

## Inhalt

<b>Vorwort</b> .....	9
<b>1. Das neue Medium und die Lexikografie: Eine Bandbreite an Möglichkeiten</b> .....	11
<b>2. Zur Verortung und Eingrenzung des Themas dieser Arbeit</b> .....	15
<b>3. Ordnende Betrachtungen zu elektronischen Wörterbüchern und lexikografischen Prozessen</b> .....	21
3.1 Zum Thema des Kapitels .....	21
3.2 Arten elektronischer Wörterbücher.....	23
3.2.1 Müller, Edelmüller und Sondermüller .....	23
3.2.2 Unterscheidungsmöglichkeiten elektronischer Wörterbücher: Bereits vorgelegte Typologien.....	27
3.2.3 Was ist ein elektronisches Wörterbuch?.....	29
3.2.4 Elektronisches Wörterbuch – Wortschatzinformationssystem .....	32
3.2.5 Automatisch erstellte vs. lexikografisch bearbeitete Wortschatzinformationssysteme .....	33
3.3 Arten von lexikografischen Prozessen.....	42
3.3.1 All in one: Die Duden-Ontologie.....	42
3.3.2 Arten von lexikografischen Prozessen nach H.E. Wiegand.....	43
3.3.3 Arten von lexikografischen Prozessen. Eine erweiterte Übersicht ....	46
3.4 Ebenen im lexikografischen Prozess.....	53
3.5 Relevanz der getroffenen Unterscheidungen .....	56
<b>4. Anforderungen an die Modellierung lexikografischer Daten</b> .....	59
4.1 Funktionalitäten von Wortschatzinformationssystemen .....	60
4.2 Unterstützung im lexikografischen Prozess.....	68
4.3 Schnittstelle zur automatischen Sprachverarbeitung .....	72
4.4 Allgemeine Eigenschaften der Datenhaltung.....	73
<b>5. Die XML-basierte Modellierung einer lexikografischen Datenbasis</b> .....	77
5.1 Begründung der Wahl von XML und verwandten Standards.....	79
5.1.1 Was ist XML?.....	79
5.1.2 Vorteile des Einsatzes von XML .....	82
5.1.3 Einige Bemerkungen zum Gesamtaufbau einer XML-basierten Redaktionsumgebung.....	87

5.1.4	XML als „Piece of a Larger Picture“ .....	89
5.1.5	Notwendigkeit von Richtlinien für die Anwendung von XML .....	91
5.2	Differenzierung verschiedener Ansätze einer DTD-Entwicklung .....	95
5.2.1	Maßgeschneiderte vs. Standard-DTD .....	95
5.2.2	Strukturierung eines bestehenden Datenbestandes vs. Neuerstellung eines Datenbestandes .....	99
5.2.3	Layoutorientierte vs. inhaltsorientierte Datenauszeichnung .....	100
5.2.4	Layoutbasierte vs. konzeptuelle Inhaltsmodellierung .....	103
5.2.5	Angemessenheit der Modellierung .....	106
5.2.6	Strenge vs. weiche Struktur .....	107
5.3	Rahmenbedingungen und Zielsetzungen des Modellierungskonzeptes .....	109
5.4	Modellierung des Inhaltsstrukturenprogramms .....	117
5.4.1	Inhaltliche Konzeption .....	117
5.4.2	Standardbasierte Ausdrucksmöglichkeit: XML-DTDs .....	132
5.4.2.1	Graphen und Bäume .....	132
5.4.2.2	Kurzeinführung in die Syntax von XML .....	135
5.4.2.3	„The Well-Dressed-DTD“: einige Voraussetzungen für eine von Menschen lesbare DTD .....	146
5.4.2.3.1	Optische Gestaltung .....	146
5.4.2.3.2	Verteilung von Elementen, Attributen und Entities .....	148
5.4.2.3.3	Modularisierung von DTDs .....	149
5.4.2.3.4	Kommentierung der DTD .....	150
5.4.2.3.5	Dokumentation der DTD .....	151
5.4.2.3.6	Sprache .....	152
5.4.3	Richtlinien für die XML-basierte Modellierung des Inhaltsstrukturenprogramms .....	152
5.4.3.1	Modellierung des hierarchischen Aufbaus des Inhaltsstrukturenprogramms .....	153
5.4.3.2	Modellierung der terminalen Elemente .....	158
5.4.3.2.1	Unterscheidung in Angaben <sub>v</sub> , Kommentare und Hinweise ...	158
5.4.3.2.2	Modellierung von Angaben <sub>v</sub> .....	160
5.4.3.2.3	Modellierung von Kommentaren .....	167
5.4.3.2.4	Modellierung von Hinweisen .....	168
5.4.3.2.5	Nicht gekennzeichnete terminale Elemente .....	169
5.5	Modellierung des Vernetzungsstrukturenprogramms .....	169
5.5.1	Inhaltliche Konzeption .....	170
5.5.2	Standardbasierte Ausdrucksmöglichkeit: XLink .....	180
5.5.3	Richtlinien für die XML-basierte Modellierung des Vernetzungsstrukturenprogramms .....	190

5.6	Mehr als ein Instruktionenbuch: Die Anwendung des Modellierungskonzeptes im lexikografischen Prozess .....	204
5.7	Flexibilität der Darstellung der entsprechend modellierten Daten ....	210
5.7.1	Grundsätzliches .....	210
5.7.2	Standardbasiertes Verfahren: Darstellung mit XSL(T) .....	212
5.7.3	Veranschaulichung anhand eines Beispiels .....	216
5.7.4	Umgang mit Verdichtung in medienneutral konzipierten lexikografischen Prozessen .....	229
5.8	Perspektive Wörterbuchverbund .....	233
<b>6.</b>	<b>Modellierung einer lexikografischen Datenbasis im Projekt <i>ellexiko</i>: ein Werkstattbericht .....</b>	<b>243</b>
6.1	Kurzbeschreibung des Projekts .....	243
6.2	Rolle des Modellierungskonzeptes bei der DTD-Entwicklung .....	246
6.3	Vorgehensweise bei der DTD-Entwicklung .....	248
6.4	Skizze der technischen Redaktionsumgebung .....	249
6.5	Die Modellierung des Inhaltsstrukturenprogramms .....	250
6.5.1	Grundsätzliche Aufteilung der DTDs .....	250
6.5.2	Grundstruktur eines Artikels im <i>ellexiko</i> -Wörterbuch (bearbeitete Artikel) .....	253
6.5.3	Anwendung der Richtlinien zur Modellierung des Inhaltsstrukturenprogramms .....	259
6.5.3.1	Modellierung des hierarchischen Aufbaus des Inhaltsstrukturenprogramms .....	259
6.5.3.2	Modellierung der terminalen Elemente .....	261
6.5.3.2.1	Unterscheidung in Angaben <sub>v</sub> , Kommentare und Hinweise .....	261
6.5.3.2.2	Modellierung von Angaben <sub>v</sub> .....	264
6.5.3.2.3	Modellierung von Kommentaren .....	271
6.5.3.2.4	Modellierung von Hinweisen .....	273
6.6	Modellierung des Vernetzungsstrukturenprogramms .....	276
6.7	Weitere Aspekte einer „Well-Dressed-DTD“ .....	287
6.7.1	Benennung der Objekte .....	287
6.7.2	Anordnung der Entities, Elemente und Attribute und Kommentierung der DTDs .....	288
6.7.3	Dokumentation der DTDs .....	289
6.8	Darstellung der Artikel und Perspektiven für die Recherche .....	292
6.9	Die DTD-Entwicklung im Spannungsfeld zwischen theoretischen Richtlinien und ihrer praktischen Umsetzung .....	298

7.	<b>Schlussbemerkung</b> .....	301
8.	<b>Literatur</b> .....	303
8.1	Wörterbücher .....	303
8.2	Forschungsliteratur .....	303

## **Vorwort**

Dieses Buch ist eine überarbeitete Fassung meiner im Juni 2005 an der Heidelberger Universität eingereichten Dissertation. Zum Gelingen dieser Arbeit haben viele Menschen beigetragen, bei denen ich mich hiermit – zumindest auszugsweise – bedanken möchte: Zunächst bedanke ich mich bei Herbert Ernst Wiegand für die interessanten, spannenden und auf gewisse Weise auch unterhaltsamen Seminare während meiner Studienzeit, die mich auf die Fährte der (Meta-)Lexikografie gebracht haben. Dabei habe ich besonders seine breit angelegte fachliche Neugierde schätzen gelernt, die auch einer solchen Arbeit auf der Grenze zwischen Metalexikografie und Texttechnologie offen gegenübersteht. Außerdem möchte ich meinen Kollegen und Kolleginnen am Institut für Deutsche Sprache für Anregungen, Unterstützung und kollegiale Kritik danken; genauso den anderen in- und ausländischen Kollegen im Arbeitsfeld der Wörterbuchforschung und Lexikografie.

Ein großes Projekt wie eine Doktorarbeit, die einige Zeit in Anspruch nimmt, wird jedoch auch immer vom persönlichen Umfeld getragen. Deshalb möchte ich mich als erstes bei meinen Eltern bedanken, die mich meinen eigenen Weg haben gehen lassen und dabei immer ein sicheres Netz für mich waren und sind. Ihnen habe ich viel zu verdanken. Ich bedanke mich nicht weniger bei meinem Mann Sebastian, der im Alltag durch Unterstützung in vielerlei Hinsicht sicher den größten Beitrag zum Gelingen der Arbeit geleistet hat. Auch möchte ich mich bei unseren Töchtern Marie und Anna bedanken, die auf ihre Weise die Entstehung der Doktorarbeit und die anschließenden Prüfungen befördert haben. Nicht zuletzt bedanke ich mich bei meinen Freundinnen und Freunden, insbesondere bei Eva und Ingrid, für ihre Zuneigung und Verbundenheit.

Und schon im Voraus bedanke ich mich bei allen Lesern des Buches für ihr Interesse und wünsche allen nach soviel Dankesworten eine hoffentlich anregende Lektüre.

Ich widme diese Arbeit meiner Familie.

Carolin Müller-Spitzer

Plankstadt, September 2007

## 1. **Das neue Medium und die Lexikografie: Eine Bandbreite an Möglichkeiten**

Der Computer kann [...] Wörterbücher rationalisieren, systematisieren, präzisieren, symmetrisieren, parametrisieren, modularisieren, homogenisieren, technisieren, operationalisieren, automatisieren und propagandieren – mit einem Wort, optimieren.

*Francis Knowles*

Die guten Leute, die sich über die technischen Fortschritte unserer Zeit so entsetzen, verwechseln, wie mir scheint, Zweck und Mittel. Wer nur um Gewinn kämpft, erntet nichts, was der Mühe wert ist. Aber die Maschine ist kein Ziel und darum ist auch das Flugzeug kein Zweck, sondern ein Werkzeug, ein Gerät, nichts anderes als der Pflug.

*Antoine de Saint-Exupéry*

Das elektronische Medium, welches so neu gar nicht mehr ist, hat die gesamte Publikationslandschaft tiefgreifend verändert. Es verheißt auch der Lexikografie eine Bandbreite an Möglichkeiten. Dabei kann der Computer jedoch nicht selbst aktiv werden, um diese Möglichkeiten zu nutzen; der Computer selbst kann nichts. Akteur bleibt der Mensch, der den Rechner als neues technisches Werkzeug zur Erreichung der von ihm definierten Handlungsziele einsetzen kann. Es gilt daher zu prüfen, wie der Einsatz des neuen Mediums so gestaltet werden kann, dass er nicht Selbstzweck, sondern sinnvoller Gebrauch dieses neuen Werkzeuges ist.

Das gilt auch für die Sprachlexikografie als eigenständige kulturelle Praxis, die hier Gegenstand der Betrachtung ist.<sup>1</sup> Hier wird das Potenzial des elektronischen Mediums für Wörterbücher besonders hoch eingeschätzt.

[...] It should not come as a surprise that nearly everyone involved in present-day dictionary making is enthusiastic about the potential of the electronic medium. ED [Electronic Dictionary, Anm. d. Verf.] dreams are indeed not without a solid basis. (Schryver 2003, S. 162)

---

<sup>1</sup> Vgl. Wiegand (1998a, S. 41). Näheres dazu siehe Kapitel 2.

Als Gründe werden – sehr allgemein formuliert – in der Forschung u.a. folgende Eigenschaften und Gebrauchsweisen von Wörterbüchern genannt: Wörterbücher werden nicht linear gelesen, sondern in ihnen wird punktuell nachgeschlagen, Wörterbuchartikel sind auf komplexe Weise miteinander vernetzt, was auf Papier oft nicht hinreichend umgesetzt werden kann, und Artikel von Sprachwörterbüchern sind in der Regel standardisiert in einzelne Positionen aufgeteilt, die meist nicht alle von einem Benutzer zur Beantwortung einer Suchanfrage zur Kenntnis genommen werden. In der Umsetzung dieser Eigenschaften wird dem elektronischen Medium ein potenzieller Mehrwert gegenüber dem Druckmedium zugesprochen.<sup>2</sup>

Neue Möglichkeiten hinsichtlich der Funktionalität werden in der Praxis jedoch nur langsam genutzt, wie es bei großen Medienveränderungen durch die Geschichte fast immer zu beobachten ist (vgl. Faulstich 1997). Sicherlich hat sich in den letzten Jahren in der lexikografischen Praxis viel verändert, sodass man nicht mehr unbedingt sagen kann, dass es sich bei den meisten elektronischen Wörterbüchern um eine Widerspiegelung der Printwörterbücher auf „reduziertem Niveau“ (Feldweg 1997, S. 110) handelt. Doch elektronische Wörterbücher ziehen gerade ihre Berechtigung daraus, sich sehr deutlich von gedruckten zu unterscheiden, indem sie die Möglichkeiten des neuen Mediums wirklich sinnvoll nutzen (vgl. auch Lehr 2000). Jedoch sind Beispiele für ganz grundsätzlich andere, neue Funktionalitäten elektronischer Wörterbücher gegenüber Printwörterbüchern, so wie sie z.B. Atkins fordert, nach wie vor kaum zu finden:

If new methods of access (breaking the iron grip of the alphabet) and a hypertext approach to the data stored in the dictionary do not result in a product light years away from the printed dictionary, then we are evading the responsibilities of our profession. (Atkins 1992, S. 521; zitiert nach Schryver 2003, S. 144)

Das neue Medium spielt in der Lexikografie jedoch nicht nur als Publikationsmedium eine Rolle. Ein wichtiger Bereich ist auch der Einsatz neuer Technologien in der Herstellung von Wörterbüchern, also im lexikogra-

---

<sup>2</sup> Einen sehr guten Überblick über das, was in der Literatur in Bezug auf die Möglichkeiten des elektronischen Mediums für die Lexikografie formuliert wird, bietet Schryver (2003). Aus der deutschsprachigen Wörterbuchforschung sind u.a. Storrer (2001) und die Beiträge aus Lemberg/Schröder/Storrer (Hg.) (2001) zu nennen. Zahlreiche Beispiele für Online-Wörterbücher sind über das Wörterbuchportal von Robert Beard und Paul J.J. Payack, [www.yourdictionary.com](http://www.yourdictionary.com) (Stand: Februar 2007), zu finden.

fischen Prozess. Dabei geht es nicht nur um die Herstellung von elektronischen Wörterbüchern, sondern genauso um die Herstellung von gedruckten Wörterbüchern oder von Wörterbüchern, die gleichzeitig in beiden Medien erscheinen. Ziel ist auch hier, im lexikografischen Prozess qualitative Verbesserungen durch den Einsatz des Computers zu erreichen, nicht nur quantitative Erleichterungen im Sinne einer schnelleren und reibungsloseren Herstellung. Doch können aus einfachen Quantitätssteigerungen nicht direkt Qualitätssteigerungen erwartet werden.

[...] Im computerunterstützten lexikografischen Prozess kann die Quantität niemals in Qualität (sozusagen 'von allein') umschlagen. Vielmehr kommt die Qualitätssteigerung im Beispiel dadurch zu Stande, dass ein aufgabengerechter und intelligenter Einsatz des Computers erfolgte, sodass die Datenverarbeitung zu einer Organisation des Materials führte, die eine bessere und zusammenhängende Wahrnehmung dieses Materials unter verschiedenen, für die Lösung der Aufgabe relevanten Perspektiven erlaubte, was eine wichtige Voraussetzung für eine adäquate Materialinterpretation durch die beiden Lexikografen darstellte. Es wäre wohl falsch zu folgern, die Qualitätssteigerung sei allein das Ergebnis der Materialinterpretation der beiden Lexikografen, auch wenn diese präzise Fragen stellten; vielmehr ist sie das Ergebnis einer gelungenen Mensch-Maschine-Kommunikation. (Wiegand 1998a, S. 160)

Das elektronische Medium bietet der Lexikografie also nicht schon an sich eine schöne neue Welt, sondern es muss etwas dafür getan werden, dass diese vielversprechenden Möglichkeiten die lexikografische Praxis verbessern und so zum Entstehen innovativer Produkte beitragen können. Dabei sind die beiden Bereiche – neue technische Möglichkeiten im lexikografischen Prozess und neues Potenzial bei der Publikation lexikografischer Daten – nicht losgelöst voneinander zu betrachten, sondern stehen in einem engen Verhältnis zueinander. Denn erst durch einen sinnvollen Einsatz des Computers im lexikografischen Prozess liegen die lexikografischen Daten in der Regel so vor, dass sie 'mediengerecht' im elektronischen Medium präsentiert werden können und – falls gewünscht – gleichzeitig auch ein gedrucktes Wörterbuch entstehen kann.

Es geht in dieser Arbeit daher allgemein um die Frage: Wie können neue technische Möglichkeiten so im lexikografischen Prozess eingesetzt werden, dass der Computer eine gute Unterstützung für die lexikografische Arbeit bietet und gleichzeitig eine Datenbasis entsteht, aus der neuartige elektronische Produkte entwickelt werden können? Eine Fragestellung ist in dieser Form jedoch noch ein zu 'weites Feld'. Ein wesentlicher Punkt im lexikogra-



fischen Prozess ist die Frage der Datenorganisation. Speziell die Aufgabe der Datenmodellierung ist, je nachdem, wie man Datenmodellierung definiert, als Aufgabe neu hinzugekommen bzw. hat sich grundlegend verändert. Meines Erachtens werden dabei in der Ausgestaltung der Datenmodellierung die entscheidenden Weichen dafür gestellt, ob die lexikografischen Daten am Ende so vorliegen, dass aus ihnen ein innovatives elektronisches Wörterbuch entwickelt werden kann und gleichzeitig eine gedruckte Präsentation dieser Daten grundsätzlich möglich ist. Darüber hinaus kann eine bestimmte Form von Datenmodellierung und entsprechender Datenerfassung wesentliche Qualitätsverbesserungen für die lexikografische Arbeitsumgebung bewirken. Allerdings ist bisher noch nicht umfassend aufgezeigt worden, wie eine solche Datenmodellierung und die darauf aufbauenden Verbesserungsmöglichkeiten in der lexikografischen Praxis aussehen können. Wie man eine solche Datenmodellierung entwickeln kann, welchen Richtlinien man dabei folgen sollte, welcher technischen Möglichkeiten man sich bei der Entwicklung bedienen kann, sind daher hier zu klärende Fragen. Das Thema dieser Arbeit ist somit darauf eingegrenzt, ein Konzept für die Modellierung einer lexikografischen Datenbasis zu entwickeln, die in benannten Bereichen Verbesserungsmaßnahmen bewirken kann, und exemplarisch aufzuzeigen, wie solche Verbesserungen aussehen können.

Diese allgemeine Themenbeschreibung soll nun genauer eingegrenzt und innerhalb der Lexikografie und Wörterbuchforschung verortet werden.

## 2. Zur Verortung und Eingrenzung des Themas dieser Arbeit

Im Laufe meiner Beschäftigung mit Wörterbüchern [...] bin ich allmählich zu der Überzeugung gelangt, dass eine möglichst weitgehende theoretische und methodologische Durchleuchtung einer althergebrachten [...], pragmatisch eingespielten Praxis zu einem vertieften Verständnis dieser Praxis, zu einem höheren Grad ihrer Beherrschung und damit in manchen Bereichen auch zu ihrer Erneuerung führen kann.

*Herbert Ernst Wiegand*

Gegenstand dieser Untersuchung ist der lexikografische Prozess, genauer die Datenmodellierung im lexikografischen Prozess. Dabei sind nicht beliebige Arten von lexikografischen Prozessen im Blickpunkt, sondern solche, die zu Sprachnachschatlagewerken führen. Unter diesen stehen wiederum die wissenschaftlichen Wörterbuchprojekte im Vordergrund, da hier eher der Raum für eine fundierte Konzeption bei der Modellierung der lexikografischen Daten gegeben ist. Für die genaue Erläuterung des Aufbaus dieser Arbeit ist es hilfreich, dieses Thema innerhalb der Lexikografie und Wörterbuchforschung zu verorten. Dazu sollen zunächst – anschließend an Wiegand – Sprachlexikografie und Wörterbuchforschung gegeneinander abgegrenzt werden.

Die wissenschaftliche Form der Sprachlexikografie kann mit Wiegand als eine „eigenständige kulturelle und wissenschaftliche Praxis“ (Wiegand 1998a, S. 62) charakterisiert werden. Sprachlexikografie als solche ist danach keine Wissenschaft, da „das Ziel der Sprachlexikographie [...] nicht darin“ besteht, „wissenschaftliche Theorien zu erarbeiten“ (ebd., S. 23). Natürlich können Lexikografen auch sprachwissenschaftliche Tätigkeiten ausführen. Dazu bringt Wiegand das Beispiel der Etymologie. Wenn ein etymologisches Wörterbuch erarbeitet würde, dann sei die „Erarbeitung einer Etymologie und damit die von etymologischem Wissen eine *sprachwissenschaftliche* Tätigkeit“, die damit „auch eine Tätigkeit eines bestimmten Lexikographen sein“ kann.

Die lexikographische Tätigkeit besteht u.a. in einer Verarbeitung des verfügbaren etymologischen Wissens zu einem lexikographischen Text. Die Erarbeitung einer Etymologie einerseits und deren Darstellung in einem Wörterbuchartikel andererseits sind (auch wenn sie von einer Person ausgeführt werden und ineinander greifen können) zwei verschiedene und deutlich unterscheidbare Dinge. (Wiegand 1998a, S. 23)

In der wissenschaftlichen Sprachlexikografie werden „in unterschiedlichem Ausmaß wissenschaftliche Methoden aus *verschiedenen* Wissenschaften angewendet und wissenschaftliche Ergebnisse aus *verschiedenen* Wissenschaften genutzt“ (Wiegand 1998a, S. 40). Nicht nur Ergebnisse z.B. aus der Lexikologie können demnach angewandt werden, sondern wie später zu sehen sein wird, auch Ergebnisse beispielsweise aus dem Bereich der Texttechnologie. Von lexikografischen Tätigkeiten deutlich zu unterscheiden sind aber solche

Tätigkeiten, die nicht als Teil eines lexikographischen Prozesses darauf ausgerichtet sind, daß Sprachnachschlagewerke entstehen, sodaß sie nicht mit dem Schreiben von Wörterbüchern in einem erkennbaren Handlungszusammenhang stehen (ebd., S. 45).

Diese Tätigkeiten solle „man nicht Lexikografie nennen“, da man „Wörterbücher nicht theoretisch schreiben“ kann (ebd., S. 45). Selbstverständlich enthält lexikografische Tätigkeit reflexive Komponenten, die aber immer noch zur praktischen Tätigkeit der Sprachlexikografie gezählt werden sollen (ebd., S. 46).

Die verschiedenen wissenschaftlichen Formen der Beschäftigung mit Sprachlexikografie sind daher nicht selbst Teil der Sprachlexikografie, sondern bilden den wissenschaftlichen Metabereich hierzu (Wiegand 1998a, S. 64). Die Sprachlexikografie ist demnach der „Gegenstandsbereich“, „aus dem für die verschiedenen *Forschungsgebiete* der Wörterbuchforschung die wissenschaftlichen Gegenstände gewonnen werden“ (ebd., S. 46). Ein Bereich der Wörterbuchforschung ist dabei auch der Einsatz des Computers im lexikografischen Prozess. „In diesem Sinne kann die Wörterbuchforschung als die Gesamtheit aller, im Schrifttum nachweisbaren, wissenschaftlichen Bemühungen aufgefaßt werden, die darauf abzielen, diejenigen theoretischen, methodischen, terminologischen, historischen, dokumentarischen, didaktischen und kulturpädagogischen Fragen zu beantworten, die sich stellen“ u.a. „beim Einsatz des Computers in der Lexikographie und Wörterbuchforschung“ (ebd., S. 79). Sinn und Ziel dieser Unterteilung von Sprachlexikografie und Wörterbuchforschung ist bei Wiegand nicht die „unfruchtbare Aufteilung in Praktiker und Theoretiker“, sondern „die Unterscheidung zweier Arten von Tätigkeiten“ (ebd., S. 88). Die Abgrenzung dieser beiden Arten von Tätigkeiten nach Wiegand ist deshalb referiert, um das hier behandelte Thema den verschiedenen Bereichen zuordnen zu können.

Einen großen Teil dieser Arbeit macht die Entwicklung eines Konzepts für die Modellierung einer lexikografischen Datenbasis aus, die nicht auf einen bestimmten lexikografischen Prozess bezogen ist, sondern auf möglichst viele konkrete lexikografische Prozesse anwendbar sein sollte. Dabei soll das Konzept so theoretisch fundiert sein, dass es nachvollziehbar wird. Die Entwicklung dieses Konzepts für eine Datenmodellierung ist daher der Wörterbuchforschung zuzurechnen, da sie nicht mit dem Ziel verbunden ist, ein konkretes Wörterbuch zu erarbeiten. Insofern ist die Arbeit auch als „Beitrag zur Neuordnung lexikographischer Theoriebildung unter Bezugnahme auf die Möglichkeiten des Rechnereinsatzes“ (Geeb 2001, S. 56) zu sehen. Da die Intention, die hinter der Entwicklung dieses Konzepts steht, die Verbesserung der lexikografischen Praxis ist, wird die Umsetzung dieses Konzepts an einem konkreten sprachlexikografischen Projekt illustriert. Diese Entwicklung einer Datenmodellierung in einem konkreten lexikografischen Prozess kann der Sprachlexikografie zugeordnet werden, denn sie ergänzt und ersetzt teilweise die früheren Instruktionbücher und Arbeitsanweisungen. Zu Letzteren bemerkt Wiegand:

Die Ergebnisse der lexikographischen Tätigkeiten werden zum lexikografischen Prozeß gerechnet, und zwar nicht nur die Endprodukte, die Wörterbücher, sondern auch – und dies gilt speziell für die wissenschaftliche Lexikographie – alle Teil- oder Zwischenergebnisse wie Arbeitsanweisungen, Instruktionen, Protokolle, Konzeptionspapiere, Probeartikel, Musterartikel, Korrekturunterlagen, Werkstattberichte usw., also alle jene Schriftstücke, in denen auch die selbstreflexive Komponente der Lexikographie als einer eigenständigen, kulturellen Praxis konkret faßbar wird. (Wiegand 1998a, S. 77)

Nun ist z.B. die Modellierung einer Datenbank nicht ein Schriftstück wie eine Arbeitsanweisung, aber sie ist Teil eines konkreten lexikografischen Prozesses; ihre Entwicklung ist daher als lexikografische Tätigkeit aufzufassen. Der Werkstattbericht im sechsten Kapitel dieser Arbeit ist damit der Sprachlexikografie zuzuordnen, da es sich hier um den Bericht über Ausschnitte aus einem konkreten lexikografischen Prozess handelt.

Das Thema dieser Arbeit ist damit sowohl ein metalexikografisches als auch ein lexikografisches: Einerseits die Entwicklung eines Konzepts für die Modellierung einer lexikografischen Datenbasis; andererseits die Anwendung dieses Konzeptes an einem konkreten Fallbeispiel. Das Modellierungskonzept soll dabei dazu dienen,

- die Lexikografen im lexikografischen Prozess, d.h. bei der Datenerarbeitung und Dateneingabe zu unterstützen,
- die so erarbeiteten lexikografischen Daten in innovativer Weise für Wörterbuchbenutzer im elektronischen Medium zugreifbar zu machen, und
- den Weg offen zu halten, die so erarbeitete lexikografische Datenbasis in gedruckter Form präsentieren zu können.

Um das Konzept zu entwickeln, werden zunächst im folgenden Kapitel einige terminologische und ordnende Überlegungen angestellt, die den Themenbereich dieser Arbeit betreffen, jedoch auch über ihn hinausweisen (vgl. auch Müller-Spitzer 2003). Sie sind deshalb notwendig, um eingrenzen und benennen zu können, welche Arten von elektronischen Nachschlagewerken hier Gegenstand der Betrachtung sind, um welche Arten von lexikografischen Prozessen es geht und welche Ebenen im lexikografischen Prozess im Rahmen dieser Arbeit unterschieden werden. Im vierten Kapitel werden in einem Forschungsüberblick Anforderungen an die Modellierung lexikografischer Daten aufgezeigt. Dieser Überblick soll zusätzlich zu den oben formulierten grundsätzlichen Zielen der Modellierung einer lexikografischen Datenbasis sehr viel detaillierter zusammentragen, was in diesem Zusammenhang von (Meta-)Lexikografen gefordert wird und welchen Ansprüchen damit das in dieser Arbeit entwickelte Konzept genügen muss. Im darauf folgenden Hauptkapitel 5 werden die Richtlinien für die XML-basierte Modellierung einer lexikografischen Datenbasis entwickelt. Wie dabei genau vorgegangen wird und wie die Gliederung des Hauptkapitels motiviert ist, wird am Anfang des Kapitels erläutert. In Kapitel 6 wird in einem Werkstattbericht exemplarisch gezeigt, wie dieses Konzept im Projekt *ellexiko* des Instituts für Deutsche Sprache<sup>3</sup> umgesetzt worden ist.

Ziel dieser Arbeit ist es damit, bekannte Forschungen aus dem Bereich der Printlexikografie mit neuen Verfahren und Techniken aus dem Bereich der XML-Modellierung zu einem eigenständigen, praxisorientierten und theoretisch fundierten Ansatz für ein Konzept der maßgeschneiderten Modellierung einer lexikografischen Datenbasis zusammenzubringen und dabei die Relevanz und Auswirkungen eines solchen Ansatzes für die und anhand der lexikografischen Praxis aufzuzeigen. Damit eng verbunden ist das Ziel,

---

<sup>3</sup> Für Projektinformationen siehe [www.ellexiko.de](http://www.ellexiko.de) (Stand: Februar 2007).

Wege aufzuzeigen, wie bessere elektronische Wörterbücher hergestellt werden können, mit der zusätzlichen Möglichkeit, aus der dafür erarbeiteten Datenbasis auch ein gedrucktes Wörterbuch herzustellen. Die Arbeit bietet damit keine Analyse des Bestehenden, sondern entwickelt einen Leitfaden für eine Modellierung, die zu einer verbesserten lexikografischen Praxis unter Ausnutzung der heutigen technischen und medialen Bedingungen führen soll.