auf die Seite ‚Konstruierte Sprache‘ weitergeleitet und somit beide Ausdrücke synonym gesetzt (http://de.wikipedia.org/wiki/K%C3%BCnstliche_Sprache [19.10.2007]).

⁵² Eine differenzierte Bestimmung von Fachsprache leistet etwa Roelcke (Roelcke 2005, S. 15ff.), wobei Registereingänge für ‚natürliche Sprache‘ und ‚künstliche Sprache‘ bei ihm gar nicht aufscheinen.

⁵³ „Natürlich-sprachlich basierte [Dokumentationssprachen]: Die Deskriptoren bestehen aus einer natürlichen Sprache, wie z. B. bei einem Thesaurus.“ (<http://de.wikipedia.org/wiki/Dokumentationssprache> [19.10.2007])

Thesaurus

Schließlich seien noch die Begriffsbestimmungen des Kernbegriffs selbst analysiert. Die beiden gedruckten Publikationen geben als Oberbegriff Dokumentationssprache an, die elektronische Publikation den häufig als Synonym verwendeten Ausdruck „kontrolliertes Vokabular“ (vgl. Bertram 2005, S. 127). Wp weist Einträge sowohl zu ‚Dokumentationssprache‘ als auch zu ‚kontrolliertes Vokabular‘ auf, ohne dass jeweils auf den anderen verwiesen oder das Verhältnis der beiden Begriffe zueinander expliziert würde. Der attributive Relativsatz zu „kontrolliertes Vokabular“ („dessen Begriffe durch Relationen miteinander verbunden sind“) ist redundant. Anders verfährt DIN, indem hier zwar eigentlich derselbe Oberbegriff, allerdings in paraphrasierter Form angesetzt wird: „eine geordnete Zusammenstellung von Begriffen und ihren [...] Bezeichnungen“.

GpID fährt fort, indem ausführlich aus DIN 1463 zitiert wird (was bei den Artikeln zum Themenkomplex Thesaurus eher die Ausnahme darstellt), und fügt zu dieser Definition noch an, dass Thesauri präskriptiv seien. Der letzte Satz verdichtet die Information nochmals zu „eine natürlich-sprachlich basierte Dokumentationssprache, die die umkehrbar eindeutige Zuordnung von Begriff und Bezeichnung der natürlichen Sprache anstrebt“.

Beinahe wörtlich findet sich diese Formulierung auch in TID, mit Verweisen auf ‚Vokabularkontrolle‘ und ‚terminologische Kontrolle‘ sowie auf ‚Deskriptor‘. Ein Hinweis auf letzteren fehlt erstaunlicherweise bei GpID; Wp bringt ihn zwar nicht im Definitionsteil des Artikels, wohl aber im Zuge der weiteren Ausführungen. Dort erfährt der Leser dann auch Näheres über die Funktion von Thesauri („zum Indexieren, Speichern und Finden von Dokumenten“). Diese Angaben sind recht ausführlich (wenngleich nicht sehr klar gegliedert und aufgebaut). GpID dagegen bringt nur sehr fachlich-abstrakt gehaltene Erläuterungen zur Funktion von Thesauri (s. Zitat oben). TID schließlich bleibt gänzlich auf einer abstrakt-definitiven Ebene. Dies ist insofern typisch für die drei Publikationen, weil sich darin die Textproduktion für je unterschiedliche intendierte Adressaten spiegelt: die Texte sind für jeweils andere Zielgruppen geschrieben.

Insgesamt lässt sich feststellen, dass TID kurze, kompakte Definitionen für Fachpublikum bietet, GpID mehr an begleitenden Hinweisen etwa zu Funktion und Gebrauch der beschriebenen Begriffe, auch unter Angabe von Beispielen, und Wp am ausführlichsten ist, wobei jedoch z.T. eine Verwischung zwischen Definition und weiteren Ausführungen zu einem Begriff stattfindet.

3.6 Mediostruktur

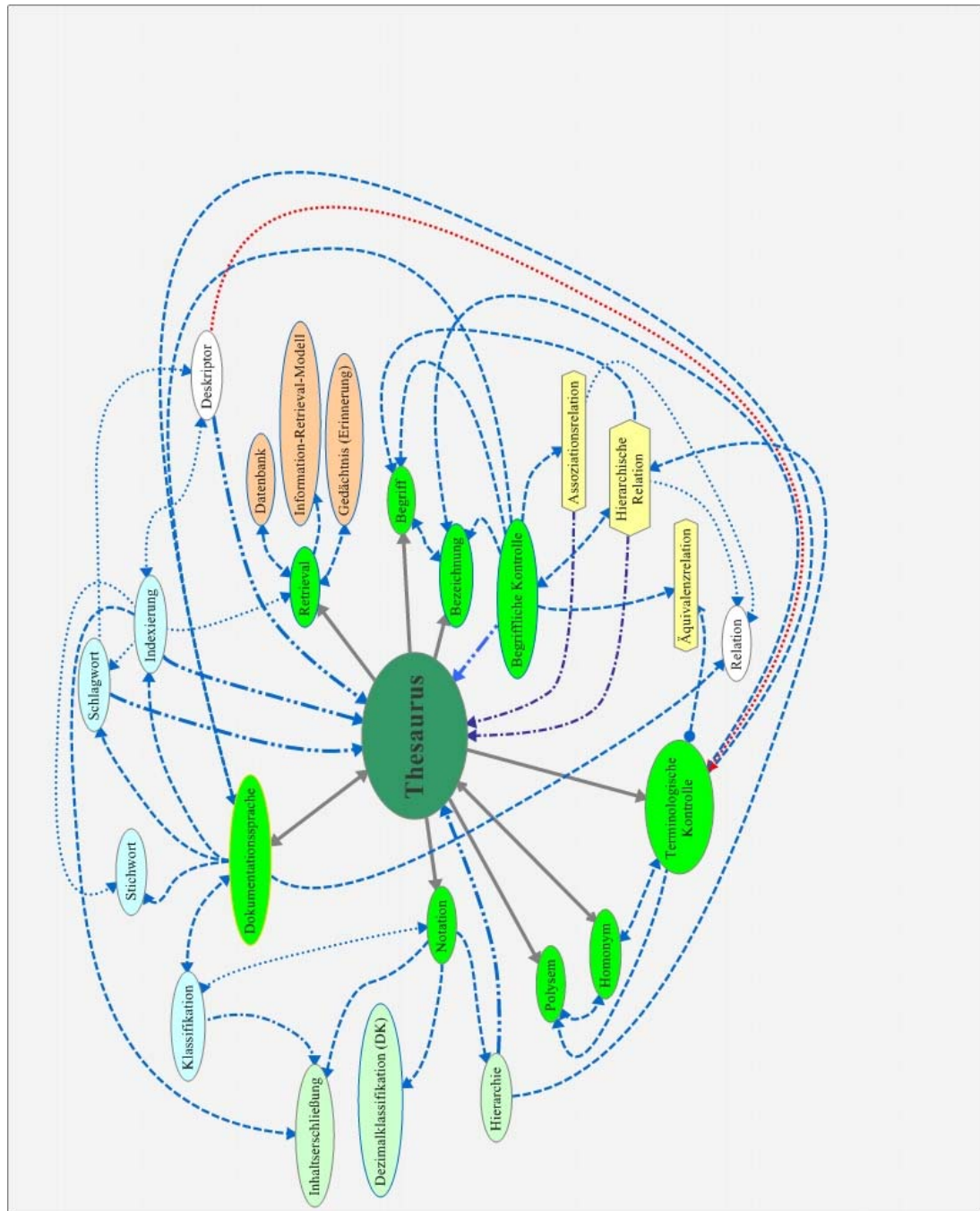
Als „fachlexikographische Mediostruktur“ bezeichnet Roelcke die Vernetzung von in Wörterbüchern enthaltenen Angaben, die als Verweis zwischen verschiedenen Artikeln oder Einträgen („artikelexterner Verweis“) oder innerhalb des selben Artikels oder Eintrags („artikelinterner Verweis“) angelegt sein können (Roelcke 2005, 137f.). Dieser Begriff lässt sich problemlos analog auf die für vorliegende Untersuchung relevanten Textsorten anwenden, wobei hier einerseits Verweise zu Synonymen sowie Ober- und Unterbegriffen („bedeutungsrelationaler Verweis“) (siehe den vorhergehenden Abschnitt), andererseits Verweise in den Definitionen und Bedeutungsbeschreibungen („onomasiologischer Verweis“) von Interesse sind. Einsatz und Häufigkeit von Verweisen hängen mit verschiedenen für die jeweilige Textsorte relevanten Faktoren zusammen. So helfen Verweise, den lückenlosen Aufbau von Definitionen hierarchisch angeordneter Begriffe zu gewährleisten, wobei allerdings gemäß der allgemeinen Definitionslehre jeweils das *genus proximum* auf der obersten Stufe der Definitionsleiter und die *differentiae specifica*e auf den untersten Stufen unbestimmt bleiben müssen. In erster Linie wichtig sind Verweise natürlich für den Benutzer des Nachschlagewerks, da sie für ihn die inhaltliche Kohärenz herstellen. Damit erweisen sie sich als Mittel zu einer hohen Textökonomie. Besonders deutlich wird die Steigerung der Nutzerfreundlichkeit bei der elektronischen Form des Verweises, dem Hyperlink. Darin besteht – neben der einfachen Zugänglichkeit – wohl der größte Mehrwert von Wikipedia. Empirische Erhebungen über den Gebrauch von Hyperlinks durch User sind – im Gegensatz zu Verweisen in gedruckten Publikationen – technisch problemlos zu bewerkstelligen.

Im Folgenden wird die Mediostruktur der zu untersuchenden Publikationen anhand des Themenkomplexes Thesaurus beispielhaft betrachtet. Das Lemma ‚Thesaurus‘ bildet dabei den Kern, von dem aus Verweise auf andere Lemmata dargestellt werden (erster Kreis). Von den Lemmata des ersten Kreises führen Verweise zu weiteren thematisch verbundenen Lemmata (zweiter Kreis). Auf die vollständige Darstellung darüber hinausführender Verweise wird aus Gründen der Übersichtlichkeit verzichtet.

GpID:

Die Verweisstruktur zum Themenkomplex Thesaurus ist in Abb. 1 dargestellt.⁵⁴ Vom Kern aus wird auf acht Lemmata bzw. die Einträge dazu verwiesen (erster Kreis). Nur in zwei von diesen Einträgen gibt es einen Rückverweis auf das Lemma ‚Thesaurus‘ (nämlich in ‚Dokumentationssprache‘ und in ‚Homonym‘). Der zweite Kreis enthält 13 Lemmata, womit das Verhältnis zwischen erstem Kreis und Kern 8:1 beträgt, das Verhältnis zwischen den Lemmata des zweiten zu denen des ersten Kreises aber deutlich niedriger ist (10:8). Mit anderen Worten: Die Lemmata des ersten Kreises benötigen zu ihrer Definition weitaus weniger Verweise zu Lemmata des zweiten Kreises als der Kern zu Lemmata des ersten Kreises, was z.T. an Verweisen innerhalb des ersten Kreises liegt. Noch deutlicher wird das bei einem Blick auf den zweiten Kreis. Zwar weisen Einträge zu einigen Lemmata hier durchaus Verweise in einen nächsten, dritten Kreis auf. Die Mehrzahl jedoch kommt ohne solche Verweise aus, sondern bildet nun ein dichtes Netz an artikelexternen Verweisen innerhalb des Themenkomplexes, gerade vom zweiten Kreis aus auch mit Rückverweisen auf den Kern ‚Thesaurus‘ (von ‚Hierarchie‘, ‚Schlagwort‘, ‚Indexierung‘).

⁵⁴ Die Abbildungen zeigen die Mediostruktur, indem Verweise durch vom jeweiligen Lemma ausgehende Pfeile dargestellt werden, Hin- und Rückverweise durch Doppelpfeil. Kern, erster und zweiter Kreis sind jeweils farblich gekennzeichnet. Durchgängige graue Pfeile stehen für Verweise vom Kern in den ersten Kreis. Andere Pfeilarten sind aus darstellungstechnischen Gründen gewählt.

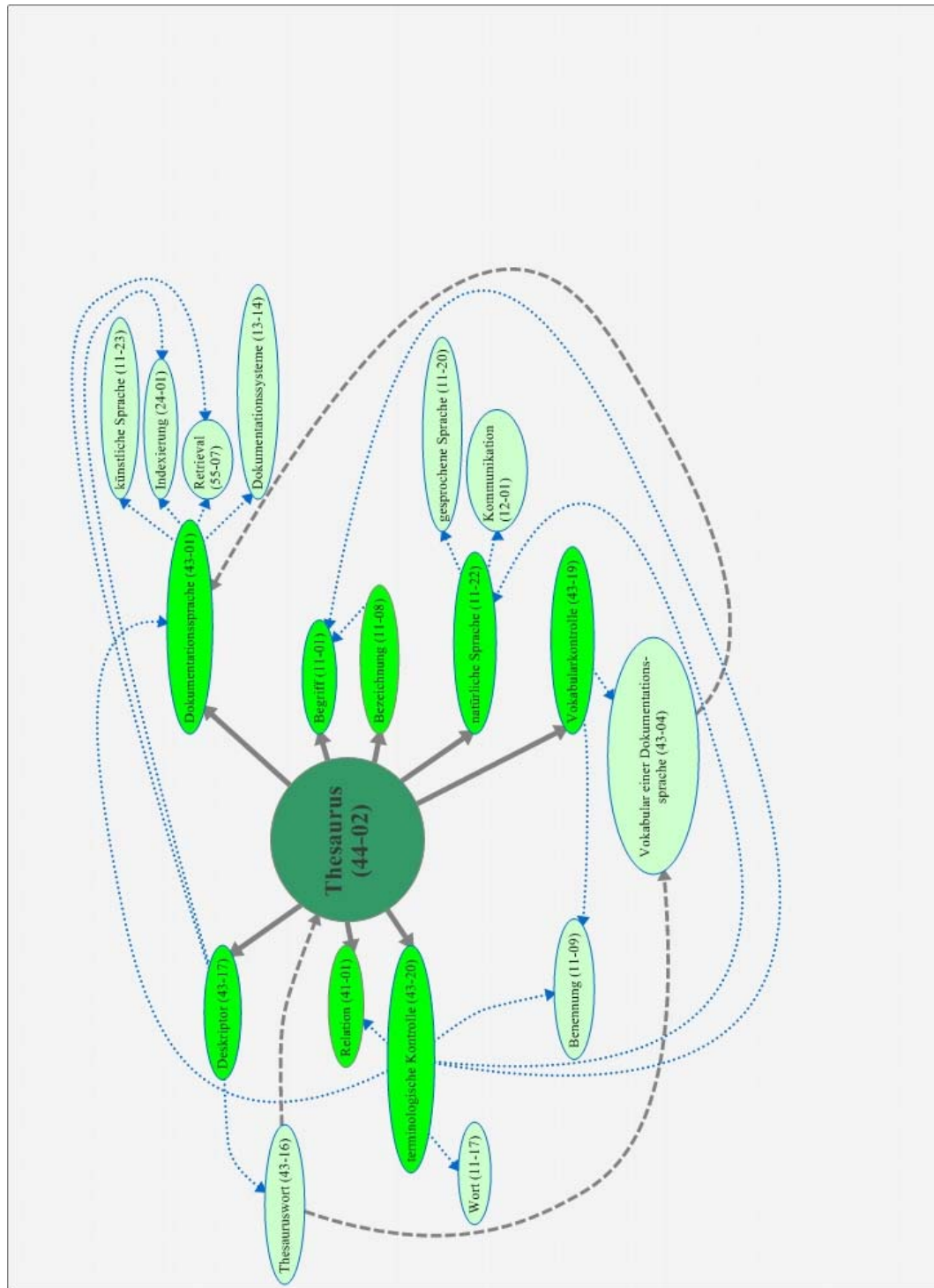


Auffällig ist bei GpID zweierlei: (1) Auf den für den Themenkomplex Thesaurus entscheidenden Begriff der Begrifflichen Kontrolle wird vom Kern aus nicht verwiesen, trotz Verweises von diesem Eintrag aus auf den Kern (also das Lemma ‚Thesaurus‘). Dass der Eintrag ‚Begriffliche Kontrolle‘ für das Thema sehr relevant ist, wird jedoch aus der Zahl der von ihm ausgehenden Verweise auf Einträge deutlich, von denen aus durchaus auf den Kern ‚Thesaurus‘ rückverwiesen wird.

Und (2): der wichtige Begriff Deskriptor scheint in GpID natürlich auf, ist in das Begriffnetz zu ‚Thesaurus‘ jedoch nur sehr lose eingebunden, da er nur zum dritten Kreis gehört, zwar seinerseits auf ‚Thesaurus‘ verweist, auf ihn selbst aber nur in ‚Schlagwort‘ und in ‚Indexierung‘ verwiesen wird.

TID:

Die Mediostruktur der Behandlung des Themenkomplexes Thesaurus im *Definitivonsteil* ist in Abb. 2 wiedergegeben. Zunächst fällt auf, dass vom Kern aus auf gleich viele Lemmata verwiesen wird wie in GpID (nämlich acht), obzwar die Lemmata dieses ersten Kreises in beiden Publikationen nur genau zur Hälfte (vier aus acht) übereinstimmen, nämlich: ‚Begriff‘, ‚Bezeichnung‘, ‚Dokumentationssprache‘ und ‚terminologische Kontrolle‘. Die zweite Hälfte der Lemmata des ersten TID-Kreises kommt bei GpID erst jenseits des zweiten Kreises (‚Deskriptor‘, ‚Relation‘, ‚Vokabularkontrolle‘) oder als eigenes Lemma gar nicht vor (‚natürliche Sprache‘). Somit decken sich die definitorischen Bestimmungen hinsichtlich der für sie herangezogenen Begriffe zu einem großen Teil. Verfolgt man jedoch vom Kern ausgehend eine der definitorischen Verweis-Ketten, so werden über diese Zwischenstufen die (unvermeidlichen) definitorischen Grenzen offensichtlich. So etwa über die Verweiskette ‚Thesaurus‘ – ‚Dokumentationssprache‘ – ‚künstliche Sprache‘ und von dort weiter zu ‚Sprache‘ (11-16) und ‚Begriff‘ (11-01). Letzterer wird sehr allgemein als „zur Umweltstrukturierung gebildete Denkeinheit“ bestimmt (TID, S. 36). Eine Definition, die mit ihrer Abstraktionshöhe der anvisierten fachkundigen Zielgruppe gerecht wird, für eine Einführung in dieses Fach jedoch wenig geeignet wäre. Auch hier erweist sich somit der Unterschied in den jeweils intendierten Benutzergruppen.

**Wp:**

Hinsichtlich seiner Makrostruktur unterscheidet sich Wikipedia grundsätzlich von gedruckten Werken, da Verweise im elektronischen Medium als Hyperlink für den User weitaus praktischer zu handhaben sind. Mühsames Nachschlagen ent-

fällt, mittels Hyperlink kann Verweisen mühelos nachgegangen werden. Mit diesem Qualitätssprung sind jedoch auch geänderte Anforderungen an eine sinnvolle Mediostruktur verbunden. Wp versucht dem Rechnung zu tragen, indem Autoren in den Richtlinien Hinweise zum sinnvollen Einsatz von Hyperlinks gegeben werden:

Links sollten sparsam und sinnführend sein. Zu viele Links lenken von den wichtigen Links ab, geben ein unruhiges Schriftbild und erschweren das Querlesen, da zu viel Aufmerksamkeit auf Unwichtiges gelenkt wird. Gute Links sollten tatsächlich nur dorthin führen, wo es entweder Erläuterungen (Fremdwörter, Fachbegriffe) oder weiterführende Informationen zum Thema gibt.⁵⁵

Für User wird jedoch schnell offensichtlich, dass diese Richtlinien oft nicht eingehalten und Hyperlinks oft über Gebühr eingesetzt werden.

Betrachtet man die im Definitionsteil (eigentlich: den beiden Definitionsteilen) zum Eintrag ‚Thesaurus‘ eingefügten elektronischen Verweise, so wird ein Aspekt dieser Problematik deutlich.⁵⁶ In den beiden ersten Absätzen, die neben Angaben zur Definition auch solche zu Etymologie und zu anderen Wortbedeutungen enthalten, finden sich auf sieben Zeilen elf Verweise, darunter neben ‚Dokumentationswissenschaft‘, ‚kontrolliertes Vokabular‘ und ‚Begriffe‘ auch auf ‚Griechisch‘ und ‚latinisiert‘, nicht jedoch für den hier zentralen Begriff Relation. Im zweiten Absatz, der nochmals eine Definition von Thesaurus im Sinne von BID enthält, fehlen dann manche dieser Links, obwohl die Ausdrücke in der Definition verwendet werden (‚Begriffe‘, ‚kontrolliertes Vokabular‘), dafür scheinen ‚Oberbegriff‘ und ‚Unterbegriff‘ auf, allerdings ohne einen Hinweis auf den verwandten Begriff Relation.

Aufgrund dieses tendenziell großzügigen Einsatzes von Hyperlinks schwillt die Anzahl der Lemmata im zweiten Kreis stark an. Klickt man etwa im Eintrag ‚Thesaurus‘ auf den ersten Link ‚Dokumentationswissenschaft‘, gelangt man zu einem Artikel, der optisch nicht weiter gegliedert ist (etwa in einen Definitions- und ei-

⁵⁵ <http://de.wikipedia.org/wiki/Wikipedia:Verlinken> [19.10.2007]

⁵⁶ Der Eintrag ist mit den hier relevanten Teilen im Anhang wiedergegeben, wo in diesem Fall die typografische Kennzeichnung der Verweise (Hyperlinks) beibehalten wird.

nen historischen Teil), im Textteil aber über 20 Links enthält (darunter etwa auf die Artikel ‚Internet‘, ‚Sputnik‘ und ‚USA‘!)⁵⁷.

Zwischen den drei Publikationen sind also im untersuchten thematischen Ausschnitt ‚Thesaurus‘ Unterschiede feststellbar, wobei GpID und TID hier Wp gegenüberstehen. Dies erklärt sich einerseits mit den Möglichkeiten, welche das Medium jeweils bietet, andererseits mit den Bedingungen der Produktion von Einträgen, die bei Wp einer ungleich geringeren Kontrolle unterliegen, was zu größerer Uneinheitlichkeit bei der Qualität der Einträge führt.

4 Zusammenfassung und Ausblick

Ziel dieser Arbeit war es, neuere Ansätze der Terminologearbeit, die im Bereich Wissensorganisation vonstatten geht, zu untersuchen. Dazu wurden zwei Publikationen ausgewählt, die innerhalb der letzten Jahre in neuen Auflagen herausgebracht worden sind, nämlich GpID (2004) und TID (2006). Diese beiden Druckwerke wurden einem elektronischen Medium gegenübergestellt, nämlich Wp; hier spiegelt die Arbeit den Stand vom Oktober 2007. Um den weiten Bereich Wissensorganisation für die Untersuchung handhabbar zu machen, wurde daraus ein Themenkomplex heraus- und exemplarisch unter die Lupe genommen. Dafür schien der Komplex Thesaurus gut geeignet, weil er ein zentrales Hilfsmittel der Wissensorganisation betrifft und einschlägige Publikationen sich diesem Thema daher im Detail widmen. Im Zuge der Arbeit hat sich erwiesen, dass die Beschränkung auf diesen thematischen Ausschnitt weit genug war, um relevanten Fragestellungen nachgehen zu können, jedoch eng genug, um diesen Ausschnitt einigermaßen eingehend zu betrachten.

Die Leitfrage der Arbeit zielte auf die Art und Weise, wie Terminologearbeit in den drei Publikationen realisiert wird und wie sich in der Folge die Produkte dieser Terminologearbeit (die Einträge in Glossar, Register mit Definitionsteil bzw. Enzyklopädie) unterscheiden. Zum Zweck der Untersuchung wurden einige der Kriterien, die Roelcke zur Analyse fachlexikografischer und terminografischer

⁵⁷ <http://de.wikipedia.org/wiki/Dokumentationswissenschaft> [19.10.2007]

Wörterbuchartikel beschreibt (Roelcke 2005, S. 132ff.) für die hier interessierenden Textsorten angepasst und angewendet.

Bei den übergreifenden Aspekten Makrostruktur, Mikrostruktur und Mediostruktur konnten auf diese Weise grundlegende Unterschiede herausgearbeitet werden.

In der Makrostruktur wird die Spezifik der jeweiligen Textsorte deutlich, da GpID als Glossar versucht, Termini umfassend zu definieren und in ihrem thematischen Zusammenhang darzustellen, gleichzeitig aber eine möglichst einfache Auffindbarkeit zu gewährleisten, während es TID um eine fachlich möglichst präzise, lückenlose Definition der Begriffe geht. Für Letzteres empfiehlt sich ein inhaltlich motivierter Aufbau, während sich für Ersteres eine alphabetische Ordnung anbietet. Wp kann als elektronisches Medium beide Vorteile miteinander verbinden.

Mikrostrukturell sind GpID-Artikel dafür deutlich länger, enthalten im Durchschnitt eine höhere Zahl an Verweisen und bieten vor allem neben reiner Definition auch verstärkt Information über Relevanz und praktischen Einsatz der dargestellten Sachverhalte, veranschaulicht oft durch Beispiele (vgl. die Artikel ‚Polyssem‘, GpID, S. 97 oder ‚Hierarchische Relation‘, GpID, S. 49f.). Zudem behandelt GpID häufig in einem Artikel mehrere zusammenhängende Begriffe (etwa im Artikel ‚Begriff‘ auch Begriffsinhalt und Begriffsumfeld, GpID, S. 9). TID dagegen widmet den letztgenannten je einen eigenen Eintrag, fasst diese aber in einer Untergruppe zusammen. Zwar definiert TID Begriffe zum Teil recht ausführlich, bleibt dabei aber auf der rein abstrakten Ebene (etwa bei ‚hierarchische Relation‘, TID, S. 60, 41-26). Auch gibt TID fallweise Beispiele an, bei den Artikeln des Themenkomplexes Thesaurus bleibt das eher die Ausnahme.

In Hinsicht auf den mikrostrukturellen Aufbau vergleicht sich Wp eher mit GpID: Artikel sind umfassender, enthalten über die Definition hinaus weiterführende Angaben und konkrete Umsetzungen (etwa Weblinks zu Thesaurus-Entwicklungsprogrammen). Diese größere Informationsfülle geht zuweilen auf Kosten einer klaren Strukturierung, wenn etwa Definition und zusätzliche Angaben ineinander übergehen oder redundant sind. Es fällt auf, dass in allen drei Publikationen kaum Quellen angegeben werden, am ehesten berufen sich Definitionen noch direkt auf die DIN (etwa im Artikel ‚Thesaurus‘, GpID, S. 120).

Hinsichtlich der Mediostruktur sind eher GpID und TID miteinander vergleichbar, während hier Wp Besonderheiten aufweist. Dies erklärt sich jedoch durch den wesentlichen Unterschied, den das Medium (bzw. der Übertragungskanal) darstellt. Verweise (hier: Hyperlinks) werden ungleich häufiger eingesetzt, ohne dass dieser Einsatz in jedem Fall sinnvoll scheint. Wenngleich einige Mängel von Wikipedia (wie in Abschnitt 2.4 dargestellt) grundsätzlicher Natur sind, so zeigt der Vergleich von Wikipedia mit den konventionellen Publikationen in Hinblick auf die Behandlung des Themenkomplexes Thesaurus, dass sie dort weniger gravierend ausfallen (siehe Abschnitt 2.5).

Über das Medium hinaus sind diese Unterschiede auch Folgen je anderer Zielsetzungen. So reicht das Spektrum der hauptsächlich angesprochenen Leser (Nutzer) von Fachexperten (TID) über fachlich vorgebildete Interessenten – Studenten, Fachleute angrenzender Fächer – (GpID) bis zu fachlich interessierten Laien (Wp).

Vor diesem Hintergrund kann die Behandlung des Themenkomplexes Thesaurus als insgesamt angemessen betrachtet werden.

Im Rahmen der vorliegenden Arbeit habe ich mich auf die *Produkte* der Terminologearbeit beschränkt. Nicht ins Auge gefasst wurden die konkreten *Prozesse*, die im Zuge der terminologischen Aufbereitung eines Bereichs wie Wissensorganisation erforderlich sind, d.h. einzelne Arbeitsschritte, wie Sammlung, Analyse und Festlegung von Termini. Auch Aspekte der Arbeitsorganisation wurden im Theoretischen Teil nur insofern berücksichtigt, als sie für grundlegende Unterschiede zwischen den behandelten Publikationen von Relevanz waren. Sollte in einem weiterführenden Schritt eine Prozessanalyse vorgenommen werden, so wären dazu etwa Interviews mit Autoren, Redakteuren oder auch Administratoren, die an der jeweiligen Publikation mitarbeiten, notwendig. Zugang zur Materialsammlung, von der die jeweilige Publikation ausgeht, könnte dabei auch weitere Aspekte beleuchten, wie die Quellenlage, explizit aufgestellte Richtlinien für das Verfassen von Einträgen und das Ausmaß, in dem die theoretischen Grundlagen für das jeweilige Projekt von Terminologearbeit herausgearbeitet und offengelegt worden sind.

Neben dem Aspekt der *Erstellung* könnte auch der Aspekt der *Nutzung* behandelt werden, etwa indem eine Soziologie der Nutzer erstellt wird. Im Rahmen dieser Arbeit wurde von antizipierten bzw. intendierten Nutzern ausgegangen, ohne den realen Nutzer empirisch zu fassen. Eine Untersuchung könnte hier herausarbeiten, inwieweit sich beide Gruppen decken oder aber unterscheiden.

Die hier geleistete Analyse der Mikrostruktur orientierte sich an inhaltlichen Fragen vor dem Hintergrund der allgemeinen Definitionslehre. Darüber hinaus könnte eine stärker sprachwissenschaftlich ausgerichtete Untersuchung auf interessante Ergebnisse kommen, was etwa textsortenspezifische Kommunikationsstrategien betrifft oder die Kohäsion innerhalb der einzelnen Artikel.

Als Grundlage für solche weiterführenden Fragestellungen dient die vorliegende Arbeit, indem sie (sozusagen „werkimmanent“) Strukturen herausgearbeitet und Vergleiche angestellt hat. Manche der dabei erzielten Ergebnisse könnten für die Redaktionsteams der jeweiligen Publikationen von Wert sein und zumindest zu punktuellen Verbesserungen beitragen.

Insgesamt scheint mir, dass die zentrale Rolle der Sprache für das menschliche Wissen innerhalb der Disziplin Wissensmanagement noch zu wenig Aufmerksamkeit erfährt. Das Augenmerk auf einen wesentlichen Aspekt davon, Terminologie und ihre Erarbeitung, zu richten, war mein Ziel bei dieser Arbeit.

5 Quellen

- Arntz, Reiner / Picht, Heribert / Mayer, Felix (2004): Einführung in die Terminologiearbeit. 5. Auflage, Hildesheim (Georg Olms Verlag)
- Bargmann, Monika (2007): Soziale Software – eine Einführung. Skriptum zur LV Online-Moderation und -Kommunikation für das Sommersemester 2007 an der FH Burgenland. Stand: 19.03.2007
- Beling, Gerd † / Port, Peter / Strohl-Goebel, Hildburg (Redaktion)(2006): Terminologie der Dokumentation und Information. Komitee Terminologie und Sprachfragen (KTS) der Deutschen Gesellschaft für Informationswissenschaft und Informationspraxis (DGI); 2., neu bearbeitete Ausgabe, Frankfurt am Main (Deutsche Gesellschaft für Informationswissenschaft und Informationspraxis)
- Bertram, Jutta (2005): Einführung in die inhaltliche Erschließung. Grundlagen – Methoden – Instrumente. Würzburg (Ergon)
- Budin, Gerhard (1996): Wissensorganisation und Terminologie. Die Komplexität und Dynamik wissenschaftlicher Informations- und Kommunikationsprozesse. Tübingen (Narr)
- Budin, Gerhard (2006): Kommunikation in Netzwerken – Terminologiemanagement; In: Andreas Blumauer, Tassilo Pellegrini (Hrsg.): Semantic Web. Wege zur vernetzten Wissensgesellschaft. Berlin (Springer)
- von Buol, Bettina (2000): Qualitätsgestützte, kooperative Terminologiearbeit. Aachen (Shaker)
- Dahlberg, Ingetraut (1994): Wissensorganisation – eine neue Wissenschaft? In: Rudolf Wille, Monika Zickwolff (Hrsg.): Begriffliche Wissensverarbeitung. Grundlagen und Aufgaben. Mannheim (Bibliographisches Institut), S. 225-238
- Dahlberg, Ingetraut (2006): Knowledge Organization: A New Science? In: Knowledge Organization 33, Würzburg (Ergon), S. 11-19
- Deutsches Terminologie-Portal: „Terminologiearbeit im Unternehmen“
<http://www.iim.fh-koeln.de/dtp/terminologieDT.html> [19.10.2007]
- DIN 1463, Teil 1 (1987): Erstellung und Weiterentwicklung von Thesauri – Ein-sprachige Thesauri. Berlin
- DIN 2330 (1979): Begriffe und Benennungen – Allgemeine Grundsätze. Berlin
- DIN 2333 (1987): Fachwörterbücher – Stufen der Ausarbeitung. Berlin
- DIN 2342 (1992): Begriffe der Terminologielehre – Grundbegriffe. Berlin
- Encyclopædia Britannica (2006): Fatally Flawed – Refuting the recent study on encyclopedic accuracy by the journal *Nature*;
http://corporate.britannica.com/britannica_nature_response.pdf [19.10.2007]

- Felber, Helmut / Budin, Gerhard (1989): Terminologie in Theorie und Praxis. Tübingen (Forum für Fachsprachen-Forschung 9)
- Giles, Jim (2005, 14. Dezember): Internet encyclopaedias go head to head. In: Nature Online, http://www.nature.com/news/2005/051212/pf/438900a_pf.html
- Giles, Jim (2007): Wikipedia 2.0 – now with added trust. In: New Scientist Magazine, Ausgabe 2622 vom 20. September 2007, S. 28f. und <http://technology.newscientist.com/article/mg19526226.200-wikipedia-20--now-with-added-trust.html> [19.10.2007]
- Glück, Helmut (Hg.) (2005): Metzler Lexikon Sprache. 3. Auflage, Stuttgart (J.B. Metzler),
- Graf, Klaus (2006, 19. Dezember): Re: (InetBib) Online-Lexikon zur Bayerischen Geschichte <http://www.ub.uni-dortmund.de/listen/inetbib/msg32061.html> [19.10.2007]
- Graff, Bernd (2007): Schmutzige Spielchen – Wikipedia und ungewollte Artikel von Werbern; In: sueddeutsche.de, 26.01.2007, <http://www.sueddeutsche.de/kultur/artikel/578/99479/article.html>
- Hansen, Hans Robert / Neumann, Gustaf (2005): Wirtschaftsinformatik 1. Grundlagen und Anwendungen. 9. Auflage, Stuttgart (Lucius & Lucius)
- Hjørland, Birger: What is Knowledge Organization (KO)? http://www.db.dk/bh/Lifeboat_KO/CONCEPTS/knowledge_organization.htm [19.10.2007]
- KÜDES (2002) [Konferenz der Übersetzungsdienste europäischer Staaten – Arbeitsgruppe Terminologie und Dokumentation]: Empfehlungen für die Terminologiearbeit. Bern (Schweizerische Bundeskanzlei, Sektion Terminologie)
- Kuhlen, Rainer / Seeger, Thomas / Strauch, Dietmar (Hrsg.) (2004): Grundlagen der praktischen Information und Dokumentation. Begründet von Klaus Liespen, Ernst Lutterbeck und Karl-Heinrich Meyer-Uhlenried, 5. völlig neu gefasste Auflage; München (K. G. Saur)
Band 1: Handbuch zur Einführung in die Informationswissenschaft und -praxis
Band 2: Glossar
- Malik, Fredmund (2004): Gefährliche Managementwörter. Und warum man sie vermeiden sollte. Frankfurt am Main (Frankfurter Allgemeine Buch)
- Oberhauser, Otto (2004): [Rezension zu: Grundlagen der praktischen Information und Dokumentation. Hrsg.: Rainer Kuhlen, Thomas Seeger, Dietmar Strauch.] In: Mitteilungen der VÖB 57, Nr. 3/4, S. 86-90 (auch unter: http://eprints.rclis.org/archive/00007437/01/vm_57_34-3.pdf) [19.10.2007]
- Roelcke, Thorsten (2005): Fachsprachen. 2. Auflage, Berlin (Erich Schmidt)

- Schulenburg, Frank / Raschka, Achim / Jungierek, Michail (2007): Der „McDonald's der Informationen“? Ein Blick hinter die Kulissen des kollaborativen Wissensmanagements in der deutschsprachigen Wikipedia; In: Bibliothek, Forschung und Praxis 31, Nr. 2, München (Saur), S. 225-229
- Voß, Jakob / Danowski, Patrick (2004): Bibliothek, Information und Dokumentation in der Wikipedia; In: Information – Wissenschaft & Praxis 55 (2004) 8, S. 457-462
- Wehn, Karin / Welker, Martin (2006, 01. September): Weisheit der Massen; <http://www.heise.de/tp/r4/artikel/23/23435/1.html> [19.10.2007]
- Wiegand, Dorothee (2007): Entdeckungsreise. Digitale Enzyklopädien erklären die Welt. In: *c't* 6/2007, S. 136-145
- Wright, Sue Ellen / Budin, Gerhard (1997): Handbook of Terminology Management. Volume 1: Basic Aspects of Terminology Management; Amsterdam/Philadelphia (John Benjamins)

6 Anhang

Ausgewählte Begriffe zum Themenkomplex Thesaurus und deren Definitionen in *Grundlagen der praktischen Information und Dokumentation* (GpID), *Terminologie der Information und Dokumentation* (TID), Wikipedia bzw. dem Wikipedia-Portal *Bibliothek, Information, Dokumentation* (Wp) und der DIN 1463 *Erstellung und Weiterentwicklung von Thesauri*. (Verweise innerhalb der Artikel auf andere Lemmata sind nicht eigens gekennzeichnet. Eine Ausnahme bildet der Wp-Artikel ‚Thesaurus‘, dessen Mediosstruktur in der vorliegenden Arbeit nicht durch eine Abbildung illustriert wird. Daher werden hier die Hyperlinks sichtbar gelassen, um so zumindest die Verweise vom Kern ‚Thesaurus‘ auf die Lemmata des ersten Kreises deutlich zu machen.)

Assoziationsrelation

GpID *engl.: associative relation*

Eine Assoziationsrelation ist eine zwischen Begriffen bzw. ihren Bezeichnungen als wichtig erscheinende Relation, die weder eindeutig hierarchischer Natur ist, noch als äquivalent angesehen werden kann (Beispiel: DIESELMOTOR und OTTOMOTOR sind verwandt durch gemeinsame Merkmale). Die Beziehungen dieser Relation haben ganz unterschiedlichen Charakter, z.B. instrumental, kausal, temporal, Antonymie, Vorgänger-Nachfolger oder Rohstoff-Erzeugnis. Der eigentliche Sinn dieser Relation besteht darin, zusätzlich zur hierarchischen Struktur Querbeziehungen zu anderen, für die Formulierung des Sachverhaltes möglicherweise geeigneten Deskriptoren anzubieten. Im Thesaurus wird diese Beziehung als Verwandter Begriff (related term) gekennzeichnet. (S. 5f.)

TID Eine Assoziationsrelation ist eine symmetrische Relation [sic!], die in Thesauri Begriffe miteinander verbindet, zwischen denen für die Recherche und Indexierung relevante Relationen bestehen, die nicht anderweitig dargestellt sind.

ANM

- (1) [Verweis auf ANM (3) unter ‚assoziative Relation‘:] In Dokumentationssprachen wird „Assoziationsrelation“ meist in einem eingeschränkten Sinn verwendet. „Assoziative Relation“ und „Assoziationsrelation“ sollten daher nicht synonym verwendet werden. [Eine assoziative Relation ist eine Relation zwischen zwei Elementen, die durch Assoziation mit einander verbunden werden. 41-17, S. 59]

Assoziationsrelationen werden meist durch Verweisungen wie „vgl.“ und „siehe auch“ ausgedrückt. (41-30, S. 61)

Wp [Der Hyperlink ‚Assoziationsrelationen‘ aus dem Artikel ‚Thesaurus‘ führt zum Artikel ‚Assoziation (Psychologie)‘.]:

‚**Assoziation** in der Psychologie und der Psychoanalyse ist die Theorie, dass der (menschliche) Geist lernt, indem er einfache, nicht reduzierbare Elemente nach bestimmten Prinzipien miteinander verknüpft.

Der Begriff der **Assoziation** dient dabei zur Erklärung des Phänomens, dass zwei (oder mehr) ursprünglich isolierte psychische Inhalte (wie z.B. Eindrücke, Gefühle oder auch Ideen), auch als *Assoziationsglieder* bezeichnet, eine so enge Verbindung eingehen, dass das Aufrufen eines Assoziationsgliedes das Auftreten eines oder mehrerer weiterer Assoziationsglieder nach sich zieht oder zumindest begünstigt.

Die Leistungen des Gedächtnisses beruhen nach der herrschenden Meinung auf eben

solchen Assoziationsketten. Damit ergibt sich als eine unabdingbare Voraussetzung des menschlichen Gedächtnisses die Fähigkeit zur Assoziation. Bedeutsam ist dies vor allem in der Gedächtnis- und Lernforschung.

Assoziation wird auch im Rahmen der technischen Mustererkennung als eine Eigenschaft von Neuronalen Netzen genannt.

http://de.wikipedia.org/wiki/Assoziation_%28Psychologie%29 [19.10.2007]

- DIN** Eine Assoziationsrelation ist eine zwischen Begriffen bzw. ihren Bezeichnungen als wichtig erscheinende Relation, die weder eindeutig hierarchischer Natur ist, noch als äquivalent angesehen werden kann.
- Dabei sollte jedoch beachtet werden, daß zwischen den Begriffen vielfältige Beziehungen bestehen. Begriffe sollten nur dann als assoziiert dargestellt werden, wenn diese Relationen für Indexierung und Retrieval hilfreich sind.
- Assoziierte Begriffe werden in den englischsprachigen Thesauri als RELATED TERM (RT) gekennzeichnet. Die deutsche Form lautet VERWANDTER BEGRIFF (VB). Zu den assoziierten Relationen gehören folgende Formen von Beziehungen: [es folgen: Relationen zum Determinationsbegriff und dem spezifizierenden Merkmalsbegriff; Gleichordnung (logische); Nebenordnung (räumliche, ontologische); Antonymie; Folge- bzw. Nachfolgebeziehungen; Affinität; jeweils mit Beispielen]. (DIN 1463, S. 6)

Deskriptor

- GpID** *engl.: descriptor*
- Ein Deskriptor ist ein Wort innerhalb eines Thesaurus, das für die Indexierung zugelassen ist. Alle anderen Elemente des Thesaurus haben den Status von Nicht-Deskriptoren (Synonymen); sie werden in den Thesaurus aufgenommen, und bilden somit einen Bestandteil des Zugangsvokabulars, können aber selbst nicht zur Indexierung und Recherche verwendet werden, sondern verweisen auf den zugehörigen Deskriptor. [S. 24]
- TID** Ein Deskriptor ist ein selbständiges Thesauruswort, das zur Indexierung und zum Retrieval verwandt wird. [43-17, S. 64]
- Wp** Als **Deskriptor** (Plural *Deskriptoren*) oder **Schlagwort** (Plural *Schlagwörter*) bezeichnet man meist vorgegebene Bezeichnungen, die zur inhaltlichen Beschreibung eines Objektes ausgewählt werden können. Die Vergabe von Deskriptoren wird auch als *Verschlagwortung* oder *Indexierung* bezeichnet. Im Gegensatz zu Stichwörtern – wichtigen Wörtern, die in einem Text vorkommen – entstammen Schlagwörter meist einem kontrollierten Vokabular, in dem festgelegt ist, welche Wörter für welche Sachverhalte verwendet werden sollen. Beispiele für solche Vokabulare sind Normdateien und Thesauri. Die Vergabe von Deskriptoren ist Bestandteil der inhaltlichen Erschließung von Dokumenten, zum Beispiel in Bibliotheken. [<http://de.wikipedia.org/wiki/Deskriptor> 11.10.2007]
- DIN** Deskriptor ist eine Bezeichnung (Vorzugsbezeichnung), die zur Inhaltskennzeichnung (Indexierung) zugelassen ist.

Dokumentationssprache

- GpID** *engl.: documentary language; documentation language*
- Eine Dokumentationssprache ist eine künstliche Sprache zur Indexierung, Speicherung und zum Retrieval innerhalb von Dokumentationssystemen. Hierzu zählen insbesondere Verfahren zur Indexierung mit Stichwörtern oder Schlagwörtern sowie der Thesaurus und die Klassifikation. (S. 29)

TID Eine Dokumentationssprache ist eine künstliche Sprache zum Zweck der Indexierung, der Speicherung und des Retrievals innerhalb von Dokumentationssystemen.

SYN Informationsrecherchesprache; IRSp (43-01, S. 63)

Wp Eine **Dokumentationssprache** ist eine konstruierte Sprache zur Nutzung innerhalb von Informations- und Dokumentationssystemen, d. h. zur Indexierung, Speicherung und zum Retrieval von Inhalten, indem einzelne Deskriptoren aus der Dokumentationssprache einem Dokument zugeordnet werden.

Man kann zwei Arten von Dokumentationssprachen unterscheiden:

- Natürlich-sprachlich basierte: Die Deskriptoren bestehen aus einer natürlichen Sprache, wie z. B. bei einem Thesaurus.
- Nicht natürlich-sprachlich basierte: Die Deskriptoren bestehen nicht aus einer natürlichen Sprache, doch die Beschreibung der Inhalte geschieht mit natürlichen Sprachelementen, wie z. B. bei einer Systematik.

Gebräuchliche Dokumentationssprachen sind

- Thesaurus
- Schlagwortsystem
- Klassifikationen

Dokumentationssprachen zeichnen sich dadurch aus, dass es sich bei ihnen um ein kontrolliertes Vokabular handelt oder durch Notationen die Begriffe eindeutig identifizierbar sind. Auf diese Weise werden im Gegensatz zur natürlichen Sprache Begriffe und Bezeichnungen eindeutig aufeinander bezogen, so dass keine Homonyme, Polyseme und Synonyme mehr auftreten. (<http://de.wikipedia.org/wiki/Dokumentationssprache>, [19.10.2007])

DIN Eine Dokumentationssprache ist eine Menge sprachlicher Ausdrücke (Bezeichnungen), die, nach bestimmten Regeln angewendet, der Beschreibung von Dokumenten zum Zweck des Speicherns und einer gezielten Wiederauffindung (Retrieval) dient. Dokumentationssprachen können dargestellt werden durch Schlagwortsysteme, Thesauri und Klassifikationen. (DIN 1463, S. 2)

Information Retrieval

GpID *engl.: retrieval*

Retrieval (auch Recherche oder Information Retrieval genannt) bezeichnet den Arbeitsvorgang des gezielten Suchens bzw. Wiederfindens von relevanten Daten und Fakten zu einer speziellen Fragestellung in gedruckten oder elektronischen Informationsmitteln. Im heutigen Sprachgebrauch wird Recherche häufig mit dem Online-Retrieval gleichgesetzt. Bei der Online-Recherche werden Suchanfragen mit Hilfe der Retrievalsprache unter Verwendung von Operatoren formuliert und von einem Rechner im Direktzugriff auf eine Datenbank durchgeführt. Retrieval beschäftigt sich mit der Suche nach Informationen und mit der Repräsentation, Speicherung und Organisation von Wissen. Information Retrieval modelliert Informationsprozesse, in denen Benutzer aus einer großen Menge von Wissen die für ihre Problemstellung relevante Teilmenge suchen. Dabei entsteht Information, die im Gegensatz zum gespeicherten Wissen problembezogen und an den Kontext angepasst ist. – In der Psychologie bedeutet Retrieval (Abruf) den Zugang zu der zuvor gespeicherten Information im Gedächtnis (Erinnerung). (S. 107)

TID Retrieval ist der Vorgang, gespeicherte Daten in Auswahl wieder aufzufinden. (55-07, S. 76)

Wp **Information Retrieval (IR)** bzw. **Informationswiedergewinnung**, gelegentlich **Informationsbeschaffung**, ist ein Fachgebiet, das sich mit computergestütztem inhaltsorien-

tiertem Suchen beschäftigt. Es ist ein Teilgebiet der Informationswissenschaft, der Computerlinguistik wie auch der Informatik. Wie der Begriff *retrieval* (deutsch *Wiedergewinnung, Auffindung*) sagt, sind Informationen in großen Datenbeständen zunächst verloren und müssen wieder gewonnen bzw. wieder gefunden werden.

DEFINITION:

Zwei Konzepte prägen das IR und grenzen es von der Suche in herkömmlichen Datenbanken ab:

- *Vagheit*: Der Benutzer kann sein "diffuses" Informationsbedürfnis nicht präzise und formal (wie z. B. in SQL in relationalen Datenbanken) ausdrücken. Die Anfrage enthält daher vage Bedingungen.
- *Unsicherheit*: Dem System fehlen Kenntnisse über den Inhalt der Dokumente (die Texte, Bilder, Video etc. enthalten können). Dies führt zu fehlerhaften und fehlenden Antworten. Probleme bei Texten bereiten z. B. Homographie (Wörter, die gleich geschrieben werden; z. B. Bank - Geldinstitut, Sitzgelegenheit) und Synonyme (Bank und Geldinstitut).

Methoden des Information-Retrieval werden in Internetsuchmaschinen (z. B. Google), aber auch in Digitalen Bibliotheken (z. B. zur Literatursuche), in Bildsuchmaschinen usw. verwendet. Auch Antwortsysteme oder Spamfilter verwenden IR-Technologien.

Generell sind am IR zwei (sich unter Umständen überschneidende) Personenkreise beteiligt (vgl. Abbildung rechts). [...]

(<http://de.wikipedia.org/wiki/Information-Retrieval> 19.10.2007)

DIN Eine Dokumentationssprache ist eine Menge sprachlicher Ausdrücke (Bezeichnungen), die, nach bestimmten Regeln angewendet, der Beschreibung von Dokumenten zum Zweck des Speicherns und einer gezielten Wiederauffindung (Retrieval) dient. (DIN 1463, S. 2)

Thesaurus

GpID *engl.: thesaurus*

Der Thesaurus als Dokumentationssprache wird in der DIN 1463 in seinen wesentlichen Merkmalen beschrieben. Dort wird der Thesaurus im informationswissenschaftlichen Sinne so definiert: „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. Er ist durch folgende Merkmale gekennzeichnet: (a) Begriff und Bezeichnung werden eindeutig aufeinander bezogen (terminologische Kontrolle), indem Synonyme möglichst vollständig erfasst werden, Homonyme und Polyseme besonders gekennzeichnet werden, für jeden Begriff eine Bezeichnung (Vorzugsbenennung, Begriffsnummer oder Notation) festgelegt wird, die den Begriff eindeutig vertritt, (b) Beziehungen zwischen den Begriffen (repräsentiert durch ihre Bezeichnungen) werden dargestellt.“ Diese Definition ist zu ergänzen um: Der Thesaurus ist präskriptiv, indem er für seinen Geltungsbereich festlegt, welche begrifflichen Einheiten zur Verfügung gestellt werden und durch welche Bezeichnungen diese repräsentiert werden. Ein Thesaurus ist damit eine natürlich-sprachlich basierte Dokumentationssprache, die die umkehrbar eindeutige Zuordnung von Begriff und Bezeichnung der natürlichen Sprache anstrebt. (S. 120)

TID Ein Thesaurus ist eine Dokumentationssprache, die die umkehrbar eindeutige Zuordnung von Begriffen und Bezeichnungen der natürlichen Sprache anstrebt und in der vollständige Vokabularkontrolle und terminologische Kontrolle ausgeübt und die Begriffe sowie die zugehörigen Relationen zu anderen Deskriptoren dargestellt werden. (44-02, S. 65)

Wp Thesaurus [Lemma]:

Ein **Thesaurus** bzw. **Wortnetz** ist in der [Dokumentationswissenschaft](#) ein [kontrolliertes Vokabular](#), dessen [Begriffe](#) durch Relationen miteinander verbunden sind. Die Bezeichnung wird gelegentlich auch für **linguistische Thesauri** oder wissenschaftliche [Wortschatz](#)-Sammlungen verwendet. Das Wort „Thesaurus“ stammt aus dem [Griechischen](#) (*thesauros* für Schatz, Schatzhaus, mit [latinisierter](#) Endung auf -us).

Thesaurus [Zwischenüberschrift]:

Als Thesaurus bezeichnet man ein Modell, das versucht, ein Themengebiet genau zu beschreiben und zu repräsentieren. Es besteht aus einer systematisch geordneten Sammlung von Begriffen, die in thematischer Beziehung zueinander stehen. Der Thesaurus ist ein kontrolliertes Vokabular (Attributwertebereich) für das jeweils zu beschreibende [Attribut](#). Es werden in erster Linie [Synonyme](#), aber auch [Ober-](#) und [Unterbegriffe](#) verwaltet.

Beispiel: Bildnis (Synonym: Abbild, Bild, Spiegelbild), Tischler (Oberbegriff: Handwerker)

Geschichte [Zwischenüberschrift]:

Der Begriff *Thesaurus* kommt aus dem Griechischen und lässt sich im Englischen erstmals 1736 nachweisen^[1]. Im allgemeinen Wortsinn bezeichnete es zunächst einen „Wissensspeicher“ wie beispielsweise ein Wörterbuch oder eine Enzyklopädie. Der vor allem im Englischen Sprachraum einflussreiche, 1852 von [Peter Mark Roget](#) veröffentlichte [Roget's Thesaurus of English Words and Phrases](#) rückte die Bedeutung des Begriffs in die Richtung eines linguistischen Thesaurus. Im Bereich des [Information Retrieval](#) wurde der Begriff erstmals 1957 von [Peter Luhn](#) verwendet als in den 1950er Jahren verschiedene Systeme zur [Indexierung](#) entwickelt wurden. Zu den ersten Thesauri, die in der Praxis zur Erschließung eingesetzt wurden, gehören das System von Du Punt (1959) und der *Thesaurus of ASTIA Descriptors* (1960). Ein einheitliches Format für Thesauri wurde 1967 mit dem *Thesaurus of Engineering and Scientific Terms* (TEST) vorgelegt. Aus den bereits von Anfang an entwickelten Regeln für den Aufbau von Thesauri entwickelten sich mit der Zeit allgemeine Standards, die die Form des klassischen Thesaurus zur Dokumentation festlegen. Dazu gehören die von Derek Austin und Dale entworfenen *UNESCO's Guidelines for the Establishment and Development of Monolingual Thesauri* deren Inhalte in den ISO-Standard 2788 (1986) einfließen.

Thesaurus zur Dokumentation [Zwischenüberschrift]:

In der [Dokumentationswissenschaft](#) hat sich der Thesaurus als geeignetes Hilfsmittel zur [Sacherschließung](#) und zum Auffinden von Dokumenten erwiesen. Dabei dienen Relationen zwischen den einzelnen Begriffen zum Auffinden bei der [Indexierung](#) (Vergabe von Schlagworten) und bei der [Recherche](#). Im Gegensatz zu einem linguistischen Thesaurus enthält ein Thesaurus zur Dokumentation ein [kontrolliertes Vokabular](#), d.h. eindeutige Benennungen ([Deskriptoren](#)) für jeden Begriff. Unterschiedliche Schreibweisen (*Photo/Foto*), [Synonyme](#) bzw. als gleichbedeutend behandelte Quasi-Synonyme, [Abkürzungen](#), Übersetzungen etc. werden durch [Äquivalenzrelationen](#) miteinander in Beziehung gesetzt. Begriffe werden außerdem durch [Assoziationsrelationen](#) und [hierarchische](#) Relationen vernetzt.

Der Thesaurus dient als Dokumentationssprache zum Indexieren, Speichern und Finden von Dokumenten. Die Relationen ermöglichen es, bei der Indexierung und Recherche passende [Benennungen](#) für gesuchte Begriffe zu finden. Bei der Suche können Thesauri durch die automatische Erweiterung der Suchanfrage durch Synonyme und Unterbegriffe hilfreich sein.

Ein Thesaurus kann damit auch allgemein zur Begriffsklärung dienen und hat im besten Fall die Funktion einer Normdatei inne. Im Gegensatz zu einer monohierarchischen Tabelle oder Datenbank kann der Thesaurus eine polyhierarchische Struktur besitzen (d.h. ein Unterbegriff kann mehrere Oberbegriffe haben).

Die Thesaurusnormen DIN 1463-1 bzw. das internationale Äquivalent ISO 2788 sehen folgende Relationsarten und dazugehörige Abkürzungen vor:

[Es folgt eine Tabelle „Kürzel und Bezeichnung“ (nach DIN 1463 bzw. ISO 2788)]

Die häufigsten Relationen in einem Thesaurus sind [Äquivalenz-](#), [Assoziations-](#) und [hierarchische Relationen](#).

In der Regel wird ein Element einer Äquivalenzrelation, also eine Benennung, als Vorzugsbenennung festgelegt. Die Nicht-Vorzugsbenennungen erhalten einen Verweis auf die ihnen äquivalente Vorzugsbenennung.

Beispiel:

Fahrzeug verweist auf die Unterbegriffe *LKW* und *PKW*.

Auto verweist auf die Vorzugsbenennung *PKW* und mit einer Assoziationsbeziehung („siehe auch“) auf *LKW*.

Siehe auch: [semantisches Netz](#)

[Es folgen Absätze zu „Thesaurus als Sammelwerk“, „Linguistische Thesauri“ sowie „Beispiele“, „Quellen“ und „Literatur“.]

<http://de.wikipedia.org/wiki/Thesaurus> [11.10.2007]

DIN 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.

Er ist durch folgende Merkmale gekennzeichnet:

- (a) Begriffe und Bezeichnungen werden eindeutig aufeinander bezogen („terminologische Kontrolle“), indem
 - a. Synonyme möglichst vollständig erfaßt werden,
 - b. Homonyme und Polyseme besonders gekennzeichnet werden,
 - c. für jeden Begriff eine Bezeichnung (Vorzugsbenennung, Begriffsnummer oder Notation) festgelegt wird, die den Begriff eindeutig vertritt,
- (b) Beziehungen zwischen Begriffen (repräsentiert durch ihre Bezeichnungen) werden dargestellt. (DIN 1463, S. 2)

7 Lebenslauf

Daten zur Person

Name Stefan Hauser
Geburtsdaten 09.04.1968, Hall in Tirol

Akademische Ausbildung

2005–2007 Fachhochschullehrgang *Angewandtes Wissensmanagement*, FH Burgenland, Eisenstadt
1995 Abschluss des Studiums der *Deutschen Philologie (Lehramt)* und von *Philosophie, Psychologie, Pädagogik (Lehramt)* an der Leopold-Franzens-Universität Innsbruck

Beruflicher Werdegang

2006- Referent für E-Learning am Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft (BFW)
2004–2006 Tätigkeit als Lektor sowie als Kursleiter für Deutsch (u.a. für die *Aktion Österreich – Tschechische Republik*)
2003/2004 Referent im Bundesministerium für auswärtige Angelegenheiten
2001–2003 Assistent des Leiters der Geschäftsstelle der Ethikkommission an der Medizinischen Fakultät der Universität Innsbruck
1997–2001 Österreichischer Lektor und Tutor an der *School of Language, Law and International Studies* der *University of Surrey*, Guildford, Großbritannien
1996/1997 Unterrichtspraktikum am Bundesgymnasium Bludenz
1993–1996 Österreichischer Lektor am Germanistischen Lehrstuhl der *Univerzita Palackého* in Olomouc/Olmütz, Tschechische Republik