

Ulrike Spree, Hamburg, Jutta Lindenthal, Lübeck, und Andje Knaack, Bonn

Wortnetz Kultur – ein Thesaurusprojekt zur kollaborativen Erschließung von Fachinformationen des kulturellen Erbes

Am Beispiel der Kooperation zwischen dem Department Information der HAW Hamburg und dem Landschaftsverband Rheinland in einem Thesaurusprojekt zur kollaborativen Erschließung von Fachinformationen des kulturellen Erbes wird gezeigt, wie unter den begrenzten Möglichkeiten der informations- und bibliothekswissenschaftlichen Fachbereiche an Fachhochschulen Forschungs- und Entwicklungsvorhaben in enger Kooperation mit der Praxis umgesetzt werden können. Es wird dargestellt, wie Domänenexperten und Informationsspezialisten beim Aufbau und der Pflege eines kontrollierten Fachvokabulars erfolgreich kooperieren können, aber auch welche Schwierigkeiten eine solche Zusammenarbeit zu überwinden hat.

Deskriptoren: Thesauruserstellung, Projekt, Kultur, Fachhochschule, Zusammenarbeit, HAW Hamburg, Landschaftsverband Rheinland

Wortnetz Kultur – a thesaurus project for the collaborative indexing of cultural heritage

The cooperation between the Department Information of the University of Applied Sciences Hamburg and the „Landschaftsverband Rheinland“ in a thesaurus-project for the collaborative indexing of cultural heritage is an example of a successful research cooperation between information scientists and domain specialists regardless of the limited financial and human resources of LIS-departments at universities of applied sciences in Germany.

Keywords: thesaurus construction, project, cultural heritage, cooperation, University of Applied Sciences Hamburg, Landschaftsverband Rheinland

Réseau de mots dans le domaine de la culture – un projet de thésaurus pour la description collaborative des informations spécialisées dans le patrimoine culturel

L'exemple de la coopération entre le Département d'Information de la HAW Hamburg (l'Université des Sciences Appliquées de Hambourg) et le Landesverband

Rheinland (le Conseil régional de Rhénanie) sur un projet de thésaurus pour la description collaborative des informations spécialisées du patrimoine culturel montre comment des projets de recherche et de développement peuvent être développés en étroite collaboration avec la pratique et cela malgré les possibilités limitées des départements des sciences de l'information et de bibliothéconomie des universités. L'article montre comment des experts dans ces domaines et des spécialistes de l'information peuvent coopérer avec succès dans la construction et le maintien d'un vocabulaire spécialisé contrôlé, mais il décrit aussi les difficultés qu'une telle coopération doit surmonter.

Mots-clés: thésaurus, projet, collaboration, patrimoine culturel, Université des Sciences Appliquées de Hambourg, Conseil régional de Rhénanie

1 Einleitung

Im August 2011 feierte Wortnetz Kultur (WNK) seinen zehntausendsten Eintrag.¹

Wortnetz Kultur ist ein Thesaurus, der zur Erschließung kulturhistorischer Fachinformationen entwickelt wird. Er soll für Indexierung und Retrieval in verschiedenen Informationssystemen eingesetzt werden, z. B. in der Bau- und Bodendenkmalpflege, aber auch zur Beschreibung von Museumsobjekten und Umweltdaten. Eine erste Anwendung findet der Thesaurus seit Oktober 2010 im Portal KuLaDig (Kultur.Landschaft.Digital), einem Online-Informationssystem über das landschaftliche Kulturerbe in Nordrhein-Westfalen und Hessen. KuLaDig ist eine

¹ Anfang November 2011 enthielt der Thesaurus 12.285 Bezeichnungen für 3.206 Begriffe sowie 5.555 hierarchische Verknüpfungen und 7.260 Assoziationsrelationen.

öffentlich zugängliche Web-Anwendung, deren Datenbestand sich noch im Aufbau befindet.²

Im Juni 2007 wandte sich Andje Knaack, Mitarbeiterin des LVR-Amtes für Bodendenkmalpflege im Rheinland (ABR)³, telefonisch an das Department Information der Hochschule für Angewandte Wissenschaften (HAW) Hamburg mit der Bitte um eine fachliche Beratung bei der Erstellung eines Vokabulars zur Erschließung archäologischer Fachinformationen. Sie berichtete, dass eine neue, web-basierte Datenbank für Bodendenkmalpflege und Denkmalpflege konzipiert werde und man für dieses Projekt die bereits bestehenden Wortlisten in ein Begriffs- system überführen möchte. Der Anlass für die Kontaktaufnahme war eine konkrete Frage: „Wie kann die Tatsache abgebildet werden, dass man einen Sarkophag während einer Ausgrabung als Befund bezeichnen kann und er nach der Bergung im Museum dann als Fund angesprochen wird?“

Aus dieser unverbindlichen Anfrage hat sich eine praxiszentrierte Zusammenarbeit zwischen dem Landschaftsverband Rheinland (LVR) und dem Department Information der HAW entwickelt. In einem späteren Entwicklungsabschnitt kamen eine freiberufliche Informationsberaterin und die Kulturgrenossenschaft digiCULT-Verbund eG als Partner hinzu.

Im Kontext des vorliegenden Schwerpunktheftes zu Forschung an Fachhochschulen ist die entstandene Zusammenarbeit in doppelter Hinsicht interessant. Erstens ist der Ablauf der Kooperation mit der Praxis typisch für die Art und Weise, wie unter den begrenzten Möglichkeiten der informations- und bibliothekswissenschaftlichen Fachbereiche – Lehrdeputat von 18 Semesterwochenstunden, aufgrund begrenzter Personalressourcen in der Forschungsinfrastruktur kaum Beratung bei der Vertragsgestaltung für Drittmittelprojekte, wenige wissenschaftliche Mitarbeiter und nur geringe Laborunterstützung für IT-Projekte sowie kleine Forschungsbudgets – an Fachhochschulen Forschung, oder genauer Forschung und Entwicklung, möglich ist. In diesem Beitrag wird an einem Beispiel dargestellt, wie unter diesen Rahmenbedingungen dennoch Forschungs- und Entwicklungsvorhaben in enger Kooperation mit der Praxis umgesetzt werden können. Zweitens wurden im Rahmen des Projektes Wortnetz Kultur wertvolle Erfahrungen gesammelt, wie Domänenexperten und Informationsspezialisten beim Aufbau und

der Pflege eines kontrollierten Fachvokabulars erfolgreich kooperieren können, aber auch welche Schwierigkeiten eine solche Zusammenarbeit zu überwinden hat. Diese Erfahrungen werden zum Ende dieses Beitrags in fünf Lessons learned gebündelt.

2 KuLaDig und Wortnetz Kultur

Der Landschaftsverband Rheinland (LVR) entwickelt ein Fachinformationssystem, über das in einer ersten Phase die Informationen einiger Einrichtungen des LVR-Dezernats Kultur und Umwelt (Industriemuseen, Landesmuseum, Denkmalpflege, Bodendenkmalpflege, Kulturlandschaftspflege, Landeskunde und Regionalgeschichte), die fachliche Inhalte dokumentieren und strukturiert aufbereiten, sowohl den Nutzern der verschiedenen LVR-Einrichtungen als auch der Fachwelt und der breiten Öffentlichkeit zugänglich gemacht werden sollen. Die präsentierten Informationen (Texte, Karten, Bilder) zu Objekten des landschaftlichen kulturellen Erbes können in KuLaDig durch eine Kartenansicht über ein Geographisches Informationssystem (GIS) lokalisiert werden.

3 Phasen der Kooperation zwischen dem Landschaftsverband Rheinland (LVR) und der HAW Hamburg

In den verschiedenen Fachabteilungen des Landschaftsverband Rheinland (LVR) wurde eine Reihe von Vokabularen zur Inhaltserschließung eingesetzt. Aufgrund ihrer unterschiedlichen Herkunft und Struktur schienen sie für eine kollaborative Nutzung nicht geeignet, sodass der LVR an eine Neuentwicklung dachte. Der LVR wünschte sich ein Instrument zur gemeinsamen Nutzung in den verschiedenen Dienststellen, um eine einheitliche Indexierung der Dokumente und eine anwendungsübergreifende, gezielte Recherche zu gewährleisten. Dabei sollten die bereits existierenden Schlagwortlisten zunächst für die Projekte KuLaDig und BODEON (Boden- und Denkmalpflege online) in ein verbindlich zu nutzendes Vokabular überführt werden. Die wichtigsten Anforderungen fasst die Projektskizze des LVR zusammen:

- inhaltlich-systematische Gruppierung der Schlagwörter

² Landschaftsverband Rheinland; Landesamt für Denkmalpflege Hessen: KuLaDig. URL: <http://www.kuladig.de/>.

Teile des Wortgutes des Thesaurus sowie die hierarchischen Beziehungen sind über die erweiterte Suche zugänglich.

³ Damals noch „Rheinisches Amt für Bodendenkmalpflege“ (RAB).

- hierarchische Organisation des Vokabulars
- Abbildung der Inhaltsgruppen durch Facetten
- Synonymliste
- Glossar.

Besonders wichtig war dem Verband eine Funktion, die es den verschiedenen LVR-Einrichtungen ermöglicht, jeweils eigene, anwenderspezifische Sichten auf das gemeinsame Vokabular zu definieren.

Die Abschnitte unserer Kooperation orientierten sich weitgehend an den Empfehlungen aktueller Standards zur Vokabularentwicklung wie ANSI/NISO Z39.19 und ISO 25964-1:2011:

- In einem ersten Schritt sollte der geeignete Vokabulartyp ermittelt und mögliche nachnutzbare Quellen gefunden werden. Diese Phase wurde mit Unterstützung durch Studierende des Department Information der HAW durchgeführt.
- Anknüpfend an diese Ergebnisse wurde das Thesauruskonzept im Rahmen einer externen Beratungsleistung entwickelt und eine geeignete Software zur Verwaltung des Vokabulars ausgewählt.
- Nun begann die eigentliche Arbeit am Vokabular, die in den Händen der Redaktion des LVR liegt. Die externe Informationsspezialistin und die Fachhochschule begleiten diesen Prozess durch kontinuierliche Beratung.
- In einer dritten Phase der Kooperation wurde am Department Information die Gebrauchstauglichkeit des ersten Prototyps der Endnutzerschnittstelle des Vokabulars durch Usability-Tests nach der Thinking-Aloud-Methode mit externen Fachwissenschaftlern und Informationsspezialisten evaluiert.
- Zur Zeit untersucht ein Student des Studiengangs Informationswissenschaft- und -management in seiner Masterarbeit die Möglichkeit der Dokumentation und Präsentation des Thesaurus-Handbuchs in einem semantischen Wiki.

4 Entwicklung eines Thesaurus-Prototyps

Im Rahmen einer Lehrveranstaltung zu Methoden der Wissensorganisation wurden von den Studierenden zunächst Vorschläge entwickelt, wie die Schlagwortlisten des ABR in ein kontrolliertes Vokabular überführt werden könnten. Für die Anforderungsanalyse wurden auf dem Wege der teilnehmenden Beobachtung (eine Art Shadowing vor

Ort) sowie mit Hilfe von Fokusgruppen und Interviews mit den Mitarbeiterinnen und Mitarbeitern des LVR-Amtes für Bodendenkmalpflege die Arbeitsabläufe, die für die Arbeit mit einem kontrollierten Vokabular von Bedeutung sind, ermittelt. Es wurde deutlich, dass ein Thesaurus den Anforderungen an eine tiefe inhaltliche Beschreibung der dokumentierten Bodendenkmäler besser gerecht werden würde als eine Klassifikation. Beim Retrieval soll der Zugriff auf einzelne Dokumente und weniger die Ordnung von Dokumentgruppen unterstützt werden. Die Verschlagwortung der Dokumente in den verschiedenen Informationssystemen des LVR-Amtes für Bodendenkmalpflege mit Hilfe einer einfachen Wortliste war nicht immer einheitlich und es entstand ein erhöhter Aufwand bei der Recherche.

Aus diesen ersten Beobachtungen und Überlegungen hat Magdalena Kaljaj in ihrer Diplomarbeit⁴ ein Thesauruskonzept entwickelt, das die in den Normen DIN 1463-1 und ANSI/NISO Z39.19 dargelegten Prinzipien der terminologischen Kontrolle sowie die Regeln zur Erstellung von Thesaurusrelationen berücksichtigt. Zur Erarbeitung der systematischen Struktur des Thesaurus wurden u. a. die semantischen Beziehungen innerhalb des Vokabulars auf der Grundlage der Referenzontologie CIDOC Conceptual Reference Model (ISO 21277) analysiert. Hier zeichnete sich bereits auch die wegweisende Bedeutung des Getty Art & Architecture Thesaurus und der English Heritage National Monuments Record Thesauri ab. Das Konzept konnte Frau Kaljaj dann in einem Prototypen veranschaulichen: Die etwa 500 vorliegenden Begriffe aus dem LVR-Amt für Bodendenkmalpflege präsentierten sich nun in der Struktur eines kleinen Thesaurus.⁵ Diplomarbeit und Demoversion des Thesaurus wurden einigen Kolleginnen und Kollegen des beteiligten LVR-Amtes zum Testen zur Verfügung gestellt.

Bei diesem Test zeigte sich, dass es teilweise Probleme bereitete, die gesuchten Begriffe über die hierarchische Einbindung oder die Klassenzuordnung aufzufinden. So vermutete z. B. keiner der testenden Kollegen den Begriff „Sarg“ in einer Klasse „Behälter“. Der „Sarg“ wurde eher im Kontext von Bestattung oder Grab gesehen. Dies verdeutlichte, dass bei den Domänenpezialisten oft eine thematische und kontextbezogene Sichtweise dominiert, während formale/funktionale Ordnungsaspekte als Grundlage der Begriffsstrukturierung zunächst eher fremd sind.

⁴ Kaljaj 2008. Online: http://opus.haw-hamburg.de/frontdoor.php?source_opus=580&la=de.

⁵ Kaljaj 2008.

Sarg	
NT	Baumsarg
	Bleisarg
	Holzsarg
	Sarkophag
CC	5.3.2 Behälter
termCreatedDate 2007-06-05	

Abb. 1: Ausschnitt aus dem Thesaurus-Prototypen.

Auf der Grundlage dieses kleinen Thesaurus, der den Domänenexperten die Funktionsweise eines Thesaurus mit dem ihnen vertrauten Vokabular veranschaulichte, war eine Diskussion über zentrale Konzepte der Thesaurustheorie wie terminologische Kontrolle, Facettenanalyse oder systematische Darstellung des Thesaurus möglich. Zudem wurde den Mitarbeitern und Entscheidern auf Seiten des Landschaftsverband Rheinland die Komplexität der Aufgabenstellung bewusst. Auch bildete sich ein Konsens über die Nützlichkeit eines Thesaurus als Mittel der Verständigung der beteiligten Abteilungen untereinander, und die Erwartungen an ein strukturiertes Vokabular auf Basis der existierenden Wortlisten konkretisierten sich.

dem Bereich Archäologie bzw. Kulturerbe. Die Recherchen nach verwandten Web-Projekten zeigte, dass weltweit bereits eine Fülle von Portalen zum Kulturerbe existiert oder geplant ist, diese jedoch noch kaum miteinander vernetzt waren. Der Zugriff auf die Informationen auch über Geoinformationssysteme, etwa mit Hilfe interaktiver Karten (Georeferenzdaten), konnte an vielen der untersuchten Beispiele beobachtet werden. Insofern unterstützten die Recherchen den vom LVR für das Kulturportal angedachten Ansatz, Georeferenzdaten für das Retrieval zu nutzen. Bei den eingesetzten kontrollierten Vokabularen handelte es sich überwiegend um Einzellösungen für spezialisierte Fachgebiete oder einzelne Web-Projekte. Es konnten keine für eine Nachnutzung frei verfügbaren Vokabulare ermittelt werden, auf die das Wortnetz Kultur unmittelbar hätte aufbauen können. Die Beispiele waren auch geeignet die Notwendigkeit des Einsatzes von einfachen und standardisierten Konventionen zur Beschreibung von Ressourcen, also die Einigung auf Metadatenstandards im Allgemeinen sowie die Verwendung von kontrollierten Vokabularen im Besonderen, zu untermauern.

Im Hinblick auf die Sicherstellung einer finanziellen Unterstützung für das Projekt wuchs jetzt die Erkenntnis, dass eine durchdachte und nachhaltige Informationsstrukturierung sowie der abgestimmte Einsatz wissensorganisatorischer Instrumente von nicht zu überschätzender Bedeutung für den Erfolg und die Nachhaltigkeit der verschiedenen geplanten Online-Informationssysteme wie z. B. KuLaDig sein würde.

Die Beispiele, die Frau Pinna und Frau Wessels zusammengetragen hatten, zeigten außerdem anschaulich, wie entscheidend die nutzerfreundliche Einbindung von Vokabularen in die Suchschnittstellen der Datenbankanwendungen ist. Im LVR gab es zu dem Zeitpunkt noch keine konkreten Vorstellungen darüber, wie ein zu entwickelnder Thesaurus, das zukünftige Wortnetz Kultur, für die Erschließung und die Informationsrückgewinnung (Recherche), etwa in einer vernetzten Datenbank oder einer Webanwendung, eingesetzt werden würde. Beim Auftakt-Workshop des Projekts „Wortnetz Kultur“ im Juni 2009 am Institut für Landeskunde und Regionalgeschichte des LVR wurde deswegen dieser Aspekt besonders berücksichtigt. Auf der Grundlage ihrer Recherchen erstellten Pinna und Wessels eine Präsentation „Vokabular meets Webprojekt“, in der sie an vier ausgewählten Webprojekten (Archäologische Ausgrabungen in Sachsen, Archäologisches Informationssystem Oberösterreich, Zentrales archäologisches Inventar von Flandern und FMIS Fornsök) Beispiele für die wissensorganisatorische Aufbereitung und den Einsatz von kontrollierten Vokabularen für die

5 Überblick über kontrollierte Vokabulare und Web-Projekte im Bereich des Kulturerbes

In der nun folgenden Phase der Konkretisierung der Anforderungen an das Wortnetz Kultur konnte die Kooperation zwischen dem LVR und der HAW von einer weiteren Unterstützung durch Studierende profitieren. Loredana Pinna und Wiebke Wessels erarbeiteten im Rahmen ihres Forschungs- und Praxisprojekts, das Teil des Masterstudiengangs Informationswissenschaft und -management an der HAW Hamburg ist, einen umfangreichen Literaturbericht zu Web-Projekten, Standards und Vokabularen, der für die weitere Projektarbeit als Best-Practice-Bericht Orientierung bieten konnte.⁶ Die Übersicht identifiziert und beschreibt 54 Vokabulare (Thesauri, Taxonomien, Wortlisten, Systematiken und ähnliches) und 41 Web-Projekte aus

⁶ Pinna; Wessels Literaturbericht 2009; dieselb.: Webprojekte, Standards, Vokabulare 2009.

Recherche vorstellten. Besonders das schwedische Projekt FMIS⁷ vermittelte den Teilnehmern des Workshops einen konkreten Eindruck davon, wie die kontrollierten Vokabulare in eine Rechercheschnittstelle eingebunden werden könnten.

An der konkreten Umsetzung des Thesaurus ist das Department Information nicht mehr direkt beteiligt, begleitet den Prozess aber weiterhin durch die Integration in das Qualitätsmanagement des Wortnetz Kultur, durch die Evaluierung der Usability des Vokabulars aus der Perspektive der Nutzergruppen der Fachanwendungen und der Einsatzmöglichkeiten für das Retrieval innerhalb der Webanwendung KuLaDig. Eine nutzerfreundliche Aufbereitung des Thesaurus-Handbuchs, verschiedener Schulungsmaterialien und Nachschlagewerke in Form eines Wiki mit semantischen Suchmöglichkeiten ist Gegenstand einer laufenden Masterarbeit im Studiengang Informationswissenschaft und -management.

6 Entwicklung des Thesauruskonzepts

Entwickler von kontrollierten Vokabularen können auf Standards⁸ und umfangreiche informationswissenschaftliche Literatur zur Thesaurustheorie und -praxis⁹ zurückgreifen. Hingegen wird der Prozess der tatsächlichen Vokabularentwicklung in Zusammenarbeit mit Anwendern, Fachnutzern, Entwicklern und Laiennutzern in den Projekten dann häufig nur als Nebenaspekt behandelt. Im Folgenden wird die Realisierung des Projekts Wortnetz Kultur vom Entwurf des Thesauruskonzepts über die Implementierung der Entwicklungsumgebung bis hin zu einem Blick in die Werkstatt der Thesaurusredaktion dargestellt.

6.1 Das Ausgangsvokabular

In den einzelnen LVR-Einrichtungen waren unterschiedliche Vokabulare (elf flache Schlagwortlisten, fünf Systematiken und ein hierarchisch strukturiertes Begriffssystem)

⁷ Riksantikvarieämbetet (Hrsg): Riksantikvarieämbetet – Fornsök [online]. – Online: <http://www.fmis.raa.se/cocoon/fornsov/search.html> [10.11.2011].

⁸ ISO 25964-1:2011, ANSI/NISO Z39.19: 2005, ISO 2788, DIN 1463-1 und -2.

⁹ Vgl. z. B. Wersig 1978, Aitchison et al. 2000, Roe u. a. 2004, Broughton 2006, Harpring 2010.

zur Indexierung benutzt worden. Wie die erste Analyse von Magdalena Kaljaj schon gezeigt hatte, konnten diese Vokabulare nicht ohne weitere Bearbeitung für eine kooperative Erschließung eingesetzt werden. Eine genauere Inspektion dieser Wortsammlungen deckte den Regelungsbedarf auf, z. B. für die Ansetzungen der Bezeichnungen oder eine Synonym- und Zerlegungskontrolle. Dieses „gewachsene“ Wortgut erwies sich aber auch als eine Fundgrube von Begriffskombinationen, die anzeigen, welche Fülle an Informationen ein Indexierer oft mit einem einzelnen Term ausgedrückt sehen möchte.

Das Fehlen von Ansetzungsregeln führte nicht nur zu ungebräuchlichen Bezeichnungen, sondern auch zu einer unbeabsichtigten Bedeutungsvielfalt der verwendeten Wort- und Satzzeichen (z. B. Gewandschm. -zubeh. allgem.). Oft sind verschiedene Begriffe in einem einzigen Eintrag aufgezählt und zu einer Äquivalenzklasse verkettet, wie z. B. Bienenhaus, Bienenzaun, Immenstand, Immenzaun; Straße/Weg/Straßentrasse/-damm/-graben. Einige Bezeichnungen waren mit Zusätzen versehen, wie sie konventionellerweise zur Disambiguierung verwendet werden. Hier wurden sie aber oft benutzt, um den Begriffskontext auszudrücken, also z. B. den topographischen Zusammenhang wie in Chaussee (Dammlage) oder den zeitlichen Bezug wie in Höhle (Paläolithikum). Da die Zusätze selbst aber nicht kontrolliert waren, haben sie manchmal weitere Synonyme produziert, z. B. Flügel (Instrument) – Flügel (Musikinstrument)¹⁰.

Das Ausgangsvokabular enthielt in hohem Maße Begriffskombinationen wie Feld-/Waldbegrenzung durch Wall und/oder Graben, ältere gelbtonige Drehscheibenware oder Doppelmalakowtürme mit dazwischenliegendem Maschinenhaus. Diese Phrasen wurden für Wortnetz Kultur auf ihre Komponenten und impliziten Relationen untersucht und zur Ermittlung relevanter Facetten herangezogen. Besonders häufig wurden Objektbezeichnungen mit Begriffen aus den Kategorien Material, Epoche/Stil/Zeit und Objektmerkmal kombiniert.

Eine der schwierigsten Aufgaben bei der Thesaurusbearbeitung ist die Zerlegungskontrolle, die auch das Redaktionsteam von Wortnetz Kultur sehr beschäftigt: Wann sollen Kombinationen als spezifische Begriffe erhalten bleiben und wann ist eine Zerlegung mit einer mindestens dreistelligen Verweisungsrelation sinnvoll? Allgemeine Regelungen dafür zu finden ist auch deshalb nicht einfach, weil die Entscheidung für ein präkombinierendes

¹⁰ Aus diesem Grund wurde für Wortnetz Kultur der Homonymenzusatz als eigenständiges Datenelement mit kontrollierten Werten definiert.

oder postkoordinierendes Verfahren nicht unabhängig von der Metadatenstruktur der Fachdatenbank getroffen werden kann. Daraus kann ein Zielkonflikt entstehen: Das Begriffssystem soll als Wissensbasis allgemein nutzbar und interoperabel sein und möglichst unabhängig von Designkriterien der jeweiligen Anwendungen aufgebaut werden, aber die Struktur der Fachdatenbanken ist häufig bereits vorgegeben und im Rahmen eines solchen Projekts oft nicht mehr zu beeinflussen. Erst im Verlauf des Projekts Wortnetz Kultur stellte sich heraus, dass die Datenbankmodelle für KuLaDig und BODEON nicht für alle gewünschten Erschließungsaussagen mächtig genug waren.

6.2 Das Thesauruskonzept

Wortnetz Kultur ist im Unterschied zu den Ausgangsvokabularen konzeptbasiert, d. h. die Vokabularelemente werden durch einen abstrakten Identifikator eindeutig adressiert (hier einen URI) und durch Bezeichnungen repräsentiert. Bevorzugte Bezeichnungen vertreten den Begriff für Indexierung und Retrieval. Kernelemente des Thesauruskonzepts sind die Deskriptoren, die in allen beteiligten Einrichtungen für die Indexierung verbindlich zu benutzen sind. Daneben gibt es Strukturelemente, die zur Gruppierung der Begriffe eingesetzt und selbst nicht zur Indexierung verwendet werden. Im Wortnetz Kultur gibt es zwei Typen dieser Gruppierungselemente: die in den Standards so genannten „Node Label“, die der Gruppierung von Geschwisterbegriffen unter einem bestimmten Aspekt dienen, und die „Kategorien“, die eine anwendungsbezogene oder thematische Clusterung der Deskriptoren ermöglichen. Diese in der Thesaurustheorie und -praxis übliche klassifikatorische oder systematische Ordnung der Terme wird in Wortnetz Kultur benutzt, um die einzelnen Fachsichten der LVR-Einrichtungen abzubilden.

Strukturelemente wie „Node Label“ haben im LVR-Thesaurus eine bedeutende Rolle gewonnen, besonders zur Gruppierung von Geschwisterbegriffen nach explizit benannten Merkmalen. Diese Reihung nebengeordneter Begriffe ist nicht nur informativ, sondern auch nützlich zum Prüfen der Vollständigkeit des Indexats. Zusätzlich können diese „Minifacetten“ für eine facettenbasierte Navigation ausgewertet werden.¹¹

¹¹ Siehe dazu Priss/Jacob 1999.

6.2.1 Bezeichnungen und ihre Attribute

Die Rolle der Bezeichnungen wurde über die übliche Kennzeichnung als Deskriptor oder Nichtdeskriptor hinaus weiter ausdifferenziert, um z. B. auch Quasisynonyme markieren zu können. Besonders wichtig für Wortnetz Kultur war die Möglichkeit zur Angabe der Sprachvarietät. Eine Fülle regionaltypischer Bezeichnungen im Rheinland und umgangssprachliche Varianten sind für den Sucheinstieg relevant und werden deshalb als synonyme oder quasisynonyme Äquivalente zur fachwissenschaftlichen Vorzugsbezeichnung aufgenommen.

6.2.2 Die Begriffsrelationen

Entgegen häufiger Praxis¹² werden die Hierarchierelationen in diesem Thesaurus standardkonform differenziert in generische, partitive und Instanzbeziehung.

Für die Genus-Spezies-Relation gilt die Bedingung der logischen, transitiven Is-A-Relation. Diese Vereinbarung wurde für Wortnetz Kultur getroffen, um

1. eine automatische Ausweitung der Suche (Explode) über die Hierarchieleiter zu unterstützen;
2. nachvollziehbare und möglichst allgemein gültige Kriterien zur konsistenten Pflege der Thesaurushierarchien an die Hand zu geben; subjektive, individuelle Sichten sollen in der Begriffsleiter ebenso vermieden werden wie „pragmatisch“¹³ motivierte Über- und Unterordnungen;
3. die Interoperabilität mit anderen Begriffssystemen zu erleichtern;
4. den Thesaurus möglichst nachnutzbar zu machen.

Teil-Ganzes-Beziehungen sind im Wortnetz Kultur z. B. für Geografika, aber auch zur Objektdokumentation von besonderer Bedeutung. Für geografische Entitäten ist eine partitive Hierarchierelation wünschenswert, um das automatische Einbeziehen von Teilen eines Gebietes zu ermöglichen, aber auch umgekehrt eine Ausweitung der Suche auf die umfassende Region zu gestalten.

Eine andere Variante der partitiven Beziehung findet man in der Welt der archäologischen Funde und Museumssubjekte, wo oft Teile eine Rolle spielen, die nicht konsolidativ für das Ganze sind. Zur Darstellung dieses Beziehungstyps wurde eine spezifische Assoziationsrelation

¹² So kennt beispielsweise der Standard-Thesaurus Wirtschaft (STW) nur die nicht weiter spezifizierte Hierarchierelation narrower term/broader term. (Wikisaurus 2011).

¹³ Fischer (1999) nennt dieses Vorgehen sehr anschaulich „floating pragmatics“. Ein „pragmatisches“ Vorgehen vertreten dagegen z. B. Soergel (1995) und Lambe (2007).

definiert („kann Teil sein von/kann als Teil haben“). Die Anwendung dieser Relation stellte die Thesaurusentwickler aber vor ein Entscheidungsproblem: Wann ist es sachlich sinnvoll und für die Suche nützlich, einen Begriff als möglichen Teil eines Ganzen auszuweisen? Für diesen Anwendungsfall müssen für das Wortnetz in Zukunft Regeln formuliert werden, die von Fachwissenschaftlern und Informationsspezialisten gemeinsam erarbeitet werden sollten.

Die „klassische“ Assoziationsrelation machte den Domänenexperten grundsätzlich Probleme. Es fehlten ihnen Kriterien, wann eine assoziative Verlinkung „zur Seite“ sinnvoll ist. Die „Verwandtschaftsrelation“ wird auch in den aktuellen Normen eher grob definiert, aber es werden typische Anwendungsfälle beschrieben und die Semantik dieser Relationen wird explizit benannt, sodass Anhaltspunkte für sinnvolle assoziative Verknüpfungen gegeben sind. Eine regelbasierte Definition der Assoziationsrelation ist jedoch wünschenswert, um den Thesaurusentwicklern Entscheidungshilfen zu geben und die Konsistenz des Begriffsnetzes zu erhöhen.

6.2.3 Spezielle Anforderungen

Die Thesaurusbegriffe sollen allen LVR-Einrichtungen für die gebundene Indexierung zur Verfügung stehen und die gemeinsame Nutzung der Indexierungsterme ist verbindlich. Einige Anwender benötigen aber darüber hinaus spezielle Termini, die nicht für die allgemeine Nutzung geeignet sind. Domänenexperten, Informationsspezialisten und Programmierer entwickelten gemeinsam eine Lösung, indem anwendungsspezifische Mapping-Relationen definiert und implementiert wurden.

7 Die Entwicklungsumgebung – Thesaurussoftware xTree

Vor allem zwei Aspekte waren – neben Kompatibilität mit Thesaurusnormen und standardisierten Austauschformaten – bei der Auswahl der Software wichtig: Sie sollte verteiltes, kollaboratives Arbeiten gut unterstützen und vor allem anwenderspezifisch erweiterbar sein.

Die Wahl fiel auf die Vokabularmanagement-Software xTree, die im Rahmen des Verbundprojekts digiCULT¹⁴ entwickelt worden war. Das Objektmodell verfügte bereits über wesentliche Eigenschaften, die die geforderten Konstrukte für Wortnetz Kultur darstellen können. Zusammen mit der Kulturgenossenschaft digiCULT-Verbund eG konnte xTree um die gewünschten Funktionalitäten erweitert werden. Die mit xTree erarbeiteten Vokabulare werden standardmäßig über einen formatneutralen Web-service (SOAP) bereitgestellt. Für Wortnetz Kultur wurde von digiCULT-Verbund eG eine Pilotimplementierung des Austauschformates vocnet¹⁵ vorgenommen. Zur Erleichterung der Kommunikation mit den Datenbankentwicklern wurde das vocnet-Modell mittels UML und OWL spezifiziert.

8 Die alltägliche Vokabulararbeit – Blick in die Werkstatt

Den Grundstock für Wortnetz Kultur bildeten die für KuLaDig gesammelten Deskriptorvorschläge. Dieser pragmatisch-empirische Ansatz bestimmt das Vorgehen beim Aufbau des Wortschatzes auch weiterhin, indem die einzelnen LVR-Einrichtungen die gewünschten Indexterme als Kandidaten bei der Thesaurusredaktion anmelden. Die Thesaurusredaktion ist zuständig und verantwortlich für Annahme oder Ablehnung der Vorschlagsterme, die Einpflege von akzeptierten Kandidaten, terminologische und begriffliche Kontrolle und die Einarbeitung der Deskriptoren in das Thesaurusnetz. Wenn nötig, werden generische Begriffe ergänzt, um das Relationennetz so weit möglich lückenlos, dicht und konsistent zu knüpfen.

Mit zunehmender Erfahrung bei der Aufnahme neuer Begriffe und Bezeichnungen zeigte sich, dass es für die

¹⁴ Das Verbundprojekt digiCULT war ein von der europäischen Union und dem Land Schleswig-Holstein von 2003 bis 2009 gefördertes interdisziplinäres Projekt zur digitalen Erfassung und Publikation musealer Sammlungen. Im Januar 2010 wurde das Projekt durch die Gründung der Genossenschaft digiCULT-Verbund eG verstetigt (s. http://digicult.museen-sh.de/pdfwrap.php?file=files/genossenschaft/digicult_eg_projekt_zur_genossenschaft.pdf).

¹⁵ Das Austauschformat vocnet (Version 0.4) ist eine Weiterentwicklung von museumvok (s. <http://museum.zib.de/museumsvokabular/index.php?main=tech-dok&ls=9&co=we>) und berücksichtigt das Datenmodell des ISO-Standards 25964-1:2011 (vgl. http://www.niso.org/schemas/iso25964/Model_2011-06-02.jpg) sowie die SKOS-Spezifikation (<http://www.w3.org/TR/2009/REC-skos-reference-20090818/>).

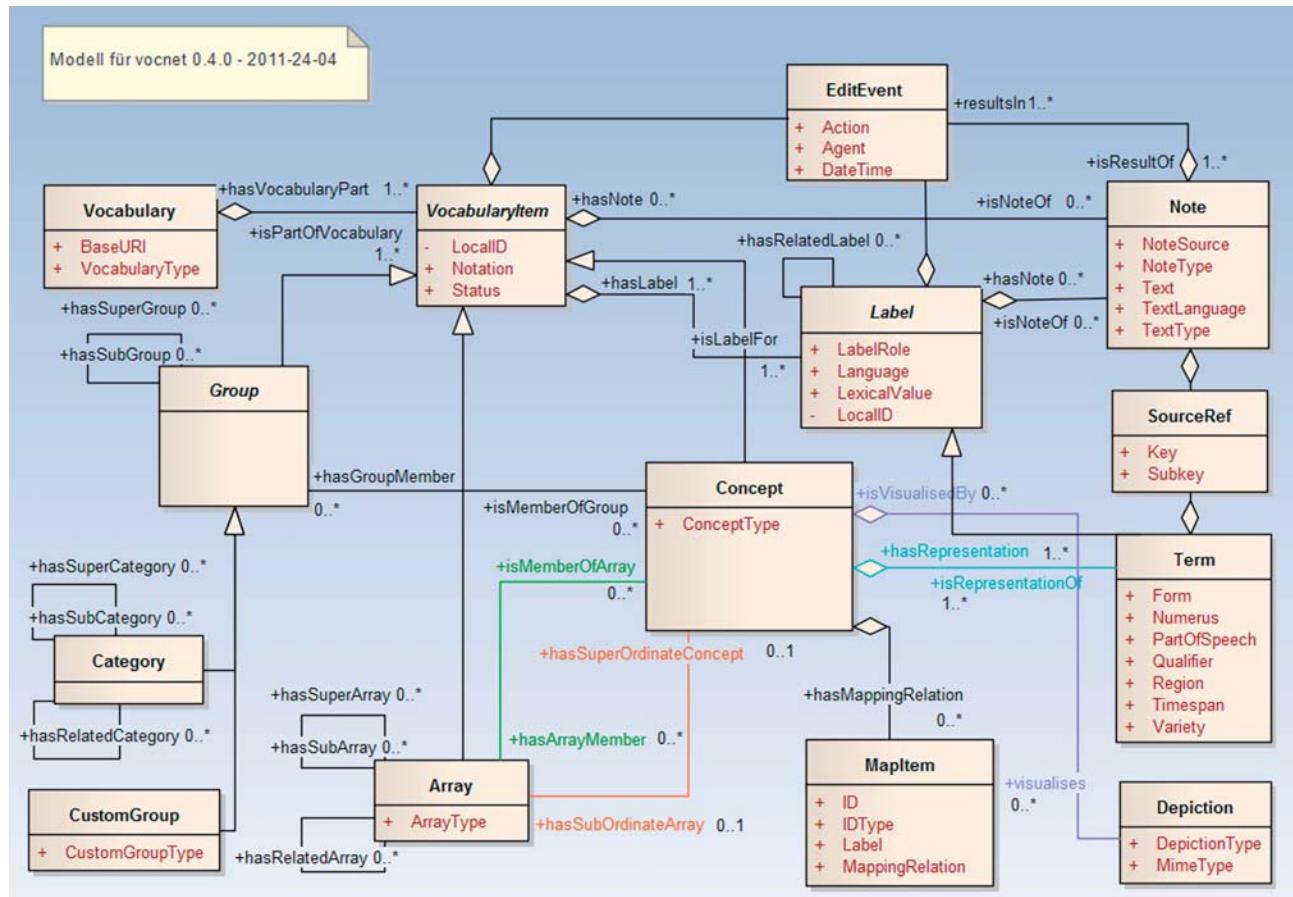


Abb. 2: UML-Diagramm des Austauschformates vocnet.

strukturelle Arbeit erforderlich ist, eine geregelte Abfolge in der Bearbeitung einzuhalten. Die Einarbeitung von Begriffen in das Thesaurusnetz setzt eine Definition der Kandidatenterme durch die Fachleute voraus. Erst auf der Basis von mit Quellen belegten Definitionen ist den Thesaurusredakteuren eine weitere Bearbeitung möglich. Die durch die verschiedenen Institutionen eingereichten Bezeichnungen müssen abgeglichen werden; dabei steht im Vordergrund zu prüfen, ob es sich tatsächlich um Neuvorschläge für Deskriptoren handelt oder ob Synonyme bereits vorhandener Begriffe als Deskriptorkandidaten angemeldet wurden. Es zeigte sich, dass Vorschläge für eine hierarchische Einbettung der Kandidaten nicht ohne Weiteres übernommen werden können, weil diese oftmals auf thematischen Überlegungen des jeweiligen Fachgebiets basieren und nicht den geforderten Regeln der Generalisierung oder Spezialisierung folgen. Die hierarchische Organisation der Begriffe wird in Wortnetz Kultur auf Grundlage der Definitionen in Allgemeinwörterbüchern wie Duden oder Lexika – hier besonders Fachlexika bzw. Fachliteratur vorgenommen, wobei auch

immer die Gemeinsame Normdatei und Spezialthesauri, wie der Art & Architecture Thesaurus, einbezogen werden. Diese Quellen, ergänzt um Informationen aus dem Wortschatz-Portal der Universität Leipzig und Synonymwörterbüchern, werden auch für die Prüfung und Ergänzung von Synonymen herangezogen. Wenn eine Abgrenzung zwischen Begriffen besonders schwierig ist, werden die Merkmale der fraglichen Begriffe mit Hilfe einer Matrix analysiert. In Einzelfällen ist es auch immer wieder erforderlich, externe Fachleute als Berater zu konsultieren.

Das Vorgehen bei der Einarbeitung von Begriffen in die hierarchische Struktur ist abhängig von den eingebrachten Listen der Kandidatenterme. Zunächst werden die als Deskriptoren akzeptierten Begriffe ohne weitere Differenzierung den entsprechenden Oberbegriffen untergeordnet. Wenn absehbar ist, dass die Reihe der Geschwisterbegriffe lang und unübersichtlich wird oder dass eine Gliederung nach Merkmalen in jedem Fall sachlich-wissenschaftlich oder für Indexierung und Suche hilfreich ist, erfolgt eine feinere Strukturierung in „Minifacetten“. Anders ist es, wenn bereits aus der vorliegenden

Kandidatenliste ersichtlich ist, dass eine Feinstruktur erforderlich ist. In diesem Fall werden nach Aufnahme des Oberbegriffs sofort passende Node Labels angelegt. Im Anschluss an die hierarchische Einordnung werden Assoziationsrelationen geknüpft und der Glossareintrag für Wortnetz Kultur vorgenommen. Des Weiteren kann es erforderlich sein, für einzelne Begriffe Scope Notes abzufassen, die den Indexierer bei der Begriffswahl unterstützen. Es besteht die Möglichkeit, Vorschläge für Glossareinträge und Scope Notes im Forum – oft durchaus kontrovers – zu diskutieren. Dies ist wichtig und wünschenswert; so kann abgesichert werden, dass die fachlichen Gesichtspunkte der einzelnen Bereiche berücksichtigt werden.

Die Verwendung der Deskriptoren aus Wortnetz Kultur ist für alle Dienststellen des LVR zur Indexierung ihrer Dokumente und Objekte verbindlich. Alle LVR-Einrichtungen, die den Thesaurus nutzen, entwickeln nach Bedarf eine eigene Systematik, um die WNK-Begriffe unter ihrer Fachsicht ordnen und auffinden zu können. Die eindeutig adressierbaren Deskriptoren (URI) werden von den Anwendern selbst der entsprechenden Systematikstelle zugeordnet. Diese themenbezogenen Sichten haben einen zusätzlichen Vorteil: Die primär nach Form oder Funktion und erst sekundär nach Kontext organisierten Hierarchien trennen thematisch zusammengehörende Deskriptoren voneinander. Unter dem Dach der systematischen Fachsichten werden die Begriffe dann in ihrem thematischen Kontext sichtbar.

Abschließend sind nochmals zentrale Problemstellungen und Herausforderungen aufgeführt, für die in der täglichen Arbeit der Spezialisten aus den verschiedenen Fachgebieten untereinander und in der Kooperation mit den Informationsspezialisten Lösungen gefunden werden müssen:

- Bedeutungen der Begriffe wurden oft nur im Kontext der jeweils eigenen Fachanwendung wahrgenommen und Homonyme dadurch übersehen.
- Zwischen Hierarchierelationen und Assoziationen zu unterscheiden, wurde als schwierig empfunden.
- Die Funktion der Thesaurusrelationen für die Suche wurde nicht immer ausreichend berücksichtigt.
- Es bestanden zu hohe Erwartungen an die Leistungen eines Indexierungsvokabulars; es war nicht von vornherein klar, dass ein Thesaurus nur die Werte für Metadatenelemente und eine paradigmatische Struktur dieser Werte bereitstellen, aber selbst nichts über die syntagmatischen Relationen des zu indexierenden Objektes aussagen kann.

Das bereits erwähnte Forum und die Einbindung externer Experten haben sich als wertvolle Instrumente bei der Lösung einiger der angesprochenen Probleme erwiesen.

8.2.1 Beratung durch externe Experten

Die tägliche Arbeit zeigt, dass es immer wieder erforderlich sein kann, externe Experten für die Definition von Begriffen heranzuziehen, sei es, weil innerhalb des Dezer-nats entsprechende Fachkompetenz nicht vorhanden ist oder weil die Sichtweisen der Fachbereiche so weit auseinandergehen, dass eine Klärung durch Fachleute angebracht ist. Ein besonders schwieriges Problem, das die Definition und Einordnung des Begriffs „Heiligtum“ betraf, konnte nur mit Hilfe der Abteilung für Religionswissenschaften an der Bonner Universität gelöst werden. Bei diesem Treffen wurden von Seiten der Universität auch die Assoziationen zu Sakralbau, „Heiliger Raum“ und „Heiliger Gegenstand“ herausgearbeitet, weiterhin wurden auch Hinweise auf Unterbegriffe und Synonyme gegeben. Bei Unklarheiten zu Begriffen aus der Imkerei und Taubenzucht wurden z. B. Anfragen an den „Deutschen Imker Bund“ und das „Deutsche Taubenhausarchiv“ gerichtet, die nach Einreichung der entsprechenden Begriffe schnell und kompetent beantwortet wurden. Bereitschaft zur Unterstützung der Thesaurusrarbeit signalisierte auch das Rheinische Mühlendokumentationszentrum. Die Zusammenarbeit mit externen Experten ist aber auch dann erforderlich, wenn Fachleute in den Einrichtungen zur Verfügung stehen. Nur so kann abgesichert werden, dass sowohl die Begriffe als auch die Definitionen nicht nur regionalspezifische Gesichtspunkte berücksichtigen. Selbstverständlich gibt es auch Anfragen, die zu keiner endgültigen Lösung führen, aber zu weiteren Überlegungen anregen. Ein Beispiel hierfür ist die Nachfrage, ob

The screenshot shows a software interface for managing a thesaurus. At the top, there's a header with 'Ident-Nr.: wk000245', 'Bezeichnung: Sarkophag (de)', and 'Typ: Begriff'. Below the header are three icons: a green plus sign, a blue pencil, and a red minus sign. The main area is a scrollable list of forum posts. The first post, by 'erstellt' on 'Freitag, 29. Januar 2010, 14:26:35', contains the following information:

- Titel:** Sarkophag nach ...
- Thema:** Node Label
- Text:** Bitte diskutieren, ob eine Untergliederung durch Node Label erwünscht ist. Zeit und Material kommen doch eigentlich aus anderen "Facetten".

Below this, there are tabs for 'Grunddaten', 'Relationen', 'Mappings', 'Anmerkungen', and 'Forum'. The 'Forum' tab is currently selected.

Abb. 3: Forumseintrag in der Thesaurusrsoftware xTree.

es ein hierarchisch strukturiertes Vokabular für „Fahrzeuge“ gibt, beim Verkehrsmuseum Dresden und beim Kraftfahrt-Bundesamt in Flensburg. In beiden Fällen legten die Mitarbeiter die Auffassungen ihrer Dienststellen dar und stellten, soweit vorhanden, Unterlagen zur Verfügung.

8.2.2 Dokumentation

Änderungen an den Begriffssätzen (Status, Änderungsdatum) werden in der Thesaurussoftware xTree protokolliert. Auch das integrierte Forum stellt eine Form der Dokumentation dar, mit deren Hilfe Entscheidungsfindungen nachvollzogen werden können. Schulungsmaterialien in Form von Handreichungen, eine Liste verwendeter Nachschlagewerke und Vokabulare, Thesaurus-Handbuch mit einer genauen Beschreibung der einzelnen Thesauruselemente und Relationen, Erläuterungen zur Indexierung, Ergebnisse der Usability-Tests und organisatorische Informationen werden in Form eines Projekt-Wikis kontinuierlich dokumentiert.

9 Testen – Einsatz von Thesauri im Information Retrieval

Thesauri sind ein wichtiges Hilfsmittel zur Erhöhung der Erschließungsqualität von Informationsbeständen. Der eigentliche Testfall für die Leistungsfähigkeit eines Thesaurus ist der Einsatz im Information Retrieval.¹⁶ Über den aktuellen Prototypen von KuLaDig steht dem Endnutzer bisher nur über die „Erweiterte Suche“ ein kleiner Ausschnitt des Vokabulars zur Verfügung. Im Rahmen von aufgabenbasierten Usability-Tests des Portals KuLaDig mit vier Testpersonen, zwei Informationsspezialisten und zwei Domänenexperten, mit der Thinking-Aloud-Methode im März 2011 wurde auch die Gebrauchstauglichkeit des Vokabulars über die derzeit bereit gestellte Schnittstelle evaluiert. Da in KuLaDig bisher erst ein kleiner Teil der mit Hilfe von Wortnetz Kultur erschlossenen Dokumente zugänglich ist, erlauben die Tests keine repräsentativen Aussagen, liefern aber bereits Hinweise für die Weiterentwicklung der Schnittstelle. Sowohl Laien als auch Domänenexperten kommentierten die Unterstützung der Suche durch ein kontrolliertes Vokabular durchweg

positiv. Die tatsächliche Nutzung der Thesaurusschnittstelle hingegen war für die Testpersonen schwierig.

- Die Testpersonen verstanden erst nach einer Erklärung das Konzept der Fachsichten.
- Die Testpersonen konnten kein kognitives Modell entwickeln, wie die Navigation zwischen den verschiedenen Anzeigefeldern funktioniert. Das Konzept der Schlagwortsuche erschließt sich den Testpersonen nicht. Die Notwendigkeit Suchbegriffe in die „Liste der Suchbegriffe“ zu übernehmen, wird nicht verstanden.
- Während alle Testpersonen unmittelbar verstanden, was mit Oberbegriff und Unterbegriff gemeint ist, wurde die Rolle der Node Labels (z. B. Burg nach Funktion; Burg nach Lage) zunächst nicht verstanden.

Die Entwicklung einer gebrauchstauglichen Nutzerschnittstelle ist eine zentrale Herausforderung für die Weiterentwicklung des Wortnetz Kultur. Erste Verbesserungen können durch eine eindeutigere Platzierung der verschiedenen Navigations- und Informationselemente auf der Seite und selbsterklärende Benennungen (Positionierung von Suchbuttons, Liste, Fachsicht) erzielt werden. Ansatzpunkte für weitere Lösungen zur Einbindung kontrollierter Vokabulare in die Suche, auf die zurückgegriffen werden kann, werden derzeit vor allem im Umfeld des Semantic Search¹⁷ entwickelt:

- Explizite Anweisungen, etwa in Form kontextsensitiver Hilfetexte;
- die Integration von Suche und Facettierung für eine weitere Einschränkung oder Erweiterung der Suchergebnisse;
- Anreicherung der Trefferanzeigen durch ergänzende Kontextinformationen, durch die vor allen den internen Nutzern eine Kontextanzeige angeboten wird, wenn sie sich aus bestimmten Fachanwendungen mit Konzepten beschäftigen.

10 Lessons learned und zukünftige Aufgaben

1. Thesaurusprojekte erfordern einen beträchtlichen zeitlichen, finanziellen und personellen Aufwand. Die einschlägigen Normen weisen nachdrücklich darauf

¹⁶ „The prime application for a thesaurus is in information retrieval.“ ISO 25964-1, S. 12.

¹⁷ Die Ideen gehen zurück auf einen Vortrag von James Kalbach an der HAW Hamburg im Januar 2011. Siehe auch: Spree 2011.

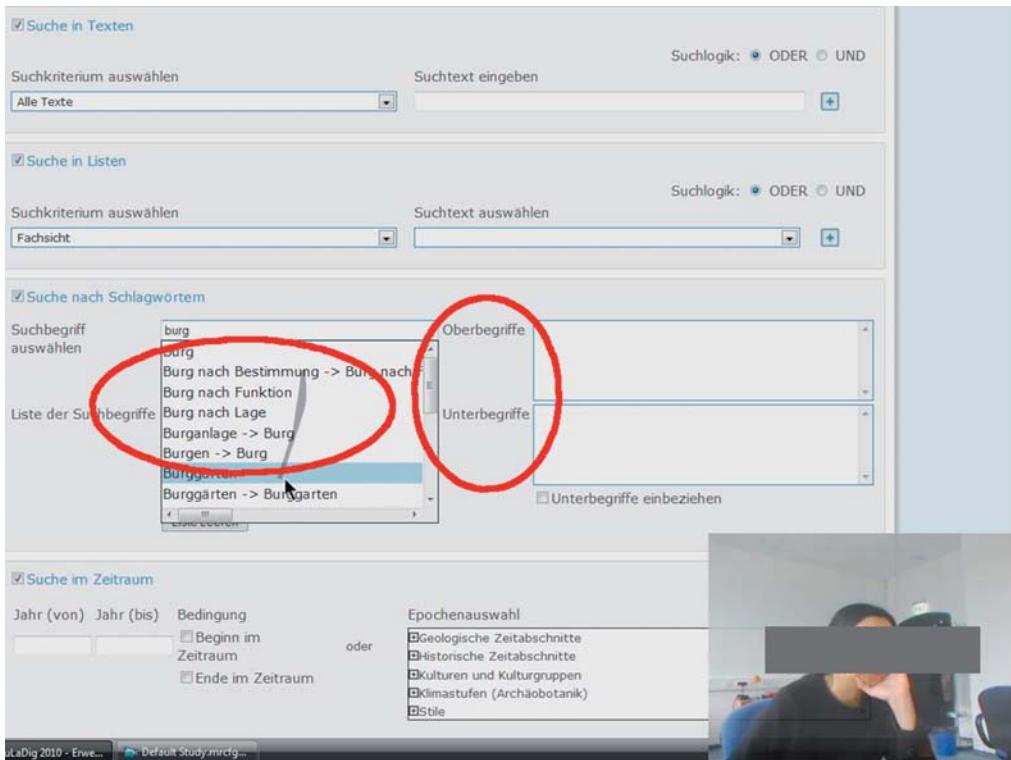


Abb. 4: Screenshot Usability-Test Erweiterte Suche KuLaDig.

hin, dass die Entwicklung eines Thesaurus eine arbeitsintensive, oft viele Jahre dauernde Angelegenheit ist, wenn sie erfolgreich sein soll. Man muss sich vor der Entscheidung, einen Thesaurus aufzubauen, sehr bewusst sein, dass erhebliche Ressourcen und Personalkapazitäten durch ein solches Projekt gebunden werden. Die Erfahrungen im Wortnetz Kultur bestätigen diese Annahme. Hier ist es die Aufgabe der Informationsspezialisten, mit Erfahrungswerten zum zeitlichen, finanziellen und personellen Ressourcenbedarf den Auftraggebern zu verdeutlichen, worauf sie sich langfristig einlassen.

2. *Kontinuierliche strukturierte Kommunikation zwischen allen Projektbeteiligten ist eine wesentliche Erfolgsvoraussetzung für die kollaborative Erschließung von Fachinformationen.* Innerhalb des Projekts Wortnetz Kultur bestand von Beginn an eine intensive Abstimmung zwischen Domänenexperten, Informationsspezialisten und den Thesaurussentwicklern. Das Forum innerhalb der Thesaurussoftware hat wesentlich dazu beigetragen, dass alle Beteiligten einen einheitlichen Diskussionsstand hatten und ihre Sichtweise angemessen kommunizieren konnten.
3. *Die Einbeziehung von externen Experten erhöht die Qualität des Vokabulars.* Die tägliche Arbeit zeigt,

dass es immer wieder erforderlich sein kann, externe Experten für die Definition von Begriffen heranzuziehen, sei es, weil innerhalb des Dezernats entsprechende Fachkompetenz nicht vorhanden ist oder weil die Sichtweisen der Fachbereiche so weit auseinandergehen, dass eine Klärung durch entsprechende Fachleute angebracht ist.

4. *Die Menschen müssen die Dinge sehen, damit sie sie verstehen können.* Best-Practice-Beispiele eignen sich gut, um mit Auftraggebern und Anwendern über die konkreten Anforderungen ins Gespräch zu kommen.
5. *Das Testen von Konzepten oder Modellen ist schwierig.* Als praktisch schwer umsetzbar erwies sich das Testen der Tragfähigkeit des Thesauruskonzepts. Es stellte sich heraus, dass das Vorhandensein einer kritischen Masse korrekter Konzepte Voraussetzung für aussagekräftige Tests ist. Ein möglicher Lösungsansatz wird im Rapid Prototyping gesehen, bei dem die Anwender in die unterschiedlichen Entwicklungsschritte, beispielsweise die Nachvollziehbarkeit der Kategoriensysteme oder der Facettierung, durch partizipative Methoden wie Card Sorting einbezogen werden. Auch die Entwicklung von praxisnahen, prototypischen Anwendungsbeispielen oder Visualisierungen ist förderlich für den Austauschprozess.

Die nächsten Schritte in der kooperativen Entwicklung des Thesaurus werden sein:

- Entwicklung von Kriterien und Regeln für die Begriffszerlegung
- Präzisierung der Äquivalenzrelation für Wortnetz Kultur
- Definition regelbasierter Assoziationsrelationen für Wortnetz Kultur
- Optimierung der Vokabularstruktur für facettenbasierte Navigation
- Workflow-Management für verteiltes Arbeiten, wenn das Vokabular bereits zur Indexierung eingesetzt wird.

11 Fazit

Will man die sich mittlerweile über mehr als vier Jahre entwickelnde Kooperation zwischen dem LVR und dem Department Information in Hinblick auf Forschung an Fachhochschulen bewerten, so wird deutlich, dass bisher der Aspekt der praktischen Entwicklungsarbeit eindeutig im Vordergrund stand. Dies ist auch für ein Thesaurus-Projekt wie Wortnetz Kultur ein angemessener Ansatz. Aus der Entwicklungsarbeit heraus ergaben sich verschiedene über das eigentliche Kooperationsprojekt hinausgehende Forschungsfragen, die sich aber unter den gegebenen personellen und zeitlichen Rahmenbedingungen nicht bearbeiten ließen. Nur im Rahmen eines umfangreicheren vergleichend angelegten Forschungsprojekts ließe sich beispielsweise erforschen, wie sich eine regelbasierte Definition von Assoziationsrelationen umsetzen lässt und unter welchen Bedingungen es sachlich sinnvoll ist, einen Begriff als möglichen Teil eines Ganzen auszuweisen. Auch ein ursprünglich geplantes unabhängiges Monitoring des Gesamtprojektes, das es erlaubt hätte, Schwachstellen im Projektaufbau zu identifizieren und ein für ähnliche Projekte nachnutzbares Phasenmodell abzuleiten, war innerhalb des begrenzten Rahmens nicht umsetzbar.

Nur durch die Einbindung der Studierenden in allen Phasen des Projektes in Form von Seminararbeiten, Forschungsprojekten und Abschlussarbeiten war es möglich, innerhalb des Entwicklungsprojektes auch über den Tellerrand hinauszuschauen und etwa alternative Vokabularmodelle oder Anwendungsszenarien vergleichend zu betrachten. An seine Grenzen stößt die Leistungsfähigkeit dieses Ansatzes dort, wo ein Forschungsvorhaben spezialisiertes Vorwissen und intensive Einarbeitung (systematische Untersuchung des Wortnetz Kultur im Retrieval und Entwicklung von Prototypen) voraussetzt. Das verlangt

eine personelle Kontinuität, die angesichts der von Semester zu Semester wechselnden Studierendengruppen nicht gewährleistet werden kann. Ohne wissenschaftliche Mitarbeiter sowie die Kooperation mit weiteren informationswissenschaftlich erfahrenen Kollegen stoßen informationswissenschaftliche Forschungsprojekte an schwerpunktmäßig ingenieurwissenschaftlich orientierten Hochschulen wie der HAW an ihre Grenzen.

12 Postskriptum

Die Frage, wie ein Sarkophag je nach Kontext einmal als Befund und einmal als Fund abgebildet werden kann, hatte die Zusammenarbeit zwischen der Hochschule und dem Landschaftsverband angestoßen. Doch gerade diese Frage ist mit dem Instrument „Thesaurus“ allein nicht zu lösen. Frau Kaljaj hatte bereits in ihrer Diplomarbeit gezeigt, dass kontextabhängige Rollen, wie z. B. „Befund“ oder „Fund“, nicht sinnvoll durch Thesaurussrelationen darstellbar sind. Vielmehr lassen sich solche syntagmatischen Zusammenhänge erst auf der Ebene des Metadatenmodells abbilden.

Literatur

- Aitchison, Jean; Bawden, David; Gilchrist, Alan: Thesaurus construction: A practical manual. 4th ed. London: Fitzroy Dearborn Publishers, 2000.
- Bechhofer, Jean; Goble, Carole: Thesaurus construction through knowledge representation, 2001. Online: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.134.9575>.
- Bekiaris, Chr.; Doerr, Martin: Documentation and Reasoning on Parts and Potential Wholes. Technical Report 260, ICS-FORTH: Heraklion, 1999. Online: http://www.ics.forth.gr/isl/publications/paperlink/parts_and_wholes_TR.pdf [22. 11. 2011].
- Broughton, Vanda: Essential thesaurus construction. London: Facet Publishing, 2006.
- British Standard – Draft for Development DD8723-5:2008. Online: <http://schemas.bs8723.org/>.
- Budin, Gerhard; Swertz, Christian; Mitgutsch, Konstantin; International Society for Knowledge Organization: Knowledge organization for a global learning society: Proceedings of the Ninth International ISKO Conference, 4–7 July 2006, Vienna, Austria: Advances in knowledge organization. 10. Würzburg: Ergon-Verl., 2006.
- Crofts, Nick; Doerr, Martin; Gill, Tony; Stead, Stephen; Stiff, Matthew (Hrsg.): Definition of the CIDOC Conceptual Reference Model, 2011. Online: http://www.cidoc-crm.org/docs/cidoc_crm_version_5.0.3.pdf [22. 11. 2011].

- Dellschaft, Klaas; Hachenberg, Christian: Repräsentation von Wissensorganisationssystemen (KOS) im Semantic Web: Ein Best Practice Guide. Universität Koblenz-Landau, 2011. Online: http://www.kim-forum.org/material/pdf/BPG_Repraesentation_von_KOS_im_Semantic_Web.pdf [20. 11. 2011].
- Doerr, Martin: Semantic Problems of Thesaurus Mapping. In: Journal of Digital Information (2001), Vol 1, No 8. Online: <http://journals.tdl.org/jodi/article/viewArticle/31/32> [3. 11. 2011].
- Doerr, Martin; Kritsotaki, Athina; Stead, Stephen: Which Period is it? A Methodology to Create Thesauri of Historical Periods, [2005]. Online: http://www.forth.gr/ics/isl/publications/paper-link/CIDOCpaper1_Doerr.pdf [10. 11. 2011].
- Elin, Jacob, Elin: Classification and Categorization: A Difference that Makes a Difference. In: Library Trends, Bd. 52, Nr. 3, S. 515–540. University of Illinois, 2004. Online: <https://www.ideals.illinois.edu/bitstream/handle/2142/1686/Jacob515540.pdf> [21. 11. 2011].
- Fischer, Dietrich: From Thesauri towards Ontologies? GMD/PSI, 1998. Online: http://www.ipsi.fraunhofer.de/topas/publications/Fischer_1998.pdf [21. 11. 2011].
- Gnoli, Claudio ; Mazzocchi, Fulvio: International Society for Knowledge Organization (Hrsg.): Paradigms and conceptual systems in knowledge organization. Proceedings of the eleventh international ISKO conference 23–26 february 2010 Rome, Italy: Advances in Knowledge Organization. 12. Würzburg: Ergon, 2010.
- Harpring, Patricia: Introduction to Controlled Vocabularies: Terminology for Art, Architecture and Other Cultural Works. Los Angeles: Getty Publications, 2010. Online: http://www.getty.edu/research/publications/electronic_publications/intro_controlled_vocab/index.html [22. 11. 2011].
- Hyvönen, Eero: Preventing Interoperability Problems Instead of Solving Them. Semantic Web Journal (2010), Nr. 2–3. Online: <http://www.semantic-web-journal.net/sites/default/files/swj33.pdf> [22. 11. 2011].
- Isaac, Antoine; Schlobach, Stefan; Matthezing, Henk; Zinn, Claus: Integrated access to cultural heritage resources through representation and alignment of controlled vocabularies. In: Library Review (2008), Nr. 57, S. 3, 187–199.
- ISO 21127:2006: Information and documentation – A reference ontology for the interchange of cultural heritage information.
- ISO 25964-1:2011: Information and documentation – Thesauri and interoperability with other vocabularies. Thesauri for information retrieval.
- Janßen-Schnabel, Elke: KuLaDig steht jetzt im Netz. In: LVR-ADR, Denkmalpflege im Rheinland (2011), Nr. 1. Online: http://www.lvr.de/media/wwwlvrde/kultur/kulturlandschaft/kuladignw/dokumente_189/dir2011_kuladig.pdf [16. 10. 2011].
- Kaljaj, Magdalena: Archäologische Fachinformation bearbeiten : Konzept eines Thesaurus für das Rheinische Amt für Bodendenkmalpflege. Hamburg, Hochschule für Angewandte Wissenschaften, Fakultät DMI, Dipl.-Arb., 2008. Online: http://opus.haw-hamburg.de/frontdoor.php?source_opus=580&la=de [23. 11. 2011].
- Knöchel, Franz-Josef: KuLaDig – Kultur. Landschaft. Digital: Zur Partnerschaft des Rheinischen Vereins mit dem Informationssystem über die Historische Kulturlandschaft und das landschaftliche Kulturelle Erbe. In: Rheinische Heimatpflege (2011), Nr. 1, S. 51–58, Online: <http://www.lvr.de/media/wwwlvrde/>
- kultur/kulturlandschaft/kuladignw/dokumente_189/20110824_kuladig_rndl.pdf
- [22. 11. 2011].
- Lambe, Patrick: Organising knowledge: Taxonomies, knowledge and organisational effectiveness. Chandos knowledge management series. Oxford: Chandos, 2007.
- Lindenthal, Jutta: Kontrollierte Vokabulare für das Information Retrieval im Landschaftsverband Rheinland. Vortrag auf dem Berliner Herbsttreffen zur Museumsdokumentation am 13. 10. 2010. Online: http://www.museumsbund.de/fileadmin/fg_doku/termine/2010_Oktober_Herbsttagung/Beitraege/Lindenthal_2010-10-13_Druck.pdf [10. 10. 2011].
- Lindenthal, Jutta: Wortnetz Kultur: Allgemeine Einführung in den Aufbau des Begriffssystems. Version 0.7, 2009-08-10.
- Lindenthal, Jutta: Wortnetz Kultur: Einführung in das Vokabularmodell (Entwurf 0.1). Version 0.1, 2009-08-17.
- Mäkelä, Eetu: View-Based User Interfaces for the Semantic Web. Aalto University, School of Science and Technology, Espoo, 2010. Online: <http://lib.tkk.fi/Diss/2010/isbn9789526034478/isbn9789526034478.pdf> [20. 11. 2011].
- Molholt, Pat: The Art and Architecture Thesaurus: Controlling Relationships through Rules and Structures. In: Bean, Carol A.; Green, Rebecca (ed.): Relationships in the organization of knowledge. Dordrecht: Kluver, 2001.
- Pinna, Loredana; Wessels, Wiebke: Literaturbericht – Nationale und internationale Ansätze zu Wissensorganisationssystemen (KOS) im Bereich Archäologie und Kulturerbe. Webprojekte, Standards und Vokabulare. 2009. – Forschungs- und Praxisprojekt im Masterstudiengang Informationswissenschaft und -management an der Hochschule für Angewandte Wissenschaften Hamburg.
- Pinna, Loredana; Wessels, Wiebke: Webprojekte, Standards, Vokabulare – State of the Art der Wissensorganisationssysteme für Archäologie und Kulturerbe. In: HAW Abstracts in Information Science and Services (2009), Nr. 1, August 2009, <http://www.bui.haw-hamburg.de/fileadmin/haiss/2009PinnaWessels.pdf> [27. 10. 2011].
- Priss, Uta; Jacob, Elin: Utilizing Faceted Structures for Information Systems Design. Indiana University Bloomington, 1999. Online: <http://www.upriss.org.uk/papers/asis99.pdf> [22. 11. 2011].
- Roe, Sandra K.; Thomas, Alan R. (Hrsg.): The Thesaurus: Review, Renaissance, and Revision. New York; London; Victoria: Haworth Information Press, 2004.
- Riksantikvarieämbetet (Hrsg.): Riksantikvarieämbetet – Fornsök Online: <http://www.fmis.raa.se/cocoon/fornsoek/search.html> [10. 11. 2011].
- Shearer, James R.: A Practical Exercise in Building a Thesaurus. In: Roe, Sandra K.; Thomas, Alan R. (Hrsg.): The Thesaurus. Review, Renaissance, and Revision. New York; London; Victoria: Haworth Information Press, 2004, S. 35–56.
- Spree, Ulrike; Feißt, Nadine; Lühr, Anneke; Piesztal, Beate; Schroeder, Nina; Wollschläger, Patricia: Semantic Search : State-of-the-Art-Überblick zu semantischen Suchlösungen im WWW. In: Lewandowski, Dirk: Handbuch Internetsuchmaschinen 2. Heidelberg: AKA Verlag, 2011, S. 77–108.
- Uddin, Nasir M.; Janacek, Paul: Performance and usability testing of multidimensional taxonomy in website search and navigation. In: Performance, measurements and metrics, (2007), Nr. 8, S. 1, 18–33.

- Wersig, Gernot: Thesaurus-Leitfaden: Eine Einführung in das Thesaurus-Prinzip in Theorie und Praxis. München u. a.: Verl. Dok., 1978. (DGD-Schriftenreihe; Bd. 8).
- Wikisaurus: Wiki des Standard-Thesaurus Wirtschaft (STW): Thesaurus-Struktur. Online: <http://zbw.eu/wikis/wikisaurus/index.php?n>Main.Thesaurus-Struktur> [22. 11. 2011].
- Will, Leonard: A data model for BS8723 thesaurus structure. Vortrag gehalten auf der Tagung der ISKO UK, 2008. Online: http://www.iskouk.org/SKOS_July2008.htm [22. 11. 2011].
- Zhitomirsky-Geffet, Maayan; Bar-Ilan, Judit; Miller, Snunit S.: A generic framework for collaborative multi-perspective ontology acquisition. In: Online Information Review (2010), Nr. 34, S. 1, 145–159.

Eingereicht am 30. November 2011, angenommen am 18. Dezember 2011.

Prof. Dr. Ulrike Spree

Hochschule für Angewandte Wissenschaften – Hamburg University of Applied Sciences – Fakultät Design, Medien und Information Department Information Finkenau 35 20081 Hamburg ulrike.spree@haw-hamburg.de

Prof. Dr. Ulrike Spree ist seit 1999 Professorin für das Fachgebiet „Informationsdienstleistung und -vermittlung in der Mediendokumentation“ an der Hochschule für Angewandte Wissenschaften Hamburg.

Jutta Lindenthal

Hövelnstraße 11
23566 Lübeck
jutta.lindenthal@googlemail.com

Jutta Lindenthal berät freiberuflich Kultureinrichtungen bei der Entwicklung von Informationssystemen. Sie ist Mitglied der Arbeitsgruppe ISO 25964 (Thesauri and interoperability with other vocabularies) und gehört der Special Interest Group des CIDOC CRM (ISO 21127) sowie der DGI-Arbeitsgruppe Komitee Terminologie und Sprachfragen an.

Dr. Andje Knaack

Landschaftsverband Rheinland
– Wortnetz Kultur –
Amt für Bodendenkmalpflege im Rheinland
Endenicher Straße 133
53115 Bonn
andje.knaack@lvr.de

Dr. Andje Knaack ist Wissenschaftliche Mitarbeiterin im Landschaftsverband Rheinland. Sie ist ausgebildete Museologin und Prähistorikerin und Mitglied des Teams „Projekt Wortnetz Kultur“.