

Ulrike Spree

## D 6 Wörterbücher und Enzyklopädien

### D 6.1 Wörterbücher und Enzyklopädien in der Informationspraxis

In einem Wörterbuch schlägt man nach, wenn man etwas nicht weiß – in der Erwartung punktuelle geprüfte Information zu erhalten. Darüber hinaus werden lexikographische Darstellungsformen jedoch nicht nur als informative Gebrauchstexte, sondern auch in der Belletristik aufgegriffen (Lit. 01).

In der bibliotheks- und informationswissenschaftlichen Literatur werden Wörterbücher und Enzyklopädien im Rahmen des Auskunftsdienstes und der in diesem Zusammenhang vorgestellten Informationsmittel und Nachschlagewerke behandelt (Lit. 02). Ein umfassender systematischer Überblick ist hier nicht angestrebt. Selbst die zahlreichen gedruckten und elektronischen Bibliographien vermögen nur noch Teilbereiche abzudecken (Lit. 03, Lit. 04, Lit. 05). In diesem Beitrag werden exemplarisch wesentliche Strukturmerkmale von Wörterbüchern und Enzyklopädiën dargestellt und die wissensorganisatorische Leistung von Wörterbüchern und Enzyklopädiën als eine seit Jahrhunderten praktizierte Form der strukturierten Informations- und Wissensaufbereitung herausgearbeitet sowie Trends der Entwicklung des lexikographischen Marktes aufgezeigt.

#### D 6.1.1 Begriffsklärung Lexikon, Wörterbuch, Enzyklopädie

Die terminologisch klare Differenzierung von Lexika und Wörterbüchern als Nachschlagewerke für sprachliche Information und Enzyklopädiën als solche, die Begriffe (Gegenstände, Sachverhalte, Dinge, Ereignisse) erklären, verwischt sich in der Praxis, da Wörterbücher aufgrund des Verweischarakters der Sprache stets auch Sachinformationen enthalten, und Enzyklopädiën auch linguistische Informationen beinhalten. Die besondere Bedeutung von Wörterbüchern und Enzyklopädiën im Zusammenhang mit Informationserschließung und -vermittlung ergibt sich aufgrund dreier gemeinsamer Charakteristika:

1. Die strukturierte Anordnung der Einträge nach einem vorgegebenen für die Zielgruppen offensichtlichen Ordnungsprinzip, die vor allem darauf abzielt, den Gebrauch als Nachschlagewerk zu erleichtern (vgl. Lit. 02).
2. Hieraus ergibt sich die atomistische Betrachtungsweise der Inhalte, die jeden Eintrag einzeln betrachtet. Wörterbücher und Enzyklopädiën sind in der Regel nicht zur linearen Lektüre vorgesehen.
3. Wörterbücher und Enzyklopädiën orientieren sich vorrangig daran, Zweifel des Nutzers in Hinblick auf den Gebrauch einzelner Worte oder Begriffe aufzulösen.

#### D 6.1.2 Historischer Rückblick: Lexika und Lexikonherstellung

Wörterbücher und Enzyklopädiën sind auch aus historischer Perspektive als pädagogisch-didaktische Instrumente der Wissensvermittlung und „Gedächtnisstützen“ für die Lernenden (vgl. Lit. 06) zu verstehen.

Vorläufer einer systematischen Darstellung des Wissens finden sich in allen antiken Hochkulturen. Der Begriff Enzyklopädie lässt sich auf das griechische *enkýklios paideia* (ἐγκύκλιος παιδεία) zurückführen, was übersetzt in etwa „Kreis des Wissens“ bedeutet und mit unserem Begriff der „allgemeinen Bildung“ korrespondiert (vgl. Lit. 06, S. 39). Bereits aus dem Mittelalter kennen wir Enzyklopädiën, Wörterbücher und glossarähnliche Wortlisten (vgl. Lit. 06). Als Spiegel des zeitgenössischen Wissensstandes erlebten Universallexika ab dem 18. Jahrhundert eine erneute Blüte. Re-

präsentatives Beispiel für Deutschland ist Zedlers „Universallexikon“ (1732-1754). Die in den Jahren 1751 bis 1780 von Denis Diderot und Jean le Rond d'Alembert herausgegebene „Encyclopédie“ ist ein den Gedanken der Aufklärung verpflichtetes Nachschlagewerk neuen Typs, dem als eine auch politische Waffe eine wichtige Rolle in der gedanklichen Vorbereitung der französischen Revolution zugeschrieben wird (vgl. Lit. 07). Parallel zur Entwicklung der Nationalstaaten und der Vereinheitlichung der Nationalsprachen seit dem Ende des 18. Jahrhunderts wurden große Wörterbuchprojekte ins Leben gerufen (Lit. 08).

Im Verlauf des 19. und 20. Jahrhunderts haben sich mehrere grundlegende Wandlungen im Selbstverständnis der Lexikonhersteller vollzogen. Lexika sollten Stichwortgeber für die gebildete Konversation (Konversationslexikon) sein. Auf die Verdoppelung des Fachwissens innerhalb eines Fachgebiets alle drei bis zehn Jahre reagierten die Lexikonverlage mit der Ausdifferenzierung der Gattung in unterschiedliche Typen. Auf der einen Seite standen die umfangreichen Enzyklopädien wie die britische „Encyclopaedia Britannica“. Solche Werke wiesen bis zu 2000 Mitarbeiter auf, darunter renommierte Fachwissenschaftler, und die Artikel sind meist signiert. Der Lexikoneintrag stellt seinen Gegenstand in größere Zusammenhänge, wodurch die Artikel auf einen beträchtlichen Umfang anwachsen können. In der Regel wird weiterführende Literatur angegeben. Auf der anderen Seite erfolgte ab 1860 eine Reorientierung an der Idee des Universallexikons mit allgemeinverständlicher Darstellung und einer hohen Stichwortzahl, die einen gezielten Zugriff auf Einzelinformationen erlaubt. Für die Inhalte der unsignierten, alphabetisch nach engem Schlagwort zusammengestellten Artikel zeichnete eine Lexikonredaktion verantwortlich.

## D 6.2 Nachschlagewerke als spezielle Formen des organisierten, strukturierten Wissens

Wesentliche Leistungen der Lexikographie, der wissenschaftlichen Beschäftigung mit Fragen der Wörterbuchentwicklung, liegen in der Entwicklung von Wörterbuchtypologien und der exakten Beschreibung lexikographischer Strukturen (vgl. Lit. 09). Die Überschneidungen und Berührungspunkte mit anderen Formen der Wissensorganisation wie Klassifikationen, Thesauri und Ontologien sind vielfältig. Anders als die genannten Ordnungssysteme und Dokumentationssprachen bieten konkrete lexikographische Nachschlagewerke jedoch stets auch den Durchgriff auf Fakteninformation, sei es linguistischer oder enzyklopädischer Art.

### D 6.2.1 Typologie lexikographischer Nachschlagewerke

Wörterbuchtypologien werden nach inhaltlichen Kriterien, aus Produktionssicht, Anwendersicht, Datensicht sowie aus der Sicht des Lebenszyklus lexikographischer Nachschlagewerke erstellt (vgl. Lit. 10). *Inhaltliche* Unterscheidungskriterien sind z. B. Anzahl der Sprachen, thematische Spezialisierung, Beziehung der Lemmata zueinander (z. B. Synonyme) oder verschiedene Informationstypen (Redewendungen, Sprichwörter) oder Sprachvariationen. Bekannte *formale Kriterien* zur Wörterbuchtypologisierung sind z. B. Publikationsform (gedruckt: Buch, Loseblattsammlung; digital: offline, online), Verweismechanismen (Hypertextualisierung) oder Formen des Zugriffs.

### D 6.2.2 Strukturierung lexikographischer Nachschlagewerke

Ungeachtet der Formenvielfalt, die sich im Zuge der (jahrhundertealten) lexikographischen Praxis ergeben hat, lässt sich eine Reihe immer wiederkehrender typischer Strukturelemente benennen. Wesentliche Strukturierungsleistungen liegen in der *Lemmatisierung*, der *Zugriffsstruktur*, *Rahmenstruktur*, *Verteilungsstruktur*, *Makrostruktur*, *Mikrostruktur* sowie der *Verweisstruktur* (vgl. Lit. 11).

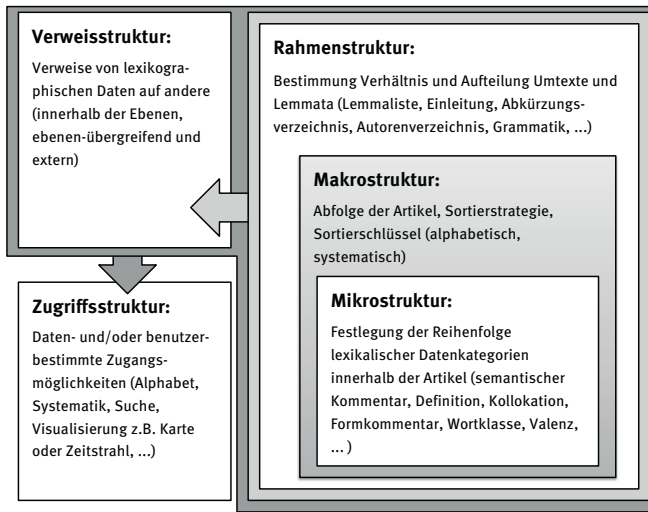


Abb. 1: Strukturelemente

Der Hauptteil eines lexikographischen Nachschlagewerks besteht aus Artikeln. Jeder Artikel besteht aus mindestens einem *Lemma* (Stichwort; gr. Empfangenes) und lexikographischen Daten zu diesen Lemmata. Im Zuge der Lemmatisierung werden die verschiedenen Wortformen zu einem *Lexem* zusammengefasst und das Lexem als Lemma in die Lemmaliste aufgenommen. Die Auswahl der Lemmata für ein lexikographisches Nachschlagewerk erfolgt i. d. R. nach folgenden Kriterien:

- Auswertung bestehender Wortlisten aus lexikographischen Nachschlagewerken.
- Persönliches (Fach-)Wissen des/der LexikographInnen und/oder hinzugezogener ExpertInnen. In der Online-Enzyklopädie Wikipedia werden Lücken in der Lemmaliste von den Mitgliedern kollaborativ in einer Liste Artikelwünsche gepflegt.
- Auswertung von Fachliteratur/-zeitschriften (einschl. ihrer Inhaltsangaben und Register).
- Korpusbasierte Lemmatisierung. Hier werden die Lexeme aus einem Textkorpus gefiltert und erlauben so den Rückgriff auf die Originaltexte (vgl. Digitales Wörterbuch der Deutschen Sprache – DWDS; Lit. 12).

**Die Rahmenstruktur** (Wörterbuchttextstruktur, Megastruktur) beschreibt die übergeordneten Elemente eines lexikographischen Nachschlagewerks in ihrem Zusammenspiel: *Umtexte* und *Lemmaliste(n)*. Umtexte ist eine Sammelbezeichnung für alle Teile eines lexikographischen Nachschlagewerks, die nicht Teil der Lemmaliste(n) sind: *Vorwort*, *Benutzerführung*, *Einleitung*, *Fachliche Einführung*, *Grammatik*, *Register* und *Anhänge* sowie ausgelagerte längere und in Bezug auf die lemmatisierten Lexeme übergreifende (Fach-)Artikel. In elektronischen lexikographischen Nachschlagewerken kann bei Rahmentexten die in gedruckter Form zwingende lineare Abfolge der Elemente aufgebrochen werden und Elemente der Benutzerführung wie Abkürzungen oder grammatikalische Informationen können kontextsensitiv angeboten bzw. aufgelöst werden.

Durch **die Verteilungsstruktur** wird die Anordnung der verschiedenen Informationen (Definition, Grammatik, Literaturlisten etc.) in den einzelnen Artikel sowie in den Umtexten bestimmt. Auch in der Gestaltung der Verteilungsstruktur sind Online-Lexika flexibler. So können z. B. grammatikalische Informationen oder Literaturangaben sowohl kontextsensitiv als auch in einem separaten Teil platziert werden.

**Die Makrostruktur** (Lit. 13) beschreibt die Abfolge, die Sortierstrategie sowie den primären Sortierschlüssel der Lemmata in einem lexikographischen Nachschlagewerk. Grundsätzlich unterscheidet man zwischen alphabetischen Makrostrukturen und systematischen Makrostrukturen.

	<b>alphabetische Makrostruktur</b>	<b>systematische Makrostruktur</b>
Auffindbarkeit für den Laien	Leichte Auffindbarkeit für den Laien, der kein System erlernen muss. Das Alphabet ist ein eindeutiges Ordnungssystem.	Der Benutzer muss die Struktur kennen. Klassifikationen sind mehrdeutige Ordnungssysteme. Ein falscher Einstieg kann dazu führen, dass die gesuchte Information nicht aufgefunden wird. Informationen sind auffindbar, auch wenn die genaue Benennung (Fachterminologie) unbekannt ist.
Eindeutigkeit der Zuordnung	Einzelne Aspekte eines Gegenstandes können auf verschiedene im Alphabet verteilte Lemmata zerstreut sein.	Ein Sachverhalt kann aus unterschiedlichen Perspektiven betrachtet werden. Um multiple Perspektiven abzubilden, muss es möglich sein, dieselben Sachverhalte polyhierarchisch zuzuordnen, was dann wieder zulasten der Eindeutigkeit der Zuordnung geschehen muss.
Erweiterbarkeit	Neues Wissen wird zu einem Einstiegslexem zusammengefasst und kann jederzeit ergänzt werden.	Systematische Ordnungen sind starr und können nicht mit dem Entwicklungstempo der Wissenschaften mithalten. Umstrukturierungen eines Ordnungssystems, sofern die logische Struktur betroffen ist, sind aufwändig.
Zusammenhang	Inhalte werden aus einem größeren Sachzusammenhang gerissen. Zusammengehöriges wird getrennt und Unzusammenhängendes wird in einen zufälligen Zusammenhang gestellt.	Artikel werden in einen inhaltlichen Zusammenhang gestellt. Die Einordnung in eine Systematik liefert zusätzliche Sachinformationen z. B. zum Grad der Spezifität eines Artikels.
Fachkenntnisse	Der Benutzer muss das Stichwort (Einstiegslemma) kennen. Mehrdeutige Lemmata können über Verweise disambiguiert werden.	Information kann auch gefunden werden, wenn die Terminologie nicht genau bekannt ist.
Systematische Einordnung	Entstehung von Redundanzen, da die Stellung innerhalb des Ordnungssystems für jeden Artikel individuell dargelegt werden muss.	Die Stellung in einem Ordnungssystem liefert Zusatzkenntnisse zur Sache.
Überblick und Hinweis auf Wissenslücken	Die alphabetische Makrostruktur liefert keinen Hinweis auf Lücken in der Lemmaliste.	Generisches Browsen ermöglicht den Überblick über die Domäne und die Identifikation von Wissenslücken.
Übersetzbarkeit in andere Sprachen	Schwierig zu übersetzen, da an Sprache gebunden, in der die Lemmatisierung erfolgt	Gute Übersetzbarkeit bei kulturunabhängigen Fachgebieten wie Technik, da gewisse Sprachunabhängigkeit gegeben. Bei kulturabhängigen wie z. B. Jura ist das auch mit einer Systematik sehr schwer, wenn auch leichter als mit einer alphabetischen Ordnung.

**Tab. 1:** Leistung alphabetischer und systematischer Makro- und Zugriffsstrukturen (in Anlehnung an Lit. 14)

In der Praxis findet man seit jeher Mischformen der systematischen und alphabetischen Makrostruktur: Alphabetische Indices und Querverweise erleichtern den punktuellen Zugriff auf systematische Makrostrukturen und systematische Übersichtstafeln und Querverweise dienen der inhaltlichen Ergänzung der alphabetischen Makrostruktur.

In datenbank- oder graphbasierten Online-Lexika lassen sich alphabetische und systematische Makrostruktur flexibler kombinieren als in gedruckten Lexika. In dem Projekt Dbpedia werden die Inhalte der Wikipedia zudem durch den Einsatz von Linked Data semantisch erschlossen. So bietet z. B. Faceted Wikipedia Search die Möglichkeit der facettierten Suche auch für einen menschlichen Nutzer. Sie erlaubt komplexe Suchanfragen wie nach deutschen Wissenschaftlern, die zwischen 1900 und 1970 in Lübeck geboren wurden.

Zur Unterstützung der intuitiven Nutzbarkeit werden die Artikel in einem lexikographischen Nachschlagewerk nach demselben zuvor festgelegten Schema (*Mikrostruktur*) aufgebaut (vgl. Lit. 15). Man unterscheidet zwischen *Formkommentar* und *semantischem Kommentar*. Zum Formkommentar gehören Angaben zu *Abkürzung, Aussprache, Rechtschreibung, Formvarianten, grammatische Angaben wie Genus und Beugungsformen, Wortart, Morphologie, Trennung* usw. Auf den Formkommentar folgt i. d. R. der semantische Kommentar und darin bei mehrsprachigen Wörterbüchern auch die Übersetzungen der Lemmata. Im semantischen Kommentar finden sich darüber hinaus Angaben zu *Antonymen, Synonymen, Bedeutungserklärung, Belegen, Fachgebieten, Wortverbindungen (Kollokationen u. a.)* etc.

Jeder Artikel eines lexikographischen Nachschlagewerks hat eine *konkrete Mikrostruktur*. Verschiedene Artikel innerhalb eines Nachschlagewerkes können unterschiedliche konkrete Mikrostrukturen haben. Auf der Grundlage der konkreten Mikrostrukturen lässt sich die *abstrakte Mikrostruktur* als Zusammenfassung aller im lexikographischen Nachschlagewerk vertretenen Mikrostrukturen ableiten. Die abstrakte hierarchische Mikrostruktur bildet die Voraussetzung für ein konsistentes Datenmodell, das die Umsetzung xml- oder rdf-basierter Lexika z. B. im Semantic Web ermöglicht.

**Die Verweisstruktur** (Lit. 16) enthält die Elemente, die von lexikographischen Daten auf andere (lexikographische) Daten verweisen. Man unterscheidet zwischen Verweisen innerhalb eines Artikels, zwischen den Artikeln eines Nachschlagewerkes untereinander und externen Verweisen. Verweise können *explizit* („siehe“) oder *implizit* sein.

**Die Zugriffsstruktur** (Lit. 11) beschreibt die Merkmale der Schritte, die ein Benutzer bis zum Auffinden der gewünschten Informationen beschreiten muss. Während bewährte Formen der Wissensorganisation auch weiterhin Bestand haben, hat sich das Spektrum von zusätzlichen Zugriffsmöglichkeiten durch die Digitalisierung und die damit möglich gewordenen graphischen Benutzeroberflächen erheblich erweitert.

Durch die Nutzung von *Hypertext* ist eine vielfältige, auch kontextsensitive Einbindung der Umtexte möglich, Browsing- und Stöberfunktionen können ausgebaut werden. Der gezielte Einsatz von *Multimedia* und *Visualisierungen* wie Mindmaps kann die Anschaulichkeit der Beiträge erhöhen und ermöglicht die Schaffung zusätzlicher Access Points. Durch Nutzung *interaktiver Funktionen* kann z. B. dem Nutzer ermöglicht werden eigene Einträge zu ergänzen, oder Nachschlagewerke können durch Foren ergänzt werden. Letztendlich zielen die aufgeführten Ansätze auf ein Verständnis von lexikalischen Einträgen als Einheiten, die dynamisch vom Nutzer im Prozess des Nachschlagens generiert werden (vgl. Lit. 17). Eine besondere Herausforderung solcher Ansätze besteht in der Entwicklung gebrauchstauglicher interaktiver Nutzerschnittstellen.

### D 6.2.3 Lexikonbenutzer und -nutzung

Um einem Benutzer über ein lexikographisches Nachschlagewerk genau die Informationen zu geben, die er zu einem gegebenen Zeitpunkt in einer gegebenen Situation benötigt, ist eine genaue Analyse der Nutzung und der Nutzer notwendig (vgl. Lit. 18, Lit. 19). Die Forschung differenziert

zwischen *Benutzervoraussetzung* (z. B. Laie, Spezialist), *Benutzersituation* (z. B. Arbeitskontext, Hobby) und *Benutzerintention* (z. B. Textrezeption, Textproduktion). Das Digitale Wörterbuch der Deutschen Sprache des 20. Jahrhunderts (DWDS) erlaubt z. B. den Einstieg über verschiedene Sichten für den Laien (LEO-Sicht) und den Wissenschaftler (z. B. Statistiksicht).

Theoretische Überlegungen zur Benutzerstruktur sind im Umfeld der Konzeption von Enzyklopädien eher ein Desiderat. Die Online-Enzyklopädie Wikipedia beispielsweise macht in den Allgemeinen Informationen über Wikipedia zwar Aussagen über die Autoren aber nicht über die intendierte Zielgruppe (vgl. Lit. 20). Kenntnisse über die tatsächliche Nutzung von gedruckten Wörterbüchern und Enzyklopädiën basieren auf der Auswertung von Verlags- und Herausgeberdaten wie Auflagenhöhen oder Subskriptionslisten und Leserbriefe sowie anekdotischen Leseerfahrungen (vgl. Lit. 08). Bei Online-Enzyklopädiën lässt sich über Webstatistiken und User-Tracking-Verfahren ein genaueres Bild der tatsächlichen Nutzung gewinnen. Die Nutzung der Wikipedia beispielsweise ist zudem über Nutzerbefragungen recht gut untersucht. Wikipedia ist zu einem Alltagsmedium geworden und wird weltweit schwerpunktmäßig von Personen mit höherem Bildungsstand von allen Altersgruppen relativ gleichmäßig genutzt, wobei die Nutzung seit 2007 stark gestiegen ist (vgl. Lit. 21, Lit. 22). Typische Nutzungssituation ist der schnelle Überblick zu einem Thema (vgl. Lit. 23). Viele Journalisten nutzen Wikipedia mittlerweile offen als Ressource (vgl. Lit. 24). Die explizit Wikipedia-kritische-Einstellung von Lehrern und Hochschullehrern macht zunehmend einem Plädoyer für die Förderung von Informationskompetenz, nicht zuletzt bei Schülern und Studierenden, und einem grundsätzlich kritischen Umgang mit Informationsressourcen Platz (vgl. Lit. 25).

### D 6.3 Lexikonproduktion

Die Bedingungen, unter denen lexikographische Nachschlagewerke produziert wurden und werden, sind äußerst vielfältig. Ungeachtet dieser Vielfalt sehen sich die Produzenten lexikographischer Nachschlagewerke mit einer Reihe von allgemeinen Problemstellungen konfrontiert.

1. Grundsätzlich gilt es Entscheidungen über die Auswahl und die Eingrenzung des Stoffes zu treffen. Prinzipiell universell angelegte Projekte wie Wikipedia stehen thematisch eingegrenzten Unternehmungen gegenüber.
2. Es gilt zu klären, wie und ob ein Zusammenhang des Wissens über Systematiken, Register oder Verweise hergestellt werden kann.
3. Lexika unterscheiden sich in der Darstellungstiefe, die von der Faktendarstellung über die Erklärung detaillierter Kausalzusammenhänge bis hin zu groben Überblicksdarstellungen schwanken kann.
4. Jedes Lexikon muss sich der Frage des Umgangs mit Parteilichkeit stellen. Lexika können bewusst Partei ergreifen, einen pluralistischen Ansatz verfolgen und verschiedene Stimmen zu Wort kommen lassen oder eine „neutrale“ Haltung anstreben.
5. Da der Anspruch auf gesicherte Informationen konstitutiv für das Selbstverständnis von Nachschlagewerken ist, gilt es die jeweils angewandten Verfahren der Verifikation zu prüfen, die ebenfalls historischen Wandlungen unterliegen.

#### D 6.3.1 Kooperation und Kollaboration

Hinter jeder Enzyklopädie steht ein langer Prozess der Aushandlung darüber, was zu einer bestimmten Zeit in einem bestimmten Kontext als wissenswert gelten soll (vgl. Lit. 26, Lit. 27). Oft genug, vor allem in der ersten Hälfte des 20. Jahrhunderts, wurden lexigraphische Nachschlagewerke und Enzyklopädiën gezielt als Mittel zur Ausübung kultureller und ideologischer Macht funktionalisiert (vgl. Lit. 28). Bei der täglichen Arbeit in den komplexen Redaktionen mit teilweise hunderten

von Mitarbeitern sind jedoch häufig weniger Konflikte über große weltanschauliche Fragen zu bewältigen als Details der redaktionellen Bearbeitung.

Das interne Qualitätsmanagement basiert auf einer Kombination von Redaktionsstrukturen (z. B. Fachredaktionen), Bearbeitungsrichtlinien und automatischen Tools (z. B. Entdecken von Edit Wars oder Vandalismus). Dennoch stößt das Wikipedia-Modell der kollaborativen Zusammenarbeit auch an seine Grenzen:

- Beachtliche Qualitätsunterschiede werden in Kauf genommen. Artikel, insbesondere zu begrifflich komplexen Fragestellungen, laufen Gefahr, im Laufe des langwierigen Diskussionsprozesses zu zerfasern.
- Ungeachtet der großen Zahl angemeldeter Nutzer werden Entscheidungen von relativ wenigen Mitarbeitern gefällt.
- Technische Hilfsmittel können helfen, auffällige Datenströme zu identifizieren, Veränderungen (verstärkter Zugriff) und formale Fehler rasch zu entdecken. Eine akribische intellektuelle Faktenüberprüfung findet nur im Einzelfall statt, vor allem wenn eine Behauptung formal durch eine verlässliche Quellenangabe informationell abgesichert zu sein scheint (vgl. Lit. 29).

### D 6.3.2 Technische Umsetzung

Alle Bearbeitungsschritte eines lexikographischen Nachschlagewerkes, ungeachtet seiner Publikation als Onlinelexikon oder als gedrucktes Lexikon, sind heute computergestützt (vgl. Lit. 30). Die wichtigsten Phasen des lexikographischen Erarbeitungsprozesses sind:

1. Vorbereitung: Konzeptentwicklung, Organisationsstruktur, Finanzplan
2. Datenakquise: Primär- und Sekundärquellen, Korpora, Illustrationen usw.
3. Computerisierung: Tagging und Annotation von Korpustexten, Suchtools, Schreibprogramme auswählen
4. Datenstrukturierung: Lemmatisierung: Datengrundlage in Bezug auf Benutzer/Fach definieren und erstellen; Lemmaauswahl in Bezug auf Benutzer und beteiligte Fachgebiete durchführen; Definition des Projekts anhand der Strukturbegriffe: Rahmenstruktur, Verteilungssystem, Makrostruktur, Mikrostruktur, Verweisstruktur, Zugriffsstruktur
5. Datenanalyse und Verfassen der Einträge: Automatische Gewinnung von Daten aus Korpora (Kollokationen, Häufigkeiten), Verfassen der Artikel, Verlinkung
6. Qualitätskontrolle: Korrektur lesen, Usability Evaluation, Definition des Lebenszyklus der Daten und ihrer Revalidierung/Löschung

Aus datenstruktureller Sicht und im Hinblick auf die Modellierung der Daten sind Datenbanklösungen, der Einsatz von Auszeichnungssprachen (vor allem *XML*, *TEI*) und zunehmend auch graph-basierte Ansätze zu unterscheiden (vgl. Lit. 31). Onlinelexika haben gegenüber gedruckten Lexika den Vorteil, dass Interface und Datengrundlage deutlich getrennt werden können. Bei entsprechender Kennzeichnung mit einer Auszeichnungssprache wie XML oder der Darstellung als RDF-Triple können die Elemente der Mikro- und der Makrostruktur für ein gezieltes Retrieval genutzt werden. Auf den RDF-Tripeln lassen sich Abfragen und Sortierfunktionen durchführen. Die Anfrage nach Name, Geburtsdatum und englischer Kurzbeschreibung von Informationswissenschaftlern, die in Frankreich geboren sind, in SPARQL (Simple Protocol and RDF Query Language) lautet:

```
PREFIX dbo: <http://dbpedia.org/ontology/>
```

```
SELECT ?name ?birth ?description ?person WHERE {
  ?person dbo:birthPlace :France .
```

```

?person <http://purl.org/dc/terms/subject> <http://dbpedia.org/re-
source/Category:Information_science> .
?person dbo:birthDate ?birth .
?person foaf:name ?name .

?person rdfs:comment ?description .
FILTER (LANG(?description) = 'en') .
}

```

Tabelle 2 zeigt das Ergebnis der Anfrage:

name	birth	description	person
"Renée-Marie-Hélène-Suzanne Briet"@en	"1894-02-01"^^xsd:date	"Renée-Marie-Hélène-Suzanne Briet, known as "Madame Documentation," was a librarian, author, historian, poet, and visionary best known for her treatise Qu'est-ce que la documentation? (What is Documentation?), a foundational text in the modern study of information science. She is also known for her writings on the history of Ardennes and the poet Arthur Rimbaud."@en	:Suzanne_Briet
"Suzanne Briet"@en	"1894-02-01"^^xsd:date	"Renée-Marie-Hélène-Suzanne Briet, known as "Madame Documentation," was a librarian, author, historian, poet, and visionary best known for her treatise Qu'est-ce que la documentation? (What is Documentation?), a foundational text in the modern study of information science. She is also known for her writings on the history of Ardennes and the poet Arthur Rimbaud."@en	:Suzanne_Briet

Tab. 2: Ergebnis SPARQL-Abfrage

### D 6.3.3 Lexikonevaluation – Qualitätskriterien

Ihrem Selbstverständnis nach sind Enzyklopädien gleichermaßen normativ und deskriptiv und verstehen sich nicht selten als „Brücke zwischen Wissenschaft und Gesellschaft“ (vgl. Lit. 08). Hieraus ergibt sich, dass ein kritischer und reflektierter Umgang mit Nachschlagewerken stets auch den Entstehungskontext eines Nachschlagewerkes berücksichtigen wird. Unterstützung für den quellenkritischen Umgang mit Informationsmitteln bieten auf einzelne Genres und den jeweiligen Informationsbedarf zugeschnittene Kriterienkataloge zur Qualitätsevaluation (vgl. Lit. 32). An ein Hilfsmittel bei der Faktenverifikation wird man andere Anforderungen stellen als an ein Lernwörterbuch. Die Tabelle 3 basiert auf einer (hier nicht vollständig wiedergegebenen) Synopse verschiedener Kriterienkataloge (vgl. Lit. 20). Die wissenschaftliche Auseinandersetzung mit verschiedenen Qualitätsaspekten der Wikipedia hat einen wichtigen Impuls zur Erforschung automatischer Verfahren der Qualitätsevaluation auf der Grundlage der Analyse messbarer Eigenschaften (Länge, Diskussion, Veränderungshäufigkeit etc.) geliefert (vgl. Lit. 32). Während viele der formalen Kriterien gut automatisch zu überprüfen sind, bleibt die vollständige Überprüfung der inhaltlichen Korrektheit angesichts einer Artikelanzahl von mehreren hunderttausend Artikeln bis zu mehreren Millionen Einträgen weiterhin ungelöst.



Kriterium	Beispiel für Testfragen	Beispiele für Qualitäts-Hinweise
Lemmatisierung	Ist die Lemmatisierung eindeutig und für die Zielgruppe nachvollziehbar?	Einsatz von expliziter oder impliziter Disambiguierung
Glaubwürdigkeit und Autorität	Findet eine Überprüfung der inhaltlichen Korrektheit statt? Findet eine Qualitätskontrolle statt? Ist die Information gut recherchiert? Wenn ein parteischer Standpunkt vertreten wird, ist dieser deutlich kenntlich gemacht?	Literaturangaben und Bibliographien sind vorhanden. Zusammensetzung der Redaktion, Autorenliste, Mitarbeit von Fachwissenschaftlern. Entscheidungsprozesse sind transparent. Im Vorwort werden die Ziele offengelegt.
Inhaltliche Tiefe	Ist die Auswahl der Lemmata und Artikel ausgewogen? Sind die Artikel aktuell? Haben die Informationen langfristig Gültigkeit?	Angaben zum Umfang Das Werk wird regelmäßig aktualisiert. Es werden aktuelle Zusatzinformationen angeboten (z. B. Jahreschroniken). Ähnliche Themen werden ähnlich behandelt.
Geltungsbereich	Ist die Zielgruppe genau angegeben?	
Klarheit und Lesbarkeit/Schreibstil	Ist der Schreibstil der Nutzergruppe angemessen? Gibt es eine einheitliche klare Mikrostruktur?	Verständliche Formulierungen, vereinfachte Darstellung von komplexen Sachverhalten Übersichtlichkeit, beschreibend statt wertend
Usability/Struktur und Zugang	Ist eine Orientierung an Benutzerstrukturvoraussetzungen erkennbar? Ist ein Index/Register vorhanden? Ist die Wahl des Zugangsalphabets angemessen? Ist ein systematischer Zugang auf die Information möglich? Wird Hypertext sinnvoll eingesetzt? Welche Suchoptionen? Welche Verweisungen/Links?	Unterschiedliche Zugänge zu den Inhalten speziell für elektronische Werke: Links (verweise innerhalb eines Artikels) Intertextuelle Links (Verweise zwischen Artikeln) extratextuelle Links (Verweise zu externen Quellen) Angebot Suche und Browsing
Zeitaufwand bei der Recherche	Wie hoch ist der Zeitaufwand bei der Recherche? Wie funktional werden Links und Animationen eingesetzt?	Einfache Suchmöglichkeiten Transparente Verweisstrukturen
Kosten	Ist das Preis-Leistungsverhältnis angemessen? Bietet der Verlag zusätzliche Mehrwertleistungen an? Welche? Wird die Nutzung durch Werbeeinblendungen beeinträchtigt?	Differenzierte Abrechnungsverfahren (Abonnement, pauschal, pay per view etc.).

**Tab. 3:** Qualitätskriterien für lexikographische Nachschlagewerke

### D 6.3.4 Lexika als Produkte auf dem Informationsmarkt

Die Produktion und Vermarktung von lexikographischen Nachschlagewerken ist gekennzeichnet durch eine Verquickung privatwirtschaftlicher Unternehmungen mit Institutionen, die teilweise oder ganz durch öffentliche Mittel finanziert werden (vgl. die Kooperation zwischen Dudenverlag und dem Institut für Deutsche Sprache Mannheim mit seinen verschiedenen Wörterbüchern, das Goethewörterbuch oder das Rechtswörterbuch). Lexikonproduzenten waren und sind aber auch im privatwirtschaftlichen Bereich Vorreiter für die Entwicklung neuer Geschäftsmodelle, etwa durch das Aufgreifen neuer Distributionsformen und Vertriebsstrategien wie Absatz in Einzellieferungen und Subskription sowie bis heute den Direktvertrieb durch Kolportage- und Reisebuchhandel (vgl. Lit. 08, S. 134ff, Lit. 33). Eine große Herausforderung für eine erfolgreiche Vermarktung sind die teilweise sehr langen Publikationszeiträume. Bearbeitungszeiträume zwischen 5 und 10 Jahren sind üblich (z. B. Wörterbuch der Lexikographie und Wörterbuchforschung oder Lexikon der Bibliotheks- und Informationswissenschaft). Über einen so langen Zeitraum gilt es durch geeignete Marketingmaßnahmen Kundeninteresse und Kaufbereitschaft aufrecht zu erhalten. Der Markt für gedruckte Wörterbücher und Enzyklopädien ist durch die Marktführerschaft weniger großer Verlage gekennzeichnet. Infolge des seit ihrer Gründung 2001 einmaligen Erfolgs der kollaborativen Online-Enzyklopädie Wikipedia veränderte sich der lexikographische Markt grundlegend. Bisher hat sich der kommerzielle Markt noch nicht erfolgsversprechend konsolidieren können, was sich daran zeigt, dass die Marktführer mit immer wieder veränderten Geschäftsmodellen wie Contentsyndication und Werbefinanzierung experimentieren. So gab Brockhaus 2008 die Einstellung der gedruckten Version bekannt, um 2011 eine neue Druckausgabe für 2014 anzukündigen (vgl. Lit. 33). Der Brockhaus wechselte verschiedentlich den Verlag, zuletzt von Langenscheidt zu Bertelsmann. Die Bertelsmann-Tochtergesellschaft wissenmedia vermarktet gedruckte und digitale Nachschlagewerke wie Brockhaus, Bertelsmann Lexikon oder Wahrig unter den jeweils eingeführten Markennamen. Im Bereich der mehrsprachigen Lexika engagieren sich vor allem die Schulbuchverlage, z. B. der Ernst Klett Verlag mit dem pons.eu-Portal. Neben der eigentlichen Werbefinanzierung findet man auch die Einbindung von Nachschlagewerken auf kommerziellen Webseiten, um den Traffic zu erhöhen und die Vertrauenswürdigkeit einer Seite zu steigern. Traditionelle Wörterbuchanbieter wie Duden (gehört seit 2009 zu Cornelsen) ergänzen ihr Portfolio durch Beratungsdienstleistungen und Sprachtechnologieangebote (vgl. Lit. 34).

## D 6.4 Forschungsfeld Lexikographie und Enzyklopädistik

Lexikographie ist ein Spezialgebiet der Linguistik. Neben der Erforschung historischer Wörterbücher, Spezialthemen der mehrsprachigen Lexikographie z. B. zu Wortgruppen oder zu Lernwörterbüchern, sind die Entwicklung von Nutzertypologien und elektronischen Wörterbüchern wichtige Forschungsfelder (vgl. Lit. 35). Die Fachlexikographie beschäftigt sich verstärkt mit Fragen der Fachsprachen und der Terminologie. In den Geschichtswissenschaften, der Philosophie und Bildungssoziologie werden Enzyklopädien aus unterschiedlichen Perspektiven als Indikatoren der sich kontinuierlich wandelnden gesellschaftlichen Bedeutung von Wissen, Bildung und Information wissenschaftlich untersucht (vgl. Lit. 03). Infolge des Erfolges der durch das Web 2.0 erst möglich gewordenen Onlineprojekte können Qualitätsfragen, die Grenzen und Möglichkeiten erfolgreicher Kollaboration und Kooperation und Fragen des Wissenszugangs- und der Zugänglichkeit erstmals auch in Form von groß angelegten quantitativen Untersuchungen untersucht werden (Lit. 32, Lit. 36).

Wörterbücher und Enzyklopädien im Kontext von Information und Dokumentation haben Gemeinsamkeiten, da sie:

- einen wesentlichen Beitrag zur Wissensorganisation leisten und gleiche oder ähnliche Strukturmerkmale aufweisen;

- die verstärkte Ausrichtung an den Nutzern durch die Erarbeitung von Nutzerstrukturbeschreibungen im Rahmen der linguistischen Lexikographieforschung ein gutes Beispiel für die Nutzerorientierung bei der Entwicklung von Informationsprodukten ist;
- und nicht zuletzt Bestrebungen die Zugänglichkeit lexikographischer Informationen durch Einsatz semantischer Technologien erhöhen und auf bereits geleistete Strukturierungsleistungen z. B. innerhalb der Textencoding Initiative (TEI) aufbauen können.

#### Danksagung

An dieser Stelle möchte ich den Studierenden der Masterkurse „Online-Lexikographie“ 2011 und 2012 am Department Information der HAW Hamburg danken. Die Arbeit an diesem Überblicksartikel wurde sehr durch die Vorarbeiten in Form von Forschungsüberblicken der Studierenden unterstützt.

## Literatur

- 01 Schmitz-Emans, Monika; Schulz, Christoph Benjamin ; Fischer, Kai Lars (Hrsg.): Enzyklopädien des Imaginären: Jorge Luis Borges im literarischen und künstlerischen Kontext. Hildesheim, Zürich, New York: Georg Olms, 2011. Online: Acta Litterarum. 2010. (<http://www.actalitterarum.de/theorie/mse/enz/index.html>; Abruf:2012-08-10)
- 02 Bates, Marcia J.: What Is a Reference Book: A Theoretical and Empirical Analysis. RQ 26 (1986), 37-57
- 03 „Allgemeinwissen und Gesellschaft“ ([www.enzyklopaedie.ch](http://www.enzyklopaedie.ch); Abruf: 2012-08-10)
- 04 Bibliothekszentrum Baden-Württemberg: FabiO: Wörterbücher, Übersetzungshilfen. (<https://wiki.bsz-bw.de/doku.php?id=linksammlungen:fabio:woerterbuecher>; Abruf: 2012-08-30)
- 05 Bibliothekszentrum Baden-Württemberg: FabiO: Lexika, Enzyklopädien. (<https://wiki.bsz-bw.de/doku.php?id=linksammlungen:fabio:lexika>; Abruf: 2012-08-10)
- 06 McArthur, Tom: Worlds of reference : Lexicography, learning and language from the clay tablet to the computer. Cambridge, New York, New Rochelle u. a.: Cambridge University Press, 1986
- 07 Selg, Anette; Wieland, Rainer: Die Welt der Encyclopédie. Frankfurt a. Main: Eichborn AG, 2001
- 08 Spree, Ulrike: Das Streben nach Wissen: Eine vergleichende Gattungsgeschichte der populären Enzyklopädie in Deutschland und Großbritannien im 19. Jahrhundert. Tübingen: Niemeyer, 2000
- 09 Herbst, Thomas; Klotz, Michael: Lexikographie. Paderborn [u. a.]: Ferdinand Schöningh, 2003
- 10 Geeb, Franziskus: Semantische und enzyklopädische Informationen in Fachwörterbüchern. Aarhus: Wirtschaftsuniversität, 1998
- 11 Bergenholtz, Henning; Tarp, Sven; Wiegand, Herbert Ernst: Datendistributionsstrukturen, Makro- und Mikrostrukturen in neueren Fachwörterbüchern. Handbücher zur Sprach- und Kommunikationswissenschaft, hrsg. v. Hugo Steger und Herbert Ernst Weigand. Band 14. Fachsprache, hrsg. v. Lothar Hoffmann, Hartwig Kalverkämper und Ernst Wiegand. Berlin, New York: Walter de Gruyter, 1998, S. 1889-1909
- 12 DWDS: Referenzkorpora. Stand 2011-06-06. (<http://www.dwds.de/resource/referenzkorpora/>; Abruf: 2012-08-10)
- 13 Wiegand, Herbert Ernst: Aspekte der Makrostruktur im allgemeinen einsprachigen Wörterbuch: alphabetische Anordnungsformen und ihre Probleme. HSK 5.1, 371-409 [= Nr. 65]
- 14 Michel, Paul: Darbietungsweisen des Materials in Enzyklopädien. Tomkowiak, Ingrid (Hrsg.): Populäre Enzyklopädien. Gedenkschrift für Rudolf Schenda. Zürich: Chronos Verlag, 2002
- 15 Wiegand, Herbert Ernst: Der Begriff der Mikrostruktur: Geschichte, Probleme, Perspektiven. HSK 5, 409-462
- 16 Wiegand, Herbert Ernst: Über die Mediostrukturen bei gedruckten Wörterbüchern. K. Hyldegaard Jensen, V Hjørnager Pedersen, A. Zettersten (Hrsg.): Symposium on Lexicography VI. Proceedings of the Sixth International Symposium on Lexicography May 7-9 1992 at the University of Copenhagen. Tübingen 1994, 315-329

- 17 Spohr, Dennis: Towards a multifunctional lexical resource: Design and implementation of a graph-based lexicon model. Berlin, Boston: de Gruyter, 2012
- 18 Fuertes-Olivera, Pedro A.; Bergenholtz, Henning (Hrsg.): e-Lexicography : The Internet, Digital Initiatives and Lexicography. London: Continuum International Publishing Group, 2011
- 19 Wiegand: Fachlexikographie. Lexicography for Special Purposes. Zur Einführung und bibliographischen Orientierung. *Lexicographica* 11/1995, 1-14
- 20 Lewandowski, Dirk; Spree, Ulrike: Ranking of Wikipedia articles revisited: Fair ranking for reasonable quality? *Journal of the American Society for Information Science and Technology* 61 (2010)
- 21 Busemann, Katrin; Gescheidle, Christoph: Web 2.0: Nutzung steigt – Interesse an aktiver Teilhabe sinkt. *Media Perspektiven* 7-8 (2010), 359-368. ([http://www.media-perspektiven.de/uploads/tx\\_mppublications/07-08-2010\\_Busemann.pdf](http://www.media-perspektiven.de/uploads/tx_mppublications/07-08-2010_Busemann.pdf); Abruf: 2012-08-15)
- 22 Waller, Vivienne: The search queries that took Australian Internet users to Wikipedia. *Information Research* 16 (2011), 48-63
- 23 Shawkat, E.: Wikipedia use. *British dental journal: BDJ; the journal of the British Dental Association* 206 (2009), 117-118
- 24 Messner, Marcus; South, Jeff: Legitimizing Wikipedia: How US national newspapers frame and use the online encyclopedia in their coverage. *Journalism practice* 5 (2011), 145-160
- 25 Stöcklin, Nando; Hielscher, Michael; Hartmann, Werner: Thematisieren statt verbieten! Kompetenter Umgang mit der Wikipedia als Unterrichtsthema. *Computer + Unterricht: Lernen und Lehren mit digitalen Medien* 21 (2011), 56-58
- 26 Darnton, Robert: Glänzende Geschäfte: Die Verbreitung von Diderots Enzyklopädie oder wie verkauft man Wissen mit Gewinn? Berlin: Wagenbach, 1993
- 27 Muggleston, Lynda: Lost for words: The hidden history of the Oxford English Dictionary. New Haven/ London: Yale University Press, 2005
- 28 Prodöhl, Ines: Die Politik des Wissens: Allgemeine deutsche Enzyklopädien zwischen 1928 und 1956. Berlin: Akademie Verlag, 2011
- 29 Zotter, Christoph: Der Experte ist tot, es lebe der Experte: Der Einfluss des Internets auf die Wissenskultur am Beispiel Wikipedia. Wien: facultas.wuw 2009
- 30 Klosa, Annette (erscheint): The lexicographical process II: Online dictionaries. Rufus H. Gouws; Heid, Ulrich; Schweickhard, Wiegand, Herbert Ernst (Hrsg.): *Dictionaries: An international encyclopedia of lexicography. Supplementary volume: Recent developments with special focus on computational lexicography*. Berlin; New York: de Gruyter. (Vorabdruck online: [http://multimedia.ids-mannheim.de/mediawiki/web/images/6/6b/Klosa\\_LexicographicalProcess\\_Preprint.pdf](http://multimedia.ids-mannheim.de/mediawiki/web/images/6/6b/Klosa_LexicographicalProcess_Preprint.pdf); Abruf: 2012-08-20)
- 31 TEI (Text Encoding Initiative): P5: Guidelines for Electronic Text Encoding and Interchange. Chapter 9: Dictionaries. Version. 2.1.0. Stand: 2012-06-17. (<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/DI.html#DIEN>; Abruf: 2012-08-20)
- 32 Hammwöhner, Rainer: Qualitätsaspekte der Wikipedia. Stegbauer, Christian, Schmidt, Jan, Schönberger, Klaus (Hrsg.): *Wikis: Diskurse, Theorien und Anwendungen. Sonderausgabe von kommunikation@gesellschaft*, Jg. 8. Online-Publikation ([http://www.soz.uni-frankfurt.de/K.G/B3\\_2007\\_Hammwoehner.pdf](http://www.soz.uni-frankfurt.de/K.G/B3_2007_Hammwoehner.pdf); Abruf: 2012-08-13)
- 33 Hünermann, Christoph: Wir hauchen dem Brockhaus neues Leben ein: Interview mit Wissenmedia Geschäftsführer Christoph Hünermann. *buchreport* (2011-01-29). ([http://www.buchreport.de/nachrichten/verlage/verlage\\_nachricht/datum/2011/01/29/wir-hauchen-dem-brockhaus-neues-leben-ein.htm](http://www.buchreport.de/nachrichten/verlage/verlage_nachricht/datum/2011/01/29/wir-hauchen-dem-brockhaus-neues-leben-ein.htm); Abruf: 2012-08-29)
- 34 Contentmanager.de: Dudenverlag. (<http://www.contentmanager.de/dienstleister/unternehmen/dudenverlag.html>; Abruf: 2012-08-22)
- 35 Wörterbuch zur Lexikographie und Wörterbuchforschung. Band 1: A - C. Wiegand, Herbert Ernst; Wolski, Werner; Storrer, Angelika; Kammerer, Matthias; Beißwenger, Michael (Hrsg.). Berlin u.a.: de Gruyter, 2010
- 36 Reagle, Joseph Michael: Good Faith Collaboration: The Culture of Wikipedia. Cambridge, Massachusetts: MIT Press, 2010