

Inhaltsverzeichnis

Inhaltsverzeichnis	1
1 Einleitung	3
2 Grundlegende Begriffe	4
2.1 Informationsmanagement.....	4
2.2 Metadaten.....	5
2.3 Metasprachen.....	6
2.3.1 SGML.....	6
2.3.2 XML.....	7
2.4 Semantik und Semantische Netze.....	7
2.5 Vision des Semantic Web.....	8
2.6 Ontologien.....	10
2.6.1 Was ist eine Ontologie?.....	11
2.6.2 Ontologien in der praktischen Anwendung.....	12
2.6.3 Ontologie-Entwicklung.....	13
3 XML Topic Maps	13
3.1 ISO/IEC 13250.....	13
3.2 Der XML Topic Maps (XTM)-Standard.....	14
3.3 Topics.....	14
3.4 Topic Typen.....	15
3.5 Topic Occurrences.....	16
3.6 Assoziationen.....	16
3.7 Scopes.....	17
3.8 Merging.....	17
4 Topic Map Tools	17
5 Das WebInfoService-Projekt und Topic Maps	19
5.1 Überblick über das Projekt.....	19
5.1.1 DiBIS.....	19
5.1.2 Ziele des Projekts.....	20
5.1.3 Projektverlauf.....	21
5.2 Ontologie im Projekt.....	23

5.3	Beispiel: Entwicklung eines Teils der DiBIS-Ontologie mit Topic Maps	23
6	Zusammenfassung	28
7	Literatur	29
8	Anhang	32

1 Einleitung

Die vorliegende Baccalaureatsarbeit ist im Zusammenhang mit dem Projekt "Kontextualisierung bei webbasierten InfoServices" entstanden, welches von Dr. Ralf Klischewski und Martti Jeenicke im WS 2002/03 an der Universität Hamburg veranstaltet wurde. Sie beschäftigt sich mit der Abbildung und Modellierung von Wissensstrukturen mit Topic Maps.

Topic Map ist ein auf XML basierender Standard (ISO/IEC bzw. TopicMaps.Org) zum Austausch von Wissensstrukturen, mit dem sich begriffsbasierte Wissensnetze realisieren lassen. Eine Topic Map besteht aus einer Sammlung von Themen (englisch: topics), ihren Beziehungen untereinander und den jeweiligen Themen zugeordneten Informationsquellen. Diese Technologie baut auf Konzepten der Wissensstrukturierung des Bibliothekswesens wie Indexen und Glossaren auf.

Im letzten Jahrzehnt ist es zu einer starken Zunahme an elektronisch gespeicherten Informationsbeständen gekommen, wodurch die Erstellung von Ordnungen und Klassifizierungen, d.h. das Informationsmanagement, immer mehr an Bedeutung gewonnen hat. Besonders die Informationsbestände des Internets werden hierbei betrachtet, und bei der Umsetzung der im Rahmen der Forschung neu entstandenen Technologien, wie unter anderem Topic Maps, sehen manche hier die Vision des "Semantic Web" vor sich, bei dem Agenten (gemeint sind "intelligente Agenten", das heißt Programme, die einen Prozess automatisch abwickeln und eventuell mit anderen Agenten kommunizieren, stellvertretend für einen Menschen) in der Lage sind, die Ordnungsstrukturen zu verstehen. In diesem Zusammenhang haben sich neue Web-Technologien gebildet.

Im Rahmen des Projektes wurden diese Semantic Web Technologien erprobt, um einen Prototyp für einen Ausschnitt des e-Government-Infoservice der Stadt Hamburg zu entwickeln. Dabei sollte das webbasierte Bürgerinformationssystem dem Bürger durch Erschließen des Kontextes seiner Anfragen speziell auf ihn zugeschnittene Informationen anbieten können. Der Nutzerkontext wird dazu während einer Websession vom Computer erschlossen.

Im Projekt wurde als Technologie das Jena-Software-Toolkit mit DAML+OIL verwendet. Topic Maps wurden am Anfang des Projekts als Alternative diskutiert.

Welche der Web Technologien für welche Anwendung im Informationsmanagement am geeignetsten sind, ist ein Thema, das momentan an verschiedenen deutschen Hochschulen untersucht wird.

Ein solcher Vergleich wird in dieser Arbeit nicht vorgenommen, sondern es wird ein Überblick über Topic Maps und ihre theoretischen Verwendungsmöglichkeiten für das Projekt erörtert.

Topic Maps lassen die Modellierung einer Vielzahl von Beziehungen zwischen Themen zu und sind besonders gut geeignet, weit gefasste Metadatenetze¹ darzustellen, was bei e-Government-Services von Bedeutung ist.

Ich werde im ersten Teil der Baccalaureatsarbeit die wichtigsten Konzepte von Topic Maps vorstellen.

Im zweiten Teil der Arbeit werde ich auf das WebInfoService Projekt eingehen.

Da Topic Maps im Zusammenhang mit den Begriffen Ontologie, Semantic Web und Informationsmanagement verwandt werden, sollen diese Begriffe zunächst erklärt werden.

2 Grundlegende Begriffe

2.1 Informationsmanagement

Informationsmanagement und Wissensmanagement (knowledge management) sind verwandte Begriffe. Von Wissensmanagement wird im Rahmen von Unternehmensstrategien gesprochen, womit gemeint ist, dass in großen Unternehmen die Wissensbestände, die ein wichtiger Produktions- und Wettbewerbsfaktor sind, einfach zugänglich gemacht werden sollten. „Viele internationale Unternehmen versprechen sich einen Vorteil durch den Aufbau einer strukturierten deklarativen Repräsentation ihres gesamten Wissens. Die deklarative Repräsentation führt dazu, dass das Wissen in unterschiedlichen Unternehmensfunktionen wieder verwendet werden kann.“ (R. G. Smith, zitiert nach J. Dorn und G. Gottlob in [Rechenberg 2002] S.988).

Beim Informationsmanagement werden Informationsressourcen betrachtet, im allgemeinen Bücher, Artikel, Daten auf lokalen oder entfernten Computern, also Informationen, die dem Unternehmen schon zur Verfügung stehen.

Klischewski versteht unter Informationsmanagement “an interdisciplinary field which draws on and combines skills and resources from librarianship and information science, information technology, records management, archives and general management. Its focus is information as a resource, independently of the physical form in which it occurs (e.g. books and articles, data stored on local or remote computers, microforms, audiovisual media, etc. - sometimes even the information in people`s head)” [Klischewski et al. 03].

¹ Metadaten: Daten über Daten; Metadatenetz = ein Netz von Daten (d. h. die Daten stehen in Verbindung), das andere Daten beschreibt

Hauptanliegen ist die effiziente Dokumentensuche (englisch: information retrieval), die vor allem bei heterogenen und von verschiedenen Institutionen herausgebrachten Dokumenten aufwendig ist.

Ein Beispiel für Informationsmanagement war das Projekt, da dort alle Informationsressourcen, die zu einer Behörde gehören, kategorisiert und verfügbar gemacht werden sollten. Es wurden nur webbasierte Informationsressourcen betrachtet, d.h. digitale Dokumente (wie Websites, PDF-Dateien, Grafik-Dateien usw.), die eine eindeutige Webadresse URI (Unified Resource Identifier) besitzen. Ein URI beinhaltet alle Internetadressen, die dem aktuellen oder einem zukünftigen Adressierungsschema entsprechen².

Alle Konzepte, die im Informationsmanagement verwandt werden, benötigen Metadaten.

2.2 Metadaten

Wenn Daten strukturiert werden, muss die Bedeutung der gespeicherten Daten beschrieben und abgelegt werden. Außer den Daten selbst (die man genauer Primärdaten nennt) gibt es also auch „Daten über Daten“, sog. Metadaten. Damit nun keine endlose Kette von Daten und ihren Beschreibungen entsteht (Metadaten, Metametadaten, Metametametadaten), muss die Semantik der Metadaten ein für allemal festgelegt werden.

Beispielsweise bilden der Titel, der Betreff, der Autor und die Dateigröße (die Dateiattribute) die Metadaten zu einer Datei [Computer 2000].

Mit Hilfe von Metadaten können verschiedenste nicht nur lokale, sondern auch über das Netzwerk zugängliche Informationsressourcen selbst, wie etwa Textdokumente, Musikstücke, Graphiken, Tabellen in Datenbanken näher beschrieben werden. Die Metadaten können bei der Identifikation, der Zuordnung zu einem bestimmten Kontext und bei der Lokalisation helfen. Im Bibliothekswesen sind bibliographische Informationen, Zusammenfassungen oder Schlüsselwörter aus einem standardisierten Vokabular Metadaten. Sie werden auch benutzt, um Zugriffsrechte festzulegen. Im World Wide Web benutzen Suchmaschinen Metadaten, um die Suchergebnisse zu sortieren [Rollet 00]. Yahoo und andere Suchmaschinen können als Metadaten-Webseiten betrachtet werden, denn sie geben Auskunft über andere Webseiten. Yahoo ist ein nach Sachgebieten gegliederter Katalog (z. B. Sport, Politik, Bücher), womit es aber ziemlich allein dasteht, wohingegen andere Suchdienste folgendermaßen benutzt werden (Volltextsuchmaschine): An die Suchmaschine wird die Anfrage gestellt: „Welche Seite enthält diese Zeichenkette?“ Da diese Metainformation nicht sehr spezifisch ist, wird die Volltextsuche allgemein nicht als optimal angesehen. Besser wären Metadaten, die Themen filtern.

² URI (Uniform Resource Identifier) sind URLs (Uniform Resource Locators) oder URNs (Uniform Resource Names) (vgl. [Young 02], S. 91).

Eine Schwierigkeit besteht darin, die Semantik der Metadaten zu definieren. Das Topic Maps Paradigma geht davon aus, dass man nicht bei der Informationsressource beginnen und diese durch Metadaten beschreiben, sondern ein Modell erstellen sollte, in dem Abstrakta für die Gegenstände, von denen die Informationsressourcen handeln, gebildet werden, und zwar in Form eines begriffsbasierten Netzwerkes. Es gibt dann nicht mehr zu jeder Informationsressource einzelne Metadaten, sondern zwei Ebenen: Die Ebene der Informationsressourcen und über dieser die des begriffsbasierten Netzwerkes als Metaebene (vgl. [Newcomb03]).

2.3 Metasprachen

Im Folgenden möchte ich einen Überblick über Sprachen zur Beschreibung von Metadaten geben.

Mit SGML nahmen die neuen Semantic Web Technologien ihren Anfang, und aus dieser Sprache haben sich HTML und XML entwickelt.

2.3.1 SGML

SGML (Standard Generalized Markup Language) ist eine allgemeine Auszeichnungssprache.

Sie fügt einem Dokument Metadaten in Form von *Markups* (englisch auch tags genannt, tags auf deutsch: Etikette) hinzu. Die Markups ermöglichen automatische Verarbeitungsprozesse. SGML ermöglicht es, Informationen zu Dokumenten zwischen verschiedenen Programmen und Computer-Plattformen auszutauschen und sie wieder zu verwenden. Nötig ist dazu bloß eine SGML verarbeitende Software.

„SGML definiert eine Basis-Syntax, erlaubt jedoch auch, eigene Elemente zu erstellen (darum der Begriff *Generalized*“ ([Young02], S 32). Durch die strikte Trennung von Inhalt des Dokuments und seiner Struktur kann dasselbe Dokument von Programmen zur Darstellung und anderen Anwendungen verschieden verwandt werden. Wenn ein Softwareprogramm die Markups interpretieren kann („wichtig findet“), dann werden sie benutzt, andernfalls von dem Programm beim Lesen ignoriert.

„SGML identifiziert Elemente des Dokumentes, wie Titel, Paragraph, Tabelle, Kapitel etc. als voneinander abgrenzbare Objekte und erlaubt dem Benutzer die Beziehungen zwischen den Objekten festzulegen und damit dem Dokument eine Struktur zu verleihen. Die Beziehungen zwischen den Elementen werden in einer Document Type Definition festgelegt.“ (vgl. [FOLDOC], Stichwort SGML).

Document Type Definitions sind Bestandteil sowohl von SGML, als auch von dem von XML abgeleiteten XML. Sie werden im folgenden Abschnitt noch einmal kurz angesprochen.

2.3.2 XML

XML ist ein extrem einfacher Dialekt von SGML. XML (**extensible Markup Language**) wurde von der XML-Arbeitsgruppe des World-Wide-Web-Consortium (W3C)³ definiert. Das Ziel der Arbeitsgruppe war es, „zu ermöglichen, generisches SGML in der Weise über das Web auszuliefern, zu empfangen und zu verarbeiten, wie es jetzt mit HTML möglich ist. XML wurde entworfen, um eine einfache Implementierung und Zusammenarbeit sowohl mit SGML als auch mit HTML zu gewährleisten“ ([XML-Spezifikation98], zitiert nach [Young02], S. 25).

Die XML-Syntax ist einfacher als die von SGML. Dies hat folgende Vorteile: Es erleichtert die Erstellung und Lesbarkeit von XML-Dokumenten. Die Entwicklung von Browsern und anderen Programmen, die XML-Dokumente verarbeiten können, wird ebenfalls erleichtert (vgl. [Young02], S 34).

In einer **Document Type Definition (DTD)** wird mit einer besonderen Syntax, im wesentlichen durch reguläre Ausdrücke, der Inhalt eines zukünftigen Dokuments im Voraus beschrieben. Reguläre Ausdrücke stammen aus der Mathematik und sind ein Konstruktionsverfahren für Sprachen (hier für Strukturen). Ein XML-Validator ist ein Programm, das den korrekten Aufbau eines Dokuments anhand der DTD überprüft. Entspricht ein Dokument in seinem Aufbau der DTD, ist es *gültig* (validiert).

Eine andere Art, den Dokumenttyp festzulegen, sind XML Schemas. Sie haben keine besondere Syntax wie die DTD, sondern werden selbst in XML verfaßt. Sie sind ausdrucksstärker als DTDs. Mit ihnen können u. a. Datentypen festgelegt, Datentypen von einander abgeleitet und komplexe Datentypen gebildet werden (vgl. [Young02], S.177).

Im folgenden Teil der Arbeit werden „high-level“⁴ Metasprachen, die auf XML und SGML basieren, betrachtet (insbesondere Topic Maps). Der Topic Map Standard ist ein Metamodell, denn er gibt die Konstruktionsregeln für die Erstellung von Domänenmodellen vor.

2.4 Semantik und Semantische Netze

Semantik heißt Bedeutung, Inhalt (eines Wortes, Satzes oder Textes) [Duden 97].

³ Das World Wide Web Consortium wurde im Oktober 1994 gegründet, um die Entwicklung des World Wide Web durch die Entwicklung allgemeingültiger Standards voranzutreiben.
<http://www.w3.org>

⁴ High-level: Der Begriff wird normalerweise für Programmiersprachen angewandt und bezeichnet Sprachen, die einen gewissen Abstraktionsgrad aufweisen im Gegensatz zu Assembler-Sprachen (low-level-Sprachen) (vgl. [FOLDOC]).

Der Begriff ist Informatikern aus der Programmierung bekannt und bezeichnet die Beziehung zwischen Wörtern und Symbolen und ihren zugeordneten Bedeutungen. Programmiersprachen sind bestimmten semantischen Regeln unterworfen.

Im KI-Bereich (KI = Künstliche Intelligenz) werden dabei nicht die Beziehungen zwischen Wörtern betrachtet, sondern die Verhältnisse von Objekten zueinander. Man versucht diese in „menschenähnlicher“ Weise darzustellen [Computer 2000]. Hiermit wird angestrebt, dass eine Maschine in der Lage ist, die Objekte und Beziehungen ähnlich zu behandeln und zu „erkennen“ wie der Mensch. Dies wird auch konzeptuelle Semantik genannt.

In der Linguistik steht die Erforschung der Bedeutungen von Wörtern, Sätzen und Texten (also von Zeichen und Zeichenfolgen) und ihrer Wechselbeziehungen im Vordergrund (vgl. [Brockhaus88], Stichwort: Semantik). Dort werden semantische Netze zur Beschreibung der semantischen Struktur von Sätzen eingesetzt. Sie entwickelten sich ursprünglich als psychologische Modelle für kognitive Strukturen, wie z.B. Assoziationen im Gedächtnis.

In Anlehnung an Baumüller (vgl. [Baumüller02], S.13) handelt es sich bei einem semantischen Netz um einen gerichteten Graphen, wobei die Knoten die Objekte (Begriffe, Konzepte) repräsentieren und die gerichteten Kanten (engl. arcs oder links) für die Beziehungen (Relationen) zwischen den Objekten stehen. Die Kanten sind typisiert, wodurch zusätzliche Informationen in das Modell integriert werden. Die an den Knoten dargestellten Begriffe lassen sich in Individualbegriffe und generische Begriffe einteilen.

In der KI wird mit semantischen Netzen eine Wissensbasis aufgebaut, aus der neues Wissen durch Schlussfolgerungen mit Hilfe von Inferenzregeln gewonnen wird. Eine Wissensbasis ist Teil eines wissenabasierten Systems. Die Wissensrepräsentation ist formalisiert, meist durch mathematische Logik. Ein Expertensystem ist ein sehr domänenspezifisches Wissensbasis-System (knowledge-based system), das für einen speziellen Zweck benutzt wird, z. B. für medizinische Diagnostik.

Die genaue Abgrenzung der Begriffe semantisches Netz, Wissensbasis und Ontologie gegeneinander ist nicht Gegenstand dieser Arbeit.

Topic Maps sind semantische Netzwerke, d.h. mit ihrer Syntax lassen sich semantische Netze aufschreiben (siehe Kapitel 3).

Ein globales semantisches Netzwerk ist die Zukunftsvorstellung vieler Arbeitsgruppen des World Wide Web Consortium (W3C).

2.5 Vision des Semantic Web

Der Begriff Semantic Web wurde von Tim Berners-Lee (Gründer des World Wide Web Consortium) geprägt und beschreibt ein "Web für Maschinen", im Gegensatz zum Web, das nur für Menschen lesbar (interpretierbar) ist (vgl. [Berners-Lee 2001]). Intelligente Agenten sollen in diesem Web eigenständig Aufgaben übernehmen.

Das Web ist die größte allgemein verfügbare, ziemlich unstrukturierte Wissensbasis. Um automatisiertes Suchen und automatisches Schließen zu ermöglichen, braucht man wie im KI-Bereich eine strukturierte Wissensbasis in einem für Maschinen verständlichen Formalismus.

Im heutigen Web geht implizites Wissen, das den Webressourcen durch ihre kontextuellen Verbindungen zueinander zuschreibbar ist, verloren.

Durch eine strukturierte Wissensbasis werden die Daten im Web für die Maschine automatisch erschließbar und nicht nur darstellbar, wodurch sie sich auch sinnvoll von dieser verarbeiten lassen, d.h. interpretiert, durchsucht, ausgetauscht werden können.

Früher konnten Browser im Web Zeichen nur wiedergeben. Sie konnten nicht interpretiert werden, da der Kontext fehlte.

Heute kann man Informationen durch Metadaten mit Hilfe von XML hinzufügen. Besonders implizite Informationen, die der Leser aus dem Kontext und Layoute einer Webseite erschließt (wie Titel, Autor etc.), müssen für Programme explizit dargestellt werden.

Auch zwischen den Informationsressourcen (üblicherweise Webseiten) bestehen semantische Beziehungen, die der Maschine so noch entgehen. Um explizit den Kontext der Informationsressourcen darzustellen, braucht man eine Sprache zur Beschreibung dieser Zusammenhänge (RDF⁵, Topic Maps).

Dies ermöglicht Suchmaschinen, effizienter zu arbeiten. Intelligen Agenten⁶ kann es so ermöglicht werden, Schlussfolgerungen zu ziehen und zu Entscheidungen zu treffen.

Das semantische Web besteht also aus einer Folge von immer abstrakter werdenden Schichten zur Vermittlung der Denkweise des Menschen an die Maschine, wobei von Uni-Code über XML-Schemas (Metasprache), RDF bis zu Ontologiesprachen (DAML⁷, Topic Maps) immer größere kontextuelle Zusammenhänge erstellt werden (vgl. [Beaulieu 2001]).

⁵ <http://www.w3.org/RDF/>

⁶ Intelligente Agenten: „Diese Programme, die agents, besitzen typischerweise einen hohen Grad an Autonomie und initiieren von selbst Aktionen, die der Erreichung ihrer Ziele dienlich sind. Dabei gehen sie aber nicht strikt nach einem gegebenen Schema vor, sondern interagieren mit ihrer Umwelt und reagieren auf Veränderungen. Die Vorgabe der Ziele erfolgt entweder in Form einer expliziten Definition durch den Benutzer oder indem der Agent das Benutzerverhalten beobachtet und selbst lernt“ [Rollet 00].

⁷ <http://www.daml.org/> DARPA Agent Markup Language

Agenten können erst in vollem Umfang „intelligent“ handeln, wenn sie eine Wissensrepräsentation haben, auf die sie zurückgreifen können. Der Agent muss ein Bild von der Welt haben, eine Domäne, in die eine Information gehört. Logische Schlussfolgerungen, Entscheidungsfindung und Beantwortung von Fragen können dann wie bei Wissensbasierten Systemen der künstlichen Intelligenz ausgeführt werden.

Man geht davon aus, dass in naher Zukunft viele Dokumente des Web mit Metadaten ausgezeichnet sein werden und so das verbesserte Suchen möglich wird (von manchen schon als Semantic Web bezeichnet). Eine Vision in noch weiterer Zukunft ist es, dass das ganze World Wide Web als eine vereinheitlichte „virtuelle Welt“ zu sehen ist in der Weise, dass hinter allen Dokumenten auf einer Metaebene dem Ganzen ein einziges standardisiertes konzeptuelles Modell zugeordnet ist (vgl. [Beaulieu 2001]).

Hierfür hat sich in der Informatik in den letzten Jahren ein Schlagwort gebildet: der Begriff Ontologie.

2.6 Ontologien

Im Folgenden wird erklärt, woher der Begriff Ontologie stammt und wie Ontologien im Informationsmanagement zur Strukturierung von Wissensbeständen genutzt werden.

Seit langem wurde, um Begriffe eines Fachgebietes zu ordnen und ihre Beziehungen darzustellen, eine Systematik aus dem Literatur- und Bibliothekswesen, ein Thesaurus, verwendet. Er ist ein herkömmliches Strukturierungsmittel für Begriffe einer Domäne.

Thesaurus⁸ (griechisch: Schatz): Der Begriff wurde früh auf den Wortschatz bezogen (vgl. das große lateinische Wörterbuch „Thesaurus Linguae Latinae“ von 1894). In der dokumentarischen Fachsprache hat sich folgende Bedeutung herausgebildet: Ein üblicher elaborierter Thesaurus für ein Fach- oder Aufgabengebiet enthält zumeist eine alphabetische Liste der Bezeichnungen für Sachbegriffe, also definierte Deskriptoren, wobei von Wörtern, die zu Nicht-Deskriptoren erklärt werden, auf Deskriptoren verwiesen wird, englisch mit dem Wort „use...“. Es wird aber auch von Deskriptoren auf Nicht-Deskriptoren verwiesen („used for...“). Praktisch geht es wieder um Synonyme oder Quasi-Synonyme, verwandte Begriffe usw. Es gibt dann noch Verweise auf („narrower term“=NT), weitere Begriffe („broader term“=BT) und auf Begriffe, zu denen sonstige Beziehungen bestehen („related term“=RT) [Kluth 1979].

Mit einem Thesaurus lassen sich durch „broader terms“ und „narrower terms“ Strukturen und damit hierarchische Relationen erzeugen und darstellen. Die Begriffe einer Anwendungsdomäne werden strukturiert und ein kontrolliertes Vokabular (Deskriptoren) wird erstellt.

Beispielsweise sind Thesauri in der Medizin: UMLS, SNOWMED, MeSH, mit denen versucht wird, die medizinische Fachsprache zu standardisieren.

⁸ Thesaurus bezeichnet auch einfach ein Synonymwörterbuch

Auch DiBIS enthält einen Thesaurus. Umgangssprachliche Begriffe, die bei einer Suchanfrage benutzt werden, werden in Hauptdescriptoren übersetzt. Auf die Anfrage nach einem §5-Schein wird über eine Relation zwischen §5-Schein und Wohnberechtigungsschein erkannt, dass es sich um Synonyme handelt und der korrekte Fachausdruck Wohnberechtigungsschein ist.

„Komplexe Modelle, die mit solchen Relationen arbeiten, werden teilweise auch als Ontologien bezeichnet“ ([Rechenberg 2002]S. 922).

2.6.1 Was ist eine Ontologie?

Ontologie (grch.): die philosophische Grunddisziplin der Lehre vom Seienden (Seinswissenschaft).

Sie handelt von dem Aufbau der Welt als Gegenstand allgemeiner Klassifizierungen. Danach ordnen sich die Elemente des Seins nach Einzelercheinungen in Gruppen, Arten und Gattungen, die sich aufgrund bestimmter Merkmale (belebt/unbelebt, körperlich/unkörperlich) unterscheiden.

Festgelegt wurde der Begriff von Aristoteles, der sich mit den metaphysischen Fragen beschäftigte: Was sind die Gemeinsamkeiten der “Seienden Objekte”? Was macht das Seiende aus? (vgl.[Guarino1995].)

Sein System beruht auf der Klassifikation in Totes, Lebendiges, Geistig-Übersinnliches und göttlich Ideelles. Diese erste Philosophie des Aristoteles, die nach dem Seienden, seinem Wesen und seiner Bestimmung fragt, wurde später allgemeine Metaphysik genannt. (vgl.[Brockhaus93], Stichwort „Ontologie“).

Seit Jahrhunderten wurde der Begriff in diesem philosophischen Zusammenhang benutzt. Im weiteren Sinne kann der Begriff aber auch nicht philosophische Einteilungen von Wesen/Dingen nach ihren charakteristischen Wesensmerkmalen bezeichnen.

Ein Beispiel hierfür ist eine sehr umfangreiche Ontologie, die Pflanzen-Klassifizierung von Linné, (1783), einem schwedischen Botaniker.

In der KI begann man mit der Erstellung künstlicher (virtueller) Welten und damit, für diese jeweils eigene Ontologien für die jeweilige Domäne zu erstellen (vgl. [Klischewski et al 03]), die nicht mehr den Anspruch erhoben, universelle Gültigkeit zu haben. Zahlreiche Ontologien wurden konstruiert, da verschiedene Bereiche der Welt modelliert wurden und die Sichtweise auf diese Ausschnitte verschieden ist. Diese Ontologien sind von der spezifischen Aufgabe geprägt, für die sie entwickelt wurden (vgl. [Ukena03], S. 20).

Im Bereich der Informatik bezeichnet Ontologie eine explizite formale Spezifikation eines konzeptuellen Modells. Konzeptualisierung kann intuitiv mit Begriffsbildung übersetzt

werden, in der KI wird der Begriff formal durch konzeptuelle Relationen erklärt. Darauf wird hier nicht weiter eingegangen.

„An ontology is an explicit specification of a conceptualization. The term is borrowed from philosophy, where an ontology is a systematic account of existence. For AI systems, what “exists” is that which can be represented. When the knowledge of a domain is represented in a declarative formalism, the set of objects that can be represented is called the universe of discourse. This set of objects, and the describable relationships among them, are reflected in the representational vocabulary with which a knowledge-based program represents knowledge” [Gruber93].

Für ein Programm im KI Bereich ist nur das existierend, was (formal) darstellbar ist.

Die Ontologie wird in einer deklarativen Sprache beschrieben, und die Dinge der Domäne werden „universe of discourse“ genannt. Die Objekte und Relationen werden durch ein sie eindeutig festlegendes Vokabular repräsentiert.

2.6.2 Ontologien in der praktischen Anwendung

Der Begriff Ontologie wird außerhalb der KI sehr pragmatisch verwandt, und es wird auch dann von einer Ontologie gesprochen, wenn keine formale, sondern nur eine informelle Spezifikation des Modells der Anwendungsdomäne vorliegt, die Ontologie beispielsweise durch einen Prosatext beschrieben wird (vgl. [Ukena 03], S. 21).

Der Begriff hat sich besonders im Wissensmanagement verbreitet, da für praktische Anwendungen in Bereichen wie Medizin, Bioinformatik, Verwaltungsorganen (e-government-services) Domänenmodelle benötigt werden. Man spricht auch von Domänenontologien.

Dem Erstellen einer Ontologie geht die Konzeptualisierung voraus, die Begriffsbildung für die Domäne. Dabei ist die Hauptschwierigkeit für eine weite Anwendungsdomäne, ein standardisiertes (kontrolliertes) Vokabular zu finden.

Dies ist gerade bei Semantic Web Technologien wichtig, da diese auf dem WWW operieren und damit eine sehr weite Anwendungsdomäne betreffen.

In verschiedenen modellierten Domänen, also verschiedenen Datenbeständen muss einem Agenten die Möglichkeit gegeben werden, zu erkennen, dass ein Begriff in diesen verschiedenen Systemen dieselbe Sache auszeichnet.

Auch in der Verwaltung ist die Einführung von allgemeingültigen Konzepten und standardisierten Begriffen durch die relative Eigenständigkeit der Kommunen und Ämter schwierig. Dieses ist aber für die Interoperabilität bei e-Government-Services notwendig. Gerade die Verwendung von Semantic Web Technologien hat die Interoperabilität beim Informationsmanagement mit im Auge. Mit der Findung einer standardisierten Sprache und

klaren Konzepten hierfür beschäftigen sich die Verwaltungsstellen verschiedener Länder, und hierzu gibt es internationale Gremien. Viele Länder haben ihre eigenen länderspezifischen Thesauren geschaffen (Australien hat seit den 70er Jahren Thesauren für Verwaltungszwecke (vgl. [Klischewski03 a]).

2.6.3 Ontologie-Entwicklung

Beim Design einer Ontologie arbeiten verschiedene Fachrichtungen, wie Linguistik, Logik, Formale Philosophie und Fachleute aus dem Bereich der zu modellierenden Domäne zusammen. Das Vorgehen hierbei, das Ontology Engineering, hat sich zu einem eigenständigen Forschungsfeld entwickelt (siehe z. B. Methontology von Gomez-Perez am Zentrum für KI in Madrid).

Wie oben gesagt, wird in der Praxis auf eine formelle Repräsentation der Ontologie verzichtet. Somit gibt es viele Möglichkeiten, Ontologien zu repräsentieren.

3 XML Topic Maps

Topic Maps dienen der Darstellung von Wissensstrukturen. Sie sollen

eine optimale Navigation ermöglichen,

den Informationspool verwalten,

eine Strukturierung der Informationen ermöglichen.

Es gibt zwei Arten von Topic Maps: Den ISO-Standard 13250 und XML Topic Maps, eine „aktuelle alternative Weiterentwicklung des ISO-Standards“ (vgl. [Mück et al.02], S. 3). Im nächsten Abschnitt wird ein Überblick über beide Standards gegeben.

3.1 ISO/IEC 13250

Anfang der 90er Jahre wurde die Entwicklung von Topic Maps durch Versuche angestoßen, umfangreiche Computer-Dokumentationen zu strukturieren und Indexe von Computer-Manuals miteinander zu verschmelzen (vgl. [Newcomb03], S. 37). Ende 1999 wurde der ISO 13250-Standard für Topic Maps verabschiedet. Die Spezifikation kann von der Webseite des ISO/IEC JTC1/SC34 heruntergeladen werden⁹. „Der Standard ist in formaler Hinsicht eine so genannte Meta-DTD“ (vgl. [Mück et al. 02], S. 89). Er besteht aus einer Menge so genannter *architectural forms*. Mit diesen kann der Benutzer eine eigene Topic Map-DTD „zusammensetzen“. Die für Topic Maps verwendeten *architectural forms* sind ein Ausschnitt von Hytime. „Topic Maps sind eine HyTime-Anwendung, das heißt sie verwenden einen zweckentsprechenden Ausschnitt dieses Standards“ (vgl. Mück et al. 02], S. 123). HyTime ist eine SGML-Erweiterung mit einer Vielzahl von *architectural forms*. Auf HyTime wird hier

nicht weiter eingegangen¹⁰. ISO 12350 Topic Maps basieren somit auf HyTime und SGML. Sie können trotzdem nicht nur in SGML, sondern auch in XML notiert werden. Denn „die Unterschiede zwischen SGML und XML sind im Zusammenhang mit Topic Maps nicht gravierend“ (vgl. [Mück et al. 02], S. 72). Das im Vorausgehenden öfter zitierte Buch „Topic Maps“ von [Mück et al.02] stellt ISO 12350-Topic Maps nicht in SGML, sondern in XML dar. Dadurch hielt ich sie am Anfang der Einarbeitung in das Thema fälschlicherweise für XML Topic Maps. Diese sind aber ein eigener Standard, und unterscheiden sich in vielen Punkten von ISO 12350 Topic Maps.

Auf die Unterschiede zwischen ISO 12350 und XML Topic Maps wird hier nicht eingegangen.

3.2 Der XML Topic Maps (XTM)-Standard

2000 wurden von der Arbeitsgruppe TopicMaps.org, XML Topic Maps (im weiteren XTM) spezifiziert¹¹. Diese sollten durch eine einfachere Syntax größeren Kreisen verständlich gemacht werden und optimal mit dem Web zusammenarbeiten. Die XTM beziehen sich eindeutig auf XML, und HyTime und SGML werden nicht mehr herangezogen (vgl. [Mück et al.02], S369). Die XTM-Syntax wird in einer DTD eindeutig festgelegt (siehe Anhang).

Aus diesen Gründen habe ich mich in dieser Arbeit nur mit XTM beschäftigt. Wenn im Folgenden von Topic Maps gesprochen wird, sind immer XTM gemeint.

Im Folgenden werde ich die wichtigsten Konzepte von XML Topic Maps vorstellen. Die Beispiele wurden in Anlehnung an [Altenburger00] gebildet. Die Erläuterungen sind der XTM Spezifikation entnommen (siehe [Pepper01]).

3.3 Topics

Ein Topic „steht für“ einen Gegenstand (englisch: subject) der realen Welt. Dieser Gegenstand kann alles sein, was der Mensch ersinnen und worüber geredet werden kann. Wenn ein Gegenstand in einer Topic Map durch ein Topic dargestellt wird, ist der Gegenstand „verdinglicht“ worden (englisch: reification), d. h. ein Topic vertritt ihn in der Topic Map, und hiermit wird der Gegenstand für den Computer erfassbar.

Jedes Topic muss eine eindeutige ID (unique identifier for element) besitzen. Diese dient lediglich zum internen Gebrauch innerhalb einer Topic Map.

Sollen Topic Maps zusammenarbeiten, muss genau bestimmbar sein, für welchen Gegenstand ein Topic „steht“. Dies wird über ein Element *subjectIdentity* festgelegt. Dieses

⁹ Auf dieser Seite der ISO/IEC findet sich die Topic Map specification:

<http://www.y12.doe.gov/sgml/sc34/>

¹⁰ HyTime specification: <http://www.y12.doe.gov/sgml/sc34/>

¹¹ <http://www.topicmaps.org/>

Element spezifiziert den Gegenstand der Realwelt, den ein Topic repräsentiert (vgl. Mück et al[02], S. 373). Es enthält zu diesem Zweck Unterelemente, die auf Topics oder Ressourcen verweisen, die den Gegenstand des Topics beschreiben. Beispielsweise kann ein subjectIdentity-Element ein *subjectIndicatorRef-Element* enthalten, das eine adressierbare Informationsressource bezeichnet, die eine eindeutige inhaltliche Beschreibung des Gegenstandes beinhaltet. Allgemein wird etwas, das die Identität eines Topics eindeutig beschreibt, als *Subject Indicator* bezeichnet. Published Subject Indicators sind der Öffentlichkeit zur Verfügung gestellte subjectIndicators. Durch sie sollen die Austauschbarkeit und die Verschmelzung von Topic Maps ermöglicht werden..

Zum Beispiel ist die deutsche Sprache durch einen Published Subject Indicator <http://www.topicmaps.org/xtm/1.0/language.xtm#de> eindeutig beschrieben.

Einem Topic können nachfolgende Eigenschaften (topic characteristics) zugeordnet sein: Name des Topic, seine Referenzen auf externe Ressourcen (englisch: occurrences) und Rollen, die es in Assoziationen spielt.

Der Name eines Topics besteht aus einem baseName und optional variant names zur Benutzung in verschiedene Anwendungszusammenhängen. Das variantName-Element enthält Parameter, die den Kontext beschreiben, in dem der jeweilige variant name zur Anwendung kommt.

Durch scopes können Gültigkeitsbereiche für Namen festgelegt werden (siehe Kapitel 3.5).

Das folgende Beispiel erzeugt ein Topic uni-hamburg und legt für dieses den Namen Universität Hamburg fest:

```
<topic id="uni-hamburg">
  <baseName>
    <baseNameString>Universität Hamburg</baseNameString>
  </baseName>
</topic>
```

3.4 Topic Typen

Einem Topic kann durch ein *instanceOf*-Element ein Typ zugewiesen werden, wobei der Typ selbst ein Topic ist. Ein Topic kann auch mehrere Typen besitzen. Sie können in der gleichen Topic Map definiert werden, in der sie benutzt werden. Sie können auch in einem anderen Dokument definiert werden, zu welchem eine Referenz erstellt wird. Typen führen eine Typhierarchie ein.

In diesem Beispiel wird gezeigt, dass uni-hamburg vom Typ Universität ist.

```
<topic id="uni-hamburg">
  <instanceOf>
    <topicRef xlink:href="#Universität"></topicRef>
  </instanceOf>
  <baseName>
```

```

    <baseNameString>Universität Hamburg</baseNameString>
</baseName>
</topic>

```

3.5 Topic Occurrences

Ein *topic* kann beliebig viele *occurrences* (deutsch: Vorkommen) aufweisen. Diese können externe Webressourcen mit relevanten Informationen (*information resources*) zu dem *topic* im allgemeinen via URIs referenzieren, wie im nachfolgenden Beispiel:

```

<topic id="uni-hamburg">
<baseName>
    <baseNameString>Universität Hamburg</baseNameString>
</baseName>
<occurrence>
    <resourceRef xlink:href="http://www.uni-hamburg.de"></resourceRef>
</occurrence>
</topic>

```

Occurrences können auch *character-data* in Form eines *ResourceData*-Elements enthalten, um einem Topic kurze Informationen in Form einer Zeichenkette hinzuzufügen.

```

<occurrence>
    <resourceData>Die Uni wurde 1921 gegründet</resourceData>
</occurrence>

```

Einer *occurrence* kann ein Typ zugeordnet werden; Eine *occurrence*, die auf ein elektronisch gespeichertes Bild verweist, könnte vom Typ *Grafik* sein. Der Sichtbarkeitsbereich für *occurrences* kann mit *scopes* eingeschränkt werden. Die technische Umsetzung der Verweise erfolgt mit *XLink*¹² (XML Linking Language).

3.6 Assoziationen

Assoziationen (englisch: *associations*) beschreiben Beziehungen zwischen Topics. Es wird keine Aussage über Symmetrie und Transitivität gemacht. Beliebig viele Topics können an einer Assoziation teilnehmen. Die beteiligten Topics werden Mitglieder (englisch: *members*) genannt. Jedem Mitglied der Assoziation kann eine Rolle zugewiesen werden. Im folgenden Beispiel wird eine Assoziation *uni-hamburg liegt in Deutschland* erstellt, wobei *uni-hamburg* die Rolle *Institution* und *Deutschland* die Rolle *Lage* zugewiesen wird. Es können verschiedene Mitglieder einer *association* die gleiche Rolle spielen, wie *uni-hamburg*, *uni-Berlin*, *uni-Erlangen*... *liegen-in* *Deutschland*, wobei sie alle die Rolle *Institution* zugewiesen bekommen.

Beispiel

¹² XML Linking Language (XLink) Version 1.0 (<http://www.w3.org/TR/xlink/#dt-hyperlink>)


```

<association>
<instanceOf>
  <topicRef xlink:href="liegt-in"></topicRef>
</instanceOf>
  <member>
  <roleSpec>
    <topicRef xlink:href="#institution"></topicRef>
  </roleSpec>
  <topicRef xlink:href="#uni-hamburg"/>
  </member>
    <member>
  <roleSpec>
    <topicRef xlink:href="#Lage"></topicRef>
  </roleSpec>
  <topicRef xlink:href="#deutschland"/>
  </member>
</association>

```

3.7 Scopes

Scopes legen den Gültigkeitsbereich für die Charakteristika eines Topics (*name*, *occurrence* und *association*) fest. Diese Topic-Eigenschaften sind damit nur in bestimmtem Kontext gültig. Wird nicht explizit ein Gültigkeitsbereich angegeben, sind die topic characteristics unbeschränkt gültig (*unconstrained scope*).

3.8 Merging

Topic Maps können mit Hilfe des *mergeMap*-Elements zusammengeführt werden, wobei das *mergeMap*-Element auf eine andere Topic Map verweist. Dabei werden Topics, die als gleich erkannt werden, zu einem Topic verschmolzen. Die Regeln, nach denen Topics verschmolzen werden, sind im XTM Processing Requirement-Anhang der XTM-Spezifikation aufgeführt. Sie sind sehr komplex. Es entsteht ein Topic, dessen Topic-Charakteristika die Vereinigung der Charakteristika der einzelnen Topics sind. Zwei Topics werden unter anderem als gleich angesehen, wenn sie entweder den gleichen Subject Indicator haben, oder wenn sie den gleichen baseName führen und außerdem im selben scope liegen.

4 Topic Map Tools

Es gibt spezielle Software-Produkte für die Verarbeitung von Topic Maps. Grob lassen sie sich nach ihrer Verwendung klassifizieren in Topic Map-Editoren (bzw. Designer) und Topic Map Browser.

Zu Topic Maps-Editoren können alle Programme gerechnet werden, die zum Erstellen von Topic Maps geeignet sind. Dazu gehören im weiteren Sinne auch Tools, die das Erstellen von Ontologien ermöglichen, aus denen sich mit Hilfe von Exportfunktionen Topic Maps generieren lassen (vgl. [Baumüller02], S. 39). „Ein Topic Map Editor (bzw. Designer) muss dabei Funktionalitäten zur Verfügung stellen, die den Anwender bei der Erstellung und Modellierung von Topic Maps unterstützen“ ([Baumüller02], S. 39); Ein Topic Map Editor sollte eine übersichtliche Benutzeroberfläche bieten. Da die Entwicklung einer Topic Maps meist ein teamorientierter Prozess ist, kann es von Vorteil sein, wenn ein Topic Map Editor, zusätzliche Funktionalität, wie Versionskontrolle und Vergabe von Nutzerrechten und Nutzerrollen, zur Verfügung stellt (vgl. [Baumüller02], S39).

Ein Topic Map Browser sollte eine optimale Visualisierung ermöglichen. Am Laboratoire d'Informatique de Paris wurde ein Browser für Topic Maps entwickelt, der Topic Maps in Explorer-Darstellung und als 3-D-Graphen anzeigt. Die 3-D-Darstellung soll eine übersichtliche Navigation in sehr großen Topic Maps ermöglichen (vgl. [Soto]).

Einige bereits erschienen kommerziellen und frei verfügbaren Topic Map Tools werden im Folgenden aufgelistet:

Frei verfügbar:

- TM4J (www.techquilla.com/tm4j.html)
- Topic Map Designer (www.topicmap-design.com)
- Tmproc (www.ontopie.net/software/tmproc)

Kommerzielle Systeme:

- Ontopia (www.ontopia.net)
- Empolis (www.empolis.com)
- Modeca (www.mondeca.com)
- Infoloom (www.infoloom.com)

(Zitiert nach [Hinkelmann02])

5 Das WebInfoService-Projekt und Topic Maps

Verwaltungsvorgänge werden zunehmend komplexer und in immer größerem Umfang elektronisch abgewickelt. Diese Dienste (englisch: e-government services) sind ein Anwendungsfeld für XML-basierte Semantic Web Technologien (die sich in einem verwandten Bereich, dem E-Business, bereits etabliert haben). Aufgrund dieser Entwicklung ist die Verwaltung der Stadt Hamburg bemüht, sich dieser e-government-Entwicklung anzuschließen und hat das WebInfoService-Projekt unterstützt.

Ziel des Projektes war es, Semantic Web Technologien durch Entwicklung eines Prototyps für die Hamburger Behörde zu evaluieren unter besonderer Berücksichtigung des Informationsmanagements und mit dem Anliegen, Benutzereingaben semantisch auszuwerten und für eine „intelligente“ Navigation nutzbar zu machen.

In diesem Kapitel wird ein Überblick über das Projekt, soweit es für diese Arbeit relevant ist, gegeben. Besonders wird die Erstellung einer Ontologie erörtert und darauf die Verwendung von Topic Maps in diesem Zusammenhang an einem Beispiel erläutert.

Am Ende des Kapitels wird darauf eingegangen, wie man Topic Maps noch für das Informationsmanagement in Bezug auf DiBIS nutzen könnte.

5.1 Überblick über das Projekt

Im Projekt wurde versucht, in Anlehnung an das Bürgerinformationssystem DiBIS einen Prototypen zu entwickeln, der dessen Funktionalität erweitert. Darum wird am Anfang dieses Kapitels DiBIS erläutert.

5.1.1 DiBIS

DiBIS steht für Direkte Bürger-Informationen-Services und ist ein unter der Adresse <http://DiBIS.dufa.de> aufrufbarer Webservice (er lässt sich auch über die Seite www.hamburg.de erreichen). WebInfoServices sind internetgestützte Informationssysteme, die als Produkt Informationen anbieten. DiBIS erkennt Schlagwörter, die im Zusammenhang mit Verwaltungsvorgängen benutzt werden, und bietet zu diesen kurze Informationen über die Dienstleistungen der Verwaltung (der Bezirksämter und Behörden), wie Zuständigkeiten, Adresse und Raumnummer des Amtes, Telefon und Telefax, Öffnungszeiten, behindertenfreundlichen Zugang, erforderliche Unterlagen und Gebühren, Besonderheiten, die zu beachten sind, Verkehrsverbindungen des Hamburger Verkehrsbundes HVV.

Die Portalseite von DiBIS bietet einen Einstieg zu den verschiedenen Angeboten (durch Auswahlangebot an Hyperlinks zu den jeweiligen Sites). Die „DiBIS-Auskunft“-Seite kann

geöffnet werden und zeigt dann einen Schlagwortkatalog mit Informationen über die Dienstleistungen der Verwaltung an.

Außerdem können eine Download-Seite mit Formularen und Broschüren im PDF-Format und eine Hilfeseite erreicht werden.

Es besteht darüber hinaus die Möglichkeit, e-mail-Anfragen an DiBIS zu stellen, und es kann eine Sprache ausgewählt werden, in der die weiteren Informationen angezeigt werden sollen.

Im Folgenden wird die Funktionalität von DiBIS an einem Beispiel erläutert.

Beispiel:

Von der Portalseite ruft man die Seite *DiBIS-Auskunft* auf.

Hier kann man in einer Maske Suchbegriffe eingeben oder alphabetisch nach Begriffen suchen, z.B. *Reisepass*.

Der Service verlangt daraufhin die Eingabe der Wohnstraße in eine neue Eingabemaske, z.B. *Weidestraße 33*.

Daraufhin werden Informationen über die zuständige Behörde angezeigt,

Z.B. *Einwohnerabteilung Barmbek/Uhlenhorst* mit Adresse, Telefonnummern, Öffnungszeiten: *Mo 9-17...*, Verkehrsverbindung *U-Bahn, Linie U2, U3, S-Bahn S1, Busse ...*, erforderliche Unterlagen: *Personalausweis oder alter Reisepass und ein Foto etc.*

DiBIS fordert, nachdem über den Schlagwortkatalog das Interessengebiet gewählt wurde, Benutzerdaten an, im Beispiel die Wohnstraße *Weidestraße*. Die Benutzerdaten werden mit dem Sachgebiet des Schlagworts in Zusammenhang gebracht, um dem Benutzer auf ihn zugeschnittene Auskünfte zu erteilen (*in der Weidestraße wohnhaft + Reisepass beantragen*).

DiBIS ermöglicht nicht Transaktionen, wie z.B. das Überweisen von Gebühren an die Behörde oder Online-Abgabe von Formularen. Es ist nur möglich, Formulare auszudrucken, der Weg zur Behörde bleibt dem Bürger nicht erspart.

Auch in anderen Städten (wie München) gibt es derartige Webinfoservices.

5.1.2 Ziele des Projekts

DiBIS sollte im Rahmen des Projektes in folgendem Sinne verändert werden:

Im Zuge des Informationsmanagements war es ein Ziel, einen durch eine Ontologie strukturierten Prototyp zu entwerfen.

Die Ontologie sollte ein „festes Gerippe“ des Prototypen bilden, an den die einzelnen Informationsressourcen „angedockt“ werden können. Eine derartige Schichtung in zwei Ebenen verbessert die Wartung, Änderbarkeit und Interoperabilität des Softwareentwurfs. Es lassen sich redaktionelle Änderungen, die die Grundstruktur nicht betreffen, dann leicht durch Austausch der Informationsressourcen durchführen, ohne die obere Ebene umstellen zu müssen.

Interoperabilität mit Webservices anderer deutscher Städte herzustellen, war ebenfalls ein wesentliches Ziel des Projektes. Das jetzige DiBIS hat eine in sich abgeschlossene IT Infrastruktur und ist so für globales e-government ungeeignet.

Ein dabei in Kauf zu nehmender Nachteil ist, dass sich die Laufzeit des Programmes gegenüber der direkten Vernetzung der Informationsressourcen durch Hyperlinks verschlechtert.

Weiterhin soll der Prototyp eine vom Kontext geleitete Navigation ermöglichen, das heißt, vom Benutzer während der Websitzung gemachte Angaben werden gespeichert und weiterhin genutzt, um spezifische, auf den Benutzer zugeschnittene Informationen auszugeben (Kontextualisierung von Benutzerdaten). Die zu speichernden Angaben sind der bisherige Navigationspfad und die Einträge in Eingabemasken während der jeweiligen Websession. Im jetzigen DiBIS werden solche Kontextinformationen nur in geringem Umfang genutzt (siehe obiges Beispiel). Mit Hilfe einer Ontologie soll eine Wissensbasis festgelegt werden, die explizit kontextuelle Zusammenhänge repräsentiert und mit den gespeicherten Eingaben dem Prototyp ermöglicht, Schlüsse zu ziehen und so „intelligente“ Antworten und Informationen auf Anfragen zu ermöglichen. Der Versuch, eine Ontologie zu erstellen, wurde zum Schwerpunkt des Projekts und erwies sich in dessen Verlauf als sehr schwierig.

Es war kein Ziel des Projekts, Benutzerdaten über eine Websession hinaus zu speichern, und auch nicht, Transaktionen online durchzuführen, da dies eine Authentifikation des Benutzers voraussetzt und nicht ohne weiteres mit den Datenschutzgesetzen vereinbar ist.

5.1.3 Projektverlauf

Das Projekt fand einmal wöchentlich über ein Jahr statt und es nahmen 10 Studenten aus höheren Fachsemestern daran teil. Einbezogen in das Projekt wurde der Entwickler von DiBIS, Herr Kruse, der als Kontaktperson zur Behörde fungierte und über das jetzige DiBIS und Anforderungen an das System von Seiten der Behörde informierte.

Die Software-Technik verlangt, dass der Entwicklung eines Softwaresystems für einen speziellen Anwendungsbereich eine detaillierte Anforderungsanalyse vorausgeht und der Anwendungshintergrund ausgeleuchtet wird, indem die zukünftigen Anwender befragt werden, das heißt die Bürger und Verwaltungsmitarbeiter.

Dies war nur in soweit möglich, als das Herr Kruse lediglich einmal befragt werden konnte. Außerdem fand ein Gespräch mit Mitarbeitern der Firma sd&m statt, die ebenfalls an einer Erweiterung von DiBIS arbeiten. Ansonsten wurde versucht, die Anforderungen an das System durch das Anfertigen von Szenarien zu klären. Dies sind skizzierte Szenen, die sich theoretisch abspielen könnten und im Projekt ausführlich diskutiert wurden.

Es wurden Szenarien aus Sicht des Bürgers zu den Lebenslagen Umzug, Heirat und Beerdigung entworfen.

Die Lebenslage *Umzug* schien den Studenten als Beispiel am besten geeignet, in eine Semantic Web- Anwendung umgesetzt zu werden.

5.1.3.1 Technische Umsetzung des Prototyps

Es handelte sich um ein exploratives Projekt, bei dem verschiedene Techniken diskutiert und praktische Erfahrungen durch die Erstellung des Prototyps gesammelt werden sollten. Es wurden die Webtechnologien Topic Maps und DAML+OIL¹³ im Projekt erläutert. Die Wahl fiel auf DAML+OIL, da hierfür ein Software Toolkit (Jena Framework¹⁴) von Hewlett Packard zur Verfügung stand. (DAML+OIL ist eine Erweiterung von RDF/RDFS, ist aber stärker auf die Erstellung von Ontologien ausgerichtet). Die Ontologie sollte von einer in Java implementierten Anwendungslogik-Schicht mit für jeden Benutzer individuell angelegten RDF-Instanzen genutzt werden, um die auf den Benutzer zugeschnittenen Websites mit dem Content Management System Coremedia zu generieren. Zu technischen Details siehe: „Umsetzung eines semantischen Modells in DAML+OIL zur Steuerung einer Web-Anwendung“ (Baccalaureatsarbeit von Moritz Kleine [Kleine03]). Die Kontextualisierung der Benutzerdaten erwies sich mit dem Jena Softwaretoolkit als so aufwendig, dass nach Ende des einen Jahres Projektarbeit kein fertiges Ergebnis vorlag.

Darum enthält diese Arbeit keinen Überblick darüber, wie die Kontextualisierung technisch verwirklicht wurde.

¹³ <http://www.daml.org/2001/03/daml+oil-index.html>

¹⁴ <http://jena.sourceforge.net>

5.2 Ontologie im Projekt

Im Projekt ist keine wirkliche Ontologie, sondern ein semantisches Netz entstanden.

Die erste Ungewöhnlichkeit der „Ontologie“ im Projekt war, dass zum Teil keine physischen Dinge konzeptualisiert wurden, und auch keine nicht physischen Dinge, wie Liebe, Totes, Geistlich-Übersinnliches, sondern stattdessen Vorgänge und Transaktionen, wie UMZUG oder eine MELDUNG. Dies allein ist aber noch kein Hinderungsgrund, von einer Ontologie zu sprechen.

Außerdem wurde bei der konzeptuellen Sicht auf die Domäne nicht eine einheitliche Perspektive eingehalten. Hierfür wird ein Beispiel gegeben. Assoziationen werden dabei wie Relationen in der Mathematik notiert, zum Beispiel: kleiner-als(ZWEI, FÜNF).

Die Folgende Relation wurde aus der Perspektive des Bürgers, der umziehen möchte, erstellt:

Umziehen-innerhalb-Hamburgs(UMZUG, UMMELDEN).

Aus einer anderen Sichtweise heraus, die den Blick auf die spätere mehrdimensionale Hyperebene der gangbaren Navigationspfade richtet, ist Folgende Relation entstanden:

Verweis-auf-Information(UMZUG, www.ich-zieh-um.de). Diese Relation hat sich ergeben, da die Studenten, die die Ontologie entwickelten, davon ausgingen, dass ein Hyperlink von der Start-Seite Umzug, zu einer allgemeinen Informationsseite über Dinge, die beim Umziehen zu bedenken sind, praktisch ist. Dies hat aber nichts mit einer Ontologie zu tun.

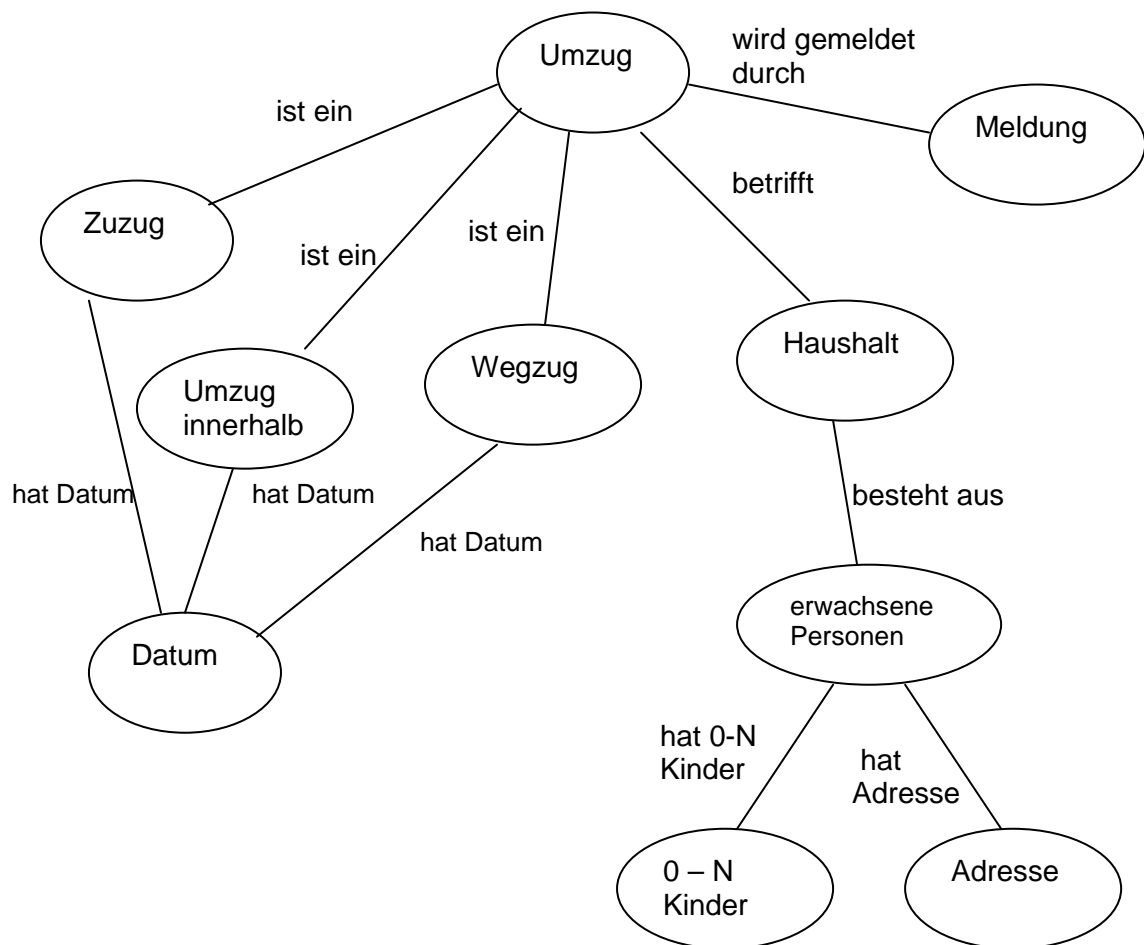
Andere Assoziationen der „Ontologie“ zeigen das, was man - ohne den Begriff zu vergewaltigen - , noch als Domänenontologie bezeichnen kann, wie z.B.

Ist-geboren-in(Person, Geburtsort) oder Öffnungszeiten-haben(Amt, Öffnungszeiten). Dieses sind feste Gegebenheiten, die nicht unter der Vision der zukünftigen Funktionalität des Systems entstanden sind. Da die Studenten im Projekt keinerlei Erfahrung mit Ontologien hatten, wurde dieses Wort im Projekt ständig für dieses semantische Netz benutzt, in dem verschiedene Elemente vermischt vorkommen, was mich als Projektteilnehmerin, sehr stark verwirrte und zu allgemeinen Verständigungsschwierigkeiten im Projekt führte. Im weiteren Verlauf des Projektes wurde versucht, die Ontologie durch „Anbauten“ und „Umbau“ anzupassen. Diese Veränderungen führten dazu, dass die Ontologie umfangreicher wurde, eine einheitliche Konzeptualisierung wurde nicht erreicht.

5.3 Beispiel: Entwicklung eines Teils der DiBIS-Ontologie mit Topic Maps

Das im Rahmen des Projektes entworfene semantische Netz wurde von M. Kleine im Rahmen seiner Baccalaureatsarbeit *Umsetzung eines semantischen Modells in DAML+OIL*

zur *Steuerung einer Web-Anwendung* [Kleine02] überarbeitet. Dabei ist eine Ontologie der Lebenslage Umzug entstanden.



(Beispiel in Anlehnung an die *Umzugsontologie* aus Moritz Kleines *Baccalaureatsarbeit* [Kleine02], S. 22).

Kleine schreibt zu dieser Umzugsontologie: "All diejenigen Begriffe, die aus behördlicher Sicht für den Bürger bei dem Anliegen *Umzug* relevant sind, werden in dieser Ontologie aufgeführt, beschrieben und zueinander in Beziehung gesetzt" (vgl. [Kleine02], S. 21). Die Abbildung auf dieser Seite gibt einen leicht abgewandelten Ausschnitt dieser Umzugsontologie wieder.

Die Ontologie wurde im Zusammenhang mit der Weiterarbeit an dem Prototypen zum Zwecke der Kontextualisierung von Benutzerzerdaten mit DAML+OIL entwickelt. Ich habe mich in dieser Arbeit mit dem Prototyp nicht weiter beschäftigt, da dies einer Einarbeitung in DAML&OIL bedurft hätte. Aus diesem Grund wird die Ontologie aus Kleines Arbeit hier als

Beispiel verwendet, ohne dass darauf eingegangen wird, welchem Zweck sie im Prototyp dient und warum sie in genau dieser Weise strukturiert ist.

Die Abbildung der Umzugsontologie zeigt: Umzug ist eine Verallgemeinerung von Zugug (nach Hamburg), Umzug innerhalb (innerhalb Hamburgs), und Wegzug (aus Hamburg). Alle diese näher spezifizierten Umzüge besitzen ein Datum. Die weiteren Beziehungen zeigt die Abbildung. Ein Umzug betrifft einen Haushalt und wird durch eine Meldung gemeldet usw. Beim Knoten „Kinder“ steht 0-N, da erwachsene Personen keine Kinder haben müssen.

Im Folgenden wird der Umzugsontologie-Ausschnitt in Topic Maps dargestellt. Das Beispiel wurde mit XMLSpy erstellt, einem kommerziellen XML-Editor. Erläuterungen finden sich in Form von Kommentaren im XML-Dokument. Zuerst werden alle Topics erstellt, dies sind die Knoten in der Umzugsontologie. Danach werden alle Assoziationen erstellt.

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XMLSPY v2004 rel. 3 U (http://www.xmlspy.com) -->
<!-- Ausschnitt aus der Projekt-DiBIS-Umzugsontologie-->

<!-- Einbinden der Topic Map-DTD (ich habe sie StevePeppers.DTD genannt)-->
<!DOCTYPE topicMap SYSTEM "D:\Baccalaureat\Topic Maps\StevePeppers.DTD">

<topicMap>

<!-- Anlegen der Topics -->

<!-- Das Topic Umzug.-->
  <topic id="Umzug">
    <baseName>
      <baseNameString>Umzug</baseNameString>
    </baseName>
  </topic>

<!-- Spezialisierte Umzüge (Zuzug, Wegzug aus Hamburg oder Umzug innerhalb von
Hamburg) sind Instanzen von Umzug, ausgedrückt durch "instanceOf". InstanceOf ist eine
Kurzform für class-instance-Beziehungen. "class-instance published subject" beschreibt
diese Beziehung (siehe XTM-Spezifikation) -->

<!-- Das Topic Zuzug -->
  <topic id="Zuzug">
    <instanceOf>
      <topicRef xlink:href="#Umzug"/>
    </instanceOf>
    <baseName>
      <baseNameString>Zuzug</baseNameString>
    </baseName>
  </topic>

<!-- Topc Wegzug-->
  <topic id="Wegzug">
    <instanceOf>
      <topicRef xlink:href="#Umzug"/>
    </instanceOf>
    <baseName>
```

```

        <baseNameString>Wegzug</baseNameString>
    </baseName>
</topic>

<!-- Topic Umzug innerhalb-->
<topic id="Umzug_innenhalb">
    <instanceOf>
        <topicRef xlink:href="Umzug"/>
    </instanceOf>
    <baseName>
        <baseNameString>Umzug innerhalb Hamburgs</baseNameString>
    </baseName>
</topic>

<!--Topic Haushalt -->
<topic id="Haushalt">
    <baseName>
        <baseNameString>Haushalt</baseNameString>
    </baseName>
</topic>

<!-- Topic Meldung -->
<topic id="Meldung">
    <baseName>
        <baseNameString>Meldung</baseNameString>
    </baseName>
</topic>

<!-- Topic erwachsene Person(en) -->
<topic id="erwachsenePersonen">
    <baseName>
        <baseNameString>erwachsene Person(en)</baseNameString>
    </baseName>
</topic>

<!-- Topic 0-N Kinder -->
<topic id="Kindermenge">
    <baseName>
        <baseNameString>0-N Kinder</baseNameString>
    </baseName>
</topic>

<!-- Assoziationen mit ihren Mitgliedern und deren Assoziationsrollen -->

<!-- Umzug betrifft Haushalt-Assoziation -->
<association id="Umzug-betrifft-Haushalt">
    <member>
        <roleSpec>
            <topicRef xlink:href="#Subjekt"/>
        </roleSpec>
        <topicRef xlink:href="#Umzug"/>
    </member>
    <member>
        <roleSpec>
            <topicRef xlink:href="#Objekt"/>
        </roleSpec>
        <topicRef xlink:href="Haushalt"/>
    </member>

```

```

</association>

<!-- Umzug wird durch Meldung gemeldet-Assoziation -->
<association id="Umzug-wird-gemeldet-durch-Meldung">
  <member>
    <roleSpec>
      <topicRef xlink:href="zu meldendender Sachverhalt"/>
    </roleSpec>
    <topicRef xlink:href="#Umzug"/>
  </member>
  <member>
    <roleSpec>
      <topicRef xlink:href="Meldungsmittel"/>
    </roleSpec>
    <topicRef xlink:href="#Meldung"/>
  </member>
</association>

<!-- Haushalt besteht aus erwachsene(r) Person(en)-Assoziation -->
<association id="Haushalt-besteht-aus-Personen">
  <member>
    <roleSpec>
      <topicRef xlink:href="#Ganzes"/>
    </roleSpec>
    <topicRef xlink:href="#Haushalt"/>
  </member>
  <member>
    <roleSpec>
      <topicRef xlink:href="#Teil"/>
    </roleSpec>
    <topicRef xlink:href="#erwachsenePersonen"/>
  </member>
</association>

<!-- Erwachsene(r) Person(en) haben 0-N Kinder -->
<association id="hatKindermenge">
  <member>
    <roleSpec>
      <topicRef xlink:href="#Eltern"/>
    </roleSpec>
    <topicRef xlink:href="#erwachsenePersonen"/>
  </member>
  <member>
    <roleSpec>
      <topicRef xlink:href="#Kinder"/>
    </roleSpec>
    <topicRef xlink:href="Kindermenge"/>
  </member>
</association>

<!-- Assoziation, spezielle Umzüge haben ein Datum -->
<association id="DatumHaben">
  <member>
    <roleSpec>
      <topicRef xlink:href="Ereigniss, das ein Datum hat"/>
    </roleSpec>
    <topicRef xlink:href="#zuzug"/>
  </member>

```

```

<member>
  <roleSpec>
    <topicRef xlink:href="Ereigniss, das ein Datum hat"/>
  </roleSpec>
  <topicRef xlink:href="#Wegzug"/>
</member>
<member>
  <roleSpec>
    <topicRef xlink:href="Ereigniss, das ein Datum hat"/>
  </roleSpec>
  <topicRef xlink:href="#UmzugInnerhalb"/>
</member>
<member>
  <roleSpec>
    <topicRef xlink:href="Datum"/>
  </roleSpec>
  <topicRef xlink:href="#Datum"/>
</member>
</association>
</topicMap>

```

Der Blickwinkel auf eine Topic Map kann durch Scopes in Anpassung an den Benutzerkontext geändert werden. So kann z.B ein englischsprachiger Bürger durch den scope englisch englische Webseiten angezeigt bekommen (vgl. [Pepper03]).

6 Zusammenfassung

Im Zusammenhang mit der Zunehmend Menge an Dokumenten im Internet werden Semantic Web Technologien immer wichtiger, um einen Überblick in der Informationsflut zu behalten.

Das Informationsmanagement ist ein wichtiges Anwendungsfeld für Topic Maps. Die im zweiten Kapitel vorgestellten Konzepte zur Wissensstrukturierung, wie Thesauren, semantische Netze und Ontologien, lassen sich mit Topic Maps darstellen. Welche Funktion die im Zusammenhang mit dem Prototypen erstellte Ontologie für eine Erweiterung von DiBIS hat, wurde hier nicht beschrieben. Es wurde lediglich eine im Rahmen von DiBIS erstellte Ontologie mit Topic Maps dargestellt. Es zeigte sich, dass eine Ontologie sich grundsätzlich mit Topic Maps repräsentieren lässt.

Da die Topic Maps, wenn man sie direkt in XML erstellt, schnell sehr umfangreich und unübersichtlich werden, scheint es mir notwendig, für ihre Erstellung Topic Map-Editoren zu verwenden. Die Untersuchung von Topic-Map-Editoren ist jedoch nicht Gegenstand dieser Arbeit.

7 Literatur

- [Altenburger00] *Authoring XTM Topic Maps*. Online Tutorial,
<http://topicmaps.it.bond.edu.au/docs/6/toc>, besucht am 14.11.2003
- [Berners-Lee et al.01] Berners-Lee, T., Henler, J., Lassila, O.: *The Semantic Web*,
Scientific American, Mai 2001
- [Baumüller02] Baumüller, A.: *Topic Maps zur semantischen Strukturierung von Wissensbasen im Knowledge Management*. Diplomarbeit,
Technische Fachhochschule Wildau
- [Beaulieu2001] Beaulieu, P. : *Le Web sémantique*, Travail final, Université de
Montreal, École de bibliothéconomie et des sciences de
l'information,
- [Brockhaus88] dtv Brockhaus Lexikon. F. A. Brockhaus, Mannheim und Deutscher
Taschenbuchverlag München
- [Brockhaus93] *Brockhaus Enzyklopädie*. F.A. Brockhaus Mannheim
- [Computer00] *Computer Fachlexikon Fachwörterbuch*, Microsoft Press
Deutschland, Unterschleißheim
- [Duden 97] Fremdwörterduden, Dudenverlag Mannheim
- [FOLDOC] *FOLDOC Free On Line Dictionary Of Computing*.
<http://foldoc.doc.ic.ac.uk/foldoc/index.html>, besucht am 13.11.11.
- [Gruber93] Gruber T.R.: *A translation approach to portable ontologies*.
Knowledge Acquisition, 5(2):199-220.
- [Hinkelmann02] Hinkelmann Knut: *Informationserschließung durch Klassifikation*
Folien zur Vorlesung *Informations- und Wissensmanagement* (WS
2001/032). Fachhochschule Solothurn.
[http://www.hsw.fhso.ch/hinkelmann/IWM/IWM2.2-
Wissensstrukturen.pdf](http://www.hsw.fhso.ch/hinkelmann/IWM/IWM2.2-Wissensstrukturen.pdf), besucht am 13.11.2003
- [Kleine03] Kleine M.: *Umsetzung eines semantischen Modells in DAML+OIL zur Steuerung einer Web-Anwendung*, Baccalaureatsarbeit.
Universität Hamburg, Fachbereich Informatik.

- [Klischewski03 a] Klischewski, R.: *Towards an Ontology for e-Document Management in Public Administration - the Case of Schleswig-Holstein*. Proceedings HICSS-36, IEEE.
- [Klischewski03 b] Klischewski, R.: *Top Down or Bottom Up? How to Establish a Common Ground for Semantic Interoperability within e-Government Communities*. In: Traunmüller, R., Palmirani, M. (ed.) *E-Government: Modelling Norms and Concepts as Key Issues*. Proceedings of 1st International Workshop on E-Government at ICAIL 2003. Bologna: Gedit edizioni, pp. 17-26.
- [Klischewski et al.03] Klischewski R., Jeenicke M.: *Semantic Web Technologies for Information Management within e-Government Services*. Universität Hamburg, Fachbereich Informatik.
- [Kluth79] Kluth R.: *Lehrbuch der Bibliothekspraxis*. Otto Harrasowitz, Wiesbaden.
- [Newcomb03] Newcomb S.R.: *A perspective on the quest for global knowledge interchange*. In [Park et al 03], S 31.
- [Park et al 03] Park, J., Hunting S.(Hrsg.): *XML Topic Maps: Creating and Using Topic Maps for the Web*. Addison Wesley Professional.
- [Pepper01] Pepper, S.;Moore, G.: *XML Topic Maps(XTM) 1.0*. XTM Specification <http://www.topicmaps.org/xtm/1.0/>, besucht am 14.11.2003.
- [Rechenberg et al 02] Rechenberg, P., Pomberger, G.(Hrsg.): *Informatik-Handbuch*, Hanser Verlag, München und Wien.
- [Rollet00] Rollet H.: *Aspekte des Wissensmanagement*. Diplomarbeit, Technische Universität Graz.
http://www2.iicm.edu/herwig/thesis/Diplomarbeit_Herwig_Rollett_ONLINE.pdf besucht am 14.11.2003.
- [Pepper03] *The Tao of Topic Maps*.
<http://www.ontopia.net/topicmaps/materials/tao.html>, zuletzt besucht am 14.11.2003
- [Soto] Soto, M.: *Information management - Topic Maps visualization*. Universite Pierre et Marie Curie, Laboratoire d'Informatique <http://www.gca.org/papers/xmleurope2000/papers/s29-03.html>, besucht am 13.11.2003

- [Ukena03] Ukena, S.: *Ontologieansätze in internet-basierten IT-Systemen am Beispiel der elektronischen Verwaltung*. Studienarbeit, Universität Hamburg, Fachbereich Informatik.
- [Widhalm et al. 02] Widhalm, R., Mück, T.: *Topic Maps*, Springer-Verlag, Berlin und Heidelberg
- [XML02] Extensible Markup Language (XML) 1.0. W3C Recommendation.
<http://www.w3.org/TR/REC-xml>, besucht am 13.11.2003
- [Young02] Young M.: *XML Schritt für Schritt*, Microsoft Press Deutschland Unterschleißheim.

8 Anhang

Topic Maps DTD aus der XTM Spezifikation (siehe [Pepper01], Annex D).

XTM 1.0 Document Type Declaration

```
<!-- ..... -->
<!-- XML Topic Map DTD ..... -->
<!-- file: xtm1.dtd
      -->
<!-- XML Topic Map (XTM) DTD, Version 1.0
```

This is XTM, an XML interchange syntax for ISO 13250 Topic Maps.

XML Topic Map (XTM)

Copyright 2000-2001 TopicMaps.Org, All Rights Reserved.

Permission to use, copy, modify and distribute the XTM DTD and its accompanying materials for any purpose and without fee is hereby granted in perpetuity, provided that the above copyright notice and this paragraph appear in all copies. The copyright holders make no representation about the suitability of the DTD for any purpose. It is provided "as is" without expressed or implied warranty.

```
Editors:  Steve Pepper  <pepper@ontopia.net>
          Graham Moore <gdm@empolis.co.uk>
Authors:  Murray Altheim <altheim@eng.sun.com>
          Michel Biezunski <mb@infoloom.com>
          Sam Hunting    <shunting@etopicality.com>
          Steven R. Newcomb <srn@coolheads.com>
Status:   Release
Version:  v1.0.1
Revision: $Id: xtm1.dtd,v 1.2 2001/02/08 16:03:12 pepper Exp $
PublicId: "-//TopicMaps.Org//DTD XML Topic Map (XTM) 1.0//EN"
```

Revisions:

```
#2001-01-21: removed baseName from occurrence
#2001-02-02: made variantName optional in variant
#2001-02-02: changed ID to #IMPLIED on association
#2001-02-02: changed ID to #IMPLIED on resourceData
#2001-02-02: changed PLUS to REP on member
```

```
-->
```

```
<!-- Use this URI to identify the default XTM namespace:
```

```
      "http://www.topicMaps.org/xtm/1.0/"
```

```
Used to identify the XLink namespace:
```

```
      "http://www.w3.org/1999/xlink"
```

```
-->
```

```
<!-- topicMap: Topic Map document element ..... -->
<!ELEMENT topicMap (topic | association | mergeMap)*>
<!ATTLIST topicMap
  id ID #IMPLIED
  xmlns CDATA #FIXED "http://www.topicmaps.org/xtm/1.0/"
```



```

xmlns:xlink CDATA #FIXED "http://www.w3.org/1999/xlink"
xml:base CDATA #IMPLIED
>
<!-- topic: Topic element ..... -->
<!ELEMENT topic (instanceOf*, subjectIdentity?, (baseName | occurrence*))>
<!ATTLIST topic
  id ID #REQUIRED
>
<!-- instanceOf: Points To a Topic representing a class ..... -->
<!ELEMENT instanceOf (topicRef | subjectIndicatorRef)>
<!ATTLIST instanceOf
  id ID #IMPLIED
>
<!-- subjectIdentity: Subject reified by Topic ..... -->
<!ELEMENT subjectIdentity (resourceRef?, (topicRef | subjectIndicatorRef*))>
<!ATTLIST subjectIdentity
  id ID #IMPLIED
>
<!-- topicRef: Reference to a Topic element ..... -->
<!ELEMENT topicRef EMPTY>
<!ATTLIST topicRef
  id ID #IMPLIED
  xlink:type NMTOKEN #FIXED "simple"
  xlink:href CDATA #REQUIRED
>
<!-- subjectIndicatorRef: Reference to a Subject Indicator ..... -->
<!ELEMENT subjectIndicatorRef EMPTY>
<!ATTLIST subjectIndicatorRef
  id ID #IMPLIED
  xlink:type NMTOKEN #FIXED "simple"
  xlink:href CDATA #REQUIRED
>
<!-- baseName: Base Name of a Topic ..... -->
<!ELEMENT baseName (scope?, baseNameString, variant*)>
<!ATTLIST baseName
  id ID #IMPLIED
>
<!-- baseNameString: Base Name String container ..... -->
<!ELEMENT baseNameString (#PCDATA)>
<!ATTLIST baseNameString
  id ID #IMPLIED
>
<!-- variant: Alternate forms of Base Name ..... -->
<!ELEMENT variant (parameters, variantName?, variant*)>
<!ATTLIST variant
  id ID #IMPLIED
>
<!-- variantName: Container for Variant Name ..... -->
<!ELEMENT variantName (resourceRef | resourceData)>
<!ATTLIST variantName
  id ID #IMPLIED
>
<!-- parameters: Processing context for Variant ..... -->
<!ELEMENT parameters (topicRef | subjectIndicatorRef)+>
<!ATTLIST parameters
  id ID #IMPLIED
>
<!-- occurrence: Resources regarded as an Occurrence ..... -->
<!ELEMENT occurrence (instanceOf?, scope?, (resourceRef | resourceData))>
<!ATTLIST occurrence

```

```

    id ID #IMPLIED
  >
  <!-- resourceRef: Reference to a Resource ..... -->
  <!ELEMENT resourceRef EMPTY>
  <!ATTLIST resourceRef
    id ID #IMPLIED
    xlink:type NMTOKEN #FIXED "simple"
    xlink:href CDATA #REQUIRED
  >
  <!-- resourceData: Container for Resource Data ..... -->
  <!ELEMENT resourceData (#PCDATA)>
  <!ATTLIST resourceData
    id ID #IMPLIED
  >
  <!-- association: Topic Association ..... -->
  <!ELEMENT association (instanceOf?, scope?, member+)>
  <!ATTLIST association
    id ID #IMPLIED
  >
  <!-- member: Member in Topic Association ..... -->
  <!ELEMENT member (roleSpec?, (topicRef | resourceRef | subjectIndicatorRef)*)>
  <!ATTLIST member
    id ID #IMPLIED
  >
  <!-- roleSpec: Points to a Topic serving as an Association Role .. -->
  <!ELEMENT roleSpec (topicRef | subjectIndicatorRef)>
  <!ATTLIST roleSpec
    id ID #IMPLIED
  >
  <!-- scope: Reference to Topic(s) that comprise the Scope ..... -->
  <!ELEMENT scope (topicRef | resourceRef | subjectIndicatorRef)+>
  <!ATTLIST scope
    id ID #IMPLIED
  >
  <!-- mergeMap: Merge with another Topic Map ..... -->
  <!ELEMENT mergeMap (topicRef | resourceRef | subjectIndicatorRef)*>
  <!ATTLIST mergeMap
    id ID #IMPLIED
    xlink:type NMTOKEN #FIXED "simple"
    xlink:href CDATA #REQUIRED
  >
  <!-- end of XML Topic Map (XTM) 1.0 DTD -->

```