

**met@BiM - Ein semantisches Datenmodell
für Baustoff-Informationen im World Wide Web**

Anwendungen für Beton mit rezyklierter Gesteinskörnung

Von der Fakultät Bauingenieur- und Vermessungswesen
der Universität Stuttgart
zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Abhandlung

Vorgelegt von

Marcus Schreyer
aus Backnang

Hauptberichter: Prof. Dr.-Ing. Hans-Wolf Reinhardt
Mitberichter: Prof. Dr.-Ing. Peter Grübl

Tag der mündlichen Prüfung: 29. Mai 2002

Institut für Werkstoffe im Bauwesen der Universität Stuttgart
2002

Vorwort

Die vorliegende Arbeit wurde während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Werkstoffe im Bauwesen (IWB) an der Universität Stuttgart verfasst. Die Grundlagen dieser Arbeit sowie die in den Beispielen verwendeten Daten aus Materialprüfungen an Beton- und Zuschlagproben stammen aus den Ergebnissen des Verbundforschungsprojekts „Baustoffkreislauf im Massivbau (BiM)“. Der Themenbereich „Umsetzungshilfen“ war dabei Bestandteil des vom Bundesministerium für Bildung und Forschung (BMBF) sowie vom Deutschen Ausschuss für Stahlbeton (DAfStb) geförderten Projekts. Mit Hilfe moderner Informations- und Kommunikationstechnologie sollte unter den Teilprojekt-Titeln „I 01 Erstellen einer Datenbank“ und „I 02 Erstellen eines Expertensystems“ eine Software zur rascheren Umsetzung der umfangreichen Forschungsergebnisse in der Baupraxis entwickelt werden. Die während der Bearbeitung des Informationssystems BiM-Online gesammelten Erfahrungen und Ideen gaben den Anlass, mich vertiefter mit dem Gebiet der semantischen Datenmodellierung zu befassen.

Aus den zahlreichen Personen in meiner beruflichen und privaten Umgebung, die mir in Gesprächen direkt und auch unbewusst bei der Erstellung dieser Arbeit geholfen haben, möchte ich einige besonders hervorheben. An erster Stelle danke ich Herrn Prof. Dr.-Ing. Hans-Wolf Reinhardt für die langjährige Betreuung und die Unterstützung, die ich bei meiner Arbeit auf diesem Gebiet von ihm immer wieder erfahren habe. Er hat auch die Übernahme des Hauptberichtes für diese Arbeit übernommen.

Für die zahlreichen wichtigen Anregungen und die enge Zusammenarbeit am IWB gilt mein besonderer Dank Herrn Akad. Rat Dr.-Ing. Joachim Schwarte, der meine Faszination für die Möglichkeiten moderner Internettechnologien als Medium zum Wissens- und Informationsaustausch teilt. Die Diskussionen mit ihm haben wesentlich zum Gelingen dieser Arbeit beigetragen.

Herrn Prof. Dr.-Ing. Peter Grübl danke ich für die Übernahme des Mitberichts. Der Software AG Darmstadt, vertreten durch ihr Systemhaus Stuttgart, habe ich für die freundliche Überlassung einer Test-Lizenz ihrer XML-Datenbank Tamino zu danken.

Allen Mitarbeitern des Instituts für Werkstoffe im Bauwesen danke ich für ihre Hilfsbereitschaft und Unterstützung. Ohne die gute Zusammenarbeit mit Kollegen aus den anderen Forschungseinrichtungen, Hochschulen und Unternehmen, die mir im Rahmen meiner Arbeit an BiM-Online bereitwillig Zugang zu ihren Forschungsergebnissen gewährt haben, hätte diese Arbeit nicht in der vorliegenden Form entstehen können. Die Diplomarbeiten von T. Abele, O. Bingel, M. Luik, T. Weikert und J. Koptik lieferten ebenfalls einen wichtigen Beitrag zu den betontechnologischen Grundlagen. Letztendlich möchte ich noch die Unterstützung meiner Familie und meiner Freunde erwähnen, ganz besonders von Cornelia und Dorothee Wirths für ihre Hilfe und Geduld bei der Korrektur des Manuskripts.

Stuttgart, im März 2002

*Im Grunde sind es doch die Verbindungen
mit den Menschen, welche dem Leben seinen
Wert geben.
(Wilhelm von Humboldt)*

Summary

met@BiM - A Semantic Data Model for Building Materials Information on the World Wide Web (Demonstrated on Concrete with Recycled Aggregates)

The met@BiM data model that is described in this dissertation allows an automated content based retrieval of information resources about concrete with recycled aggregates in the World Wide Web (WWW). The proposed concept builds thematically and methodologically upon the online system “BiM-Online”, which documents the results of the Germany wide cooperative research project entitled Life Cycle of Concrete and Masonry Structures (BiM) and makes these results accessible to professionals via an interactive web interface.

This work has been motivated by the belief that the WWW can not be used for professional purposes in the communication and information intensive building branch if its contents don not become accessible in a more “engineering like” manner. “Engineering like” in this context means, that more precise queries must be possible on the contents published in the Web. Detailed queries should be able to consider all information resources in the Web regardless of their formatting. Even resources that can hardly be accessed or recognized by full-text search engines, such as logically connected paragraphs of web pages and multimedia (binary) web elements, should be included. Furthermore, the filter criteria in Internet queries must become more sophisticated, so that queries concerning values, like e.g. mixtures in which more than 30% of the aggregate is composed by brick, can be executable over web contents about concrete technology topics.

To realize such a functionality, semantic relations between terms like “aggregates” and “composition” must be established in a computer interpretable (“machine readable”) manner. In addition to this, the possibility to enter comparison operators like “less then” or “greater then” in combination with numeric values and units should be created. For local queries in information systems, such functionalities can be realized with the help of database supported (“dynamic”) web applications. This has been shown in BiM-Online. In contrast to the mostly relational queries used in local queries, external full text search engines in the web generally use computer linguistic analyses of the HTML code of web pages. Out of the poorly semantically structured HTML code, such search engines can only derive very limited information about the content of these documents. The database structure behind dynamic web pages, which contains information about relations between particular web page elements, is hidden from traditional full text based engines. Queries over inhomogeneously structured database supported web sites therefore cannot be realized. The semantic meaning of terms used in a database normally will not be defined explicitly, but can be derived from the names of tables, fieldnames and other similar labels. The ability to abstract the context of a term from its name enables human beings to search in data collections with different structures. To realize functionalities like shown above in the web, this ability has to be transmitted to the computer.

The Information Technology (IT) approach of this dissertation thus deals with web contents in a more structured publication form. In the Semantic Web Initiative, which is lead by the “father” of the WWW Tim Berners-Lee, the World Wide Web Consortium (W3C) develops standards for semantic data models to enable a machine readable description of content based relations in various knowledge domains. Although this does not mean, that the Semantic Web will develop into a second WWW, the Semantic Web will extend the existing by a semantic data layer, which creates an environment for “intelligent” internet services, thereby improving the collaboration between computer systems and increasing the professional usability of the WWW.

Objective

The basic idea of the Semantic Web is to code statements about web contents and content relations in a machine readable way. This is achieved by assigning web contents with structured text descriptions called “metadata”. To realize the machine readability of this metadata is an interdisciplinary task. In the Semantic Web Initiative, a general concept to formally describe such metadata models has been established under the name “Resource Description Framework (RDF)”. To apply this concept to the design of specific metadata models in different knowledge domains, experts in the various fields will have to contribute.

This work shows the development and application of a RDF metadata model in the field of building material information or, more precisely, technical information about concrete with recycled aggregates. Built upon the well established, but very generally defined, metadata elements of the Dublin Core Metadata Standard that are used to describe web documents, the presented met@BiM model offers a finer structure to describe the characteristic properties of information sources about concrete technology in the Web. The development of a complete model, in the sense of targeting the definition of all possible building material properties or presentation formats for this information type, has not been the aim of this work. The scope of met@BiM is limited to metadata elements that allow a better retrieval of web contents about concrete with recycled aggregates. A complete model can only arise, after a metadata model, which is initially incomplete, is applied to different tasks and successively extended. met@BiM is intended to contribute to the establishment of such a comprehensive model for technical information about concrete.

Before metadata structures like met@BiM can be generally used in web applications, some problems had to be solved. First, the met@BiM model had to be coded as a machine readable RDF Schema. The definitions of information resources about concrete technology and their properties in such a RDF Schema were then used to describe concrete information sources in the web, like e.g. results of material tests, testing diagrams or descriptions of mixtures, in a RDF/XML syntax. Additionally, the user needs support to create the often quite extensive and complicated RDF code in the form of software programs. Therefore the design of a met@BiM RDF editor, especially adapted for the input of concrete technological data has also been a part of this work.

To demonstrate the advantages of these metadata structures on practical applications, a corresponding software environment was created in which the met@BiM data model has been implemented. The lack of software to query RDF/XML data structures in the required way, made the achievement of this target difficult. Thus, for the implementation of met@BiM in a prototype system, some software components had to be developed first or, as it was the case with the XML database Tamino, commercial software products had to be adapted. With this software environment, the following questions can be answered:

- How can statements related to concrete technology be modeled using RDF/XML descriptions?
- Which additional functionalities can be realized with met@BiM RDF data structures compared to full text queries?
- Does the additional effort that is necessary to create RDF descriptions, stand in an acceptable proportion to the benefit of the information? How can this effort be reduced?
- Because of the growth of knowledge fields and the need to sometimes define new branches, a data model should reflect this requirement for flexibility. How flexible is the proposed met@BiM model regarding the integration of new metadata elements or information resources?
- How could a topic specific querying user interface for met@BiM records look like?

Methodology

The idea to apply the metadata concept to web content related to concrete technology builds the central approach of this work. With the description of web contents for engineers, the described retrieval problems due to inhomogeneous information collections should be solved. Metadata does not represent a new data type. Rather information resources, like web applications, single web pages or just graphics and sections of pages, will be described and labeled with structured text data and therefore be accessible for text based search technology.

A prerequisite for the design of metadata structures is an analysis of the corresponding information resources in the knowledge domain that is to be described. In the case of the met@BiM the information resources in the domain BiM-Online were grouped by resource types and brought in relation to each other. By doing this the first syntax independent ontology has been created with an emphasis on the modeling of information resources about concrete technology. In the field of knowledge modeling, ontologies formally define the semantic meaning of elements or terms that are used in metadata descriptions through their relations to one another. Ontologies are themselves data models of a higher level and play an important role in the interoperability of metadata models in the Semantic Web. For the definition of the elements of met@BiM, currently accepted standards, such as the Dublin Core, were considered. To make more detailed descriptions of resources possible, this standard had to be extended with additional elements.

To apply the generated metadata model in a standardized syntax for the WWW, two technologies have been used that are in the spotlight of the Semantic Web activities at the moment: the eXtensible Markup Language (XML) and the Resource Description Framework (RDF). XML allows web authors to structure their web pages by using layout commands, which are so-called tags. Tags are labels like *<author>* or *<title>* that enclose every element in a XML document. These XML tags can be labeled in a meaningful way for human readers, but for a computer, XML tags offer no clue, about how to interpret the meaning of these structural elements. In opposite to that, RDF was developed with the intention to model the meaning and relations of web contents. Statements that are contained in information sources are reformulated as predicate logic subject-predicate-object triples and assigned to the resources in the form of metadata records in RDF descriptions. These RDF metadata records can be coded in XML syntax. RDF is therefore often named a XML application.

In the concept of RDF there exists a distinction between the RDF recommendation for the design of RDF descriptions and the candidate recommendation for the RDF Schema standard. RDF schemes define the resources and properties as well as the relations between such elements in RDF descriptions. The language used to express RDF schemes is usually RDF/XML syntax. RDF schemes therefore are equivalent to formalized ontologies and thus build through the definition of metadata elements a basis for the design of consistent RDF descriptions in a decentralized network like the WWW.

The method proposed in the RDF Schema standard has been used to model the concrete technology ontology of the BiM-Online knowledge domain. The first result of this work has been a serialized RDF Schema of the BiM-Online metadata model. The unambiguous affiliation of metadata elements and resources to a RDF Schema can be guaranteed by the namespace concept of XML. The RDF Schema of the met@BiM model thus is the foundation for a system predominant description of web contents relating to concrete technology. The elements defined within this RDF Schema have further been used to describe content such as scientific reports, testing results, concrete mixtures and chemical or mathematical formulas. The generated RDF descriptions in XML syntax were then integrated into the source code of web contents (web pages, courses etc.) in the form of so-called XML “data islands”. The querying of the data was realized in the absence of a standardized RDF querying language by

using the X-Query language of the Software AG's Tamino XML server. X-Query is a querying language that is based on the XPath specification published by the W3C. By using X-Query, the most important advantages of machine readable RDF/XML metadata structures for the retrieval of building material information have been demonstrated.

Results

It has been shown that the met@BiM RDF Schema presented in this work can serve as a foundation for the creation of machine readable RDF descriptions of concrete technology information resources in the WWW. Additionally, the scheme allows internet services, to access resources that exist in differently organized systems and to classify them automatically in a semantic context. The hierarchical ontology for BiM-Online consists of the following information resource types: document, course, concrete data, mathematical formula, chemical formula.

To offer the possibility for detailed context assignment for concrete technology information resources, the ontology branch assigned to the RDF schema element *concrete data* had to be refined. To achieve this refinement, new classifications to group diagrams, tables, etc. that are assigned to the resource type *concrete data* had to be developed. A starting point for such a further classification is provided by the composition of concrete: Concrete is a building material that consists of different components. In the diagrams and tables the properties of these components are often related to one another or to the properties of the concrete as a whole. Consequently, these components can be used to give material properties a semantically more precise context. Furthermore, a distinction has to be made between the properties of fresh and hardened concrete, the testing properties and also the composition of the mixtures. Considering the components and the production steps, the structure presented in this dissertation has been developed.

The degree to which the ontology has been detailed was adapted to the number and scope of the material tests of the BiM project. A further extension or refinement of the chosen structure can be made without requiring changes to the existing datasets. The descriptions of all information resources in BiM-Online are based on the metadata elements of the Dublin Core scheme. The definitions for every single Dublin Core data element were specified for information resources of BiM-Online. Therefore, the topic related metadata elements of met@BiM can be considered as an extension of the general Dublin Core metadata standard. The elements of the met@BiM scheme are assigned in the form of a complex data underneath the Dublin Core element *dc:subject*. This combination of specific met@BiM RDF Schema elements with elements provided by the Dublin Core RDF Schema has been applied on selected information resources. The detailed reflections and steps needed to design a complete RDF description are also described in this work for every example shown.

In spite of the opportunity for more detailed metadata, the method that is shown in this work allows a relatively fast description of information sources about concrete. Software tools, like the met@BiM RDF editor, have been created to help the user (e.g. a researcher or engineer) generate the RDF code. This reduces the time needed to enter data. Because of the dependence of a user interface for such an editor and the data structures that it should generate, the RDF editor had to be specially developed for met@BiM. Its user interface lets the user concentrate on inputting of the data, because the generation of the corresponding RDF code is automatically done by the editor program. Another major influence on the effort required to produce the RDF code stems from the flexibility of being able to describe statements about concrete technology in multiple-stages. met@BiM allows different stages to detail the descriptions and therefore delegates the decision about the expenditure to describe resources to the authors. This way, the user has the choice to model only the most important

building material properties or if he further chooses to enter values for these properties in the form of a range of values or specific discrete values. Physical units and dimensions of values are modeled by the met@BiM element *metabim:einheit* (German for “metabim:unit”). This solution to model the entire expression in a single metadata element as text data is at first glance an appropriate solution, which is optimized for the purposes of modeling BiM-Online content. Because a more sophisticated solution for the representation of units and dimensions for a range of scientific fields is required, it can be expected that there will soon be proposals for a standardized modeling of this problem.

Another task was to develop a more “intelligent” search engine prototype for met@BiM web resources, which consists of software modules for a predominantly automatic data collection, and data storage and a user interface for queries. Oriented on the methods of current search engines, the RDF descriptions distributed over the Web should be accumulated in a central index database. Using the solution realized in the prototype system, a user can load the descriptions into the index by simply entering the URL of a RDF enriched web page into a HTML form. Because the RDF descriptions are part of the source code of web pages, another software module is needed to separate the RDF/XML fragments from the HTML code. For this purpose, a special preprocessor module has been programmed. It has been realized as a server sided PHP application and is located between the HTML form for the collection of web pages and the actual data interface X-Node of the Tamino database. Tamino is a commercial product from Software AG. In the prototype system, Tamino is used as a central data storage and management software for the collected RDF/XML data islands. It provides a web oriented querying interface that can be used to filter certain XML objects out of the index database and to process them afterwards in tandem-arranged programs. Because RDF is relatively new, there is still a need for matured software products that really make use of the opportunities that RDF data structures offer. To realize the prototype system, RDF descriptions have been stored as XML documents in the index database and were filtered with a XML querying language (X-Query).

The actual user interface of the prototype system, which is build upon this querying interface, is designed as a combination of option fields, with default entries of the met@BiM scheme, and input fields for values, units etc.. It allows one to formulate queries using two different approaches depending on the desired specification grade. On the one hand, keyword based queries can be executed across all information resources using metadata descriptions defined according to the BiM-Online ontology. The query can also be limited to certain resource types via option fields. In this case, resulting lists offer hits for different resource types (if the search not limited to one resource), where each hit contains information about the context of the listed resource. The context will be displayed in the form of a chain of specialized ontology elements respective resource-type names. For more detailed queries, more specialized query masks exist, which allow the formulation of more precise queries on documents, concrete technology content, etc..

The records of the hit lists display the most important properties of the information resources which they refer to. In addition to obligatory details like title, author, resource description, etc., these records were extended by semantic details, such as the resource type or context explained above and the listing of further material properties that are also described in the information resource. With these semantic details the user is able to classify the resource offered by the search engine and can judge the relevance of the content before he downloads it. Another feature is the opportunity to show information sources that do not directly result from the query, but are related to the web resources that are shown in the according hit list. By following the hyperlink “*Weitere Inhalte hierzu... (Related contents)*” additional web content will be displayed that is semantically connected to the actual record via the data in the metadata elements *dc:relation* or *dc:source*. Via these two elements, content based relations

like the corresponding mixture for the displayed testing results can be modeled and subsequently automatically employed.

The examples in this dissertation show the advantages of RDF metadata sets for a web based retrieval of building material information for engineering purposes. The presented functionalities are summarized below:

More detailed queries with comparison operators (e.g. “<”, “>”) for numerical concrete technology data can be formulated by giving the particular material property, the algebraic comparison operator, the value and possibly its unit and dimension. According to conventional relational databases, more than one material property can be combined in queries across XML data structures using Boolean operators. Especially for the query of building material diagrams or tables with test results that show relations between concrete parameters, the combination of different parameters is a useful functionality to specify in a query.

As explained in the preceding section, the met@BiM RDF descriptions allow programs to automatically derive the semantic context of web content from the relation of building material properties and their corresponding concrete technological information resources defined in the met@BiM RDF scheme. Therefore, in the case of a large number of hits for a query, a search engine is able to request particular specifications of query terms by displaying the collected contexts using resource-types property combinations to choose from.

Another semantic functionality that can build upon the met@BiM data model is the possibility, that programs can derive explicitly modeled or implicitly inherent semantic relations from the RDF data structures shown in this work. An example of explicitly modeled information in this context is data, such as additional concrete parameters, that are shown in the same diagram and modeled in the same RDF description. Another example would be displaying the web address of the superior system, in which the web document is integrated. Because all met@BiM RDF descriptions of a testing series will relate to the same RDF description for a concrete mixture, the URL in the element *dc:relation* can be used to deduce an implicit connection between particular testing results.

With the help of case scenarios, the additional functionalities of RDF based queries over current full text queries can be demonstrated within the prototype system.

Prospects

The final chapter of this dissertation is dedicated to the question of how topic specific metadata models like met@BiM can become integrated into the global Semantic Web. To demonstrate the advantages of a combination of full text and semantic retrieval strategies, a vision of a future semantic search engine is introduced. Queries in such a system would require two steps. At the beginning, the search service would clarify the context of the query on a dialog bases. If necessary, the query would then be delegated to more specific services for the corresponding knowledge domain.

The organizational prerequisites that are necessary for the development of semantic metadata models in particular topic areas are examined. In the vision of a global Semantic Web, the met@BiM metadata model proposed in this work covers the semantic description of web contents for the knowledge domain concrete technology with particular regards to concrete with recycled aggregates. A wide gap exists between the sometimes very specialized elements of the met@BiM RDF Schema and the often very generally defined elements of the Dublin Core standard. A superior metadata scheme for the building industry is missing here.

Because of the decentralized development of metadata models and the thematic overlaps that occur in many knowledge domains, concurrent metadata definitions will evolve. To avoid these uneconomical parallel developments, it would be advantageous to coordinate these efforts in an international board with the intention to develop a commonly accepted metadata model for the building industry right from the beginning.

While this is an early investigation about web based metadata models for the building industry, further investigations are necessary to generalize the met@BiM model in certain areas and to complete it with additional elements. This is particularly true for the met@BiM elements for units of measurement. Additional metadata elements for concrete will arise in projects with a different emphasis, e.g. emphasis on the economical aspects of the concrete production, the processing of concrete and its usage and disposal.

It is hoped that future investigations will close the above mentioned gap during the growth of a Semantic Web and that the proposed metadata model in this dissertation will be a contribution towards a World Wide Web that offers an improved usability concerning civil engineering specific information.

Inhalt

1	Einführung.....	1
1.1	Ausgangssituation	1
1.2	Die Idee des „Semantic Web“	2
1.3	Wozu ingenieurgerechte Datenmodelle	7
1.4	Aufbau und Zielsetzung dieser Arbeit	9
2	Grundlagen des WWW.....	12
2.1	Entwicklung des World Wide Web.....	12
2.2	URIs - Das Adressierungskonzept des WWW.....	14
2.3	Publizieren in HTML	15
2.4	Die „Hypothek“ von HTML	18
2.5	Datenbankgestützte Web-Sites.....	21
2.6	XML – die Grundlage zukünftiger Publikationsstandards.....	23
2.6.1	Grundprinzipien von XML	23
2.6.2	XML im World Wide Web.....	28
2.6.3	XML-Anwendungen mit Relevanz für das Bauingenieurwesen.....	32
2.7	Recherchen im World Wide Web	35
2.7.1	Redaktionelle Katalogsysteme und Portalsites	35
2.7.2	Suchmaschinen	37
3	Baustofftechnologische Informationen im WWW	41
3.1	Webbasierte Informationsverarbeitung im Bauwesen	41
3.2	Rechercheaspekte	46
3.2.1	Quellen für baustofftechnologische Informationen	47
3.2.2	Darstellungsformen für baustofftechnologische Informationen im Web	48
3.3	Anforderungen an Datenmodelle für ingenieurgerechte Recherchen	51
4	Metadatenmodelle zur Beschreibung von Baustoff-Informationen	54
4.1	Grundlagen zu Metadaten	54
4.1.1	Was sind Metadaten.....	55
4.1.2	Vorteile der Verwendung von Metadaten.....	57
4.1.3	Überlegungen zum Umfang und der Komplexität der Datenstrukturen.....	57
4.1.4	Speicherung von Metadaten	58
4.2	Metadatenmodell nach Dublin Core	60
4.3	Entwurf von Metadatenstrukturen in BiM-Online	63
4.3.1	Informationsressourcen in BiM-Online.....	65
4.3.2	Anwendung der Dublin Core Metadaten in BiM-Online	67
4.3.3	Detailliertes Metadatenmodell für Daten zu rezykliertem Beton	69

5	Das Resource Description Framework (RDF) – Metadaten im Internet.....	74
5.1	Allgemeine Architekturasspekte.....	74
5.2	RDF-Datenmodell	76
5.3	RDF-Syntax.....	77
5.4	RDF-Schemata	81
5.5	Interoperabilität durch RDF	83
5.6	Anwendungsgebiete von XML/RDF	85
5.7	Aktuelle Softwaretools	86
5.7.1	RDF-Parser	86
5.7.2	RDF-Editoren	87
5.7.3	RDF-Datenbanken	87
5.7.4	Implementation in Mozilla / Netscape Communicator.....	88
5.7.5	Implementation in aktuellen Suchmaschinen	88
6	RDF-Anwendungsszenarien zu „Beton mit rezyklierter Gesteinskörnung“	90
6.1	RDF-Schema für met@BiM	90
6.2	Entwurf der RDF-Datensätze	96
6.2.1	RDF-Beschreibung von Dokumenten und Online-Kursen.....	96
6.2.2	RDF-Beschreibung von Gleichungen.....	102
6.2.3	RDF-Beschreibung von Betonprüfdaten und Rezepturen	105
6.3	Diskussion der met@BiM-RDF-Modelle	116
6.4	met@BiM-RDF-Editor	118
6.4.1	Programmstruktur	119
6.4.2	Eingabeoberflächen für die einzelnen Ressourcentypen	119
6.5	Implementierung von met@BiM	123
6.5.1	Speicherung der RDF-Datensätze im HTML-Quellcode	123
6.5.2	Datenerfassung über den Preprozessor	125
6.5.3	Aufbau einer RDF/XML-Indexdatenbank mit Tamino	126
6.5.4	Abfragemodul für met@BiM-Datensätze	127
6.6	Abfrageszenarien für met@BiM-Datensätze	131
7	Schlussbetrachtung	138
7.1	Einbindung von met@BiM in ein Semantic Web.....	138
7.2	Zusammenfassung.....	141
8	Literaturverzeichnis.....	144
Anhang A	met@BiM-Schema	154
Anhang A1	Ontologie betontechnologischer Informationsressourcen	154
Anhang A2	Schema Festbeton	154

Anhang A3	Schema Frischbeton.....	155
Anhang A4	Schema Betonzuschlag	156
Anhang A5	Schema Zement	157
Anhang A6	Schema Herstellung.....	157
Anhang A7	Schema Zusatzmittel.....	157
Anhang B	XML-Serialisierung des met@BiM-RDF-Schemas.....	158
Anhang C	XML-Serialisierung des Dublin Core-Metadaten-RDF-Schemas.....	169
Anhang D	Beispiele für RDF-Beschreibungen	173
Anhang D1	Einfluss der Rohdichte rezyklierter Zuschläge auf die Betondruckfestigkeit (Tab. 6.1)	173
Anhang D2	Carbonatisierung von Beton mit unterschiedlichen Brechsandanteilen im Zuschlag (Abb. 6.7)	175

1 Einführung

Die Aufgabenstellung dieser Arbeit entwickelte sich aus den Erfahrungen, die am Institut für Werkstoffe im Bauwesen (IWB) der Universität Stuttgart bei der Erstellung des Online-Informationssystems BiM-Online [BiM-Online2001] gesammelt werden konnten. Das vom Bundesministerium für Bildung und Forschung (BMBF) und dem Deutschen Ausschuss für Stahlbeton (DAfStb) geförderte Verbundforschungsprojekt „Baustoffkreislauf im Massivbau (BiM)“ hatte zum Ziel, die Grundlagen für eine optimierte Kreislaufwirtschaft im Bereich mineralischer Baumaterialien zu schaffen. Die Materialeigenschaften von Beton mit rezyklierter Gesteinskörnung sowie die technischen Abläufe, die mit der Aufbereitung der Materialien einhergehen, mussten so definiert werden, dass die Voraussetzungen für eine wirtschaftliche Nutzung erfüllt wurden. In das deutschlandweit und interdisziplinär durchgeführte Forschungsprojekt waren zwei Teilprojekte integriert, welche die Dokumentation und die Bereitstellung der Ergebnisse für die Baupraxis mit den Mitteln moderner Informationstechnologie zur Aufgabe hatten.

Als Ergebnis entstand eine datenbankgestützte Web-Site, die über mehrere inhaltspezifische Abfrageoberflächen einen raschen und gezielten Zugriff auf die Forschungsergebnisse von BiM ermöglicht. Die Verwendung von relationalen Datenbanken zur Speicherung komplexer Datenmodelle ist jedoch für die Recherchierbarkeit und Referenzierbarkeit der darin verwalteten Informationen im Web mit gewissen Nachteilen verbunden, deren Relevanz in jüngerer Zeit immer stärker erkannt wird.

Die Ursache für diesen Informationsverlust, der beim Zugriff von außen auf BiM-Online entsteht, liegt in den unterschiedlichen Datenmodellen der Datenbank des Webservers [Reinhardt2000] und der Hypertext Markup Language (HTML) für Dokumente des World Wide Web begründet. Die Inhalte von HTML-Webseiten sind in ihrer Darstellung auf den menschlichen Betrachter ausgerichtet. HTML bietet wenig Möglichkeiten, die inhaltliche Struktur der Dokumente für die automatische Interpretation durch Programme aufzubereiten. Eine automatische Kategorisierung von Webinhalten durch Suchmaschinen ist daher ohne erweiterte Beschreibungsmöglichkeiten nicht möglich.

So lange lediglich wenige Inhalte zum Thema „Kreislaufwirtschaft im Massivbau“ im Web existieren, bleiben die Abfrageoberflächen von BiM-Online und damit die Inhalte des Systems für den Anwender leicht auffindbar. Durch das explosive Wachstum des WWW wird sich die Notwendigkeit jedoch verstärken, verbesserte Suchdienste mit der Möglichkeit feinerer, inhaltsbezogener Rechercheparameter zur semantischen Ausfilterung irrelevanter Inhalte bereit zu stellen.

1.1 Ausgangssituation

Wie kaum eine andere technische Entwicklung zuvor hat der Internetdienst World Wide Web (WWW) unseren beruflichen und privaten Umgang mit dem Computer verändert. Im Zentrum dieser Entwicklung steht dabei die „Virtualisierung“ von Informationen in Datennetzen, wodurch diese unabhängig von Ort und Zeit verfügbar werden. Überwiegend auf Informationen beruhende Dienstleistungen und Geschäftsabläufe wurden auf die multimediale Oberfläche des WWWs übertragen und nutzen dessen anwenderfreundliche und weltumspannende Netzwerkinfrastruktur. Die große wirtschaftliche Bedeutung, die das Internet in der industrialisierten Welt bereits innerhalb weniger Jahre erlangt hat, zeigt, welcher Stellenwert dieser „Ressource Information“ und vor allem ihrer effizienten Nutzung zur Optimierung von Arbeitsprozessen heute zugesprochen wird.

Der wirtschaftliche Umgang mit großen Informationsmengen erfordert jedoch einen entsprechenden Organisationsaufwand, innerhalb dessen Informationen erfasst, publiziert und

bewertet werden. Information wird zum wirtschaftlichen Gut, dass ge-„managed“ werden muss. Die vormalig unter dem Sammelbegriff „Elektronische Datenverarbeitung“ (EDV) bezeichneten Dienstleistungen werden heute häufig unter dem Begriff der Informationstechnologie (IT) als fester Bestandteil von Produktions- und Geschäftsprozessen neu definiert.

Auch das Bauwesen wird durch diesen Paradigmenwechsel vor neue Herausforderungen gestellt. In der Bauausführung stehen nur noch geringe Entwicklungspotentiale zur Verfügung. Bauunternehmen und Planer sehen sich einem durch die Globalisierung verstärkten internationalen Wettbewerb gegenüber, der in den klassischen Aufgabenfeldern des Bauwesens nur noch vergleichsweise geringe wirtschaftliche Wachstumsraten ermöglicht. In der Konsequenz zeigte sich in der Branche in den vergangenen Jahren eine deutliche Tendenz hin zu einer Ausweitung des bisherigen Dienstleistungsangebots. Facility Management, Real Estate Management und die bauherrenfreundliche Bündelung von Bau- und Finanzierungsdienstleistungen über Generalübernehmer seien hier als Beispiele genannt. Diese Dienstleistungen werden in einem hohen Maße durch den Austausch von Informationen geprägt, die heute überwiegend in elektronischer Form verarbeitet werden. Das Internet wurde hierbei im Bauwesen, im Gegensatz zu anderen Branchen, bis Anfang 2000 nur in sehr geringem Umfang als Plattform für Dienstleistungsangebote (Informations- und Kommunikationsangebote) eingesetzt. Dabei bietet die Übertragung von Geschäftsabläufen auf die elektronischen Medien, das sogenannte „eBusiness“, besonders für die kommunikations- und wissensintensive Baubranche Chancen zur Rationalisierung zeit- und damit kostenintensiver Prozesse.

Mit zunehmender Informationsmenge im World Wide Web wird jedoch die Frage nach der Erschließbarkeit der gespeicherten Informationen in diesem als „Datenpool der Menschheit“ beschriebenen Netz ein immer drängenderes Problem. Jeder Webanwender, der schon einmal mit einer komplexeren Recherche an eine der Suchmaschinen des WWW herangetreten ist, kennt dieses Problem: Entweder die Suchanfrage ist zu detailliert formuliert, oder sie ist zu allgemein gehalten, so dass der Anwender die große Zahl der gefundenen Ressourcen nicht sinnvoll recherchieren kann. Die Technologie der Volltextsuchmaschine scheint beim aktuellen Umfang und der raschen Verfallszeit von Informationen im Web an ihrer Leistungsgrenze angelangt. Auch die Frage, wie die wachsende Menge multimedialer Webinhalte effektiv recherchiert werden kann, ist noch unzureichend gelöst. Ist das World Wide Web, das sich in den vergangenen Jahren eben noch als professionelles Informations- und Kommunikationsmedium etablieren konnte, damit inzwischen schon wieder auf dem Weg, zu einer Informationswüste zu veröden?

Forschungsarbeiten, die v.a. in den 80er und 90er Jahren auf dem Gebiet der Künstlichen Intelligenz (KI) entstanden versuchten, die Erschließung großer Informationsmengen durch die Entwicklung „lernender“ Programme zu lösen [CYC1997]. Diese Programme waren mit einem formalisierten Grundwissen, einer „Wissensbasis“, ausgestattet und sollten semantische Zusammenhänge selbstständig erkennen können. Analog zur Assoziationsfähigkeit des menschlichen Gehirns sollten die Programme die zur Lösung von Problemstellungen benötigten Informationen automatisch finden bzw. erfragen und entsprechend zuordnen können. Die aus diesen Arbeiten hervorgegangenen Ergebnisse waren jedoch, gemessen an den zuvor hochgesteckten Erwartungen, nicht immer befriedigend. Der aktuelle Ansatz des „Semantischen WWW“ setzt daher nicht bei den Analysewerkzeugen an, sondern bei den Daten selbst bzw. deren Strukturierung.

1.2 Die Idee des „Semantic Web“

Die Idee der maschinellen Erschließbarkeit von Informationen beschäftigt auch die „Väter des Internets“, insbesondere Tim Berners-Lee, der heute in leitender Funktion am World Wide

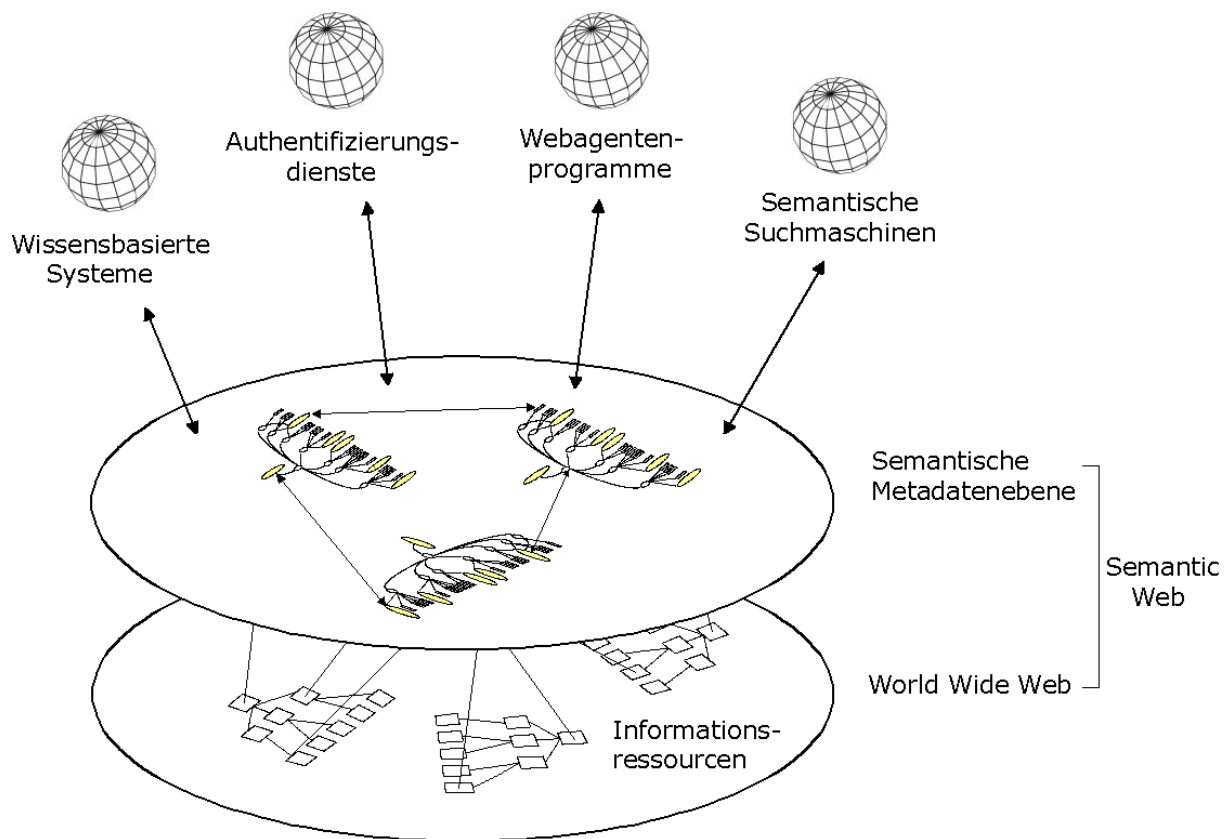


Abb. 1.1 Das Semantic Web erweitert das WWW mit seinen Informationsressourcen um eine maschinenlesbare Metadatenebene für „intelligendere“ Internetdienste.

Web Konsortium (W3C) an neuen Standards für Datenmodelle des World Wide Web der nächsten Generation arbeitet. Diese Standards werden unter dem Sammelbegriff des „Semantic Web“ zusammengefasst. Sie verfolgen die Strategie, dass eine verbesserte und vor allem maschinelle Erschließung großer Datenbestände auf der Basis strukturierterer Daten erfolgen sollte.

Die Ausgangssituation für zukünftige Datenmodelle bildet dabei das WWW aktueller Prägung: Der überwiegende Teil der heutigen Webangebote wurde primär für den menschlichen Betrachter entworfen und nicht für die Verarbeitung durch Computerprogramme optimiert. Das multimediale Konzept des WWW lässt Autoren nahezu völlige Freiheit bezüglich der Präsentation von Informationen. Dieselbe Aussage kann sowohl in verschiedenen Datenformaten z.B. als Text, als bildhafte Darstellung in einem bestimmten Grafikformat als auch als Audiodatei publiziert werden. Diese kreative Freiheit der Präsentationsform von Webinhalten führt jedoch bei der informationstechnologischen Verarbeitung zu Problemen.

Computer können Webseiten sehr rasch hinsichtlich ihres Layouts und enthaltener Programmanweisungen erfassen, sie besitzen jedoch bislang keine Möglichkeit, auf deren Sinngehalt zu schließen. Wer heute mit einer Suchmaschine bspw. nach „Mozart“ sucht, bekommt als Treffer sowohl Seiten mit der Biografie des Komponisten, als auch das Opernprogramm oder Produktinformationen zu marzipangefüllten Schokoladekugeln. In welchem Zusammenhang der Begriff auf der Webseite verwendet wird, kann von der Volltextsuchmaschine nicht ermittelt werden.

Das Semantic Web strukturiert die sinntragenden Inhalte von Webseiten und schafft damit ein Umfeld, in welchem Softwareagenten – dies sind selbstständig operierende Webdienste – komplexe Aufgaben für den Anwender durchführen können. Derartige Agenten werden nicht nur die Stichworte („Keywords“) „Mozart, Don Giovanni, Staatsoper, Stuttgart“ einer Webseite erkennen, wie dies auch schon von heutigen Suchmaschinen geleistet wird, sondern solche Programme werden dem Webangebot darüber hinaus auch entnehmen können, dass die *Oper mit dem Namen Don Giovanni komponiert von Mozart im Großen Haus an der Staatsoper Stuttgart an 5 Terminen im Mai* dieses Jahres aufgeführt wird und dass die lokale Reservierungsdatenbank einen *Datumsbereich* im yyyy-mm-dd Format entgegennimmt und *verfügbare Karten* zurückgibt.

Alle diese Zusammenhänge wären im Semantic Web in Form von den Webinhalten zugeordneten Beschreibungen, sogenannten Metadatensätzen, maschinenlesbar codiert. Diese Beschreibungen erlauben es Programmen, selbständig Schlussfolgerungen bezüglich der Struktur und Bedeutung von Webinhalten zu ziehen. Mit dem Semantic Web wird dabei kein zweites World Wide Web entstehen, vielmehr wird das Bestehende durch eine semantische Datenschicht erweitert (Abb.1.1). In einem ersten Schritt werden dabei bestehende Webinhalte um Metadatensätze ergänzt. Auf diese Weise können die schon im WWW enthaltenen Informationen in einer formalisierten Art und Weise strukturiert werden, was sowohl die Zusammenarbeit von Computersystemen als auch die zwischen Systemen und Menschen verbessert. Die renommierten Softwarehäuser, die sich an der Entwicklung des Semantic Web beim W3C beteiligen, und die Intensität, mit der dieses Thema bei Internetsoftwareentwicklern diskutiert wird, lassen hoffen, dass schon in naher Zukunft die so geschaffenen semantischen Datenstrukturen die Grundlage für neuartige und leistungsfähigere Funktionalitäten bilden werden.

Im Zentrum des Semantic Web stehen zwei Technologien: Die eXtensible Markup Language (XML) und das Resource Description Framework (RDF). XML ermöglicht es jedem Webautor, eine beliebige Strukturierung seiner Webseite anhand von Steuerbefehlen, sogenannten Tags, vorzunehmen. Tags sind Bezeichner in Form von <Autor> oder <Titel> welche die einzelnen Elemente eines XML-Dokuments einfassen. Diese XML-Tags lassen sich zwar für menschliche Leser sinnvoll benennen, geben jedoch einem Computer keine Hinweise für das Verständnis der Bedeutung der Strukturelemente. Im Gegensatz dazu wurde RDF mit der Absicht entworfen, die Bedeutung von Webinhalten und deren inhaltliche Zusammenhänge abzubilden. Diese mit Hilfe des RDF Formalismus erzeugten Metadatensätze können ebenfalls in der XML-Syntax codiert werden. RDF wird daher auch als eine XML-Anwendung bezeichnet.

Damit das Semantic Web funktionieren kann, müssen Computer Zugang zu derartig strukturierten Informationssammlungen bekommen sowie über einen Satz von Inferenzregeln verfügen, mit welchen sich Schlussfolgerungsautomatismen erstellen lassen. Auf dem Forschungsgebiet der Künstlichen Intelligenz wurden solche Modelle unter der Bezeichnung „Wissensrepräsentationsmodelle“ bereits lange vor der Entwicklung des World Wide Web erforscht. Bisherige Wissensrepräsentationsmodelle zeichnen sich jedoch durch ihre zentralisierte Organisationsform aus: Es wird eine exakte, eindeutige Definition für die verwendeten Bezeichnungen innerhalb des Systems gefordert. Solche zentralisierten Modelle geraten bei schnell wachsenden, sehr großen Informationsmengen wie dem WWW sehr schnell an ihre Grenzen und lassen sich dann nicht mehr verwalten. Weiterhin sind die möglichen Fragestellungen, die solchen Systemen gestellt werden können, eng begrenzt. Damit wird sichergestellt, dass der Computer stets zuverlässig eine Antwort liefern kann. Dieses Problem erinnert an Gödels mathematisches Theorem [Gödel1962, Hofstadter1979], welches besagt, dass ein System, das ausreichend komplex ist, um nützliche Aussagen zu erlauben, auch nicht beantwortbare Fragen, wie z.B. Paradoxien, einschließt. Um derartige

Probleme zu vermeiden, besaßen bisherige Wissensrepräsentationssysteme jeweils ihre eigene, sehr eng definierte Regelsprache, die auf ihre jeweilige Datenbasis angewandt wurde. Dies hatte jedoch zur Folge, dass selbst dann, wenn die Daten von einem System in ein anderes übernommen werden konnten, die Übernahme für die im System definierten Inferenzregeln nicht ohne weiteres möglich war.

Im Semantic Web hingegen wird das Auftreten von Paradoxien und unbeantwortbaren Fragen für die vielseitige Anwendbarkeit des Modells in Kauf genommen [Berners-Lee2001]. Diese Entscheidung für ein universelles, jedoch bereits vom Konzept her unvollständiges, System weist starke Parallelen zur Anfangsphase des World Wide Web auf. Durch die Realisierung des Semantic Web auf den bestehenden Informationsstrukturen des WWW ist es auch denselben Anforderungen wie dieses unterworfen. Wie im gesamten Internet gilt auch im Semantic Web das Paradigma der Dezentralisierung mit all seinen Konsequenzen. Das WWW wurde beispielsweise durch seine dezentrale Organisation des Ideals der völligen Konsistenz aller Verknüpfungen beraubt, die sich bspw. in der Fehlermeldung „Error 404: (File) Not found“ äußert, wenn ein Hyperlink sein Ziel nicht mehr findet. Gleichzeitig erlaubte dieser Verzicht auf eine zentrale Verwaltungsinstanz jedoch auch ein unbeschränktes Wachstum der Inhalte und trug somit wesentlich zum Erfolg des WWW bei. Eine weitere wichtige Eigenschaft des World Wide Web ist seine Universalität. Entsprechend müssen sich die semantischen Beschreibungen auch auf beliebige Themengebiete in beliebiger Detailliertheit anwenden lassen.

Natürlich muss ein Webdienst, welcher Informationen aus verschiedenen Quellen kombiniert, zur korrekten Interpretation der Informationen auch Wissen über deren Bedeutung erlangen können. Nur so können Programme z.B. erkennen, dass zwei unterschiedliche Begriffe mit derselben Bedeutung einander entsprechen. Im Idealfall muss dem Programm ein Weg offen stehen, die allgemein definierte Bedeutung von Begriffen abzufragen. Ein solcher Weg soll über eine weitere Komponente des Semantic Web bereitgestellt werden: Den Ontologien. In der Philosophie wird dieser Begriff für eine Disziplin verwendet, die sich mit der Natur des Seins – der Typisierung des Existierenden – beschäftigt. In der Informatik wird dieser Begriff etwas weniger abstrakt für übergeordnete Datenmodelle verwendet, in denen Beziehungen zwischen Begriffen für ein bestimmtes Wissensgebiet, einer sogenannten Wissensdomäne, formal definiert werden [Gruber1993].

Die Entsprechung zu einer Ontologie findet man im Konzept des Semantic Web bei dem Resource Description Framework Schema (RDFS) wieder. Innerhalb des RDFS werden, ähnlich wie im objektorientierten Paradigma [Hartmann2000, Heuer1997], Klassen von Objekten und die Beziehungen dieser untereinander abgebildet (Abb. 1.2). Ein einfaches Beispiel für solche Beziehungen wäre der Begriff einer *Adresse*, die wiederum dem Begriff *Stadt* zugeordnet wäre und der Postleitzahl *PLZ*, die ebenfalls dem Begriff *Stadt* zugeordnet wäre. Die Beziehungen zwischen den Begriffen können dabei über die Zuordnung von Eigenschaften und die Vererbung dieser Eigenschaften an untergeordnete Klassen modelliert werden. Ein entsprechendes Programm könnte anhand solcher formaler Wissensmodelle aus der Angabe der *PLZ* auf die entsprechende *Stadt* schließen und weiterhin darauf, dass es sich um eine *Adresse* in Deutschland handelt. Daraufhin könnte es den Datensatz entsprechend den deutschen Standards formatieren. Auch in diesem Fall „versteht“ ein Computerprogramm nicht wirklich den Begriff *PLZ*. Es wird ihm aber ermöglicht, aus den Beziehungen der Ontologie gewisse Schlussfolgerungen hinsichtlich des Umgangs mit den Daten zu ziehen.

Ontologien werden die Funktionalitäten des WWW auf vielfache Weise erweitern: Grundsätzlich können sie dazu benutzt werden, eine bessere Erschließung von Webinhalten über entsprechende Recherchedienste zu ermöglichen. Weiterentwickeltere Programme werden Ontologien dazu nutzen, die in Metadatensätzen abgebildeten semantischen Beziehungen in ihre aufgabenspezifischen Weltmodelle einzuordnen und Inferenzregeln auf

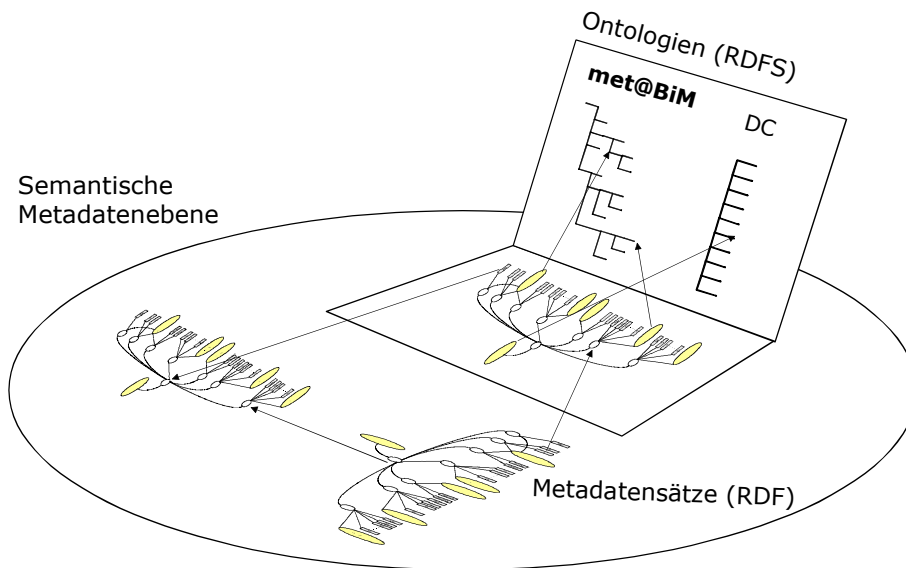


Abb. 1.2 Ontologien definieren die Bedeutung der in den Metadatensätzen verwendeten Elemente. Diese übergeordneten Datenmodelle sind selbst Bestandteil des Semantic Web und beschreiben formal die begrifflichen Zusammenhänge eines bestimmten Wissensbereichs.

die so zusammengestellten Wissensbasen auszuführen. Die automatische Interpretation, d.h. die Zu- bzw. Einordnung von Informationen in einen semantischen Kontext, bildet die Grundlage für zielgerichtete Aktionen intelligenter Computersysteme.

Ein einheitliches Konzept zur Beschreibung von semantischen Strukturen erlaubt entsprechenden Webdiensten weiterhin, auch in unterschiedlichen Online-Informationssystemen automatisch Zusammenhänge zwischen den abgebildeten Daten zu erkennen. Schickt der Anwender im Eingangsbeispiel das Agentenprogramm mit dem Auftrag ins Netz, noch verfügbare *Opernkarten* für die *Abendvorstellungen* in *Stuttgart* zu recherchieren, so genügen dem Agenten diese Angaben von Veranstaltungsart, Stadt und Datum. Als Rückmeldung erwartet er den *Spielort*, *Uhrzeit* und die *Anzahl* der freien Karten. Er übergibt die Abfrageparameter an die jeweiligen Reservierungssysteme, innerhalb derer diese Variablen möglicherweise einen anderen Namen besitzen. Die Unterstützung der Datenübergabe durch Ontologien ermöglicht es dem Agenten jedoch, die Werte automatisch an die entsprechenden Abfrageparameter des lokalen Systems weiterzugeben bzw. die Rückgabewerte des lokalen Systems wieder zuzuordnen. Komplexe, mehrstufige Recherchen werden so mit maschineller Hilfe für den Benutzer verkürzt. Die Weiterverwendbarkeit von Webinhalten in unterschiedlichen Systemen wird durch eine semantische Strukturierung erheblich erleichtert. Die Nutzung elektronisch publizierter Informationen kann auf diese Weise verbessert und ihr qualitativer Wert erhöht werden.

Vergleichbar mit der Einführung des WWW, sind auch die Möglichkeiten eines Semantic Web sehr breit angelegt. Es lässt sich heute noch nicht mit Sicherheit sagen, welche Anwendungsschwerpunkte sich für diese Technologie herauskristallisieren werden. Möglicherweise werden auf dieser Basis Anwendungen entstehen, an die wir heute noch nicht einmal denken. Der Einfluss semantischer Automatismen auf die Funktionalität von Objekten des täglichen Lebens, wie beispielsweise der Interaktion zwischen vernetzten Haushaltsgeräten, wird von Visionären der IT Branche bereits in Szenarien ausgemalt. In dieser Arbeit werden die Möglichkeiten des Semantic Web vor allem unter dem Aspekt einer

besseren Erschließung technischer Informationen über Baustoffe untersucht. Wie im folgenden Abschnitt näher erläutert wird stellt diese Thematik im WWW aktueller Prägung ein noch unbefriedigend gelöstes Problem dar, welches einer professionellen Nutzung des Mediums zur Informationsrecherche im Bauingenieurbereich bislang im Wege steht.

1.3 Wozu ingenieurgerechte Datenmodelle

Im beruflichen Alltag eines Bauingenieurs gestaltet sich die Nutzung des WWW zur Informationsrecherche noch wenig professionell. Die von den Suchmaschinen im Web verwendete Methode der Volltextsuche ist für viele technische Fragestellungen ein nur unzulängliches Hilfsmittel. Dies zeigt sich vor allem bei detaillierten oder komplexeren Abfragen, innerhalb derer mehrere Parameter kombiniert verwendet werden. Bei einer Recherche nach Informationen zu bestimmten Baumaterialien z.B. findet die Volltextsuche zunächst eine Liste von Webseiten. Im günstigsten Fall findet der Anwender auf diese Weise eine auf Baumaterialien spezialisierte Web-Site, wie z.B. eine Online-Datenbank. Solch ein System besitzt für gezieltere Abfragen eine detailliertere Abfrageoberfläche, welche meist auf relationalen oder objektorientierten Datenstrukturen basiert [Hartmann2000]. Der gesamte Rechercheprozess erstreckt sich somit über zwei Systeme, die jeweils unterschiedliche Datenmodelle mit eigenen Erschließungsmechanismen besitzen (Abbildung 1.3).

Die für die Problemstellung des Nutzers relevanten Inhalte müssen von diesem manuell und oft mit großem Zeitaufwand aus den Trefferlisten einer Suchmaschine herausgefiltert werden. Gleiches gilt für die Recherche nach Begriffen, die in mehreren Anwendungskontexten verwendet werden. Die Suchmaschine besitzt ohne die Angabe entsprechender Parameter auf der Webseite keine Möglichkeit zu erkennen, in welchem Kontext z.B. der Suchbegriff „Bond Model“ auf den erfassten Webseiten verwendet wird. Der Anwender erhält sowohl Webseiten zu Darstellerinnen einschlägiger Kinofilme als auch über die Modellierung von Verbundeigenschaften zwischen Werkstoffen.

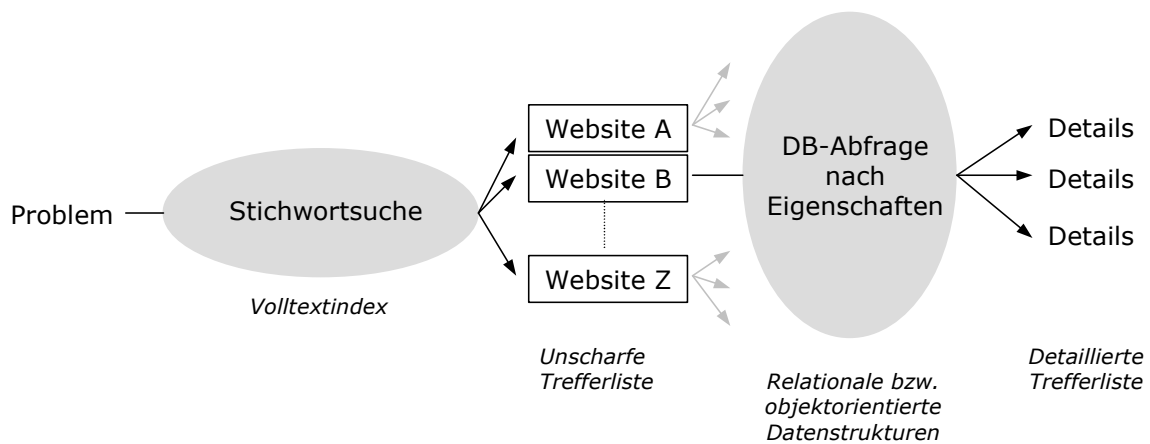
Exaktere Recherchen in sehr großen Datennetzen wie dem WWW werden dann möglich, wenn Webinhalte mit den im vorangegangenen Abschnitt angesprochenen semantischen Beschreibungen versehen werden. Eine Suchmaschine könnte dann Webinhalte anhand der ihnen zugeordneten Eigenschaften identifizieren: Beispielsweise könnten für die Recherche nach einer bestimmten Betonrezeptur die Parameter für die Zuschlagzusammensetzung mit „Ziegelsplitt“ und für die Druckfestigkeitsklasse mit „C25/30“ belegt werden.

Eine weitere Verkürzung der Recherchewege wird durch die Ausführung von Inferenzen auf den semantischen Datenstrukturen ermöglicht. Fehlende Parameter können durch diese „intelligenten“ Algorithmen selbständig aus den semantischen Zusammenhängen geschlossen werden. Dies erlaubt eine Minimierung der benötigten Eingabeparameter.

Ein Problem, das besonders die Nutzung des Web im Ingenieurbereich betrifft, ist das Fehlen von relativen Abfragen mit den Vergleichsoperatoren „<“ „>“ „=“ etc., wie sie in technischen Datenbanken üblich sind. Auch auf Sonderzeichen zur Erzeugung mathematischer Symbole oder bauspezifischer Dimensionsangaben kann im technischen Bereich nicht verzichtet werden. Bis die entsprechenden XML-Standards (s.a. Abschnitt 2.6.3) zur Codierung mathematischer Symbole ausreichend verbreitet sind, bieten sich zur Recherche von Variablen bekannter Größen in Formeln (z.B. E-Modul, Sigma etc.) textbasierende Konstrukte (vgl. Kapitel 4) zu deren Beschreibung an. Der Entwurf semantischer Modelle für die einzelnen Wissensbereiche ist eine fachspezifische Aufgabenstellung. Die Informatik steuert hierfür den Formalismus bei, mit welchem Informationsstrukturen abgebildet werden können.

Ein Verständnis der Zusammenhänge in einer Wissensdomäne, der Denkweise der Anwender sowie des Fachjargons ist für die Gestaltung sinnvoller Datenmodelle unabdingbar. Aus

a



b

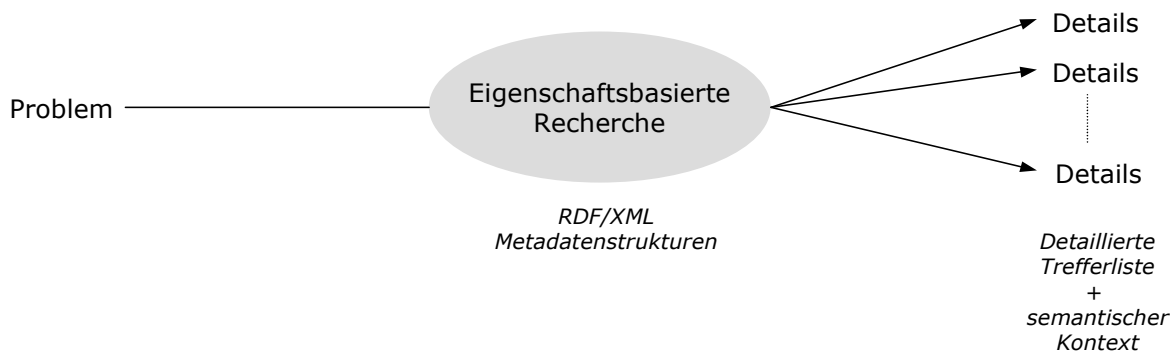


Abb. 1.3 Konventionelle Recherchen im Web erfolgen oft mehrstufig über unterschiedliche Datenstrukturen und Systeme hinweg (a). Semantische Datenstrukturen verkürzen und präzisieren diese Vorgänge (b).

diesen Kenntnissen heraus werden die Parameter Detailliertheit, Tiefe und Umfang des Modells bestimmt. Für die Gestaltung der Datenstrukturen besitzt daher die Fähigkeit des Entwicklers, spätere Nutzerperspektiven zu antizipieren, einen wesentlichen Einfluss auf die Anwendbarkeit des Datenmodells.

Große Verbesserungspotentiale bietet das Semantic Web auf dem Gebiet des Wissenstransfers zwischen Bauforschung und Praxis. Neue Erkenntnisse in der Baustoff-Forschung entstehen meistens zunächst zentralisiert innerhalb von Forschungsprojekten. Sind zu einem Themengebiet noch wenig Informationen vorhanden, ist eine Online-Publikation der Forschungsarbeiten besonders effektiv. Wie das Beispiel des Informationssystems BiM-Online zeigt, erfolgt die Veröffentlichung der Ergebnisse im WWW zunächst nur auf einem oder wenigen Servercomputern der Forschungsgruppe. Durch die Konzentration der Informationen auf wenige Stellen im Web ist die Auffindbarkeit und Abdeckung dieser Wissensdomäne mit herkömmlichen Volltextsuchmaschinen anfangs sehr gut. Verbreitert sich aber der Personenkreis, der sich mit dem neuen Werkstoff beschäftigt, entstehen dadurch auch weitere, voneinander unabhängige Informationsquellen im Internet. Sekundäre Informationsquellen entstehen im WWW, sowie weitere Systeme, in denen lediglich auf bestimmte Aspekte des Baustoffes eingegangen wird bzw. die Thematik für eine besondere Zielgruppe, wie z.B. den

Käufern des Baustoffes, aufbereitet wurde. Mit der wachsenden Zahl solcher Informationsinseln wird eine effiziente Recherche immer schwieriger. Allein über Suchmaschinen ist im Internet bereits nach relativ kurzer Zeit eine vollständige Erschließung der über einen neuen Baustoff verfügbaren Informationen nicht mehr sichergestellt.

Nicht nur für den Wissenstransfer in die Praxis, sondern auch im Bereich der Hochschullehre selbst kommen Webtechnologien zunehmend zum Einsatz [Schwarte2002; Koptik2001]. Die Verbesserungen für den Zugriff auf Informationsquellen im Semantic Web vereinfachen die Nutzung entsprechender eLearning Angebote in unterschiedlichen Softwareumgebungen. Interdisziplinäre Standards [SCORM2002] zur semantischen Beschreibung von Online-Kursen wurden hier bereits erarbeitet. Allerdings enthalten diese Standards bislang nur geringe Möglichkeiten für fachspezifische Erweiterungen der Datenstrukturen. Aktuelle Forschungsarbeiten untersuchen, inwiefern sich das multimediale WWW zur Verbesserung der Vermittlung von Vorlesungsinhalten nutzen lässt [Koptik2001]. Auch in der Baustoffindustrie und deren Verbänden zeichnen sich Bestrebungen ab, internetbasierende Informationssysteme im Bereich der Wissensvermittlung einzusetzen.

Neben Verbesserungen bei der Informationsrecherche eröffnen sich durch semantische Datenmodelle auch Chancen für neue Entwicklungen im Bereich der Bausoftware. Semantische Strukturen erlauben eine engere Verzahnung von lokalen Programmen mit dem Informationsangebot des Web. CAD-Programme können z.B. eine Auswahl von Fertigteilhersteller-Webseiten als Datenbasis nutzen und ohne mehrstufige Recherchen auf die dort angebotenen Detailzeichnungen bei Bedarf zugreifen. Auch verbesserte Software-Assistenten, die ihre Empfehlungen aufgrund der jeweils neuesten Bauvorschriften ermitteln, wären in einem Semantic Web wirtschaftlicher zu realisieren. Voraussetzung hierfür ist, dass diese Dokumente und Daten in entsprechenden Webangeboten in maschinenlesbarer Form zur Verfügung gestellt werden.

Auch im Bereich der Softwareschnittstellen eröffnen semantische Datenmodelle Potential für Verbesserungen. Die erleichterte maschinelle Weiterverarbeitung semantisch definierter Informationen vereinfacht nicht nur die Konfiguration von Schnittstellen zwischen Programmen mit ähnlicher Aufgabenstellung, bspw. dem CAD-Datenaustausch, sondern auch zwischen Software für unterschiedliche Projektphasen (z.B. Planungs-, baubetriebliche Software, Facility Management). Das sogenannte Datenmapping, die Zuordnung sich inhaltlich entsprechender Datensätze, wird auf diese Weise nicht mehr zur Domäne komplex codierter Standards.

1.4 Aufbau und Zielsetzung dieser Arbeit

Die Realisierung des Semantic Web ist eine interdisziplinäre Aufgabe. Die Informationstechnologen sind gefordert, einen universell anwendbaren Formalismus zur Beschreibung der Semantik von Webinhalten zur Verfügung zu stellen. Dieser Forderung wurde vom W3C (vgl. Abschnitt 1.1) durch die Veröffentlichung der Resource Description Framework (RDF) Model and Syntax Specification bzw. der Resource Description Framework Schema (RDFS) Specification 1.0 nachgekommen.

Für die Anwendung dieser Standards sind nun die einzelnen Fachbereiche gefordert. In dieser Arbeit wird ein Datenmodell für das Gebiet der Baustoff-Informationen, genauer der Publikation von betontechnologischen Daten zu Betonen mit rezyklierten Zuschlägen, entwickelt. In einer eigens entwickelten Softwareumgebung wird die Nutzung eines solchen semantischen Metadatenmodells an konkreten Beispielen untersucht. Da RDF-Beschreibungen rasch eine hohe Komplexität erreichen, wird für die effiziente Erzeugung fachspezifischer RDF-Datensätze ein entsprechend angepasstes Softwarewerkzeug benötigt. Diese Aufgabenstellung bildet den Ausgangspunkt dieser Arbeit.

Nach einer Einführung in die der Arbeit zugrundeliegenden Problemstellungen und Konzepte werden in Kapitel 2 die informationstechnologischen Grundlagen der Publikation und Nutzung von Informationen im WWW erläutert, besonders im Hinblick auf die Recherchierbarkeit technischer Inhalte aus dem Ingenieurbereich. In Kapitel 3 werden die verschiedenen Formen der webbasierten Informationsbearbeitung im Bauwesen zusammengetragen. An einem praxisnahen Szenario wird die Notwendigkeit kurzer Recherchewege als Voraussetzung für eine intensivere Internetnutzung innerhalb der Vergabephase, der Bauplanung und Bauausführung dargelegt. Dabei wird auch die besondere Rolle der baustofftechnologischen Informationen bei der Informationssuche innerhalb der genannten Bauphasen herausgestellt und die Art und Weise beschrieben, wie diese heute im WWW veröffentlicht werden. Ziel dieses Kapitels ist es, aus der Analyse des Status Quo die Anforderungen an neue Datenmodelle für ingenieurgerechte Recherchen nach baustofftechnologischen Webinhalten zu formulieren.

Kapitel 4 erläutert zunächst in allgemeiner Form das Konzept der Metadaten. Dabei wird auf den Dublin Core-Metadatenstandard als zur Zeit wichtigste Anwendung des Metadatenkonzepts ausführlicher eingegangen. Aufbauend auf den allgemeinen Definitionen dieses eingeführten Standards wird das met@BiM-Datenmodell zur Beschreibung der kennzeichnenden Eigenschaften betontechnologischer Webinhalte entwickelt. Es wird im Besonderen ein Metadatenmodell zur Erschließung der in BiM-Online gespeicherten Webinhalte zu Beton mit rezyklierten Zuschlägen entworfen. Es ist nicht beabsichtigt, im Rahmen dieser Arbeit ein „vollständiges“ Datenmodell im Sinne der Definition aller Eigenschaften und Informationsressourcen der Betontechnologie zu definieren. Ein solches kann letztendlich nur dann entstehen, wenn durch Anwendung des Datenmodells in Systemen mit unterschiedlichen Aufgabenstellungen dieses sukzessiv um weitere Elemente erweitert wird. Das in dieser Arbeit entwickelte Datenmodell leistet hierfür jedoch einen ersten Beitrag. Mit der Beschreibung des Resource Description Framework (RDF) in Kapitel 5 wird ein geeigneter Formalismus zur Abbildung der Metadaten im WWW vorgestellt. RDF erlaubt für bestehende Ontologien Ergänzungen sowohl in horizontaler Richtung durch Hinzufügen neuer Elemente bzw. Eigenschaften, als auch die Detaillierung von Elementen durch Verfeinerung einzelner Ontologiebereiche. Diese Flexibilität ermöglicht die Integration des hier vorgestellten Modells in übergeordnete oder themenverwandte zukünftige Metadatenstandards im WWW.

Zielsetzung von Kapitel 6 ist es, die Nutzung der zunächst syntaxunabhängig formulierten met@BiM-Metadatenstrukturen in einer konkreten Webanwendung zu demonstrieren. Dabei geht es nicht nur um die Beschreibung von met@BiM als RDF-Schema, sondern auch um dessen Anwendung zum Entwurf konkreter Datensätze. Die Ermittlung der Eigenschaftswerte aus den Informationsquellen und der Aufbau der hier entwickelten RDF-Beschreibungen wird an einigen Beispielen erläutert.

Der Anwender muss bei der Erzeugung des bezüglich seiner Syntax sehr umfangreichen RDF-Codes von Softwarewerkzeugen unterstützt werden. Da die Gestaltung einer sinnvollen Editorsoftware von dem zugrundeliegenden Datenmodell abhängt, musste für met@BiM ein spezieller RDF-Editor entworfen werden. Dieses Programm wird in seinem Aufbau und seiner Anwendung in der Arbeit vorgestellt.

Um die Vorteile der so erzeugten Datenstrukturen an praxisnahen Szenarien demonstrieren zu können, wurde weiterhin eine entsprechende Softwareumgebung entworfen, in die das met@BiM-Datenmodell implementiert wurde. Diese erfasst und verwaltet die in einer Demonstrationsanwendung auf mehreren Webservern verteilten Datensätze über Webinhalte zu rezykliertem Beton und erlaubt ein ingenieurgerechtes Erschließen der betontechnologischen Informationsressourcen.

Die zukünftige Einbindung des met@BiM-Datenmodells in ein universelles Semantic Web sowie die Schlussbetrachtung zu den innerhalb der Arbeit erreichten Verbesserungen für die Recherche betontechnologischer Webinhalte bilden in Kapitel 7 den Schlusspunkt der Arbeit.

2 Grundlagen des WWW

Das Verständnis der in dieser Arbeit vorgeschlagenen Konzepte erfordert zunächst einige grundlegende Kenntnisse über die Publikationsmöglichkeiten für fachspezifische Informationssammlungen im Web. Innerhalb dieses Kapitels werden deshalb die wichtigsten Begriffe und Technologien des Internets in ihren Zusammenhängen eingeordnet und, soweit in diesem Rahmen möglich, kurz erläutert. Der Schwerpunkt liegt hierbei auf Technologien, die innerhalb des Internetdienstes WWW ihre Anwendung finden.

2.1 Entwicklung des World Wide Web

Der Beginn der Entwicklung des Internetdienstes WWW führt zurück bis in das Jahr 1969. In diesem Jahr wurde der Vorläufer des Internet unter dem Namen „ARPAnet“ [ARPAnet] in Betrieb genommen. Es handelte sich dabei um ein militärisch veranlassetes Projekt des US-Verteidigungsministeriums für ein Kommunikationsnetzwerk, das auch im Falle des Ausfalls einzelner Knotenpunkte durch einen nuklearen Krieg noch funktionstüchtig sein sollte. Die ersten Überlegungen hierzu beinhalten auch heute das Grundprinzip des Internets: Ein Computernetzwerk ohne zentrale Kontrolle, bei dem grundsätzlich von einzelnen unzuverlässigen Verbindungen ausgegangen wird. Zu diesem Zweck werden die Nachrichten mit Adresse und Absender versehen und in Pakete eingeteilt. Die Pakete werden von Knoten zu Knoten weitergegeben, wobei der Weg nicht feststeht. Jedes Paket hat seinen individuellen Weg, je nach Verfügbarkeit der Knoten.

1972 gab es bereits 37 Knotenpunkte. Als Übertragungsprotokoll für die Nachrichtenübermittlung wurde der auch heute noch benutzte Standard TCP/IP (Transmission Control Protocol over Internet Protocol) eingeführt. Zu diesem Zeitpunkt wurde das Netz bereits überwiegend zur Kommunikation zwischen den Forschern genutzt und weniger als reines Rechnernetzwerk. Angeschlossen waren nicht mehr nur militärische Forschungseinrichtungen, sondern auch staatliche Organisationen und Universitäten. Anfang der 80er Jahre spaltete sich das Militär ab und gründete das MILNET. Aus dem verbliebenen Netz bildete sich das National Science Foundation Network, welches zunächst 5 große Universitäten der USA zu Forschungszwecken miteinander verband. Eine kommerzielle Nutzung des Netzes war zu diesem Zeitpunkt noch nicht erlaubt. Nur wenige Jahre später begann auch die Wirtschaft das enorme Potential dieses Netzwerks zu entdecken.

Mit steigender Anzahl der im Netz integrierten Computer wurde der Aufruf der Serverrechner über die IP-Nummer immer unübersichtlicher. Den eindeutigen IP-Nummern der Rechner innerhalb von TCP/IP Netzwerken wurde deshalb ein sogenannter Domännennamen zugeordnet. Diese Zuordnung erfolgte mittels eines weiteren Internetdienstes, des Domain Name Service (DNS). Die Server wurden systematisch nach diesen „Domains“ eingeteilt, die Informationen über die Betreiberorganisation und Standort der jeweiligen Server liefern. Ende der 80er Jahre existierten die Domänen *gov* (staatlich), *mil* (militärisch), *edu* (Universitäten), *com* (kommerziell), *org* (nichtkommerziell) und *net* für Gateways zwischen den einzelnen Netzen. Trotz des einsetzenden starken Wachstums im nun als Internet bezeichneten Netzwerk blieb diese Einteilung im Wesentlichen erhalten. Im Laufe der vergangenen Jahre verwischten sich die Grenzen zwischen den Endungen immer mehr und es kamen weitere hinzu. Heute besitzen die meisten Staaten eine entsprechende Endung für die auf ihrem Staatsgebiet verwalteten Server, wie z.B. *de*. Auch weitere, landesunabhängige Domänen wie *tv* (Fernsehen) oder *info* (Informationssysteme) kamen hinzu.



Abb. 2.1 Bildschirmansicht aus den Anfängen des WWW am CERN [Berners-Lee1990]

Entwicklung von Internet Websites

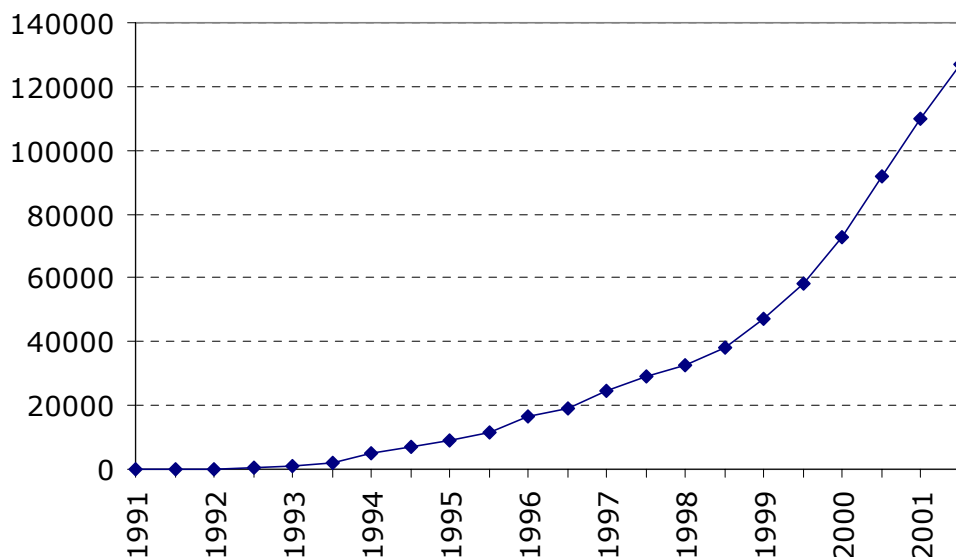


Abb. 2.2 Entwicklung der Web-Sites Anzahl seit 1991 [ISC2001]

In seiner Topologie handelte es sich beim Internet bald nicht mehr um ein einzelnes Netzwerk. Vielmehr wurde durch die gemeinsamen Protokolle und Standards eine Infrastruktur geschaffen, die mehrere größere lokale Netzwerke, z.B. der Universitäten, zu einem globalen Netzwerk zusammenschloss. Bis 1991 wurden im Internet sechs Dienste genutzt. Dies waren Electronic Mail (Email), Terminal Emulation (Telnet), File Transfer (FTP), Newsgroups (Usenet), Wide Area Information Server (WAIS) und das Textdateninformationssystem Gopher.

In den Jahren 1989/90 entwickelte Tim Berners-Lee am CERN (European Organisation for Nuclear Research / Centre Européen pour les Recherches Nucléaires) in Genf den auf Hypertext basierenden Internetdienst World Wide Web [Cailliau1995], um die Fülle der Informationen besser nutzen zu können. Mittels des Konzepts der „Unified Resource Locator“ (URL) wurde die eindeutige Adressierung einzelner Dokumente und Dateien ermöglicht, ohne dass es eine Rolle spielte, ob die Dokumente dabei auf weit entfernten Servern lagen [Schwarte1997b]. Durch den Hypertext konnte man den Benutzer von einem Dokument zum anderen weiterleiten, ohne dass irgendwelche schwer verständlichen Kommandos eingegeben werden mussten. Man benötigte zur Nutzung lediglich ein Hilfsprogramm, einen sogenannten „Webbrowser“, um die angeforderten Informationen am Bildschirm (Abb. 2.1) darstellen zu können. Für die unterschiedlichen Computer- und Betriebssystemplattformen standen bald die entsprechenden Versionen der ersten Browser zur Verfügung. Mit der Entscheidung des CERN, die hinter dem WWW stehende Technik offenzulegen und keinerlei Patentansprüche geltend zu machen, begann der endgültige Siegeszug des WWW.

Die Publikation der Dokumente im WWW erfolgt in der Hypertext Markup Language (HTML). Als Übertragungsprotokoll wurde das Hypertext Transmission Protocol (HTTP) eingeführt. Sowohl HTML, das derzeit in der Version 4.01 vorliegt, als auch die anderen frei verfügbaren Webstandards werden ständig weiterentwickelt und vom World Wide Web Consortium (W3C) als Empfehlung herausgegeben.

In kürzester Zeit entwickelte sich das WWW zum wichtigsten Internetdienst (Abb. 2.2). Von 500 WWW-Servern wuchs die Zahl bis Okt. 2001 an auf über 127,9 Mio..

2.2 URIs - Das Adressierungskonzept des WWW

Für die ständige Verfügbarkeit von Webinhalten müssen die entsprechenden Dateien auch auf einem permanent mit dem Internet vernetzten Servercomputer vorgehalten werden. Die Identifikation eines Webserver im WWW kann über die bereits in vorigen Abschnitten erwähnte IP-Adresse des Computers erfolgen. Der Nutzer kommt mit dieser etwas kryptischen Angabe jedoch nicht häufig in Kontakt. Ihm steht eine andere, etwas komfortablere Methode der Adressierung zur Verfügung: Das Domain Name System (DNS)-Konzept. Es unterteilt die Adressierungen mit Hilfe eines hierarchischen Systems und splittet deren Bestandteile in wiederum durch Punkte getrennte Einzelteile auf. Der am weitesten rechts stehende Teil entspricht dabei der größten Hierarchierungsstufe und wird als „Top-Level-Domain“ bezeichnet. Die ursprünglichen Top-Level-Domains des Internet wurden bereits im vorangegangenen Abschnitt aufgelistet. In Deutschland werden in der Regel Domainnamen mit der Top-Level-Domain „de“ verwendet. Für die „de“-Domain verwaltet die DENIC eG die eigentlichen Adressen, die „Second-Level-Domains“. Die Rechte zur Nutzung eines solchen Second-Level-Domain-Namens lassen sich bei der DENIC erwerben.

In der dritten Hierarchiestufe eines WWW-Rechnernamens wird meist der jeweils zuständige Internetdienst kenntlich gemacht. Die Bestandteile des fiktiven Rechnernamens *dienst.name.de* haben somit folgende Bedeutung:

dienst	Internetdienst, mit dem auf die Adresse zugegriffen werden kann, z.B. <i>www</i> oder <i>ftp</i> .
name	Second-Level-Domain. Frei wählbarer Namensbestandteil, der weiter untergliedert werden kann, z.B. <i>www.vorname.name.de</i> .
de	Top-Level-Domain. Neben Länderkürzeln wie <i>de</i> sind weitere Top-Level-Domains gebräuchlich, z.B. <i>com</i> , <i>net</i> , <i>org</i> , <i>edu</i> , <i>info</i> und <i>tv</i> , bzw. in Zukunft verfügbar, z.B. <i>name</i> und <i>museum</i> .

Solche Adressen innerhalb des WWW, die als Unified Resource Locator oder kurz URL bezeichnet werden, bestehen neben dem Rechnernamen aus der Protokollkennzeichnung "http://" sowie Angaben bezüglich einzelner Pfade oder Dokumente, die hinter dem Top-Level-Domain aufgeführt werden, z. B. "http://www.name.de/ast/unterast/Dokument.htm":

http://	Protokollbezeichnung.
www.name.de	WWW-Rechnername (siehe oben).
/ast/unterast/Dokument.htm	Pfad- und Dokumentenname innerhalb des Dateisystems des Webservers.

Die Protokollbezeichnung ist durchaus nicht überflüssig, wie vielfach angenommen wird. Sie wird vielmehr vom Browser erforderlichenfalls automatisch ergänzt, da es sich hierbei um das WWW-Standardprotokoll (Hypertext-Transfer-Protokoll) handelt.

Jede Datei, Dokument oder Multimediadatei, besitzt durch dieses Konzept einen eindeutigen URL, welche die Auffindbarkeit der im Web abgelegten Information gewährleistet. Auch dynamisch auf dem Server erzeugte Informationsobjekte, wie z.B. die Antwortseiten auf Datenbankabfragen, können häufig über URLs abgerufen werden. Die Besonderheiten datenbankgenerierter Webseiten werden im Abschnitt 2.5 beschrieben.

URLs stellen eine besondere Anwendung des Uniform Resource Identifier (URI) Standards [RFC2396] dar, einem allgemeinen Konzept zur eindeutigen Bezeichnung von abstrakten oder auch physikalischen Webinhalten. Während sich der Anwendungsbereich eines URL auf Netzwerkadressen beschränkt, beschreibt die allgemeiner gefasste URI-Definition den Aufbau eines eindeutigen Zeigers, ohne die konkrete Anwendung im Rahmen eines Netzwerkprotokolls oder Mediums wie das Internet zu beschränken. Eine weitere Anwendung eines URI ist der Uniform Resource Name (URN). Ein URN stellt die dauerhafte Form eines URI dar, die selbst dann bestehen bleibt, wenn sich der tatsächliche Standort eines Objekts ändern sollte. Eine ausführliche Erläuterung des Adressierungskonzepts im Internet findet sich in [Schwarte1997b].

2.3 Publizieren in HTML

Die „HyperText Markup Language“ (HTML) gehört zu den drei Kerntechnologien des WWW [HTML401Rec]. Die am CERN maßgeblich von Tim Berners-Lee entwickelte Hypertext-Plattform benötigte 1991 zur Erstellung sowie für den Austausch von Dokumenten eine einfache Editiersprache. Als Untermenge des wesentlich komplexeren und umfangreicheren SGML von 1969 wurde HTML als eine solche „Dokumentbeschreibungssprache“ entwickelt. Sie sollte den Austausch von im ASCII-Zeichensatz codierten Textdokumenten systemübergreifend möglich machen. Eine grafische Oberfläche war zunächst nicht vorgesehen. Komplexere Formatierungen und eventuell erforderliche Sonderzeichen sollten durch

entsprechende Formatierungsbefehle dargestellt werden. Zur Unterscheidung von den darstellbaren Texten sind diese Formatierungsbefehle in spitze Klammern < und > eingeschlossen, die während der Textaufbereitung durch einen Browser eine besondere Auswertung der von ihnen eingeschlossenen Zeichenfolgen erwirkt. Die von spitzen Klammern eingeschlossenen, im Webbrowser nicht dargestellten Formatierungsbefehle werden als sogenannten HTML-„Tags“ (vgl. Abschnitt 1.2) bezeichnet. Auch die gegenseitige Verknüpfung von Dokumenten über Hyperlinks geschieht durch Verwendung entsprechender Tags.

Die in spitzen Klammern eingeschlossenen HTML-Formatierungsbefehle gliedern das Dokument in einer hierarchischen Art und Weise. Hierzu werden die HTML-Tags paarweise geschachtelt, wobei den von derartigen Tag-Paaren eingeschlossenen Dokumentteilen gewisse Eigenschaften zugewiesen werden. Die Bedeutung der zahlreichen existierenden HTML-Befehle zu definieren, ist eine Aufgabe des World Wide Web Consortiums [W3C], dem Vertreter aller größeren Softwareunternehmen mit Bezug zum Internet angehören. Jede Version der entwickelten Browser verfügt über neue Möglichkeiten des Webdesigns, die in der Vergangenheit häufig erst nachträglich Eingang in die entsprechenden Standards fanden.

Die Integration von multimedialen Objekten für illustrierte HTML-Dokumente, wie wir sie heute kennen, steigerte die Attraktivität des WWW als Informationsmedium. Um multimediale Objekte wie Grafiken, Sounddateien oder auch im Browser ausführbare Programme wie z.B. Java-Applets in Webseiten integrieren zu können, werden diese über entsprechende Befehls-Tags mit Verweisen auf die Adresse der Objekte eingebunden. Der tatsächliche Speicherort kann sich dabei auf einem beliebigen Rechner im WWW befinden. Gerade multimediale Webseiten bestehen somit nicht aus einer einzigen, homogenen Datei, sondern sind vielmehr eine Verknüpfung mehrerer, auch unabhängig voneinander nutzbarer Multimediaobjekte zu einer zusammenhängenden Repräsentation im Browser (s. Abb.2.3). Tatsächlich war eines der Leitmotive der Schöpfer des WWW „To link everything with everything online“. Durch seine Multimedialität erlaubt das Web Präsentationsformen, die sich in Broschüren und anderen Druckerzeugnissen nicht in vergleichbarer Form verwirklichen ließen. Der Zugang zu den Informationen, die über ständig mit dem Internet vernetzten Webservern zur Verfügung gestellt werden, ist dabei zeit- und ortsunabhängig möglich. Auf diese Weise kann sich jeder Nutzer, unabhängig von seinem Aufenthaltsort seinen Interessen gemäß ausführlich informieren.

Das Layout zur Darstellung von Webseiten ist dabei bildschirmorientiert. Im Gegensatz zur Textverarbeitung spielen im WWW die oft national unterschiedlichen Papierformate keine Rolle bei der Gestaltung von Webseiten. Um die Erzeugung von HTML-Seiten zu erleichtern, erfolgt die Programmierung meist mit einem sogenannten HTML-Editor. Diese inzwischen sehr ausgereiften Programmierwerkzeuge unterstützen den Autor durch eine grafische Benutzeroberfläche und ausgeklügelte Funktionen zur Verwaltung komplexer Web-Sites.

HTML-Seiten enthalten jedoch nicht nur für die Darstellung am Bildschirm bestimmte Informationen. In den sogenannten Kopfdaten eines HTML-Dokuments lassen sich unter anderem Informationen über den Titel einbinden sowie „Meta“-Informationen über den Autor, Inhalt, Keywords etc.. Von den Meta-Informationen wird in Kapitel 4 später in der Arbeit noch ausführlicher die Rede sein. Diese Angaben dienen hauptsächlich einer rudimentären Kategorisierung in den Suchmaschinen. Detailliertere Angaben über die Inhalte sind über diese Standardelemente von HTML nur beschränkt möglich.

Die Navigation zwischen den einzelnen Seiten des WWW erfolgt über Hyperlinks. Einzelne Textabschnitte oder Grafiken werden hierfür mit einer Adresse verknüpft und grafisch hervorgehoben. Sie erlauben dem Benutzer durch simple Mausklicks zwischen den miteinander verlinkten Dokumenten hin- und herzuschalten. Diese Funktionalität, die charakteristisch für das WWW ist, wird als „Web-Surfen“ bezeichnet.



Abb. 2.3 Aufbau einer HTML-Seite: Die Seite beginnt mit einem „Head“-Abschnitt mit Titel, Metadaten etc.. Der Seiteninhalt nach `<Body . . .>` enthält den Text sowie Formatierungsbefehle und URLs zu eingebundenen multimedialen Inhalten.

Die freie Vernetzbarkeit der Dokumente erlaubt, im Gegensatz zu klassischen linearen Informationsmedien wie dem Buch, zusätzlich eine nichtlineare Strukturierung der Informationen (Abb. 2.4). Solche Dokumentstrukturen, zu denen auch das WWW zählt, werden auch als „Hypertexte“ bezeichnet. Der multimediale Aspekt des World Wide Web ist in dem Begriff der Hypermedia enthalten. Hypermedia erweitert die Idee des Hypertextes über den bloßen Text hinaus (Abb. 2.5). Hypermedia bezieht die unterschiedlichen Formen von Medienelementen in HTML-Dokumenten mit ein.

Da zur Nutzung der Inhalte von Hypertexten eine Strukturierung der einzelnen Webseiten zueinander sinnvoll ist, erfolgt mit der Verknüpfung häufig gleichzeitig eine baumartige Hierarchierung zu einem System übergeordneter und vertiefender Seiten. Eine solche strukturierte Sammlung miteinander verknüpfter Webseiten, die sich alle hierarchisch auf eine einzelne Leitseite, die sogenannte „Homepage“ zurückführen lassen, wird als „Web-Site“ bezeichnet. Diese Homepage bekommt der Benutzer zumeist automatisch als Startseite angezeigt, wenn er den entsprechenden Domainnamen im Browser anwählt.

Von besonderer Bedeutung bei der Gestaltung von interaktiven Systemen im WWW sind die HTML-Formular-Elemente. Mit Hilfe dieser Elemente ist es möglich, gezielt

Benutzereingaben anzufordern und zur Weiterverarbeitung an ein entsprechendes Programm weiterzugeben. Die Gestaltung dieser Elemente orientierte sich an den Bedienungsobjekten moderner grafischer Betriebssystemoberflächen, mit denen die Anwender in aller Regel bereits hinlänglich vertraut sind.

Die Formelarauswertung geschieht entweder durch clientseitige Prozesse, die auf der Anwenderseite im Webbrowser ausgeführt werden, oder auf dem Webserver [Reinhardt2000]. Zur Programmierung dieser Prozesse stellt HTML den Webautoren keine Funktionen zur Verfügung. Wie eingangs erwähnt, handelt es sich bei HTML um eine reine Seitenformatierungssprache und nicht um eine Prozess-Sprache. Für das World Wide Web entstanden aus diesem Grund eine Reihe weiterer Programmiersprachen, die sich grob in clientseitige Sprachen (z.B. Javascript [Koch1999], Java [Sun2002]) einteilen lassen, die in den Webbrowsern auf den Computern der Benutzer interpretiert werden, und serverseitige Programmiersprachen (z.B. Active Server Pages (ASP) [Johnson1997], PHP [Krause2000], JavaServer Pages (JSP) [SunJSP]), die auf dem Servercomputer zentral ausgeführt werden. Bei den zur Programmierung des Webserver verwendeten Sprachen handelt es sich heute überwiegend um Skriptsprachen. Dabei kann der Programmcode in den HTML-Code der Webseiten integriert werden. Bei Aufruf einer solchen Seite wird vor dem Versenden dieser Programmcode vom Webserver ausgeführt. Der Anwender erhält dann ein HTML-Dokument. Ihm bleibt der auf dem Server ausgeführte Code verborgen.

Die Entscheidung, welche Abläufe client- bzw. serverseitig ausgeführt werden sollen, ist abhängig von den programmierten Funktionen und muss fallweise getroffen werden [Reinhardt2000]. Dabei haben clientseitige Prozesse den Nachteil, dass die Programm-Module zunächst vom Server über das Internet übertragen werden müssen, was bei umfangreicheren Programmen zu sehr langen Zugriffszeiten führen kann. Auch sind trotz aller Standardisierungsbemühungen die Ergebnisse clientseitiger Programme in unterschiedlichen Webbrowserprogrammen nicht identisch, so dass evtl. mehrere an die jeweiligen Besonderheiten der Browser angepasste Programmversionen auf dem Server vorgehalten werden müssen.

Serverseitige Lösungen besitzen jedoch den Nachteil, dass bei zunehmender Nutzung auch die Belastung des Servercomputers sehr stark ansteigt. Häufig ist es deshalb sinnvoll, beide Möglichkeiten miteinander zu kombinieren und beispielsweise alle, die Darstellung der Inhalte betreffenden, Programmkomponenten clientseitig für den Browser des Benutzers zu programmieren, während auf der Serverseite vor allem Datenbankalgorithmen ausgeführt werden, die eine Informationsauswahl vornehmen, bevor die Daten zum Nutzer gesendet werden.

2.4 Die „Hypothek“ von HTML

Trotz seines großen Erfolgs als Publikationsstandard und der bis dato unerreichten Verkürzung der Informationswege zeigten sich im WWW schon früh an ersten Unzulänglichkeiten die Grenzen von HTML. Ein grundlegendes Problem stellte hierbei die mangelnde Erweiterbarkeit von HTML dar. Durch die zentrale Verabschiedung der jeweils neuen Versionen durch das W3C und die Notwendigkeit, diese Erweiterungen in neue Browsergenerationen zu implementieren, ist die Flexibilität von HTML bezüglich der Einführung neuer Tags sehr eingeschränkt.

Für Bauingenieure besonders schmerzlich ist dabei der eingeschränkte Zeichensatz, der sich mit HTML darstellen lässt. HTML erlaubt damit keine zufriedenstellende Darstellung von mathematischen und chemischen Formeln. Die Publikation entsprechender Gleichungen erfolgt zumeist in Form von Grafiken, die in den Text geladen werden. Diese Umwandlung bedeutet aber zusätzliche Arbeitsschritte für den Autor von Webdokumenten und führt in der

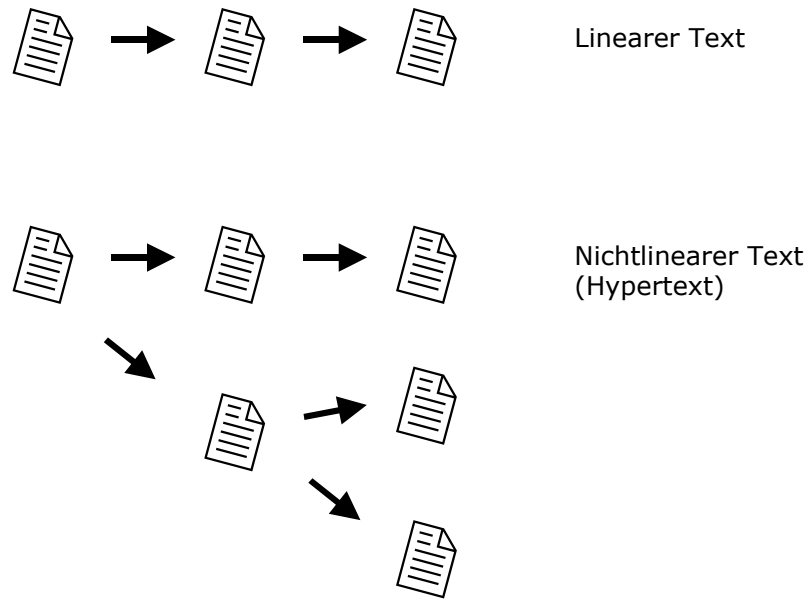


Abb. 2.4 Linearer Text und nichtlinearer Hypertext

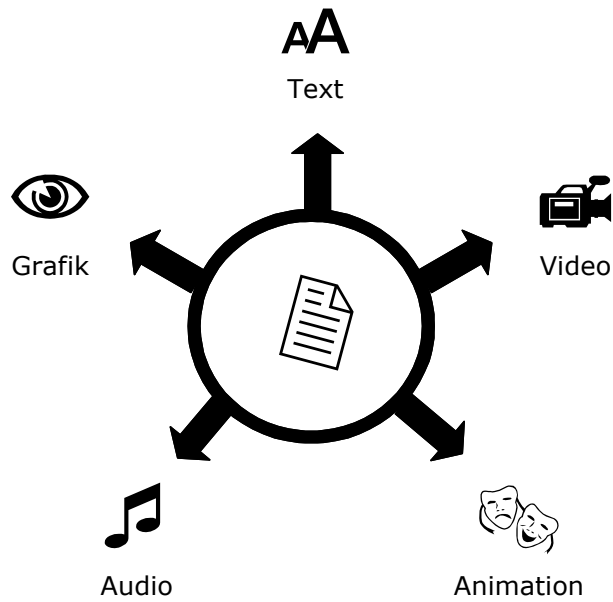


Abb. 2.5 Hypermedia kombiniert die Konzepte Multimedia und Hypertext

Regel zu einer inhomogenen Darstellung. Vor allem jedoch können Programme, wie z.B. Suchmaschinen, diese Grafiken nicht als Gleichungen erkennen und entsprechend interpretieren. Die Bedeutung der Gleichungselemente bleibt maschinellen Recherchen daher verborgen. Für den Ingenieurbereich stellt dies einen wesentlichen Mangel bezüglich des Nutzwerts derartig erstellter Webinhalte dar.

Die fehlende Strukturierung der Inhalte von Webseiten in HTML beschränkt sich dabei nicht auf Gleichungen. Ein Autor besitzt lediglich über die sogenannten Meta-Tags die Möglichkeit, Aussagen über den Inhalt des Dokuments in Form von Stichwortlisten, Kurzbeschreibungen oder diversen Dokumentattributen, wie z.B. den Autor und das Erstellungsdatum der Seite zu machen. Eine für Suchmaschinen erkennbare Gliederung des Dokuments sowie eine feinere Beschreibung der Inhalte sind in HTML hingegen nicht möglich. Weiterhin kennt HTML nur den Datentyp „text“, so dass die eingangs erwähnten Recherchemöglichkeiten mit vergleichenden Operatoren (z.B. „<“, „>“, „=“ etc.) nicht realisiert werden können.

In einem explosionsartig wachsenden WWW verhindert das Fehlen von semantischen Strukturelementen jedoch eine detailliertere und gezieltere Erschließung der Online publizierten Informationen. Automatische Agentenprogramme könnten bspw. gezielt Informationen aus unterschiedlichen Webangeboten zusammenstellen, die wiederum von anderen Programmen weiterverarbeitet und ausgewertet werden. In einem HTML-basierten Web wäre eine solche „maschinenlesbare“ Web-Site nicht realisierbar, da hier die einzelnen Inhalte einer Seite von einer Maschine nicht ausreichend genau unterschieden und identifiziert werden könnten.

Auch für die Publikation von Online-Inhalten wären Funktionalitäten wünschenswert, die HTML den Autoren nicht bieten kann. So ist es nicht möglich, die Benutzerrechte von HTML-Dokumenten über das Leseverbot hinaus einzuschränken. Weitere Stufen wie das Einschränken der Druckrechte oder die Möglichkeit, das Kopieren von Textabschnitten aus dem Browser in die Textverarbeitung zu verhindern, bietet HTML ebenfalls nicht. Trotz der Erweiterung von HTML im Zuge der Version 4.01 um dynamische Vektorgrafik zu Dynamic HTML (DHTML) blieben die grafischen Funktionalitäten von DHTML hinter den Ansprüchen der Webdesigner zurück. Diese Unzulänglichkeiten wurden bislang von kommerziell geprägten, proprietären Softwarekomponenten abgedeckt.

Für möglichst direkte Verweise und damit kurze Informationswege wäre es zudem sinnvoll, Hyperlinks zu bestimmten Elementen von Webseiten gezielt setzen zu können. Bei sehr umfangreichen Dokumenten würde dies eine erneute lokale Suche auf der Webseite vermeiden. Dies ist zwar bislang schon über das Setzen von entsprechenden Markierungen im HTML-Quellcode der Dokumente möglich, jedoch erfordert das Modifizieren des Quellcodes den Schreibzugriff auf die Datei. Da es sicher nicht sinnvoll wäre, den Schreibzugriff auf Webdokumente allgemein zu gestatten, bleibt die Möglichkeit, Sprungziele in HTML-Dokumenten zu setzen, bislang das Privileg des Autors oder zumindest auf sehr wenige Personen beschränkt.

Sind die Inhalte erst einmal im HTML-Format publiziert, ist deren Weiterverarbeitung in anderen Formatierungen nur unter großem Aufwand möglich. HTML optimiert die Darstellung für Präsentationen im Browser mit dem Bildschirm als Ausgabegerät. Mit der alltäglichen Nutzung webbasierter Informationsdienste bekamen auch Mobilgeräte wie Mobiltelefone und PDAs etc. Zugang zum Netz, also Geräte, die schon aufgrund ihrer Größe andere Darstellungsmöglichkeiten als PC-Bildschirme besitzen. Nicht nur die Aufbereitung von Inhalten für andersartige Ausgabegeräte, sondern bereits das Exportieren aus einer Web-Site in eine andere ist eine meist zeitraubende, häufig manuell geprägte Tätigkeit. Automatismen zur Umformatierung können auch hier die Strukturelemente eines Dokuments

nicht erkennen, um diesem eine neue Formatierung zuzuweisen. Bereits zur Pflege bestehender Inhalte werden auch Kenntnisse über HTML-Befehle benötigt, sofern die Web-Site nicht über ein Content-Management-System (CMS) [Goldfarb1999] erstellt wird.

Die Ursache für diese schlechte Verwaltbarkeit und Recherchierbarkeit von Informationen liegt in der fehlenden Trennung von Inhalt und Formatierung in HTML. HTML-Dokumente enthalten bezüglich ihrer Inhalte lediglich Formatierungsbefehle, welche die Darstellung für den menschlichen Leser optimieren. Computer benötigen jedoch strukturierende Elemente, welche die maschinelle Verarbeitung von Informationen ermöglicht. Diese stellt der HTML-Standard nicht zur Verfügung. Eine Lösung für diese vielfältigen Problemstellungen müsste daher eine große Flexibilität gegenüber zukünftigen Erweiterungen aufweisen und gleichzeitig zu bestehenden Standards kompatibel sein.

2.5 Datenbankgestützte Web-Sites

Um eine bessere Verwaltbarkeit der Inhalte einer Web-Site realisieren zu können, müssen diese zunächst von den formatierenden HTML-Tags getrennt gespeichert werden. In datenbankgestützten Web-Sites erfolgt diese Speicherung in einem Datenbanksystem. Die Formatierung der Texte und Grafiken in HTML-Dokumente erfolgt durch Zusammenfügen der in der Datenbank gespeicherten Datensätze mit „leeren“ HTML Formatvorlagen auf dem Webserver.

Man unterscheidet am Zeitpunkt der Generierung zwei Varianten maschinell erzeugter HTML-Seiten: Bei der ersten Möglichkeit wird die Erzeugung einer Seite durch die Abfrage des Nutzers ausgelöst [Schwarte1997a]. Eine solche Webseite wird auch als dynamischer Hypertext bezeichnet. Das Baustoff-Informationssystem BiM-Online [Reinhardt2000, Reinhardt1999a], welches den Ausgangspunkt dieser Arbeit bildet, wurde zum großen Teil aus solchen „dynamischen“ Webseiten aufgebaut. Die Kombination von Datenbanken und serverseitig ausgeführten Programmkomponenten erlaubt die Erstellung von stark interaktiven Webangeboten mit intensivem Dialog zwischen Anwender und System (Dialogsysteme) [Schreyer1999a]. Ein ebenfalls sehr sinnvoller Anwendungsfall sind Informationsangebote mit sich schnell verändernden Daten, wie z.B. Aktienkursen. Der Prozess der Generierung einer Seite erzeugt jedoch eine vergleichsweise hohe Prozessorlast auf dem Datenbank-Servercomputer. Bei hohen Zugriffszahlen kann dies auch leistungsfähige Computer an ihre Grenzen bringen.

Eine Alternative ist das sogenannte „staged serving“. Dabei werden die HTML-Dokumente nur nach Veränderung der Inhalte aus der Datenbank neu erzeugt. Die Eingabe der Inhalte erfolgt über Formulare, deren Inhalte dann in vorgegebenen Dokumentvorlagen formatiert und als statische HTML-Seiten gespeichert werden. Die hohe Prozessorlast für die Generierung des HTML-Codes erfolgt somit seltener. Durch die Verwendung von Formatvorlagen lässt sich auch ein einheitliches Aussehen der auf diese Weise generierten HTML-Seiten innerhalb einer Web-Site gewährleisten.

Eine Mischung zwischen diesen beiden Varianten wird mit „serversided caching“ bezeichnet. Dabei legt der Server dynamisch erzeugte Webseiten in einem Zwischenspeicher, dem Cache, ab und hält sie dort bis zur nächsten Änderung vor. Für weitere Zugriffe erzeugt der Server nur dann die HTML-Seiten dynamisch, wenn sich diese noch nicht im Cache des Servers befinden. Auf diese Weise lässt sich ebenfalls die Hardware entlasten.

Die getrennte Speicherung von Webseiteninhalten in der Datenbank und der Formatierung in entsprechenden Vorlagen erlaubt es, einmal eingegebene Inhalte mit relativ geringem Aufwand für unterschiedliche Ausgabeformate vorzuhalten. Mit entsprechend angepassten Dokumentvorlagen für Mobilgeräte (z.B. Wireless Application Protokoll (WAP)) oder

Druckversionen können somit auf denselben Inhalten unterschiedliche Ausgabemedien erstellt werden (vgl. Content-Management-Systeme (CMS) [Goldfarb1999]).

Weiterhin erlaubt die Speicherung der Inhalte in einem Datenbanksystem die Zuweisung von numerischen Datentypen, die zusätzlich zu textbasierten Recherchen auch die Möglichkeit vergleichender Datenabfragen auf die lokalen Inhalte gestatten. Diese zusätzlichen Funktionalitäten datenbankgestützter Web-Sites erfordern jedoch auf der Serverseite eine entsprechend aufwendigere Softwarearchitektur. Eine typische Softwarearchitektur für eine datenbankgestützte Web-Site, wie sie in BiM-Online [Reinhardt2000] verwendet wurde, ist schematisch in Abb. 2.6 dargestellt.

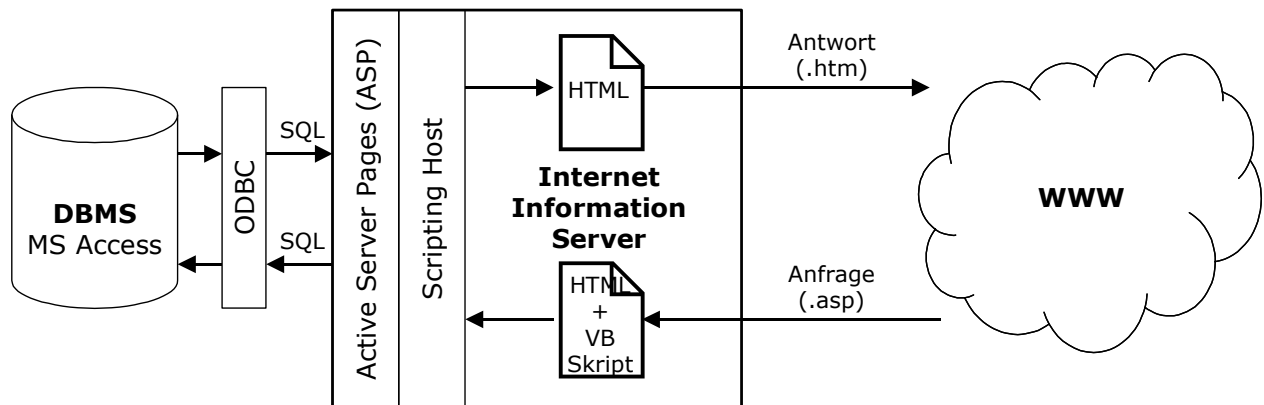


Abb. 2.6 Softwarearchitektur der datenbankgestützten Web-Site BiM-Online

Bei den Datenbankmanagementsystemen handelt es sich zumeist um Systeme, die auf dem relationalen Datenmodell von T. Codd [Codd1970] basieren. Die Daten werden in diesem heute sehr verbreiteten Datenmodell in Tabellen verwaltet, die wiederum miteinander über logische Verknüpfungen, den Relationen, in Beziehung gesetzt werden können. Für die Abfrage der Daten hat sich aufgrund der hohen Verbreitung dieses Datenbanktyps die Standard Query Language (SQL) [SQL1999, Schicker2000] als Abfragesprache durchgesetzt. In jüngster Zeit werden auch häufig Datenbanksysteme eingesetzt, die in ihrem Speichermodell dem objektorientierten Paradigma folgen. Diese Systeme werden als objektorientierte (OO) Datenbanken [Heuer1997] bezeichnet und eignen sich besonders für die Verwaltung multimedialer Informationen. Die Segmentierung der oft komplexen und stark vernetzten Multimediadaten in mehrere relationale Tabellen kann in OO-Datenbanken vermieden werden. [Tewes1996].

Die Datenbankschnittstellen sitzen zwischen dem Webserver und dem Datenbanksystem. Sie übermitteln die Anfragen der Benutzer an die Datenbank und steuern das Auslesen der Daten. Auch bei den Datenbankschnittstellen haben sich Standards herausgebildet. Das Vorschalten einer standardisierten Schnittstelle besitzt den Vorteil, dass bei Bedarf das Datenbanksystem ausgetauscht werden kann, ohne dass man die Programmlogik hierfür modifizieren muss. Die beiden wichtigsten Schnittstellenstandards sind der Open DataBase Connectivity (ODBC) von Microsoft und dessen Java-Variante Java Database Connectivity (JDBC). Entsprechende Softwaremodule zur Anbindung der jeweiligen Datenbanktreiber sind in die Betriebssysteme integriert (Microsofts ODBC) beziehungsweise im Web als Open Source Software [OSI2001] frei verfügbar (JDBC). Eine dritte Möglichkeit, von der vor allem im Bereich der Open Source Software Gebrauch gemacht wird, ist die direkte Integration von Datenbankschnittstellen in die Skriptsprachen des Webserver. Die Skriptsprachen schließlich steuern

die Prozesse zwischen der WWW-Oberfläche der Web-Site, die am Webserver generiert wird, und der Datenbank.

Bei hochbeanspruchten Web-Sites lässt sich die in Abb. 2.6 abgebildete Aufteilung zwischen Datenbanksystem und Webserver auch für die Programmlogik weiterbetreiben. Die Programmlogik wird aufgrund ihrer Stellung zwischen Datenbanksystem und Webserversoftware häufig als sogenannte Middleware bezeichnet. Auf diese Weise entstehen schon auf Seiten der Serversysteme vernetzte Softwarearchitekturen, deren Prozesse durch entsprechende Programmiersprachen serverseitig miteinander verknüpft werden.

Trotz der vielen Vorteile für die Funktionalität von Webangeboten, welche die Verwendung von Datenbanken mit sich bringt, werden dadurch wesentliche Unzulänglichkeiten bzgl. der Recherche und Weiterverwendung dieser Inhalte nicht behoben. Die innerhalb der Datenbanken in relationalen Datenstrukturen abgespeicherten Informationen bleiben für allgemeine Suchdienste, die von außerhalb der Web-Site auf das System zugreifen wollen, unzugänglich. Die formularartigen Abfragemasken solcher Systeme können von Suchmaschinen bislang nicht erkannt und der Bedeutung der einzelnen Eingabefelder entsprechend bedient werden. Automatischen Suchrobotern dieser Dienste, auf die im übernächsten Abschnitt ausführlicher eingegangen wird, ist somit die vollständige Erschließung von in Datenbanken gespeicherten Inhalten heute noch nicht möglich.

2.6 XML – die Grundlage zukünftiger Publikationsstandards

In den vorhergehenden Kapiteln wurde wiederholt darauf hingewiesen, dass die mangelnde inhaltsbezogene Strukturierung von HTML-Dokumenten das Kernproblem für die Entwicklung detaillierterer Rechercheinstrumente darstellt. Einen hinsichtlich einer strukturierten Abbildung von Informationen im WWW verbesserten Ansatz verfolgt die Extensible Markup Language (XML).

2.6.1 Grundprinzipien von XML

Bei XML handelt es sich nicht um eine weitere Programmiersprache, sondern um ein erweiterbares („extensible“) Konzept, welches die Schaffung beliebiger Sprachen zur Beschreibung von Daten innerhalb von WWW-Dokumenten erlaubt. XML wird deshalb auch als Meta-Sprache bezeichnet [Behme2000]. Im Vergleich zu HTML erlaubt das XML Konzept dem Benutzer, eigene Steuerbefehle bzw. Tags zu definieren. Dokumente können damit nicht nur durch ihre Formatierung, sondern auch inhaltlich strukturiert werden.

Vom Standpunkt seiner historischen Entwicklung her betrachtet, handelt es sich bei XML ursprünglich um ein Dokumentenformat. XML gilt heute als die wichtigste Sprache für den Austausch, die Verwaltung und die Repräsentation von strukturierten Informationen. Sie wird derzeit von praktisch allen maßgeblichen Softwareentwicklern im Bereich des Internets als die zentrale Zukunftstechnologie angesehen. Die XML zugrundeliegende Idee ist dabei nicht neu, sondern basiert auf einem inzwischen über 35 jährigen Konzept des „generic coding“. W. Tunnicliffe machte 1967 den Vorschlag, den Informationsgehalt eines Dokuments von seiner äußeren Form zu trennen [vgl. Behme2000]. In diesem Zeitraum veröffentlichte S. Rice seine Idee der „editorial structure tags“ [vgl. Behme2000], woraus später das „generic markup“ entstand. Der Begriff des „markup“ stammt aus dem Verlagswesen und aus einer Zeit, in der die elektronische Dokumentenverarbeitung noch unbekannt war. Auf die inhaltliche Überprüfung und Korrektur eines Manuskripts folgte die Bearbeitung durch den Layouter, der die Entscheidungen über die typografischen Festlegungen traf. Diese wurden in Form von handschriftlichen Markierungen und Anweisungen (Markup) in das Manuskript eingefügt und anschließend im Satz berücksichtigt.

Ebenfalls auf den Ideen von Tunnicliffe und Rice basierend, entwickelten 1969 C. Goldfarb, E. Mosher und R. Lorie bei IBM die Auszeichnungssprache Generalized Markup Language (GML). GML [vgl. Behme2000] enthielt erstmals das Konzept eines formal definierten Dokumenttyps mit einer hierarchischen Struktur. Die Entwicklung zu einem Standard begann 1978, als das American National Standard Institute (ANSI) ein Komitee mit dem Ziel der Entwicklung einer standardisierten Textbeschreibungssprache auf der GML-Basis gründete. Nach mehreren Normentwürfen wurde die International Organization for Standardization (ISO) in das Projekt mit einbezogen. Als Ergebnis der Arbeit wurde 1986 die Standard Generalized Markup Language (SGML) als internationaler ISO-Standard 8879 veröffentlicht.

Ende der 80er Jahre entschied sich T. Berners-Lee dafür, als Auszeichnungssprache für Dokumente des von ihm als World Wide Web bezeichneten Dokumentensystems eine einfache, auf SGML basierende Sprache zu benutzen. Diese erhielt später den Namen Hypertext Markup Language (HTML). In den darauf folgenden Jahren entstanden u.a. mit den HTML-Formularen und den Cascading Style Sheets (CSS) zahlreiche Erweiterungen dieses Standards. Der erste XML-Entwurf wurde im November 1996 im Rahmen der SGML'96-Konferenz in Boston vorgestellt. Die Sprache wird von einer Arbeitsgruppe des W3C betreut, dessen Leitung J. Bosak (Sun Microsystems) inne hat. Bei XML handelt es sich um eine Teilmenge von SGML [vgl. Clark1997.], dessen Standard im Ruf steht, sehr komplex und für viele Anwendungszwecke zu unhandlich zu sein. Die komplexe Syntax von SGML enthält zum großen Teil nur selten verwendete Eigenschaften. Genau an dieser Stelle setzte man beim Entwurf von XML an: Die Sprache führt das Konzept des strukturierenden Markups fort, verzichtet dabei aber auf unnötigen Ballast für die Anwendung in Hypertext-Dokumenten.

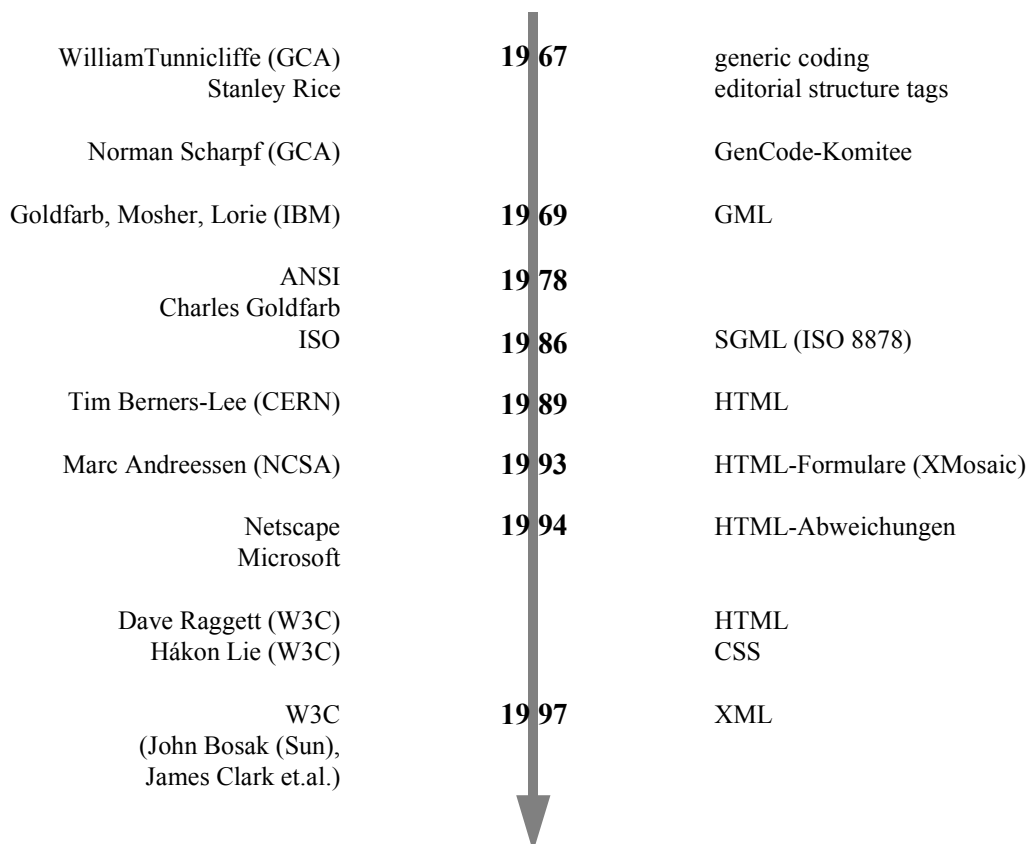


Abb. 2.7 Historische Entwicklung der Textbeschreibungssprachen bis XML [Behme2000]

In einem XML-Dokument werden Inhalte, z.B. der Text, die Bilder usw., zunächst ohne Formatierungsangaben gespeichert. Die einzelnen Inhalte werden dabei in eine hierarchisch durch Tags gegliederte Elementstruktur eingeordnet. Über aussagefähige Elementnamen und aus der Stellung der einzelnen Elemente zueinander heraus lässt sich damit die logische Gliederung eines Dokuments in einer Objekthierarchie aus „abstrakten Datentypen“ [Behme2002] abbilden.

Infolge der Dominanz des WYSIWYG-Prinzips („What you see is what you get“) moderner Textverarbeitungssoftware ist diese Strukturkomponente zumeist weniger offensichtlich. In Textverarbeitungsprogrammen begegnet sie dem Anwender heute lediglich in Form von Dokumentvorlagen für typische Dokumentarten bzw. -inhalte. Ein Zeitschriftenartikel setzt sich z.B. aus Kapiteln, Absätzen, Verweisen auf weitere Quellen, Bildunterschriften usw. zusammen. All diese logischen Informationen, die der Leser auf den ersten Blick aus der Formatierung heraus erkennt, müssen einem Computerprogramm über entsprechende Angaben erst deutlich gemacht werden. Die für das visuelle Erscheinungsbild maßgebenden Elemente der Formatierung wie Layout, Schriftart usw. sowie die Inhalte selbst komplettieren schließlich die logische Aufteilung eines Dokuments (Abb. 2.8).

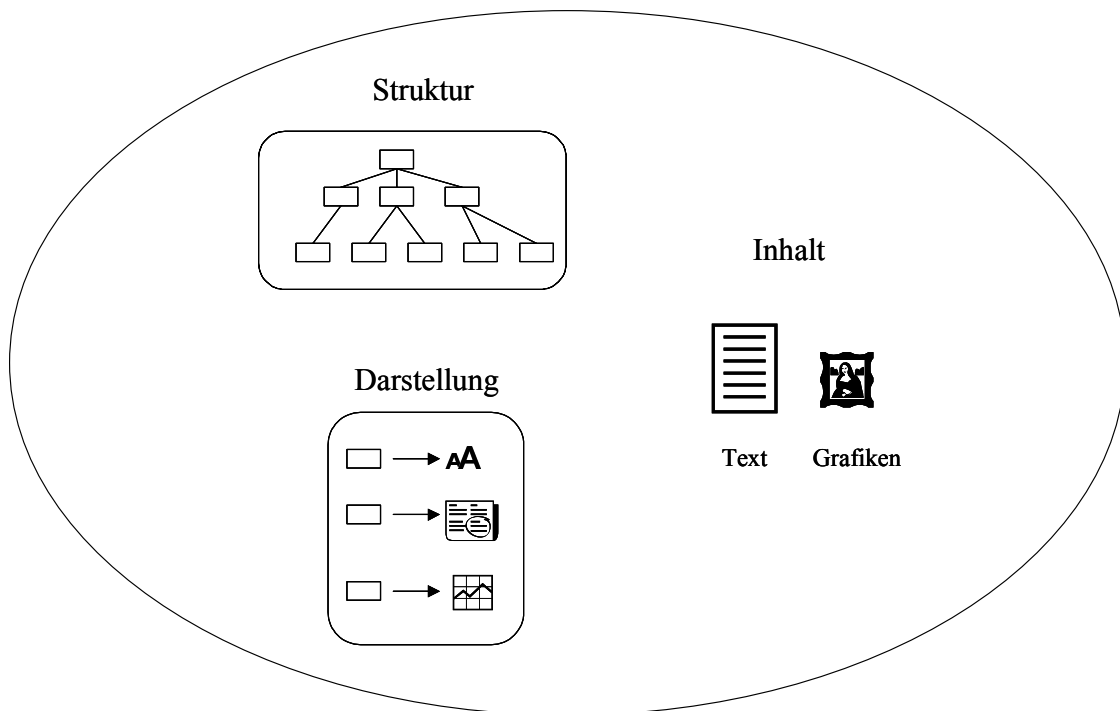


Abb. 2.8 XML trennt Dokumente in drei Komponenten [Behme2002]

Die Möglichkeit, eigene Tags zu definieren erlaubt es, dass die Elementstrukturen in XML-Dokumenten frei erweitert oder entwickelt werden können. Der strukturelle Aufbau sowie die logischen Elemente einer bestimmten Klasse von Dokumenten (Dokumenttyp) wird in einer Dokumenttyp-Definition (DTD) oder einem XML-Schema-Dokument beschrieben. Eine DTD definiert, welche Elemente innerhalb eines XML-Dokuments wie häufig und an welcher Stelle innerhalb des Elementbaums vorkommen dürfen bzw. müssen. Während DTDs ursprünglich auch für SGML-Dokumente benutzt wurden und in ihrer Syntax nicht der XML-Syntax entsprechen, handelt es sich bei XML-Schemata um Dokumentdefinitionen, die selbst in XML-Syntax verfasst sind. Sie erlauben über die Funktionalität von DTDs hinaus auch Festlegungen wie die, welches Datenformat z.B. die Daten zwischen den Tags besitzen

dürfen. DTDs und XML-Schema-Definitionen legen damit den prinzipiellen Aufbau der von ihnen bestimmten Dokumentarten fest.

An einem Beispiel wird die Aufgabe einer DTD deutlich. Eine Email besteht z.B. aus einem Empfänger, einem Absender, einem Thema sowie einer Nachricht, und zwar in dieser Reihenfolge. Alle diese Elemente enthalten ausschließlich reinen Text, der im folgenden Beispiel Parsed Character Data (PCDATA) genannt wird.

```
<!-- E-Mail-DTD -->
<!ELEMENT email      (empfaenger, absender, thema, nachricht)>
<!ELEMENT empfaenger (#PCDATA)>
<!ELEMENT absender  (#PCDATA)>
<!ELEMENT thema     (#PCDATA)>
<!ELEMENT nachricht (#PCDATA)>
```

Eine entsprechende Email sieht dann als XML-Dokument wie folgt aus:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE email SYSTEM "email.dtd">
<email>
  <empfaenger>Joachim Schwarte</empfaenger>
  <absender>Marcus Schreyer</absender>
  <thema>Kaffeepause?</thema>
  <nachricht>
    Hallo Joachim,

    wie wär's mit einer kleinen Kaffeepause? Wir könnten ja bei der
    Gelegenheit gleich das Konzept für die Präsentation am Montag
    durchgehen...

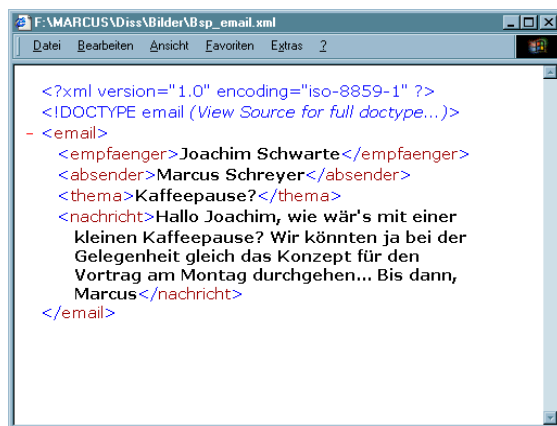
    Bis dann, Marcus
  </nachricht>
</email>
```

Das obere XML-Dokument ist auch ohne DTD eigentlich recht gut lesbar. Tatsächlich erlaubt XML auch, Dokumente ohne DTD oder ein XML-Schema zu schreiben. Welchen Vorteil hat also in diesem Fall die Verwendung der DTD?

Sobald man eine Reihe von gleichartigen Texten schreibt, beispielsweise mehrere Emails, dient eine DTD als Vorlage oder Muster. Ein Editorprogramm zur Erstellung von Emails kann dieses Muster schon vor dem Schreiben auswerten und einem Autor entsprechende Eingabefelder für die Inhalte anbieten oder beim Schreiben die DTD mit dem aktuellen Dokument vergleichen und auf strukturelle Fehler hinweisen. Wenn ein XML-Dokument von der Struktur her mit den Festlegungen in der zugeordneten DTD übereinstimmt, wird es als „gültig“ bezeichnet. Der Vorgang der strukturellen Überprüfung eines XML-Dokuments wird aus dem Englischen abgeleitet als Validierung bezeichnet. Ein anderer wichtiger Begriff im Zusammenhang mit der Erstellung von XML-Dokumenten ist die „Wohlgeformtheit“. Es ist eine zentrale Anforderung an alle XML-Dokumente, dass sie wohlgeformt sein müssen, d.h. sie müssen sich exakt an die Vorgaben der XML-Spezifikation halten. HTML ist bzgl. des Umgangs mit Leerzeichen und der Groß-/Kleinschreibung von Tags weit weniger restriktiv.

Mit der Auszeichnung von Inhalten allein lässt sich noch keine für den Endanwender brauchbare Darstellung erzeugen. Das oben dargestellte Beispiel eines Email-Dokuments wird in einem aktuellen XML-fähigen Browser als Baumstruktur samt Inhalten angezeigt (Abb. 2.9). Um die Inhalte in eine vernünftige Darstellung zu bringen, existiert eine weitere, auf XML-basierende Sprache: Die Extensible Stylesheet Language (XSL).

a



b

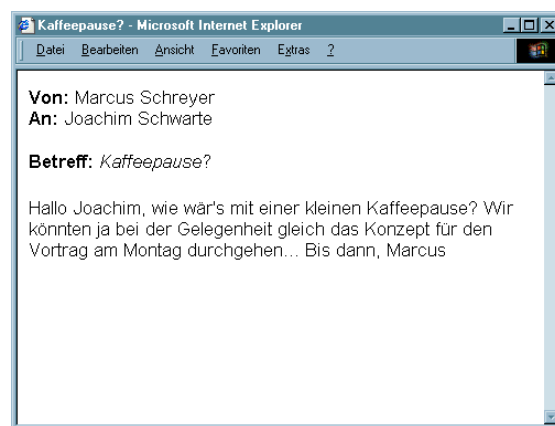


Abb. 2.9 Anzeige des XML-Dokuments ohne (a) und mit (b) einer XSL-Formatvorlage (Style Sheet) zur Formatierung der Darstellung in einem Webbrowser

Über unterschiedliche XSL-Vorlagen lässt sich derselbe Inhalt jeweils in unterschiedlichen Formatierungen darstellen. Diese Trennung besitzt besonders bei der schon angesprochenen Darstellung identischer Inhalte auf unterschiedlichen Ausgabegeräten einen großen Nutzwert. Das in Abb. 2.9 dargestellte Email-Dokument könnte durch entsprechende XSL-Dateien sowohl zur Darstellung auf einem Mobiltelefon als auch in einem Emailprogramm automatisch aufbereitet werden.

Seit dem Erscheinen von XML sind viele Anwendungen entstanden, die XML nicht als Dokumentenformat, sondern allgemeiner als Datenformat bzw. Datenaustauschformat ansehen. Derartige Datendokumente sind nicht für menschliche Augen, sondern zur automatischen Verarbeitung durch Programme bestimmt. Bis Anfang 1999 wurde XSL nicht nur als Sprache benutzt, welche die Formatierung von XML-Dateien für die Darstellung auf Ausgabemedien übernehmen sollte, sondern auch zur Transformierung von XML-Dokumenten von einer XML-Datenstruktur in eine andere. Im selben Jahr legte die entsprechende Arbeitsgruppe des W3C dann zwei verschiedene Entwürfe vor: Einen, der die Formatierung betraf [XSL2000], sowie einen, der die möglichen Transformationen [XSLT2001] beschrieb.

Bei XSLT handelt es sich um einen Bestandteil der Stilsprache XSL, der für die Wandlung von XML-Dokumenten der einen Art in eine andere vorgesehen ist. Die Wandlung erfolgt dabei anhand von Schablonen („template rules“). Sie regeln, welche Elemente aus der Eingabestruktur in die Ausgabe gelangen, in welcher Form sie das tun oder ob sie vielleicht gar nicht im resultierenden Baum vorkommen. Sind die Daten einmal in das XML-Format überführt, lassen sich mit dieser Methode relativ einfach Schnittstellen zu anderen Programmen erstellen.

Diese Flexibilität des XML Konzepts führt dazu, dass im IT-Bereich vielen Branchen Standards für den Datenaustausch, wie Standard for the Exchange of Product Model Data (STEP) oder der Electronic Data Interchange (EDI) in universelleren XML-Varianten neu definiert und abgelöst werden. Gerade der EDI-Standard, der vor allem von großen Unternehmen mit Milliardeninvestitionen zur Verbesserung des Datenaustauschs zwischen Firmen und Geschäftspartnern entwickelt wurde, hat sich in seiner ursprünglichen Version als sehr teuer und schwierig zu implementieren erwiesen. Die benötigten Transformationsprogramme mussten an die verwendete Plattform und den jeweiligen Geschäftspartner angepasst werden, enthielten strenge Informationen über den Umgang mit den Daten und waren deshalb unflexibel und wenig zukunftssicher [Goldfarb1999]. Mit XML als

universellem Datenformat können die Datenstrukturen von den Geschäftsregeln, d.h. den Regeln zum Umgang mit den Daten, getrennt werden. Folglich genügt es, im wesentlichen die neue Zuordnung der Datensätze zu definieren, so dass jeder Geschäftspartner seine eigenen Geschäftsregeln anwenden kann. Der Entwicklungsaufwand für Schnittstellen wird hierdurch um ein Vielfaches vereinfacht. Wie im Abschnitt 2.6.3 dargestellt wird, spielt XML als Datenformat auch im Bauwesen inzwischen eine zentrale Rolle in der aktuellen Diskussion über den Datenaustausch.

2.6.2 XML im World Wide Web

Genau wie HTML basiert XML auf der Standard Generalized Markup Language (SGML). Zu HTML gibt es jedoch bezüglich der Publikation von Informationen im WWW einige zentrale Unterschiede.

Im Gegensatz zur Erstellung von HTML-Dokumenten kann in XML die logische Struktur der Inhalte abgebildet werden. Da sich in XML-Dokumenten eigene Tags definieren lassen, können bestimmte Abschnitte eines Dokuments, wie z.B. Angaben zu auf der Seite abgebildeten Produkten und Preisen, auch mit entsprechenden Tags `<Preis>19.99</Preis>` `<Währung>EUR</Währung>` ausgezeichnet werden. Aufgrund der dezentralen Organisation des WWW sollte bei der Definition von XML-Tags berücksichtigt werden, dass ein gleichnamiges Tag bereits in einem anderen Zusammenhang innerhalb von anderen Dokumenten verwendet werden kann. Um Konflikte durch gleichnamige Tags zu vermeiden, sollten zur Unterscheidung für die Tag-Bezeichnungen in Online-Publikationen Namensräume verwendet werden. Bei den Namensräumen handelt es sich um ein durch „:“ getrenntes Präfix vor den Tagnamen. Innerhalb von XML-Dokumenten, in denen solche Namensraumkonstrukte verwendet werden, wird dem Präfix eine eindeutige URI zugeordnet (Abb. 2.10). Mehr über das Konzept der Namensräume findet sich in [Namespaces1999]

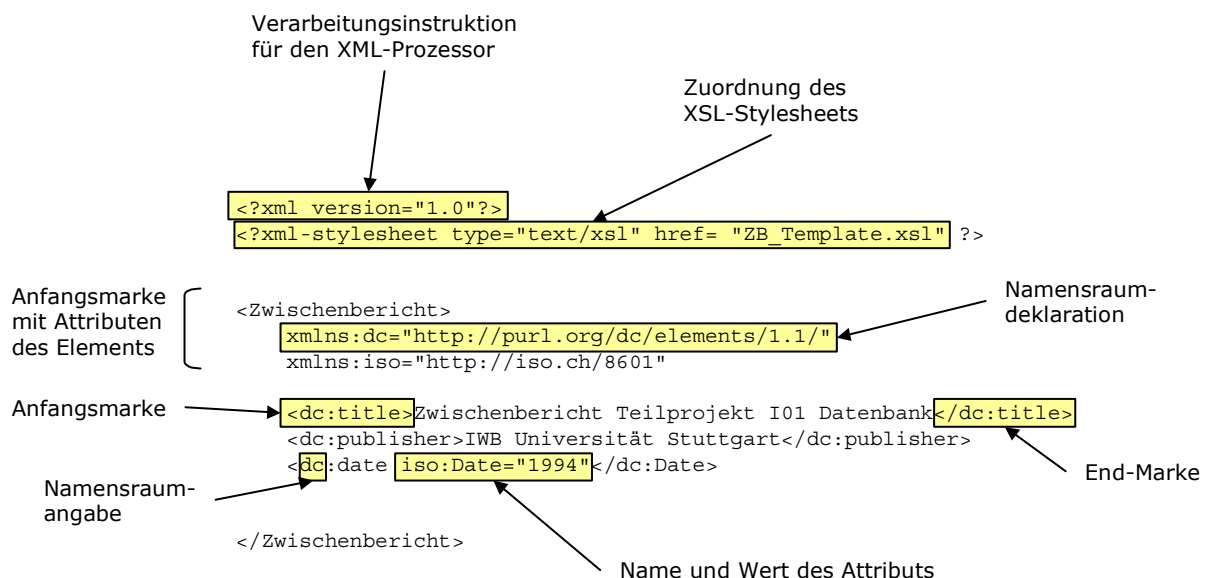


Abb. 2.10 Die Syntax eines XML-Dokuments

Die Formatierung von XML-Dokumenten für einen HTML-Browser erfolgt durch Zuordnen einer Formatvorlage-Datei (Style Sheet). Sie beschreibt, welche Elemente mit den entsprechenden HTML-Formatbefehlen zugeordnet werden. Zur Zeit kann die Formulierung in zwei

unterschiedlichen Standards erfolgen: Die einfachste Art und Weise, XML-Dokumente in einem HTML-Browser anzuzeigen, sind die Cascading Style Sheets (CSS) [CSS1999, CSS1998]. Da dieser Standard für Formatvorlagen bereits in der HTML 4.0 Spezifikation [HTML401Rec] eingeführt wurde, ist er vielen Webautoren bereits vertraut. Konsequenter, da ebenfalls in XML-Syntax codiert, ist allerdings die Nutzung von XSL Style Sheets zur Formatierung von XML-Elementen.

Die Ausführung des Formatierungsvorganges erfolgt bei der Verwendung von XSL in einem sogenannten XSL-Prozessor. Diese Programmkomponente kann sowohl in den Client, z.B. den Webbrowser (Abb. 2.11), als auch bereits im Webserver integriert sein (Abb. 2.12). Grundsätzlich muss zum Einlesen eines XML-Dokuments in den Rechner Speicher eine zusätzliche Softwarekomponente vorgeschaltet sein, die als XML-Parser bezeichnet wird. Der XML-Parser erzeugt aus dem XML-(Text)Dokument eine baumartige Repräsentation, die von anderen Programmen, wie z.B. einem XSL-Prozessor, bearbeitet werden kann. Eine brauchbare, clientseitige Darstellung von XML-Dokumenten ist zum heutigen Zeitpunkt lediglich mit Microsofts Internet Explorer (MSIE) möglich. Ein XML-Parser (MSXML [MSXML2000]) sowie ein XSL-Prozessor sind Bestandteil der neuesten Versionen (ab MSIE 5.0) dieses Webbrowsers. Ebenso wie bei den im vorigen Abschnitt beschriebenen datenbank-gestützten Web-Sites stellt die clientseitige Formatierung eine Entlastung des Webserver dar.

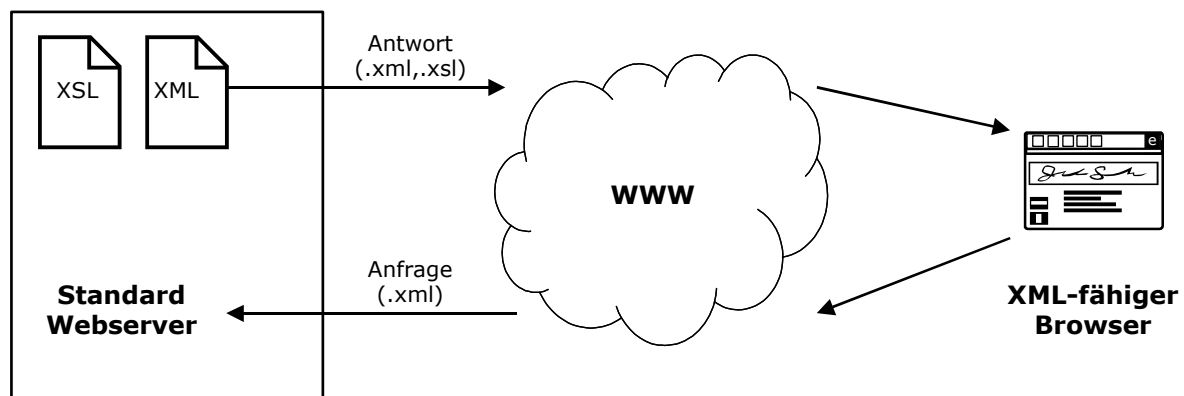


Abb. 2.11 Bei der clientseitigen Formatierung eines XML-Dokuments muss im Webbrowser ein XML-Parser sowie ein XSL-Prozessor implementiert sein, um die Dokumente richtig darstellen zu können.

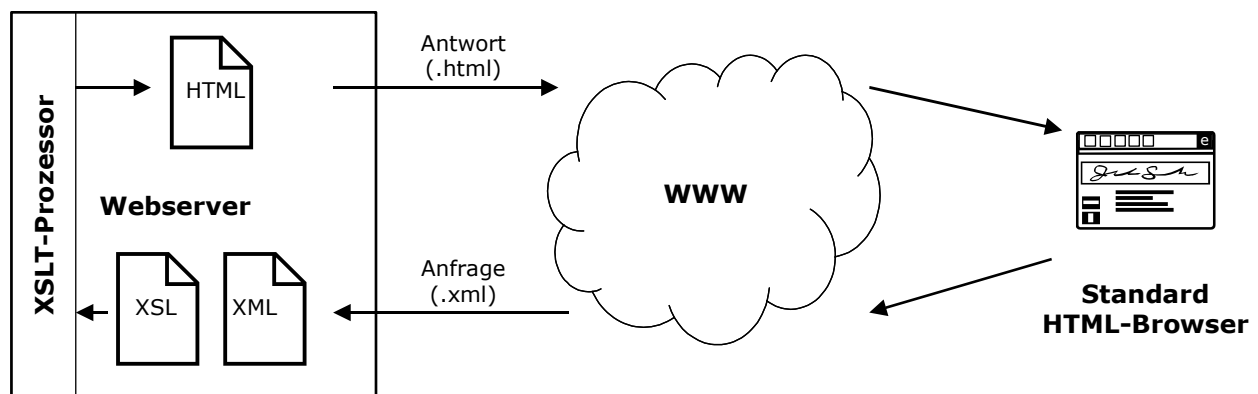


Abb. 2.12 Die Formatierung auf dem Server ist zwar unabhängiger von den Fähigkeiten des Webbrowsers, bedeutet jedoch auf der Serverseite eine hohe Rechenlast

Eine zweite Möglichkeit eröffnet die Installation des XML-Parsers und eines XSL-Prozessors auf dem Webserver. Erfolgt die Formatierung bereits auf dem Server, dann versendet dieser lediglich HTML-Seiten an die Clients. Der Anwender bemerkt in diesem Fall überhaupt nicht, dass es sich bei der Seite ursprünglich um ein XML-Dokument handelt. Diese Vorgehensweise besitzt den Vorteil, dass die Nutzung eines Systems unabhängig von den XML Fähigkeiten der Browser ist. Der Nachteil dieser Lösung besteht darin, dass bei jedem Seitenaufruf HTML-Code aus den Datenquellen erzeugt werden muss, was vor allem bei größeren Datenbeständen und vielen zu generierenden Zieldokumenten eine erhebliche Serverleistung in Anspruch nimmt.

Die dritte Möglichkeit stellt das sogenannte „Content Negotiation“ dar. Dabei „klärt“ der Server vor dem Versenden der Daten die Fähigkeit des Clients zur Interpretation von XML ab. Erst wenn der Webbrowser keinen XML-Code interpretieren kann, schickt er HTML-Daten über das Netz.

Besonders interessant für Webanwendungen sind die erweiterten Verknüpfungsmöglichkeiten („Linking“) von XML. Außer unidirektionalen Hyperlink-Verknüpfungen, wie sie in HTML möglich sind, erlaubt XML weitergehende Verweise zwischen Webinhalten. Aus einem XML-Dokument heraus lassen sich z.B. direkt einzelne Elemente in anderen Dokumenten ansprechen, z.B. ein bestimmter Absatz oder eine Grafik. Im Gegensatz zu HTML muss dabei am Zieldokument keine besondere Markierung vom Autor angebracht werden. Das Sprungziel wird anhand der Struktur des Dokuments ermittelt. Auch werden von einem Punkt aus mehrfache Links ermöglicht werden, die dann auf dem Bildschirm dem Anwender zur Auswahl angeboten werden. Die beiden für die Erstellung von Hyperlinks relevanten XML-Standardentwürfe sind XLink [XLink2000] und XPointer [XPointer2001].

In seinen Möglichkeiten ist XML damit HTML weit überlegen. In der folgenden Aufzählung sind die wichtigsten Verbesserungen bezüglich der Verwendung von XML im Web zusammengefasst:

- Präzisere Suchanfragen mit qualitativ besseren Ergebnissen werden möglich.
- Flexible Webanwendungen werden realisierbar, bei denen Suchanfragen oder Datensätze an andere Anwendungen, Objekte oder Server im Internet übergeben werden können.
- XML-Daten lassen sich einfach an unterschiedliche Darstellungen anpassen, bspw. abhängig von der Bildschirmgröße oder den Programmeinstellungen.
- Bei XML handelt es sich um eine auf das Web optimierte Technologie mit offenen Standards.

Wird XML also in den nächsten Jahren den HTML-Standard bezüglich der Publikation von Informationen im World Wide Web verdrängen?

Trotz der vielen Vorteile und neuen Funktionalitäten wird XML den HTML-Standard zunächst nur in bestimmten Bereichen ablösen. HTML ist als Sprache zur Erstellung von Webinhalten inzwischen sehr stark verbreitet und hat sich bewährt. Von den Unternehmen wurde eine Menge Geld in IT-Systeme mit HTML-Oberflächen investiert. Es ist daher nicht zu erwarten, dass für die Publikation von Webinhalten XML die Sprache HTML bald ablösen wird. Der Entwickleraufwand für einfache Webseiten ist bei XML deutlich höher als bei HTML. Der Grund dafür liegt darin, dass zur sinnvollen Nutzung der Vorteile die Inhalte in XML auch logisch strukturiert werden müssen und die Ausgabe dieser Inhalte zusätzlich über XSL o.ä. programmiert werden muss. Für eine Webseite wären dann schon zwei Dateien, je eine für XML und eine weitere für XSL, notwendig.

Langfristig sollen jedoch alle Informationen im WWW auf XML als gemeinsamem Datenformat aufbauen. Aus diesem Grund wurde HTML in XML als Extensible HyperText

Tabelle 2.1 Arbeitsgruppen innerhalb des XML-Konzepts des W3C

Titel	Thema	Status*
XML 1.0 Extensible Markup Language	Definiert das XML-Datenmodell und die XML-Syntax	REC
XML-Schema	Legt die Struktur von XML-Dokumenten fest und die in den Datenelementen verwendeten Datenformate. Die Syntax von XML-Schema Dokumenten ist ebenfalls XML.	REC
XSL 1.0 Extensible Stylesheet Language	Formatiert die Darstellung von XML-Dokumenten. Ermöglicht die Übersetzung von XML-Dokumenten in HTML.	REC
XLink und XPointer	XLink stellt eine Syntax für Hyperlinks innerhalb von XML-Dokumenten dar. Mit XLink lassen sich sowohl unidirektionale Links wie in HTML, als auch komplexere Hyperlinks konstruieren. XPointer benützt den sogenannten Fragment Identifier (#), um auf Objekte innerhalb von XML-Dokumenten oder XML-Anwendungen zu verweisen. Es erlaubt den direkten Zugriff auf einzelne Elemente, Eigenschaften dieser Elemente und Angaben zu deren Position im XML-Baum.	WD REC
XPath 1.0 XML Path Language	XPath ist eine Sprache, welche die Adressierung von Teilen eines XML-Dokuments erlaubt. Sie wird sowohl von XSL als auch von XPointer genutzt.	REC
XMLQuery	Stellt ein Datenmodell für XML-Dokumente bereit, auf welches eine Abfragesprache mit zugehörigen Abfrageoperatoren ausgeführt werden kann. Die beiden Letztgenannten sollen ebenfalls in XML Query definiert werden. Die Abfragen sollen sowohl auf einzelne Dokumente, als auch auf Dokumentgruppen angewendet werden können. Die Rückgabewerte können ganze Dokumente bzw. Teile von Dokumenten sein.	WD
DOM Level 1&2 Document Object Model	Das Document Object Model (DOM) ist eine plattform- und programmiersprachenneutrale Schnittstelle, die es Programmen erlaubt, den Inhalt, die Struktur und die Formatierung von Dokumenten dynamisch zu verändern. Das dabei entstehende neue Dokument kann weiterbearbeitet werden bzw. kann ein veränderter Dokumententeil wieder in das ursprüngliche Dokument integriert werden.	REC

* REC = Recommendation, WD = Working Draft

Markup Language (XHTML) [XHTML2000] neu definiert. Bei XHTML handelt es sich daher um keine neue Version von HTML, sondern es wurde die HTML 4.0 Spezifikation in XML konformer Syntax erstellt (vgl. Abb. 5.1). Verbesserungen und Erweiterungen, wie z.B. durch die in Abschnitt 2.6.3 beschriebenen XML-basierenden Sprachen für Gleichungen und Formelzeichen, werden in XHTML dadurch möglich.

Auf dem Konzept von XML bauen noch einige weitere Standards auf, welche die unterschiedlichen Aspekte der Dokumentendarstellung und –manipulation betreffen. Die Standardisierung im Bereich von XML ist noch nicht abgeschlossen. Die Vielzahl an XML-Arbeitsgruppen innerhalb des W3C, die sich z.T. mit konkurrierenden Entwürfen zu Empfehlungen beschäftigen (Empfehlungen besitzen beim W3C standardähnlichen Status), erschwert den Überblick über die Entwicklung in diesem Bereich. Das W3C unterscheidet von der Vorbereitung bis zur Verabschiedung einer Empfehlung folgende Phasen [W3CProcess2001]:

- Working Draft (WD)
- Last Call Working Draft
- Candidate Recommendation (CR)
- Proposed Recommendation (PR)
- Recommendation (REC)

Eine Übersicht über die wichtigsten Aktivitäten im W3C auf dem Gebiet der XML-Standardisierung gibt Tab. 2.1. Die angeführten Standards werden dabei den o.g. Phasen zugeordnet.

2.6.3 XML-Anwendungen mit Relevanz für das Bauingenieurwesen

Das XML-Konzept selbst wird mit seinen Standards innerhalb der Arbeitsgruppen des W3C weiterentwickelt. Dies gilt auch für einige branchenübergreifende Anwendungen wie MathML oder SVG, die im folgenden noch erläutert werden. Die anwendungsspezifischen Tag-Vokabulare von XML, die sogenannten „XML-Anwendungen“ (vgl. Abb. 5.1), werden häufig von unabhängigen Interessengruppen entwickelt. Es gibt bereits eine ganze Reihe von Unternehmen, die sich in Gremien mit dem Ziel zusammengeschlossen haben, XML als Standard für Datenformate für einen verbesserten Datenaustausch zu etablieren.

In branchenbezogenen Arbeitsgruppen und Ausschüssen entstehen auf diese Weise Markup-Sprachen mit einem definierten, fachspezifischen XML-Vokabular. Nachfolgend werden einige allgemeine XML-Anwendungen vorgestellt, die für die technischen Informationen des Bauwesens besondere Relevanz besitzen. Für die meisten der hier genannten XML-Anwendungen sind inzwischen Programme zur Darstellung entsprechend codierter Webinhalte im Internet frei verfügbar.

MathML – Mathematical Markup Language [MathML2001]

Mit MathML wird die Darstellung mathematischer Formeln und Gleichungen im Web möglich. Auch Sonderzeichen, wie z.B. Integrale, Brüche und Summenzeichen, die in HTML nicht möglich sind, lassen sich mit MathML über neue Tags darstellen. MathML wurde vom W3C auf der Basis von XML entwickelt und 1998 verabschiedet. Es gibt inzwischen einige Zusatzprogramme für Webbrowser im WWW, mit denen sich entsprechend codierte Inhalte darstellen lassen. Auch von bekannten Mathematik-Programmen wird MathML als Ausgabeformat inzwischen unterstützt.

CML – Chemical Markup Language [CML2002]

Ähnlich wie MathML ist CML für wissenschaftliche Anwendungen und Texte im Netz gedacht. CML fällt in den Bereich der Darstellung zwei- und dreidimensionaler Moleküle. Durch selbstdefinierte Tags wie <ATOMS> oder <FORMULA> (Definition für chemische Formeln) sollen Moleküle grafisch dargestellt werden.

SVG – Scalable Vector Graphics [SVG2001]

SVG besitzt seit 2001 den Status einer Recommendation des W3C. Es handelt sich um eine Sprache, mit der zweidimensionale Vektorgrafiken auf der XML-Syntax aufbauend in HTML- und XML-Dokumente integriert werden können. Bisher konnten im Web nur Pixelgrafiken eingesetzt werden, die an eine festgelegte Größe gebunden sind. Wird eine Pixelgrafik in der Größe verändert, entsteht ein Problem, denn es müssen entweder Pixel hinzugefügt oder herausgelöscht werden, was mit einem beträchtlichen Qualitätsverlust einhergeht. Vektorgrafiken können hingegen aufgrund ihrer eindeutigen mathematischen Definition beliebig skaliert werden.

X3D – Extensible 3D [Web3D2002]

Die VRML (Virtual Reality Modelling Language) wurde als VRML97-Standard spezifiziert, um dreidimensionale Abbildungen im Web zu realisieren. Sie erwies sich jedoch als recht komplex und wird nun, auf der Basis von XML, unter dem Namen X3D neu definiert.

Die Entwicklung bzw. Neudefinition von Datenmodellen mit der XML als universeller Grundlage hat auch vor dem Bauwesen nicht halt gemacht. Besonders im Projektsteuerungs- und CAD-Bereich findet die Idee eines Datenaustauschformats auf der Basis von XML Beachtung.

aecXML – architecture, engineering and construction XML-Framework [aecXML2000]

Diese Gruppe arbeitet unter der Federführung des US-amerikanischen Bausoftwareunternehmens Bentley Systems an der Entwicklung eines XML-Schemas, das die in Kommunikationsprozessen übermittelten Informationen zwischen den beteiligten Planern, Bauunternehmen, Gebäudebetreibern, Baustellen sowie weiteren Beteiligten definiert. Ziel ist, dass den unterschiedlichen Softwareprogrammen, die von den Beteiligten genutzt werden, mit aecXML ein gemeinsames Datenformat zur Nachrichtenübermittlung, der Koordination und Synchronisation logisch zusammenhängender Projektinformationen zur Verfügung gestellt wird.

Versprechungen, dass Daten lediglich einmal eingegeben werden müssen und diese dann auch über Unternehmens- bzw. Bürogrenzen hinweg verwendet werden können, sollen mit dieser Technologie endlich eingelöst werden. Gleichzeitig sollen die aecXML konformen Standards die Recherche nach Bauprodukten und Bauvorschriften im Internet detaillierter möglich machen, als dies mit der heute verfügbaren Volltextsuche möglich ist. Dokumentenbasierende Abläufe, wie die Ausschreibung und Vergabe von Bauleistungen über das Internet sollen dadurch weitestgehend automatisierbar werden. Durch die gemeinsame Strukturierung aller Planungsinformationen in XML lassen sich diese als Bauwerksdokumentation auch für den Betrieb des Bauwerks einfacher weiterverwenden. Zur Integration in allgemeine eBusiness Netzwerke soll aecXML in Microsofts BizTalk Konzept integriert werden, welches einen allgemeinen, branchenübergreifenden XML-Standard für den elektronischen Informationsaustausch in B2B-Systemen darstellt.

bcXML – building and construction XML [bcXML2002]

Innerhalb des europäischen Forschungsprojekts IST 10303 (eConstruct) wird ein Standardisierungsvorschlag mit ähnlicher Zielrichtung wie aecXML entwickelt. Das Projekt wird von einem internationalen Gremium aus Vertretern europäischer Hochschulen sowie der Bau- und Bausoftwareindustrie geleitet.

Es soll zunächst ein sogenanntes „controlled vocabulary“ für das Bauwesen entworfen und in mehrere europäischen Sprachen umgesetzt werden. Mit diesem Vokabular soll dem elektronischen Handel mit Bauprodukten und anderen internetbasierenden Geschäftsprozessen im Bauwesen eine Grundlage für Datenkommunikationsprozesse in die Hand gegeben werden. Es handelt sich dabei um ein die Einzeldisziplinen des Bauwesens übergreifendes Konzept. bcXML soll sich sowohl auf Begriffe aus den Entwurfs- und Planungsphasen, als auch der Bauausführung und dem Bereich der angebundenen Bürosoftware erstrecken. Hervorzuheben ist, dass bcXML in seinem Entwurf schon bestehende Klassifikationen für das Bauwesen berücksichtigt. Das neu zu schaffende Vokabular soll zunächst zu den in Abb. 2.13 dargestellten Standards kompatibel sein und damit gleichzeitig die einzelnen Bereiche, welche diese Klassifikationen abdecken, mit einem gemeinsamen Konzept verknüpfen.

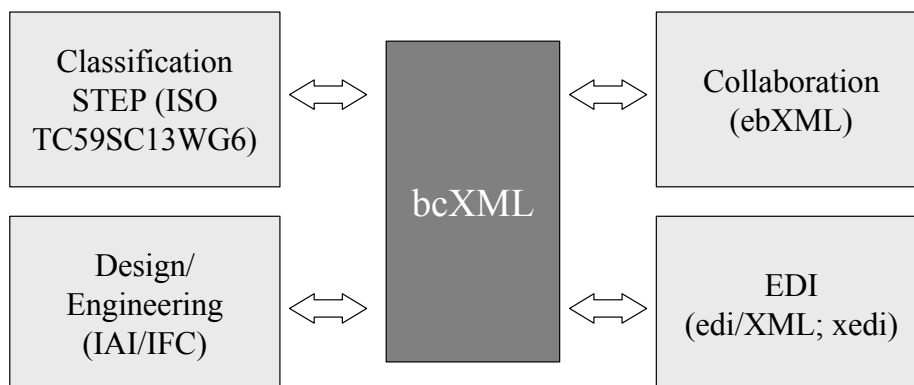


Abb. 2.13 bcXML integriert die abgebildeten Standards zu einem gemeinsamen XML-Konzept für das Bauwesen [Steinmann2001]

bcXML selbst besteht dabei aus mehreren Komponenten. Im Modul „bcDictionary“ werden standardisierte Beschreibungen von Bauprodukten und Dienstleistungen am Bau definiert. Dabei handelt es sich lediglich um eine Definition der Begrifflichkeiten selbst, nicht um eine hierarchische oder andersartige Ordnung der Begriffe zueinander. Die Strukturierung erfolgt in „bcTaxonomy“ anhand einer „Meta“-Struktur, innerhalb der die Begriffe als Objekte (unterteilt in Produkt, Dienstleistung und Ressource), Eigenschaft, Maßeinheit oder Assoziation charakterisiert werden. Das entstehende Objektmodell wird mit der international anerkannten Klassifikation der International Alliance for Interoperability (IAI) abgestimmt, innerhalb welcher auch der aecXML-Standard entwickelt wird [aecXML2000]. Zur Modellierung der Kommunikation zwischen Geschäftspartnern wird der branchenübergreifende ebXML-Standard integriert. Erste Testapplikationen sollen in Zusammenarbeit mit der Nemetschek AG im Bereich der Fertigteilindustrie erfolgen.

Als auf XML-basierende Datenstandards für Bausoftwareanwendungen bieten aecXML und bcXML die Chance, den Datenaustausch zwischen Programmen und Baubeteiligten aller Fachrichtungen unter Einbeziehung moderner Internettechnologie nachhaltig zu verbessern. Bezüglich der Recherchierbarkeit von fachlichen Informationen im Web lassen die genannten Standards jedoch noch einige Fragen offen. All diese Konzepte zum standardisierten Aufbau von XML-Beschreibungen orientieren sich an der Voraussetzung, dass die Daten bereits als XML-Dokumente vorliegen. Dies ist für die meisten der heute im WWW verfügbaren Dokumente jedoch nicht der Fall. So existieren im WWW auch heute schon

Informationssysteme für das Bauwesen, die ihre Inhalte in HTML oder anderen gebräuchlichen Datenformaten vorhalten. Aus den weiter oben erwähnten Gründen ist auch nicht zu erwarten, dass die Publikation von Dokumenten im WWW in den kommenden Jahren ausschließlich auf der Basis von XML erfolgen wird.

Auch das Problem der semantischen Heterogenität bestehender XML-Tags wird durch die Einführung von Tag-Vokabularen alleine nicht gelöst. Wird in bestehenden Datenmodellen ein XML-Element mit derselben Bedeutung, jedoch einer anderen Bezeichnung verwendet, so wird dies von standardgemäßer Software nicht als gleichbedeutend erkannt. Es fehlt innerhalb der bislang vorgestellten XML-Anwendungen daher eine Möglichkeit, den Sinngehalt bzw. die Semantik von XML-Tags, die sich dem Menschen durch den Elementnamen erschließt, auch für den Computer erkennbar zu codieren.

Diese Problemstellung ist Teil der Disziplin Wissensverarbeitung in der Informatik. Im Zuge der Forschungen hinsichtlich künstlich intelligenter (KI) Software in den 80er und 90er Jahren des vergangenen Jahrhunderts wurden Konzepte zur Modellierung semantischer Zusammenhänge entwickelt. Einige der damals gemachten Erfahrungen spiegeln sich in einer XML-Anwendung wider, der heute als eine der zukunftssträchtesten Entwicklungen für das WWW angesehen wird: Das Resource Description Framework (RDF). Diese XML-Anwendung wurde auch als Formalismus für das in dieser Arbeit vorgestellte met@BiM-Datenmodell verwendet und wird deshalb ausführlicher in Kapitel 5 dieser Arbeit beschrieben.

2.7 Recherchen im World Wide Web

Die dezentrale Organisation des WWW erforderte schon in einer frühen Phase Konzepte, die es ermöglichten, die wachsende Informationsmenge organisiert zu erschließen. Ohne entsprechende Suchdienste und Webkataloge wäre eine sinnvolle Nutzung des WWW in seinem heutigen Umfang kaum mehr vorstellbar.

Trotz der erstaunlichen Leistungsfähigkeit moderner Suchmaschinen zeigen Studien [Bergmann2000], dass mit den heute verwendeten Technologien nur ein Bruchteil der im WWW vorhandenen Informationsmenge genutzt werden kann. Um die Ursachen hierfür sowie für die in Abschnitt 1.3 beschriebenen Unzulänglichkeiten bzgl. der professionellen Nutzung des Web als Informationsmedium für Bauingenieure zu verstehen, müssen zunächst die Funktionsweisen der Recherchedienste betrachtet werden. Diese lassen sich im wesentlichen den nachfolgend beschriebenen beiden Typen zuordnen.

2.7.1 Redaktionelle Katalogsysteme und Portalsites

WWW-Katalogsysteme bestehen aus hierarchisch aufgebauten Rubriken, sogenannten Taxonomien. Die Zuordnung von Web-Sites in die entsprechende Kategorie erfolgt durch einen Redakteur. Das wohl bekannteste Beispiel für einen allgemeinen Webkatalog ist der Internet-Dienst Yahoo!. Autoren, die ihre Web-Site in diesem Online-Katalog aufgeführt haben wollen, müssen in einem HTML-Formular zunächst einen Vorschlag zur Zuordnung der Site zu einer Rubrik angeben. Die Bearbeitung und das Einpflegen der Site in die Rubrikenhierarchie erfolgt dann durch den Redakteur bei Yahoo!.

Die Oberfläche des Systems geht über die Gruppierung von Webinhalten hinaus. Sie bietet zusätzlich Zugang zu eCommerce-Angeboten wie Online-Shops und Marktplätzen, sowie diverse Kommunikationsangebote wie Diskussionsforen, Web-eMail und Web-SMS. Außerdem besteht die Möglichkeit, die vielfältigen Angebote der Nutzeroberfläche den eigenen Bedürfnissen anzupassen, zu „personalisieren“. Der Zugriff auf die Inhalte ist flexibel gestaltet und kann sowohl über vorgegebene Strukturen als auch über die Freitextsuche

erfolgen. Ein solcher Webkatalog stellt damit eine Informationsschnittstelle dar, welche die in einer Vielzahl von unterschiedlichen Webangeboten enthaltenen Informationen und Dienstleistungen über eine Systemoberfläche zugänglich macht. Ein solcher Knotenpunkt im Web wird von vielen Anwendern als Einstiegsseite bei der Nutzung des WWW benutzt und daher auch als Portal-Site bezeichnet [Cox]. Die Vorteile von Portalen liegen in

- der Auswahl qualitativ hoher, für die jeweiligen Rubriken relevanter Webadressen,
- der Vorgabe einer Strukturierung, um das Problem des „Information-Overloads“ zu reduzieren,
- der Möglichkeit, menschliche Assoziationen zu berücksichtigen und Inhalte kontextspezifisch einzuordnen.

Dieses Konzept wird auch von zumeist größeren Unternehmen erfolgreich zur Verbesserung der unternehmensinternen Kommunikation und Information genutzt. Unternehmens- oder branchenspezifische Portale werden als sogenannte Enterprise Information Portals (EIP) bezeichnet. Der logische Aufbau eines solchen EIP ist in Abb. 2.14 schematisch skizziert.

Je nachdem, welche Zielgruppe durch das Portal in ihrer Tätigkeit unterstützt werden soll bzw. wer die Informationen bereitstellt, unterscheidet man die in Tab. 2.1 aufgelisteten Typen von Portalen [Cox].

Tab 2.2 Typen von EIP Systemen

Typ	Aufgabe
B2E Portals (Business-to-Employee)	Besserer Zugang innerhalb eines Unternehmens zu Unternehmensinformationen und -anwendungen, Verbesserung der horizontalen und vertikalen Kommunikation innerhalb von Unternehmen
B2B Portals (Business-to-Business)	Verbesserung der Geschäftsabläufe zwischen Unternehmen. V.a. in den Bereichen Einkauf bzw. Bestellung
B2C Portals (Business-to-Consumer)	Erschließen neuer Vertriebswege im WWW. Zumeist über einen Online-Shop, in dem der Kunde über das WWW Waren bestellen kann.

Jedoch auch die Informationserschließung über Portale birgt Nachteile in sich. Die Wege, die der Anwender auf seiner Suche nach Informationen wählt, sind abhängig von seiner persönlichen Denkweise und somit sehr individuell. Eine zu straff organisierte Verwaltung der Inhalte führt hier häufig zu Schwierigkeiten, die Inhalte in den Strukturen des Systems aufzufinden bzw. sie darin einzupflegen. Oft lassen sich Inhalte mehreren Rubriken gleichzeitig zuordnen oder es ist unklar, auf welcher Hierarchieebene ein Angebot einzuordnen ist. Die Denkweise, die dem Entwurf solcher Taxonomien (Rubrikenhierarchien) von Portalen zugrunde liegt, ist häufig von einer hierarchischen Schachtelung der Inhalte geprägt [Goldmann2001]. Dabei ist das Assoziative eine der Grundeigenschaften des Hypertextes und entspricht der heutigen Vorstellung von den menschlichen Denkvorgängen weitaus besser.

Ein anderes Problem betrifft die Pflege und Aktualisierungsgeschwindigkeit solcher Systeme, wenn der Umfang der Informationen sehr groß wird. Bei einer schnell wachsenden Datenbasis können Redakteure mit der Erfassung nicht Schritt halten. Sich rasch ändernde Webinhalte stellen bei sehr großen Informationsangeboten ebenfalls ein fast unlösbares Problem dar für Systeme, die wie Online-Portale auf manuelle Kategorisierung der Inhalte angewiesen sind.

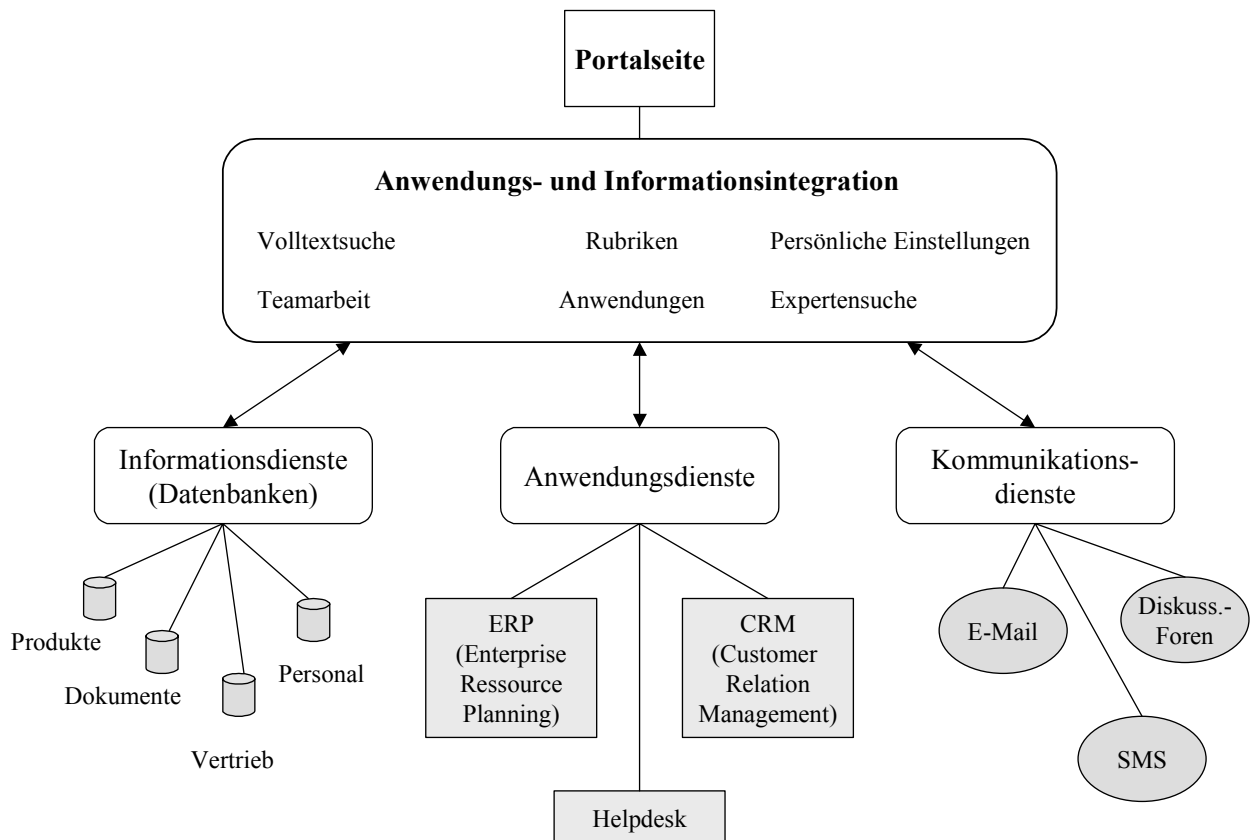


Abb. 2.14 Schematischer Aufbau eines Enterprise Information Portal. Integrierende Systemebenen fassen technisch heterogene und verteilte, jedoch funktionell zusammengehörende Dienste zusammen. Der Zugang erfolgt in der Regel über eine einzelne Portalseite.

Für die Recherchierbarkeit der Inhalte einzelner Webseiten eignen sich allgemeine Web-Portalsysteme hingegen nicht. Eine feinere Detaillierung wäre hinsichtlich des erforderlichen Aufwands für die Redakteure auch schwerlich möglich.

2.7.2 Suchmaschinen

Suchmaschinen im WWW sind komplexe Systeme. Bei den Prozessen, die innerhalb von Suchmaschinen ablaufen, unterscheidet man:

- Erfassen der Dokumente (Crawler)
- Aufbereiten der Dokumente zur Speicherung in einem Index (Indexierung)
- Bereitstellen einer leistungsfähigen Suchfunktion für diesen Index (Searcher)
- Überwachung der Aktualisierungsintervalle für die Indexeinträge

Die einzelnen Funktionen einer typischen Suchmaschine verteilen sich daher auf mehrere Komponenten und Datenstrukturen, die in Abb. 2.15 gezeigt werden. Das in der Grafik abgebildete Schema stellt dabei eine idealisierte Abbildung der Kernkomponenten einer Suchmaschine dar. Die heutigen Systeme besitzen häufig noch zusätzliche Komponenten bzw. Datenstrukturen, welche ihre besonderen Leistungsmerkmale verwirklichen.

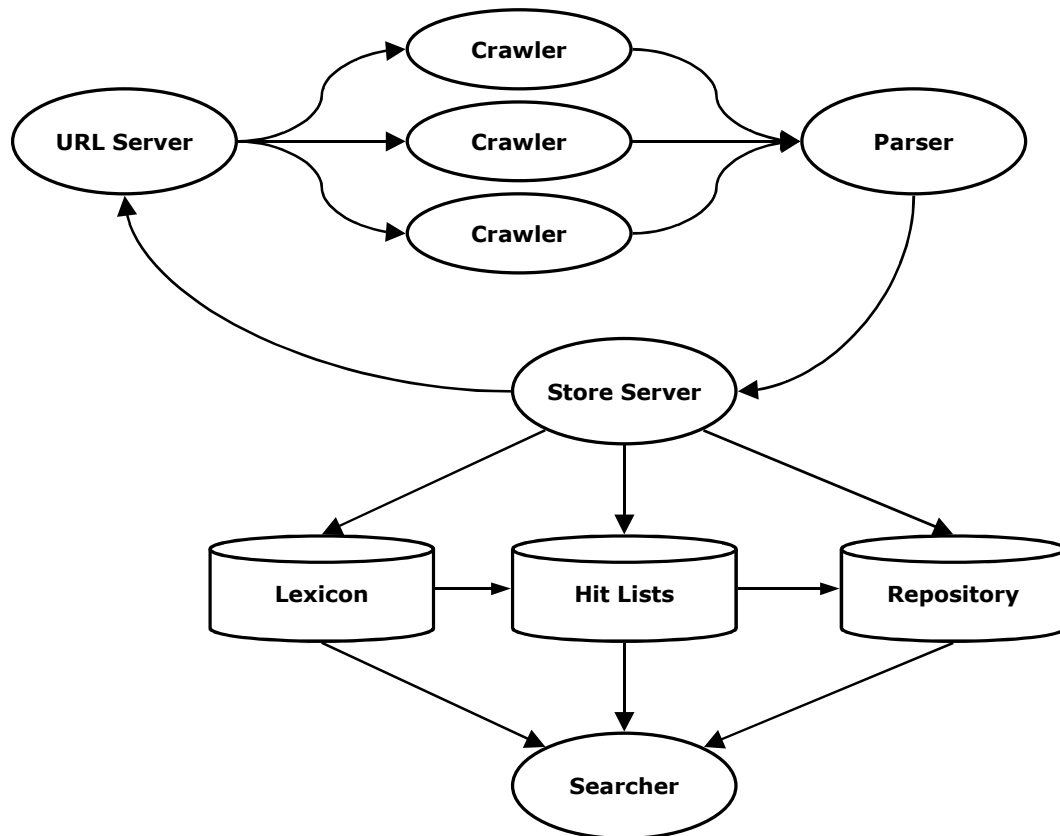


Abb. 2.15 Aufbau der Suchmaschine Google [Brin1998, Wichmann1999]

Der URL Server, ein oder mehrere Crawler, der Parser sowie der Store Server bilden die Webcrawler-Komponente. Sie durchläuft das Web, indiziert alle gefundenen Seiten und bereitet diese zur Speicherung in den Datenstrukturen vor. Der URL Server verwaltet die URLs, welche noch aus dem Internet geholt, indiziert und verarbeitet werden müssen. Er gibt diese an die Crawler weiter, sobald neue Kapazitäten frei werden. Für die effiziente Organisation der Datenerfassung ist es wichtig zu entscheiden, in welcher Reihenfolge die URLs abgearbeitet werden sollen. Diese Entscheidung wird daher bereits im URL Server getroffen. Unabhängig von der Reihenfolge besteht aber das Problem, dass nur Seiten erfasst werden können, auf die auch Links existieren. Isolierte Seiten oder Bereiche des WWW, auf die kein Verweis führt, werden auf diesem Wege nicht gefunden. Aus diesem Grund bieten die meisten Suchmaschinen an, URLs manuell dem URL Server zu übergeben, z.B. durch die Möglichkeit der Anmeldung von Seiten.

Crawlerprogramme werden häufig auch als Robots oder Spiders bezeichnet. Es handelt sich bei diesen Programmen um selbstständig arbeitende Webbrowser, die automatisch die Hyperlinks von Webseiten weiterverfolgen und dabei den gesamten Inhalt oder Teile der Webseiten zur Erfassung über den Parser an die Indexdatenbank der Suchmaschine weiterleiten. Bei großen Suchmaschinen, die einen möglichst umfassenden Index erzeugen wollen, ist es sinnvoll, gleichzeitig mehrere Crawler auf verschiedenen Systemen laufen zu lassen. In [Brin1998] wird berichtet, dass ein Crawler typischerweise 300 Verbindungen gleichzeitig offen hat und das Gesamtsystem in Spitzenzeiten mit 4 Crawlern etwa 100 Seiten pro Sekunde erfasst. Problematisch bei dieser Komponente ist die enorme Netzlast, welche besonders beim Crawler-System und bei den Zielserversn entsteht.

Nachdem eine Seite aus dem Internet geholt wurde, muss sie vom Parser zur Analyse vorbereitet werden. Der Parser erstellt nun aus jeder Seite ein am baumartigen Aufbau von

HTML-Dokumenten orientiertes Abbild der Seite. Dieses übergibt er dann dem Store Server zur Verarbeitung.

Die Aufgabe des Store Servers ist es, aus der vom Parser erhaltenen Abbildung der Webseite die für die Suchmaschine wichtigen Informationen zu extrahieren. Anhand dieser Informationen erfolgt die Auswahl der URLs. Bei der URL-Auswahl muss besonderer Wert darauf gelegt werden, dass keine dynamisch generierten Seiten angefordert werden. Es wäre fatal für die Suchmaschine, wenn sie sich in der Anfragemaske einer anderen Suchmaschine oder in dynamisch generierten Seiten eines Online-Spiels verfangen würde. Im Text der geholten WWW-Seite wird dann nach neuen Wörtern gesucht und bei Bedarf das Lexikon erweitert. Zu jedem Wort der Seite wird in den Hit Lists vermerkt, wie oft es in dieser Seite vorkommt. Schließlich wird ein Teil der Seite im Repository abgespeichert. Manche Suchmaschinen speichern nur den Titel, andere die ersten 20 Zeilen. Mittlerweile geht man immer mehr dazu über, die gesamte Seite komprimiert zu speichern.

Im Lexikon werden alle Begriffe gespeichert, die im Netz angetroffen wurden und nach denen der Benutzer suchen kann. Füllworte, die einem Satz keinen Inhalt beisteuern (Artikel, „und“, „oder“ usw.), werden dabei ausgefiltert. Die Hit Lists enthalten zu jedem Wort des Lexikons eine Menge von Zeigern auf die Seiten im Repository, in denen sie auftreten. So können bei einer Suchanfrage zu einer gegebenen Menge von Suchwörtern schnell die entsprechenden Zeiger zusammengestellt und die Ergebnismenge an URLs berechnet werden.

Der Searcher ist schließlich der Teil der Suchmaschine, den der Anwender zu Gesicht bekommt. Er besteht aus einem Webserver, der das Frontend der angebundenen Datenbank darstellt und auf der Startseite die Anfragemaske bietet. Hat der Searcher die Suchanfrage erhalten, erstellt er mit Hilfe des Lexikons und der Hit Lists die Ergebnismenge, welche mit Hilfe eines Ranking-Algorithmus sortiert wird. Dabei bekommen im einfachsten Fall die Seiten das höchste Ranking, welche die meisten Suchwörter am häufigsten enthalten. Die Treffer auf die als relevant eingestuften Seiten werden zur besseren Übersicht des Anwenders dann mit Informationen aus dem Repository angereichert.

Die Abfragemöglichkeiten von Volltextsuchmaschinen sind dabei unterschiedlich. Bei den meisten Suchdiensten lassen sich Suchbegriffe mit den Ausdrücken „UND, ODER, NICHT“ der Booleschen Algebra logisch verknüpfen [Trinkwalder2000]. Auch das nur teilweise Angeben von Begriffen sowie das Auffinden von Inhalten mit ähnlichen Schlüsselbegriffen über Thesauri ist bei vielen Suchmaschinen inzwischen möglich.

Moderne Suchmaschinen erweitern ihre Repositories um weitere Dateitypen. So werden in der Suchmaschine Google auch Worddokumente, Powerpoint-Präsentationen und andere proprietäre Datentypen erfasst. In einigen Suchdiensten wie Altavista und Google wird darüber hinaus auch eine Suche nach multimedialen Inhalten angeboten. Bild- und Sounddateien sowie Verzeichnispfade lassen sich hierbei über eine grafische Oberfläche recherchieren. Die Abfrage selbst ist dabei jedoch textorientiert. Die Zuordnung der Schlüsselbegriffe zu den multimedialen Inhalten erfolgt über den Dateinamen bzw. durch Auslesen des „alt“-Attributs, mit dem in HTML eingebundenen Multimediainhalten ein kurzer Text zugeordnet werden kann. Ein weiterer Weg führt über die in der Nähe der Abbildung stehenden Texte. Der Suchalgorithmus geht bei der Analyse der Webseite davon aus, dass Begriffe, die im Dokument in der Nähe der Abbildung stehen, auch inhaltlich mit dieser zusammenhängen.

Die automatische Erfassung von Webdokumenten kann von den Web-Site Autoren durch Einbindung manuell erzeugter Metadaten ergänzt werden. Bisher machen Suchmaschinen nur wenig Gebrauch von diesen Möglichkeiten und wenn, dann auch nur von einzelnen Elementen. Altavista und Infoseek z.B. berücksichtigen nur die HTML META-Elemente „Description“ und „Keywords“. Allerdings ist davon auszugehen, dass mit zunehmender

Bedeutung der Metadatenelemente des Dublin Core (vgl. Abschnitt 4.2) diese auch in entsprechendem Maße von den Suchmaschinen berücksichtigt werden.

Trotz dieser Erweiterungen der Inhaltserfassung sind technische Recherchen, wie sie in den Fragestellungen von Ingenieuren an Informationssysteme häufig auftreten, hierdurch nicht möglich. Auch machen die Datenmengen in dynamisch erzeugten, datenbankgestützten Web-Sites inzwischen den größeren Teil der WWW-Inhalte aus [Bergmann2000]. Diese sind nur bis zu den systemspezifischen Abfrageoberflächen für allgemeine Suchmaschinen zugänglich. Der größte Teil themenspezifischer und qualitativ hochwertiger Informationen (Abb. 2.16) ist somit über allgemeine Suchmaschinen, wie sie oben beschrieben werden, nicht zugänglich. Gleiches gilt für kommerzielle Informationsangebote mit passwortgeschütztem Zugang.

Schätzungen über den Umfang des heutigen WWW, welche die in den Suchmaschinen gespeicherten Inhalte zugrunde legen, belaufen sich auf 19 Terabyte. Doch schon die 60 umfangreichen Online-Datenbanken, die in [Bergmann2000] analysiert wurden, besitzen einen Umfang von zusammen 748,5 Terabyte. In der in Nature veröffentlichten Studie [Lawrence2000] wird daher die These vertreten, dass über die herkömmlichen Suchmaschinen lediglich 0,03% der im Web verfügbaren Inhalte erfasst werden.

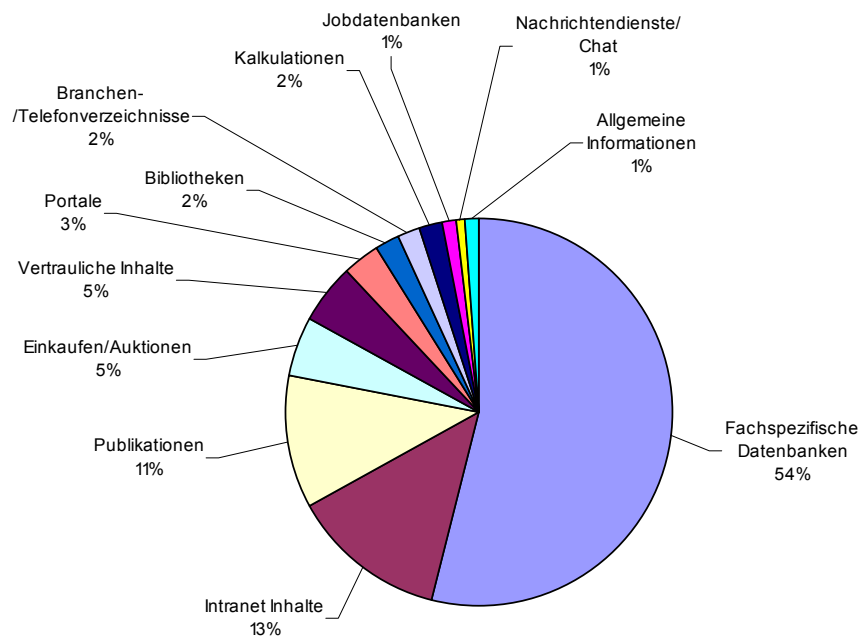


Abb. 2.16 Verteilung der Informationstypen, die laut [Bergmann2000] nicht über konventionelle Suchmaschinen erreicht werden können.

3 Baustofftechnologische Informationen im WWW

Die webbasierte Informationsverarbeitung im Bauwesen steht erst am Anfang ihrer Entwicklung. Eine investitionsfeindliche Baukonjunktur, die für viele Unternehmen der Baubranche eine Neustrukturierung der Tätigkeitsfelder sowie z.T. drastische Gewinneinbrüche mit sich brachte, ließ die Bauwirtschaft bei der Einführung durchgängiger IT-Strategien im Vergleich zu anderen Branchen zurückfallen. Dabei wächst die Zahl derer stetig an, die sich Informationen zu für sie interessanten Themen vorwiegend über das Internet einholen. Das bedeutet, kein Wirtschaftsbereich wird sich künftig ohne digitale Medien am Markt behaupten können.

Nach einer Studie der Industrie- und Handelskammer Stuttgart im Jahr 2000 sind etwa 70% aller Unternehmen mit eigenen Webangeboten im Internet vertreten. Während in der Kreditwirtschaft bereits 95% der Unternehmen das Internet nutzen, sind es im Baugewerbe nur 40%. Für eine Branche, in der Computer für das Erledigen vielfältiger Aufgaben in allen Bereichen zu unverzichtbaren Werkzeugen geworden sind, ist dies sehr wenig.

Die Gründe hierfür sind dabei allerdings nicht nur in der anhaltend schwachen Konjunktur, sondern auch in den Besonderheiten der Branche zu suchen. Im folgenden Abschnitt werden diese Besonderheiten näher betrachtet und anhand eines Szenarios werden die Vorteile einer durchgängigen Nutzung des WWW veranschaulicht.

3.1 Webbasierte Informationsverarbeitung im Bauwesen

Im Gegensatz zu anderen großen Industriebranchen, wie der oben genannten Kreditwirtschaft oder der Automobilbranche, wird die Baubranche überwiegend von mittelständischen, meist hochspezialisierten Unternehmen geprägt. Bei fast jedem Bauprojekt, mit seinen individuellen Problemstellungen für Planung und Ausführung, ändert sich die Zusammensetzung der Projektbeteiligten.

Durchgängige Standardisierungen der Arbeitsprozesse laufen diesen Besonderheiten entgegen und erzeugen für die beteiligten Unternehmen Kosten, die diese aufgrund ihrer geringen Firmengröße häufig nicht tragen können. In der Folge haben sich entsprechende Standards anderer Branchen, wie der weiter oben schon erwähnte EDI, in der Baubranche nicht durchsetzen können. Selbst wenn in größeren Unternehmen durch die Benutzung einheitlicher Software zwischen den hausinternen Abteilungen ein Quasistandard eingeführt wurde, ist dieser für den Datenaustausch mit Projektpartnern in vielen Fällen nicht anwendbar.

Dabei erfordert erfolgreiches Bauen von den Beteiligten interdisziplinäres Denken und Handeln [FOGIB1997]. Für die Planung und Bauausführung bedeutet dies, nicht nur das Wissen der einzelnen, beteiligten Fachrichtung isoliert zu betrachten. Die größte Schwierigkeit besteht nämlich selten in der Optimierung hinsichtlich einzelner Aspekte des Bauens, wie z.B. der technischen Machbarkeit einer Bauaufgabe oder der Kostenoptimierung. Wichtig ist vielmehr, mit der möglichst zeitnahen Bereitstellung und Koordination aufgabenübergreifenden Schnittstellenwissens über eine Entscheidungsgrundlage zu verfügen, mit Hilfe derer die Interaktionen zwischen den Fachrichtungen berücksichtigt werden können. Diese Informationen, die von mehreren Nutzern parallel in ihrem jeweiligen Anwendungskontext benötigt werden, sind es, die den allgemein sehr hohen Kommunikationsaufwand beim Bauen verursachen. Ein typisches Beispiel für Schnittstellenwissen sind Baustoff- bzw. Bauproduktinformationen. Betontechnologische Daten werden beim Betonhersteller benötigt, bei der Planung des Tragwerks und in der Bauausführung vom Bauleiter bzw. dem Controller. Diese Informationen veralten mit der Weiterentwicklung der Produkte bzw. der Baustoffe jedoch sehr schnell und müssen daher ständig überprüft bzw. aktualisiert werden. Der

Informationsbedarf ist besonders groß, wenn über einen neuen Werkstoff noch wenig Erfahrungen vorliegen.

Am fiktiven Beispiel des Baus eines öffentlichen Bürogebäudes soll diese Aussage veranschaulicht werden. Der öffentliche Auftraggeber sieht sich aufgrund des Kreislaufwirtschaftsgesetzes [KrWG1994] dazu verpflichtet, für die Betonbauteile Beton mit rezyklierter Gesteinskörnung vorzuschreiben. Um die Vorteile der Nutzung des WWW für die Beteiligten im obigen Szenario aufzuzeigen, werden die notwendigen Rechercheprozesse zunächst ohne und danach mit WWW-Nutzung betrachtet.

Nach der Erteilung des Auftrags macht sich der Architekt an die Entwurfsplanung. Er möchte im überdachten Innenhof des Gebäudes eine Wand in Sichtbetonbauweise ausführen. Bei der Anfertigung der Planungsunterlagen kommen ihm Zweifel, ob sich dies mit dem besonderen Beton realisieren lässt. Außerdem würde er gerne wissen, ob das Erscheinungsbild der Oberfläche sich von Normalbetonbauteilen unterscheidet. Da er selbst für Recherchen keine Zeit übrig hat, beauftragt er seine Büroassistentin mit der Literatursuche. Diese findet über einen Fachverlag und die Hochschulbibliothek schließlich einige Veröffentlichungen in Fachzeitschriften zum Thema Beton mit rezyklierter Gesteinskörnung. Nachdem einige Tage verstrichen sind, erhält der Architekt die aus der Bibliothek entliehenen Aufsätze und arbeitet sie auf der Suche nach Aussagen zu o.g. Fragen durch.

Dem Tragwerksplaner, der die Pläne mit dem Vermerk „Wand in Sichtbeton n. statischen Erfordernissen“ erhält, geht es nicht anders. Auch er muss zunächst anhand verfügbarer Fachliteratur klären, welche Besonderheiten er bei der Bemessung beachten muss, welche Anwendungsgebiete für diesen Beton zulässig sind und ob Zulassungen für entsprechende Betone existieren. Die Literaturhinweise des Architekten helfen ihm dabei nur bedingt, da dort seine Fragen nicht ausführlich genug behandelt werden.

Nach Abschluss der Genehmigungsplanung erfolgt die Ausschreibung des Projekts. Der Bauunternehmer, der die Ausschreibungsunterlagen in Papierform erhält, übernimmt zunächst manuell die Einzelpositionen in sein LV-Programm. Beim Eingeben der Positionen kontaktiert er seinen Betonlieferanten in der Hoffnung, näheres über die Qualitätssicherungsmaßnahmen, zusätzliche Baustelleneinrichtungen etc. für Beton mit rezyklierter Gesteinskörnung zu erfahren. Dieser kann ihm jedoch lediglich anbieten, sich selbst bei seinen Zulieferern kundig zu machen und ihn in den nächsten Tagen zurückzurufen.

Nachdem der Unternehmer den Zuschlag erhalten hat, muss für den Beton ein Lieferant gefunden werden, der bereits Erfahrung mit dem Recyclingwerkstoff gesammelt hat. Nach mehreren Telefonaten erhält er schließlich von einem Baustoffverband die Adresse eines Betonwerks mit dem gesuchten Betontyp im Sortenverzeichnis. Der Preis pro m³ Beton erscheint ihm zwar etwas hoch, er ist aber zunächst froh, überhaupt einen Lieferanten im näheren Umkreis gefunden zu haben. Erst später erfährt er, dass auch ein anderes Betonwerk seit kurzem den Werkstoff anbietet – zu günstigeren Preisen. Die erforderlichen Mengen werden dem LV entnommen, in die Bestellscheine eingetragen und gefaxt.

Der Betontechnologe im Betonwerk erhält den Bestellschein und gibt die Bestellung in das Buchhaltungsprogramm ein. Bei der Zusammenstellung der Betonrezeptur kommen ihm Zweifel, ob mit der üblichen Vorgehensweise auch bei Verwendung rezyklierter Zuschläge Sichtbetonqualität hergestellt werden kann. Er nimmt sich vor, bis zum Betoniertermin in der folgenden Woche sich in der Fachliteratur Klarheit zu verschaffen.

Der Polier auf der Baustelle zweifelt an der Verdichtbarkeit des Betons. Er befürchtet Entmischungsvorgänge im Frischbeton, wenn dieser, wie für Sichtbeton notwendig, gründlich mit der Rüttelflasche verdichtet wird. Er wendet sich mit seinen Bedenken an den Bauleiter,

der daraufhin mit ihm zusammen die zusammengetragene Fachliteratur nach entsprechenden Angaben durchgeht.

Einige Jahre nach der Fertigstellung des Bauwerks treten an einigen Stellen Risse mit Abplatzungen auf. Auch der mit der Schadensanalyse beauftragte Gutachter recherchiert in Bibliotheken nach Fachliteratur, um Hinweise auf Anwendungsfehler zu finden.

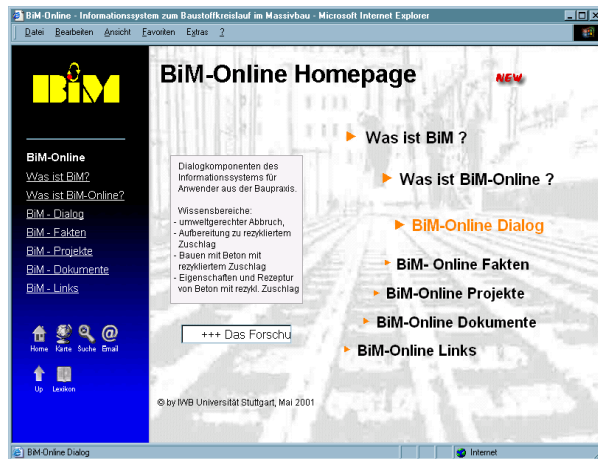
Wie das Beispiel zeigt, ist die konventionelle Recherche in Druckerzeugnissen sehr zeitaufwendig. Selbst wenn man die richtigen Schlüsselbegriffe für die Recherche besitzt, besteht immer noch die Gefahr, dass die gewünschte Zeitschrift oder das Buch gerade nicht verfügbar ist. Nicht nur die Suche nach entsprechender Literatur, sondern auch die Suche in der Literatur nach relevanten Informationen benötigt viel Zeit. Die meisten Quellen werden vermutlich nur einen geringen Anteil an relevanten Informationen enthalten, die zunächst manuell durch das Durcharbeiten der Literatur herausgefiltert werden müssen. Da der Aufwand, an Quellen zu gelangen, sehr groß ist, beschränkt man sich häufig auf eine relativ schmale Literaturgrundlage. Die Gefahr, dabei ein wichtiges Dokument übersehen zu haben, ist recht groß. Wie könnten dieselben Vorgänge aussehen, wenn die Beteiligten konsequent die Möglichkeiten genutzt hätten, die ihnen das Internet bietet?

Der Auftraggeber hätte in diesem Fall den Architekten auf das Informationssystem BiM-Online [BiM-Online2001, Reinhardt2000] zum Thema Beton mit rezyklierter Gesteinskörnung im WWW hinweisen können. Mit seiner Fragestellung hätte der Architekt unter der URL <http://www.b-i-m.de> das System konsultieren und die Systemkomponente BiM-Dialog [Schreyer1999a] mit Informationen für Baupraktiker anwählen können. Dort füllt er sein Interessenprofil (Abb.3.2) aus und gibt als Suchbegriff für die Wissensbasis „Sichtbeton“ an. Das System antwortet ihm mit der Empfehlung der Lektion „Besondere Betone“ [Reinhardt2000]. Dort wird der Architekt bereits nach wenigen Mausklicks fündig und erhält die Aussagen, dass Sichtbetonbauteile bei sorgfältiger Verarbeitung und hoher Zuschlagqualität mit rezyklierten Zuschlägen hergestellt werden können. Die angegebenen Literaturquellen [Reinhardt2000] bestellt er über die Webseite seines Fachverlages zur Ansicht, um die Angaben zu überprüfen. Bei seiner weiteren Recherche im Web findet er noch andere Bauprojekte [Reinhardt2000], bei denen Beton mit rezyklierter Gesteinskörnung verwendet wurde, mit Angaben zu Planern und ausführenden Firmen. Seinen Plänen, die er über den Browser in das Kommunikationssystem auf dem Projektwebserver lädt, fügt er einen Hyperlink hinzu und sendet eine E-Mail an die anderen Projektpartner, in der er auf das Online-Informationssystem aufmerksam macht.

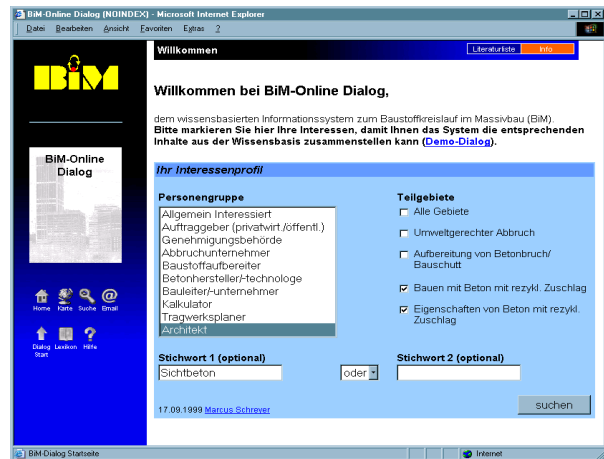
Der Tragwerksplaner liest am nächsten Morgen diese Hinweise, während er die Pläne des Architekten zur Weiterbearbeitung aus dem Projektwebserver lädt. Auch er findet über das Interessenprofil sehr schnell Antworten auf seine Fragen zur Bemessung und zu den Anwendungsmöglichkeiten von Beton mit rezyklierter Gesteinskörnung. Er notiert sich die Aussage, dass Innenbauteile aus Beton bis zu einer Festigkeitsklasse C30/37 ohne Zulassung im Einzelfall möglich ist, und bestellt sich über einen Verweis auf der Web-Site des DIN/DAfStb [DIN2002] gleich die Zugriffsrechte auf die elektronische Version der entsprechenden Richtlinie „Beton mit rezykliertem Zuschlag“ [DAfStb1999]. Einige interessante Artikel innerhalb von BiM-Online druckt er sich aus, um sie in der Kaffeepause später am Tag zu lesen.

Nach Abschluss der Genehmigungsplanung erfolgt die Ausschreibung des Projekts. Der Bauunternehmer, der über die Web-Site der Stadtverwaltung von der Ausschreibung erfahren hat, lädt sich die vorbereiteten elektronischen Ausschreibungsunterlagen als Grundstruktur für sein Leistungsverzeichnis vom Projektwebserver herunter. Die Frage nach zusätzlicher Baustelleneinrichtung klärt auch er in BiM-Online-Dialog [Reinhardt2000]. Nachdem der Unternehmer den Zuschlag erhalten hat, recherchiert er nach Angaben zu Lieferanten und Preisen für die benötigten Baustoffe.

a



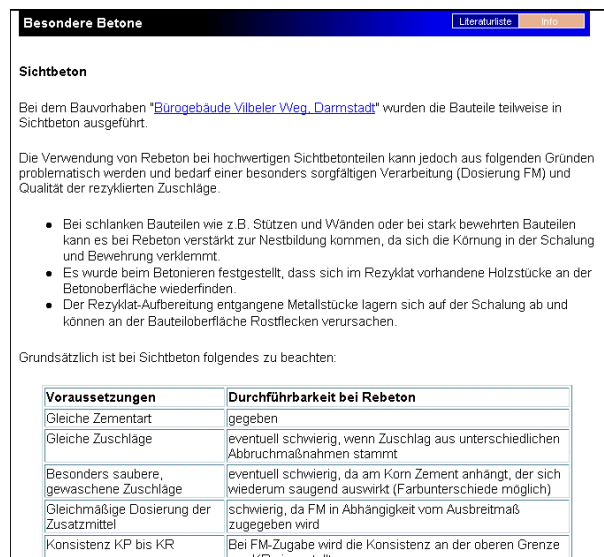
b



c



d



e

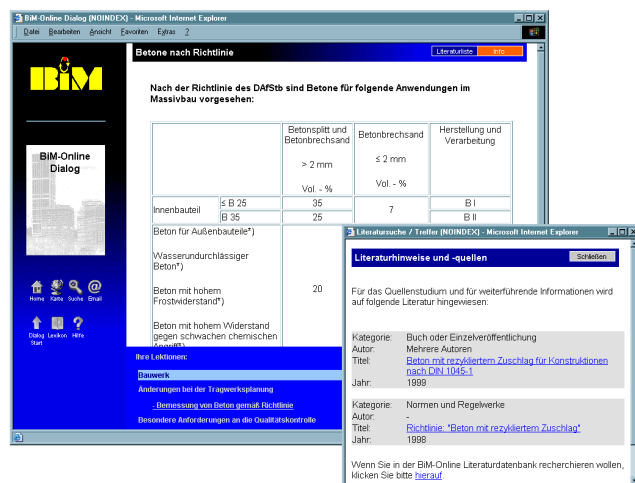


Bild 3.2 a-e [Reinhardt2000]

- Aufruf von BiM-Dialog über die BiM-Online Homepage
- Startseite von BiM-Dialog mit Interessenprofil
- Gemäß den Eingaben wird eine Auswahlliste entsprechender Lektionen erzeugt
- Dialogfenster innerhalb einer Lektion mit der gesuchten Information
- Der Anwender kann über die Schaltflächen „Literaturliste“ und „Info“ weiterführende Informationen abrufen

Da er in BiM-Online hierzu keine Angaben erhält, recherchiert er über eine Suchmaschine nach Betonwerken. Er erhält daraufhin einige Seiten von Betonwerken, die den gesuchten Beton im Sortenverzeichnis führen. Nachdem er die Preise der Werke im näheren Umkreis abgefragt und telefonisch bzw. per E-Mail nachverhandelt hat, bestellt er beim günstigsten Anbieter. Als später am Tag sein Polier mit seinen Bedenken bezüglich der Verdichtung des Betons zu ihm kommt, recherchieren sie noch einmal in BiM-Online. Auch dem Betontechnologen sowie dem Gutachter mit ihren jeweiligen Fragestellungen kann durch das System geholfen werden.

In dem hier beschriebenen Szenario sind die Problemstellungen zwar konstruiert, das BiM-Online-System [Reinhardt2000] und die anderen Softwarewerkzeuge existieren jedoch bereits und werden von Baustoff-Forschern aber auch von Praktikern genutzt. Durch die Publikation qualitativ hochwertiger Informationen im Web wird dieses Fachwissen unabhängig von örtlichen und zeitlichen Beschränkungen, wie z.B. der Bibliotheksöffnungszeit und des Ausleihstatus, verfügbar. Die Informationswege werden hierdurch drastisch verkürzt. Rechercheprozesse lassen sich mit weniger Zeitaufwand und umfassender durchführen. Durch die zentrale Verwaltung der Projektdaten in einem Kommunikationssystem gibt es weniger Redundanzen. Durch eine zentrale Koordination lassen sich Fehler beim Datenaustausch und der Umformatierung schneller erkennen und vermeiden.

Allerdings gibt es noch zu wenige Informationssysteme für das Bauwesen, die wie BiM-Online ein Themengebiet ähnlich vollständig abdecken. Es ist jedoch anzunehmen, dass die Zahl solcher Informationssysteme zu Themen der Bauforschung in Zukunft steigen wird. Wie das folgende Kapitel zeigt, profitiert nicht nur der Wissenstransfer in der Baupraxis von solchen Systemen, sondern auch die Forschungsarbeit selbst. Mit einer größer werdenden Informationsmenge wird jedoch die Auffindbarkeit der in den Systemen enthaltenen Informationen immer wichtiger. Bislang erfolgt die Recherche in zwei Stufen: In einem ersten Schritt führt eine mehr oder weniger ausgefeilte Suchmaschinenabfrage zu einer Liste von Webseiten. Da Volltextsuchmaschinen, wie z.B. AltaVista oder Google, lediglich vage Informationen über die Inhalte der Webseiten besitzen, muss die Trefferliste noch nach relevanten Webangeboten durchsucht werden. Findet man solch spezifische Informationssysteme wie BiM-Online, bieten diese erneut detailliertere Abfragemasken an, die sich an den relationalen oder objektorientierten Datenstrukturen der Wissensbasis orientieren [Schreyer2001a]. Es fehlt bislang ein systemübergreifendes Datenmodell für derartige Informationsquellen, welches sowohl die semantische Einordnung der Web-Sites selbst, als auch deren Inhalte erlaubt.

Außer Informationssystemen zu Themen der Bauforschung gibt es inzwischen eine Reihe von Webangeboten zur Unterstützung der unterschiedlichen Prozesse des Bauens. Tabelle 3.1 gibt eine grobe Zusammenstellung von typischen, branchenspezifischen Dienstleistungen wieder, die für das Bauwesen im Web angeboten werden.

Allen Systemen gemeinsam ist, dass sie außer einem Internetzugang mit einem aktuellen, kostenlos verfügbaren Webbrowser, keine weiteren Softwarevoraussetzungen benötigen. Im Gegensatz zu konventionellen Programmen, die lokal auf den Festplatten installiert werden und normalerweise eng mit dem Betriebssystem und der Hardwareumgebung verwoben sind, erfordert die Nutzung der Software keine zusätzlichen IT-Investitionen. Für die kostenpflichtigen Angebote fallen lediglich Gebühren an, die für die Dauer der Nutzung bezahlt werden müssen.

Tab. 3.1 Einige Dienstleistungen für die Baubranche im WWW

Dienstleistung	Beispiele
Handel mit Bauprodukten (eCommerce)	• Baustoffmarktplätze
Kommunikation zwischen Geschäftspartnern (B2B)	• Online-Ausschreibungen • Projektkommunikationssysteme
Projektdatenaustausch/-verwaltung	• Projektkommunikationssysteme
Fachinformationsdienste	• Online-Zugang zu Fachzeitschriften, Normen etc. • Informationssysteme z. Themen der Bauforschung • Branchennachrichten • Baurechtliche Datenbanken • Bauportale
Mietsoftware über das WWW (Application Service Provider - ASP)	• Bemessungsprogramme für Details • Ausschreibungssoftware • Software Erstellung von Sicherheits- und Gesundheitsplänen (SiGe)
Webbasierte Hochschullehre (eLearning)	• Lehrmaterialien im Web

3.2 Rechercheaspekte

Das WWW spielt als Informationsmedium auch für den Bereich der Baustoffe eine wachsende Rolle. Dass ein Bedarf an derartigen Informationsangeboten besteht, zeigen die Ergebnisse verschiedener Untersuchungen, die in [Scherer1997, Christians1999] vorgestellt werden. Diese gehen von einem Arbeitszeitanteil von 8-30% aus, der allein für die Informationssuche innerhalb von Planungs- und Bauprozessen benötigt wird. Bei den gesuchten Informationen handelt es sich dabei zumeist um Angaben zu Eigenschaften und Angeboten von Bauprodukten [BuildOnline2000], aber auch um Lösungen für Detailfragen der Anwendung unter Berücksichtigung von technischen Dokumentationen und gültigen Normen.

Nicht nur in der Baupraxis, sondern auch in der Forschung ist die Kenntnis bestimmter technologischer Kennwerte unverzichtbar. Die Bemessung und Beurteilung der Gebrauchstauglichkeit von Bauteilen bzw. die Einordnung von Probekörpern für weitere wissenschaftliche Arbeiten hängen davon ab [Rinder2001].

Während sich das Web in seiner heutigen Form sehr gut zur Befriedigung allgemeiner Informationsbedürfnisse eignet, verlaufen Recherchen nach solch fachspezifischen Webinhalten für Bauingenieure häufig wenig zufriedenstellend. Die oftmals hoch komplexen Fragestellungen aus dem Bereich der Baustoffkunde zeigen schnell die Grenzen von Portalen und Suchmaschinen (vgl. Abb. 2.7) auf.

Die Ursachen solcher Beschränkungen liegen im konzeptionellen Aufbau von HTML begründet und wurden in Abschnitt 2.4 erläutert. Solange Dokumente und darin enthaltene multimediale Komponenten nicht mit strukturierten Beschreibungen bezüglich ihrer Inhalte versehen werden, besitzen Suchdienste keine Möglichkeit, auf den Kontext der Webinhalte zu schließen und semantische Abfragen durchzuführen.

Aufgrund der großen Menge der schon in HTML veröffentlichten Informationen ist es jedoch wenig sinnvoll, diese bisher praktizierte Publikationsform zu ignorieren und lediglich ein neues, strukturierteres Datenformat vorzuschlagen. Besser wäre eine Lösung, welche die

bisher im Web vorhandenen Inhalte mit einbeziehen kann und auch für zukünftige Entwicklungen offen ist. Dies gelingt jedoch nur, wenn für die semantische Beschreibung von Webinhalten ein vom Darstellungsformat unabhängiges Konzept angewandt wird. Dadurch können semantische Datensätze auch nachträglich zu bereits bestehenden Webinhalten hinzugefügt werden und diese damit besser auffindbar gemacht werden.

Vor dem Entwurf solcher Datenstrukturen ist es wichtig zu wissen, in welcher Form bzw. in welcher Strukturierung die Daten vorliegen und welche inhaltlichen Aspekte für fachspezifische Recherchen relevant sind.

3.2.1 Quellen für baustofftechnologische Informationen

Informationen zu Baustoffen und Bauprodukten entstehen in der Regel im Bereich der wissenschaftlichen und industriellen Baustoff-Forschung. Vor dem Hintergrund einer Produkt- oder Verfahrensidee werden in Forschungsarbeiten die Randbedingungen für die Umsetzung der Idee untersucht. Ergebnisse dieser Arbeiten fließen anschließend in die technischen Regelwerke ein und schaffen somit die rechtlichen Voraussetzungen für die Produktentwicklung in der Industrie. Produktspezifisches Wissen wird wiederum in der Produktzulassung sowie weiterer Dokumentation zu den einzelnen Baustoffen festgehalten. In der Planung und Ausführung von Bauprojekten wird auf das dort dokumentierte Wissen schließlich wieder Bezug genommen.

Zwischen den einzelnen Bereichen Baustoff-Forschung, Normung und der Anwendung der Regelwerke in der Planung und Ausführung von Bauprojekten finden jeweils Wissenstransfers statt. Die in der Vergangenheit vorwiegend praktizierte Methode der Wissensvermittlung besteht in der Verwendung konventioneller Printmedien. Wie am Szenario im vorangegangenen Abschnitt dargestellt wurde, ist die Recherche in papiergebundenen Informationsquellen mit einem hohen Zeitaufwand bei häufig vergleichsweise geringer Effizienz verbunden. Zur Vermittlung der Ergebnisse zwischen den einzelnen Bereichen wird aus diesem Grund vermehrt das World Wide Web als Kommunikationsmedium eingesetzt. Neben dem Vorteil kürzerer Informationswege bietet der Einsatz von Informationssystemen im Web gegenüber konventionellen Publikationsmethoden noch eine Reihe weiterer Vorteile:

- Zentrale Verwaltung und leichte Aktualisierbarkeit der Inhalte
- Geringe Kosten, da Druck- und Verteilerkosten entfallen, sowie eine einmalige zentrale Bereitstellung nahezu unbegrenzten Zugriff erlaubt
- Zeit- und ortsunabhängiger Zugang zu den Informationen
- Verbesserte Nutzungsmöglichkeiten der Dokumente durch Volltextrecherche etc.
- Erweiterte Gestaltungsmöglichkeiten für die Präsentation der Inhalte, z.B. durch multimediale Komponenten

Das Informationssystem BiM-Online ist ein Beispiel dafür, den Wissenstransfer aus der Bauforschung mit den Methoden moderner Informationstechnologie zu beschleunigen. Inzwischen hat auch die Baustoffindustrie das World Wide Web zur Vermittlung baustoff- bzw. bauproduktbezogener Informationen an ihre Kunden entdeckt.

Der erste Schritt bei der Entwicklung eines Informationssystems besteht dabei häufig in der Umwandlung bereits in gedruckter Form vorhandener Informationen in ein webfähiges Format. Dem Anwender wird dadurch ein alternativer Zugang zu den Informationen über die Web-Site des Unternehmens ermöglicht. Im weiteren wird dieses Angebot dann durch erläuternde oder ergänzende Informationen ausgebaut. Wächst der Inhalt einer solchen Web-Site über ein gewisses Maß hinaus, ist zur Sicherstellung einer sinnvollen Nutzung eine an die

Bedürfnisse der Anwender angepasste, komplexere Oberfläche mit entsprechenden Funktionalitäten für den Zugriff auf die Inhalte erforderlich.

Typische Inhalte von Webangeboten mit baustoffbezogenen Informationen sind in den einzelnen Bereichen:

- Bauforschung: Forschungsberichte, Fachpublikationen, multimediale Präsentationen
- Normtexte, Produktzulassungen etc., die zumeist von Fachverlagen zur Verfügung gestellt werden,
- Baustoff-Firmen/-Verbände: Datenblätter zu Produkten, multimediale Präsentationen, Software
- Planungsbüros und Bauunternehmen: Projektpräsentationen
- Hochschulen: Lehrinhalte aus Skripten, Vorlesungen und Übungen.

Der Hochschulbereich ist ein weiteres Einsatzgebiet für webbasierende Informationssysteme über Baustoffe. Hier soll der Transfer von Forschungs- und Praxiswissen aus der Hochschule zu den Studenten mit eLearning-Techniken [Schwarte2002] verbessert werden. Ein zum Zeitpunkt der Verfassung dieser Arbeit aktuelles Forschungsprojekt zur webbasierten Baustofflehre im Hochschulbereich wird unter [Koptik2001] beschrieben.

3.2.2 Darstellungsformen für baustofftechnologische Informationen im Web

Baustoffe werden im Allgemeinen anhand ihrer Eigenschaften und ihrer Zusammensetzung beschrieben. Diese Parameter werden in Versuchen ermittelt oder vom Hersteller vorgegeben. Für die Angabe von Baustoffeigenschaften stehen zunächst mehrere Darstellungsformen zur Verfügung:

- Textbeschreibungen (Aussagen, Aufzählungen, Absätze)
- Abbildungen (Fotos, Zeichnungen, Skizzen etc.)
- Diagramme (Darstellung von Relationen, Korrelationen, funktionalen Zusammenhängen)
- Tabellen (Darstellung geordneter Zahlenwerte)
- Mathematische oder chemische Gleichungen

In einigen Bereichen ist die Darstellungsform bestimmter Eigenschaften in den Normen weitgehend vorgeschrieben und kann nicht frei gewählt werden.

Die zunehmende Verwendung digitaler Foto- und Videokameras machen darüber hinaus die Erzeugung und den Umgang mit digitalisierten, multimedialen Daten zur alltäglichen Selbstverständlichkeit. Der Einsatz von digitaler Fotografie bzw. Film macht auch vor der Baustoff-Forschung nicht Halt. Die Nutzung von Digitalkameras und digitalen Videokameras ist inzwischen auch in den Baustofflabors aufgrund der kostengünstigen Dokumentationsmöglichkeiten weit verbreitet. Die oben angeführten Darstellungsformen müssen daher heute um Videos und teilweise interaktive Animationen ergänzt werden. Die letztgenannten Darstellungsformen finden im Bereich der Baustoffe überwiegend zur Veranschaulichung der Versuchsdurchführung, der Herstellungsabläufe sowie für zeitabhängiger Effekte Verwendung.

Alle der hier aufgeführten Darstellungsformen lassen sich in HTML-Dokumenten abbilden. Im WWW werden Grafiken und multimediale Inhalte als separate Dateien abgespeichert und erst im Browser in das HTML-Dokument integriert (s.a. Abschnitt 2.2). Eine Webseite besteht daher zumeist aus mehreren Dateien: Dem eigentlichen HTML-Dokument sowie den über Verknüpfungen im HTML-Code eingebundenen Dateien für Grafiken etc.. Für die einzelnen multimedialen Inhalte haben sich auch entsprechende Dateiformate herausgebildet, die für die rasche Übermittlung der Inhalte über das Web optimiert wurden. So werden

Grafiken und Fotografien aus Speichergründen nicht in textbasierenden Dateiformaten gespeichert, sondern in den komprimierten, binären Formaten GIF und JPEG. Gleiches gilt für Videos (MPEG, Real Media, AVI), Animationen und eingebundene Programme (JAVA), deren Speicherung ebenfalls in bevorzugten Dateiformaten erfolgt.

Rückschlüsse von den einzelnen Dateitypen auf ihre Inhalte sind dabei nicht immer möglich. So werden komplexer aufgebaute Tabellen häufig nicht in HTML codiert, sondern als Grafikdateien in Webseiten eingebunden. Für den menschlichen Betrachter macht dies keinen Unterschied. Aufgrund der binären Codierung der Grafikdateien gehen die in den Tabellen enthaltenen Texte hierbei für die Volltextsuche jedoch verloren.

Dasselbe gilt für chemische und mathematische Gleichungen, die in HTML-Dokumenten abgebildet sind. Werden die Gleichungen in ihrem Aufbau komplexer oder enthalten sie Sonderzeichen außerhalb des ASCII-Standards, dann lässt sich deren Formatierung nicht mehr mittels HTML-Befehlen bewerkstelligen. Die Gleichungen werden dann ebenfalls als Grafiken in die Webseite integriert. Eine Verbesserung bei der Darstellung von Formeln ist allerdings durch die oben erwähnten XML-Anwendungen MathML und CML (vgl. Abschnitt 2.6.3) in den kommenden Jahren zu erwarten.

Webinhalte zu Baustoffen sind daher nicht an bestimmte Dateiformate gebunden und auch nicht auf Textdaten beschränkt. Für die Recherche bedeutet dies, dass ein entsprechender Suchdienst sich bei den zu erfassenden Datenquellen ebenfalls nicht auf Textdaten beschränken darf. Ein solcher Dienst würde weiterhin zusätzliche Angaben zu den Inhalten von Grafikdateien in Textform benötigen, da Angaben zu den Datei-Inhalten aus den binären Dateiformaten selbst nicht ableitbar sind.

Betrachtet man die Verwendung von binären Datenquellen zur Darstellung von technischen Sachverhalten, so stellt man fest, dass diese Informationen transportieren, die häufig für Ingenieure von besonderem Wert sind. Ingenieuren wird oft eine visuell orientierte, anschauliche Denkweise unterstellt. Diese Ansicht wird durch die Tatsache unterstützt, dass in den Fachpublikationen gerade dieser Berufsgruppe sehr häufig Diagramme zur Veranschaulichung komplexer Sachverhalte benutzt werden. Im Bereich der Baustoffe gibt es hierfür genügend Beispiele:

- Abhängigkeiten einzelner Baustoffeigenschaften voneinander, z.B. die Arbeitslinie von Werkstoffen mit Spannungs- /Dehnungsverlauf
- Darstellung der Abhängigkeiten von unterschiedlichen Versuchsparametern und anderen Einflussgrößen

Diagramme dienen allgemein zur Darstellung von (funktionalen) Abhängigkeiten zwischen mehreren Kenngrößen, die im Diagramm verschiedene Zahlenwerte annehmen können. Je nach Art und Anzahl dieser Kenngrößen sowie der Art der Abhängigkeit existieren verschiedene Diagrammtypen. Die wichtigsten Diagrammtypen für die Naturwissenschaft und Technik sind Koordinatendiagramm, Balkendiagramm und Kreisdiagramm [Duden1997]. Jeder Diagrammtyp kann durch vielfältige Variationen im Layout, in der Achsenskalierung oder der Zahl der verknüpften Größen (Dimension) der Aufgabe angepasst werden (Abb. 3.3).

Die gewählte Form der grafischen Darstellung eines Sachverhalts hängt dabei von den Aussagezielen des Autors ab und ist daher sehr individuell. Während ein Mensch keine Schwierigkeiten hat, eine Aussage trotz unterschiedlicher Darstellung wiederzuerkennen (Abb. 3.4), ist dies für ein Computerprogramm noch nicht immer möglich. Auch für Fotografien und bewegte Abbildungen gibt es noch keine Software, die den Inhalt von Abbildungen sicher „erkennen“ und thematisch zuordnen könnte.

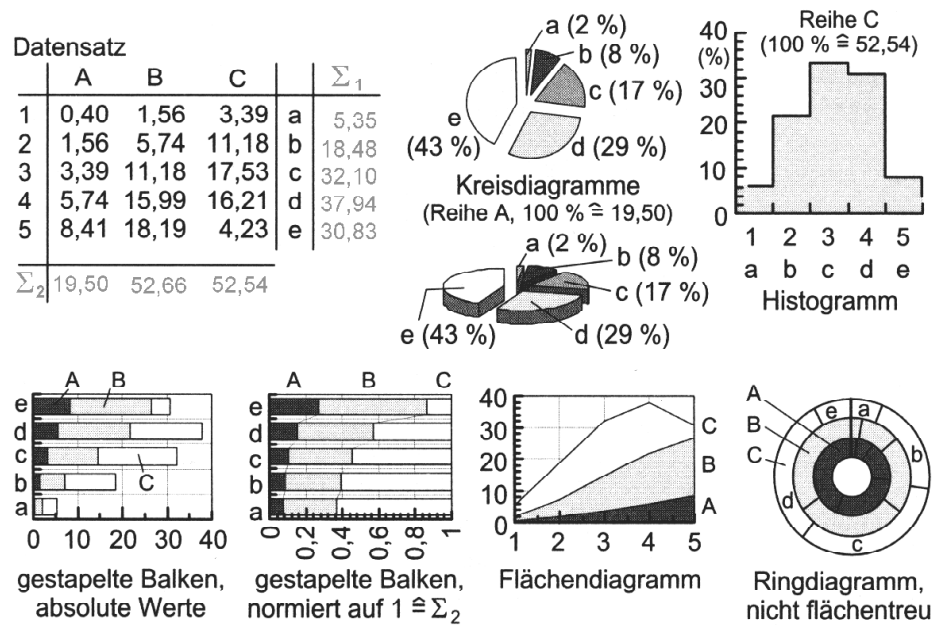


Abb. 3.3 Unterschiedliche Variationen desselben Diagrammtyps [vgl. Duden1997]

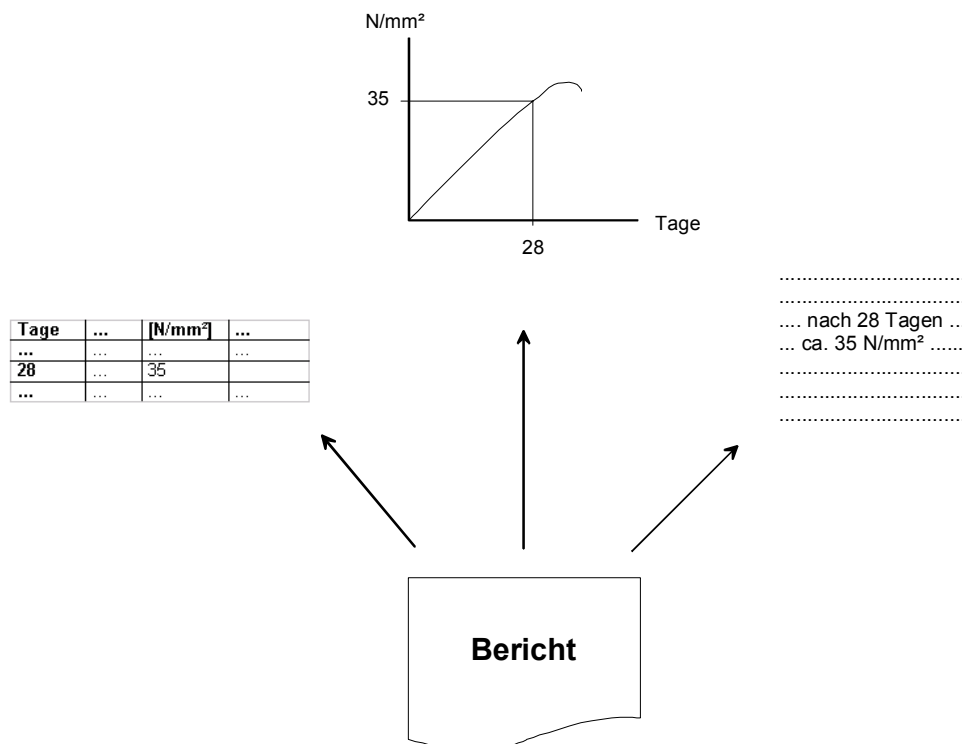


Abb. 3.4 Dieselbe Aussage kann sowohl in unterschiedlichen Datenformaten als auch in unterschiedlichen Formatierungen abgebildet werden. Wie ein Sachverhalt dargestellt wird, hängt sehr von den Aussagezielen des Autors ab.

Volltextsuchmaschinen sind heute die einzigen Systeme im WWW, die automatisch Webinhalte erschließen und über Stichwortzuweisungen einordnen. Sie können jedoch lediglich textbasierende Daten erfassen. HTML-Dokumente enthalten nur wenig Informationen über die eingebundenen Grafiken, multimedialen Elemente oder Java-Programme in ihrem Quellcode. Lediglich der Dateinamen und evtl. ein kurzer Text zum Inhalt kann als Informationsquelle für Volltextrecherchen herangezogen werden. Grafische oder multimediale Inhalte bleiben Suchmaschinen jedoch verborgen. Dabei bilden Diagramme in der Baustofftechnologie oft mehrdimensionale Zusammenhänge ab, die sie in einer für den menschlichen Betrachter kompakten Art und Weise zusammenfassen (vgl. Abschnitt 6.2.3).

Ohne zusätzliche Beschreibungsmöglichkeiten, die von Suchdiensten zur semantischen Erschließung genutzt werden können und die der inhaltlichen Komplexität solcher multimedialer Informationsquellen entsprechen, wird kaum jemand in der Lage sein, den dort verborgenen Informationsreichtum in einer wirtschaftlichen Weise nutzen zu können. Die wachsende Bedeutung multimedialer Datenformate zur Speicherung von technischen Informationen verschärft dabei den Bedarf nach einem entsprechenden webfähigen Beschreibungskonzept. Denn es ist naheliegend, für die Verbreitung digitaler Informationsquellen die Infrastruktur des WWW zu nutzen.

3.3 Anforderungen an Datenmodelle für ingenieurgerechte Recherchen

Webinhalte zu baustofftechnologischen Themen werden im WWW auf einer Vielzahl von Web-Sites gespeichert, deren Inhalte, Gestaltung und technologische Realisierung ganz bei dem jeweiligen Anbieter liegt. Eine zentrale Verwaltung solcher Internetinhalte gibt es zunächst nicht. Diese radikale Dezentralisierung stellt für die meisten Internetnutzer allerdings ein großes Problem dar. Der typische „Web-Surfer“ hat den Anspruch, dass ihm nach der Einwahl in das Netz zunächst sehr wohl ein zentralisierter Dienst bereitgestellt wird. Für diese Internet-Zentralen hat sich der Begriff „Portal“ eingebürgert (vgl. Abschnitt 2.7.1). Dieses offensichtlich zwangsläufige Nutzerverhalten ist der Grundstein des Erfolgs bekannter Online-Dienstleister wie Yahoo [Yahoo2002], Lycos [Lycos2002], Web.de [Web.de2002], AOL [AOL2001], T-Online [T-Online2002] usw.. Derartige Dienste richten ihre Angebote auf das breite Publikum aus, dessen Informationsbedürfnis hierdurch inzwischen in sehr hoher Qualität befriedigt werden kann.

Völlig anders stellt sich die Situation hingegen dar, wenn sehr spezielle Informationen über das Internet transportiert und verbreitet werden sollen. Zwar existieren im WWW inzwischen zahlreiche qualitativ hochwertige Informationsangebote zu speziellen Themengebieten innerhalb des Bauingenieurwesens, wie z. B. baunetz [baunetz], netinform [netinform] und mybau.com [mybau2002]. Diese können jedoch nur jeweils einen bestimmten, redaktionell ausgewählten Bereich an Online-Informationsquellen abdecken. Der höchstmögliche Nutzwert all dieser Angebote würde erst dann erreicht, wenn sie zum einen intensiv untereinander vernetzt wären und zum anderen eine die eigenen Systemgrenzen übergreifende und umfassende Recherche ermöglichen würden.

Aufgrund des Umfangs und des raschen Wachstums von Webinhalten ist eine redaktionelle Erfassung von Informationsquellen langfristig nicht sinnvoll. Die erfassten Daten zu den in den Informationsquellen enthaltenen Informationen wären häufig nicht mehr aktuell, da es organisatorisch fast unmöglich wäre, Änderungen in derselben Geschwindigkeit manuell zu erfassen und redaktionell aufzuarbeiten. Stets würde der Informationsstand eines solchen Recherchesystems weit hinter der wirklichen Situation im WWW zurückbleiben. Angaben zu baustoffspezifischen Webangeboten bzw. zu deren Inhalten müssen daher automatisch verarbeitet werden können. Ein Datenmodell zur Beschreibung von baustofftechnologischen

Webinhalten sollte daher möglichst alle zur Recherche und semantischen Einordnung benötigten Merkmale in maschinenlesbarer Form bereitstellen.

Die Inhalte, die sich heute im Web befinden, sind in der Mehrheit innerhalb von HTML-Dokumenten publiziert. Diese besitzen, wie in Abschnitt 2.4 dargestellt, eine nur gering ausgeprägte inhaltliche Strukturierung, die von Suchdiensten zu über die Volltextrecherche hinausgehenden Recherchemöglichkeiten genutzt werden können. Für Konzepte zur maschinenlesbaren Beschreibung von Webinhalten bedeutet dies, dass diese sich daher auch nachträglich in bzw. zu den bestehenden Informationsangeboten in HTML ergänzen lassen müssen, ohne dass deren Benutzung eingeschränkt wird. Im besten Fall entsteht dem Autor durch die Ergänzung seiner Webinhalte mit standardisierten Beschreibungen ein Zusatznutzen, ohne dass er dadurch Nachteile für die Nutzbarkeit seines bisherigen Informationsangebots in Kauf nehmen muss.

Die Beschreibungen von Webinhalten mit semantischen Datensätzen, konform dem in dieser Arbeit vorgestellten Datenmodell met@BiM, müssen daher wie eine homogene Datenschicht für zukünftige Suchdienste über den heterogenen Webinhalten funktionieren. Auf diese Weise kann die Darstellung der Informationen individuell bleiben, während die für Suchdienste relevante semantische Struktur in einer standardisierten Form erfolgen kann. In der Folge bedeutet dies jedoch auch, dass zunächst viele Informationen doppelt publiziert werden müssen - einmal im Informationsobjekt (Webseite, Diagramm etc.) selbst, ein anderes mal in dessen maschinenlesbarer Beschreibung.

Dieser Aufwand, den die Erzeugung der beschreibenden Datensätze von Informationsanbietern erfordert, muss dabei in einem angemessenen Verhältnis zum Nutzwert der Informationen stehen. Ein Autor kann zwar bei der Beschreibung der Webinhalte von entsprechenden Softwareprogrammen gezielt unterstützt werden, so dass z.B. mehrfache oder für die Anwender irrelevante Eingaben vermieden werden können. Der Aufwand wird von Anbietern von Informationsquellen jedoch nur dann betrieben werden, wenn die maschinenlesbaren Beschreibungen ihrer Webinhalte auch einen deutlich höheren Nutzwert und damit Vorteile für die Anwender aus der Baupraxis einbringen.

Für eine professionelle Nutzung des WWW als Informationsmedium muss dieses dabei Anforderungen gerecht werden, die über die aktuellen Möglichkeiten der Volltextrecherche im Web hinausgehen. Recherchesystemen müssen hierfür detailliertere Angaben zu den Inhalten von Webdokumenten zur Verfügung stehen. Die Darstellungsform, welche zur Präsentation der Informationen vom Anbieter gewählt wurde, darf hierbei auf deren Auffindbarkeit keinen Einfluss haben. Das bedeutet im Umkehrschluss, dass ein entsprechendes Beschreibungskonzept sowohl auf Webressourcen in Textform, d.h. auf einzelne Webseiten, Sammlungen von Webseiten oder auch nur Textabschnitte, als auch auf binäre Webressourcen wie Java-Programmen und Multimediapräsentationen anwendbar sein muss. Es ist wichtig zu betonen, dass die Freiheit des Autors, die Präsentation seiner Aussagen so zu gestalten, wie er es für sinnvoll erachtet, durch das Datenmodell keinesfalls eingeschränkt werden darf. Würde Autoren z.B. die Formatierung ihrer Diagramme, Grafiken oder Tabellen vorgeschrieben, um den Anforderungen eines standardisierten Datenmodells zu genügen, wäre der Erfolg des Konzepts von vornherein als sehr fragwürdig zu beurteilen.

In technischen Wissensgebieten ist auch die Möglichkeit numerischer Abfragen unter Verwendung von Vergleichsoperatoren erforderlich. Da Maßeinheiten und Dimensionen, welche den Werten zugeordnet werden, das Ergebnis solcher Abfragen beeinflussen, muss ein Datenmodell zur Beschreibung von Baustoff-Informationen diese ebenfalls berücksichtigen.

Zur Ausfilterung irrelevanter Treffer ist es zudem sinnvoll, Angaben zum Kontext in welchem Webinhalte stehen, maschinenlesbar zum Bestandteil beschreibender Datensätze zu machen. Über die Zuordnung von Webinhalten zu einem oder mehreren Kontexten lassen sich auch

logische Verknüpfungen zu weiteren Informationsquellen ableiten. Findet z.B. ein Planer den benötigten Produktkatalog einer Baustoff-Firma im WWW, gelingt es ihm bei einer gut strukturierten Web-Site zwar sehr schnell, wesentliche bauphysikalische und mechanische Eigenschaften in Erfahrung zu bringen. Jedoch schon der Wunsch, Preise von vergleichbaren Produkten über das Web zu finden, erfordert von ihm einige Stunden der Recherche.

Da sich Wissensgebiete stets weiterentwickeln und neue Teilgebiete entstehen, muss ein entsprechendes Datenmodell diese Anforderung ebenfalls berücksichtigen. Soll das Konzept zukunftsfähig sein, ist die Erweiterbarkeit des Datenmodells in jeder Richtung ein zentrales Erfordernis. Die Flexibilität des Datenmodells muss sich auch auf zukünftige Typen von Informationsobjekten erstrecken, die sich ebenfalls durch das Datenmodell beschreiben lassen müssen. Neue Elemente müssen sich in das Modell ohne grundsätzliche Änderungen an bereits vorhandenen Beschreibungen im Web integrieren lassen, da Anwendern ein ständiges Überarbeiten bereits vorhandener Beschreibungen nicht zugemutet werden darf. Für die breite Akzeptanz bei Informationsanbietern, die für den Erfolg eines solchen Datenmodells unabdingbar ist, wäre dieser zusätzliche Aufwand sehr kontraproduktiv. Einmal verfasste Beschreibungen müssen, sofern sie nicht inhaltlich überarbeitet werden müssen, in ihrer Struktur stabil bleiben.

Diese z.T. widersprüchlichen Anforderungen, dass das Datenmodell auf der einen Seite bzgl. der Formatierung und Erweiterbarkeit flexibel und universell anwendbar sein muss, auf der anderen Seite jedoch sehr exakte Recherchen erlauben soll, sind mit klassischen Datenmodellen im WWW nicht zu verwirklichen. Hierfür ist die Einführung eines besonderen Datentyps notwendig, der im folgenden Kapitel beschrieben wird: Den Metadaten.

4 Metadatenmodelle zur Beschreibung von Baustoff-Informationen

Um Recherchen in sehr großen Informationsnetzen effizient durchführen zu können, werden Inhalte von Webseiten von Suchmaschinen zunächst komprimiert und in eine einheitliche Datenstruktur gebracht. In aktuellen Volltextsuchmaschinen wird diese Aufgabe von dem in Abschnitt 2.7.2 beschriebenen Crawler erbracht, welcher die schwach strukturierten Inhalte der Webseiten in eine homogene Indexdatenbank überführt. Bezüglich der inhaltlichen Aussagen in Webseiten lassen sich aus der automatischen Analyse des HTML-Quellcodes mit computerlinguistischen Methoden jedoch nur unscharfe Zusammenhänge ableiten [Fühles-Ubach1997].

Ein anderer Weg der Datenerfassung besteht darin, schon auf Seiten der Informationsanbieter eine homogen strukturierte Beschreibung der publizierten Webinhalte bereitzustellen, die dann automatisch von Suchdiensten erfasst und interpretiert werden kann. Dabei muss nicht unbedingt die Struktur oder gar Darstellung der Webinhalte selbst verändert werden. Ein solches Konzept würde eine Überarbeitung der Vielzahl bereits bestehender Informationsangebote voraussetzen, was die Akzeptanz stark in Frage stellen würde. Es genügt, eine Datenschicht einzuführen, die sich „über“ die Inhalte legt und den Suchmaschinen das zur Verfügung stellt, was diese zur Erfassung benötigen: Regelmäßig strukturierte, maschinenlesbare Datensätze, welche die kennzeichnenden Merkmale und die Adresse der eigentlichen Webinhalte beschreiben. Dieses Prinzip der Einführung einer zusätzlichen Datenschicht ist nicht neu, sondern jedem Benutzer von Bibliotheken bereits bekannt: Auch in Bibliotheken erfolgt die Recherche nach Literaturtiteln nicht in den Bücher selbst, sondern über Karteikarten bzw. heute über Meta-Datensätze in den Datenbanken der Bibliothek. Ein solches Vorgehen der indirekten Recherche anhand von Metadaten lässt sich auch zur Verbesserung der Auffindbarkeit von Baustoff-Informationen im Web anwenden.

Dafür ist es jedoch zunächst wichtig, die relevanten Merkmale von betontechnologischen Webinhalten in einem Metadatenchema zu definieren. Standards für Metadatenmodelle, die zur allgemeinen Beschreibung von Webinhalten angewandt werden können, existieren bereits. Mit dem Dublin Core-Metadatenstandard (vgl. Abschnitt 4.2) wird in dieser Arbeit solch ein Modell vorgestellt, das inzwischen in vielen auf Metadaten basierenden Informationssystemen Anwendung findet. Viele Elemente dieses Standards lassen sich auch zur Beschreibung betontechnologischer Inhalte benutzen. Für die angestrebten Recherchemöglichkeiten müssen jedoch darüber hinaus noch weitere, und zwar fachspezifische Elementstrukturen ergänzt werden. Dieser Schritt wird im Abschnitt 4.3 zum Entwurf des met@BiM-Datenmodells innerhalb dieses Kapitels ausführlicher erläutert.

Da die Beschreibung der Metadatenmodelle unabhängig von einer bestimmten Syntax erfolgen kann, werden die Elemente und Strukturen von met@BiM zunächst in Abbildungen und Tabellen allgemein verständlich dargestellt. Die Anwendung zur Beschreibung konkreter, betontechnologischer Webinhalte wird im sechsten Kapitel anhand von Beispielen aufgezeigt.

4.1 Grundlagen zu Metadaten

Die Idee der Anwendung des Metadatenkonzepts auf betontechnologische Webinhalte stellt den zentralen Ansatz dieser Arbeit dar, mit dem die eingangs erläuterten Probleme der Recherchierbarkeit ingenieurwissenschaftlicher Webinhalte behoben werden sollen. Für ein Verständnis der Überlegungen beim Entwurf und der Anwendung von Metadatenstrukturen ist jedoch zunächst ein Verständnis des Metadatenkonzepts selbst erforderlich. In den folgenden Abschnitten wird der abstrakte Begriff der Metadaten in seinen vielfältigen Ausprägungen erläutert und der Aufbau, der Entwurf sowie die Anwendung dieses Konzepts anhand von Beispielen beschrieben.

4.1.1 Was sind Metadaten

Wörtlich übersetzt, handelt es sich bei Metadaten um „Daten über Daten“. Tatsächlich definiert der Begriff „Metadaten“ nicht einen neuen Datentyp, sondern vielmehr strukturierte Beschreibungen von Informationsobjekten in Textform. Metadaten lassen sich sowohl zur Beschreibung von Webinhalten als auch zur Beschreibung von physikalischen oder abstrakten Objekten der Realität heranziehen, die nicht über das Web zugänglich sind. Metadaten machen daher sowohl die Informationsobjekte des Internets als auch reale Objekte für Suchmaschinen erfassbar.

Alle Informationsobjekte besitzen drei Hauptmerkmale, die anhand von Metadaten abgebildet werden können [Gilliland2000]:

- Inhalt – bezieht sich auf den Inhalt des Objekts. Dabei handelt es sich um ein inhärentes Merkmal eines Informationsobjekts.
- Kontext – beschreibt die Umstände der Publikation: Wer ist der Verfasser? Mit welcher Absicht, wo und wie wurde das Objekt erstellt?
- Struktur – bezieht sich auf die semantischen Beziehungen innerhalb eines einzelnen Informationsobjekts bzw. zwischen Informationsobjekten.

Um die inhaltlichen Aspekte eines Informationsobjekts sinnvoll beschreiben zu können, sind Fachkenntnisse der jeweiligen Wissensdomäne unverzichtbar. Datenschemata werden von den Anwendergruppen dann angenommen, wenn ihr Aufbau möglichst einfach und der anzugebende Inhalt in den Elementen, z.B. durch deren Bezeichnung, offensichtlich ist. Auch die hierarchische Einstufung von inhaltsbezogenen Metadatenelementen erfordert Kenntnisse über deren Bedeutung sowie der jeweiligen Zusammenhänge. Um zu vermeiden, dass die in den jeweiligen Fachgebieten entwickelten Schemata aufgrund von semantischen Doppelbelegungen inkompatibel zueinander werden, muss die Entwicklung von Metadatenstandards koordiniert werden und auf mehreren Ebenen erfolgen.

Während für die inhaltlich unabhängigen Metadatenelemente allgemeine, d.h. fachübergreifende Standards definiert werden können, muss die Definition von Elementen zur Beschreibung der Inhalte eines Informationsobjekts in Gremien mit fachspezifischer Ausrichtung erfolgen. Auf diese Weise können dort fachspezifische Metadatenstandards zur Beschreibung der charakteristischen inhaltlichen Merkmale von Informationsobjekten entstehen, die den allgemeinen Standards untergeordnet sind.

In den Kulturwissenschaften sowie im Bibliothekswesen werden schon seit den 60ern Jahren standardisierte Klassifikationen verwendet, um Literatur themenbezogen zu ordnen, zu beschreiben oder auffindbar zu machen. Beispiele für solche Klassifikationen sind auf internationaler Ebene im Bibliothekswesen der MARC (Machine-Readable-Cataloging Format) und der LCSH (Library of Congress Subject Headings). Auch für das Bauwesen wurden solche Themenkataloge zur Ordnung von Informationsquellen wie Fachbücher, Planungsunterlagen etc. in Fachbibliotheken und Unternehmen entworfen. In einer Veröffentlichung [Büchin1973] von 1973 werden z.B. Kategorien für Quellen des Bauwesens und in [Müller1966] von 1966 für Quellen zum Thema Beton vorgeschlagen. Da Klassifikationen ursprünglich für die Aufstellung von Büchern gedacht waren, gilt auch hier das Prinzip der Monohierarchie, d. h. eine Klasse darf nur genau einer anderen Klasse untergeordnet sein, weil ein einzelnes Buch auch nur an einem Ort stehen kann. Wie bei den Datenbanken mit hierarchischer Struktur entsteht auch bei Klassifikationen das Problem der schwierigen Erweiterbarkeit und der Unflexibilität gegenüber nachträglichen Veränderungen. Analytische Klassifikationen weisen daher als Hilfskonstruktion häufig eine Klasse "Sonstiges" zur Erweiterung oder auch nur "Anderes" auf [Burkart1990]. Ein weiteres Problem von Monohierarchien zeigt sich in ihrer praktischen Anwendung: Häufig werden

Quellen in Klassen eingeordnet, unter der sie von den Nutzern nicht erwartet werden. Oder es handelt sich um interdisziplinäre Quellen, die in mehreren Klassen eingeordnet werden könnten.

Metadatenbeschreibungen der Eigenschaften von Quellen entschärfen diese Probleme, in dem sie mit der Möglichkeit, einem Informationsobjekt mehrere klassifizierende Elemente zuzuordnen, dessen Zugehörigkeit zu mehreren Klassen erlauben. Bezüglich der Recherche in elektronischen Datenbanken ist es für die Auffindbarkeit von Informationsobjekten wichtig, ihnen möglichst alle sinnvollen Kontexte zuzuordnen, in die die Quelle eingeordnet werden kann. Das definierte Vokabular der o.g. Klassifikationen kann jedoch dabei helfen, die Angabe des Kontexts in den Metadaten einheitlich zu gestalten und damit sprachliche Unschärfen über unterschiedliche Systeme hinweg zu vermeiden.

Mit der Nutzung von Computern zur Recherche hat auch die Struktur von digitalen Informationsquellen an Bedeutung gewonnen. Die Entwickler von auf Metadaten basierenden Suchsystemen erkannten, dass je feiner der strukturelle Aufbau eines Informationsobjekts ist, desto besser lässt es sich für die Suche, Manipulation und für Bezüge zu anderen Informationsobjekten heranziehen. Man spricht in diesem Zusammenhang auch von einer höheren Qualität eines Informationsobjekts. Die Erfassung, Dokumentation und eventuelle Nachbesserung dieser Struktur erfordert jedoch weitere, spezifischere Metadaten. Diese lassen sich in die in Tabelle 4.1 beschriebenen Kategorien von Metadatentypen gruppieren.

Tab. 4.1 Metadatentypen und ihre Funktion [Gilliland2000]

Typ	Definition	Beispiele
Administrativ	Metadaten zur Verwaltung und dem Management von Informationsressourcen	- Copyright etc. - Dokumentation der Nutzungsrechte - Speicherort - Versionsinformationen
Beschreibend	Metadaten zur Beschreibung oder der Identifizierung von Informationsressourcen	- Katalogisierende Informationen - Fachspezifische Indexe - Infos für die Recherchierbarkeit (Stichworte etc.) - Hyperlink-Beziehungen zwischen Datensätzen - Anmerkungen von Nutzern
Konservierend	Metadaten im Zusammenhang mit der Erhaltung von Informationsressourcen	- Dokumentation des physischen Zustands von Quellen - Dokumentation der Erhaltungsmaßnahmen
Die Nutzung beschreibend	Metadaten, welche das Anwendungsgebiet beschreiben	- Nutzung und Nutzerverfolgung - Informationen über die Weiterverarbeitung der Inhalte
Technisch	Metadaten zur Beschreibung der Systemfunktionalitäten	- Hardware- und Softwaredokumentation - Digitalisierungsinformationen (Formate, Komprimierungsraten, Zoom) - Authentifizierungsinformationen (Passworte, Kryptograph. Schlüssel)

4.1.2 Vorteile der Verwendung von Metadaten

Wurde bis Mitte der 90er Jahre der Ausdruck „Metadaten“ vorwiegend in den Anwendungsgebieten der geographischen Datenverarbeitung und im Datenmanagement verwendet, erlangt dieses Konzept in der heutigen Zeit durch die automatisierte Informationsverarbeitung eine stärkere Bedeutung. Vor allem zur Verbesserung der Recherchefunktionen in Informationssystemen mit großen und verteilten Datenmengen finden Metadatenmodelle Anwendung. Die Auffindbarkeit von Informationen in großen Netzen, zu denen auch das World Wide Web gehört, ist zentrales Kriterium der Nutzbarkeit sowie der Qualität der dort gespeicherten Informationen.

Der Umweg bei der Recherche über die Metadaten birgt für das Auffinden von typischen Ingenieurinhalten im Web einige Vorteile:

Unabhängigkeit vom ursprünglichen Repräsentationsformat: Den größten Vorteil stellt die gewonnene Unabhängigkeit vom ursprünglichen Repräsentationsformat dar. Folglich ist es für die Erfassbarkeit von Webinhalten unwichtig, ob die originäre Information in einem bestimmten Dateiformat oder einer bestimmten Darstellungsform als Grafik, Textabschnitt oder Tabelle vorliegt. Damit werden auch heterogene, in verteilten Systemen verwaltete Informationen über einen zentralen Dienst recherchierbar. Lokale Speicherstrukturen, strukturelle Merkmale wie die inhaltliche Gliederung und Darstellungsmerkmale bleiben unverändert. Metadaten ermöglichen daher eine sehr schonende Art der Aufbereitung von Informationsobjekten für die Integration in andere Systeme und eine verbesserte Recherchierbarkeit. Umformatierung der Informationsressourcen selbst in neue, unflexible Schemata sind sonst häufig mit Informationsverlusten verbunden.

Multimediale Quellen: Auf diese Weise können auch multimediale Informationsquellen über textbasierte Suchmaschinen recherchiert werden. Somit werden Diagramme und Videos sowie andere Quellen in binären Dateiformaten auf vertraute Weise recherchierbar.

Datentypen: Über bestimmte Metadatenelemente lassen sich den Webinhalten Datentypen zuordnen. Rechercheprogramme können diese Angaben auswerten und z.B. bei numerischen Abfragen nur entsprechende Quellen berücksichtigen. Auch die Zuordnung von physikalischen Einheiten ist mit entsprechenden Angaben in Metadatenelementen möglich. Dies erlaubt im technischen Bereich exaktere Recherchen.

Kontext: Metadaten erlauben für unterschiedliche Benutzerperspektiven die Bereitstellung von verschiedenen intellektuellen Zugriffswegen auf Informationsobjekte. Damit kann der Quelle ein Kontext zugeordnet werden, d.h. sie kann semantisch beschrieben werden. Dies kann auch durch Einordnung von Informationsressourcen im Rahmen einer Ontologie erfolgen.

Authentizität: Über Metadatenelemente können Angaben über Urheber, Herkunft und Vollständigkeit von Informationsobjekten definiert werden.

Die Erzeugung von Metadaten ist jedoch immer mit einem zusätzlichen Aufwand verbunden. Zwar können bestimmte Metadaten, vor allem zu formalen Merkmalen der Informationsobjekte, auch von Programmen automatisch generiert werden. Die Erzeugung von Metadaten zu den inhaltlichen Aspekten von Informationsquellen ist jedoch zumeist nur manuell möglich.

4.1.3 Überlegungen zum Umfang und der Komplexität der Datenstrukturen

Der Entwurf von Metadatenstrukturen ist geprägt von der Abwägung zwischen dem Aufwand der Metadatenerzeugung und dem benötigten Detaillierungsgrad der Beschreibung. Für die Auffindbarkeit von Informationsressourcen im Web ist es nicht notwendig, alle inhaltlichen

Aspekte vollständig in den Metadaten zu repräsentieren. Es kommt vielmehr darauf an, sich intensiv mit den unterschiedlichen Benutzerperspektiven auseinanderzusetzen. Metadaten-Schemata müssen daher zweckorientiert entworfen werden. Werden Metadaten hauptsächlich zum Auffinden von Informationen genutzt, z.B. durch Abfragen mit komplexen Kontextangaben, dann müssen die inhaltlichen Zusammenhänge auch entsprechend detailliert abgebildet werden. Sollen hingegen die Metadaten vorwiegend technische Formatangaben zu Informationsressourcen bereitstellen, wie dies z.B. bei Systemen zur Verwaltung von Fotos oder Audiodateien häufig der Fall ist, liegt der Schwerpunkt auf dieser Informationsgruppe. Nur so können die maßgeblichen Merkmale identifiziert werden, unter denen die Informationsressource später gesucht wird. Mit der Reduktion der Metadatenätze auf wenige wesentliche Elemente reduziert sich auch der Eingabeaufwand für Metadaten erheblich.

Da die Meta-Informationen häufig über Benutzereingaben erfolgen, sollte es das Ziel sein, ein ausgewogenes Verhältnis zwischen der Detailliertheit der Beschreibungen und den für die zu erwartenden Abfragen erforderlichen Angaben zu schaffen.

Dabei ist zu berücksichtigen, dass mit zunehmendem Detaillierungsgrad auch der administrative Aufwand zur Verwaltung der dadurch größeren Informationsmenge stark ansteigt. Der Entwurf eines Metadatenschemas stellt damit einen Optimierungsprozess dar zwischen dem Anspruch an möglichst gezielte, individuelle Abfragemöglichkeiten, die sich über entsprechend detaillierte Metadatenstrukturen realisieren lassen, und dem Aufwand zur Erstellung und Pflege der entsprechenden Strukturen. Hilfreich ist es in diesem Zusammenhang auch, wenn die Bezeichnungen der Metadatenelemente selbsterklärend gewählt werden, so dass die Webautoren die zugehörigen Inhalte assoziieren können.

Da Webinhalte auch mehrere Beschreibungen gleichzeitig besitzen können, besteht zudem stets die Möglichkeit, Metadatenelemente nachträglich hinzuzufügen oder einzelne Elemente durch weitere Metadaten hierarchisch zu verfeinern. Um allerdings Redundanzen und Widersprüche innerhalb eines Schemas zu vermeiden, müssen vor dessen Entwurf einige Fragen sorgfältig untersucht werden:

- Welche bereits bestehenden Metadaten schemata können angewandt werden, um den Anforderungen der Nutzer, der geplanten Systemfunktionalitäten und der Autoren gerecht zu werden?
- Welche Metadatenmerkmale sind für diese Anforderungen unabdingbar, wie detailliert müssen die einzelnen Metadaten typen strukturiert sein?
- Werden die benutzten Metadatenstandards auch in den jeweils aktuellsten Versionen verwendet?

Besonders dann, wenn die zu beschreibenden Informationsressourcen zukünftig in das WWW oder in andere Systeme integriert werden sollen, ist es wichtig, beim Entwurf neuer Metadatenelemente schon vorhandene Standards zu berücksichtigen. Gerade für die allgemeinen Merkmale von Webinhalten existieren bereits einige Metadatenstandards, wie beispielsweise der Dublin Core-Metadatenstandard. Dieser in Abschnitt 4.2 beschriebene Standard genießt eine wachsende Popularität bei Informationsanbietern und Suchmaschinenbetreibern im Web, so dass er, wenn immer möglich, gegenüber individuellen Entwürfen bevorzugt werden sollte.

4.1.4 Speicherung von Metadaten

Metadaten werden häufig zusammen mit der Informationsressource abgespeichert. In HTML-Dokumenten erfolgt die Zuordnung von Metadatenelementen wie „Keywords“ über besondere HTML-Tags, die in den Kopfdatenbereich der Webseite integriert werden können.

Dasselbe Prinzip lässt sich auch auf einige Dateiformate für Grafiken anwenden. So lassen sich mit Hilfe entsprechender Editoren auch Metadatenelemente in bestimmte Bereiche des binären Codes von PNG (Portable Network Graphic format) und JPEG (Joint Photographic Experts Group) Dateien integrieren, um beispielsweise digitale Fotos mit Meta-Informationen zu versehen [Lafon2000]. Diese bleiben jedoch dem menschlichen Betrachter der Informationsressourcen normalerweise verborgen. Für die maschinelle Weiterverarbeitung und Auffindbarkeit der Inhalte über mehrere Systeme hinweg ist die Speicherung der Metadaten direkt bei der Informationsressource von Vorteil. Ändert sich der Speicherort der Quelle, so ist auf diese Weise sichergestellt, dass die zugehörigen Metadaten nicht verloren gehen.

Der Nachteil dieser verteilten Speicherung der Metadaten besteht jedoch im höheren Verwaltungsaufwand. Um die aktuellen Metadaten eines Informationssystems zu ermitteln, muss ein Programm zunächst auf alle Informationsquellen zugreifen und die darin gespeicherten Metadaten erfassen. Dieses Prinzip wird auch von den in Abschnitt 2.7.2 beschriebenen Suchmaschinen angewandt. Da der Prozess der Metadatenerfassung verhältnismäßig viel Zeit benötigt und vor allem bei parallelen Anfragen eine hohe Netzbelastung verursacht, nimmt man größere Aktualisierungsintervalle in Kauf. Das heißt, die beim Erfassungsvorgang gesammelten Daten werden, zusammen mit der Information über den Speicherort der zugehörigen Informationsressource, in einem zentralen Zwischenspeicher abgelegt, dem sogenannten Repository.

Eine weitere Möglichkeit der Speicherung von Metadaten besteht in der Trennung der Metadaten von den Informationsobjekten. Metadaten sind bzgl. ihres Speicherorts nicht an die zugehörige Informationsressource gebunden. In Fällen, in denen der Metadatenerzeuger nicht gleichzeitig der Autor von Webinhalten ist und er daher gar nicht die Möglichkeit hat, die Informationsressourcen direkt mit Metadaten anzureichern, müssen diese separat gespeichert und mit Hyperlinks oder ähnlichen Zeigern auf die Adresse der Informationsressource ausgestattet werden. Auf diese Weise ist es auch Dritten, wie z.B. Fachverlagen oder Fachinformationsdiensten, möglich, Metadatensammlungen zu spezifischen Inhalten im Web aufzubauen und über ausgefeilte Recherchefunktionen dem Fachpublikum zur Verfügung zu stellen, ohne Schreibrechte für diese Webinhalte besitzen zu müssen. Diese Möglichkeit eröffnet einen Ansatzpunkt für inhaltliche Verknüpfungen zwischen Informationssystemen.

Die Speicherung externer Metadaten kann z.B. als ASCII-Text in entsprechenden Dateien erfolgen oder, wie bei BiM-Online (Abb. 4.1), auch in einem relationalen Datenbanksystem. Auch bei dieser Entscheidung sollte man die Erfassung der Metadaten durch Suchmaschinen berücksichtigen. Diese lässt sich bei Verwendung eines Datenbanksystems auf konventionellen Wegen nur unzureichend realisieren.

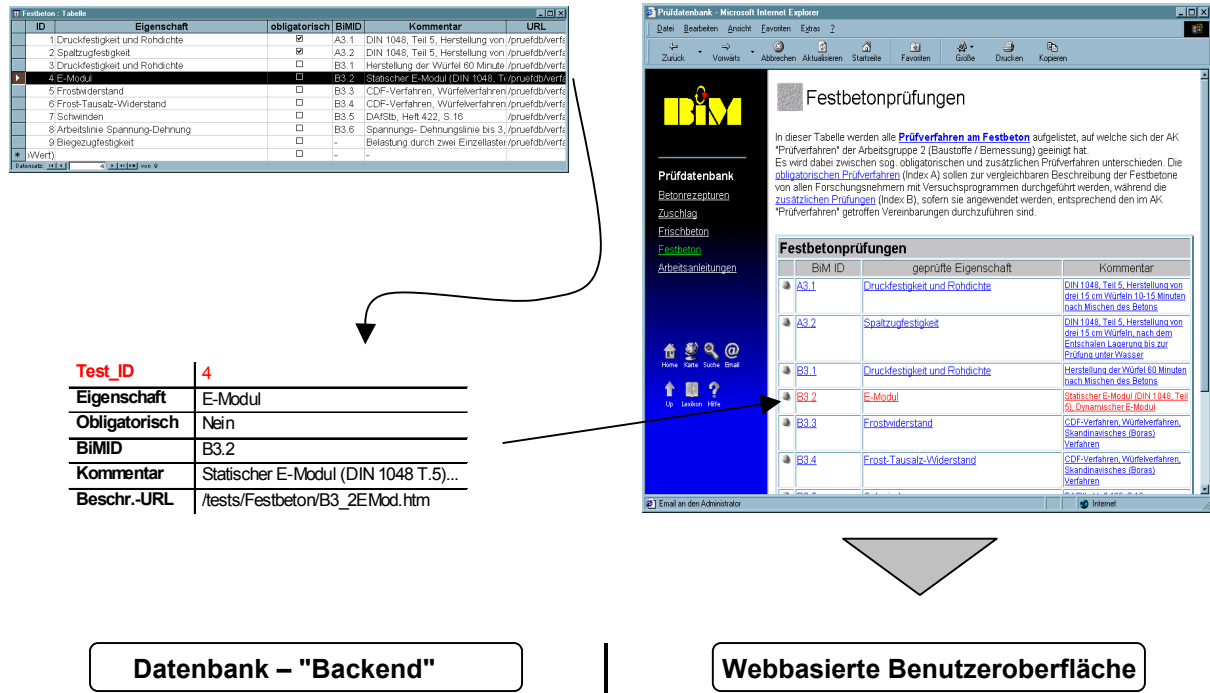


Abb. 4.1 In BiM-Online werden die Metadatensätze im Datenbank-Backend verwaltet. Die Speicherung erfolgt in relationalen Tabellen [Schreyer2001a].

4.2 Metadatenmodell nach Dublin Core

Mitte der 90er Jahre wurde ein weltweit standardisierter Metadatensatz entworfen mit dem Ziel, die Erschließung von elektronischen Dokumenten zu verbessern und neben der einfachen Volltextsuche auch qualifiziertere Recherchefunktionen zur Verfügung stellen zu können. Dazu trafen sich 1995 neben Bibliothekaren auch Wissenschaftler anderer Fachrichtungen aus aller Welt im US-amerikanischen Dublin.

Beim Entwurf ließen sich die Experten vor allem von den folgenden Prinzipien leiten:

- **Einfachheit:** Dublin Core soll insbesondere auch für Nicht-Bibliothekare benutzbar und verständlich sein.
- **Semantische Kompatibilität:** Die Metadaten sollen sich über Grenzen zwischen unterschiedlichen Fachbereichen hinweg semantisch entsprechen, um eine qualifizierte Recherche im gesamte Internet zu ermöglichen.
- **Internationaler Konsens:** Die Dublin Core Metadata Initiative (DCMI) stützt sich auf die aktive Beteiligung von etwa 30 Ländern in Nordamerika, Europa, Australien und Asien.
- **Erweiterbarkeit:** Der vorliegende Metadatensatz zeichnet sich durch die Möglichkeit aus, feinere semantische Strukturen durch flexible Erweiterungen abzubilden.

Zum Standard von Dublin Core gehören neben den eigentlichen Metadatenelementen natürlich vor allem ihre semantischen Beschreibungen, was eine Recherche mit Hilfe dieser Angaben überhaupt erst sinnvoll macht. Jedes Element von Dublin Core ist zunächst durch die zehn Attribute gekennzeichnet, die in der ISO/IEC 11179 für die Beschreibung von Datenelementen vorgesehen sind: *Name*, *Identifier*, *Version*, *Registration Authority*, *Language*, *Definition*, *Obligation*, *Datatype* sowie *Maximum Occurrence* und *Comment*.

Einige von diesen besitzen allerdings Werte, die für alle Elemente von Dublin Core identisch sind:

- Version: 1.1
- Registration Authority: Dublin Core Metadata Initiative
- Obligation: Optional
- Datatype: Character String
- Maximum Occurrence: Unlimited

Die Hauptelemente des Dublin Core sind in der nachfolgenden Tabelle zusammengefasst.

Tabelle 4.2 Elemente des Dublin Core Datensatzes, Version 1.1

Name	Identifier	Erläuterung
Title	Title	Name der Quelle
Creator	Creator	Verantwortliche Körperschaft (Person, Institution) für den Inhalt des Dokuments
Subject and Keywords	Subject	Thema der Quelle
Description	Description	Überblick über den Inhalt der Quelle (z.B. Abstract, Inhaltsverzeichnis)
Publisher	Publisher	Für die Verfügbarkeit der Quelle verantwortliche Körperschaft
Contributor	Contributor	Sonstige beteiligte Person(en) oder Organisation(en)
Date	Date	Datum, das mit dem Lebenszyklus der Quelle assoziiert ist (z.B. Erstellungsdatum, Veröffentlichungsdatum)
Resource Type	Type	Die Art oder das Genre der Quelle (z.B. generelle Kategorien, Funktionen)
Format	Format	Physikalische oder digitale Manifestation der Quelle (z.B. Datenformat, Größe, Systemvoraussetzungen)
Resource Identifier	Identifier	Eindeutige Referenz der Quelle innerhalb des Kontextes (z.B. URI inkl. URL, ISBN)
Source	Source	Referenz zur ursprünglichen Quelle
Language	Language	Sprache, in der der Inhalt der Quelle verfasst ist
Relation	Relation	Referenz auf eine verwandte Quelle
Coverage	Coverage	Abgedeckter geographischer Raum bzw. Zeitraum
Rights Management	Rights	Information über Rechte, die an der Quelle oder deren Teilen gehalten werden

Die Implementation der Dublin Core Metadaten in Internetseiten kann wie schon im vorangegangenen Abschnitt angesprochen anhand des HTML-Tags `<META NAME="...">` erfolgen. Den Namen der Dublin Core (DC) Elemente wird dabei das Präfix „DC.“

vorangestellt, um Suchmaschinen die Zugehörigkeit des Meta-Elements zum Dublin Core-Standard zu signalisieren.

Eine Dublin Core konforme Version der Angabe von Metadaten im HTML-Code sieht daher wie folgt aus:

```
<html>
<HEAD>
...
<META NAME="DC.Title" CONTENT=" BiM-Online - Informationssystem zum
Baustoffkreislauf im Massivbau">
<META NAME="DC.Identifizier" CONTENT="http://www.b-i-m.de">
<META NAME="DC.Creator" CONTENT="Marcus Schreyer">
<META NAME="DC.Publisher" CONTENT="Universit&uml;t Stuttgart, IWB">
<META NAME="DC.Language" CONTENT="ger">
<META NAME="DC.Description" CONTENT="BMBF-Forschungsvorhaben zum
Baustoffrecycling für Tragwerke des Massivbaus.">
<META NAME="DC.Subject" CONTENT="BiM BiM-Online Baustoffrecycling
Recyclingbeton Bauschutt Abbruch Baustoffaufbereitung">
</HEAD>
...
```

Alternativ können die DC-Metadatenelemente auch als RDF (Ressource Description Framework) Datensätze codiert werden. Auf diese Variante wird innerhalb der Erläuterung des RDF Konzeptes in den folgenden Kapiteln vertiefter eingegangen.

Eine Verfeinerung der 15 DC-Elemente bzw. eine formale Definition für die Inhalte und deren Darstellung wurde durch die Einführung von sogenannten Qualifiern erzielt. Die in [Dublin] dargestellten Qualifier Elemente stellen jedoch keine verbindlichen Metadatenätze dar, sondern sollen lediglich die Einbindung bereits bestehender, spezifischerer Datenschemata demonstrieren. In den dort publizierten Anwendungsbeispielen werden beispielsweise für die Inhalte des DC-Elements *Subject* die bibliothekarischen Klassifikationen (Vokabulare) des LCSH (Library of Congress Subject Headings) bzw. des MeSH (Medical Subject Heading) vorgeschlagen. Für die Inhalte der DC-Elemente *Source*, *Identification*, *Relation* sollten Adressangaben gemäß der URI Richtlinie [RFC2396] des W3C erfolgen. Die Sprache sollte in *Language* nach ISO 639-2/RFC 1766 [ISO639-2, RFC1766] angegeben werden. Beispiele für die Verfeinerung von Elementen sind die Unterteilung des *Description* Elements in die Qualifier *Table of Contents* und *Abstract* oder des DC-Elementes *Date* in *Created*, *Valid*, *Available*, *Issued* und *Modified*. Die Genehmigung neuer Qualifier Elemente erfolgt zentral durch das Dublin Core Metadata Initiative (DCMI) Usage Committee.

Zur Entlastung der Webautoren von mühsamer Tipparbeit wurden sogenannte Metadateneditoren entwickelt, welche die bspw. über ein Web-Formular eingegebenen Daten automatisch in die entsprechende Syntax umwandeln. Solch einen Metadateneditor bietet unter anderem das UKOLN (UK Office for Library Networking) an der Universität Bath in Großbritannien mit seinem DC-dot Online-Editor. DC-dot [UKOLN2000] unterstützt die Erfassung innerhalb eines HTML-Formulars durch Vorschlagswerte (siehe Abb. 4.2), die es automatisch aus dem Quellcode der HTML-Seite erzeugt. Der Eingabeaufwand zur Beschreibung einer Webseite konform dem DC-Metadatenstandard reduziert sich hierbei auf ein Minimum. Die im Browser angezeigte Metadatenbeschreibung, die wahlweise in HTML oder RDF-Syntax ausgegeben werden kann, muss vom Autor nur noch in den Quellcode der Webseite integriert werden.

Abb. 4.2 Benutzeroberfläche des Metadateneditors DC-dot [UKOLN2000] zur Erzeugung von Dublin Core konformen Metadatensätzen in HTML oder RDF

4.3 Entwurf von Metadatenstrukturen in BiM-Online

Schon der erste Entwurf von BiM-Online [BiM-Online2001] sah vor, die Zugänglichkeit der in den digitalisierten Forschungsberichten enthaltenen Informationen durch die Verwendung von Metadatensätzen zu verbessern. Hierfür wurde zur Erfassung der Ergebnisse aus Materialprüfungen ein lokaler Satz von Metadatenelementen (Tab. 4.3) entwickelt, über den die in den veröffentlichten Berichten enthaltenen Aussagen, Diagramme und Tabellen zu Prüfergebnissen beschrieben werden konnten. Den einzelnen Stellen wurden über ihre URI in relationalen Tabellen entsprechende Datensätze zugeordnet. Auf Grundlage dieser Metadaten konnte nun eine Benutzeroberfläche für die datenbankunterstützten Webseiten erstellt werden, die einen sehr effizienten Zugriff auf die in mehr als 180 Berichten verteilten Forschungsdaten bot. Recherchen, wie z. B. die gezielte Suche nach σ/ε -Diagrammen von rezykliertem Beton, reduzieren sich durch diese Strukturierung auf wenige Mausklicks.

Tab. 4.3 Ursprünglicher Aufbau der Metadatentabellen zur Speicherung der BiM-Prüfdatensätze [Schwarte2000]

Feldnamen	Beispiel	Funktion
Prüfklasse	Festbeton	Eindeutige Beschreibung des Prüfverfahrens
PrüfID	399	
Prüfbeschreibung	„Abhängigkeit der Spannungs-Dehnungsbeziehung von der Zuschlagart“	
Teilprojektkürzel	E03	Interne Bezeichnung des Teilprojekts, in dem die Prüfung durchgeführt wurde
Monat	09	Datum der Prüfung
Jahr	98	
URL	http://www.b-i-m.de/Berichte/E03/E03z0998.htm#4	Webadresse (URL) des Berichtes sowie Referenznummer der Markierung (Ankertag) im Quelltext
Relevanz-Prüfgruppe	5	Relevanzgrad der Prüfung in Bezug auf das Teilprojektthema

Durch die Verwendung eines relationalen Datenbanksystems zur Speicherung der Metadaten blieb die Nutzung der Metadatenstrukturen jedoch auf das lokale System beschränkt. Außerdem wurden beim Entwurf systemübergreifende Standards, wie der o.g. Dublin Core-Metadatenstandard, noch nicht berücksichtigt. Die entstandene lokale Metadatenstruktur schränkt somit die Vernetzbarkeit der BiM-Online-Informationsobjekte mit anderen Systemen ein.

In einer grundlegenden Überarbeitung der ursprünglichen Metadatenstruktur sollen sich daher folgende Punkte niederschlagen:

- Das Metadatenmodell sollte sich auf alle Informationsressourcen zu Beton mit rezyklierter Gesteinskörnung anwenden lassen und nicht nur auf Ergebnisse zu Materialprüfungen.
- Das Metadatenmodell sollte auch von externen Systemen genutzt werden können, um auf die in BiM-Online enthaltenen Informationsressourcen zugreifen zu können. Auf diese Weise wird eine Integration der Inhalte von BiM-Online mit den Inhalten anderer, thematisch verwandter Informationsquellen im Web möglich.
- Beim Entwurf der Metadatenstrukturen für die BiM-Online-Informationsressourcen sollen, soweit möglich, Elemente eingeführter Metadatenstandards berücksichtigt werden.
- Die Metadatenbeschreibungen sollten außerdem detailliertere Informationen zur Recherche der Informationsquellen enthalten. Auch die Zusammenhänge zwischen einzelnen Webinhalten sollten abbildbar sein.
- Die Integration des Metadatenmodells für Webinhalte zu Beton mit rezyklierter Gesteinskörnung in übergeordnete Modelle, z.B. für Webinhalte zum Bauwesen oder auch fachübergreifend sollten möglich sein.

Ziel ist es, ein Metadatenmodell zur Verfügung zu stellen, das eine detailliertere Beschreibung von betontechnologischen Webinhalten erlaubt, dabei jedoch allgemein bzw. systemübergreifend anwendbar bleibt.

4.3.1 Informationsressourcen in BiM-Online

Die ersten Schritte beim Entwurf eines Metadatenmodells bestehen zunächst in der Definition und Analyse der zu beschreibenden Informationsressourcen. Zur Beantwortung der Frage, welche Inhalte von BiM-Online überhaupt als Informationsressourcen definiert werden können, kann zunächst auf die ursprüngliche Systemstruktur zurückgegriffen werden.

Im ursprünglichen System wurden bereits für die in den Dokumenten enthaltenen Prüfdaten aus Materialprüfungen sowie für die in BiM-Online-Dialog als „Lektionen“ bezeichneten interaktiven Kurse Metadatenstrukturen entworfen. Da die Erschließung der Dokumente selbst über eine lokal auf dem Webserver installierte Volltextsuche erfolgte, wurde zur Beschreibung der Forschungsberichte und Fachpublikationen ursprünglich auf Metadaten verzichtet. Gerade diese Inhalte stellen jedoch die wichtigste Informationsressource des Systems dar, aus der sich andere, in den Dokumenten enthaltene, Informationsobjekte ableiten lassen. Bei diesen in den Dokumenten enthaltenen Informationsobjekten mit baustofftechnologischer Relevanz handelt es sich um:

- Betontechnologische Prüfdaten
- Betonrezepturen
- Chemische und mathematische Formeln
- Literaturverweise

Vor dem Entwurf von Metadatensätzen ist es sinnvoll, die zu beschreibenden Informationsobjekte zunächst in Ressourcentypen zusammenzufassen, welche jeweils einheitliche Merkmale besitzen. Tab. 4.4 bildet solch eine Zuordnung der Inhalte von BiM-Online zu den entsprechenden Ressourcentypen ab.

Tab. 4.4 Informationsressourcen in BiM-Online mit zugeordneten Ressourcentypen und Speicherformaten.

BiM-Inhalte	Ressourcentypen		Speicherformate
Fachpublikation Zwischenbericht Schlussbericht <i>Projektbeschreibung</i> <i>Mitarbeiterbeschreibung</i>	Dokument		HTML-Dokumente
Betontechnologischer Prüfdatensatz Betonrezeptur <i>Literaturhinweis</i>	Beton-Daten	Datensatz	HTML-Textfragmente, JPEG/GIF-Grafikdateien
Mathematische Formel Chemische Formel	Mathematische Formel Chemische Formel	Formel	HTML-Textfragmente, JPEG/GIF-Grafikdateien
Lektion	Kurs		HTML-Dokumente

Im zweiten Entwurfsschritt werden die Ressourcentypen in eine hierarchische Beziehung (Ressourcenbaum) zueinander gebracht. Ein solcher Ressourcenbaum stellt eine einfache Form einer Ontologie dar, wie sie in Abschnitt 1.2. angesprochen wurde. Die semantischen

Informationen sind dabei in den in der Hierarchie enthaltenen Klassen-Subklassen Beziehungen enthalten. Diese Struktur stellt den Kontext dar, innerhalb dessen die konkreten Webinhalte eingeordnet werden können.

Um die Kontextzuweisung für betontechnologische Informationsressourcen detaillierter gestalten zu können, wurde der Ontologie-Ast unterhalb des Elements „Beton“ noch weiter verfeinert. Für die Verfeinerung der Baumstruktur musste zunächst nach Möglichkeiten für weitere Gruppierungen der mit „Beton-Daten“ bezeichneten Diagramme, Tabellen etc. gesucht werden. Einen Ansatzpunkt bildet dabei die Zusammensetzung des Betons: Beton ist ein Baustoff, welcher sich aus mehreren Bestandteilen zusammensetzt. In den Tabellen und Diagrammen werden häufig die Eigenschaften dieser Bestandteile in Beziehung zueinander oder zu den Eigenschaften des Betons abgebildet. Im Umkehrschluss lassen sich die dort beschriebenen Bestandteile zur inhaltlichen Einordnung heranziehen. Des weiteren muss zwischen den Eigenschaften von Frisch- und Festbeton, den zugehörigen Randbedingungen der Prüfung, wie auch der Zusammensetzung der Mischungen unterschieden werden. Aus diesen Überlegungen zu den Komponenten des Betons selbst und seinen Herstellungsstufen (Herstellung bzw. Rezeptur, Frischbeton, Festbeton) heraus wurde die in Abb. 4.3 dargestellte Struktur entwickelt.

Die Ressourcentypen an den Knoten des Ontologiebaums bilden die „Objektklassen“, welchen die Metadatenelemente zugeordnet werden können. Innerhalb einer Ontologie dürfen die Ressourcentypen nur einmal auftreten. Die Metadatenelemente hingegen, welche die Eigenschaften der jeweiligen Ressourcentypen repräsentieren, können so allgemein definiert werden, dass sie mehreren Ressourcentypen zugeordnet werden können. Diese Flexibilität reduziert die Anzahl der benötigten Metadatenelemente bei der Beschreibung, da diese nicht für jeden Ressourcentyp erneut definiert werden müssen (vgl. Abschnitt 6.2.3). Zudem bietet die getrennte Definition von Metadaten und Ressourcentypen Vorteile für den Rechercheprozess. Mit Hilfe einer solchen Ontologie lassen sich Recherchen in großen Datenmengen semantisch präzisieren. Wird z.B. für einen Suchbegriff eine große Zahl von entsprechenden Informationsobjekten gefunden, kann die Trefferzahl durch die Beschränkung auf einen bestimmten inhaltlichen Kontext leichter eingegrenzt werden. Dies erlaubt auch eine größere Interaktivität bei Suchvorgängen. Eine Suchmaschine könnte bei sehr großen Trefferzahlen gezielt nach Spezifizierungen der Suchbegriffe fragen, in dem sie dem Anwender die gefundenen Kontexte bzw. Ressourcentypen-Eigenschaft-Kombinationen zur Auswahl anbietet (vgl. Abschnitt 6.6). Die Simulation eines Dialogs mit einer solchen semantischen Suchmaschine wird in Abschnitt 7.1 dargestellt.

Mit dem in dieser Arbeit vorgestellten met@BiM-Datenmodell werden ausgewählte Informationsressourcen der in Tab. 4.4 dargestellten Ressourcentypen von BiM-Online beschrieben. An ihnen soll exemplarisch der Aufbau und die Nutzung einer Ontologie mit dem Schwerpunkt auf betontechnologischen Webinhalten aufgezeigt werden. Deshalb werden die mehr administrativen Inhalte, wie z.B. die in Tab. 4.4 kursiv gedruckten Literaturhinweise, die Mitarbeiter- sowie die Projektbeschreibungen in dieser Arbeit nicht weiter betrachtet. Bezüglich Metadatenstrukturen für Personendaten oder auch Videofilme wird auf die Veröffentlichungen von Dawson/Howes [vCard1998] und Hunter et al. [Hunter1999] verwiesen. Der Detaillierungsgrad des entworfenen Ontologiebaums wurde auf den Umfang der in BiM durchgeführten Werkstoffprüfungen abgestimmt. Eine Erweiterung oder Verfeinerung der gewählten Struktur um weitere Ressourcentypen kann jedoch auch noch nachträglich erfolgen, ohne dass dies Änderungen in bereits bestehenden Datensätzen zur Folge haben muss.

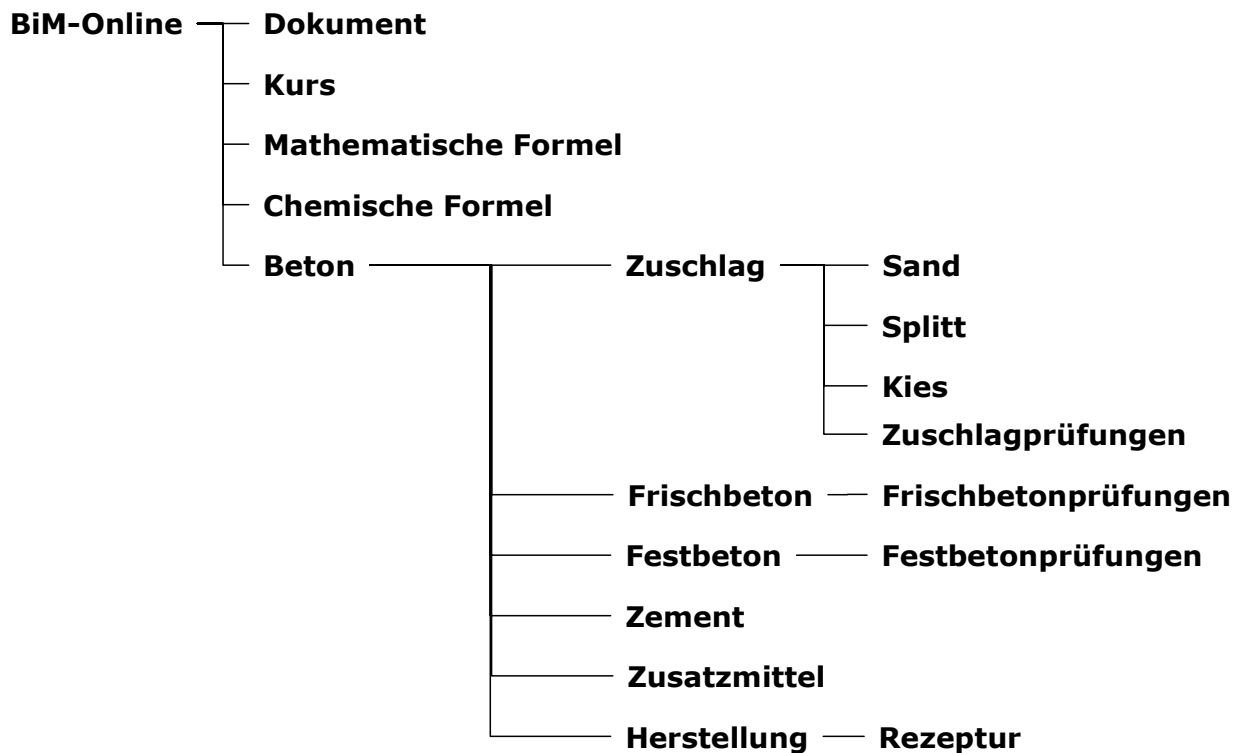


Abb. 4.3 Ontologie der BiM-Online Informationsressourcen

4.3.2 Anwendung der Dublin Core Metadaten in BiM-Online

Zur Erstellung eines Metadatenmodells müssen die kennzeichnenden Merkmale der im vorigen Kapitel definierten Informationsobjekte zusammengestellt werden. Die Auswahl der in der folgenden Tabelle angegebenen Merkmale erfolgte mit dem Ziel, die Auffindbarkeit und Zuordenbarkeit der Informationsobjekte zu verbessern. Es ist durchaus möglich und in vielen Fällen auch sinnvoll, zusätzliche Metadatenelemente für Urheberrechte oder technische Details der Webinhalte zu erfassen. Welche Metadatenelemente zum Einsatz kommen, wird letztendlich immer von der Zielsetzung der Modellierung bestimmt werden.

Die Grundlage für die Beschreibung aller Informationsobjekte in BiM-Online bilden die Elemente des Dublin Core-Schemas. Die Inhalte der einzelnen Dublin Core (DC)-Metadatenelemente werden dabei in Tab. 4.5 nochmals für die folgenden Ressourcentypen konkretisiert:

- Dokument
- Kurs
- Beton-Daten
- Mathematische Formel
- Chemische Formel

Die Elemente, die dem DC-Metadatenstandard entstammen, werden durch ein „dc:“ Präfix als zugehörig zum Dublin Core-Namensraum gekennzeichnet. Entsprechend wird den in met@BiM neu definierten Metadatenelementen das Namensraumpräfix „metabim:“ vorangestellt. Eine ausführliche Erläuterung des schon in Abschnitt 2.6.2 angesprochenen Namensraumkonzepts befindet sich in [Namespaces1999]. Die kombinierte Verwendung der

Tab. 4.5 Anwendung der Dublin Core-Metadatenelemente zur Beschreibung der BiM-Online-Informationsressourcen

Anwendung der Dublin Core-Metadatenelemente in BiM-Online		
Name	Identifier	Erläuterung
Type	dc:type	Typ der Ressource. In BiM-Online werden 5 Typen verwendet: • Dokument (Webdokument) • Kurs (Lektion/Online-Kurs) • Beton-Daten (Betontechnologische Daten), • Mathematische Formel, • Chemische Formel
Title	dc:title	Titel der Ressource • Dokumententitel (Dokument) • Titel des Kurses bzw. der Lektion (Kurs) • Charakteristische Bezeichnung des Prüfergebnisses, z.B. die Tabellen-/Bildüber-/unterschrift (Beton-Daten) • Charakteristische Bezeichnung der Rezeptur, z.B. die Tabellen-/Bildüber-/unterschrift (Beton-Daten) • Charakteristische Bezeichnung der Gleichung (Formel)
Format	dc:format	MIME-Dateityp der Ressource
Identifier	dc:identifier	Eindeutiger Zeiger auf die Ressource • URL + # (Fragment Identifier) + lokaler Anchor-Tag (Beton-Daten) • URL + # + math + lokaler Anchor-Tag (Mathematische Formel) • URL + # + chem + lokaler Anchor-Tag (Chemische Formel) • URL des Dokuments (Dokument, Lektion/Kurs) • URL der Startseite des Kurses (Kurs)
Date	dc:date	Datum der Erzeugung der Ressource
Creator	dc:creator	Namen der Autoren der Ressource
Publisher	dc:publisher	Name des verantwortlichen Herausgebers. Dies können z.B. Personen sein oder auch Körperschaften wie Unternehmen, Hochschulen etc.
Source	dc:source	Zeiger auf die übergeordnete Speichereinheit • URL des Dokuments (Beton-Daten, Formel) • URL der Web-Site bzw. Rubrik (Dokument, Kurs)
Relation	dc:relation	Zeiger auf andere Ressourcen, die semantisch mit der aktuellen zusammenhängen. • Zeiger auf Ressource mit Betonrezepturbeschreibung bei Prüfdaten (Beton-Daten) • Zeiger auf Ressourcen mit Varianten der Betonrezeptur, z.B. Versuchsreihe (Beton-Daten) • Zeiger auf weitere Formeln im Gleichungssystem (Formel) • Zeiger auf weitere Berichte/Publicationen zum Thema (Dokument) • Zeiger auf weitere Lektionen zum Thema (Kurs)
Description	dc:description	Kurzer Text zur Erläuterung der Besonderheiten der Ressource
Subject	dc:subject	Angaben zum Inhalt der Ressource • Stichwortliste (Dokument, Kurs) • „Komplexe“ Inhalte aus untergeordneten met@BiM-Elementen (Beton-Daten) • Ausgeschriebene Bezeichnungen der mathematischen oder chemischen Komponenten der Gleichung (Formel)

fachspezifischen met@BiM-Elemente und des Dublin Core-Standards ermöglicht es externen Suchmaschinen, welche die DC-Metadaten berücksichtigen, zumindest diese Informationen in den Metadatensätzen korrekt zu verarbeiten.

Zwischen den in Tab. 4.5 aufgelisteten DC-Metadatenelementen gibt es keine Hierarchierung. Alle Elemente können unabhängig von einander den o.g. Ressourcentypen zugeordnet werden. Eine Sonderstellung besitzt das DC-Element „dc:subject“. Werden betontechnologische Datensätze beschrieben, so werden die hierfür vorgesehenen Elemente aus met@BiM in Tab. 4.6 und 4.7 dem Metadatenelement *dc:subject* untergeordnet. Im Gegensatz zu den anderen DC-Elementen kann *dc:subject* daher auch komplexe Inhalte, d.h. weitere untergeordnete Elemente, beinhalten. Zur besseren Übersichtlichkeit werden diese met@BiM-Elemente zur verfeinerten Beschreibung betontechnologischer Daten in Tab. 4.6 des folgenden Abschnitts dargestellt.

4.3.3 Detailliertes Metadatenmodell für Daten zu rezykliertem Beton

Mit den Elementen des Dublin Core-Standards alleine lassen sich fachspezifische Merkmale in Webressourcen nicht detailliert genug beschreiben. Um die in Baustoffdiagrammen und Tabellen enthaltenen Aussagen ausreichend genau abbilden zu können, werden daher weitere Metadatenelemente benötigt. In der in Abb. 4.3 dargestellten Ontologie wurden mit der Verfeinerung des Astes unterhalb von „Beton-Daten“ zusätzliche Informationsressourcen definiert. In einer detaillierteren Betrachtung müssen nun die Merkmale dieser untergeordneten Informationsobjekte beschrieben werden. Die Beschreibung der betontechnologischen Merkmale anhand von Metadatenelementen wird auf diese Weise auf spezifischere Ressourcen konkretisiert und dadurch übersichtlicher in der Handhabung.

Die Auswahl der Baustoffeigenschaften, welche durch die in Tab. 4.6 beschriebenen Metadatenelemente des met@BiM-Namensraums repräsentiert werden, erfolgte anhand der in BiM-Online vorhandenen Datenbasis von ca. 540 Prüfergebnissen bzw. Rezepturen zu Beton mit rezyklierter Gesteinskörnung. In Grafiken wie Abb. 4.4 und im Anhang A wurden den baustofftechnologischen Informationsressourcen charakteristische Eigenschaften zugeordnet.

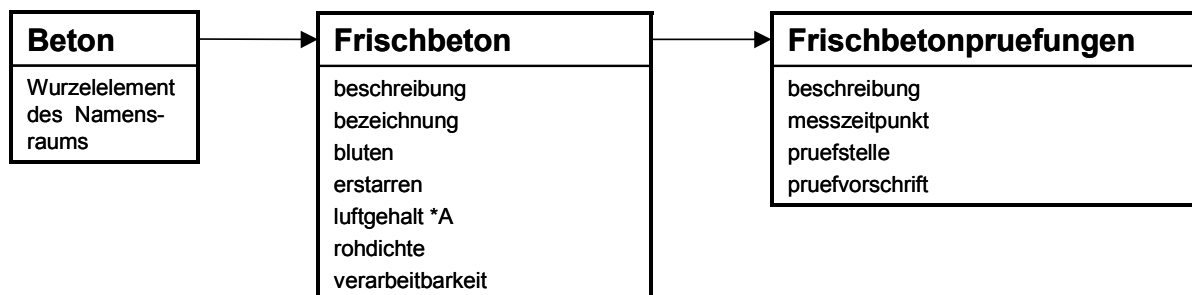


Abb.4.4 Zuordnung der Baustoffeigenschaften bzw. Metadatenelemente zu den Ressourcen unterhalb des Elements „Beton-Daten“ am Beispiel Frischbeton

Bei der Zuweisung der Metadaten zu den unterschiedlichen Ressourcentypen stellte sich recht schnell heraus, dass einige der Elemente zur Beschreibung mehrerer Informationsobjekte herangezogen werden konnten. Auffällig waren solche Dopplungen auch bei direkt einander übergeordneten Ressourcen, wie z.B. den Metadatenelementen zur Beschreibung der Eigenschaften von Betonzuschlag (nach DIN EN 206-1 bzw. DIN 1045-2 wird Zuschlag als „Gesteinskörnung“ bezeichnet) und dessen Komponenten Sand, Kies und Splitt. Alle vier genannten Ressourcen lassen sich hierbei anhand derselben Metadatenelemente beschreiben (s.a. Anhang A).

Tab. 4.6 met@BiM-Metadatenelemente für betontechnologische Ressourcen

met@BiM-Metadatenelemente zur detaillierten Beschreibung von Betoneigenschaften		
Name	Identifier	Erläuterung
Abschlämbbare Bestandteile	metabim:abschlaemmbarbestandteile	Angaben zu den abschlämbbaren Bestandteilen (DIN 4226, Teil 3)
Alter	metabim:alter	Alter der Probe
Außenbauteileignung	metabim:aussenbauteileignung	Eignung der Betonmischung für Außenbauteile (DAfStb-Richtlinie „Beton mit rezykliertem Zuschlag“)
Beschreibung	metabim:beschreibung	Kurze Erläuterung der Besonderheiten
Bezeichnung	metabim:bezeichnung	Produktname, Probenbezeichnung etc.
Biegezugfestigkeit	metabim:biegezugfestigkeit	Zugfestigkeit des Probekörpers ermittelt im Biegezugversuch
Bluten	metabim:bluten	Angaben zum Bluten des Betons beim Verdichten
Carbonatisierung	metabim:carbonatisierung	Angaben zur Carbonatisierungstiefe
Chloridgehalt	metabim:chloridgehalt	Angaben zum Chloridgehalt
Druckfestigkeit	metabim:druckfestigkeit	Druckfestigkeit des Probekörpers
Druckspannung	metabim:druckspannung	Angaben zur Druckspannung im Probekörper während eines Versuchs
E-Modul	metabim:emodul	Angaben zum E-Modul der Probe
Erhärtungsstörende Stoffe	metabim:erhaertungsstoerendestoffe	Angaben zum Gehalt erhärtungsstörender Stoffe (DIN 4226, Teil 3)
Erstarren	metabim:erstarren	Angaben, die das Erstarren des Frischbetons beschreiben
Farbe	metabim:farbe	Farbe des Probekörpers
Festigkeitsklasse	metabim:festigkeitsklasse	Einstufung der Probe in eine Festigkeitsklasse
Frost-Tausalz Widerstand	metabim:frosttausalzwiderstand	Frost-Tausalz Widerstand der Probe (DIN 4226, Teil 1 bzw. DAfStb Heft 422)
Frostwiderstand	metabim:frostwiderstand	Frostwiderstand der Probe (DIN 4226, Teil 1 bzw. DAfStb Heft 422)
Gesamtwassergehalt	metabim:gesamtwassergehalt	Gesamtwassergehalt der Betonmischung, z.B. Zugabewasser und Kernfeuchte des Zuschlags
Herstellername	metabim:herstellername	Name des Herstellers, z.B. Betonwerk, Zuschlaghersteller etc.
Hoher Frost-Tausalz Widerstand	metabim:hoherfrosttausalzwiderstand	Eignung der Betonmischung für die erhöhte Beanspruchung durch Frost-Tausalz (DIN 4226, Teil 1 bzw. DAfStb Heft 422, DAfStb-Richtlinie „Beton mit rezykliertem Zuschlag“)
Hoher Frostwiderstand	metabim:hoherfrostwiderstand	Eignung der Betonmischung für die erhöhte Beanspruchung durch Frost (DIN 4226, Teil 1 bzw. DAfStb Heft 422, DAfStb-Richtlinie „Beton mit rezykliertem Zuschlag“)
Hoher Widerstand gegen chemischen Angriff	metabim:hoherwiderstandggchem-angriff	Eignung der Betonmischung für die erhöhte Beanspruchung durch schwachen chem. Angriff (DIN 1045, DAfStb-Richtlinie „Beton mit rezykliertem Zuschlag“)
Kernfeuchte	metabim:kernfeuchte	Angaben zur Kernfeuchte des Zuschlags
Kornfestigkeit	metabim:kornfestigkeit	Angaben zur Druckfestigkeit der Zuschlagkörner

Kornform	metabim:kornform	Angaben zur Kornform (DIN 4226, Teil 3)
Kornrohddichte	metabim:kornrohddichte	Angaben zur mittleren Rohddichte der Zuschlagkörner
Kornzusammensetzung	metabim:kornzusammensetzung	Angaben zu den einzelnen Komponenten des Zuschlags innerhalb einer Fraktion
Kriechen	metabim:kriechen	Verformungen der Probe durch Kriechen
Lastverformungen	metabim:lastverformungen	Verformungen unter aufgebrachter Last
Luftgehalt	metabim:luftgehalt	Angaben zum Luftgehalt von Frischbeton
Menge	metabim:menge	Absolute Zugabemenge (allg.)
Messzeitpunkt	metabim:messzeitpunkt	Messzeitpunkte der Wertermittlung
Organische Stoffe	metabim:organischestoffe	Angaben zum Gehalt an organischen Stoffen
Probekörpergeometrie	metabim:probekörpergeometrie	Kurze Beschreibung der Probekörpergeometrie z.B. Form, Maße
Prüfstelle	metabim:prüfstelle	Name und Angaben zur prüfenden Stelle
Prüfvorschrift	metabim:prüfvorschrift	Der Prüfung zugrundegelegte Vorschrift oder Richtlinie
Quellen	metabim:quellen	Verformungen durch Quellen
Rohddichte	metabim:rohddichte	Rohddichte des Probekörpers
Schüttdichte	metabim:schuettdichte	Angaben zur Schüttdichte
Schwinden	metabim:schwinden	Verformungen durch Schwinden
Sieblinie	metabim:sieblinie	Gewählte Sieblinienklassifizierung
Spaltzugfestigkeit	metabim:spaltzugfestigkeit	Zugfestigkeit des Probekörpers ermittelt im Spaltzugversuch
Spannbetoneignung	metabim:spannbetoneignung	Eignung der Betonmischung für die Verwendung in Spannbetonbauteilen (DIN 4227 T.1)
Sulfatgehalt	metabim:sulfatgehalt	Angaben zum Sulfatgehalt
Temperaturbeanspruchung	metabim:temperaturbeanspruchung	Temperaturbeanspruchung der Probe, z.B. vor Lastversuchen
Verarbeitbarkeit	metabim:verarbeitbarkeit	Angaben zur Verarbeitbarkeit des Frischbetons. Dies können sowohl die numerischen Werte aus Ausbreitversuchen sein als auch daraus abgeleitete Angaben wie z.B. die Konsistenzklasse
Wasseraufnahme	metabim:wasseraufnahme	Angaben zur Wasseraufnahme des Zuschlaggemisches
Wasserundurchlässigkeit	metabim:wasserundurchlässigkeit	Eignung der Betonmischung für den erhöhten Anspruch an die Wasserundurchlässigkeit (DIN 1045-2 bzw. 1045, DAfStb-Richtlinie „Beton mit rezykliertem Zuschlag“)
Wirkung	metabim:wirkung	Kurze Erläuterung der Wirkung, die durch die Zugabe auf den Beton erzielt wird
w/z-Wert	metabim:wzwert	Wasser-Zement Wert der Mischung
Zementgehalt	metabim:zementgehalt	Zementgehalt der Mischung
Zugspannung	metabim:zugspannung	Angaben zur Zugspannung im Probekörper während eines Versuchs
Zulassungsbezeichnung	metabim:zulassungsbezeichnung	Bezeichnung der Zulassung für die Mischung
Zuschlaggehalt	metabim:zuschlaggehalt	Anteil des Zuschlags in der Mischung
Zusammensetzung	metabim:zusammensetzung	Beschreibung der unterschiedlichen Komponenten des Zuschlaggemisches

In Tab. 4.6 sowie Tab. 4.7 werden alle in met@BiM verwendeten Metadatenelemente zur detaillierten Beschreibung von Webinhalten mit betontechnologischen Daten definiert. Der aktuelle Umfang des von Dopplungen bereinigten Metadatenschemas beträgt 55 Elemente.

Einige der Baustoffeigenschaften, wie z.B. die Außenbauteileignung von Betonmischungen oder der Anteil erhärtungsstörender Stoffe im Zuschlag, basieren auf nationalen Normen und Richtlinien. Diese waren zwar zum Zeitpunkt der Durchführung der Untersuchungen gültig, wurden jedoch inzwischen bzw. werden im Rahmen der europäischen Normung ersetzt. Damit sind auch die Bezeichnungen der Eigenschaften, die durch diese Normen definiert werden, heute nicht mehr aktuell. Auch weitere Baustoffeigenschaften, die aus anderen Normgebungen in die europäischen Vorschriften mit einfließen, fehlen in dieser Liste. An dieser Stelle sei deshalb erneut auf die zwangsläufige Unvollständigkeit eines solchen Metadatenschemas hingewiesen, da je nach Anwendungszweck der Schwerpunkt auf anderen Ressourceneigenschaften liegen wird. Es ist jedoch so aufgebaut, dass es durch Hinzufügen neuer Metadatenelemente oder betontechnologischer Informationsressourcen in der Ontologie in Abb. 4.3 erweitert werden kann.

Die meisten baustofftechnologischen Eigenschaften wurden jedoch in ihrer Definition bewusst nicht bezüglich Konformität zu bestimmten Normen eingeschränkt. Sind beispielsweise mehrere Prüfverfahren zur Ermittlung einer Baustoffeigenschaft möglich, wie dies bei der Konsistenz von Frischbeton der Fall ist, so sollten entsprechende Angaben zum verwendeten Prüfverfahren bzw. Normabschnitt innerhalb des Metadatensatzes gemacht werden.

Mit den o.g. Elementen lassen sich die in den Tabellen, Diagrammen etc. angegebenen betontechnologischen Parameter beschreiben. Für die Recherche solcher Webinhalte wäre es jedoch darüber hinaus sinnvoll, wenn sich auch zumindest die kennzeichnenden Werte, welche diese Eigenschaften in den Diagrammen annehmen, im Datenmodell abbilden ließen. Eine zweckmäßige Position für Wertangaben innerhalb der met@BiM-Hierarchie wäre die Anordnung unterhalb des entsprechenden Elements zur Beschreibung der zugehörigen betontechnologischen Eigenschaft. Bei der Beschreibung der Werte stellt sich jedoch das Problem des Erfassungsaufwands: Diagramme und Tabellen stellen ja an sich bereits eine zusammenfassende Darstellung der Werte dar. Es ist daher nicht sinnvoll, in einer weiteren Verfeinerung der Beschreibung alle Werte erneut herauszulesen und als Metadatensätze zu modellieren.

Um deshalb den Umfang und die Komplexität der Beschreibungen in Grenzen zu halten, können Werteangaben auch in Bereichen zusammengefasst werden. Mit der Angabe einer Unter- bzw. Obergrenze des Wertebereichs muss zwar eine gewisse Unschärfe in Kauf genommen werden, die Recherchierbarkeit der Inhalte wird jedoch dadurch nicht wesentlich beeinflusst.

Die Zuordnung der physikalischen Maßeinheit und Dimension eines Wertes erfolgt in met@BiM über das Element mit dem Identifier *metabim:einheit*. Die in Abb. 4.5 gewählte Lösung, den gesamten Ausdruck innerhalb eines einzigen Metadatenelements als Text zu beschreiben, stellt zunächst eine zweckorientierte, an die Erfordernisse zur Modellierung der BiM-Online-Inhalte angepasste Lösung dar. Aufgrund der allgemeineren Anwendbarkeit einer komplexeren Lösung ist jedoch zu erwarten, dass hierfür bald entsprechende Vorschläge zu einer standardisierten Metadatenmodellierung der Maßeinheit und Dimension von Werten mit einem dafür vorgesehenen Namensraum veröffentlicht werden.

Die Modellierung der Aussagen zu den Elementwerten kommt nicht ohne eine Hierarchierung aus. Wertebereiche z.B. sind komplexe Metadatenelemente, die durch die beiden Elemente Unter- bzw. Obergrenze definiert werden. Das Element *metabim:einheit* wird innerhalb von met@BiM immer direkt der betontechnologischen Eigenschaft zugeordnet. Damit definiert dieses Element die Messeinheit für alle Werte der übergeordneten Eigenschaft und verkürzt somit den Umfang der Wertebeschreibung. Die Modellierung der in Tab. 4.7 zusammenge-

stellten met@BiM-Elemente wird am Beispiel der Zuordnung des Wertebereichs von 21.5 N/mm^2 bis 39.8 N/mm^2 für die *Druckfestigkeit* einer Probe in Abb. 4.5 dargestellt.

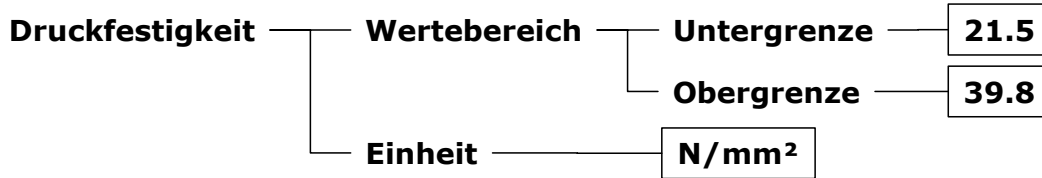


Abb. 4.5 Zuordnung von Werten bzw. Wertebereichen für die betontechnologischen Eigenschaften in met@BiM

Tab. 4.7 met@BiM-Metadatenelemente zur Verfeinerung der Elementwerte

met@BiM-Metadatenelemente zur Beschreibung der Elementwerte		
Name	Identifizier	Erläuterung
Einheit	metabim:einheit	Ausdruck der mathematischen Dimension und physikalischen Kenngröße eines Wertes
Anteil	metabim:anteil	Relativer Wert. Werden, z.B. zur Beschreibung der Zusammensetzung eines Baustoffes, die Anteile der einzelnen Komponenten benötigt, kann deren Höhe über dieses Element angegeben werden.
Wertebereich	metabim:wertebereich	Komplexes Element, welches zwei numerischer Grenzwerte enthält, innerhalb welcher sich die in der Quelle beschriebenen Werte bewegen
Untergrenze	metabim:untergrenze	Unterer Grenzwert eines Wertebereichs
Obergrenze	metabim:obergrenze	Oberer Grenzwert eines Wertebereichs

5 Das Resource Description Framework (RDF) – Metadaten im Internet

In diesem Kapitel wird mit dem Resource Description Framework ein Konzept vorgestellt, welches die Anwendung des Metadatenkonzepts auf Informationsressourcen im Web ermöglicht.

Zum Verständnis der RDF-Syntax sowie der grafischen Modelle werden zunächst die Grundlagen des Konzepts an einfachen Beispielen erläutert. Die Bedeutung und der Einsatz von Ontologien für die standardisierte Modellierung der im Web verteilten Informationsressourcen wird in den folgenden Abschnitten kurz dargestellt. Am Ende dieses Kapitels stehen Überlegungen zu Anwendungsschwerpunkten von RDF sowie ein Überblick über einige wichtige Softwareentwicklungen zur Nutzung der RDF-Datenstrukturen.

5.1 Allgemeine Architekturaspekte

In Abschnitt 2.6 wurden zahlreiche Vorteile wie die bessere Recherchierbarkeit, Flexibilität, Erweiterbarkeit und Skalierbarkeit genannt, die XML für die Modellierung und Realisierung verteilter Informationssysteme bringen soll. Unter Verwendung von aussagekräftigen XML-Tag-Namen erfasste und semantisch ausgezeichnete Daten lassen sich miteinander unterscheidbar vermischen und nutzersensitiv darstellen. Es bleiben allerdings noch eine Reihe von Problemen auf höheren Abstraktionsebenen bestehen, die insbesondere bei der Integration heterogener Systeme sichtbar werden. Während es Menschen häufig keine Schwierigkeiten bereitet, aus XML-Elementnamen auf die Bedeutung des Inhalts zu schließen, fehlt Computersystemen ein formales Konzept, aus dem sie die Semantik der Datenelemente ableiten können.

Deshalb wurde ein auf XML aufbauender Ansatz (Abb. 5.1) ins Leben gerufen, der unter anderem zur Bewältigung der Heterogenität verteilter Informationssysteme, insbesondere deren semantischer Heterogenität, beitragen kann.

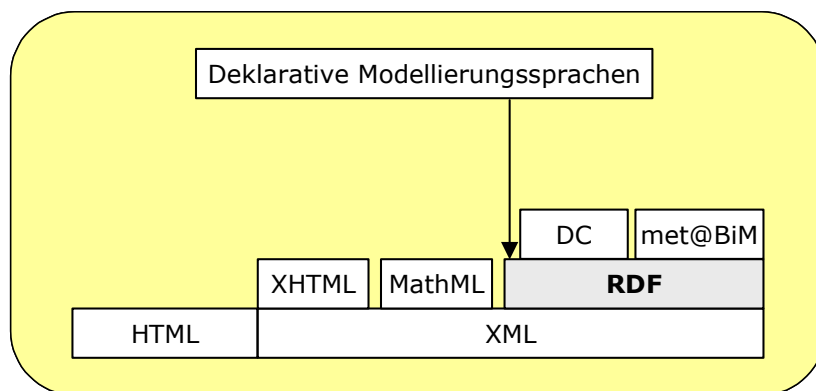


Abb.5.1 RDF wird im WWW als XML-Anwendung genutzt, auf deren Grundlage sich anwendungsspezifische Datenmodelle wie Dublin Core oder met@BiM erstellen lassen

Heterogenität ist ein bekanntes Problem aus dem Bereich verteilter Datenbanken. Aber auch für andere verwandte Gebiete, wie z.B. elektronische Bibliotheken, hat es zunehmend an Relevanz gewonnen [Paepke1998]. Obwohl die Interoperabilität der Informationssysteme durch Standardschnittstellen (z.B. JDBC) und Standardabfragesprachen (z.B. SQL) verbessert

werden kann, bleibt das schwierige Problem der semantischen Heterogenität zu lösen. Semantische Heterogenität besteht in erster Linie darin, dass unabhängig entwickelte Informationssysteme mit verschiedenen Datenmodellen (Schemata) operieren. Diese können auch bei teilweiser inhaltlicher Übereinstimmung miteinander gar nicht oder nur mit einem erheblichen manuellen Aufwand kombiniert werden. Als übergeordnete, integrierende Metadatensprache lässt sich hier das Resource Description Framework (RDF) einsetzen [Miller1998].

RDF kann XML als gemeinsame Syntax für Austausch und Verarbeitung von Informationen verwenden. Es schreibt jedoch keine bestimmte Syntax für Datenmodelle vor. Durch Ausnutzung der Eigenschaften von XML legt RDF eine Struktur zur eindeutigen Angabe der Semantik fest, die konsistente Aufbereitung, Austausch und automatische Verarbeitung von Daten erlaubt. RDF ist eine Entwicklung des W3C-Konsortiums [RDF1999] im Rahmen der Semantic Web Initiative [SemanticWeb2001]. Die Arbeiten an den RDF-Standards sind noch nicht abgeschlossen. Die Initiative wird von den Web-Standardisierungsgremien (HTML, PICS), der bibliothekarischen Gemeinschaft (Warwick Framework [Lagoze1996], Dublin Core [Dublin]), dem Bereich der strukturierten Dokumente (SGML, XML) und nicht zuletzt der Wissensrepräsentation vorangetrieben. Weitere Technologien, die zum Design von RDF beitrugen, sind die objekt-orientierte Programmierung, die Modellierungssprachen (z.B. Unified Modelling Language (UML) [UML1997, Chang1998]) und die Datenbanken. Der Entwurf von RDF wird von führenden Unternehmen der Computerindustrie wie IBM, Nokia, Netscape, Microsoft, Oracle, Apple u.a. unterstützt.

Primär wurde RDF als eine Infrastruktur zum Austausch und zur Verarbeitung von Metadaten über Webressourcen entworfen. Dieser Ansatz kann dahingehend verallgemeinert werden, dass anstelle von Webressourcen beliebige Arten von Ressourcen aus diversen Domänen verwendet werden können, wie z.B. Überweisungsaufträge, CAD-Modelle, gedruckte Dokumente usw..

Im Wesentlichen besteht die RDF-Spezifikation aus drei Komponenten. In [RDF1999] werden das Datenmodell und die Syntax festgelegt, mit deren Hilfe die Informationen repräsentiert werden. [RDFS2000] definiert Schemata, die als Sammlungen von Klassen und Eigenschaften aufgefasst werden können. Die Schemata sind typischerweise anwendungs- oder fachdomänenspezifische Wissensmodelle bzw. Ontologien. Hierarchische Organisation von Klassen und Eigenschaften sowie Erweiterbarkeit durch Vererbung sollen eine maschinelle Erschließung von (Meta)Daten ermöglichen. RDF kann als ein einfaches Wissensrepräsentationssystem [Russel1995] betrachtet werden, auf dessen Grundlage ein Inferenzmechanismus aufgebaut werden kann.

Abbildung 5.2 veranschaulicht die Rolle von RDF. Bei der Kommunikation zwischen Anwendungen dient RDF als eine standardisierte Datenschicht, welche lokal verarbeitete Informationen in ein RDF-Modell umsetzt. Dieses wird anschließend serialisiert, d.h. in eine Zeichenfolge umgewandelt. XML wurde dabei als Serialisierungsstandard gewählt. Es ist denkbar, dass im Laufe der Zeit weitere Serialisierungsformate entwickelt werden.

Es soll dennoch darauf hingewiesen werden, dass RDF das Potential besitzt, ein bedeutendes Werkzeug insbesondere für verteilte sowie offene webbasierte Informationssysteme zu werden. In den folgenden Abschnitten wird deshalb das Modell, die Syntax und das Schemakzept von RDF skizziert. Dadurch kann die Semantik des met@BiM zugrundeliegenden Datenmodells besser verstanden werden.

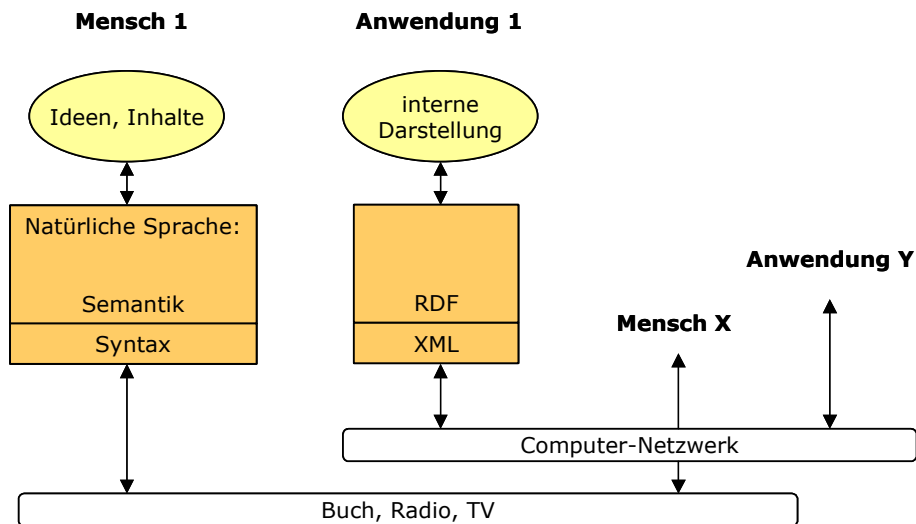


Abb. 5.2 Repräsentationsschichten für den Informationsaustausch zwischen Kommunikationspartnern [Melnik1999]

5.2 RDF-Datenmodell

Das RDF-Datenmodell basiert auf etablierten Prinzipien der Datenrepräsentation. Wie aus dem Namen „Resource Description Framework“ hervorgeht, spielt der Begriff Ressource dabei eine zentrale Rolle. Im Zusammenhang mit dem WWW wird der Begriff der Ressource für ein diskretes, im Sinne von „abgrenzbares“, Informationsobjekt verwendet, welches anhand seiner Eigenschaften beschrieben werden kann. In RDF können Ressourcen ein einzelnes elektronisches Dokument, eine komplette Web-Site, Grafiken, Videos, Dienste oder auch andere Datensätze sein, also ganz allgemein etwas, das durch eine URI (vgl. Abschnitt 4.2) identifiziert werden kann. Dank der Erweiterbarkeit von URIs kann eine URI-Referenz für jede denkbare Entität eingeführt werden. Somit lässt sich der Begriff Ressource sehr weit fassen und beispielsweise auch auf Objekte anwenden, die nicht direkt über das Web zugänglich sind oder abstrakt sind, wie z.B. Vorlesungen u.ä..

Einer Ressource können Eigenschaften (properties) zugeordnet werden (z.B. Name oder Adresse einer Person). Der Wert einer Eigenschaft kann durch ein Literal ausgedrückt und somit als Attribut der Ressource betrachtet werden (Abb. 5.3). Literale können atomare Werte wie Unicode-Zeichenketten oder Zahlen sein, können aber auch beliebige strukturierte XML-Daten beinhalten (eine Unicode-Spezifikation kann in [Unicode2001, XML2000] gefunden werden). Neben Literalen kann eine Eigenschaft auch eine Beziehung mit einer weiteren Ressource herstellen. Eine Menge von Eigenschaften einer Ressource wird als RDF-Beschreibung bezeichnet.

Eine gegebene Ressource mit einer Eigenschaft und dem Wert dieser Eigenschaft stellt eine RDF-Aussage dar. Analog zu einer natürlichsprachlichen Aussage werden die drei Komponenten einer RDF-Aussage auch als *Subjekt* (Ressource), *Prädikat* (Eigenschaft) und *Objekt* (Wert der Eigenschaft) bezeichnet. Im allgemeinen kann eine RDF-Aussage wie folgt gelesen werden:

Subjekt „hat“ *Prädikat* „mit dem Wert“ *Objekt*

RDF-Aussagen können daher als Tripel (p ; s ; o) aufgefasst werden, wobei p das Prädikat, s das Subjekt und o das Objekt bezeichnen. Die RDF-Aussagen werden graphisch als gerichtete, markierte Graphen dargestellt. In diesen Graphen repräsentieren Knoten die Ressourcen (als Ovale) und gerichtete Kanten (als Pfeile) repräsentieren Eigenschaften (Graphenbezeichnung). Literale werden als Rechtecke dargestellt. Das Verhältnis zwischen Ressourcen und Eigenschaften kann an einem Beispiel illustriert werden. Betrachten wir den Satz:

Vorlesung <http://iwb.uni-stuttgart.de/Vorlesungen/WS02-001> mit der Bezeichnung „Mineralische Bindemittel“ wird von einem Individuum gehalten, das Otto Graf heißt und dessen E-Mail Adresse graf@iwb.uni-stuttgart.de lautet.

Dieser Sachverhalt enthält gleich mehrere Aussagen, die grafisch als RDF-Modell in Abbildung 5.4 veranschaulicht werden. Die Eigenschaft *rdf:type* ist vordefiniert und dient zur Typisierung der Ressource. Mit ihrer Hilfe wird das RDF Klassensystem, das sogenannte RDF-Schema (siehe nächster Abschnitt), aufgebaut. Die unbenannte Ressource in Abbildung 5.4 kann durch einen eindeutigen Bezeichner analog zur Vorlesung (WS02-001) benannt werden. Die Wahl des Bezeichners liegt in den Händen des Entwicklers des Informationssystems.

Über die Zuordnung von Eigenschaften zu Ressourcen hinaus gestattet das RDF-Datenmodell auch die Formulierung von Aussagen über andere RDF-Aussagen. Somit ist es z.B. möglich auszudrücken, dass „das Institut für Werkstoffe im Bauwesen bestätigt, dass die Vorlesung Mineralische Bindemittel von Otto Graf gehalten wird“. Damit solche Aussagen modelliert werden können, müssen Aussagen selbst als Ressourcen behandelt werden. RDF ermöglicht dieses als Reifikation bezeichnete Konstrukt durch die Einführung von weiteren vordefinierten Eigenschaften. Mit der Möglichkeit zur Reifikation erfüllt RDF die Kriterien einer Modellierungssprache, die zum Aufbau von Wissensbasen genutzt werden kann. Damit würde die Ausführung von Inferenzalgorithmen und damit auch selbstständig agierender Programme wie Webagenten oder Expertensysteme auf den Inhalte des WWW ermöglicht. Eine vollständige Diskussion dieser Vorgehensweise würde im Rahmen dieser Arbeit zu weit führen. Eine detaillierte Beschreibung der Ausdrucksfähigkeit des RDF-Datenmodells sowie die vollständige RDF Grammatik sind in [RDFS2000] zu finden.

5.3 RDF-Syntax

Die RDF-Modell Spezifikation [RDF1999] beschreibt einen syntaxunabhängigen abstrakten Rahmen für die Datenmodellierung. Zur Aufbereitung und zum Austausch von Daten wird jedoch eine konkrete Syntax bzw. Codierung benötigt. Durch die Verwendung von XML können Instanzen des RDF-Modells in maschinen- und menschenlesbarer Form abgespeichert und ausgetauscht werden. Die Darstellung eines RDF-Modells in XML ist aus logischer Sicht nicht eindeutig. Genauso, wie man ein und denselben Sachverhalt in der natürlichen Sprache auf verschiedene Weise ausdrücken kann, kann ein und dasselbe RDF-Modell in unterschiedlicher Art und Weise codiert werden. Zudem wird einem Neuling die Erschließung von RDF-Beschreibungen durch drei mögliche XML Kurzformen der RDF-Syntax erschwert, die in der Spezifikation [RDF1999] vorgeschlagen werden. Zwei syntaktische Darstellungen sind genau dann gleichwertig, wenn die resultierenden RDF-Modelle (Graphen oder Mengen von Aussagen) übereinstimmen.

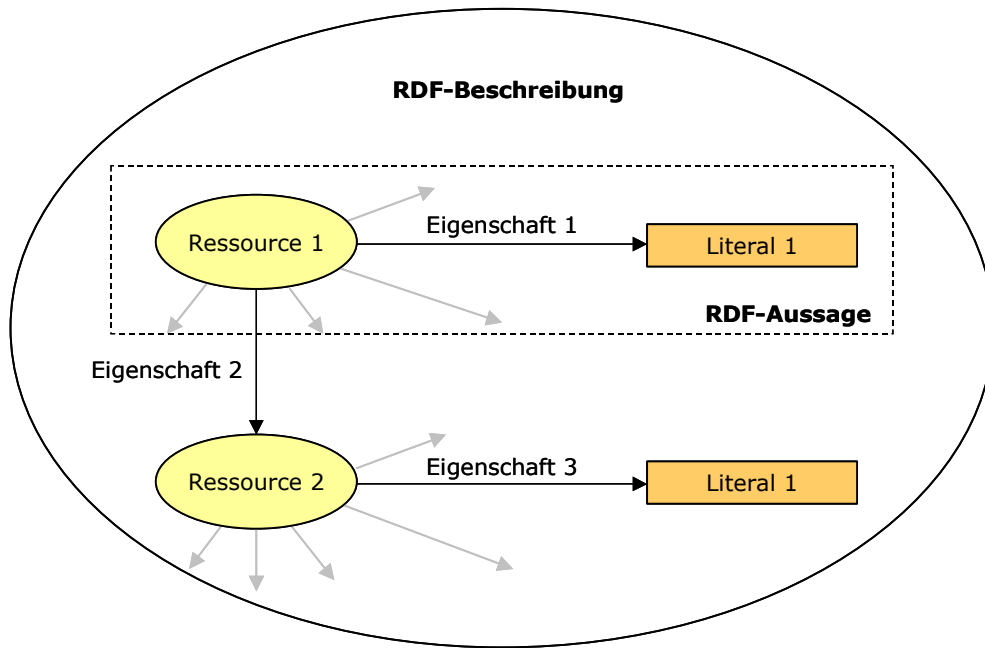


Abb. 5.3 Das RDF-Datenmodell

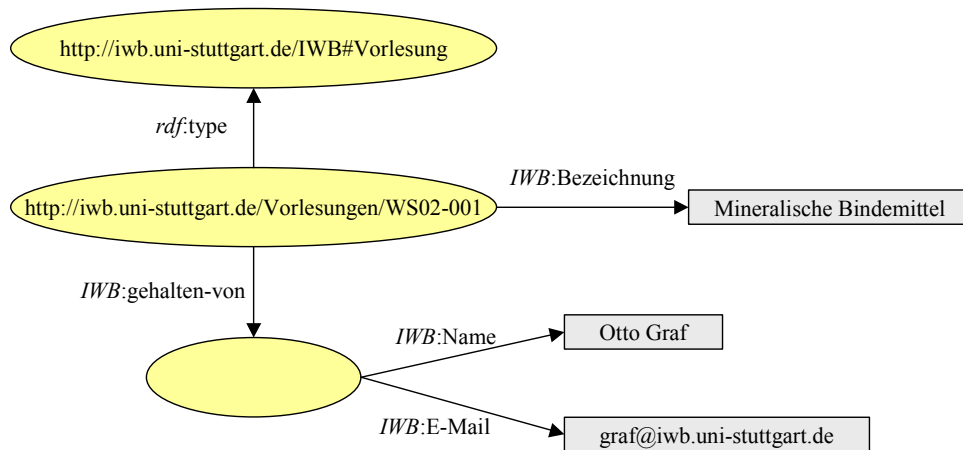


Abb. 5.4 RDF-Diagramm mit einer unbenannten Ressource

Die RDF-Syntax erschließt sich am anschaulichsten anhand von Beispielen. Die nachfolgende RDF-Beschreibung stellt eine Serialisierung des Modells aus Abbildung 5.4 dar. Beschrieben werden zwei Ressourcen: Die *Vorlesung WS02-001* und ein Individuum namens „*Otto Graf*“. Die Eigenschaft *gehalten-von* stellt eine Beziehung zwischen *WS02-001* und der Ressource her, durch welche die Person Otto Graf repräsentiert wird. Das Individuum wird durch zwei weitere atomare Werte (Literele) beschrieben, welche die Eigenschaften *Name* bzw. *E-Mail* darstellen.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/TR/WD-RDF-Syntax#"
  xmlns:IWB="http://iwb.uni-stuttgart.de/IWB#">

  <rdf:Description
    about=" http://iwb.uni-stuttgart.de/Vorlesungen/WS02-001">

    <rdf:type resource=" http://iwb.uni-stuttgart.de/IWB#Vorlesung">

    <IWB:Bezeichnung>Mineralische Bindemittel</IWB:Bezeichnung>

    <IWB:gehalten-von>
      <rdf:Description>
        <IWB:Name>Otto Graf</IWB:Name>
        <IWB:E-Mail>graf@iwb.uni-stuttgart.de</IWB:E-Mail>
      </rdf:Description>
    </IWB:gehalten-von>

  </rdf:Description>

</rdf:RDF>
```

Das XML-Element *rdf:RDF* signalisiert, dass dieses Element eine RDF-Beschreibung enthält. *rdf:Description* fasst eine Reihe von Aussagen über die Ressource zusammen, die durch das Attribut *about* spezifiziert wird. Die folgende Ergänzung zum obigen Beispiel besagt, dass die unbenannte Ressource in Zeile 13 *<rdf:Description>* vom Typ „*Person*“ ist:

```
<rdf:Description>
  <rdf:type
    resource="http://iwb.uni-stuttgart.de/IWB#Person"/>
  <IWB:Name>Otto Graf</IWB:Name>
  <IWB:E-Mail>graf@iwb.uni-stuttgart.de</IWB:E-Mail>
</rdf:Description>
```

Dabei wird unterstellt, dass die Beschreibung der Klasse „*Person*“ durch Verfolgen der URL *http://iwb.uni-stuttgart.de/IWB#Person* erhalten werden kann. Diese Klassenzuordnung wird im nächsten Abschnitt ausführlicher behandelt. Die obige Angabe ist übrigens äquivalent zu

```
<IWB:Person>
  <Name>Otto Graf</Name>
  <E-Mail>graf@iwb.uni-stuttgart.de</E-Mail>
</IWB:Person>
```

Diese leicht nachvollziehbare Syntax wird bereits seit langem für die semantische Auszeichnung der XML- bzw. SGML-Daten gebraucht. Mit RDF ist es jedoch möglich, dieser Beschreibung eine präzise und eindeutige logische Interpretation zuzuordnen.

Um die Zuweisung von mehreren Ressourcen oder Werten desselben Typs etwas zu verdeutlichen, wird in Abb. 5.5 eine ungeordnete Liste von drei Dozenten des Instituts für

Werkstoffe im Bauwesen repräsentiert. Hierbei wird die Verwendung der vordefinierten Eigenschaften (*rdf:_1*, *rdf:_2*, ..., *rdf:_n*) illustriert.

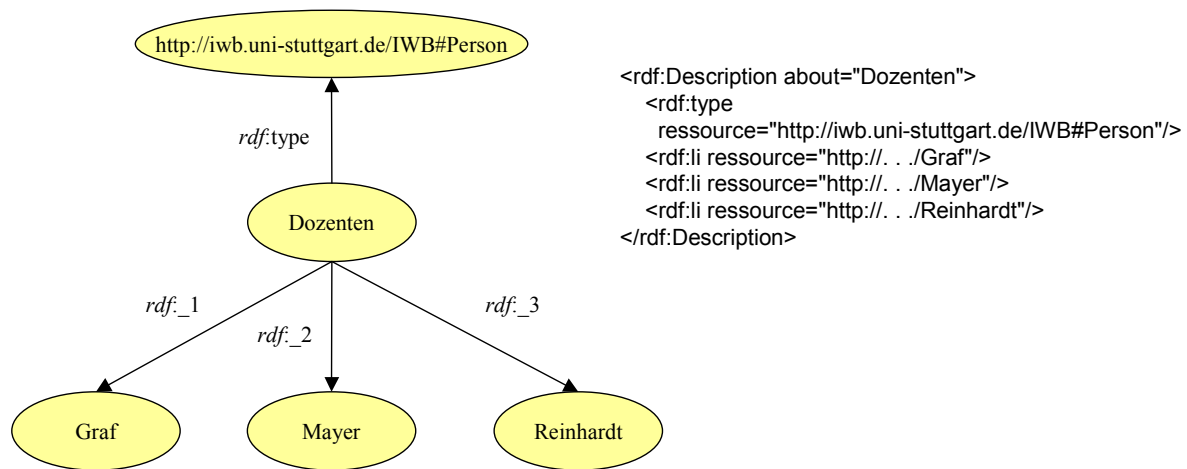


Abb. 5.5 Beispiel: Ressourcen- und Wertelisten in RDF

In der noch aktuellen Fassung der RDF Recommendation wird für die Modellierung solcher Listen die Verwendung von Container-Typen vorgeschlagen. Auf dem W3C-Server liegt jedoch ein vieldiskutierter Working Draft [RDF2001] vor, in welcher auf diese Container-Typen für unterschiedlich geartete Listen verzichtet werden soll. Da Container-Typen für die Anwendungen in dieser Arbeit keine Rolle spielen, wurde in den hier vorgestellten Beispielen von diesem Konstrukt kein Gebrauch gemacht.

Wie in Abschnitt 2.6 beschrieben, lassen sich vernetzte Informationssysteme aus strukturierten Informationsressourcen auch lediglich unter Verwendung des XML-Standards aufbauen. Bereits ohne RDF lässt sich das Datenmodell durch Namensräume von anderen XML-Datentypen abgrenzen. Für andere Anwendungen bleibt das Datenmodell jedoch undurchsichtig. Die Informationsobjekte mit ihren Attributen und Relationen können von Anwendungen, denen das verwendete Datenmodell nicht bekannt ist, maschinell nicht eindeutig identifiziert werden. Nicht einmal die Grenzen von Informationsobjekten können festgestellt werden. Betrachten wir beispielsweise folgende Personenbeschreibung:

```

<Person>
  <Name>Otto Graf</Name>
  <Adresse>
    <Ort>Stuttgart</Ort>
    <PLZ>70182</PLZ>
  </Adresse>
  <Forschungsschwerpunkte>
    <ul>
      <li>Mineralische Bindemittel</li>
      <li>Bruchmechanische Methoden</li>
      <li>...</li>
    </ul>
  </Forschungsschwerpunkte>
</Person>

```

Dank der Menschenlesbarkeit von XML können wir annehmen, dass *Name*, *Adresse* und *Forschungsschwerpunkte* vergleichbar zu den RDF-Eigenschaften sind und dass die Adresse als eine Art unbenannte strukturierte Ressource mit Attributen *Ort* und *PLZ* aufgefasst werden

kann. Obwohl der Inhalt der Forschungsschwerpunkte strukturell ähnlich aussieht, kann man bei den HTML nachempfundenen Elementen *ul* oder *li* kaum von Eigenschaften der Ressource *Forschungsschwerpunkte* sprechen. Automatisch lassen sich solche Zusammenhänge jedenfalls nicht feststellen.

Am obigen Beispiel wird die Aufgabe von RDF abermals betont. Bereits ohne jegliche Kenntnis der darunterliegenden Semantik können alle Ressourcen, Eigenschaften und atomaren Werte eines in XML serialisierten RDF-Datenmodells gesammelt, und die Beziehungen zwischen Ressourcen können automatisch aufgelöst werden. Die Bedeutung der Ressourcen und ihrer Beziehungen kann anhand von RDF-Schemata entschlüsselt werden, die im nächsten Abschnitt beschrieben werden.

5.4 RDF-Schemata

Das RDF-Modell erlaubt eine präzise, themenübergreifende und maschinell nachvollziehbare logische Erfassung von Informationen. Es bietet jedoch keine Hilfe bei der Wahl der Grundbegriffe und -beziehungen, die bei der Beschreibung bestimmter Sachverhalte benötigt werden. Somit können Computerprogramme zwar genau feststellen, worüber gesprochen wird, wissen aber nicht, was damit gemeint ist. Anhand der Nonsens-Aussage „die Beschafften von Reton mit gleichliertem Reschlag sind zuwertig eigenbar“ kann diese Interpretationsschwierigkeit veranschaulicht werden: So können wir zwar feststellen, dass es um Reton mit gleichlierten Reschlag“ geht. Das bringt den Benutzer aber nicht viel weiter.

Dieses Verständnis soll durch RDF-Schemata [RDFS2000] erreicht werden, die den Standardisierungsgrad einer W3C Candidate Recommendation (vgl. 2.6.2) besitzen. RDF-Schemata beschreiben Begriffsontologien, die in der Regel anwendungs- oder domänen-spezifisch sind. Damit wird ein Typisierungssystem für RDF-Modelle zur Verfügung gestellt. RDF-Schemata sind im Grunde genommen Entity-Relationship-Diagramme. Die Schema Definitionen selbst lassen sich als RDF-Modelle repräsentieren.

Für das korrekte Verständnis von RDF-Aussagen ist die eindeutige Interpretation ihrer Bestandteile, d.h. der Bedeutung der Prädikate und der beteiligten Ressourcen, ausschlaggebend. In einem globalen Medium wie dem WWW kann man sich nicht darauf verlassen, dass z.B. alle potentiellen Kommunikationspartner mit dem „Dozent“ eine identische Bedeutung verknüpfen. Damit die Kommunikation nicht scheitert, sollte man so präzise wie nur möglich sein. Jeder verwendete Begriff sollte demzufolge einen Hinweis darauf enthalten, wo seine Definition zu finden ist. RDF-Schemata sind nichts anderes, als ein „Wörterbuch“, in dem eine Menge von Begriffen samt ihrer Gebrauchsanweisung beschrieben werden kann. Der Namensraum-Mechanismus von XML erlaubt es, jedem Begriff das zugehörige Schema zuzuordnen. Eine menschen- und maschinenlesbare Beschreibung des entsprechenden RDF-Schemas kann durch Auflösung der Namensraum URIs gewonnen werden. Im folgenden Beispiel kann das IWB-Schema von <http://iwb.uni-stuttgart.de/IWB> heruntergeladen werden:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/TR/WD-RDF-Syntax#"
  xmlns:IWB="http://iwb.uni-stuttgart.de/IWB#">
  <IWB:Dozent>
    <Name>Otto Graf</Name>
    <E-Mail>graf@iwb.uni-stuttgart.de</E-Mail>
  </IWB:Dozent>
</rdf:RDF>
```

Das RDF Grundschemata (abgekürzt durch RDFS), das in [RDFS2000] definiert wird, spezifiziert eine Reihe von RDF-Ressourcen und -Eigenschaften, die Begriffe wie „Klasse“

und „Eigenschaft“ präzisieren. Diese Grundbegriffe dürften für sehr viele Datenmodelle von Bedeutung sein. Im Unterschied zu objektorientierten (OO-) Programmiersprachen [Hartmann2000] wie z.B. Java oder C++, die sich primär auf Klassen konzentrieren, verfolgt RDF einen eigenschaftsbasierten Ansatz. In OO-Sprachen wird typischerweise für jede Klasse eine Menge von Attributen bzw. Relationen festgelegt, die in Klasseninstanzen verwendet werden dürfen. Für eine Klasse „Buch“ könnten beispielsweise Attribute „Autor“, „Titel“ und „Erscheinungsjahr“ definiert werden. Dieser Ansatz funktioniert gut, wenn das Datenmodell bereits im Voraus konzipiert werden kann.

Um eine höhere Flexibilität der Modellierung zu ermöglichen, konzentrieren sich RDFS auf Eigenschaften. Statt eine Klasse durch ihre Eigenschaften zu definieren, wird eine Eigenschaft durch Klassen beschrieben, die sie verknüpfen kann. Zum Beispiel lässt sich festlegen, dass die Eigenschaft „Erscheinungsjahr“ die Klasse „Buch“ als Definitionsbereich und die Klasse „ISO-Datum“ als Wertebereich hat. Der Vorteil eines eigenschaftsbasierten Ansatzes besteht darin, dass es sehr einfach ist, über bestehende Ressourcen etwas auszusagen. Und dies zählt zu den Grundprinzipien des WWW.

Das Grundschemata erweitert die elementare Typisierung von Ressourcen, die im RDF-Modell durch die Eigenschaft *rdf:type* realisiert wurde. Neben den Begriffen *Class* und *Property* werden im Grundschemata transitive Relationen *subClassOf* und *subPropertyOf* sowie eine kleine Anzahl weiterer Ressourcen und Eigenschaften eingeführt. Außerdem wird ein Ausdrucksmittel zur Charakterisierung und Restriktion von Definitions- und Wertebereichen eingeführt. Das Grundschemata dient als ein fundamentales Typensystem für RDF mit dessen Hilfe komplexe Ontologien entworfen werden können. Die Definition von Datentypen wie den ganzen Zahlen, den Fließkommazahlen, Datumsformaten usw. findet im Rahmen des RDF Grundschematas nicht statt. Diese sollen von den zuständigen Standardisierungsgremien, wie z.B. einer Task Force des W3C, der ISO oder der DIN, in entsprechenden Namensräumen definiert werden.

Das folgende Schema beschreibt eine Klasse *Person* und eine Eigenschaft *Geburtsdatum*:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/TR/WD-RDF-Syntax#"
  xmlns:rdfs="http://www.w3.org/TR/WD-RDF-Schema#">

  <rdfs:Class rdf:ID="Person">
    <rdfs:comment xml:lang="de">Klasse für die
      Repräsentation von Menschen</rdfs:comment>
    <rdfs:subClassOf
      rdf:resource="http://www.erde.org/fauna#Saeugetier"/>
  </rdfs:Class>
  <rdfs:Property rdf:ID="Geburtsdatum">
    <rdfs:comment xml:lang="de">Das Geburtsdatum einer Person</rdfs:comment>
    <rdfs:domain rdf:resource="#Person"/>
    <rdfs:range
      rdf:resource="http://www.w3.org/TR/xmlschema-datatypes/#date"/>
  </rdfs:Property>

</rdf:RDF>
```

Hierbei ist *Person* eine Unterklasse von *Säugetier*. Der Definitionsbereich von *Geburtsdatum* ist *Person*, der Wert soll ein Datum analog der XML Schema Recommendation Teil 2 sein. In dieser Recommendation wird dann auf das in ISO 8601 festgelegte JJ-MM-DD Format verwiesen. Die beiden Definitionsbereiche werden jeweils durch die Eigenschaften *rdfs:domain* und *rdfs:range* spezifiziert, die im RDFS-Grundschemata definiert sind. Zu

beachten ist, dass die Angabe der Kodierung (ISO-8859-1) für die korrekte Erkennung der Umlaute für dieses Beispiel notwendig ist. Das zugehörige grafische RDF-Modell für das obige Schema-Beispiel wird in Abbildung 5.6 dargestellt.

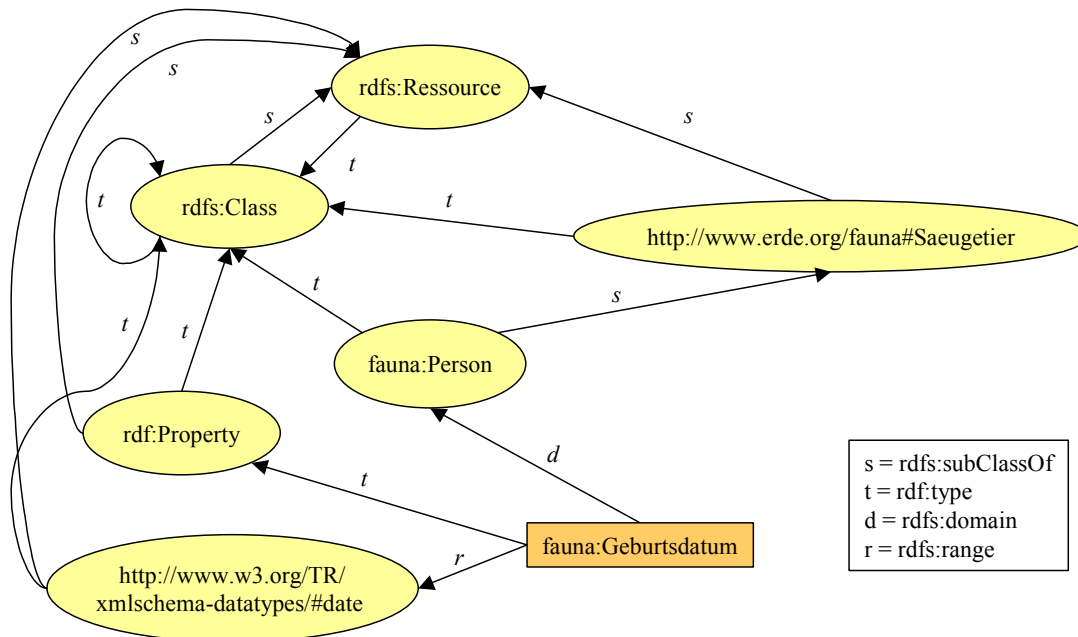


Abb.5.6 Beispiel: RDF-Schema (zur besseren Übersichtlichkeit ohne `rdfs:comment`)

5.5 Interoperabilität durch RDF

Mittels RDF kann das bereits in der Einführung angesprochene Problem der semantischen Heterogenität zum Teil entschärft werden. Namenskonflikte, die durch Homonyme, d.h. gleichnamige Begriffe mit verschiedenen Bedeutungen, verursacht werden, können bereits durch die Verwendung von Namensräumen eliminiert werden. Gleichnamige Begriffe, die in verschiedenen Namensräumen auftreten, werden als unterschiedlich angesehen. Stimmen auch die Namensräume überein, handelt es sich um gleichbedeutende Begriffe.

Demgegenüber können Synonyme, also Eigenschaften oder Klassen, die trotz verschiedener Bezeichnungen die gleiche Bedeutung tragen, in unabhängig entwickelten Ontologien durchaus entstehen. Das folgende Beispiel versucht zu zeigen, wie dieser Namenskonflikt in RDF behandelt werden kann.

Eine große Anzahl von Informationssystemen, die in Personalverwaltungen, Bibliotheken, Banken, Versicherungen usw. eingesetzt werden, verwenden personenbezogene Informationen. Die Modellierung des Datenobjekts „Person“ erfolgt dabei in den einzelnen Systemen zumeist auf lokal unterschiedliche Art und Weise. Was macht eine Person aus? Kann man anhand einer festgelegten (Mindest)Anzahl von Attributen wie Name, Geburtsdatum oder Nummer des Personalausweises den Datentyp *Person* zufriedenstellend beschreiben, so dass dieser im gesamten Anwendungsgebiet übergreifend eingesetzt werden kann? Anwendungsübergreifend bedeutet in diesem Fall, dass jedes in Frage kommende Programm bei einem gegebenen Informationsobjekt feststellen kann, dass es sich dabei um eine Person handelt, auch wenn dieses durch einen unterschiedlichen Namen repräsentiert wird. Es soll also das Objekt nicht nur identifizieren und von anderen unterscheiden, sondern auch seine Bedeutung ermitteln können.

Während konventionelle objektorientierte Ansätze es erforderlich machen würden, im Voraus alle brauchbaren Eigenschaften von Informationsressourcen zu definieren [Hartmann2000], bietet RDF hier ganz andere Möglichkeiten. Durch RDF-Schemata lassen sich beispielsweise Eigenschaften *Geburtsdatum* und *Adresse* festlegen, welche als Definitionsbereich die Klasse *Person* haben. Eine Anwendung, die zusätzlich die Sozialversicherungsnummer der Person benötigt, kann im eigenen Namensraum in ähnlicher Weise eine Eigenschaft *Sozialversicherungsnummer* definieren. Dadurch können zunächst viele Synonyme entstehen, wie z.B. *BfA:Sozialversicherungsnummer* oder *LVA:SV-NR*. Sollte sich aber der Begriff *Sozialversicherungsnummer* etablieren, könnte die Unterklassen-Zugehörigkeit der jeweiligen Eigenschaften zu einer Standardeigenschaft wie z.B. *DIN:Sozialversicherungsnummer* nachträglich spezifiziert werden. Bereits existierende Daten werden dadurch nicht betroffen. Eine Krankenkassen-Anwendung kann nun aber feststellen, dass mit *LVA:SV-NR* eine spezielle *DIN:Sozialversicherungsnummer* gemeint ist. Im Laufe des Evolutionsprozesses können sich populäre Ontologien herauskristallisieren, genauso, wie ein beliebter Pfad durch ein Feld entsteht. Die durch den RDF Ansatz gewährte Autonomie beim Design der Ontologien ist dabei ausschlaggebend.

Bisher wurde hauptsächlich auf die Namenskonflikte eingegangen. Semantische Heterogenität umfasst allerdings noch eine Fülle von anderen Schwierigkeiten. Datenkonflikte sind hierfür ein weiteres Beispiel. Bei einer Bestelltransaktion wird der Preis eines Artikels benötigt. Die Eigenschaft *Preis*, die in unserem Beispiel vom fiktiven Baustoffhersteller eBAU definiert wurde, wird von einem vordefinierten Begriff *Price* abgeleitet, wodurch ihre Bedeutung identifiziert werden kann. Für eine englischsprachige eCommerce-Applikation, die diese Daten erhält, ist jedoch nicht ohne weiteres erkennbar, um welche Währungseinheiten es sich bei der Preisangabe handelt. Diese Unklarheit kann wie folgt beseitigt werden:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/TR/WD-RDF-Syntax#"
  xmlns:eBau="http://www.ebau.de/"
  xmlns:IWF="http://iwf.org/">

  <eBau:Bestellung>
    <Artikel>Zement CEM I 42,5R</Artikel>
    <Preis
      rdf:value="20"
      IWF:units="http://iwf.org/Currencies/EUR"/>
  </eBau:Bestellung>

</rdf:RDF>
```

Dabei bezieht sich der Preis nicht auf einen atomaren Wert, sondern auf eine intermediäre unbenannte Ressource, welche die Eigenschaften *rdf:value* und *IWF:units* besitzt. Das obige Beispiel stellt einen typischen Fall einer mehrstelligen Relation dar. An dieser Relation nehmen eine Ressource vom Typ *eBau:Bestellung*, der Literal *20* sowie die Ressource *http://iwf.org/Currencies/EUR*, welche die Euro-Währung darstellt, teil.

Auch das Problem der Heterogenität bezüglich der Abfragesprachen und Protokolle könnte in Angriff genommen werden, indem die Datenabfragesprachen, wie z.B. SQL und Protokolle in RDF umformuliert werden. Beispiele hierfür werden in [Melnik1999] beschrieben. Die syntaktische Denkweise wird hier in Richtung der semantischen Modellierung verlagert. Auf diese Weise können abstraktere Sprachen entwickelt werden, die sich universeller in unterschiedlichen Softwareumgebungen anwenden lassen. An einer generischen Anfragesprache für RDF-Modelle wird zur Zeit intensiv gearbeitet (siehe z.B. [RDFQuery1998, RQL2001]).

Abschließend soll der Zusammenhang zwischen RDF und XML noch einmal verdeutlicht werden. Wie RDF ermöglicht auch XML den strukturierten Austausch und die Manipulation von Daten, weist aber gegenüber anderen Datenformaten eine Reihe von Vorteilen auf. XML ergänzt RDF insofern, als es als Serialisierung für RDF-Datenmodelle verwendet werden kann. Dadurch vererben serialisierte Datenmodelle die positiven Eigenschaften von XML (vgl. Abschnitt 2.6). XML eignet sich in erster Linie zum Einsatz in gut abgrenzbaren Problembereichen. Folgende Funktionalität wird allerdings erst durch Verwendung von RDF möglich [Melnik1999]:

- Beschreibung von Datenschemata und komplexen Ontologien bzw. Modellierung eines Weltausschnitts. In XML können dagegen nur die Instanzen der Datenmodelle repräsentiert werden. Durch XML-DTDs oder -Schemata können nur eingeschränkte strukturelle (hauptsächlich hierarchische) Beziehungen spezifiziert werden (vgl. Abschnitt 2.6).
- Integration von unabhängig entwickelten Ontologien.
- Maschinelle Erschließung der Ressourcen (Datenobjekte) und Eigenschaften (Beziehungen). In XML können die Grenzen und Beziehungen zwischen Informationsressourcen im allgemeinen nicht maschinell ermittelt werden.
- Maschinelle Erschließung der Semantik der Objekte und ihren Beziehungen zueinander (durch RDF-Schemata). Die Semantik von XML-Elementen kann dagegen nur vom Menschen aus einer aussagekräftigen Bezeichnung heraus abgeleitet werden und eventuell aus den in DTDs enthaltenen natürlichsprachlichen Beschreibungen. Maschinenlesbarkeit von XML besteht in der verlustfreien Extraktion aller in XML vorliegenden Daten.
- RDF ist logisch vollständiger als XML. In RDF lassen sich Aussagen über Aussagen sowie über die Ontologien selbst treffen. Um eine ähnliche Funktionalität in XML oder in anderen Modellen, die in objektorientierten Datenbanken und Programmiersprachen verwendet werden, zu erreichen, müsste man ein zusätzliches Metamodell über diese Sprachen legen.

RDF stellt ein zugängliches, erweiterbares und vielseitig anwendbares Instrument zur Verfügung, mit dessen Hilfe Kooperation zwischen Anwendungen gefördert werden kann. Es bietet einen erweiterbaren Rahmen für ontologische Integration von Wissensbereichen, Organisationen, Institutionen und Applikationen.

5.6 Anwendungsgebiete von XML/RDF

Wie schon weiter oben angesprochen, ist anzunehmen, dass XML die heutige Web-Sprache HTML in vielen Bereichen ablösen wird. XML wird als die Zukunft der Web-Entwicklung betrachtet, die sich im Evolutionsprozess der Datenformate entwickelt hat [Behme2000]. XML-basierte Sprachen zum Setzen von mathematischen Formeln [MathML2001] und zur Datenauszeichnung in der Chemie [CML2002] (s.a. Abschnitt 2.6.3) finden allmählich ihren Eingang in die Praxis.

Vielversprechend ist auch die Anwendung von XML auf andere strukturierte Dokumente, wie z.B. Gesetzestexte oder Dokumentationen. Hypertextuelle inhaltliche Aufbereitung von Daten kann besonders für Lern-, Diagnose- und Expertensysteme von Vorteil sein [Maurer1993]. Die Trennung der Daten von der Ausgabeformatierung ermöglicht die in Abschnitt 3.3 geforderte nutzersensitive Präsentation, die durch eine semantische Auszeichnung der Inhalte durch RDF-Beschreibungen noch erweitert wird. So kann z.B. eine Sprachwahl unterstützt werden und die Benutzerinteraktion nach den Präferenzen, dem Kenntnisstand usw. der Anwender individuell gestaltet werden.

Besondere qualitative Verbesserungen erhofft man sich von XML/RDF für die webbasierte Datenbanktechnologie [Florescu1998]. Eine explizite Strukturauszeichnung erübrigt die aufwendige Konvertierung der Information in maschinenlesbare Form, z.B. bei der automatischen Auswertung der Ergebnisse, die von einer Suchmaschine geliefert werden. Beim Herantasten an das Problem der semantischen Integration kann insbesondere durch RDF ein Durchbruch erzielt werden. Client-basierte Verarbeitung begünstigt die Rechenlastverteilung und Skalierbarkeit von Anwendungen und sie kann zur Effektivitätssteigerung von Design-Anwendungen in CAD/CAM, Architektur u.a. beitragen. Intelligente Agenten, die nach persönlichen Wünschen strukturierte Informationen zusammensuchen, TV-Programme oder maßgeschneiderte Zeitungen erstellen oder die günstigsten Immobilien oder Autos finden, sind ein weiterer Einsatzbereich für XML/RDF.

Eine Vorreiterrolle bei der Entwicklung der neuen Web-Sprache spielt das Bibliothekswesen, angetrieben durch die Notwendigkeit, Metadaten über digitale Objekte zu verwalten, auszutauschen und langfristig zu archivieren. Der als RDF-Schema modellierte Dublin Core-Metadatenstandard [Dublin] wurde einer der ersten XML basierten Standards auf diesem Gebiet.

Eine bunte Palette von Anwendungsbereichen wie verteilte Dokumentenverwaltung, Workflow-Systeme, Management des intellektuellen Eigentums, automatische Ressourcenerforschung und viele andere können von XML/RDF profitieren. XML und RDF sollen die sklavische manuelle Arbeit bei Austausch und Manipulation von Wissen eliminieren [Dertouznous1997].

5.7 Aktuelle Softwaretools

Dieser Abschnitt enthält einen Überblick über vorhandene Softwareapplikationen zur Entwicklung und Nutzung von RDF. Einige der hier vorgestellten Softwareprogramme wurden zur Erstellung des in der Arbeit vorgestellten Demonstrationssystems herangezogen. Es handelt sich dabei fast durchweg um Programme aus dem universitären Bereich, die als Open Source Software im WWW zum Herunterladen zumeist kostenlos zur Verfügung stehen. Als Programmiersprache wurde Java verwendet oder eine Software mit serverseitigen Programmiertechnologien erstellt.

Leider ist der Funktionsumfang der Programme bisher noch sehr eingeschränkt. So erzeugen die meisten der hier vorgestellten Applikationen zur Abfrage von RDF-Beschreibungen als Ergebnis eine Reihe von Aussagen in Form von Tripel (vgl. Abschnitt 5.2), die in Programmierumgebungen für das logische Programmieren (z.B. Prolog) weiterverarbeitet werden können. Die Erstellung von Anwenderoberflächen in konventionellen Programmierumgebungen wird jedoch noch kaum unterstützt. Aufgrund des starken Engagements der großen Softwareentwickler in der RDF-Entwicklung ist nach Abschluss der Standardisierung des RDF-Schema Standards mit einer Verbesserung der Anwenderfreundlichkeit der angebotenen Softwareprodukte zu rechnen.

5.7.1 RDF-Parser

Der Parser ist die zentrale Programmkomponente in RDF-Applikationen. Seine Aufgabe ist die Erschließung der in XML/RDF-Syntax geschriebenen Datenmodelle für Anwendungen, in dem die Aussagen der Syntax in Tripel überführt werden. Aufgrund seiner Stellung im Mittelpunkt der Softwarearchitektur von RDF-Anwendungen und der Tatsache, dass RDF-Parser zumeist selbst auf XML-Parsern (vgl. Abschnitt 2.6.2) aufbauen, wird diese Komponente in einigen Veröffentlichungen [vgl. Henning1999] auch als RDF-Prozessor bezeichnet.

Während der Erschließung überprüfen aktuelle Parser den RDF-Code auf syntaktische Korrektheit. Dieser Vorgang wird, analog zu entsprechenden XML-Parsern, als Validierung bezeichnet. So wurden alle in der Dissertation enthaltenen RDF-Beispiele mit dem öffentlich zugänglichen, in Java geschriebenen ARP (Another RDF-Parser) von J. Carroll aus den HP-Labs in Bristol getestet. Auf der entsprechenden Seite [RDFValid2001] der W3C Web-Site wurde zusätzlich das Visualisierungs-Programm GraphViz der AT&T Labs an den Parser angebunden. Dieses Java-Programm erzeugt aus den Tripel grafische RDF-Modelle, die sich leichter überprüfen lassen.

Der ARP-Parser stammt aus dem als Jena Java RDF API and toolkit bezeichneten Entwicklerpaket der HP-Labs. Jena enthält zusätzlich eine Programmierschnittstelle, einen RDF/XML Writer sowie eine als RDQL bezeichnete Abfragesprache, mit der RDF-Modelle in einer Datenbank abgefragt, bearbeitet und für objektorientierte Java-Methoden zugänglich gemacht werden. Weitere Parser sind der CARA der mathematischen Fakultät an der Universität Osnabrück [Cara], der Raptor von Dave Beckett [Beckett2001] sowie SIRPAC [Reggiori2001], dessen Weiterentwicklung allerdings inzwischen eingestellt wurde.

5.7.2 RDF-Editoren

RDF-Editoren unterstützen den Anwender bei der Erzeugung und Bearbeitung der RDF-Beschreibungen. Allgemeine RDF-Editoren, wie z.B. RDFedt (RDF-Editor for Windows) [Winkler2001] von J. Winkler, erleichtern die Syntaxerstellung durch eine grafische Oberfläche zur Erstellung der XML/RDF-Elementbäume. Sie unterstützen in den vorliegenden Versionen den Anwender jedoch nicht ausreichend hinsichtlich einer selbsterklärenden, einfachen Werteeingabe für diese Elemente.

Spezifische RDF-Editoren, die sich in ihrer Bedieneroberfläche an bestimmten RDF-Schemata orientieren, eignen sich daher eher zur Erfassung größerer Datenmengen in RDF. Beispiele für derartige Editoren sind z.B. DC-dot (vgl. Abb. 4.2) [UKOLN2000] zur Erzeugung von Metadatensätzen im Dublin Core Metadatenformat und auch der in Abschnitt 6.3 beschriebene met@BiM-Editor (vgl. Abb. 6.9). Spezifische Editoren liegen jedoch häufig als Online-Applikationen im Web vor und erlauben aufgrund der Sicherheitsbeschränkungen für Webanwendungen keine Bearbeitung von lokal auf dem Benutzerrechner gespeicherten Datensätzen. Sie dienen daher meist zur Erfassung von Metadaten, während die daraus erzeugte RDF-Syntax anschließend in einem separaten Fenster angezeigt und über Kopieren/Einfügen in die Dateien eingefügt werden muss.

5.7.3 RDF-Datenbanken

RDF-Datenbanken dienen der Speicherung bzw. der Verwaltung von RDF-Datensammlungen und bieten mit einer zumeist systemspezifischen Abfragesprache auch eine Zugriffsmöglichkeit auf die gespeicherten Inhalte. Sie stellen damit schon recht komplexe Programmpakete zur Nutzung von RDF-Datenstrukturen dar. Die Firma Intellidimension entwickelt mit RDF Gateway für MS Windows NT/2000 [Intellidimension] ein Softwarepaket für semantische Abfragen auf verteilte RDF-Datensätze mittels einer SQL-ähnlichen Abfragesprache (RDFQL). Die Software liegt zur Zeit als Betaversion zum kostenfreien Download vor und bietet eine grafische Oberfläche sowie Schnittstellen zu Serververwaltungs- und Datenbanktools von MS Windows. Ein vergleichbares Projekt für die Linux Plattform ist RDFDB [Guha] von R.V. Guha. Bei beiden Softwarepaketen besteht die Ergebnisausgabe von Abfragen in Tripel.

5.7.4 Implementation in Mozilla / Netscape Communicator

Die wesentlichen Anwendungsgebiete des RDF sind mit dem universellen Informationsraum World Wide Web verbunden. Demzufolge scheint es naheliegend, dass auch die heute am weitesten verbreiteten Softwaresysteme, die WWW-Browser, eine entsprechende Unterstützung anbieten. Die Firma Netscape hat große Anteile an der Entwicklung des RDF. Einen entsprechend großen Stellenwert hat RDF in der zukünftigen Version des WWW-Browsers von Netscape. Mit der neuen Komponente NavCenter verspricht man sich die Verbindung von Informationen aus dem Internet und dem Desktop. RDF dient dabei als Basistechnologie, um mit Informationen aus verschiedensten Quellen, wie Webseiten, lokalen Office-Dokumenten, Emails, Lesezeichen, Datenbanken etc. einheitlich umgehen zu können. Die Zusammenführung von lokalen und entfernten Informationen werden zum Beispiel genutzt, um Lesezeichen zentral zu administrieren aber gleichzeitig den Wünschen des Benutzers zu entsprechen. Durch die Integration von RDF wird Netscape zudem in die Lage versetzt, Dienste wie Platform for Privacy Preferences (P3P), digitale Unterschriften (DSig) oder Platform for Internet Content Selection (PICS) den Benutzern zugänglich zu machen.

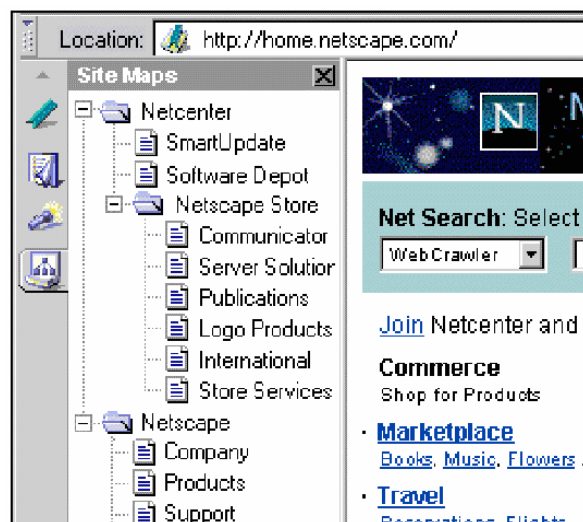


Abb. 5.7 Sitemaps Funktionalität in einer frühen Version des Mozilla-Browsers

Bereits der aktuelle Communicator (ab Version 4.06) nutzt RDF für das sogenannte Smart Browsing, um mit der aktuellen Webseite thematisch assoziierte Dokumente zur Verfügung zu stellen.

5.7.5 Implementation in aktuellen Suchmaschinen

Die automatische Erfassung und Integration von in Webseiten integrierte RDF-Beschreibungen wird von heutigen Suchmaschinen noch nicht unterstützt. Trotzdem spielt RDF bereits heute eine bedeutende Rolle bei der Katalogisierung von Webinhalten innerhalb der wichtigsten Suchdienste im WWW.

Beim Open Directory Project (ODP) [DMOZa] handelt es sich nach eigenen Angaben um das größte und vollständigste manuell erstellte Webverzeichnis im WWW. Im Gegensatz zu den kleinen Stäben von professionellen Online-Redakteuren anderer Webkataloge, wie z.B. Yahoo!, versucht man im Rahmen des ODP mit einer großen Gruppe freiwilliger Redakteure, qualitativ hohe Webinhalte zu bestimmten Themenbereichen zu erfassen und die Verweise

aktuell zu halten. Die Philosophie des seit Ende der 90er Jahre bestehenden Projekts gründet, ähnlich der erfolgreichen Open Source-Bewegung, auf gegenseitiger Hilfe: ODP-Redakteure sind Recherchierende im Web, die eine Verzeichnisstruktur für andere Recherchierende erstellen.

Die erstellten Zuordnungen von Webinhalten in Verzeichnisstrukturen zu einzelnen Themenbereichen werden im RDF-Format abgespeichert und stehen Nutzern kostenlos zur Verfügung [DMOZb]. Der ODP-Webkatalog ist die Grundlage der Webverzeichnisstrukturen der populärsten Portale und Suchmaschinen, wie z.B. AOL Search [AOL2001], Netscape Search [Netscape2001], Google [Google2002], Lycos [Lycos2002], HotBot [HotBot2002] u.v.a.. In Google z.B. werden bei Seiten, die auch in ODP gelistet sind, Verweise auf das entsprechende ODP-Verzeichnis eingeblendet. Dadurch kann der Anwender rasch zu thematisch verwandten Webinhalten gelangen. Zur Nutzung der ODP-Daten auf privaten Sites oder Firmenportalen stehen zahlreiche kommerzielle und kostenlose Programme im Web zur Verfügung.

6 RDF-Anwendungsszenarien zu „Beton mit rezyklierter Gesteinskörnung“

Zur Nutzung der in Abschnitt 4.3 zunächst syntaxunabhängig formulierten met@BiM-Metadatenstrukturen in einer konkreten Webanwendung müssen diese in eine Syntax überführt werden. Für diesen Schritt wird im Rahmen dieser Arbeit das im vorhergehenden Abschnitt erläuterte Resource Description Framework herangezogen. Anhand dieses Wissensrepräsentationsformalismus lassen sich semantische Datenmodelle konstruieren, mit denen die inhaltlichen Aussagen von betontechnologischen Webinhalten beschrieben werden können. Das Vorgehen hierbei sowie der Aufbau solcher Datenstrukturen wird an Beispielen vorgestellt.

Um die Anwendung der erzeugten RDF-Datenstrukturen an praxisnahen Szenarien demonstrieren zu können, wurde weiterhin eine entsprechende Softwareumgebung entworfen. Der Anwender muss schon bei der Erzeugung des bezüglich seiner Syntax zumeist sehr umfangreichen RDF-Codes von Softwarewerkzeugen unterstützt werden. Für diese Aufgabe wurde zunächst ein Editor-Programm entwickelt, welches mit einer formularartigen Eingabeoberfläche den Eingabeaufwand für den Autor deutlich reduziert. Die Erzeugung des RDF-Codes erfolgt automatisch durch das Programm. Aufgrund der Abhängigkeit zwischen Datenstruktur und den beschriebenen Inhalten musste für met@BiM ein spezieller RDF-Editor entwickelt werden.

Ferner wurden Komponenten zur weitgehend automatisierten Online-Datenerfassung, der Datenspeicherung und eine Abfrageoberfläche entwickelt. Auf der Grundlage dieser Softwarearchitektur wurden das met@BiM-Datenmodell und ausgewählte RDF-Beschreibungen von Webinhalten zu rezykliertem Beton in das Testsystem implementiert. Anhand einiger Anwendungsszenarien lässt sich der Nutzen von Metadatenstrukturen für das ingenieurgerechte Erschließen baustoffbezogener Informationen im Web aufzeigen.

6.1 RDF-Schema für met@BiM

Die Nutzung des in Kapitel 4 entwickelten Metadatendatenmodells innerhalb von Webanwendungen setzt dessen Codierung in einer Syntax voraus. Aufgrund der in Abschnitt 2.6 sowie 5 erläuterten Vorzüge von XML und RDF liegt es nahe, dass zur Formalisierung des met@BiM-Metadatenmodells RDF-Beschreibungen in XML-Syntax verwendet werden. Vor der Beschreibung konkreter Webinhalte müssen zunächst innerhalb eines RDF-Schemas die dafür benötigten Metadatenelemente sowie die Ressourcentypen aus der met@BiM-Ontologie in ihren Bedeutungen und Beziehungen zueinander definiert werden.

Alle diese Informationen über met@BiM lassen sich in einem RDF-Schema in maschinenlesbarer Form abbilden. Die Zugehörigkeit der Metadatenelemente und Ressourcen zu einem RDF-Schema wird dabei über das Namensraumpräfix „metabim“ festgelegt (vgl. Abschnitt 2.6.2). Ein RDF-Schema des met@BiM-Modells bildet somit die Grundlage für eine systemübergreifende Beschreibung von betontechnologischen Webinhalten. Webanwendungen, die auf RDF-Datensätze mit Elementen des „metabim“ Namensraums treffen, können das in einer XML-Datei codierte Schema dazu nutzen, auf die Beschreibungen der dort deklarierten Ressourcentypen (RDF Resources) und Metadatenelemente (RDF Properties) zuzugreifen. Aus den ebenfalls dort deklarierten Beziehungen der Eigenschaften zu bestimmten Ressourcen ist es entsprechenden Rechercheprogrammen auch möglich, automatisch den semantischen Kontext von Webinhalten zu bestimmen (vgl. Abschnitt 6.5).

Damit das RDF-Schema von unterschiedlichen Systemen genutzt werden kann, muss es im Web öffentlich zugänglich sein. Dies bedeutet, dass die XML-Datei, die das RDF-Schema enthält, eine stabile URI besitzen sollte. Die URI des met@BiM-RDF-Schemas wurde mit „<http://www.b-i-m.de/2001/01/metabim#>“ innerhalb der Webdomäne von BiM-Online

gewählt. Alle RDF-Beschreibungen, die Elemente des „metabim“-Namensraums enthalten, sollten zu deren korrekter Interpretation in RDF-Parsern auf diese URI verweisen. Gemäß den Vorgaben des W3C hinsichtlich der Fortschreibung von RDF-Schemata sollten bestehende Versionen nicht gelöscht werden, sondern den darauf verweisenden Programmen erhalten bleiben. Neue Versionen des Schemas werden in entsprechend bezeichneten Verzeichnispfaden unterhalb der Webserveradresse von BiM-Online veröffentlicht werden. Verweise auf die letzte bzw. jeweils aktuelle Version des Schemas wurden in den Kommentar am Anfang der in Anhang B vollständig abgebildeten XML-Datei integriert. Weiterhin enthält dieser Abschnitt Angaben zu Inhalt, der Version, Kontaktmöglichkeiten sowie früheren Versionen des Schemas.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!-- ISO-8859-1 Angabe für korrekte Interpretation der Umlaute -->
<!--

RDF-Schema declaration for ...
comments etc. to <schreyer@iwb.uni-stuttgart.de>

This Version: http://www.b-i-m.de/2001/01/metabim#
Latest Version: http://www.b-i-m.de/2001/01/metabim#
Previous Version: -
Editors:
Marcus Schreyer, University of Stuttgart

-->
<!DOCTYPE rdf:RDF>
<rdf:RDF xml:lang="de"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:dc="http://purl.org/dc/elements/1.1/"
>

<!-- Description of the Schema -->

<rdf:Description rdf:about="http://www.b-i-m.de/2001/01/metabim#">
  <dc:title>The met@BiM (metabim) Element Set v1.0</dc:title>
  <dc:creator>Marcus Schreyer</dc:creator>
  <dc:publisher>Institute of Construction Materials (IWB), University of
    Stuttgart (Germany)</dc:publisher>
  <dc:description>The met@bim metadata vocabulary is a simple vocabulary
    intended to facilitate discovery of resources concerning
    construction materials in the WWW. It's part of my work
    on a semantic description of the contents of BiM-Online
    (http://www.b-i-m.de)</dc:description>
  <dc:identifier>http://www.b-i-m.de/2001/01/metabim#</dc:identifier>
  <dc:language>DE</dc:language>
  <dc:relation>http://www.b-i-m.de</dc:relation>
  <dc:date>2002-02-01</dc:date>
</rdf:Description>

...

```

Nach den Kommentarzeilen zeigt das Element *rdf:RDF* an, dass es sich bei den folgenden Zeilen um einen RDF-Code handelt (vgl. Abschnitt 5.4). Das vordefinierte Element *xml:lang* sagt aus, dass die Inhalte des XML-Dokuments in deutscher Sprache verfasst sind. In XML lassen sich Dokumente auch mehrsprachig durch die Zuweisung alternativer Elementbezeichnungen und -definitionen verfassen. Ein zukünftiges, international verwendbares Metadatenmodell wäre daher mit verhältnismäßig geringem Programmieraufwand auf der Grundlage

eines bestehenden Modells zu realisieren. Die folgenden Elemente mit dem Präfix *xmlns:* deklarieren die innerhalb der Datei verwendeten Namensräume und ordnen ihnen das entsprechende Präfix zu.

Im nächsten Abschnitt wird zunächst die Schemadatei selbst unter Verwendung von Elementen des Dublin Core-Metadatenstandards beschrieben. Eine vollständige XML-Serialisierung des RDF-Schemas, in dem die Dublin Core-Metadatenelemente definiert werden, ist im Anhang C dieser Arbeit abgedruckt. Im Gegensatz zu den für menschliche Leser bestimmten Angaben im Kommentar handelt es sich bei der Beschreibung der *met@BiM*-Schemadatei mit DC-Metadaten um strukturierten XML-Code, der von entsprechenden Webanwendungen automatisch ausgewertet werden kann. Bei der Angabe der Werte für die beiden Elemente *dc:language* und *dc:date* mussten die im DC-Standard vorgegebenen Formatierungshinweise beachtet werden. So wird die im Schema verwendete deutsche Sprache mit dem ISO 639 [ISO639-2] konformen Kürzel *de* als Sprache angegeben. Weiterhin erfolgte die Angabe des Datums konform ISO 8601 [ISO8601] im internationalen numerischen Format JJJJ-MM-TT.

Nach diesen Zeilen beginnt die eigentliche Definition der *metabim*-Elemente mit der Deklaration der innerhalb des Schemas beschriebenen RDF-Ressourcen. Es handelt sich dabei um diejenigen Ressourcentypen, die in dem in Abbildung 4.3 beschriebenen Ontologiebaum dargestellt werden. Der Aufbau der Beschreibungen ist dabei recht gleichförmig: Zunächst erfolgt die Benennung der Ressource innerhalb des *ID* Attributs des RDF-Elements *Class*. Die RDF-Ressourcen innerhalb des *met@BiM*-Schemas wurden zur besseren Unterscheidung von den als RDF-Eigenschaften deklarierten Metadatenelementen in Großbuchstaben geschrieben.

Mit dem innerhalb der RDF-Schema-Spezifikation vordefinierten Element *rdf:subClassOf* (vgl. Abschnitt 5.4) lässt sich die in Abbildung 4.3 grafisch dargestellte hierarchische Struktur in serialisierter Form, d.h. hier mit dem XML/RDF-Code, nachbilden. *rdf:subClassOf* definiert dabei die aktuelle Ressource als Unterklasse der im Attribut *rdf:resource* angegebenen Ressource. Der Verweis auf Ressourcen erfolgt, wie in RDF üblich, über die URI der Ressource. Da RDF-Ressourcen, wie in den in dieser Arbeit gezeigten RDF-Schemata, normalerweise in XML-Dateien gespeichert werden, besteht die URI meist aus der URL der XML-Datei sowie dem Ressourcennamen nach dem Fragment Identifier (#). Die Ressourcen *DOKUMENT*, *KURS*, *MATHEMATISCHEFORMEL*, *CHEMISCHEFORMEL* sowie *BETON* besitzen innerhalb des *met@BiM*-Schemas keine übergeordneten Ressourcen und werden daher über *rdf:subClassOf* direkt der obersten Ressource *rdf:Resource* untergeordnet. Dadurch dass man in RDF-Schemata Ressourcen über mehrfache *rdf:subClassOf* Zuordnungen auch gleichzeitig an andere Schemata anbinden kann, bleibt die Möglichkeit einer Integration in zukünftige übergeordnete RDF-Standards für das Bauwesen grundsätzlich erhalten. Um z.B. die betontechnologischen Ressourcen und Properties als Unterklasse der Ressource *WERKSTOFFE* zu deklarieren, die in einem RDF-Schema mit der URL „<http://www.DIN.de/Bauwesen/bau-rdfschemata>“ definiert wird, genügt es im XML-Element *rdf:Class ID="BETON"* um einen weiteren *rdf:subClassOf* Eintrag zu ergänzen:

...

```
<rdf:Class ID="BETON">
  <rdfs:subClassOf
    rdf:resource="http://www.w3.org/2000/01/rdf-schema#Resource"/>
  <rdfs:subClassOf
    rdf:resource="http://www.DIN.de/Bauwesen/bau-rdfschemata#WERKSTOFFE"/>
  <rdfs:label>Betontechnologische Daten</rdfs:label>
  <rdfs:comment>In dieser Klasse werden Ressourcen und Properties zur
    Modellierung von Daten aus der Betonherstellung sowie
```

```

        Materialprüfungen zu Beton definiert
    </rdfs:comment>
</rdf:Class>

```

...

Auch wenn anderslautende Ressourcen mit identischer Bedeutung in solchen übergeordneten Standards definiert werden sollten, stellt das RDF-Schema-Konzept mit RDF-Eigenschaften vom Typ *rdfs:seeAlso* Relationen zur Verfügung, welche die Beibehaltung bestehender Datenstrukturen unter gleichzeitiger Einbettung in die neuen Schemata erlauben.

Für die *rdfs:subClassOf* Zuordnungen unterhalb der obersten Ebene der Ontologie verweist die URI auf Ressourcen innerhalb des met@BiM-RDF-Schemas. Die URI solcher Ressourcen wie *FESTBETON*, *FRISCHBETONPRUEFUNGEN*, *ZEMENT* etc. setzt sich daher zusammen aus der URL des met@BiM-RDF-Schemas <http://www.b-i-m.de/2001/01/metabim>, dem Fragment Identifier-Symbol # und dem folgendem Ressourcennamen.

Bei den beiden RDF-Eigenschaften *rdfs:label* und *rdfs:comment* handelt es sich um Elemente zur Dokumentation des RDF-Schemas. Die RDF-Eigenschaft *rdfs:label* erlaubt es, einem Element einen für menschliche Leser leichter verständlichen Namen zuzuordnen. Dies ist dann besonders sinnvoll, wenn es sich bei der gewählten Ressourcenbezeichnung um ein ansonsten unverständliches Akronym handelt. Innerhalb von *rdfs:comment* kann eine kurze Beschreibung der Ressource erfolgen, in welcher die Anwender des Schemas Hinweise auf die semantische Definition der Ressource bekommen können.

Eine besondere Rolle innerhalb der in met@BiM definierten Ressourcen spielt hierbei die Ressource *EINHEITEN*. Sie benötigt im Gegensatz zu den anderen Ressourcen keine *rdfs:subClassOf* Zuordnung, da sie lediglich die physikalischen und mathematischen Maßeinheiten und Dimensionen festlegt, die den Werten der met@BiM-Werkstoffeigenschaften zugewiesen werden dürfen. Bei den in dieser Arbeit vordefinierten Werten für dieses Metadatenelement handelt es sich, wie bereits in Abschnitt 4.3.3 erwähnt, um eine Behelfslösung. Entsprechende standardisierte Modelle zur allgemeineren Beschreibung von technischen Maßeinheiten sind jedoch bereits in Arbeit [bcXML2002].

...

```

<rdf:Class ID="EINHEITEN">
  <rdfs:label>EINHEITEN</rdfs:label>
  <rdfs:comment>Einheiten für Werte der met@BiM Properties</rdfs:comment>
  <EINHEIT rdf:ID="mm"/>
  <EINHEIT rdf:ID="cm"/>
  <EINHEIT rdf:ID="m"/>
  <EINHEIT rdf:ID="mm2"/>

```

...

Die Beschreibung der Betoneigenschaften bzw. der RDF-Eigenschaften des met@BiM-RDF-Schemas folgt auf die entsprechende Kommentarzeile *<!-- Deklaration der Properties -->*:

...

```

<!-- Deklaration der Properties -->

<rdf:Property ID="abschlaemmbarebestandteile">
  <rdfs:domain rdf:resource="#ZUSCHLAG"/>
  <rdfs:domain rdf:resource="#KIES"/>
  <rdfs:domain rdf:resource="#SAND"/>

```

```

    <rdfs:domain rdf:resource="#SPLITT"/>
    <rdfs:comment>
        Abschlämbbare Bestandteile (vgl. DIN 4226, Teil3)</rdfs:comment>
</rdf:Property>

<rdf:Property ID="alter">
    <rdfs:domain rdf:resource="#FESTBETON"/>
    <rdfs:comment>Alter der Probe</rdfs:comment>
</rdf:Property>

<rdf:Property ID="auszenbauteileignung">
    <rdfs:domain rdf:resource="#REZEPTUR"/>
    <rdfs:comment>
        Eignung der Betonmischung für Außenbauteile (vgl. DAfStb-Richtlinie
        "Beton mit rezykliertem Zuschlag")</rdfs:comment>
</rdf:Property>

...

```

Ähnlich den Ressourcendefinitionen erfolgt die Definition der Metadatenelemente für die Betoneigenschaften innerhalb eines XML-Tags mit der Bezeichnung *rdf:Property*. Die Eigenschaftsnamen, in met@BiM alle konsequent kleingeschrieben, werden dabei als *ID* Attribut der XML-Tags angegeben. Die Verknüpfung der Betoneigenschaften mit den Ressourcen, zu deren Beschreibung sie definiert wurden, erfolgt über das Tag *rdfs:domain* (vgl. Abschnitt 5.4). Befindet sich die Ressource, auf die *rdfs:domain* verweist, innerhalb der met@BiM-RDF-Schemadatei, kann auf die Angabe der kompletten URL verzichtet werden. In diesem Fall vereinfacht sich die URI im Attribut *rdfs:resource* auf den Ressourcennamen nach dem Fragment Identifier #. Das *rdfs:comment* Tag umschließt bei den RDF-Eigenschaften, wie schon bei den Ressourcendefinitionen, eine Erläuterung der im Element definierten Werkstoffeigenschaft. Bei einigen Eigenschaften, wie z.B. dem Anteil an abschlämbbaren Bestandteilen, zu deren Beschreibung die Angabe einer zugrundeliegenden Vorschrift notwendig ist, wird hier auf die entsprechenden Normen und Richtlinien verwiesen. Durch solche Verweise wird die Vergleichbarkeit der innerhalb des Elements angegebenen Werte sichergestellt. Grundsätzlich wurde jedoch versucht, die Baustoffeigenschaften möglichst allgemein zu definieren, um das Modell nicht zu unflexibel gegenüber neuen Prüfverfahren oder Normänderungen werden zu lassen. Sollten einzelne Materialeigenschaften eine weitere Unterteilung der met@BiM-Betoneigenschaften in spezifischere Elemente erfordern, ist dies bei Bedarf nachträglich immer noch möglich.

Am Ende des RDF-Schemas folgen schließlich RDF-Elemente zur Modellierung der Werte der Betoneigenschaften. Mit diesen lassen sich sowohl einfache Wertzuweisungen vom Typ *Druckfestigkeit ist 21.5* als auch sogenannte „strukturierte Werte (structured values)“ erzeugen. Bei strukturierten Werten wird einer Eigenschaft dabei nicht nur ein atomarer Wert zugewiesen, sondern ein komplexer Ausdruck, der selbst mehrere Elemente enthält. Auf diese Weise lassen sich den (Beton-) Eigenschaften vordefinierte Maßeinheiten zuordnen, ein bestimmter Datentyp für die Werte festlegen oder auch deren Schreibweise.

In einem RDF-Schema werden solche, anderen RDF-Eigenschaften untergeordnete Elemente, zunächst ebenfalls als RDF-Eigenschaften definiert. Werden sie anderen Eigenschaften in der Hierarchie untergeordnet, spricht man von sogenannten „Subproperties“. In met@BiM werden Subproperties entweder ohne Angabe einer Zuordnung oder mit einer *rdfs:subPropertyOf* Relation versehen. Im ersten Fall kann die RDF-Eigenschaft ohne Beschränkung im Zusammenhang mit allen RDF-Eigenschaften verwendet werden. Im zweiten Fall ist ihre Anwendung lediglich im Zusammenhang mit ganz bestimmten *rdf:Property* Elementen zulässig.

```

...

<!-- Deklaration der Werteeigenschaften -->

<rdf:Property ID="einheit">
  <rdfs:range rdf:resource="#EINHEITEN"/>
  <rdfs:comment>
    Ausdruck der physikalischen Kenngröße und mathematischen Dimension,
    die dem Wert zugewiesen wird</rdfs:comment>
</rdf:Property>

<rdf:Property ID="anteil">
  <rdfs:comment>
    Relativer Wert der beschriebenen Eigenschaft</rdfs:comment>
</rdf:Property>

<rdf:Property ID="wertebereich">
  <rdfs:comment>Wertebereich der beschriebenen Eigenschaft</rdfs:comment>
</rdf:Property>

<rdf:Property ID="obergrenze">
  <rdfs:subPropertyOf rdf:resource="#wertebereich"/>
  <rdfs:comment>Obergrenze des Wertebereichs</rdfs:comment>
</rdf:Property>

<rdf:Property ID="untergrenze">
  <rdfs:subPropertyOf rdf:resource="#wertebereich"/>
  <rdfs:comment>Untergrenze des Wertebereichs</rdfs:comment>
</rdf:Property>

...

```

Innerhalb der komplexen Schachtelung strukturierter Werte muss auch der Wert selbst in ein vordefiniertes RDF-Element eingefasst werden, um ihn von den Angaben zu Maßeinheiten und Datentypen unterscheiden zu können. Dafür ist in der RDF-Spezifikation das Element *rdf:value* (vgl. [RDF1999] Abschnitt 2.3) vorgesehen. Beispiele für die Anwendung strukturierter Eigenschaftswerte zur Zuordnung von Einheiten und Angabe von Wertebereichen befinden sich in Abschnitt 6.2 sowie in Anhang D dieser Arbeit.

Maßeinheiten für die Werte werden über das *rdf:Property* Element *einheit* zugeordnet. Die zulässigen Werte für dieses Element werden bestimmt von der mit der *rdfs:range* Relation zugewiesenen Ressource *#EINHEITEN* (s.o.).

Das *rdf:Property*-Element *anteil* erlaubt es, Wertepaare zu erzeugen, welche die Angabe relativer Werte erlauben. Solche relativen Werte treten häufig im Zusammenhang mit Angaben zur Zusammensetzung von Stoffgemischen auf, wie z.B. der Zuschlagzusammensetzung. Die eigentlichen, im Element *rdf:value* eingefassten Werte zur Beschreibung der Zuschlagzusammensetzung bilden in dieser Konstruktion die Bezeichnungen der Zuschlagkomponenten, wie z.B. „Ziegelbruch“, „Beton B45“, „Rheinkies“ etc.. Um diesen Zuschlagkomponenten für Aussagen, wie z.B. *Zuschlaggemisch enthält einen Anteil von Ziegelbruch von 20%*, den Relativwert „20%“ zuordnen zu können, wird das *rdf:Property*-Element *anteil* verwendet werden (vgl. Abbildung 6.7 Carbonatisierung).

Insbesondere größere Datenmengen, wie sie in detaillierten Diagrammen und Tabellen zu Versuchsreihen häufig dokumentiert werden, lassen sich in Form von Wertebereichen auf rationelle Art und Weise erfassen. Es werden dabei lediglich die Grenzwerte in den *rdf:Property* Elementen *untergrenze* bzw. *obergrenze* angegeben, die wiederum über das komplexe *rdf:Property* Element *wertebereich* zusammengefasst werden. Einige Beispiele für

die Codierung von Wertebereichen in RDF/XML-Syntax befinden sich im Abschnitt 6.2.3 sowie in Anhang D2.

Auf die explizite Zuweisung von Datentypen wurde innerhalb des met@BiM-RDF-Schemas verzichtet, da sie für die beabsichtigte Funktionalität der met@BiM-Anwendungen nicht unbedingt notwendig ist. Werden in einer Abfrage numerische Vergleichsoperatoren bspw. auf Textdaten angewandt, so können die darauf folgenden Fehlermeldungen schon im Suchalgorithmus abgefangen werden. Eine Beschränkung der zulässigen Werte schon bei der Metadateneingabe durch Angabe von Datentypen im RDF-Schema ist jedoch nicht die einzige Möglichkeit, Werten einen Datentyp zuzuweisen. Auch in den RDF-Datensätzen selbst können Verweise auf die in [XMLS2001b] definierten Datentypen eingefügt werden (vgl. [RDF1999] Abschnitt 7.3).

6.2 Entwurf der RDF-Datensätze

Auf der Grundlage der im met@BiM-RDF-Schema festgelegten Datenstrukturen können nun RDF-Beschreibungen für konkrete Informationsressourcen erstellt werden. Die RDF-Beschreibungen sollen dabei alle wesentlichen Merkmale der beschriebenen betontechnologischen Webinhalte (Ressourcen) widerspiegeln.

In den folgenden Abschnitten wird daher die Anwendung der Elemente des Dublin Core bzw. des met@BiM-Schemas zur Beschreibung der einzelnen betontechnologischen Ressourcentypen in BiM-Online an Beispielen demonstriert. Um die Lesbarkeit des Kapitels nicht zu sehr zu beeinträchtigen, wurden weitere Beispiele in den Anhang D am Ende dieser Arbeit ausgelagert. Die Modellierung einer Informationsquelle führt nicht immer zwangsläufig zu identischen Ergebnissen. Die Reihenfolge der Ressourceneigenschaften, die Vollständigkeit und Detailliertheit der Angaben liegen hierbei im Ermessen des Autors der RDF-Beschreibungen. Bei den im Folgenden vorgestellten Beispielen wurde jedoch der Versuch unternommen, einen Kompromiss zwischen den widersprüchlichen Anforderungen einer möglichst verlustfreien Beschreibung der Informationsressource und der Reduzierung des Eingabeaufwands für die Metadaten auf ein Minimum zu finden.

Zur Recherche der in den Abschnitten 6.2.1 und 6.2.2 dargestellten Ressourcentypen *metabim:DOKUMENT*, *metabim:KURS* und den Gleichungen reicht eine Beschreibung mit den Metadatenelementen des Dublin Core-Standards aus. Die Anwendung der DC-Elemente bezüglich betontechnologischer Forschungsberichte und Fachveröffentlichungen, den BiM-Online-Dialog Lektionen sowie den in diesen enthaltenen chemischen und mathematischen Gleichungen wird in diesen beiden Abschnitten beschrieben. In 6.2.3 liegt der Schwerpunkt hingegen auf der Anwendung der Elemente des met@BiM-RDF-Schemas. Diese werden dort zur Beschreibung der in technischen Diagrammen und Betonrezepturen abgebildeten betontechnologischen Aussagen und Zusammenhänge herangezogen.

6.2.1 RDF-Beschreibung von Dokumenten und Online-Kursen

Bei der Beschreibung von BiM-Online-Ressourcen der Typen *metabim:DOKUMENT* sowie *metabim:KURS* konnte die Erfahrung gemacht werden, dass beide bezüglich ihrer Recherchierbarkeit mit denselben DC-Elementen in ausreichender Detailliertheit beschrieben werden können. Bezüglich der Angaben zu einzelnen Eigenschaften sollten jedoch einige Besonderheiten der Ressourcentypen berücksichtigt werden (vgl. Tab. 4.5).

Dokumente

Bei den in BiM-Online gespeicherten Dokumenten handelt es sich überwiegend um Zwischenberichte zum Arbeitsstand der einzelnen Teilprojekte. Die Zwischenberichte wurden

nach ihrer Fertigstellung als HTML-Versionen der Textverarbeitungsdateien in das BiM-Online-Informationssystem eingebunden. Bezüglich der Gliederung der Berichte und der Formatierung der Inhalte wurden den Autoren keine Vorgaben auferlegt, so dass die dargestellten betontechnologischen Informationen in sehr inhomogener Form vorliegen. Die meisten dieser Berichte enthalten eine Vielzahl von Aussagen zu durchgeführten Materialprüfungen, die in Form von grafischen Abbildungen oder umfangreichen Tabellen präsentiert werden. Um die Metadatenstrukturen eines Dokuments nicht zu unübersichtlich werden zu lassen, erfolgt die RDF-Beschreibung des Dokuments als ganzes in einem eigenen Datensatz und zwar getrennt von den RDF-Beschreibungen zu den Inhalten. Die Metadaten des gesamten Berichts werden dann durch die in den nachfolgenden Abschnitten 6.2.2 sowie 6.2.3 vorgestellten detaillierteren RDF-Beschreibungen der Dokumentinhalte ergänzt. Durch die Aufteilung können zunächst alle wesentlichen Informationen erfasst werden, ohne dass einzelne RDF-Datensätze zu komplex werden (Abb. 6.1). Die Zusammengehörigkeit der RDF-Beschreibungen der Berichte und der Dokumentinhalte lässt sich über das *dc:source* Element von entsprechenden Programmen wiederherstellen. *dc:source* sollte bei den RDF-Datensätzen zu den Dokumentinhalten stets auf das zugrundeliegende Dokument verweisen.

Durch die Beschränkung der RDF-Beschreibung auf die im Dublin Core-Standard verwendeten Dokumenteigenschaften ergibt sich der in XML codierte RDF-Datensatz für einen Zwischenbericht wie in Abb. 6.1 zu:

```
<?xml version="1.0" encoding="ISO-8859-1"?>

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description about="http://www.b-i-m.de/berichte/c01/c01z0998.htm">

    <dc:title>BiM Zwischenbericht 09/98 Teilprojekt C01: Einflüsse der
      Aufbereitung von Bauschutt für eine Verwendung als
      Betonzuschlag
    </dc:title>
    <dc:creator>Prof. Reichel</dc:creator>
    <dc:description>Untersuchung des Einflusses einer Vorsortierung und
      einer Nassaufbereitung von Mischgranulat und
      Betongranulat auf die Eignung der Recyclinggranulate
      als Sekundärzuschlag für Beton
    </dc:description>
    <dc:publisher>MPA Dresden</dc:publisher>
    <dc:date>1999-11-15</dc:date>
    <dc:type>http://www.b-i-m.de/2001/01/metabim#DOKUMENT</dc:type>
    <dc:format>text/html</dc:format>
    <dc:identifier>http://www.b-i-m.de/berichte/c01/c01z0998.htm
    </dc:identifier>
    <dc:source>http://www.b-i-m.de</dc:source>
    <dc:Language>de</dc:Language>
    <dc:relation>http://www.b-i-m.de/berichte/c01/c01z0197.htm</dc:relation>
    <dc:relation>http://www.b-i-m.de/berichte/c01/c01z0198.htm</dc:relation>

  </rdf:Description>
</rdf:RDF>
```

Innerhalb des dargestellten RDF-Datensatzes werden Elemente der RDF-Spezifikation (Präfix *rdf:*) und des DC-Metadatenstandards (Präfix *dc:*) verwendet. Die Deklaration der verwendeten Namensräume erfolgt wie bei den RDF-Schema-Dateien am Anfang der RDF-Beschreibung.

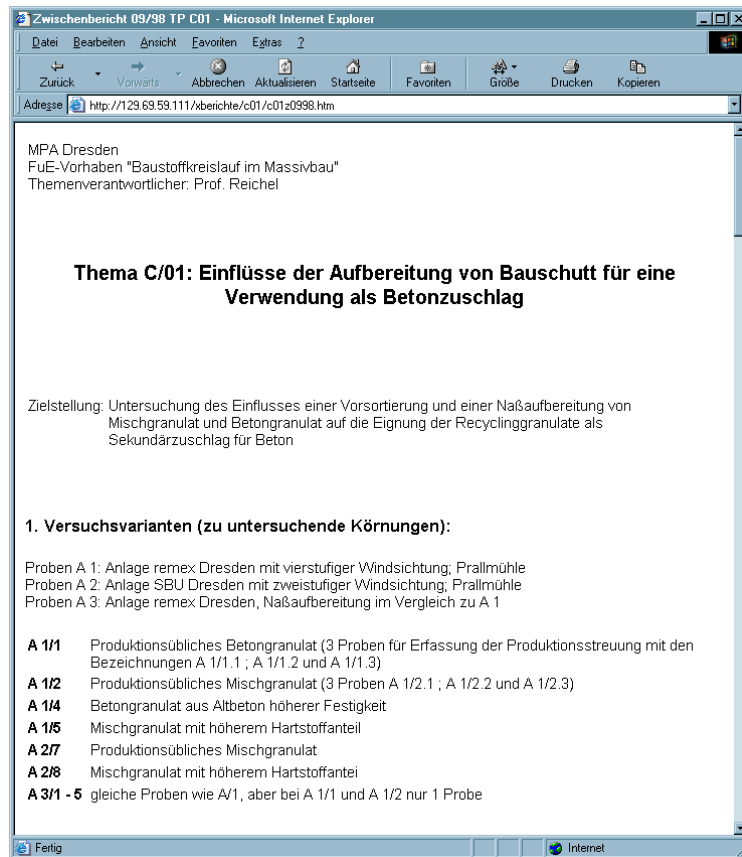


Abb. 6.1 Typischer Zwischenbericht zum Arbeitsstand eines Teilprojekts in BiM-Online als HTML-Dokument.

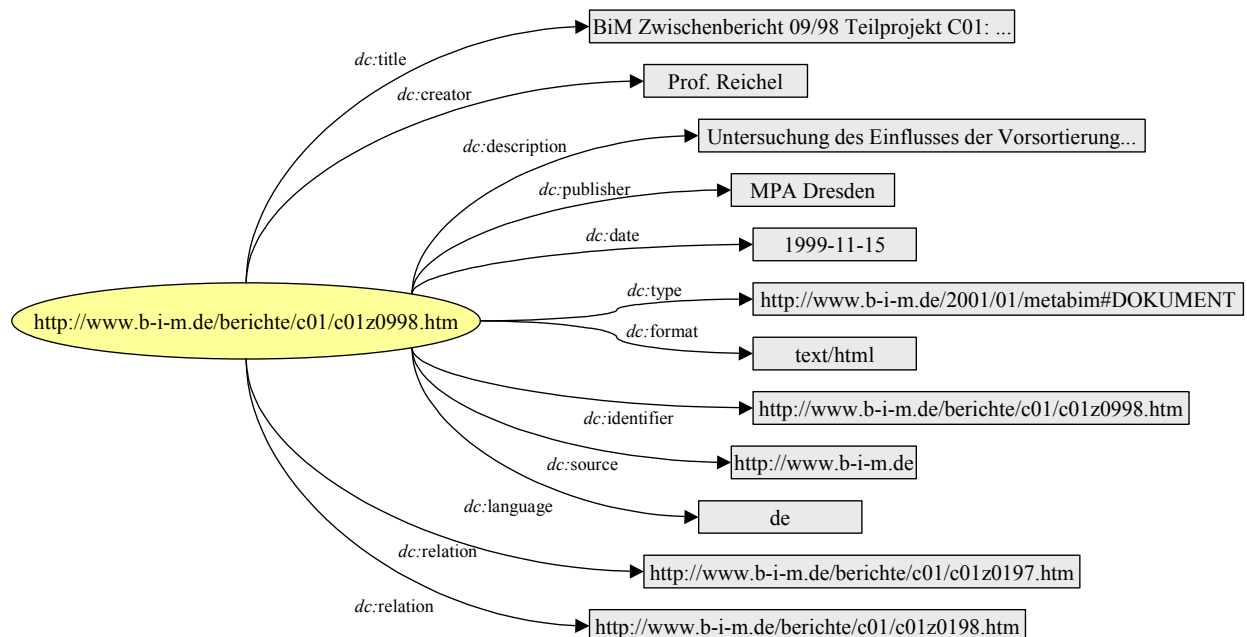


Abb.6.2 RDF-Graph der Beschreibung des Zwischenberichts aus Abb. 6.1, bestehend aus Dublin Core-Metadatenelementen.

Die RDF-Metadatenätze beginnen in der sechsten Zeile in `<rdf:Description about= ...>` mit dem Verweis auf die Ressource, die beschrieben werden soll. Als Zeiger auf die Ressource wird dabei die URL des Dokuments „<http://www.b-i-m.de/berichte/c01/c01z0998.htm>“ verwendet. Durch die Verwendung der eindeutigen URL als Bezeichnung der Ressource kann sichergestellt werden, dass der Name im Web nicht von mehreren Ressourcen gleichzeitig verwendet wird.

Innerhalb des `rdf:Description` Tags folgen die DC-Elemente. Nach Angaben zum Titel, Autor, Inhalt etc. des Dokuments wird der Bericht innerhalb des Elements `dc:type` als `met@BiM-Ressource` vom Typ `metabim:DOKUMENT` anhand der auf diese Ressourcendefinition verweisenden URI deklariert. Zur Angabe des Datenformats, in dem das Dokument vorliegt, werden die Medien Typen nach [RFC2046] verwendet. Im Falle des BiM-Online Zwischenberichts wird das Dokument als HTML-Dokument „`text/html`“ ausgewiesen.

Die Angaben des `dc:identifier` Elements unterscheiden sich bei Webdokumenten normalerweise nicht vom Ressourcennamen, da auch hier ein eindeutiger Wert zur Bezeichnung der Ressource verlangt wird, der auf die aktuelle Ressource zeigt. Bei BiM-Online-Dokumenten verweist `dc:source` auf die URL des übergeordneten Systems, in welches der Bericht eingebunden ist. Die beiden `dc:relation` Elemente schließlich zeigen auf weitere Berichte des Teilprojekts aus vorangegangenen bzw. späteren Forschungszeiträumen. Dadurch lassen sich wichtige semantische Querbezüge zu anderen Dokumenten modellieren und über DC konforme Softwareprogramme nutzen.

Kurse

Bei den Online-Kursen handelt es sich um interaktive Webanwendungen, deren Aufgabe in einer zielorientierten Wissensvermittlung besteht. Solche Ziele können z.B. Handlungsanweisungen für konkrete Problemstellungen sein, aber auch die Vermittlung didaktisch strukturierten Fachwissens im Rahmen von eLearning Systemen für die Vorbereitung auf eine Prüfung. Die Kurse bilden normalerweise inhaltlich abgeschlossene Dialoge zwischen Anwender und System, an deren Ende eine Aussage in Form einer Empfehlung oder Bewertung durch das System steht.

Zur Realisierung solcher Kurse im WWW stehen bezüglich der verwendeten Softwaretechnologien unterschiedliche Möglichkeiten offen. Die wichtigsten Varianten, die bzgl. einer Online-Publikation von Kursen in Frage kommen, sind:

- Abgeschlossenes, evtl. kompiliertes Programm, wie z.B. ein Java Applet
- Serverseitige Online-Applikation, z.B. bestehend aus dynamisch aus einer Datenbank erzeugten Dialogseiten
- Abfolge statischer, über Hyperlinks miteinander verknüpfter Webdokumente

Charakteristisch für alle diese Varianten ist die Möglichkeit einer nichtlinearen Wissensvermittlung im Gegensatz zur Linearität gedruckter Medien. Jeder Kurs kann somit aus unterschiedlichen Pfaden bestehen, die abhängig von den Angaben des Anwenders einen anderen Verlauf nehmen. Bei den hier verwendeten Beispielen handelt es sich um Lektionen der Rubrik BiM-Online-Dialog, die aus technischer Sicht eine Mischung von Kursen des zweiten als auch des dritten Typs der Aufzählung darstellen. Weitere Angaben zu Aufbau, Leistungsfähigkeit und Inhalt der BiM-Online Lektionen befinden sich in [Reinhardt1999b].

Bei der hier gewählten RDF-Beschreibung der Kurse in BiM-Online-Dialog ist zu berücksichtigen, dass diese lediglich die Beschreibung eines Kurses als Ganzes erlaubt. Dies verringert zwar den Eingabeaufwand für die Metadaten, da die Beschreibung nicht für jeden

Dialogschritt erneut erfolgen muss, als Ergebnisse von Recherchen werden jedoch lediglich die Startseiten der Lektionen zurückgegeben:

```
<?xml version="1.0" encoding="ISO-8859-1"?>

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">

  <rdf:Description about="http://www.b-i-m.de/dialog/beton/brechs1_1.htm">

    <dc:title>Beton - Einfluss rezyklierter Zuschläge auf die Betoneigen-
      schaften - Brechsand
    </dc:title>
    <dc:creator>Matthias Luik; Marcus Schreyer</dc:creator>
    <dc:description>BiM-Online Lektion: Eigenschaften von Beton mit
      rezykliertem Zuschlag
    </dc:description>
    <dc:publisher>Universität Stuttgart, IWB</dc:publisher>
    <dc:date>2001-11-21</dc:date>
    <dc:type>http://www.b-i-m.de/2001/01/metabim#KURS</dc:type>
    <dc:format>text/html</dc:format>
    <dc:identifizier>http://www.b-i-m.de/dialog/Beton/brechs1_1.htm
    </dc:identifizier>
    <dc:source>http://www.b-i-m.de/dialog/</dc:source>
    <dc:Language>de</dc:Language>
    <dc:relation>http://www.b-i-m.de/dialog/Beton/betonspl1_1.htm
    </dc:relation>
    <dc:relation>http://www.b-i-m.de/dialog/Beton/ziegelspl1_1.htm
    </dc:relation>
    <dc:relation>http://www.b-i-m.de/dialog/Beton/gips1_1.htm</dc:relation>
    <dc:relation>http://www.b-i-m.de/dialog/Beton/asphalt1_1.htm
    </dc:relation>
    <dc:subject>Betoneigenschaften Brechsand Verarbeitbarkeit Luftporenge-
      halt Druckfestigkeit Elastizitätsmodul Spaltzug- und Biege-
      zugfestigkeit Kriechen Schwinden Frost-Tausalz-Widerstand
      Frost-Tau-Widerstand Carbonatisierung ...
    </dc:subject>

  </rdf:Description>
</rdf:RDF>
```

Die Vorgehensweise bei der Wahl der Werte für die Metadatenelemente des Ressourcentyps *metabim:DOKUMENT* lässt sich auf die Beschreibung der Kurse übertragen. Zu beachten ist, dass im Element *<rdf:Description about:...>* sowie in *dc:identifizier* sinnvollerweise die URL der Startseite der Lektion angegeben werden sollte. In die Titelangabe wurden zusätzlich zur Bezeichnung der Lektion noch weitere Angaben zum Themengebiet angefügt, welche das Auffinden des Kurses innerhalb des übergeordneten BiM-Online-Dialog Systems erleichtern. Der Ressourcentyp in *dc:type* verweist auf *metabim:KURS* des met@BiM-RDF-Schemas. Die *dc:relation* Elemente verweisen zu den Startseiten weiterer Lektionen des Themenbereichs „Einfluss rezyklierter Zuschläge auf die Betoneigenschaften“. Das zur Beschreibung der in Abb. 6.3 dargestellten Lektion zum Einfluss von Brechsand auf die Betoneigenschaften gehörende grafische RDF-Modell zeigt Abb. 6.4.



Abb. 6.3 Startseite der Lektion zum Einfluss von Brechsand auf die Betoneigenschaften

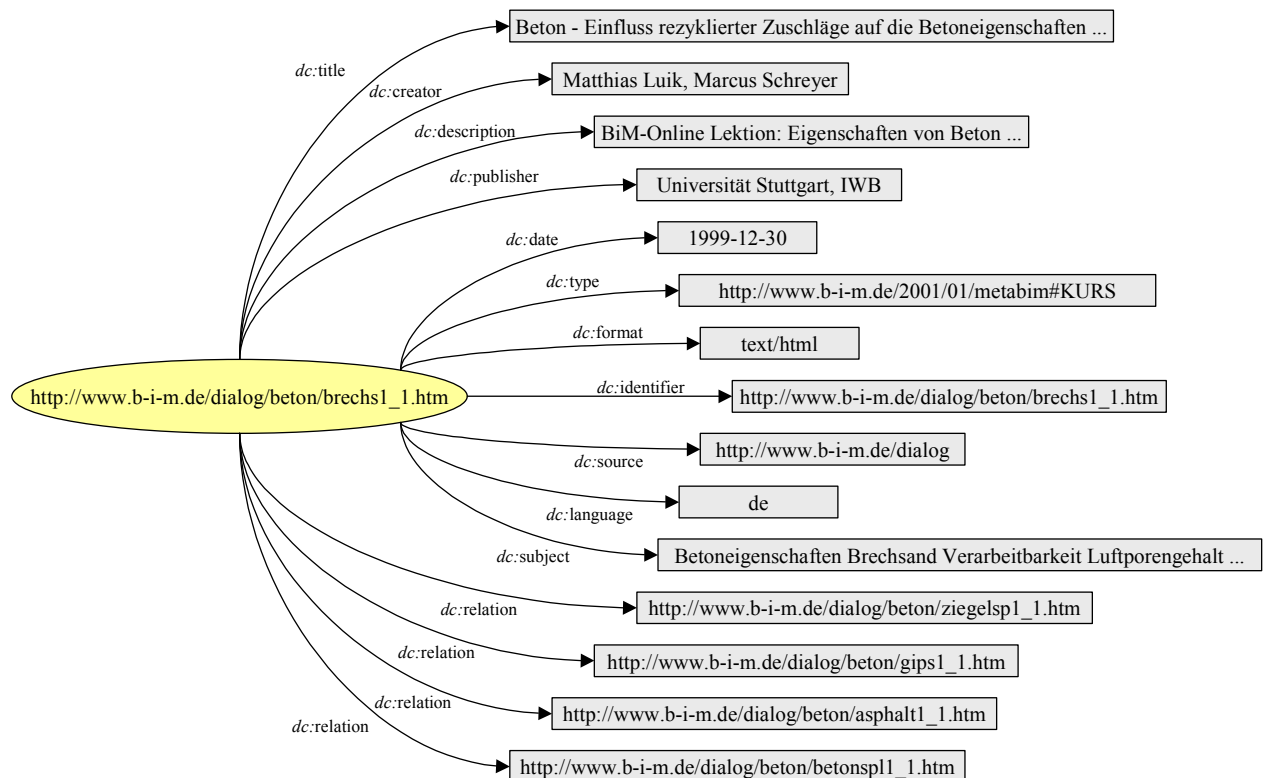


Abb. 6.4 RDF-Graphenmodell zur Beschreibung der BiM-Online-Dialog Lektion aus Abb. 6.3

6.2.2 RDF-Beschreibung von Gleichungen

Der Umgang mit chemischen und mathematischen Gleichungen bildet die Grundlage typischer Ingenieur Tätigkeiten. So erfolgt z.B. die Nachweisführung im konstruktiven Ingenieurbau anhand eines, in der Regel genormten Formelwerks. Die in der Norm angegebenen Gleichungen stellen wiederum häufig das Ergebnis bzw. die Grundlage wissenschaftlicher Arbeiten innerhalb der Bauforschung dar. Auch im Rahmen des Forschungsprojekts Baustoffkreislauf im Massivbau wurden aus empirischen Untersuchungen heraus Gleichungen entwickelt, um die Zusammenhänge zwischen einzelnen Parametern quantifizieren zu können. Weiterhin werden in den Forschungsberichten auch Formeln aus der Literatur zitiert und auf die Fragestellungen der Teilprojekte bezogen.

Bislang wird der direkte Zugriff auf diesen Ressourcentyp im WWW erschwert durch die eingeschränkten Möglichkeiten der Darstellung von Gleichungen in HTML-Dokumenten (vgl. Abschnitt 2.4 bzw. Abschnitt 3.2.2). Auch fortschrittlichere Darstellungsformen mit auf XML-basierenden Markup Sprachen (vgl. Abschnitt 2.6.3) erlauben zwar eine differenziertere Erschließung der Gleichungen, bieten jedoch den allgemeinen Suchmaschinen des Web nur wenig Anhaltspunkte bzgl. der Bedeutung der Formeln bzw. der einzelnen Formelelemente. Damit Gleichungen im Rahmen des met@BiM-Datenmodells leichter auffindbar werden, liegt es nahe, das vermutlich wichtigste Recherchekriterium, die Parameter einer Gleichung, in einer geeigneten Art und Weise zu beschreiben. Am einfachsten erfolgt dies anhand von stichwortartigen Erläuterungen der in der Formel miteinander in Beziehung gesetzten Parameter. Diese Vorgehensweise wird im Folgenden an den Beispielen einer mathematischen sowie einer chemischen Gleichung gezeigt.

Mathematische Gleichungen

Abhängig von ihrer Komplexität wurden in den Dokumenten von BiM-Online die enthaltenen Formeln entweder als Grafiken oder in HTML-Codierung integriert. Eine Codierung in MathML, der im Vergleich zu den hier dargestellten Beispielen der Medientyp „text/xml“ [IANA2002] im Element *dc:type* zugeordnet werden müsste, wurde in BiM-Online zur Darstellung nicht verwendet.

Als Beispiel für diesen Ressourcentyp wird hier die Gleichung für das Kriechmaß von Beton aus [Grübl1998] benutzt. Zur Beschreibung des Kriechverhaltens wird im Forschungsbericht die Gleichung

$$\varepsilon_{k,t} = \varepsilon_{\text{tot},t} - \varepsilon_{s,t} - \varepsilon_{\text{el},t0}$$

aus [Wesche1993] verwendet.

Zur Referenzierung von Ressourcen, die sich innerhalb von HTML-Dokumenten befinden, muss auf im Dokumentquelltext gespeicherte Markierungen mit sogenannten Anchor-Tags verwiesen werden (vgl. Abschnitt 2.4). URIs mit lokalem Anchor-Tagnamen, wie sie in den Elementen *<rdf:Description about=“...>*, *dc:identifier* sowie *dc:relation* nach dem # Symbol angegeben werden, können daher nicht auf beliebige Stellen im Dokument verweisen. Innerhalb des BiM-Online-Systems wurden daher diejenigen Dokumente, die Ressourcen vom Typ *metabim:MATHEMATISCHEFORMEL* enthielten, an den entsprechenden Stellen im Quelltext mit einem „mathID“ benannten Anchor-Tag markiert, wobei ID für eine fortlaufende Nummerierung der mathematischen Gleichungen im Dokument steht. Eine von den Datei-Schreibrechten in das Dokument unabhängige Möglichkeit der Referenzierung von Dokumentenbestandteilen wird erst durch die Codierung von Dokumenten in XML oder

XHTML und der Softwareunterstützung des XPointer-Standards (vgl. Abschnitt 2.6.2) ermöglicht werden.

Da es sich bei der oben angeführten Gleichung um eine Computergrafik handelt, wird in *dc:format* als Medientyp „image/gif“ angegeben. Im Titel wird die Formel mit einer charakteristischen Bezeichnung versehen, wie sie Gleichungen häufig in Publikationen in Klammern oder im Formelverzeichnis zugeordnet wird. *dc:source* verweist auf das Dokument als übergeordnete Speichereinheit der Gleichung. Im Element *dc:relation* sollen inhaltlich mit der Gleichung verknüpfte Ressourcen angegeben werden. Beispiele hierfür sind weitere Formeln eines Gleichungssystems (vgl. Tab. 4.5) oder, wie hier, ein Diagramm innerhalb des Dokuments, in dem der durch die Gleichung ausgedrückte Zusammenhang grafisch dargestellt wird. In *dc:subject* schließlich werden die einzelnen Parameter der Gleichung in lesbarer Form benannt und beschrieben. Das dies, wie im hier abgebildeten Fall, in mehreren *dc:subject* Elementen erfolgt, hat seinen Grund in der leichteren Lesbarkeit dieser Darstellung.

```
<?xml version="1.0" encoding="ISO-8859-1"?>

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  <rdf:Description
    about="http://www.b-i-m.de/berichte/e03/e03z0998.htm#math1">

      <dc:type>http://www.b-i-m.de/2001/01/metabim#MATHEMATISCHEFORMEL
    </dc:type>
    <dc:format>image/gif</dc:format>
    <dc:identifizier>http://www.b-i-m.de/berichte/e03/e03z0998.htm#math1
    </dc:identifizier>
    <dc:date>2001-10-18</dc:date>
    <dc:title>Berechnung des Kriechmaßes</dc:title>
    <dc:creator>Prof. Dr.-Ing. P. Grübl, Dipl.-Ing. M. Rühl</dc:creator>
    <dc:publisher>Technische Universität Darmstadt, Institut für Massivbau
    </dc:publisher>
    <dc:source>http://www.b-i-m.de/berichte/e03/e03z0998.htm</dc:source>
    <dc:relation>http://www.b-i-m.de/berichte/e03/e03z0998.htm#8
    </dc:relation>
    <dc:subject>Kriechmaß zum Zeitpunkt t</dc:subject>
    <dc:subject>Gesamtverformung zum betrachteten Zeitpunkt t</dc:subject>
    <dc:subject>Schwindmaß an unbelasteten Probekörpern zum betrachteten
      Zeitpunkt t</dc:subject>
    <dc:subject>Elastische Verformung des Probekörpers zur Zeit t0 infolge
      der aufgetragenen kriecheerzeugenden Spannung</dc:subject>

  </rdf:Description>
</rdf:RDF>
```

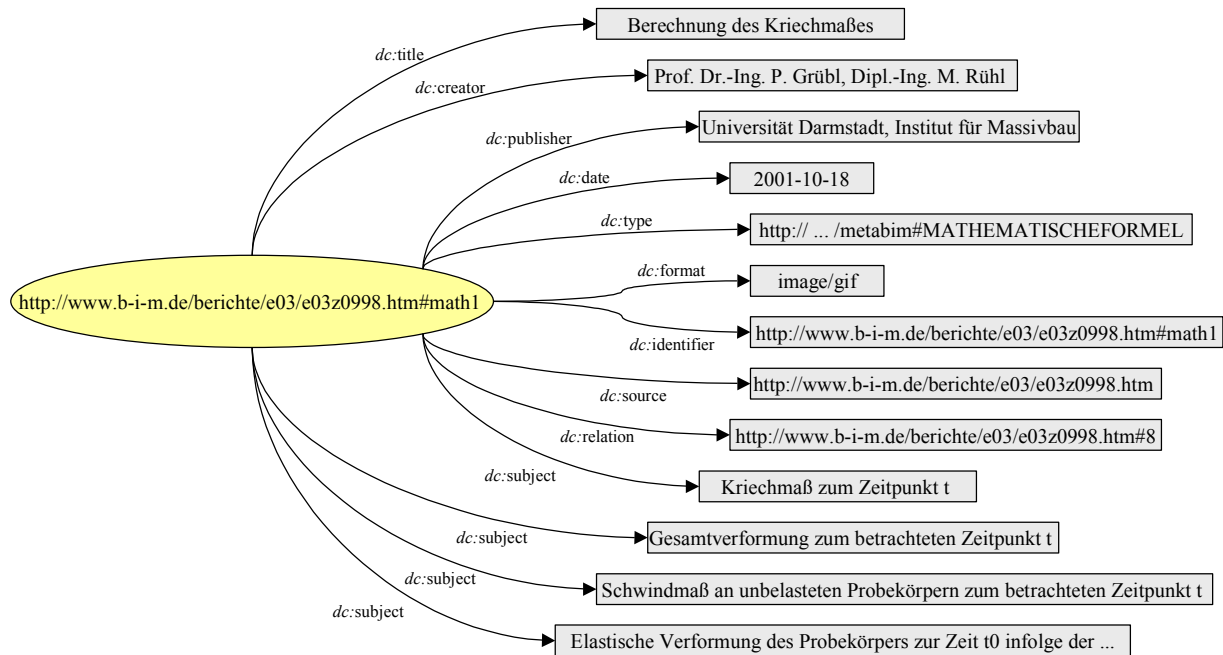
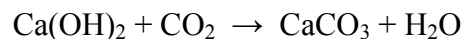


Abb. 6.5 RDF-Graphenmodell zur Beschreibung der Gleichung für das Kriechmaß

Chemische Gleichungen

Die Beschreibung chemischer Gleichungen erfolgt vom Prinzip her identisch zu mathematischen Formeln. Die innerhalb des BiM-Online-Systems veröffentlichten chemischen Formeln liegen ausschließlich als „text/html“ bzw. als „image“ Medientypen vor. Ähnlich wie MathML hat sich auch hier die auf XML-basierende Markup Sprache CML noch nicht durchgesetzt.

Am Beispiel der Carbonatisierungsgleichung, die Bestandteil eines Dokuments des BiM-Online-Lexikons ist [Reinhardt2000], sollen die Unterschiede der RDF-Modellierung erläutert werden.



Als Zeiger auf die oben abgebildete Gleichung wird ebenfalls eine URI verwendet, bestehend aus der URL des Dokuments sowie des lokalen Anchors. Analog zu den mathematischen Formeln verweist die in den Elementen `<rdf:Description about=“...>` sowie in `dc:identifizier` verwendete URI auf ein Anchor-Tag mit dem Namen „chemID“. Die einzelnen Parameter bzw. chemischen Verbindungen der Formel erscheinen in ausgeschriebener Form als `dc:subject` Elemente der RDF-Beschreibung.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
```

```
<rdf:RDF
```

```
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
```

```
  <rdf:Description about="http://www.b-i-m.de/lexikon/inhalt/
```

```
    carbonatisierung.htm#chem1">
```

```
    <dc:type>http://www.b-i-m.de/2001/01/metabim#CHEMISCHEFORMEL
```

```

</dc:type>
<dc:format>text/html</dc:format>
<dc:identifizier>http://www.b-i-m.de/lexikon/inhalt/
    carbonatisierung.htm#chem1
</dc:identifizier>
<dc:date>2000-01-11</dc:date>
<dc:title>Carbonatisierung von Beton</dc:title>
<dc:creator>Prof. Dr.-Ing. H.-W. Reinhardt, Dr.-Ing. M. Jooß
</dc:creator>
<dc:publisher>Universität Stuttgart, IWB</dc:publisher>
<dc:source>http://www.b-i-m.de/lexikon/</dc:source>
<dc:subject>Calciumhydroxid</dc:subject>
<dc:subject>Kohlendioxid</dc:subject>
<dc:subject>Calciumcarbonat</dc:subject>
<dc:subject>Wasser</dc:subject>

</rdf:Description>
</rdf:RDF>

```

6.2.3 RDF-Beschreibung von Betonprüfdaten und Rezepturen

Während die Beschreibung von Dokumenten, Kursen und Gleichungen mit Standard-elementen des Dublin Core-Metadatenstandards erfolgen kann, gestaltet sich die Abbildung der fachspezifischen Aussagen, die in den in ihrer Informationsdichte komplexeren technischen Diagrammen, Tabellen etc. enthalten sind, deutlich aufwendiger.

Diagramme und Tabellen werden dazu genutzt, die Beziehungen zwischen betontechnologischen Kenngrößen und den diesen zugeordneten Werten zu visualisieren und hervorzuheben. Die Abbildung dieser zumeist mehrdimensionalen Zusammenhänge erfordert im Gegensatz zum relativ einfach aufgebauten DC-Metadatenatz eine komplexere Struktur der Metadaten. Dies ist auch deshalb erforderlich, damit detailliertere Abfragen mit Parameterkombinationen etc. auf die Metadatenätze realisiert werden können. Um die vielfältigen Eigenschaften des Betons und seiner Ausgangsstoffe in einem übersichtlich strukturierten Metadatenmodell beschreiben zu können, wurden mit den *met@BiM*-Metadatenelementen sowie der Ontologie in Kapitel 4 (vgl. Abb. 4.3) bereits ein entsprechender Ansatz vorgestellt.

Da es sich vor allem bei den Diagrammen zu betontechnologischen Informationsquellen um komprimierte Darstellungen größerer Datenmengen handelt, führten dies im Vergleich zu den Ressourcentypen *metabim:DOKUMENT*, *metabim:KURS* oder den Formeln zu umfangreicheren RDF-Beschreibungen. Dem kann jedoch durch die Nutzung von Wertebereichen (vgl. Abb. 4.5) begegnet werden, deren RDF-Codierung in den folgenden Abschnitten beschrieben wird.

Die allgemeine Beschreibung des Kontexts bzw. der Struktur von Ressourcen erfolgt auch bei den betontechnologischen Daten anhand der Elemente des Dublin Core. *met@BiM* ergänzt dieses allgemeine Konzept, in dem es dem DC-Element *dc:subject* einen komplexen Inhalt aus fachspezifischen Metadatenelementen des *met@BiM*-Namensraums zuordnet.

Betonprüfdaten

Im Gegensatz zu den bisher beschriebenen Ressourcentypen handelt es sich bei den Ergebnissen betontechnologischer Prüfungen um einen nur mittelbar referenzierbaren Ressourcentyp. Während die URI-Zeiger der Ressourcenbeschreibungen bei Dokumenten, Kursen und den Gleichungen direkt auf die entsprechenden digitalen Ressourcen verweisen, stehen bei den betontechnologischen Daten mehrere Möglichkeiten der Modellierung zur Verfügung. Zum einen erlaubt das Konzept der URIs [RFC2396] die Verwendung von

Webadressen, die nicht auf digitale Objekte verweisen (vgl. Abschnitt 5.2). Auf diese Weise können auch „realen“ Objekten, wie der Betonprobe, an der die Versuchswerte ermittelt werden, eine solche „virtuelle“ URI zugeordnet werden. Diese muss lediglich den allgemeinen Anforderungen an eine URI entsprechen und darf noch nicht anderweitig belegt sein. Ein Problem dieser Vorgehensweise besteht in der Unsicherheit, ob die gewählte URI tatsächlich nicht schon einmal im WWW zur Bezeichnung einer anderen Ressource herangezogen wurde und somit ein Zeiger auf zwei unterschiedliche Datensätze verweisen würde.

Ein anderer Weg orientiert sich an den Abbildungen der Beton-Daten in Form von Diagrammen, Tabellen oder Beschreibungen in Textabschnitten von Webdokumenten. Da solche im Web publizierte Datenquellen bereits über URIs verfügen, lassen sie sich als eindeutige Zeiger verwenden. Die weitere Referenzierung der Rezepturen, Beton- und Zuschlagsproben erfolgt in dieser Arbeit stellvertretend anhand der Beschreibung der Diagramme etc., in welchen die Ergebnisse abgebildet werden. Genaugenommen werden damit in den RDF-Beschreibungen die Betoneigenschaften den Abbildungen zugeordnet. Für die URI der Ressource in `<rdf:description about=“...>`, auf welche sich die `met@BiM`-Metadatenelemente beziehen, ergibt sich daraus als Konsequenz, dass sie auf die in den Webdokumenten enthaltenen Abbildungen bzw. Textabschnitte verweisen müssen.

Bei der Beschreibung der Informationsressourcen werden die abgebildeten Betoneigenschaften in den Mittelpunkt gestellt. Dieses Hervorheben der Eigenschaften bei der Modellierung von RDF-Beschreibungen entspricht dabei der Situation bei der Datenrecherche, die ebenfalls anhand von Angaben zu den Eigenschaften des gesuchten Objekts erfolgt. Bereits die von der Definition der Ressourcen zunächst unabhängig erfolgende Definition der Properties (vgl. Abschnitt 6.1 und Anhang B) im RDF-Schema unterstützt dabei diesen eigenschaftszentrierten Modellierungsansatz. Die möglichst allgemein gehaltenen Definitionen der Betoneigenschaften im `met@BiM`-RDF-Schema erlauben es, dass z.B. Eigenschaften wie `metabim:frostwiderstand` in Zusammenhang mit unterschiedlichen Subressourcen der `met@BiM`-Ontologie (vgl. Abb. 4.3) verwendet werden können. So kann das Eigenschaftselement innerhalb einer RDF-Beschreibung je nach Kontext einmal auf die Ressource `metabim:FESTBETON` und ein anderes Mal auf `metabim:ZUSCHLAG` bezogen werden. Die Zuweisung des konkreten Kontexts erfolgt letztendlich in der RDF-Beschreibung einer Informationsquelle. Dort wird innerhalb des `rdf:type` Elements über einen Zeiger auf die URI der entsprechenden Ressource verwiesen.

Bei einer Recherche bezüglich des Frostwiderstands kann ein System nun aus der so formulierten Eigenschafts-Ressource-Paarung in der RDF-Beschreibung herauslesen, in welchem Zusammenhang die Baustoffeigenschaft in der Informationsquelle verwendet wird. Diese Information ermöglicht es z.B., dem Anwender völlig neue und exakte semantische Funktionen zur Präzisierung seiner Recherche anzubieten (vgl. Abschnitt 6.6). Gleichzeitig lässt sich durch die weitgehende Entkopplung der Materialeigenschaften von den Baustoffen der Umfang der benötigten Metadatenelemente reduzieren.

Im folgenden Beispiel einer Tabelle zum Einfluss rezyklierter Zuschläge im Sandanteil auf die Betondruckfestigkeit wird die Beschreibung der Aussagen einer betontechnologischen Informationsressource demonstriert. In Tab. 6.1 wird die für Beton mit rezyklierter Gesteinskörnung charakteristische mittlere Kornrohddichte der Zuschläge mit der nach DIN 1048-5 ermittelten Würfeldruckfestigkeit in Beziehung gesetzt. Als weitere Parameter werden in der Tabelle das Alter der Probewürfel sowie der w/z-Wert der Mischung angegeben. Insgesamt werden 20 Messergebnisse als diskrete Einzelwerte abgebildet und in Versuchsreihen mit jeweils identischen w/z-Werten gruppiert.

Eine RDF-Beschreibung der betontechnologischen Aussagen dieser Tabelle kann wie folgt aufgebaut werden: Zu Beginn der Beschreibung erfolgt die bereits aus den vorigen Ressour-

Tab. 6.1 Einfluss der Rohdichte rezyklierter Zuschläge auf die Betondruckfestigkeit [vgl. Müller1998a, Müller1998b]

Zement	w/z	Beton	Zuschlagrohddichte ¹⁾ [kg/m³]	Druckfestigkeit ^{1) 2)} [N/mm²]
CEM I 32,5R	0,40	AB40	2620	59,0
		RZ40	2480	57,1
	0,55	N55	2620	42,6
		RZ1	2430	32,8
		RZ2	2180	37,0
		RZ3	2310	35,1
		RZ4	2400	35,6
		RZ5	2000	29,8
	0,60	RZ10	2300	23,3
		RZ12	2490	30,1
	0,70	AB70	2620	30,0
		RZ70	2480	28,2
	0,75	RZ6	2180	20,8
		RZ7	2310	20,3
CEM I 42,5R	0,33	N1h	2620	82,5
		RZ1h	2430	65,1
		RZ2h	2180	64,2
		RZ3h	2310	65,6
		RZ4h	2400	67,1
		RZ5h	2000	57,2

¹⁾ Mittelwert aus 3 Proben²⁾ Druckfestigkeit am Würfel nach 28 Tagen

centypen bekannte Zuweisung der Namensraumpräfixe. Im Vergleich mit den vorangegangenen Beispielen muss bei der Verwendung von Elementen des met@BiM-RDF-Schemas zusätzlich mit *xmlns:metabim*="..." ein Verweis auf den entsprechenden Namensraum eingefügt werden. Wie weiter oben in diesem Abschnitt beschrieben, werden die Metadatenelemente für die Baustoffeigenschaften in der RDF-Beschreibung auf die URI der Informationsquelle bezogen. Da in den BiM-Online-Forschungsberichten Absätze, die betontechnologische Informationsressourcen enthalten, mit einem nummerierten Anchor-Tag gekennzeichnet sind, setzt sich die URI in der Zeile *<rdf:Description about="..."* aus der URL des Dokuments und dem Fragment Identifier der Tabelle zusammen. Auf die Namensraum- und Ressourcendeklaration folgen dann die Metadatenelemente des Dublin Core. Hervorzuheben sind hier zum einen die Werte des Ressourcentyps in *dc:type*, ein Element, dessen Wert auf die Ressource *metabim:BETON* zeigt. Zum anderen die *dc:relation* Elemente, innerhalb derer auf weitere, inhaltlich zusammenhängende Diagramme und Tabellen anhand von URLs verwiesen wird. Die vier in diesem Beispiel angegebenen

Adressen führen zu den Betonrezepturen, die zur Herstellung der Probekörper für die Druckfestigkeitsversuche verwendet wurden. Um die möglichen inhaltlichen Verbindungen zwischen Prüfergebnissen auf ein sinnvolles Maß zu reduzieren, wurden die *dc:relation* Elemente bei Prüfdaten aus BiM-Online konsequent mit URLs zu den jeweiligen Betonrezepturen der Probekörper belegt. Da alle Prüfergebnisse einer Rezeptur dadurch einen identischen Verweis auf die Rezepturadresse besitzen, ist auch nachträglich die Zusammenstellung entsprechender Ergebnisse über die Abfrage dieses gemeinsamen Elements möglich.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:metabim="http://www.b-i-m.de/2001/01/metabim#">
  <rdf:Description about="http://www.b-i-m.de/berichte/d03/d03z0998.htm#4">

    <dc:type>http://www.b-i-m.de/2001/01/metabim#BETON</dc:type>
    <dc:format>image/gif</dc:format>
    <dc:identifier>http://www.b-i-m.de/berichte/D03/d03z0998.htm#14
    </dc:identifier>
    <dc:date>1999-10-26</dc:date>
    <dc:title>Einfluss der Rohdichte rezyklierter Zuschläge auf die
      Betondruckfestigkeit
    </dc:title>
    <dc:creator>Dipl.-Ing. Ch. Müller, Dipl.-Ing. U. Wiens</dc:creator>
    <dc:publisher>Rheinisch-Westfälische Technische Hochschule Aachen,
      Institut für Bauforschung (ibac)</dc:publisher>
    <dc:source>http://www.b-i-m.de/berichte/d03/d03z0998.htm</dc:source>
    <dc:relation>http://www.b-i-m.de/berichte/d03/d03z0998.htm#TabelleA1
    </dc:relation>
    <dc:relation>http://www.b-i-m.de/berichte/d03/d03z0998.htm#TabelleA2
    </dc:relation>
    <dc:relation>http://www.b-i-m.de/berichte/d03/d03z0998.htm#TabelleA3
    </dc:relation>
    <dc:relation>http://www.b-i-m.de/berichte/e04/e04z0998.htm#Tabelle1
    </dc:relation>
    <dc:description>Ziel ist die Bewertung der bei der Aufbereitung von
      Bauschutt anfallenden Recyclingzuschläge hinsichtlich
      der Eignung als Betonzuschlag.</dc:description>
```

Die betontechnologischen met@BiM-Elemente werden als komplexe Inhalte unterhalb des *dc:subject* Elements in die RDF-Beschreibung integriert. Damit RDF-Parser die komplexe Struktur des *dc:subject* Elements richtig interpretieren, muss *dc:subject* um das Attribut *parseType*="Resource" ergänzt werden. Im auf das *dc:subject* Tag folgenden Absatz werden zunächst dimensions- und einheitenlose Werte für die zugrundeliegende Prüfvorschrift sowie die verwendeten w/z-Werte in RDF-Syntax beschrieben. Am Beispiel des *metabim:wzwert* Tags lässt sich ablesen, dass auch met@BiM-Elementen unterhalb von *dc:subject* ein *parseType* Attribut zugeordnet wird, wenn dem Parser ein komplexer Inhalt aus weiteren RDF-Elementen angezeigt werden soll. Bei den untergeordneten Elementen dieser Hierarchieebene des XML-Baumes handelt es sich um die Kontextzuordnung der Baustoffeigenschaft zu einer entsprechenden met@BiM-Ressource, die innerhalb des Elements *rdf:type* ... erfolgt. Die Werte von Metadatenelementen auf dieser Hierarchieebene werden entweder innerhalb eines *rdf:value*-Elements eingefasst, sofern es sich um Einzelwerte handelt, oder sie werden wie im Falle des w/z-Werts und der Zementbezeichnung mit *rdf:li* als Listenwerte (vgl. Abschnitt 5.3) codiert.

...

```

<dc:subject parseType="Resource">

  <metabim:pruefvorschrift parseType="Resource">
    <rdf:type resource=
      "http://www.b-i-m.de/2001/01/metabim#FESTBETONPRUEFUNGEN"/>
    <rdf:value>DIN 1048-5</rdf:value>
  </metabim:pruefvorschrift>

  <metabim:wzwert parseType="Resource">
    <rdf:type resource=
      "http://www.b-i-m.de/2001/01/metabim#REZEPTUR"/>
    <rdf:li>0.33</rdf:li>
    <rdf:li>0.4</rdf:li>
    <rdf:li>0.55</rdf:li>
    <rdf:li>0.6</rdf:li>
    <rdf:li>0.7</rdf:li>
    <rdf:li>0.75</rdf:li>
  </metabim:wzwert>

  <metabim:bezeichnung parseType="Resource">
    <rdf:type resource=
      "http://www.b-i-m.de/2001/01/metabim#ZEMENT"/>
    <rdf:li>CEM I 32,5R</rdf:li>
    <rdf:li>CEM I 42,5R</rdf:li>
  </metabim:bezeichnung>

```

...

Sollen den Werten Maßeinheiten zugeordnet werden, so werden diese innerhalb des *metabim:einheit*-Tags auf der Ebene der Werteangaben in die Beschreibung eingefügt. Die hier verwendeten Werte für das Element *metabim:einheit* werden im met@BiM-RDF-Schema innerhalb der Ressource *metabim:EINHEITEN* vordefiniert (vgl. Anhang B und Abschnitt 6.1).

...

```

<metabim:messzeitpunkt parseType="Resource">
  <rdf:type resource=
    "http://www.b-i-m.de/2001/01/metabim#FESTBETONPRUEFUNGEN"/>
  <metabim:einheit>d</metabim:einheit>
  <rdf:value>28</rdf:value>
</metabim:messzeitpunkt>

<metabim:druckfestigkeit parseType="Resource">
  <rdf:type resource=
    "http://www.b-i-m.de/2001/01/metabim#FESTBETON"/>
  <metabim:einheit>Nmm2</metabim:einheit>
  <rdf:li>59.0</rdf:li>
  <rdf:li>57.1</rdf:li>
  <rdf:li>42.6</rdf:li>
  <rdf:li>32.8</rdf:li>
  <rdf:li>37.0</rdf:li>
  <rdf:li>35.1</rdf:li>
  <rdf:li>35.6</rdf:li>
  <rdf:li>29.8</rdf:li>
  <rdf:li>23.3</rdf:li>
  <rdf:li>30.1</rdf:li>
  <rdf:li>30.0</rdf:li>
  <rdf:li>28.2</rdf:li>

```

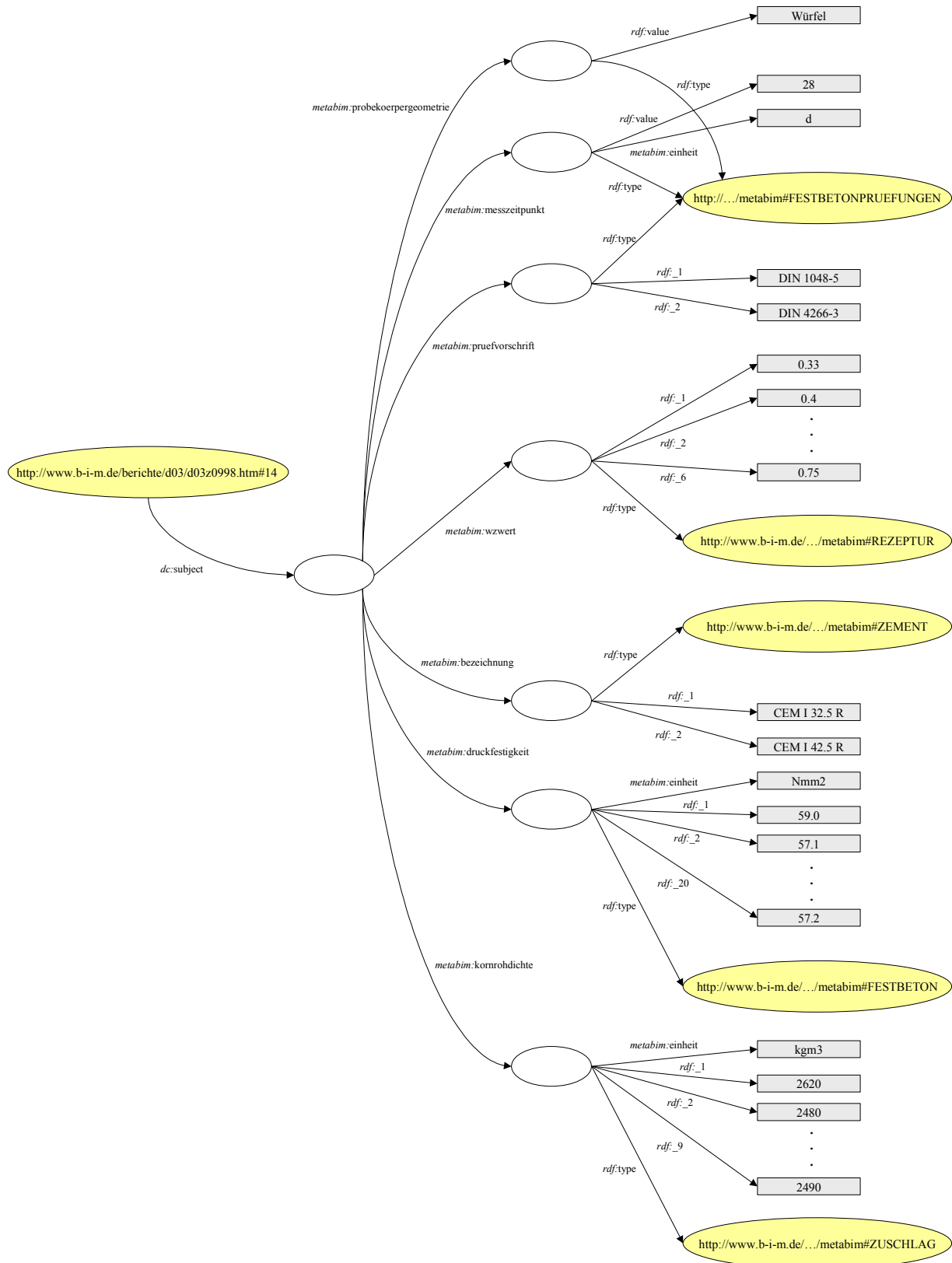


Abb. 6.6 RDF-Graph des Elements *dc:subject* sowie der untergeordneten met@BiM-Elemente zu den Aussagen in Tab. 6.1

```

<rdf:li>20.8</rdf:li>
<rdf:li>20.3</rdf:li>
<rdf:li>82.5</rdf:li>
<rdf:li>65.1</rdf:li>
<rdf:li>64.2</rdf:li>
<rdf:li>65.6</rdf:li>
<rdf:li>67.1</rdf:li>
<rdf:li>57.2</rdf:li>
</metabim:druckfestigkeit>

```

...

Der RDF-Graph des *dc:subject*-Elements wird in Abb. 6.6 dargestellt. Die vollständige RDF-Beschreibung der Tabelle 6.1 mit allen Parametern befindet sich im Anhang D1.

Die Nutzung des Wertebereichkonstrukts soll an einer stetigen Messreihe dargestellt werden. Als Beispiel dient die Entwicklung der Carbonatisierung von Beton in Abb. 6.7. Verschiedene Betonproben, die sich im Anteil an rezykliertem Gesteinskörnung aus Brechsand unterscheiden, wurden über einen Zeitraum von 182 Tagen bezüglich des Carbonatisierungsfortschritts untersucht und mit einem Referenzbeton aus Rheinkiessand verglichen.

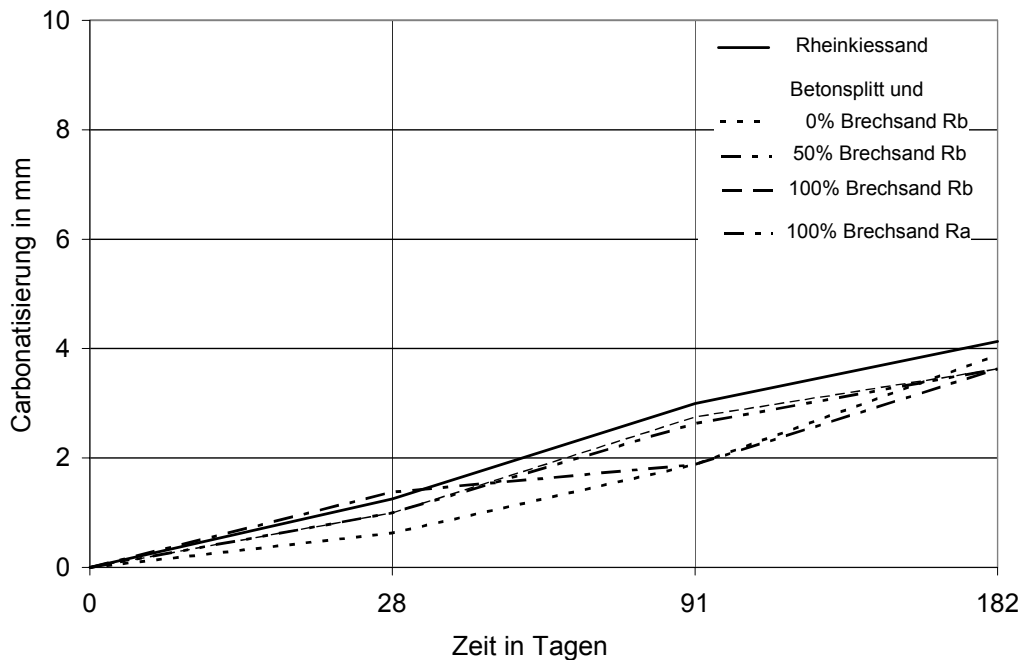


Abb. 6.7 Carbonatisierung von Beton mit rezyklierter Gesteinskörnung im Sandanteil [Siebel1998]

Im Unterschied zur Tab. 6.1 liegen in Abb. 6.7 zunächst keine diskreten Wertereihen vor. Der Aufwand für das Herauslesen solch diskreter Wertepaare scheint bzgl. der Beschreibung des Carbonatisierungsfortschritts auch wenig sinnvoll, da entsprechend exakt formulierte Abfragen auf das Diagramm nicht zu erwarten sind. Bei den zu erwartenden Abfragen wird es sich wohl kaum um detaillierte wertbehaftete Recherchen bzgl. der dargestellten Parameter handeln, wie bspw. „Suche alle Ressourcen, in denen zum Zeitpunkt *t* die Carbonisierungstiefe *c* nicht überschreitet“. Vielmehr sind Abfragen wie „Suche alle Ressourcen, in denen der Parameter Carbonisierungstiefe über den Zeitraum *t* untersucht wird“ wahrscheinlich, in denen der Schwerpunkt der Fragestellung auf Ergebnissen zu bestimmten Parameterkombinationen liegt.

In solchen Fällen sowie bei sehr umfangreichen Wertereihen ist die Anwendung des met@BiM-Elements *metabim:wertebereich* sinnvoll. Zugunsten der Verringerung des Erfassungsaufwands wird mit diesem Element eine unscharfe Beschreibung der Parameterwerte in Kauf genommen, da lediglich der untere und obere Grenzwert des abgebildeten Parameters angegeben werden. Die algebraische Auswertung, ob sich der Abfragewert innerhalb des Wertebereiches befindet oder nicht, erfolgt auf Seiten der abfragenden Software. Die Anwendung von *metabim:wertebereich* wird im folgenden Ausschnitt der RDF-Beschreibung von Abb. 6.7 zu den Eigenschaften „Carbonatisierungstiefe“ sowie „Messzeitpunkt“ demonstriert:

...

```
<metabim:carbonatisierung parseType="Resource">
  <rdf:type resource=
    "http://www.b-i-m.de/2001/01/metabim#FESTBETON"/>
  <metabim:einheit>mm</metabim:einheit>
  <metabim:wertebereich parseType="Resource">
    <metabim:untergrenze>0</metabim:untergrenze>
    <metabim:obergrenze>4.1</metabim:obergrenze>
  </metabim:wertebereich>
</metabim:carbonatisierung>

<metabim:messzeitpunkt parseType="Resource">
  <rdf:type resource=
    "http://www.b-i-m.de/2001/01/metabim#FESTBETONPRUEFUNGEN"/>
  <metabim:einheit>d</metabim:einheit>
  <metabim:wertebereich parseType="Resource">
    <metabim:untergrenze>0</metabim:untergrenze>
    <metabim:obergrenze>180</metabim:obergrenze>
  </metabim:wertebereich>
</metabim:messzeitpunkt>
```

...

Möchte man relative Werte, wie z.B. in Abb. 6.7 den Anteil an Brechsand in den untersuchten Betonproben, beschreiben, ist es notwendig, mindestens zwei Werte, nämlich die Bezeichnung der Komponente und die Höhe des Anteils, miteinander zu koppeln. In met@BiM wurde dies über das Element *metabim:anteil* realisiert. Die Höhe des Anteils wird zusammen mit der Komponentenbezeichnung (hier der Brechsand) als Attributwerte dem Element *rdf:value* bzw. *rdf:li* zugeordnet:

...

```
<metabim:zusammensetzung parseType="Resource">
  <rdf:type resource=
    "http://www.b-i-m.de/2001/01/metabim#ZUSCHLAG"/>
  <metabim:einheit>%</metabim:einheit>
  <rdf:li rdf:value="Brechsand aus B15" metabim:anteil="100"/>
  <rdf:li rdf:value="Brechsand aus B45" metabim:anteil="0"/>
  <rdf:li rdf:value="Brechsand aus B45" metabim:anteil="50"/>
  <rdf:li rdf:value="Brechsand aus B45" metabim:anteil="100"/>
  <rdf:li rdf:value="Rheinkiessand" metabim:anteil="100"/>
</metabim:zusammensetzung>
```

...

Das grafische RDF-Modell für die betontechnologischen Aussagen ist in Abb. 6.8 dargestellt. Der vollständige RDF-Code zur Beschreibung des Diagramms befindet sich im Anhang D2.



Abb. 6.8 RDF-Graph von *dc:subject* zu den Aussagen in Abb. 6.7 unter Verwendung der Wertebereiche zur Datenabbildung

Betonrezepturen

Die Beschreibung der betontechnologischen Daten aus Werkstoffprüfungen wurde in diesem Abschnitt mit den RDF-Beschreibungen von Betonmischungen zusammengefasst. Dies macht deshalb Sinn, da es sich bei den Rezepturen lediglich um einen bestimmten Satz von Betoneigenschaften handelt, der als Anforderung an eine Mischung formuliert wird. Diese Werkstoffkennwerte charakterisieren den Beton im Wesentlichen hinsichtlich seiner Verwendung sowie der Verarbeitbarkeit. Im Allgemeinen werden folgende Eigenschaften beschrieben:

- Betondruckfestigkeitsklasse
- Verarbeitbarkeit
- w/z-Wert
- Sieblinie
- Zementtyp
- Besondere Anforderungen (Wasserundurchlässigkeit, Außenbauteileignung etc.)

Tab. 6.2 Betonzusammensetzungen in kg/m³ nach [Reinhardt1998]

		Serie 1	Serie 2	Serie 3
Zement CEM I 32,5 R		240	240	350
Wasser		144	144	175
w/z - Wert		0,6	0,6	0,5
Rezyklat	0 – 2 mm	-	429	-
	2 – 4 mm	-	187	-
	4 – 8 mm	285	281	261
	8 – 16 mm	415	429	380
Sand	0 – 2 mm	667	-	610
	2 – 4 mm	274	-	251

Für alle in der Aufzählung angeführten Eigenschaften werden Metadatenelemente im met@BiM-Schema definiert. Wie in den anderen RDF-Beschreibungen von BiM-Online-Ressourcen müssen jedoch auf der übergeordneten Hierarchieebene zunächst die allgemeinen DC-Metadaten angegeben werden. Unter Berücksichtigung der in Tabelle 4.5 beschriebenen Erläuterungen zur Anwendung der Dublin Core-Elemente auf Betonrezepturen ergibt sich für die in Tab. 6.2 beschriebenen Rezepturen folgender RDF-Code für die DC-Metadaten:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:metabim="http://www.b-i-m.de/2001/01/metabim#">
  <rdf:Description about="http://www.b-i-m.de/berichte/e10/e10z0998.htm#6">

    <dc:type>http://www.b-i-m.de/2001/01/metabim#BETON</dc:type>
    <dc:format>text/html</dc:format>
    <dc:identifier>http://www.b-i-m.de/berichte/e10/e10z0998.htm#6
    </dc:identifier>
    <dc:date>1999-10-09</dc:date>
    <dc:title>Betonzusammensetzungen mit rezykliertem Zuschlag aus
      Leichtbeton</dc:title>
    <dc:creator>Prof. Dr.-Ing. H.-W. Reinhardt, Dipl.-Ing. J.Kümmel
    </dc:creator>
    <dc:publisher>Universität Stuttgart, Institut für Werkstoffe im
      Bauwesen (IWB)</dc:publisher>
    <dc:source>http://www.b-i-m.de/berichte/e10/e10z0998.htm</dc:source>
    <dc:relation>http://www.b-i-m.de/berichte/e10/e10z0198.htm#8
    </dc:relation>
    <dc:relation>http://www.b-i-m.de/berichte/e10/e10z0998.htm#1
    </dc:relation>
    <dc:description>Ziel des Vorhabens ist die Wiederverwendung von gebro-
      chenem Leichtbeton als rezyklierter Gesteinskörnung.
      Die Betonrezeptur wurde für Schwind- und Kriechversuche
      entworfen.</dc:description>

  ...

```

Der Beschreibung einer Betonrezeptur wird ebenso wie der Beschreibung eines Betonprüfergebnisses im Element *dc:type* der Ressourcentyp *metabim:BETON* zugewiesen. Auch die Belegung der anderen DC-Elemente erfolgt fast identisch zu den vorigen Beispielen.

Die *dc:relation* Elemente verweisen allerdings nicht auf die aus der Mischung gewonnenen Betonprüfdaten, da eine solche Vorgehensweise in den meisten Fällen zu einer unübersichtlichen Anzahl von *dc:relation* Elementen geführt hätte. Stattdessen wird darin auf URIs verwiesen, die auf weitere Varianten der Rezepturen zeigen (z.B. aus einer Versuchsreihe).

Die Aussagen zu betontechnologischen Daten der Tab. 6.2 werden innerhalb des *dc:subject*-Elements zusammengefasst und vervollständigen die RDF-Beschreibung:

...

```
<dc:subject parseType="Resource">

  <metabim:bezeichnung parseType="Resource">
    <rdf:type resource=
      "http://www.b-i-m.de/2001/01/metabim#ZEMENT"/>
    <rdf:value>CEM I 32,5R</rdf:value>
  </metabim:bezeichnung>

  <metabim:zementgehalt parseType="Resource">
    <rdf:type resource=
      "http://www.b-i-m.de/2001/01/metabim#REZEPTUR"/>
    <metabim:einheit>kgm3</metabim:einheit>
    <rdf:li>240</rdf:li>
    <rdf:li>350</rdf:li>
  </metabim:zementgehalt>

  <metabim:gesamtwassergehalt parseType="Resource">
    <rdf:type resource=
      "http://www.b-i-m.de/2001/01/metabim#REZEPTUR"/>
    <metabim:einheit>kgm3</metabim:einheit>
    <rdf:li>144</rdf:li>
    <rdf:li>175</rdf:li>
  </metabim:gesamtwassergehalt>

  <metabim:wzwert parseType="Resource">
    <rdf:type resource=
      "http://www.b-i-m.de/2001/01/metabim#REZEPTUR"/>
    <rdf:li>0.5</rdf:li>
    <rdf:li>0.6</rdf:li>
  </metabim:wzwert>

  <metabim:zuschlaggehalt parseType="Resource">
    <rdf:type resource=
      "http://www.b-i-m.de/2001/01/metabim#REZEPTUR"/>
    <metabim:einheit>kgm3</metabim:einheit>
    <rdf:li>1641</rdf:li>
    <rdf:li>1326</rdf:li>
    <rdf:li>1502</rdf:li>
  </metabim:zuschlaggehalt>

  <metabim:bezeichnung parseType="Resource">
    <rdf:type resource=
      "http://www.b-i-m.de/2001/01/metabim#ZUSCHLAG"/>
    <rdf:li>Rezyklat</rdf:li>
    <rdf:li>Sand</rdf:li>
  </metabim:bezeichnung>

  <metabim:zusammensetzung parseType="Resource">
    <rdf:type resource=
      "http://www.b-i-m.de/2001/01/metabim#ZUSCHLAG"/>
```

```

    <metabim:einheit>mm</metabim:einheit>
    <rdf:li>0/2</rdf:li>
    <rdf:li>2/4</rdf:li>
    <rdf:li>4/8</rdf:li>
    <rdf:li>8/16</rdf:li>
  </metabim:zusammensetzung>

</dc:subject>

</rdf:Description>
</rdf:RDF>

```

Die in diesem Abschnitt erläuterte Methode der strukturierten Beschreibung von Webinhalten mit met@BiM- und DC-Elementen erweitert die Recherchierbarkeit der Informationsquellen und ermöglicht neue Suchfunktionalitäten. Um den erforderlichen Beschreibungsaufwand insgesamt gering zu halten, werden jedoch auch gewisse Einschränkungen in Kauf genommen. Die Vorteile und die Einschränkungen der vorgeschlagenen Methode sollen im folgenden Abschnitt einander gegenübergestellt werden.

6.3 Diskussion der met@BiM-RDF-Modelle

Die Leistungsfähigkeit von Metadatenschemata, wie dem hier vorgeschlagenen met@BiM, lässt sich an den beiden konträren Parametern, nämlich der Vollständigkeit der Darstellung der Informationsressource und der möglichst einfachen Anwendbarkeit zur Modellbildung messen. Es wäre nicht erstrebenswert, das Datenmodell lediglich bezüglich eines der beiden Parameter zu optimieren. Eine möglichst vollständige Abbildung aller denkbaren Abfrageparameter wäre eine der beiden Extrempositionen. Solch ein vollständiges Datenmodell wäre jedoch bei vielschichtigen Informationsressourcen, wie den o.g. Diagrammen, Tabellen, Dokumenten etc., mit einem hohen Erzeugungsaufwand verbunden und bezüglich seines strukturellen Aufbaus sehr komplex.

Vielmehr gilt es, zwischen den beiden Extremen das richtige Maß für die Detaillierung der Datenstrukturen zu finden, die weder den Autor bei der Erzeugung in ihrer Komplexität überfordern noch die (realistischen) Erwartungen der Anwender bei fachlichen Recherchen unbefriedigt lassen. Um dieses Ziel zu erreichen, wird für den Entwurf eines fachspezifischen Metadatenmodells neben informationstechnologischen Kenntnissen auch das Wissen um die Denkweise der Anwender und um ihre potentiellen Abfragen benötigt.

Welche erweiterten Recherchemöglichkeiten bzw. welche Entscheidungsspielräume bzgl. der Modellierung von Webinhalten eröffnet unter diesen Gesichtspunkten die mit dem met@BiM-Schema vorgeschlagene Methode? Zunächst ermöglicht eine Metadatenbeschreibung die Recherchierbarkeit betontechnologischer Ressourcen unabhängig von den gewählten Darstellungsformaten (vgl. Abschnitt 3.3). Durch die Integration von Metadatenelementen des Dublin Core-Standards wird die breite Nutzbarkeit der RDF-Beschreibungen auch dann sichergestellt, wenn bis auf weiteres keine Suchdienste für fachspezifische Metadatenstrukturen zur Verfügung stehen.

Weiterhin werden durch die met@BiM-Elemente Abfragen ermöglicht, in denen mehrere Baustoffeigenschaften mit Operatoren der Booleschen Algebra verknüpft werden können. Somit lassen sich Abfragen nach Informationsquellen ausführen, in denen bspw. die Festbetoneigenschaft Druckfestigkeit in Zusammenhang mit der Trockenrohdichte der rezyklierten Zuschläge dargestellt wird. Durch die Angabe von (Grenz-)Werten für die Baustoffeigenschaften innerhalb solcher Abfragen lässt sich das Rechercheziel weiter eingrenzen. Zum Auffinden der Tab. 6.1 könnte man z.B. eine Abfrage auf alle Informationsquellen

beschränken, die eine Druckfestigkeit mit einem Wert größer 35 N/mm² enthalten und diesen Parameter in Zusammenhang mit der Trockenrohdichte der Zuschläge betrachten.

Trotz dieser detaillierten Recherchemöglichkeiten erlaubt die in dieser Arbeit vorgestellte Vorgehensweise eine relativ rasche Beschreibung von betontechnologischen Informationsquellen. Ein wesentlichen Einflussfaktor auf den Erzeugungsaufwand steckt, neben der Softwareunterstützung durch einen Editor, in der Möglichkeit, mit met@BiM betontechnologische Aussagen in mehrstufiger Detailliertheit beschreiben zu können. met@BiM überlässt dabei die Detaillierung und damit die Entscheidung über den betriebenen Beschreibungsaufwand dem Autor der Metadaten. So können lediglich die (wichtigsten) Baustoffparameter über die Integration der entsprechenden Metadatenelemente in den RDF-Code angegeben werden, jedoch auch zusätzlich die Werte dieser Parameter als Wertebereiche oder Wertereihen. Bei der Entscheidung, welche Aussagen in den Metadatenelementen abgebildet werden sollen, ist jedoch stets zu berücksichtigen, dass selbstverständlich nur Informationen gefunden werden können, die in den Metadaten beschrieben wurden oder aus ihnen ableitbar sind.

Wie schon im vorangegangenen Abschnitt erläutert wurde, lässt sich für die mit met@BiM-Elementen beschriebenen Webinhalte aus der Verknüpfung der Baustoffeigenschaften mit den betontechnologischen Ressourcen im met@BiM-RDF-Schema automatisch ein semantischer Kontext ableiten. Durch konsequente Verwendung gültiger URIs als Zeiger kann dieser Kontext im WWW auch eindeutig definiert werden. Mit der Angabe von URIs, die auf andere Informationsressourcen oder RDF-Beschreibungen verweisen, lassen sich auch explizit weitere inhaltliche Verknüpfungen erzeugen. Weiterhin können auf spezifische Wissensgebiete spezialisierte Dienste Algorithmen enthalten, die aus den Wertangaben innerhalb bestimmter Schemaelemente implizite Zusammenhänge herstellen können. Aus dem Auftreten von Zeichenketten wie „Uni“, „TH“ oder „TU“ im Metadatenelement *dc:publisher* lässt sich dann z.B. schließen, dass es sich bei der Informationsressource um eine Hochschulveröffentlichung handelt. Die Möglichkeiten, welche sich hieraus für komplexere Schlussfolgerungen aus den Definitionen der met@BiM- und DC-Elemente ergeben, werden im Abschnitt 6.6 dieser Arbeit angesprochen. Schwerpunkt dieser Untersuchung ist es jedoch, ein grundlegendes semantisches Datenmodell zu entwickeln, welches als Basis für solche inhaltspezifischen Funktionalitäten genutzt und ausgebaut werden kann.

Bei der Beschreibung von betontechnologischen Webinhalten durch met@BiM-RDF-Modelle wird allerdings auch auf die Abbildung bestimmter Aussagedetails zugunsten eines geringeren Eingabeaufwands verzichtet. So ergibt sich aus der getrennten Beschreibung der Baustoffparameter, dass Zusammenhänge zwischen den Einzelwerten nicht abgebildet werden. D.h. anhand der RDF-Beschreibung von Abb. 6.7 kann nicht ausgesagt werden, dass bei einem Probenalter von 28 Tagen die Carbonatisierungstiefe 0.6 mm betrug, jedoch dass in dem Diagramm Proben im Alter von 28 Tagen untersucht wurden und im Diagramm Carbonatisierungstiefen von 0 bis 4.1 mm (d.h. vermutlich auch 0.6 mm) gemessen wurden. Weiterhin werden bei der Verwendung von Wertebereichen zur Beschreibung der Betoneigenschaften die Abfragen unscharf. Eine Recherchesoftware kann zwar feststellen, ob sich ein Abfragewert innerhalb des angegebenen Wertebereichs der Ressource befindet, jedoch ist nicht sichergestellt, dass dieser Wert auch tatsächlich gemessen und dargestellt wird. In einigen Fällen können somit auch unpassende Quellen als Treffer angezeigt werden. Die in diesem Abschnitt genannten Möglichkeiten und Einschränkungen des met@BiM-RDF-Schemas sind in der Tab. 6.3 noch einmal einander gegenübergestellt.

Ein wichtiger Punkt ist auch, dass die manuelle Programmierung des komplexen RDF-Codes den Webautoren schwerlich zugemutet werden. Vielmehr muss einem Metadatenautor der Aufwand zur Erzeugung einer RDF-Beschreibung seiner Webressourcen so weit wie möglich abgenommen werden. Im folgenden Abschnitt wird mit dem met@BiM-RDF-Editor eine zu diesem Zweck entwickelte Software beschrieben.

Tab. 6.3 Möglichkeiten und Einschränkungen des met@BiM-RDF-Schemas zur Beschreibung von betontechnologischen Informationsquellen

Nutzungsmöglichkeiten und -grenzen von met@BiM-Datenstrukturen	
Vorteile	• Recherchierbarkeit von betontechnologischen Informationsressourcen ist unabhängig von deren Darstellungsformat. • Durch Integration des DC-Metadatensatzes bleibt ein Teil der Funktionalität erhalten, selbst wenn vorerst noch keine darüber hinausgehenden RDF Suchmaschinen zur Verfügung stehen. • Kombinierte Abfrage von Baustoffeigenschaften mit Booleschen Operatoren (und/oder/nicht) ist möglich. Beispiel: Finde alle Tabellen oder Diagramme, in denen die Parameter X UND (Y ODER Z) einander gegenübergestellt werden. • Innerhalb von Abfragen können den Baustoffeigenschaften (Grenz-) Werte und Maßeinheiten zugeordnet werden. Beispiel: Finde alle Diagramme, in denen die Parameter X mit dem Wert n UND (Y mit dem Wert n ODER Z mit dem Wert p) einander gegenübergestellt werden. • Die gewählte Vorgehensweise erlaubt eine rasche Beschreibung von Informationsquellen anhand der Skalierbarkeit der Beschreibungsdetailiertheit <i>Beispiel: 1.Stufe - Beschreibung der abgebildeten Baustoffeigenschaften; 2.Stufe - Beschreibung der Werte in Wertebereichen; 3.Stufe - Beschreibung der Einzelwerte.</i> • Der semantische Kontext von Webinhalten kann aus den entwickelten RDF-Beschreibungen automatisch abgeleitet werden • Implizit und explizit im Datenmodell enthaltene semantische Verknüpfungen sind automatisch nachvollziehbar
Einschränkungen	• Die Metadatenerzeugung bedeutet immer einen Mehraufwand • Abfragen mit diskreten Wertangaben für Baustoffeigenschaften werden bei der Verwendung von Wertebereichen unscharf. • Es wird keine Kopplung von Werten der einzelnen Parameter implementiert, da der hierdurch verursachte große Erfassungsaufwand in keinem Verhältnis zur erwarteten Nutzung steht. Jeder Parameter wird separat beschrieben. <i>Beispiel: Finde alle Diagramme, in denen die Baustoffeigenschaften X und Y das gemeinsame Wertepaar n, m besitzen.</i>

6.4 met@BiM-RDF-Editor

Im Abschnitt 6.2 wurde gezeigt, dass der RDF/XML-Code bezüglich seiner Syntax sehr umfangreich und komplex anwachsen kann. Sollen Autoren ihre Webdokumente mit Metadatenbeschreibungen versehen, darf dies nicht bedeuten, dass sie zusätzlich eine neue Programmiersprache erlernen müssen. Ein RDF-Editorprogramm reduziert hier den Aufwand für den Anwender auf die Eingabe der eigentlichen Daten in eine formularartige Oberfläche. Die Erzeugung der RDF/XML-Syntax selbst erfolgt automatisch durch den Editor.

In den folgenden Abschnitten wird der Aufbau und die Anwenderoberfläche des met@BiM-RDF-Editors vorgestellt. Es handelt sich dabei um eine Software, die speziell zur Erzeugung

von RDF-Beschreibungen für die in BiM-Online enthaltenen Ressourcentypen entwickelt wurde. Die für BiM-Online-Ressourcentypen spezifischen Eingabeoberflächen erlauben eine zügige Erfassung der Metadaten, ohne nennenswerte Einarbeitungszeiten vorauszusetzen.

6.4.1 Programmstruktur

Beim met@BiM-RDF-Editor handelt es sich um einen Online-Editor, d.h. das Programm läuft auf einem Webserver und wird über einen Webbrowser bedient. Die Datenerfassung erfolgt mit Hilfe von HTML-Formularen (vgl. 2.3), in denen der Anwender seine Eingaben vornehmen kann. Der RDF-Generator, welcher in der Programmiersprache PHP erstellt wurde und der die zentrale Programmkomponente des Editors darstellt, erzeugt dann aus den Eingaben automatisch den RDF/XML-Code für die Informationsressourcen. Der Quellcode der RDF-Beschreibung wird im unteren Teil des Browserfensters angezeigt (Abb. 6.9) und kann z.B. über die Kopieren/Einfügen-Funktion in modernen Betriebssystemen in den Quellcode von HTML-Dokumenten integriert werden.

Die RDF-Ressourcen und -Eigenschaften des met@BiM-Schemas sind dabei nicht in den Programmcode des Editors fest integriert, sondern werden in separaten Textdateien getrennt von den ausführbaren Programmteilen gespeichert. Angaben zu den met@BiM-Ressourcentypen, den Baustoffeigenschaften sowie den vordefinierten Maßeinheiten lassen sich in solchen übersichtlichen Textdateien leicht editieren. Dies vereinfacht die Übernahme von Änderungen und Erweiterungen des Schemas in den Editor erheblich. Besonders während der Entwicklung des met@BiM-Schemas an den BiM-Online-Inhalten erwies sich diese Flexibilität als sehr hilfreich, da nun die Einführung einer neuen Baustoffeigenschaft bzw. (Sub-) Ressource in den Editorauswahllisten nur noch den Aufwand eines weiteren Eintrags in der Textdatei bedeutete.

Um die Eingabeoberfläche übersichtlich und effizient zu gestalten, musste diese jedoch an die Besonderheiten der Ressourcentypen bezüglich ihrer charakteristischen Parameter angepasst werden.

6.4.2 Eingabeoberflächen für die einzelnen Ressourcentypen

Die Bedienung des met@BiM-RDF-Editors erfolgt in drei Stufen. Nach dem Aufruf des Programms unter der URL „<http://www.b-i-m.de/metabim/editor/>“ wird der Anwender zunächst aufgefordert, den Ressourcentyp auszuwählen, den er beschreiben möchte. Er hat dabei die Auswahl zwischen:

- Betontechnologische Prüfergebnisse
- Betonrezepturen
- Dokumente, Kurse oder Formeln

Die Gruppierung der Ressourcentypen in dieser Auswahl erfolgte anhand der Metadatenelemente, die zur Beschreibung der Ressourcen nach Abschnitt 6.2 benötigt werden. Im zweiten Schritt wird dem Anwender dann in der oberen Hälfte eines zweigeteilten Browserfensters die Eingabeoberfläche für die Metadaten angezeigt. Die Eingabeformulare besitzen alle einen tabellarischen Aufbau (Abb. 6.9): In der linken farblich abgesetzten Spalte, werden die Namen der Metadatenelemente angezeigt. Rechts davon befinden sich jeweils die Eingabefelder bzw. die Auswahlfelder von Metadatenelementen mit vordefinierten Werten. Am rechten Rand der Tabelle sollen Erläuterungen zur Bedeutung der Metadatenelemente dabei helfen, ein Grundmaß an Konsistenz bei der Werteeingabe sicherzustellen. In ihrem vertikalen Aufbau werden die Unterschiede zwischen den drei Eingabeformularen jedoch deutlicher. Alle drei beginnen zunächst mit den Metadatenelementen des Dublin Core

Standards. Unterschiede treten hier lediglich bei den Erläuterungen zum Inhalt der Metadatenfelder (vgl. Tab. 4.5) sowie bei der Belegung der Auswahlfelder von *dc:type* und *dc:format* auf.

met@BiM-Metadateneditor für Dokumente, Kurse (Lektionen) und Formeln V 1.2

Ressource:	http://129.69.59.111/dialog/beton/chemwid1_1.htm	(URI je nach Ress.-typ: URL; URL + #kurs; URL + #math oder #chem+ ID)
dc:type:	Lektion / Kurs	(Ressourcentyp)
dc:format:	Textfragment / ML-Code (HTML, MathML...)	(Ressourcenformat)
dc:identifier:	http://129.69.59.111/dialog/beton/chemwid1_1.htm	(XPointer-Adresse)
dc:title:	Festbetoneigenschaften - Widerstand gegen chemische Angr	(Überschrift/Titel)
dc:creator:	Matthias Luik; Marcus Schreyer	(Autor: Vorname Nachname;)
dc:publisher:	Universität Stuttgart	(Institution/Firma)
	Institut für Werkstoffe im Bauwesen (IWB)	(Abteilung)
dc:source:	http://129.69.59.111/dialog/beton/druckf1_1.htm	(URL der Website / Rubrik / Dokument)
dc:relation:	http://129.69.59.111/dialog/beton/emodul1_1.htm (1)	
	http://129.69.59.111/dialog/beton/sbzug1_1.htm (2)	(URI's zugehöriger Ressourcen: z.B. Dokument - and. Z-Berichte; Kurs - and. Kurse z.Thema;
	http://129.69.59.111/dialog/beton/kriechen1_1.htm (3)	

```
<Description about="http://129.69.59.111/dialog/beton/chemwid1_1.htm">
<dc:type>http://www.b-i-m.de/2001/01/bmd#KURS</dc:type>
<dc:format>text</dc:format>
<dc:identifier>http://129.69.59.111/dialog/beton/chemwid1_1.htm</dc:identifier>
<dc:date>21.03.2002</dc:date>
<dc:title>Festbetoneigenschaften - Widerstand gegen chemische Angriffe</dc:title>
<dc:creator>Matthias Luik; Marcus Schreyer</dc:creator>
<dc:publisher>Universität Stuttgart</dc:publisher>
<dc:publisher>Institut für Werkstoffe im Bauwesen (IWB)</dc:publisher>
<dc:source>http://129.69.59.111/dialog/beton/druckf1_1.htm</dc:source>
<dc:relation>http://129.69.59.111/dialog/beton/emodul1_1.htm</dc:relation>
<dc:relation>http://129.69.59.111/dialog/beton/sbzug1_1.htm</dc:relation>
```

Abb. 6.9 Browseransicht des met@BiM-RDF-Editors. Die Ausgabe des RDF/XML-Codes erfolgt im unteren Fenster.

Der wichtigste Unterschied zwischen den Eingabeoberflächen des Editors besteht jedoch in der Gestaltung des Elements *dc:subject* zur Beschreibung der baustoffspezifischen Aussagen. Für die Ressourcentypen *metabim:DOKUMENT*, *metabim:KURS* sowie für die Gleichungen wurde der Bereich von *dc:subject* lediglich für mehrfache Eingaben erweitert. Im RDF-Code dieser Ressourcentypen werden diesem Element keine weiteren Metadatenelemente untergeordnet, sondern lediglich atomare Werte zugewiesen. Anders stellt sich die Situation bei den Eingabefeldern für die Rezepturen und die Prüfdaten dar. Um den Anwender bei der Eingabe der Metadaten für Betonrezepturen zu unterstützen, wurde für diesen Ressourcentyp ein eigenes Eingabeformular entworfen, welches für alle charakteristischen Eigenschaften von Betonrezepturen bereits insgesamt 106 Eingabefelder enthält (Abb. 6.10). Auch hier sorgen Erläuterungen und Auswahlfelder dafür, dass bei den Dateneingaben eine konsistente Wertemenge entsteht. Die Eingabefelder zur Beschreibung einer Betonrezeptur sind dabei

zum großen Teil optional. D.h. auch bei Benutzung des met@BiM-Editors liegt es im Ermessen des Anwenders, wie detailliert er seine Webinhalte beschreiben möchte.

Noch komplexer ist der Entwurf einer Eingabeoberfläche für die betontechnologischen Daten. Ein Formular, welches alle sinnvollen Kombinationen zwischen betontechnologischen Ressourcentypen und Baustoffmetadaten enthalten würde, wäre für die Beschreibung aller 161 möglichen Baustoffparameter von met@BiM und den diesen zugeordneten Werten sicherlich nicht praktikabel. Es galt daher, eine möglichst flexible Lösung zu finden, die durch ihre Strukturierung den Anwender bei der Beschreibung der Versuchsparameter unterstützt und gleichzeitig den Umfang der in met@BiM enthaltenen Beschreibungsmöglichkeiten nicht einschränkt.

Als Ergebnis entstand ein Eingabeformular, welches die Zuordnung von bis zu 5 Baustoffparametern, bestehend aus met@BiM-Ressourcen-Eigenschaftpaaren, zu einer Informationsressource erlaubt. Die Betonressourcen und Baustoffeigenschaften des met@BiM-Schemas werden dabei in Auswahlfeldern angezeigt (Abb. 6.11), deren Inhalte aus den oben angesprochenen Textdateien geladen werden. Für jeden dieser Parameter lassen sich in beigeordneten Eingabefeldern Werte angeben. Die Werteingabe kann entweder in Form von Wertebereichen oder über die Angabe von bis zu 15 Einzelwerten erfolgen. Sollten in einem Diagramm oder einer Tabelle mehr als 5 Baustoffparameter beschrieben werden, was bei Testanwendungen mit Informationsquellen aus BiM-Online häufiger vorkam, so lassen sich diese weiteren Parameter innerhalb eines erneuten Durchlaufs im Editor generieren und nachträglich in den zuvor erzeugten RDF/XML-Code an entsprechender Stelle ergänzen.

Die Generierung des Codes wird nach erfolgter Eingabe der Metadaten per Mausklick auf die „starten“ Schaltfläche am unteren Ende des Formulars ausgelöst. Der in der unteren Fensterhälfte angezeigte RDF/XML-Code kann schließlich mittels Kopieren/Einfügen in einer eigenen Textdatei oder im Quellcode eines Dokuments abgespeichert werden. Die Aufteilung des Browsers in einen Formularbereich und einen Ausgabebereich besitzt den Vorteil, dass bei der Erfassung größerer Datenmengen die vorigen Eingaben innerhalb des Formulars auf dem Bildschirm erhalten bleiben. Gerade bei umfangreichen Tabellen und Diagrammen reduziert dies den Eingabeaufwand erheblich, da sich z.B. die DC-Metadaten für unterschiedliche Informationsressourcen innerhalb eines Dokuments nur wenig ändern. Die Möglichkeit, einzelne Bereiche des Formulars zurücksetzen zu können, war dann hilfreich, wenn durch versehentlich im Formular stehen gelassene Einträge Eingabefehler entstanden waren. So lassen sich bspw. Angaben zu einem bestimmten Parameter gezielt löschen, ohne das gleich alle anderen Eingaben ebenfalls zurückgesetzt werden müssen.

Insgesamt wurde durch die gewählte tabellarische Anordnung der Metadatenelemente eine übersichtliche Gruppierung der Eingabemöglichkeiten im Browserfenster erreicht. Die Eingabe von ca. 125 Probedatensätzen aus dem Zwischenberichtsjahrgang Herbst 1998 konnte auf diese Weise in ca. 30-40 Arbeitsstunden bewältigt werden.

met@BiM-RDF-Editor für Betonrezepturen - Microsoft Internet Explorer

dc:subject: Bezeichnung

Produkt-/Rezepturbezeichnung [1] Bezeichnung der Betonrezeptur

Anforderungen [Anforderungen - Inhalte löschen](#)

Konsistenz [2] Einheit Werte von bis Klasse Konsistenzangabe entweder als Ausbreitmaß oder als Konsistenzklasse

Druckfestigkeit [3] Klasse Klasse von bis Klasse

Besondere Anforderungen [Bes. Anforderungen - Inhalte löschen](#)

Eigenschaft ☒ wasserundurchlässig [4] ☐ hoher Frost-Tau-Widerstand [5] ☐ hoher Frost-Tausalz-Widerstand [6] ☐ hoher Widerstand gegen schwachen chemischen Angriff [7] ☐ Außenbauteil [8] ☐ Spannbetoneignung [9] Besondere Anforderungen, denen die Betonrezeptur genügt

Zusätzliche Rezepturangaben [Zus. Rezepturangaben - Inhalte löschen](#)

w/z-Wert [10] Wert Werte von bis

Zementtyp [11]

Abb. 6.10 Eingabeoberfläche für Metadaten zu Betonrezepturen

met@BiM-RDF-Editor für Baustoffprüfdaten - Frischbeton - Verarbeitbarkeit - Microsoft Internet Explorer

dc:subject: Parameter 1 [Parameter 1 Inhalte löschen](#)

Zuschlag (SubResource 1) Kornrohdichte (Property 1) (Beschreibung des betontechn. Parameters anhand eines (Sub) Resource-Property-Paares des met@BiM-Schemas)

Angabe von Werten [Angabe von Werten - Inhalte löschen](#)

Einheit Werte von bis (Numerische Grenzwerte der in der Ressource beschriebenen Daten)

ID	Num. Werte	Textwerte	Anteile
1	2620		
2	2480		
3	2430		
4	2180		
5	2310		
6	2400		
7	2000		
8	2490		
9			
10			
11			
12			
13			
14			
15			

Angabe von Einzelwerten [Angabe von Einzelwerten - Inhalte löschen](#) (Einzelwerte der in der Ressource beschriebenen Daten)

Parameter 2 [Parameter 2 Inhalte löschen](#)

Abb. 6.11 Eingabeoberfläche zur Erfassung betontechnologischer Prüfdaten

6.5 Implementierung von met@BiM

Die Implementierung von met@BiM in ein Informationssystem stellt den praktischen Anwendungsteil der Dissertation dar. In diesem Abschnitt wird der Aufbau einer Softwareumgebung erläutert, die anhand von ausgewählten RDF-Beschreibungen aus BiM-Online die Nutzung von met@BiM-RDF-Datensätzen in einem fachspezifischen Informationssystem bzw. einer Suchmaschine simulieren soll.

Im Rahmen des webbasierenden Testsystems, welches die Phasen Datenerfassung, Datenspeicherung sowie eine Abfrageschnittstelle umfasst, sollen hierbei folgende Fragestellungen untersucht werden:

- Lässt sich die Datenerfassung der verteilt im Web gespeicherten RDF-Beschreibungen in einer zentralen Indexdatenbank mit Hilfe einer Schnittstelle weitgehend automatisieren?
- Wie könnten fachspezifische Abfrageoberflächen für met@BiM-Datensätze aussehen?
- Welche zusätzlichen Funktionalitäten im Vergleich zur Volltextrecherche lassen sich mit den met@BiM-Datenstrukturen verwirklichen?

Aufgrund der frühen Anwendungsphase von RDF fehlt es noch an ausgereifter Software, mit der sich die Möglichkeiten von RDF-Datenstrukturen vollständig ausschöpfen lassen. Für das Testsystem wurden die in XML-Syntax verfassten RDF-Beschreibungen als XML-Dokumente in einer Datenbank gespeichert und unter Verwendung einer XML-Abfragesprache (X-Query) ausgefiltert. Die bisher überwiegend aus universitären Projekten entwickelten RDF-Parser und darauf aufgebauten RDF-Datenbanken (vgl. Abschnitt 5.7) beschränken sich in ihrer Programmausgabe auf Subjekt-Objekt-Prädikat Tripel der Aussagenlogik. Für eine anschauliche Weiterverarbeitung solcher Systemantworten in einer Benutzeroberfläche ist dieses Ausgabeformat zu rudimentär. Im hier vorgestellten Demonstrationssystem werden deshalb einige RDF-typische Funktionalitäten, wie z.B. die Interpretation des Graphen *rdf:Type* bei der Zuordnung von Betoneigenschaften zu den betontechnologischen Informationsressourcen, über entsprechende Recherchealgorithmen ermöglicht, die jedoch nicht allgemein auf andere RDF-Schemata übertragbar sind.

Die Datenbasis besteht hierbei im Wesentlichen aus den Inhalten von 13 ausgewählten Zwischenberichten verschiedener Teilprojekte aus „Baustoffkreislauf im Massivbau“ sowie Online-Kursen aus dem Bereich „BiM-Online-Dialog“. Im Testsystem wurde auf administrative Funktionalitäten, wie z.B. die Integration einer Benutzerverwaltung für die Zugriffssicherheit, eine komfortable Datenverwaltung sowie auf Updatefunktionen für die Indexdatenbank, verzichtet.

6.5.1 Speicherung der RDF-Datensätze im HTML-Quellcode

In Abschnitt 4.1.4 wurde bereits die Möglichkeit angesprochen, dass Metadatenätze sowohl in Informationsobjekten als auch getrennt von ihnen gespeichert werden können. Im Rahmen des Demonstrationssystems wurden die RDF-Datensätze als sogenannte XML-Fragmente in den Quellcode der HTML-Dateien integriert:

```
<html>
  <body>
    ...
  </p>
  <p><b><br>
  Tabelle 1: Versuchsprogramm je Betonsorte</b></p>
```

```

<a name="1"></a>
<!-- RDF-Metadaten Descriptions in HTML
<?xml version="1.0" encoding="ISO-8859-1"?>
  <rdf:RDF
    ...
    <rdf:Description about="...
      <dc:title>...
    ...
  </rdf:RDF>
-->
</body>
</html>

```

RDF/XML-Fragment
mit vorangestelltem
HTML-Anchor-Tag

Die Metadatenbeschreibungen in RDF/XML-Code, die sich mit Hilfe des met@BiM-RDF-Editors erzeugen lassen, werden als Kommentare zwischen den beiden Tags `<!-- . . . . -->` an den Stellen im Quellcode eingebunden, an welchen sich auch die Informationsressourcen befinden. Die Kommentar-Tags sind notwendig, um die Bildschirmanzeige der Werte von XML-Elementen als Text im Browser zu unterdrücken. Neben Kommentarzeichen wird der RDF/XML-Code auch durch ein HTML-Anchor-Tag (`<a name="...">`) ergänzt, das vom Preprozessor des Systems (vgl. Abschnitt 6.5.2) aus dem Wert des DC-Elements `dc:identifier` generiert wird. Das eingefügte Anchor-Tag erlaubt es, Informationsquellen innerhalb eines HTML-Dokuments auch ohne XPointer-Funktionalität (vgl. Abschnitt 2.6.2) direkt über die URL und den Fragment Identifier anzuspringen. Eine lästige Suche nach bestimmten Ressourcen innerhalb eines umfangreichen Dokuments kann somit vermieden werden.

Die hier beschriebene Vorgehensweise zur Integration von RDF-Beschreibungen in HTML-Dokumenten stellt allerdings nur eine Behelfslösung dar, da sie bezüglich der Interpretation des RDF-Codes durch eine XML/RDF-Software zu Widersprüchen führen kann. Inhalte von Quelltext-Kommentaren werden nach [HTML401Rec] grundsätzlich von Browsern oder Parsern nicht interpretiert, sondern bezüglich ihrer maschinellen Verarbeitung ignoriert. Innerhalb von Kommentar-Tags eingefügte RDF-Beschreibungen könnten daher auch von zukünftigen RDF-Crawlern oder ähnlichen Suchprogrammen ignoriert werden. In der RDF-Spezifikation wird in Abschnitt 7.7 deshalb noch eine weitere Methode der Integration in HTML vorgeschlagen. Bei dieser werden die Werte der RDF-Eigenschaften als Attribute der XML-Elemente codiert, so dass die Werte sich nicht mehr zwischen den spitzen Klammern, sondern innerhalb derselben befinden. Für die aktuellen Webbrowser stellen die dann verbleibenden XML-Elemente unbekannte Tags dar, die sie bezüglich der Darstellung im Browserfenster ignorieren. Auf die Kommentar-Tags kann bei dieser Methode jedoch verzichtet werden, so dass für zukünftige Programme kein Widerspruch zwischen interpretierbarem RDF-Code und der Auszeichnung als zu ignorierendem Kommentar besteht. Daher ist diese, in [RDF1999] beschriebene Methode, vor der im Rahmen des Demonstrationssystems realisierten Vorgehensweise zu bevorzugen.

Mit beiden Vorgehensweisen lassen sich die im met@BiM-RDF-Editor erzeugten RDF-Beschreibungen im Quelltext der HTML-Dokumente verbergen. Während für den Betrachter in einem Webbrowser kein Unterschied zwischen Webseiten mit RDF-Beschreibungen und normalen Webseiten zu erkennen ist, können entsprechende Suchdienste den RDF-Code in solchen im Web verteilten Dokumenten lesen und zur Erfassung der Informationsquellen in ihrer Indexdatenbank nutzen. Durch die Möglichkeit, mit RDF-Beschreibungen angereicherte Dokumente über eine weitgehend automatische Schnittstelle online zu erfassen, konnte eine möglichst realitätsnahe Situation mit einer verteilten Speicherung der RDF-Beschreibungen im Web simuliert werden.

6.5.2 Datenerfassung über den Preprozessor

Analog zum Vorgehen aktueller Suchmaschinen sollen die verteilt im Netz gespeicherten RDF-Beschreibungen in einer zentralen Indexdatenbank zusammengefasst werden. Der Vorgang der Datenerfassung soll dabei möglichst automatisch ablaufen.

Die im Demonstrationssystem realisierte Lösung bietet Anwendern über ein HTML-Formular die Möglichkeit, durch Angabe der URL einer Webseite die darin enthaltenen RDF-Beschreibungen in die Datenbasis des Systems zu laden. Da die RDF-Beschreibungen jedoch wie oben beschrieben in den HTML-Code eingebettet wurden, benötigt man zunächst eine Software, welche aus dem Quellcode der Dokumente die entsprechenden RDF/XML-Fragmente herauslesen kann. Zu diesem Zweck wurde für die Datenerfassungsschnittstelle ein spezielles Preprozessormodul programmiert. Es handelt sich dabei um ein als PHP-Skript erstelltes Programm, welches sich in Abb. 6.12 zwischen dem HTML-Formular und der eigentlichen Datenschnittstelle X-Node der Tamino Datenbank befindet. Das Preprozessormodul liest zunächst den gesamten Quellcode des Webdokumentes ein und reduziert ihn soweit, dass lediglich die RDF-Beschreibungen übrigbleiben. Diese RDF-Beschreibungen werden dann als RDF/XML-Fragmente zu einem vollständigen, gültigen RDF/XML-Dokument zusammengefasst. Das Wurzelement der aus allen Fragmenten eines Dokuments akkumulierten RDF-Beschreibungen erhält die Bezeichnung *rdf:RDF*. Zur korrekten Speicherung der RDF/XML-Datensätze in Tamino müssen diese gültig bezüglich einer als „Tamino Schema“ bezeichneten DTD sein (vgl. Abschnitt 2.6.1). Die Vorgaben dieser DTD orientieren sich dabei am Aufbau der in Abschnitt 6.2 beschriebenen Beispiele. Nur wenn die an Tamino gesandten XML-Dokumente diesen Vorgaben entsprechen, lassen sich die angestrebten Vorteile, z.B. von Abfragen mit vergleichenden Operatoren auf die Datensätze, in der XML-Datenbank auch anwenden.

Innerhalb von Tamino werden die RDF/XML-Dokumente erneut in einem XML-Baum zusammengefasst. Diese als RDF-Index bezeichnete XML-Struktur enthält auf diese Weise alle in das System geladene RDF-Beschreibungen und dient als zentrale Datenbasis für Recherchen auf dem Testsystem.

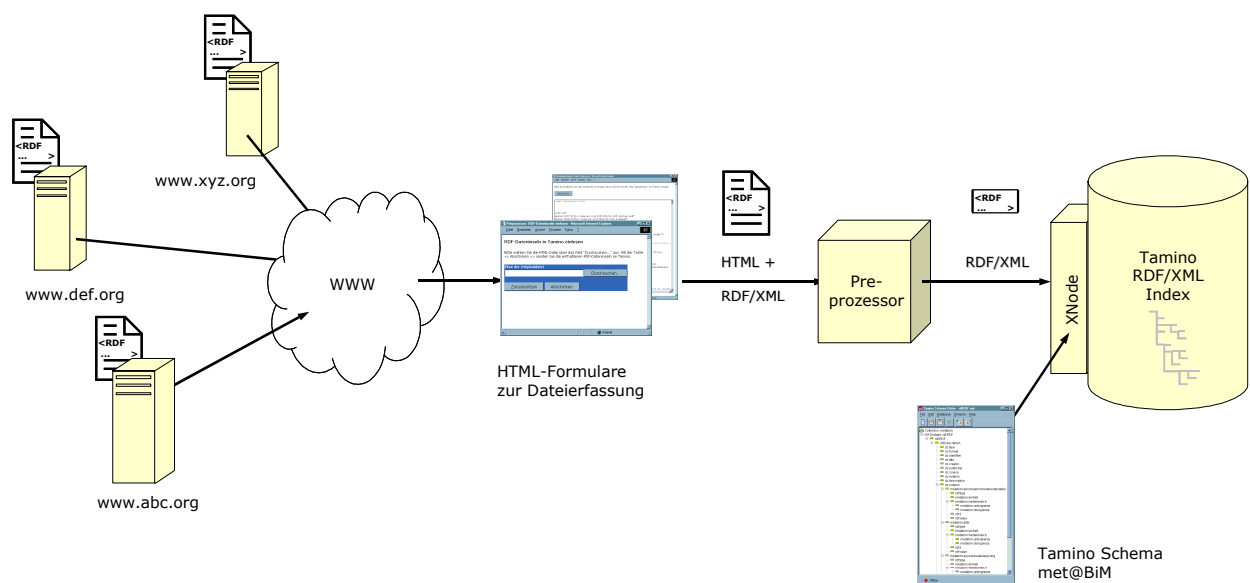


Abb. 6.12 Ablauf der Datenerfassung. Der Preprozessor extrahiert dabei die in den Webressourcen enthaltenen RDF-Beschreibungen

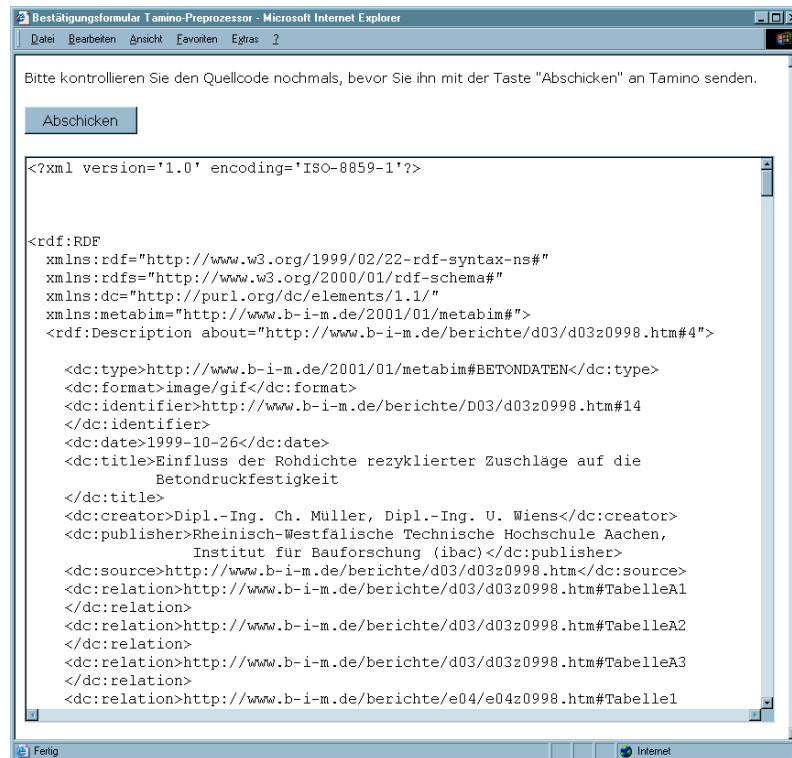


Abb. 6.13 Vor dem Abspeichern in Tamino wird das aus den RDF/XML-Fragmenten erzeugte XML-Dokument noch einmal angezeigt.

6.5.3 Aufbau einer RDF/XML-Indexdatenbank mit Tamino

Für die zentrale Speicherung der aus unterschiedlichen Webdokumenten erfassten RDF-Daten wird der Informations-Server Tamino verwendet. Bei Tamino handelt es sich um ein Produkt der Software AG [SoftwareAG2001] zur Speicherung und Verwaltung komplexer XML-Datenstrukturen. Es stellt den Anwendern weborientierte Abfrageschnittstellen zur Verfügung, um aus der gespeicherten Datenbasis gezielt bestimmte XML-Objekte herauszufiltern und in nachgeschalteten Programmen weiterzuverarbeiten.

Die über den Webbrowser zu bedienende grafische Oberfläche von Tamino unterstützt in der Version 2.3.1.1 lediglich administrative Funktionen der Datenbanken. Werkzeuge oder Vorlagen für grafische Benutzeroberflächen zur Datenabfrage bzw. Datenverwaltung, wie sie bspw. Bestandteil von MS Access sind, müssen vom Anwender selbst programmiert werden. Tamino wird im Testsystem als Aggregations-Server für die RDF-Datensätze verwendet. D.h. die über die im vorigen Kapitel beschriebene Datenerfassungsschnittstelle und den Preprozessor erfassten Daten werden in Tamino zu einem XML-Indexbaum zusammengefasst und verwaltet. Die Speicherung der RDF-Datensätze in Tamino ist deshalb möglich, weil es sich bei RDF/XML um eine wohlgeformte XML-Syntax (vgl. Abschnitt 2.6.1) handelt. Datensätze, die als XML-Objekte in der Tamino Datenbank abgespeichert werden, müssen, wie im vorigen Abschnitt angesprochen, gültig bezüglich einer DTD sein. Die Überprüfung der Gültigkeit von Datensätzen erfolgt in Tamino anhand einer als „Tamino Schema“ bezeichneten DTD. Darin werden die Elementnamen sowie die Eigenschaften der Dokumente definiert, die der Server als gültige XML-Dokumente akzeptiert. Zur Erzeugung solcher Schemata stellt Tamino einen grafischen Editor zur Verfügung.

Das Speicherkonzept von Tamino ist hierarchisch aufgebaut und orientiert sich am Adressierungskonzept des Internets bzw. am Aufbau von XML-Dokumenten nach dem Document Object Model (DOM) [DOM1] des W3C. An oberster Stelle steht die URI der

Datenbank, die sich aus der URL des Tamino Servers sowie der Erweiterung „/Datenbankname“ zusammensetzt. Innerhalb einer Datenbank können XML-Objekte, die nach demselben Tamino Schema aufgebaut sind, in sogenannten „Collections“ zusammengefasst und manipuliert werden. Zur Speicherung der RDF-Beschreibungen des met@BiM-Testsystems in Tamino genügte es, lediglich ein Tamino Schema für alle met@BiM-Ressourcentypen zu definieren. Diese Vereinfachung ergab sich aus dem sich ergänzenden Aufbau der RDF-Schemata für die met@BiM-Ressourcentypen: Die Metadaten-schemata aller Ressourcentypen des Testsystems lassen sich aus dem Schema für die Ressource *metabim:BETON* ableiten. Elemente, die nicht in allen Ressourcentypen vorkommen, müssen hierbei als optional gekennzeichnet werden.

Alle RDF-Datensätze, die in ihrem Aufbau den Vorgaben des in Abbildung 6.14 ausschnittsweise dargestellten Tamino Schemas entsprechen, werden als XML-Objekte innerhalb der Collection *metabim* abgespeichert. Innerhalb der Datenbank entsteht auf diese Weise eine Baumstruktur, bei der die über den Preprozessor erfassten RDF-Beschreibungen, unter dem Wurzelement *rdf:RDF* der Collection, die einzelnen Hauptäste bzw. Instanzen des Schemas bilden. Tamino-spezifische XML-Elemente mit der *ino:* Präfix werden den Datensätzen bei der Speicherung im Datenbanksystem automatisch hinzugefügt

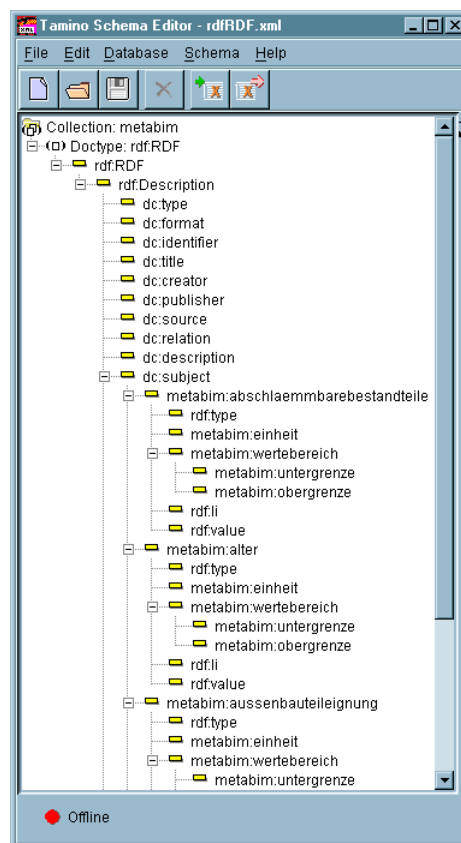


Abb. 6.14 Struktur der zur Speicherung der RDF-Beschreibungen in Tamino verwendeten DTD, dargestellt im Tamino Schema Editor.

6.5.4 Abfragemodul für met@BiM-Datensätze

Die Abfrageschnittstelle stellt dem Anwender die eigentliche Benutzeroberfläche für den Zugriff auf die Inhalte der Indexdatenbank zur Verfügung. Die Benutzeroberfläche des Testsystems wurde formularbasiert entworfen. Für den Aufbau der Abfrageformulare wurde

eine Kombination aus Auswahlfeldern mit vorgelegten met@BiM-Schemaelementen und Eingabefeldern für die Werte bzw. Wertebereiche entworfen. Die Inhalte der Auswahlfelder zur Angabe der Betonparameter sowie der Maßeinheiten werden dabei mit den für den Editor in Textdateien abgelegten Werten des met@BiM-RDFS vorgelegt. Änderungen im met@BiM-Schema können somit an zentraler Stelle in diesen Dateien vorgenommen werden und werden sofort in allen betroffenen Komponenten des Testsystems berücksichtigt. Die Übernahme solcher Änderungen direkt aus den entsprechenden RDF-Schemata wurde im Rahmen des Testsystems noch nicht verwirklicht. Für zukünftige RDF-Softwareprogramme stellen solche automatischen Verknüpfungen von Benutzeroberflächen und RDF-Schema-Dateien eine sinnvolle Funktion hinsichtlich konsistenter Programmoberflächen dar.

Die Benutzeroberfläche des Testsystems erlaubt, abhängig vom gewünschten Detaillierungsgrad der Recherche, die Formulierung von Abfragen in zwei grundsätzlich unterschiedlichen Abfrageoberflächen: Stichwortbasierende Recherchen lassen sich analog zu Volltextrecherchen über alle in der met@BiM-Ontologie definierten Informationsressourcen bzw. deren Metadatenbeschreibungen ausführen (Abb. 6.15). Einschränkungen auf bestimmte Ressourcentypen sind über die zugehörigen Kontrollkästchen auswählbar. Die Trefferlisten bestehen dann, sofern nicht eingeschränkt, aus Treffern zu den unterschiedlichen Ressourcentypen, die für den jeweiligen Treffer in der Zeile *Kontext* als Kette sich spezialisierender Ontologieknoten dargestellt werden. Für detailliertere Recherchen stehen weiterhin spezifischere Suchmasken zur Verfügung (Abb. 6.16), deren Eingabefelder die Formulierung gezielter Abfragen auf die Dokumente, betontechnologischen Inhalte etc. ermöglichen.

Die Trefferlisten (Abb. 6.17) bilden im oberen Bereich der Seite noch einmal die Spezifikationen aus den Suchmasken ab. Nach *Gesucht in* folgt der Ontologiebereich, in dem die Recherche ausgeführt wurde. In der Anzeige der eingegebenen Rechercheparameter werden nochmals die Benutzerangaben einschließlich eventueller (Grenz-)Werte wiederholt.

Die Trefferangaben enthalten die wichtigsten Eigenschaften der Informationsquellen, auf die sie verweisen. Neben obligatorischen Angaben wie dem Titel, Autor, einer Beschreibung u.ä. werden die Trefferbeschreibungen durch semantische Angaben, wie dem oben erwähnten Ressourcentyp im Abschnitt *Kontext* sowie der Aufzählung weiterer, in der Informationsquelle beschriebener Baustoffeigenschaften, ergänzt. Dem Anwender wird durch diese semantischen Angaben eine inhaltliche Einordnung und somit eine Einschätzung der Relevanz des angezeigten Treffers bezüglich seines Problems erleichtert, ohne dass er hierzu die jeweilige Informationsquelle aus dem WWW laden muss.

Eine weitere Besonderheit stellt das Angebot von Informationsquellen dar, die nicht direkt aus der Abfrage resultieren, jedoch inhaltlich mit dem jeweiligen Treffer in Beziehung stehen. Über den Hyperlink *Weitere Inhalte hierzu...* lassen sich im Testsystem inhaltlich mit der Informationsquelle verknüpfte Webinhalte anzeigen. Hierfür werden die Angaben in den Metadatenelementen *dc:relation* sowie *dc:source* herangezogen. Über diese Elemente lassen sich automatisch semantische Verknüpfungen herstellen, wie bspw. die Adresse der Mischungsbeschreibung, die dem angezeigten betontechnologischen Prüfergebnis zugrunde liegt. Eine weitere Darstellung dieser Funktionalität befindet sich im folgenden Abschnitt.

Der Zugriff auf die in der Trefferliste angeführten Informationsquellen erfolgt per Mausklick auf die angegebene URL. Aufgrund des mit der Einbindung der XML-Fragmente in die Dokumente integrierten Anchor-Tags (vgl. Abschnitt 6.5.1) können Informationsressourcen (z.B. Diagramme, Tabellen) direkt angesprungen werden. Eine lokale Suche nach der Informationsquelle innerhalb der Dokumente kann somit entfallen (Abb. 6.18).



Abb. 6.15 Rechercheoberfläche für die Stichwortsuche in met@BiM-Informationsressourcen



Abb. 6.16 Rechercheoberfläche mit spezifischer Abfragemaske für Ressourcentyp betontechnologische Daten

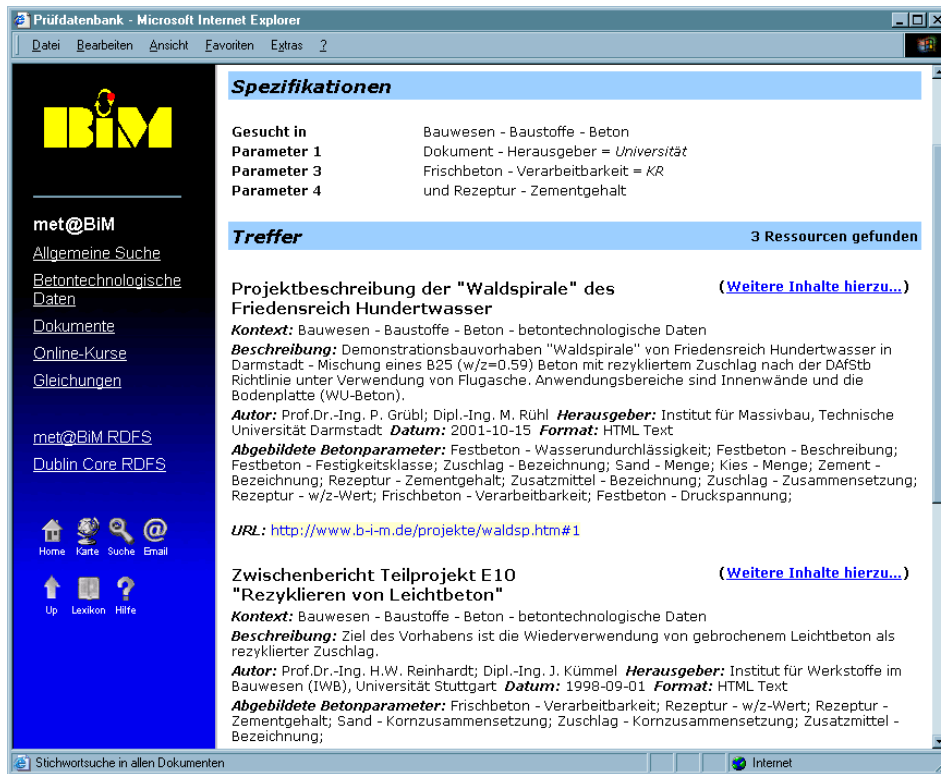


Abb. 6.17 Trefferliste einer Abfrage innerhalb des Testsystems

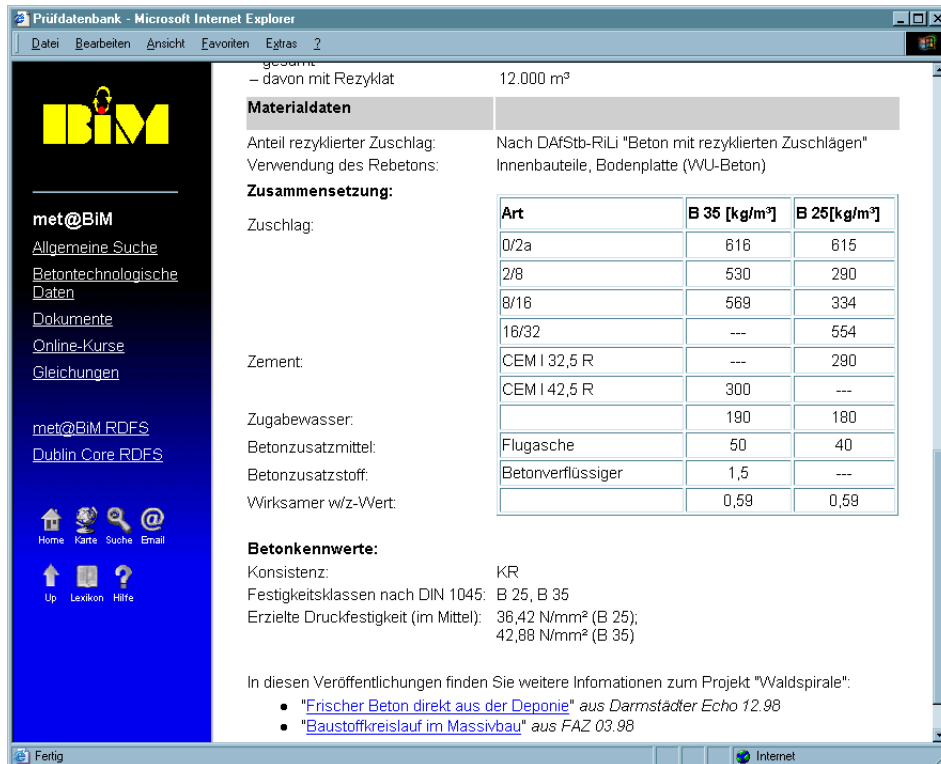


Abb. 6.18 Die bei den Treffern abgebildeten Hyperlinks führen direkt zur Informationsquelle

Bezüglich der technologischen Realisierung von XML-Abfragen wird die Tamino-Komponente „Query-Interpreter“ genutzt. Die XML-Abfragesprache, die hierbei verwendet wird, besitzt den Namen X-Query. Bei X-Query handelt es sich um eine Abfragesprache, die auf der XPath-Spezifikation (vgl. Abschnitt 2.6.2) des W3C basiert. Abfragen können als sogenannte http-Abfragen direkt in die Adressenleiste des Webbrowsers eingegeben werden. Anhand eines Beispiels soll der Aufbau einer solchen X-Query-Abfrage gezeigt werden: Die Abfrage soll auf die Datenbank *bim* auf dem Tamino Server mit der IP-Adresse *129.69.59.111* ausgeführt werden, welche die Collection *metabim* und darin XML-Objekte des Dokumenttyps *rdf:RDF* enthält. Es sollen dort nun alle XML-Objekte herausgelesen werden, die einen Knoten *metabim:verarbeitbarkeit* besitzen, der wiederum einen untergeordneten Knoten *rdf:value* mit dem Wert *KR* enthält. Für eine entsprechende Abfrage wird folgende URL benötigt:

```
http://129.69.59.111/tamino/bim/metabim/rdf:RDF?_XQL=/rdf:RDF/
metabim:verarbeitbarkeit[rdf:value='KR']
```

Als Serverantwort sendet Tamino die entsprechenden XML-Objekte als XML-Dokument an den Webbrowser zurück. Dort werden die XML-Daten dann mit Hilfe einer XSL Datei formatiert und dargestellt. Dieser Dialog wird in Abbildung 6.19 nochmals schematisch skizziert.

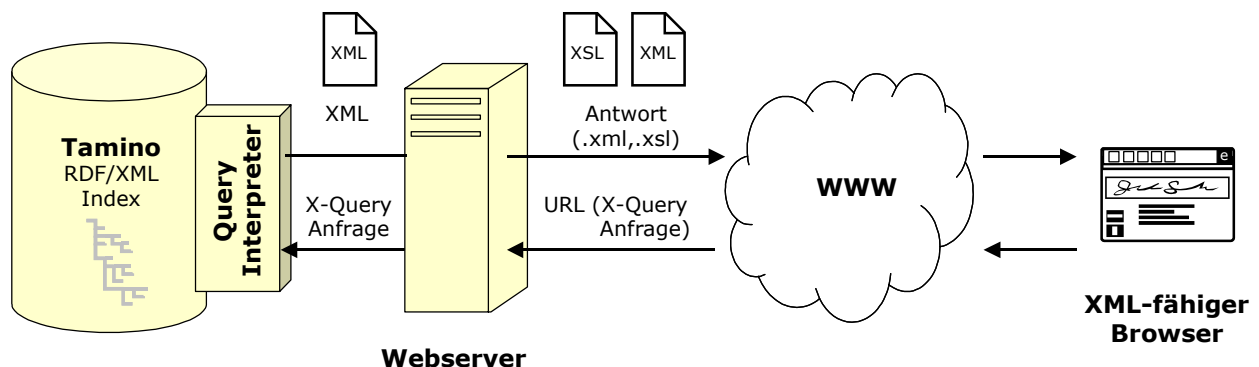


Abb. 6.19 Abfragedialog zwischen dem Webbrowser und dem Tamino Server im Testsystem

6.6 Abfrageszenarien für met@BiM-Datensätze

Die Vorteile des in den vorangegangenen Abschnitten beschriebenen Datenmodells lassen sich am anschaulichsten anhand von exemplarischen Abfragen darstellen. Die Anwendungsmöglichkeiten, die in den folgenden Beispielen für das Testsystem gezeigt werden, decken jedoch die Potentiale metadatenbasierender Dienste keinesfalls ab. Maschinenlesbare, semantische Datenstrukturen im World Wide Web bilden die Grundlage einer Vielzahl von automatisierten Funktionalitäten (vgl. Abschnitt 1.2), deren Anwendungsmöglichkeiten sich anhand der wenigen, bislang im Web realisierten, Systeme nur in Ansätzen abschätzen lassen.

Die in dieser Arbeit vorgestellten Anwendungsbeispiele beschränken sich zunächst darauf, einige der im einführenden Kapitel dargestellten wesentlichen Einschränkungen der

Recherche nach baustoffbezogenen Webinhalten durch die Nutzung des met@BiM-Datenmodells zu lösen. Diese sind im einzelnen:

- Semantischer Kontext für Webressourcen
- Exaktere, einheitenbehaftete Abfragen unter Verwendung relativer Operatoren
- Komplexe Abfragen mit Kombination mehrerer Parameter
- Nutzung der expliziten und impliziten semantischen Verknüpfungen in den Metadatenbeschreibungen der Webressourcen

Anhand der in den Beispielen genutzten Funktionalitäten des Testsystems sollen diese gegenüber aktuellen Volltextsuchmaschinen erweiterten Möglichkeiten anhand ausgewählter Inhalte von BiM-Online veranschaulicht werden.

Unterschiedliche Ressourcentypen und Kontextangaben

Für Recherchen mit relativ unkonkreten Zielvorstellungen hat sich in Informationssystemen die stichwortbasierende Recherche etabliert. Auch in Metadatenmodellen macht diese Funktionalität Sinn, wenn man die Recherche nicht auf bestimmte Eigenschaften oder Informationsressourcen beschränken möchte.

Ein Vorschlag für den Aufbau eines Stichwort-Suchformulars für das Testsystem ist in Abb. 6.20 dargestellt. Zusätzlich zur Angabe des Stichworts lässt sich hier noch der Suchbereich durch Anklicken von Kontrollkästchen auf bestimmte Ressourcentypen einschränken.

Suche bezüglich Webinhalten zu Beton

Bauwesen - Baustoffe - Beton

Stichwortsuche

Für eine Eingrenzung der Stichwortsuche markieren Sie bitte:

☐ Suche in Betonprüfdaten und Betonrezepturen
☐ Suche in Dokumente
☐ Suche in Kurse
☐ Suche in Gleichungen

Abb. 6.20 Suchformular der allgemeinen Stichwortsuche im Testsystem

Findet das System in einer Recherche zum Stichwort „Frostwiderstand“ mehrere Treffer in unterschiedlichen Ressourcentypen der met@BiM-Ontologie, so kann es anhand des RDF-Schemas aus dem Eigenschaftsnamen sowie der zugehörigen Ressourcenbezeichnung den jeweiligen Kontext ermitteln und dem Anwender anzeigen. Die Darstellung des Kontexts erfolgt als Kette von sich spezialisierenden Knoten des Ontologiebaumes. Beispiele für solche Ketten von Ressourcentypen befinden sich nach der fettgedruckten *Kontext*-Angabe innerhalb der Trefferangaben in Abb. 6.21. Ergänzend werden im Rahmen des Testsystems den aus dem met@BiM-RDFS ermittelten Kontextangaben die feststehenden Knoten Bauwesen - Baustoffe - Beton, aus Gründen der Anschaulichkeit, vorangestellt. Bei genauerer Betrachtung der Trefferanzeige erkennt man, dass es sich bei den abgebildeten Treffern nacheinander um ein Betonprüfergebnis (Ressourcentyp *metabim:BETON*), einen Online-Kurs (Ressourcentyp *metabim:KURS*) sowie ein Dokument handelt. Sollten Dopplungen auftreten, wie z.B. die Anzeige eines Dokuments als Treffer sowie aller in demselben Dokument enthaltenen entsprechenden Betonprüfergebnisse, lassen sich diese durch Algorithmen in der Trefferausgabe ausfiltern.



Abb. 6.21 Trefferausgabe der Stichwortabfrage. Die unterschiedlichen Typen von Informationsressourcen werden in der *Kontext*-Zeile angezeigt.

Nutzung von Grenzwerten zur Beschränkung der Treffermenge

Im technischen Bereich ist die Verwendung von Vergleichsoperatoren zur Detaillierung von Abfragen mit numerischen Werten eine häufig vermisste Funktionalität von Webrecherchen. Um solche Datenabfragen formulieren zu können, werden folgende Elemente benötigt:

- Beschreibung der Eigenschaft
- Der algebraische Vergleichsoperator
- Der Wert
- Evtl. Angaben zur Maßeinheit und Dimension des Werts

Die standardisierte Beschreibung der gesuchten Eigenschaften anhand von Ressource-Eigenschaft-Paaren sowie der Maßeinheit wurde bereits im Abschnitt 6.2 erläutert. Die Bezeichnungen, die für Eigenschaften und Ressourcen verwendet werden können, werden in den Metadatenschemata, wie z.B. dem met@BiM-RDF-Schema, definiert. Bezüglich einer nutzerfreundlicheren, natürlichsprachlichen Eingabe von Begriffen ist eine Zuordnung von Benutzerangaben zu den entsprechenden RDF-Schemaelementen mit Hilfe von sogenannten Thesauri in Zukunft jedoch denkbar.

Suche bezüglich betontechnologischer Daten Hilfe

Bauwesen - Baustoffe - Beton - betontechnologische Daten

Allgemeine Parameter		Wert	
(1)	-		
(2)	und -		
Betonparameter		Operator	Wert/ Anteil
(3)	Zuschlag	>	Ziegel
	Zusammensetzung		30
(4)	und	=	
	-		
	-		

starten

Abb. 6.22 Suchformular für betontechnologische Daten

Im Rahmen des Testsystems wurde das Problem einer konsistenten Bezeichnung der Betoneigenschaften durch die Benutzung von Auswahlfeldern in den Suchmasken gelöst. Zur Beschreibung eines Betonparameters stehen dem Anwender jeweils zwei übereinander angeordnete Auswahlfelder zur Verfügung, die mit für das spezifische Wissensgebiet (met@BiM) sinnvollen Bezeichnungen vorbelegt sind (vgl. Abb. 6.22): Im unteren der beiden Felder wird der Eigenschaftsname ausgewählt, während im oberen Feld der Ressourcentyp angegeben werden kann, auf den sich die Eigenschaft beziehen soll. Darüber hinaus wäre auch das Weglassen der Ressourcenzuordnung denkbar, so dass im abgebildeten Beispiel für die Betoneigenschaft *metabim:zusammensetzung* auch Treffer mit den Ressourcentypen *metabim:SAND*, *metabim:KIES* sowie *metabim:Splitt* gefunden werden können. In diesem Fall würde der im abgebildeten Beispiel wenig hilfreichen Kontextangabe eine sinnvolle Bedeutung zur Relevanzbestimmung des Treffers zukommen.

Die in Abb. 6.22 dargestellte Abfrage soll alle die Prüfergebnisse zur Anzeige bringen, bei denen der Zuschlag einen Ziegelanteil von über 30% besitzt. Die Treffer, die für diese Abfrage in der Datenbasis gefunden wurden, sind in Abb. 6.23 abgebildet.

Komplexere Abfrage mehrerer Baustoffeigenschaften

Analog zu SQL-Abfragen auf relationale Datenbasen (vgl. Abschnitt 2.5) lassen sich auch mehrere Abfrageparameter miteinander verknüpfen und auf XML-Datenstrukturen ausführen. Besonders bei der Suche nach Webinhalten, in welchen die Abhängigkeiten mehrerer Betonparameter voneinander untersucht werden, stellt die Kombinationsmöglichkeit von Parametern eine sinnvolle Funktionalität dar. Im Unterschied zu SQL-Abfragen bestehen die Parameter im met@BiM-Testsystem jedoch aus Ressource-Eigenschaft-Paaren und nicht lediglich aus dem Namen des Tabellenfeldes. In dem Beispiel der Abb. 6.24 werden die beiden Parameter *Festbeton - Druckspannung* sowie *Festbeton - Lastverformungen* miteinander kombiniert, um z.B. eine Spannungs-Dehnungslinie für einen Beton zu finden. Als weitere Einschränkung werden nur solche Prüfergebnisse gesucht, bei denen die Druckspannung den Wert 29 N/mm² übersteigt.

Tatsächlich wird in der Datenbasis des Testsystems ein entsprechender Webinhalt gefunden. Wie der Anwender den Trefferangaben entnehmen kann, handelt es sich um einen Vergleich der Spannungs-Dehnungslinien der Zylinderdruckfestigkeit eines Referenzbetons und eines Betons mit rezyklierter Gesteinskörnung aus Altbetonsplitt.



Abb. 6.23 Ergebnis einer relativen Abfrage nach Prüfergebnissen, bei denen der Anteil an Ziegelsplitt innerhalb des Zuschlags mit mehr als 30% angegeben wurde.

Suche bezüglich betontechnologischer Daten [Hilfe](#)

Bauwesen - Baustoffe - Beton - betontechnologische Daten

Allgemeine Parameter		Wert	
(1)	-		
(2)	und -		
Betonparameter		Operator	Wert/Anteil
(3)	Festbeton	>	29
	Druckspannung		N/mm²
(4)	und	=	
	Festbeton		
	Lastverformungen		-

[starten](#)

Abb. 6.24 Die abgebildete Abfrage kombiniert die Betonparameter Festbetondruckspannung sowie Lastverformungen miteinander.



Abb. 6.25 Treffer zur Abfrage in Abb. 6.24

Semantische Verknüpfungen

Eine bemerkenswerte Funktionalität des met@BiM-Datenmodells stellt die Möglichkeit dar, dass Programme aus den Metadatenstrukturen explizit angegebene, aber auch implizit in den Parameterwerten verborgene Zusammenhänge mit inhaltlich verknüpften Webinhalten automatisch ableiten können.

Zu den expliziten Informationen gehören in diesem Zusammenhang die Angaben zu weiteren Betonparametern, die in der Informationsquelle abgebildet werden (z.B. bei Betonprüfdaten), oder auch zum übergeordneten System, in welches das Webdokument oder der Online-Kurs integriert ist. Diese Angaben können von Suchmaschinen, wie der im Rahmen des Testsystems beschriebenen, direkt aus den Metadatenelementen *dc:source* bzw. *dc:relation* der RDF-Beschreibungen herausgelesen werden.

Darüber hinaus lassen sich jedoch noch weitere Angaben aus den in der RDF-Beschreibung vorhandenen Metadaten ableiten. So kann z.B. auf weitere Prüfergebnisse einer Versuchsreihe über die URL der Betonrezeptur im Element *dc:relation* geschlossen werden, die allen RDF-Beschreibungen der Versuchsreihe gemeinsam ist. Dieses „intelligente“, im Sinne von typischerweise eher von einem Menschen als von einem Automaten erwartete Verhalten, wird dabei in der Suchmaschine auf der Grundlage des RDF-Datenmodells erzeugt. Der Anwender bekommt diese semantische Funktionalität im Testsystem über ein separates Browserfenster angezeigt, für dessen Inhalte erneut eine Abfrage auf den RDF/XML-Indexbaum in der Tamino Datenbank durchgeführt wird. Das in Abb. 6.26 dargestellte Fenster wird über den Hyperlink „*Weitere Inhalte hierzu...*“ innerhalb der Trefferlisten aufgerufen. Die Inhalte des Fensters sind dabei abhängig vom Typ der Informationsquelle: Handelt es sich bei dem Treffer um Prüfergebnisse, so werden, wie in Abb. 6.26, die zugehörige Betonrezeptur sowie weitere Prüfdaten aus der Versuchsreihe angezeigt. Bei Online-Kursen oder Webdokumenten lassen sich in einem solchen Fenster bspw. alle im Dokument enthaltenen betontechnologischen Daten oder weitere Dokumente einer Berichtsreihe angeben.

Das abgebildete Beispiel stellt erneut lediglich eine von vielen denkbaren Anwendungen dar, bei denen aus den detaillierten Beschreibungen der Webinhalte automatisch weitere Aussagen abgeleitet werden. Die in den nächsten Jahren technologisch erreichbare Zukunftsvision in

diesem Bereich beschreibt das in Abschnitt 1.2 skizzierte Semantic Web, in welchem spezialisierte Softwareprogramme, sogenannte Agentenprogramme, die Inhalte des WWW selbstständig nach relevanten Inhalten bezüglich ihres Rechercheauftrags untersuchen. Fortgeschrittene Agenten können dabei z.B. selbstständig „erkennen“, dass der Wert „30 N/mm²“ dem Wert „3 kN/cm²“ entspricht, oder auch über fachspezifische Thesauri die Äquivalenz der *Betondruckfestigkeit* mit der Eigenschaft f_c nach ENV 206 erkennen.

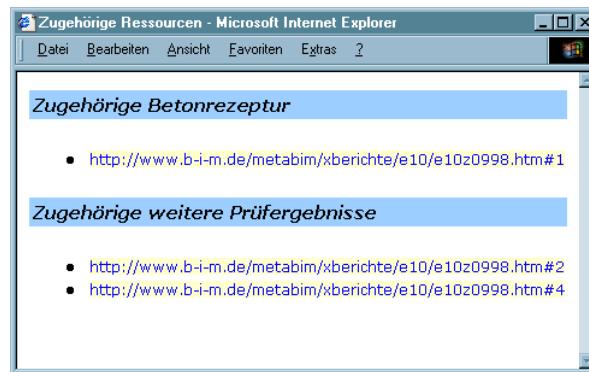


Abb. 6.26 Separates Browserfenster mit Angaben zu weiteren, inhaltlich mit dem jeweiligen Treffer verknüpften Webinhalten

7 Schlussbetrachtung

Der aktuelle Entwicklungsstand des in dieser Arbeit angewandten Wissensrepräsentationskonzepts für Webinhalte kann mit dem Stand von Hypertextsystemen verglichen werden, bevor das World Wide Web seine Bedeutung als globales Informationsmedium erlangte: Wie erste Anwendungsbeispiele zeigen, handelt es sich bei der Idee des Semantic Web offensichtlich um einen zukunftsweisenden Weg. Damit sich jedoch das volle Potential dieser Idee entfalten kann, müssen die Einzelanwendungen zu einem allgemeinen, globalen System verbunden werden. Die Betrachtung der Fragestellungen, auf welche Art und Weise dies geschehen könnte und welche Rolle hierbei fachspezifische Metadatenmodelle wie met@BiM einnehmen, bildet den Inhalt dieses abschließenden Kapitels.

7.1 Einbindung von met@BiM in ein Semantic Web

Trotz des RDF als gemeinsamem Formalismus zur Beschreibung von Webinhalten, wirft die dezentrale Entwicklung von Metadatenmodellen in den einzelnen Fachdisziplinen Fragen auf:

- Wie kann das erneute Entstehen von Insellösungen verhindert werden?
- Wie ließe sich ein baustoffspezifisches Webportal, welches analog zu dem in dieser Arbeit vorgestellten met@BiM-Testsystem aufgebaut ist, in eine allgemeine Webrecherche integrieren?
- Welche Folgen ergeben sich aus inhaltlich konkurrierenden bzw. sich überschneidenden Metadatenmodelle?
- Welche organisatorischen Voraussetzungen müssen für die Entwicklung und Pflege fachspezifischer Metadatenmodelle geschaffen werden?

Die XML-Codierbarkeit von RDF-Beschreibungen erlaubt im Gegensatz zu in Datenbanken gespeicherten Metadaten die Volltexterfassung der dort beschriebenen Inhalte. Um zusätzlich die Bedeutung der Metadatenelemente zu erfassen und über entsprechende semantische Funktionalitäten maschinell nutzen zu können, muss diese entweder fest in entsprechende Programme implementiert sein oder über Beziehungen zu standardisierten Metadatenelementen logisch abgeleitet werden können. Eine zukünftige, semantische Suchmaschine könnte somit beide Erschließungstechniken kombinieren [Berners-Lee1998]: In einem ersten Schritt würde aus einer konventionellen Volltextsuche heraus versucht werden, eine möglichst vollständige Liste von Webinhalten zu erzeugen. Im zweiten Schritt wird diese Liste dann mit Hilfe von Angaben zum Kontext hinsichtlich der Problemstellung semantisch ausgefiltert. Die Recherche innerhalb eines Semantic Web erfolgt hier über mehrere, sich stufenweise spezialisierende Recherchedienste. Allgemeine Suchmaschinen spezifizieren in einem Dialog die Benutzerrecherche und verweisen dann auf spezialisiertere Suchdienste für das entsprechende Wissensgebiet. Wie solch eine stufenweise aufgebaute Recherche mit Hilfe der Kontextangaben in einer semantischen Suchmaschine aussehen könnte, wird in Abb. 7.1 skizziert.

Durch die dezentrale Entwicklung der Metadatenmodelle und die thematischen Überschneidungen vieler Wissensgebiete wird die Entstehung konkurrierender Metadatendefinitionen in einem Semantic Web unvermeidbar sein. Die Nutzbarkeit der Datenmodelle muss dies allerdings nicht beeinträchtigen. Gleichbedeutende Metadatenelemente in unterschiedlichen RDF-Schemata lassen sich z.B. mit dem RDF-Schemaelement *rdfs:seeAlso* als Synonyme modellieren. Auch bezüglich der Weiterentwicklung von Metadatenmodellen lassen sich solche Konstrukte dazu nutzen, die Überarbeitung sämtlicher schon bestehender Datenstrukturen zu vermeiden. Suchdienste, die auf derartige Verknüpfungen treffen, können automatisch ihre Recherche auf das entsprechende Element ausweiten. RDF ergänzt hierbei

die Möglichkeit von XML, eigene Tags zur Auszeichnung von Webinhalten zu entwerfen, um Funktionen zur Einbindung der Tags in einen logischen, maschinenlesbaren Zusammenhang.

Auf lange Sicht gesehen ist die Parallelentwicklungen von Metadatenmodellen zumindest innerhalb eines Fachbereichs unwirtschaftlich und wenig sinnvoll. Es ist zu erwarten, dass einzelne Metadatenschemata zur Beschreibung häufiger genutzt werden als andere, so dass sich zumindest häufig beschriebene Webinhalte nur noch wenige Metadatenschemata etablieren werden. Um unkoordinierte Parallelentwicklungen zu vermeiden und gleich von Anfang an ein möglichst akzeptiertes Metadatenmodell für den Bereich des Bauwesens zu entwickeln ist es daher sinnvoll, diese Bemühungen innerhalb eines internationalen Gremiums zu konzentrieren. Zu den Aufgaben eines solchen Gremiums würde z.B. gehören:

- Entwurf einer Basisontologie für das Bauwesen
- Publikation von Metadatenstandards für die Ontologie bzw. die Ontologiebereiche
- Koordination der Vorschläge für Erweiterungen der Ontologie sowie zu Metadatenelementen aus den einzelnen Fachdisziplinen
- Bereitstellung von spezifischen Metadateneditoren zur Auszeichnung der Webinhalte

Für eine hohe Akzeptanz der Metadatenstrukturen bei den Webautoren sollte vor allem auf deren leichte Anwendbarkeit geachtet werden. Dabei ist es wichtig, dass z.B. die gewählten Bezeichnungen der Elemente sowie der semantischen Verknüpfungen den im Fachbereich üblichen Gepflogenheiten entsprechen. So sollte die Ontologie in ihrer Struktur das begriffliche Ordnungsschema des Fachgebiets widerspiegeln und die Metadatenelemente eine den Nutzern vertraute Definition erhalten. Weiterhin muss für Webautoren der Aufwand zur Erstellung der Datensätze durch leistungsfähige Metadateneditoren auf ein Minimum reduziert werden. Dies ist dann möglich, wenn Editoren spezifisch auf die zu beschreibenden Inhalte abgestimmte Benutzeroberflächen besitzen, in denen der Anwender sehr gezielt bezüglich der anzugebenden Metadaten abgefragt wird. In einem globalen Medium wie dem WWW sollte auch die internationale Anwendbarkeit der Metadatenmodelle bereits im Entwurf berücksichtigt werden. Das RDF-Konzept unterstützt den Prozess der Standardisierung in mehreren Sprachen insofern, als das eine Internationalisierung von Metadatenstrukturen durch Anwendung von Konstrukten wie *rdfs:seeAlso* auf einzelne Elemente auch nachträglich sehr einfach erfolgen kann.

Innerhalb der Vision eines globalen Semantic Web stellt das in dieser Arbeit vorgestellte met@BiM-Metadatenmodell einen Vorschlag zur semantischen Beschreibung von Webinhalten für den Wissensbereich Betontechnologie mit Berücksichtigung der besonderen Eigenschaften von Beton mit rezyklierter Gesteinskörnung dar. Zwischen den teilweise sehr spezialisierten Elementen des met@BiM-RDF-Schemas und den sehr allgemeinen Elementen von Dublin Core klafft jedoch noch eine große Lücke: Es fehlt an einem übergeordneten Metadatenschema für das Bauwesen.

Es ist zu hoffen, dass diese Lücke im Zuge des Wachstums eines Semantic Web geschlossen wird und dass das hier vorgestellte Metadatenmodell einen Beitrag liefert, die Nutzbarkeit des World Wide Web als Informationsquelle für bauspezifische Inhalte zu verbessern.

a

7.2 Zusammenfassung

Das World Wide Web kann von der kommunikations- und wissensintensiven Baubranche erst dann professionell zur Informationsrecherche genutzt werden, wenn sich dessen Inhalte „ingenieurgerechter“ erschließen lassen. Für die im Web publizierten Inhalte mit Bezug zum Bauwesen bedeutet dies, dass auf diese Informationsressourcen präzisere Recherchen ermöglicht werden müssen. D.h., es müssen

- detailliertere Abfragen auch von multimedialen Ressourcen, die einer Volltextrecherche nur schwer zugänglich sind
- sowie wertbehaftete Abfragen auf bestimmte Eigenschaften von Webinhalten.

ermöglicht werden. Die im WWW übliche Erschließung über Volltextrecherchen und Katalogsysteme ist für derartige Fragestellungen zu unscharf. Es ist daher ein neuer Ansatz zur Beschreibung von Informationsquellen notwendig, der über eine stärkere Strukturierung der Webinhalte auch die Publikationsform mit einschließt und diese um sinntragende Elemente ergänzt.

Die in dieser Arbeit verwendete Methodik zur Formalisierung betontechnologischer Aussagen von Webinhalten basiert auf der Idee des Semantic Web, die federführend vom „Erfinder“ des WWW, Tim Berners-Lee, vorangetrieben wird. In diesem WWW der nächsten Generation werden die eigentlichen Webinhalte um eine maschinenlesbare Beschreibung der inhaltlichen Aussagen der Informationsressourcen erweitert. Die Beschreibung des semantischen Kontexts erfolgt anhand von Metadaten, die beliebigen Webinhalten zugeordnet werden können. Ein allgemeiner Formalismus für den Entwurf von Metadatenmodellen für Webinhalte wurde unter der Bezeichnung Resource Description Framework (RDF) durch das World Wide Web Consortium (W3C) 1999 verabschiedet. Bezüglich der Anwendung der in dieser Spezifikation vorgeschlagenen Methode zur Entwicklung von Datenmodellen in den einzelnen Wissensgebieten sind die jeweiligen Fachbereiche gefordert.

In dieser Arbeit wurde mit met@BiM ein Metadatenmodell für das Gebiet der Baustoff-Informationen, genauer für betontechnologische Daten zu Beton mit rezyklierten Zuschlägen, entwickelt. Die Definition der Metadatenelemente erfolgte zunächst syntaxunabhängig. In einem weiteren Schritt wurden sie dann in einer RDF/XML-Codierung als RDF-Schema serialisiert, um sie in maschinenlesbarer Form auf Informationsressourcen im WWW anwenden zu können. Der allgemein anerkannte Standard der Dublin Core Metadata Initiative bildet den Ausgangspunkt für die Definition der Metadatenstrukturen. Diese wurden um spezifische Metadatenelemente ergänzt, die sich auf die Ressourcen bzw. Objektklassen einer Ontologie für betontechnologische Webinhalte beziehen. Während DC-Metadatenelemente den Schwerpunkt auf die allgemeinen, inhaltsunabhängigen Merkmale von Informationsquellen legen, ermöglichen die met@BiM-Elemente eine detaillierte Beschreibung der betontechnologischen Aussagen in den Informationsquellen.

Am Beispiel der Inhalte des webbasierenden Informationssystems BiM-Online mit den Forschungsergebnissen des Projekts „Baustoffkreislauf im Massivbau (BiM)“ wird die Anwendung des Metadatenmodells demonstriert. Es wird gezeigt, dass sich mit den vorgeschlagenen met@BiM-Elementen bislang fehlende Recherchemöglichkeiten im WWW für betontechnologische Inhalte realisieren lassen. Insbesondere wird die Realisierbarkeit von ingenieurtypischen Abfragen mit Parameterkombinationen unter Verwendung Boolescher Operatoren, der Angabe von Grenzwerten sowie weiteren numerischen Vergleichoperatoren auf die Datenstrukturen dargestellt.

Die Anwendung des in dieser Arbeit vorgeschlagenen Metadatenmodells kann zudem stufenweise erfolgen: So können auch lediglich die in den Informationsquellen dargestellten

betontechnologischen Eigenschaften ohne weitere Angaben zu Eigenschaftswerten beschrieben werden. Auf diese Weise lassen sich rasch die wichtigsten fachspezifischen Parameter von Webinhalten innerhalb weniger Metadatenelemente erfassen. Auch ausführlichere Beschreibungen mit Angaben zu Werten der Materialeigenschaften können durch Zusammenfassen von Werten zu Wertebereichen im Aufwand reduziert werden. Es liegt dabei im Ermessen des Anwenders, wie detailliert und damit aufwendig er die Beschreibung der publizierten Webinhalte anhand der met@BiM-Elemente gestalten möchte.

Da diese Arbeit die erste Untersuchung auf dem Gebiet der webbasierten Metadatenmodelle für die Baubranche darstellt, werden noch weitere Arbeiten benötigt, um das Modell in einigen Bereichen zu verallgemeinern oder es durch weitere Elemente zu ergänzen. Insbesondere gilt dies für die in met@BiM verwendeten Metadatenelemente zur Modellierung der Maßeinheiten. Weitere Metadatenelemente zu Beton werden in Arbeiten entstehen, die sich mit dem Werkstoff aus anderer Perspektive beschäftigen, z.B. mit Schwerpunkt auf den wirtschaftlichen Aspekten der Herstellung, Verarbeitung sowie der Nutzungsphase und der Entsorgung.

Um die praxisnahe Anwendung und den Nutzen der in dieser Arbeit vorgeschlagenen Metadatenelemente aus Dublin Core und met@BiM zu veranschaulichen wurde eine Softwareumgebung entworfen, in der die Phasen der Datenerfassung der verteilt gespeicherten Metadaten, der Datenspeicherung der RDF/XML-Datensätze sowie der Gestaltung von Abfrageoberflächen dargestellt werden. Da RDF-Beschreibungen rasch eine hohe Komplexität erreichen, muss der Anwender schon bei der Erzeugung der Metadatenätze von entsprechenden Editoren unterstützt werden. Zum heutigen Zeitpunkt stehen derartige Programme noch nicht in der benötigten allgemeinen Form zur Verfügung. Gleiches gilt für die Speicherung und Abfrage von RDF-Datenstrukturen. Zur Implementierung von met@BiM in ein Testsystem mussten daher zunächst Systemkomponenten neu entwickelt oder, wie im Falle der XML-Datenbank Tamino, kommerzielle Softwareprodukte entsprechend angepasst werden. Die hieraus entstandene Softwareumgebung umfasst folgende Komponenten:

- RDF-Editor zur Beschreibung von betontechnologischen Prüfdaten, von Rezepturbeschreibungen, Dokumenten, Online Kursen, mathematischen und chemischen Gleichungen
- Datenerfassungsschnittstelle zur automatischen Online-Erfassung bzw. Extraktion der in den Webressourcen enthaltenen RDF-Beschreibungen
- XML-Datenbank Tamino zur Speicherung des RDF/XML-Indexes
- Formularbasierende Abfrageoberfläche für inhaltsspezifische Recherchen auf dem Index

Ziel war es, die Umsetzung und die Vorteile dieser metadatenbasierten Methode gegenüber einer konventionellen Volltextsuche zu demonstrieren. Anhand der Abfragen, die sich innerhalb des Testsystems auf die Datenbasis ausführen lassen, konnte gezeigt werden, dass einige wesentliche Einschränkungen der Recherche nach baustoffbezogenen Webinhalten durch die Nutzung des met@BiM-Datenmodells gelöst werden können:

- Der semantischer Kontext von Webressourcen wird automatisch erkannt
- Exaktere, einheitenbehaftete Abfragen unter Verwendung relativer Operatoren sind möglich
- Komplexere Abfragen mit Kombinationen mehrerer Parameter sind möglich
- Die Nutzung der expliziten und impliziten semantischen Verknüpfungen in den Metadatenbeschreibungen, z.B. durch wissensbasierte Webagenten [Russell1997], ist realisierbar

Die im Rahmen des Testsystems entwickelte Softwareumgebung könnte um administrative Funktionen ergänzt in ein zukünftiges Webportal integriert werden, um bspw. als Suchmaschine für fachspezifische Recherchen nach Webinhalten zu betontechnologischen Themen genutzt zu werden. Eine Erweiterung eines solchen Portals auch auf Webinhalte zu anderen Werkstoffen ist durch die Übertragbarkeit der in dieser Arbeit vorgestellten Methoden zur formalisierten Beschreibung von Baustoffwissen gegeben.

Sobald es Softwareentwicklern gelingt, bezüglich der unterschiedlichen RDF-Schemata flexible und benutzerfreundlichere Werkzeuge zu entwickeln, ergeben sich interessante Perspektiven für den Entwurf komplexerer, wissensbasierter Anwendungen auf dem vorgeschlagenen met@BiM-Modell. Zunächst ist zwar die Ausbildung fachspezifischer, semantischer Wissensnetze zu erwarten, da die hierfür benötigten Softwarekomponenten lediglich mit einer begrenzten Zahl von RDF-Schemata umgehen müssen. Aufgrund des gemeinsamen Formalisierungskonzepts des RDF sowie des hohen Engagements der Softwareindustrie in diesem Bereich wird einem Zusammenwachsen der einzelnen Wissensnetze zu einem allgemeinen Semantic Web auf lange Sicht jedoch nichts mehr im Wege stehen.

8 Literaturverzeichnis

- [Abele1999] Abele, T.: *BiM-Online: Rezeptur und Eigenschaften von Betonen aus rezyklierten Zuschlägen (Re-Beton)*. Diplomarbeit Uni Stuttgart, IWB. 1999.
- [aecXML] aecXML Whitepaper: *A Framework for Electronic Communications for the AEC Industries*
- [aecXML2000] International Alliance For Interoperability: *aecXML*. URL: <http://www.iai-na.org/domains/aecxml.html>; 2000
- [AOL2001] America Online, Inc.: *AOL Search*. URL: <http://search.aol.com>, 2001
- [ARPAnet] Kristula, D.: *The History of the Internet*. URL: <http://www.davesite.com/webstation/net-history.shtml>; 2001
- [Baunetz] *Baunetz - Der Online Dienst für die Baubranche*. URL: <http://www.baunetz.de>
- [bcXML2002] eConstruct: *bcXML*. URL: <http://www.econstruct.org>; 2002
- [Beckett2001] Beckett, D.: *Raptor RDF-Parser Toolkit*. URL: <http://www.redland.opensource.ac.uk/raptor/>; 2001
- [Behme2000] Behme, H.; Mintert, S.: *XML in der Praxis*. 2.Auflage. München: Addison-Wesley Verlag; 2000, ISBN 3-8273-1636-7
- [Bergmann2000] Bergmann, M. K.: *White Paper - The Deep Web: Surfacing Hidden Value*. Bright Planet.com, URL: <http://www.completeplanet.com/Tutorials/DeepWeb/index.asp>, 2000
- [Berners-Lee1990] Berners-Lee, T.: *The WorldWideWeb browser*. URL: <http://www.w3.org/People/Berners-Lee/WorldWideWeb.html>; 1990
- [Berners-Lee1998] Berners-Lee, T.: *Semantic Web Road map*. URL: <http://www.w3.org/DesignIssues/Semantic.html>; 1998
- [Berners-Lee2001] Berners-Lee, T.; Hendler, J.; Lassila, O.: *The Semantic Web*. In: Scientific American, 2001, Issue 5
- [BiM-Online2001] Schreyer, M.; Schwarte, J. Reinhardt, H.-W.: *BiM-Online: Das Informationssystem zum Forschungsprojekt „Baustoffkreislauf im Massivbau“*. URL: <http://www.b-i-m.de>; 2001
- [Bingel1999] Bingel, O.: *BiM-Online: Rezyklierter Zuschlag aus aufbereitetem Betonbruch / Bauschutt*. Diplomarbeit Uni Stuttgart, IWB. 1999.
- [Braun1999] Braun, N.: *Technische Informationssysteme im Bauwesen - Die Betontechnologie digital und multimedial*. Diplomarbeit. TU Berlin. Fachgebiet Baustoffwesen und Bauökologie, 1999
- [Bray1998] Bray, T.: *RDF and Metadata*. URL: <http://www.xml.com/xml/pub/98/06/rdf.html>; 1998
- [Brin1998] Brin, S.; Page, L.: *The anatomy of a large-scale hypertextual Web search engine*. URL: <http://decweb.ethz.ch/WWW7/1921/com1921.htm>; 1998

- [Büchin1973] Büchin, K.; Dallinger, E.; Döge, B.; Sauer, H.: *BRD/SFB - Handbuch für Klassifikation und Codierung im Bauwesen*. In: Schriftenreihe des Instituts für Baukonstruktion der Universität Stuttgart, Heft 1, 1973
- [BuildOnline2000] Build Online: *Baufachleute in den US wickeln Geschäfte online ab*. URL: <http://www.build-online.com/de/news/inews.asp?CountryId=4&ArticleId=918>; 2000
- [Burkart1990] Burkart, M.: *Dokumentationssprachen* In: Buder, M.; Rehfeld, W.; Seeger, T. (Hrsg.): *Grundlagen der praktischen Information und Dokumentation: Ein Handbuch zur Einführung in die fachliche Informationsarbeit*. Bd. 1; 3. Ausgabe. München [u.a.]: Saur, 1990
- [Cailliau1995] Cailliau, R.: *A little History of the World Wide Web*. URL: <http://www.w3.org/History.html>; 1995
- [Cara] Kokkellink, S.: *Cara - Carmen RDF API*. URL: <http://zoe.mathematik.uni-osnabrueck.de/RDF/>
- [Chang1998] Chang, W. W.: *A Discussion of the Relationship Between RDF-Schema and UML*. URL: <http://www.w3.org/TR/1998/NOTE-rdf-uml-19980804>; 1998
- [Christians1999] Christians, M.; von Erdély, A.: *Der Stand des Wissensmanagements in der Bauindustrie*. Diplomarbeit. Universität Stuttgart. Institut für Baubetriebslehre, 1999
- [Clark1997] Clark, J.: *Comparison of SGML and XML*. World Wide Web Consortium Note 15-December-1997. URL: <http://www.w3.org/TR/NOTE-sgml-xml-971215>; 1997
- [CML2002] *Chemical Markup Language (CML)*. URL: <http://www.xml-cml.org>; 2002
- [Codd1970] Codd, E.F.: *A Relational Model of Data for Large Shared Data Banks*. In: Communications of the ACM Vol. 13, No. 6, 1970, pp. 377-387
- [Cox] Cox, B.: *IBM Global Services White Paper on Enterprise Information Portal Strategy*. URL: <http://edms.solutions.ibm.com/eip.pdf>
- [CSS1998] Bos, B.; Lie, H. W.; Lilley, C.; Jacobs, I.: *Cascading Style Sheets, level 2 CSS2 Specification. W3C Recommendation 12-May-1998*, URL: <http://www.w3.org/TR/REC-CSS2/>; 1998
- [CSS1999] Lie, H. W.; Bos, B.: *Cascading Style Sheets, level 1. W3C Recommendation 17 Dec 1996, revised 11 Jan 1999*. URL: <http://www.w3.org/TR/REC-CSS1>; 1999
- [CYC1997] CYCROP, Inc.: *The CYC Technology*. URL: <http://www.cyc.com>; 1997
- [DAfStb1999] Deutscher Ausschuss für Stahlbeton: *Richtlinie: Beton mit rezykliertem Zuschlag*. Berlin: Beuth Verlag, 1999
- [DeGrassi1999] De Grassi, M.; Giretti, A.; Zorgno, A.M.; Caneparo, L.: *The virtual Construction Enterprise*. URL: <http://www.ce.berkeley.edu/~tommelein/CEMworkshop.htm>; 1999
- [Dertouznous1997] Dertouznous, M.: *What Will Be*. Harper Edge, 1997
- [DIN2002] *Web-Site des Deutschen Instituts für Normung e.V.*. URL: <http://www.din.de>; 2002

- [DMOZa] Netscape: *dmoz open directory project*. URL: <http://dmoz.org>
- [DMOZb] Netscape: *Open Directory RDF Dump*. URL: <http://dmoz.org/rdf.htm>
- [DOM1] Wood, L.; et al.: *Document Object Model (DOM) Level 1 Specification Version 1.0. W3C Recommendation 1. October, 1998*. URL: <http://www.w3.org/TR/1998/REC-DOM-Level-1-19981001/>; 1998
- [Dublin] *Dublin Core Metadata Initiative*. URL: <http://purl.org/DC/>
- [Duden1997] Friedrich, C.: *Duden - Schriftliche Arbeiten im technisch naturwissenschaftlichen Studium: Ein Leitfaden zur effektiven Erstellung und zum Einsatz moderner Arbeitsmethoden*. Mannheim; Leipzig; Wien; Zürich: Dudenverlag, 1997, ISBN 3-411-06271-1
- [Florescu1998] Florescu, D.; Levy, A.; Mendelzon, A.: *Database Techniques for the World-Wide Web: A Survey*. In: ACM SIGMOD Record. 1998
- [FOGIB1997] FOGIB: *Ingenieurbauten - Wege zu einer ganzheitlichen Betrachtung. Band I: Motivation*. Abschlußbericht der DFG-Forschergruppe FOGIB an der Universität Stuttgart, Institut für Konstruktion und Entwurf II, 1997
- [Fühles-Ubach1997] Fühles-Ubach, S.: *Analysen zur Unschärfe in Datenbank- und Retrievalsystemen unter besonderer Berücksichtigung der Redundanz*. Dissertation. Humboldt-Universität Berlin, Philosophischen Fakultät I, 1997
- [Gilliland2000] Gilliland-Swetland, A. J.: *Introduction to Metadata - Setting the Stage*. URL: <http://www.getty.edu/gri/standard/intrometadata>; 2000
- [Gödel1962] Gödel, K.G.: *On Formally Undecidable Propositions*. Basic Books, New York, 1962
- [Goldfarb1999] Goldfarb, F.; Prescod, P.: *XML-Handbuch*. München; London; Mexiko; New York; Singapur; Sydney; Toronto: Prentice Hall; 1999, ISBN 3-8272-9575-0
- [Goldmann2001] Goldmann, M.: *Der menschliche Faktor*. URL: <http://news.zdnet.de/story/0,,t101-s2093142,00.html>; 2001
- [Google2002] Google. URL: <http://www.google.com>; 2002
- [Gruber1993] Gruber, T. R.: *A Translation Approach to Portable Ontology Specifications*. In: Knowledge Acquisition Vol. 5, p. 199-220, 1993
- [Grübl1997] Grübl, P.: *Baustoffkreislauf im Massivbau*. In: Bauingenieur 72, 1997, Heft 10, S. 425-432, URL: <http://www.b-i-m.de/public/ext/10.htm>
- [Grübl1998] Grübl, P.; Rühl, M.: *German Committee for Reinforced Concrete (DAfStb) - Code: Concrete with Recycled Aggregates*. Sustainable Construction: In: University of Dundee, Concrete Technology Unit, London (UK): Use of Recycled Concrete Aggregate, 11-12 November 1998, Dundee, UK, URL: <http://www.b-i-m.de/public/ext/25.htm>
- [Grübl1998] Grübl, P.; Rühl, M.: *Teilprojekt E03: Der Einfluss von Recyclingzuschlägen aus Bauschutt auf die Frisch- und Festbetoneigenschaften und die Bewertung hinsichtlich der Eignung für Baustellen- und Transportbeton nach DIN 1045*. Baustoffkreislauf im Massivbau. Zwischenbericht September 1998, TU Darmstadt, Institut für Massivbau, URL: <http://www.b-i-m.de/Berichte/E03/E03z0998.htm>; 1998

- [Guha] Guha, R. V.: *rd fDB: An RDF Database (v. 0.46)*. URL: <http://www.guha.com/rdfdb>
- [Hartmann2000] Hartmann, D. (Hrsg.) u.a.: *Objektorientierte Modellierung in Planung und Konstruktion*. DFG Forschungsschwerpunkt. 2000
- [Henning1999] Henning, D.; Schön, E.: *Untersuchung zu Notwendigkeit, Struktur und zum Funktionsumfang von Tools für den Umgang mit dem Resource Description Framework*. Studienarbeit. Technische Universität Ilmenau, Fakultät für Informatik und Automatisierung, Institut für Praktische Informatik und Medieninformatik, 1999
- [Heuer1997] Heuer, A.: *Objektorientierte Datenbanken - Konzepte, Modelle, Standards und Systeme*. 2. Ausgabe. Bonn : Addison-Wesley-Longman, 1997, ISBN 3-89319-800-8
- [Hofstadter1979] Hofstadter, D.R.: Gödel, Escher, Bach: *An eternal golden braid*. Basic Books, New York, 1979
- [HotBot2002] Lycos - HotBot. URL: <http://www.hotbot.lycos.de>; 2002
- [HTML401Rec] Raggett, D.: *HTML 4.01 Specification. W3C Recommendation 24 December 1999*. URL: <http://www.w3.org/TR/html4/>; 1999
- [Hunter1999] Hunter, J.; Mather, T.: *Veggie - An indexing, Browsing, Search and Retrieval System for Audiovisual Libraries*. URL: <http://archive.dstc.edu.au/RDU/staff/jane-hunter/INFOG99/paper.html>; 1999
- [IANA2002] iana (Internet Assigned Numbers Authority): *Text Media-Types*. URL: <http://www.iana.org/assignments/media-types/text/>; 2002
- [Ianella1998] Ianella, R.: *An Idiot's Guide to the Resource Description Framework*. URL: <http://archive.dstc.edu.au/RDU/reports/RDF-Idiot/>; 1998
- [IBMEDM2000] IBM EDM Solutions: *Portals*. URL: <http://edms.solutions.ibm.com/portal.htm>, 2000
- [Iken1994] Iken; Lackner; Zimmer: *Handbuch der Betonprüfung - Anleitungen und Beispiele*. Beton-Verlag, 1994
- [Intellidimension] Intellidimension: *RDF Gateway Beta 2 (v0.6)*. URL: <http://www.intellidimension.com/RDFGateway/beta2/>
- [ISC2001] Internet Software Consortium: *Internet Domain Survey, July 2001*. URL: <http://www.isc.org/ds/WWW-200107/index.html>; 2001
- [ISO639-2] ISO 639-2 Registration Authority: *Codes for the Representation of Names of Languages-Part 2: Alpha-3 Code*. URL: <http://lcweb.loc.gov/standards/iso639-2/langhome.html>
- [ISO8601] Wolf, M.; Wicksteed, C.: *Date and Time Formats*. URL: <http://www.w3.org/TR/NOTE-datetime>, 1997
- [Johnson1997] Johnson, S.: *Using Active Server Pages*. Indianapolis: Que Corp., 1997, ISBN 0-7897-1389-6
- [Jung2000] Jung, F.: *XML Backgrounder: Technology and Applications*. Software AG, URL: http://www.softwareag.com/tamino/download/e-XML_Backgrounder_0800.pdf, 2000

- [Koch1999] Koch, S.: *JavaScript: Einführung, Programmierung und Referenz*. 2.Auflage. Heidelberg:dpunkt-Verlag, 1999, ISBN 3-932588-26-6
- [Koptik2000] Koptik, J.: *BiM-Online - "Umweltgerechte Planung und Durchführung von Abbrucharbeiten"*. Diplomarbeit Uni Stuttgart, IWB. 2000.
- [Koptik2001] Koptik, J.; Schwarte, J.: *WIBA-NET: Multimediales Netzwerk zur Wissensvermittlung im Fach „Werkstoffe im Bauwesen“*. In: Mitteilungen des Instituts für Werkstoffe im Bauwesen (IWB) - Jahresbericht 2000/01 Activities, Band 2001/3 S.149-156, 2001, ISSN 0932-5921
- [Krause2000] Krause, J.: *PHP4: Grundlagen und Profiwissen*. München: Hanser Verlag, 2000, ISBN 3-446-21546-8
- [KrWG1994] *Gesetz zur Förderung der Kreislaufwirtschaft und Sicherung der umweltverträglichen Beseitigung von Abfällen*. KrW-/AbfG - Kreislaufwirtschafts- und Abfallgesetz vom 27. September 1994
- [Lafon2000] Lafon, Y.; Bos, B.: *Describing and retrieving photos using RDF and HTTP*. W3C Note, 28 September 2000. URL: <http://www.w3.org/TR/photo-rdf/>; 2000
- [Lagoze1996] Lagoze, C.; Lynch, C.; Daniel Jr., R.: *The Warwick Framework: A container architecture for aggregating sets of metadata*. In: Technical Report TR 96-1593. Computer Science Department, Cornell University, June, 1996
- [Lawrence1999] Lawrence, S.; Giles, C. L.: *Accessiblity of Information on the Web*. In: Nature Vol. 400, Issue July 8, 1999, p. 107-109.
- [Luik1999] Luik, M.: *BiM-Online: Rezeptur und Eigenschaften von Betonen aus rezyklierten Zuschlägen (Re-Beton)*. Diplomarbeit Universität Stuttgart, Institut für Werkstoffe im Bauwesen (IWB). 1999.
- [Lycos2002] Lycos URL: <http://www.lycos.com>; 2002
- [MathML2001] Carlisle, D.; et al.: *Mathematical Markup Language (MathML) Version 2.0*. URL: <http://www.w3.org/TR/MathML2>; 2001
- [Maurer1993] Maurer, F.: *Hypermediabasiertes knowledge Engineering für verteilte wissensbasierte Systeme*. Infix, 1993
- [Melnik1999] Melnik, S.: *ISLE: Konzeption und Realisierung eines verteilten XML-basierten Informationssystems*. Universität Leipzig, Fakultät für Mathematik und Informatik, Institut für Informatik, URL: <http://dol.uni-leipzig.de/pub/1999-4>; 1999
- [Miller1998] Miller, E.: *An Introduction to the Resource Description Framework*. In: D-Lib Magazine, 1998, Issue 5
- [MSXML2000] Microsoft Corporation: *What's New in the Microsoft XML-Parser Version 3.0 Release*. URL: <http://msdn.microsoft.com/library/default.asp?url=/nhp/Default.asp?contentid=28000438>; 2000
- [Müller1966] Müller, C.: *Alphabetisches Sachverzeichnis zur Schrifttumkartei Beton nach der Dezimalklassifikation*. Berlin; München: Ernst & Sohn, 1966
- [Müller1998a] Müller, C.; Wiens, U.: *Teilprojekt D03: Bewertung der bei der Aufbereitung von Bauschutt anfallenden Recyclingzuschläge hinsichtlich der Eignung als Betonzuschlag*. Baustoffkreislauf im

- Massivbau. Zwischenbericht September 1998, RWTH Aachen, ibac, URL: <http://www.b-i-m.de/Berichte/D03/D03z0998.htm>
- [Müller1998b] Müller, C.; Wiens, U.: *Teilprojekt E04: Mehrfachrecycling. Baustoffkreislauf im Massivbau*. Zwischenbericht September 1998, RWTH Aachen, ibac, URL: <http://www.b-i-m.de/Berichte/E04/E04z0998.htm>;
- [mybau2002] MYBAU. URL: <http://www.mybau.com>, 2002
- [Namespaces1999] *Namespaces in XML*. URL: <http://www.w3.org/TR/1999/REC-xml-names-19990114/>; 2001
- [netinform] TÜV Süddeutschland: *netinform Wissensmanagement - Das Industrieportal des TÜV Süddeutschland*. URL: <http://www.netinform.de>
- [Netscape2001] Netscape: *Search*. URL: <http://search.netscape.com>; 2001
- [OSI2001] The Open Source Initiative. URL: <http://www.opensource.org>; 2001
- [Paepcke1998] Paepcke, A.; Chang, C.-C. K.; García-Molina, H.; Winograd, T.: *Interoperability for Digital Libraries Worldwide*. In: Communications of the ACM. Vol. 41, Issue 4, 1998
- [RDF1999] Lassila, O.; Swick, R.: *Resource Description Framework (RDF) Model and Syntax Specification. W3C Recommendation 22 February 1999*. URL: <http://www.w3.org/RDF>
- [RDF2001] Beckett, D.: *Refactoring RDF/XML Syntax. W3C Working Draft 06 September 2001*. URL: <http://www.w3.org/TR/2001/WD-rdf-syntax-grammar-20010906/>
- [RDFQuery1998] Malhotra, A.; Sundaresan, N.: *RDF Query Specification. December 3, 1998*. URL: <http://www.w3.org/TandS/QL/QL98/pp/rdfquery.html>
- [RDFS2000] Brickley, D.; Guha, D. W.: *Resource Description Framework (RDF) Schema Specification 1.0. W3C Candidate Recommendation 27 March 2000*. URL: <http://www.w3.org/TR/2000/CR-rdf-schema-20000327>
- [RDFValid2001] W3C: *RDF Validation Service*. URL: <http://www.w3.org/RDF/Validator>; 2001
- [Reggiori2001] Reggiori, A.: *RDFStore SiRPAC parser on-line demonstration*. URL: <http://rdfstoredemo.jrc.it/>; 2001
- [Reinhardt1998] Reinhardt, H.-W.; Kümmel, J.: *Teilprojekt E10: Rezyklieren von Leichtbeton. Baustoffkreislauf im Massivbau*. Zwischenbericht September 1998, Universität Stuttgart, IWB, URL: <http://www.b-i-m.de/Berichte/E10/E10z0998.htm>
- [Reinhardt1999a] Reinhardt, H.-W.; Schwarte, J.; Schreyer M.: *Baustoffkreislauf im Massivbau: "BiM-Online - Ein neues Informationskonzept"*. In: DAfStb-Forschungskolloquium. Tagungsband zu DAfStb/DIN/BMBF - Fachtagung "Beton mit rezykliertem Zuschlag für Konstruktionen nach DIN 1045-1", Deutscher Ausschuss für Stahlbeton im DIN Deutsches Institut für Normung e.V., 27. und 28. Mai 1999, Berlin. 1999, S. 175-195.

- [Reinhardt1999b] Reinhardt, H.-W.; Schreyer, M.; Schwarte, J.: *BiM-Online - Ein Informationssystem zum Thema Baustoffrecycling*. In: ARCONIS - Wissen zum Planen und Bauen und zum Bauprodukt Vol. 4, 1999, H. 3, S. 18-21.
- [Reinhardt1999c] Reinhardt, H.-W.; Schreyer, M.; Schwarte, J.: *Baustoffkreislauf im Massivbau – Anwendung des Informationssystems BiM-Online in Fallbeispielen*. 2. Auflage. Universität Stuttgart, Institut für Werkstoffkunde, 1999
- [Reinhardt2000] Reinhardt, H.-W.; Schreyer, M.; Schwarte, J.: *BiM-Online - Das Interaktive Informationssystem zu "Baustoffkreislauf im Massivbau": Hervorgegangen aus den Teilprojekten "Strukturierte Speicherung der Forschungsergebnisse in einer Datenbank" und "Auswertung der Ergebnisse mit einem Expertensystem"*. Heft 500: Schriftenreihe des DAfStb. Berlin, Wien, Zürich: Beuth-Verlag, 2000.
- [Reinhardt2001] Reinhardt, H.-W.; Ernst, H.; Lünser, H.; Schwarte, J.; Koptik, J.; Schreyer, M.: *Internet für Bauherren, Architekten und Ingenieure*. 1. Aufl.. Stuttgart: Wirtschaftsministerium Baden-Württemberg, 2001.
- [RFC1766] Alvestrand, H.: *Tags for the Identification of Languages*. URL: <http://www.ietf.org/rfc/rfc1766.txt>; 1995
- [RFC2046] Freed, N.; Borenstein, N.: *Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types*. URL: <http://www.ietf.org/rfc/rfc2046.txt>; 1996
- [RFC2396] Berners-Lee T., et al.: *Uniform Resource Identifiers (URI): Generic Syntax*. URL: <http://www.ietf.org/rfc/rfc2396.txt>; 1998
- [Rinder2001] Rinder, T.: *Mechanische Kennwerte von hochfestem Beton*. In: Mitteilungen des Institut für Werkstoffe im Bauwesen. Bd. 2001/3: Jahresbericht 2000/2001 Activities. Stuttgart: 2001, S. 136-141. ISSN 0932-5921
- [RQL2001] Karvounarakis, G.; Christophides, V.: *The RDF Query Language (RQL)*. Institute of Computer Science (Greece). URL: <http://139.91.183.30:9090/RDF/RQL/>; 2001
- [Russell1997] Russell, S.; Norvig, P.: *Artificial Intelligence: A modern Approach*. Prentice Hall, 1995
- [Scherer1997] Scherer, R.J.; Hauser, M.: *Perspektiven für die Nutzung der Künstlichen Intelligenz im Bauwesen im Hinblick auf Expertensysteme*. In: Jahrbuch 1997, VDI-Gesellschaft Bautechnik; 1997
- [Schicker2000] Schicker, E.: *Datenbanken und SQL: eine praxisorientierte Einführung mit Hinweisen zu Oracle und MS-Access*. 3. durchges. Aufl. Stuttgart ; Leipzig ; Wiesbaden: Teubner, 2000
- [Schön1998] Schön, E.: *Das Resource Description Framework (RDF) - ein neuer Weg zur Verwaltung von Metadaten im Netz*. Technische Universität Ilmenau, Fakultät für Informatik und Automatisierung, Institut für praktische Informatik und Medieninformatik, URL: <http://www-ia.tu-ilmenau.de/~schoen/forschung/mik1998/rdf-einfuehrung.pdf>, 1998
- [Schreyer1999a] Schreyer, M.; Schwarte, J.: *A Virtual Advisor for Recycling Processes in Concrete Construction*. In: Otto-Graf-Journal (1999), Vol. 10, p. 76-86.

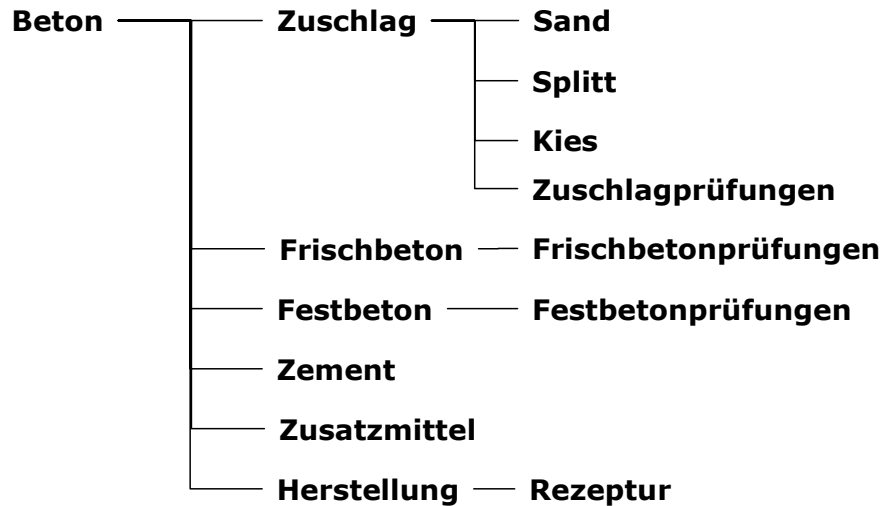
- [Schreyer1999b] Schreyer, M.; Schwarte, J.: *Ein virtueller Berater für Kreislaufprozesse im Massivbau*. In: IWB-Mitteilungen. 1999/2: Jahresbericht 1998-1999 Activities. Stuttgart: 1999, S. 73-78. - ISSN 0932-5921
- [Schreyer2001a] Schreyer, M.; Schwarte, J.: *New Building Materials - Knowledge Transfer via the WWW*. In: Association of Finish Civil Engineers RIL (Editor): Symposium Report on the 2nd Worldwide ECCE Symposium. Information and Communication Technology in the Practice of Building and Civil Engineering, RIL - Association of Finish Civil Engineers; VTT - Technical Research Centre of Finland, Building Technology, 6-8 June 2001, Espoo, Finland. Helsinki: Painopörssi Oy, 2001, p. 99-104. - ISBN 951-758-417-2
- [Schreyer2001b] Schreyer, M.; Schwarte J.: *Weborientierte Aufbereitung von Forschungsergebnissen am Beispiel innovativer Baustoffe*. In: Mitteilungen des Institut für Werkstoffe im Bauwesen. Bd. 2001/3: Jahresbericht 2000/2001 Activities. Stuttgart: 2001, S. 142-148; ISSN 0932-5921
- [Schwarte1997a] Schwarte, J.; Schreyer, M.: *Online Datenbanken im Intra-/Internet mit Windows NT*. In: Internet World (1997), Heft 8, S. 117-121.
- [Schwarte1997b] Schwarte, J.: *Uniform Resource Locators (URLs) - Mehr als nur Dateiadressen*. In: Internet World (1997), Heft 8, S. 54-55.
- [Schwarte1999] Schwarte J.; Schreyer M.: *Moderne Informationstechnik in der Baustoff-Forschung und -Lehre*. In: Werkstoffe und Werkstoffprüfung im Bauwesen - Festschrift zum 60. Geburtstag von H.-W. Reinhardt. Hamburg: Libri, 1999, S. 352-362.
- [Schwarte2000] Schwarte, J.; Schreyer, M.: *Bauingenieurwissen im Internet: Anforderungen-Erfahrungen-Perspektiven*. In: Bautechnik 77(2000), Heft 10, S. A24-A28.
- [Schwarte2002] Schwarte, J.: *"Werkstoffe im Bauwesen II" - Online*. In: 100-online e-letter (2002), URL: http://www.uni-stuttgart.de/100-online/e-letter/01_2002/gesamt/A_086.html
- [SCORM2002] Advanced Distributed Learning Network (ADLNet): *Sharable Content Object Reference Model (SCORM)*. URL:<http://www.adlnet.org>; 2002
- [SemanticWeb2001] Miller, E.; Swick, R.; Brickley, D.; McBride, B.; Hendler, J.; Schreiber, G.: *Semantic Web*. URL: <http://www.w3.org/2001/sw/>; 2001
- [Siebel1998] Siebel, E.; Kerkhoff, B.: *Teilprojekt E02: Einfluß von Recyclingzuschlägen aus Altbeton auf die Eigenschaften insbesondere die Dauerhaftigkeit des Betons*. Baustoffkreislauf im Massivbau. Zwischenbericht September 1998, FIZ Düsseldorf, URL: <http://www.b-i-m.de/Berichte/E02/E02z0998.htm>; 1998
- [SoftwareAG2001] Software AG: *Tamino XML Server - White Paper*. URL: <http://www.softwareag.com/tamino/download/tamino.pdf>; 2001
- [SQL1999] ISO (International Organization for Standardization): *ISO/IEC 9075-2:1999, Information technology --- Database languages --- SQL - Part 2: Foundation (SQL/Foundation)*. Geneva: International Organization for Standardization, URL: <http://www.iso.ch/cate/d26197.html>; 1999

- [SQL1999] JCC Consulting, Inc.: *SQL Standard*. URL: http://www.jcc.com/sql_std.html; 1999
- [Steinmann2001] Steinmann, R.; Tolman, F.; Stevens, J.; van Rees, R.; Böhm, M.; Zarli, A.: *bcXML - An XML Vocabulary for Building and Construction*. In Symposium report; 2nd Worldwide ECCE Symposium „Information and Communication Technology (ICT) in the Practice of Building and Civil Engineering“; 6-8 June; 2001; Espoo (Fin); p. 47-54
- [Sun2002] The Java Tutorial: *A practical guide for programmers*. URL: <http://java.sun.com/docs/books/tutorial/>; 2002
- [SunJSP] JavaServer Pages: *Dynamically Generated Web Content*. URL: <http://java.sun.com/products/jsp/>
- [SVG2001] Ferraiolo, J.: *Scalable Vector Graphics (SVG) 1.0 Specification*. W3C Recommendation 04 September 2001. URL: <http://www.w3.org/TR/SVG/>; 2001
- [Tewes1996] Tewes, J.: *Multimedia und Internet schieben ODBMS an*. In: Computerwoche Focus: Datenmanagement in verteilten Systemen, Band 18, Heft 3, 1996
- [T-Online2002] T-Online URL: <http://www.t-online.de>; 2002
- [Trinkwalder2000] Trinkwalder, A.: *Wortdetektive - Volltext-Suchmaschinen für Festplatte und Intranet*. In: c't, Heft 20, 2000, S.168-176
- [TUDMassiv1998] *Eignungsprüfung für Beton mit rezykliertem Zuschlag entsprechend der DAfStb-Richtlinie „Beton mit rezykliertem Zuschlag“*. In: Prüfungsbericht Nr. 273.98.2 (3. Dezember 1998) Bauvorhaben Waldspirale, Darmstadt. TU Darmstadt, Institut für Massivbau, 1998
- [Turk1999] Turk, Z.: *Four Questions about Construction Information Technology*. URL: <http://www.ce.berkeley.edu/~tommelein/CEMworkshop.htm>; 1999
- [UKOLN2000] UKOLN: *DC-Dot: Dublin Core metadata editor*. URL: <http://www.ukoln.ac.uk/metadata/dcdot/>; 2000
- [UML1997] *UML Semantics*. <http://www.rational.com/uml/>; 1997
- [Unicode2001] The Unicode Consortium: *The Unicode Character Database*. URL: <http://www.unicode.org/Public/3.1-Update/UnicodeCharacterDatabase-3.1.0.html>; 2001
- [vCard1998] Dawson, F.; Howes, T.: *vCard MIME Directory Profile*. Internet RFC 2426, 1998
- [W3C] *The World Wide Web Consortium*. URL: <http://www.w3.org>
- [W3CProcess2001] Jacobs, I.: *World Wide Web Consortium Process Document*. 19 July 2001. URL: <http://www.w3.org/Consortium/Process/>; 2001
- [Web.de2002] *Web.de*. URL: <http://web.de>; 2002
- [Web3D2002] Web 3D Consortium: *X3D draft specification (Milestone 3 February 24, 2002)*. URL: <http://www.web3d.org>; 2001

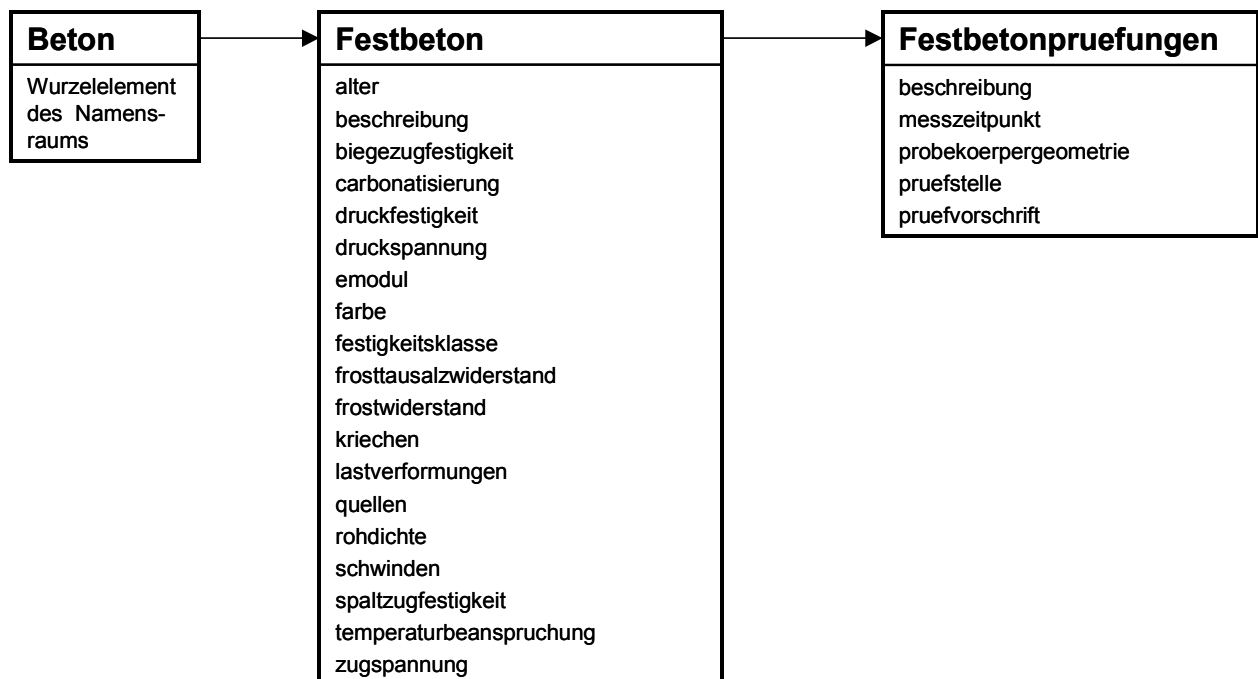
- [Weikert1998] Weikert, T.: *Entwicklung eines Online-Lexikons über die Themenbereiche Betontechnologie und Baustoffkreislauf für das WWW*. Diplomarbeit Universität Stuttgart, IWB. 1998.
- [Wesche1993] Wesche, K.: *Baustoffe für tragende Bauteile*. Band 2: Beton, Mauerwerk, 3. Auflage, Wiesbaden: Bauverlag, 1993
- [Wichmann1999] Wichmann, A.: *Aufbau und Techniken von Suchmaschinen für das WWW*. Seminararbeit: Rheinische Friedrich-Wilhelms-Universität Bonn
URL: <http://www-student.cs.uni-bonn.de/~wichmann/writings/webcrawlers/>; 1999
- [Winkler2001] Winkler, J.: *RDFed*. URL: http://www.jan-winkler.de/dev/d_rdfc.htm; 2001
- [XHTML2000] Pemberton, S. et al.: *XHTML™ 1.0: The Extensible HyperText Markup Language*. URL: <http://www.w3.org/TR/xhtml1/>; 2000
- [XLink2001] DeRose, S.; Maler, E.; Orchard, D.: *XML Linking Language (XLink)*. URL: <http://www.w3.org/TR/xlink/>; 2001
- [XML2000] Bra, T.; Paoli, J.; Sperberg-McQueen, C. M.: *Extensible Markup Language (XML) 1.0. Empfehlung des W3C, 10. Februar 1998*. URL: <http://www.w3.org/TR/2000/WD-xml-2e-20000814/>; 2000 (übersetzt von Behme, H.; Mintert, S.)
- [XMLS1999] Malhotra, A.; Maloney, M.: *XML-Schema Requirements*. URL: <http://www.w3.org/TR/1999/NOTE-xml-schema-req-19990215/>; 1999
- [XMLS2001a] Thompson, H. S.; Beech, D.; Maloney, M.; Mendelson, N.: *XML-Schema Part 1: Structures*. URL: <http://www.w3.org/TR/2001/REC-xmlschema-1-20010502/>; 2001
- [XMLS2001b] Biron, P. V.; Malhotra, A.: *XML-Schema Part 2: Datatypes*. W3C Recommendation 02 May 2001. URL: <http://www.w3.org/TR/2001/REC-xmlschema-2-20010502/>; 2001
- [XPointer2001] DeRose, S.: *XML Pointer Language (XPointer) Version 1.0. W3C Candidate Recommendation 11 September 2001*. URL: <http://www.w3.org/TR/xptr/>; 2001
- [XSL2000] Adler, S. et al.: *Extensible Stylesheet Language (XSL)*. URL: <http://www.w3.org/TR/2000/CR-xsl-20001121/>; 2000
- [XSLT2001] Kay, M.: *XSL Transformations (XSLT) Version 2.0. W3C Working Draft 20 December 2001*. URL: <http://www.w3.org/TR/xslt20/>; 2001
- [Yahoo2002] *Yahoo!* URL: <http://www.yahoo.com>; 2002

Anhang A met@BiM-Schema

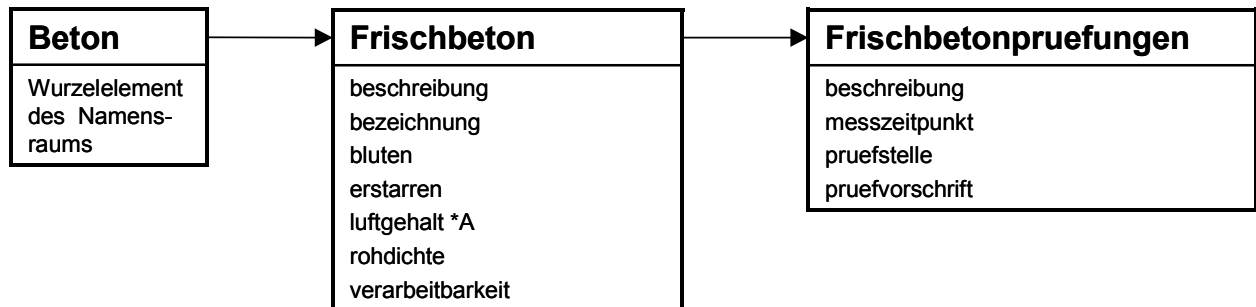
Anhang A1 Ontologie betontechnologischer Informationsressourcen



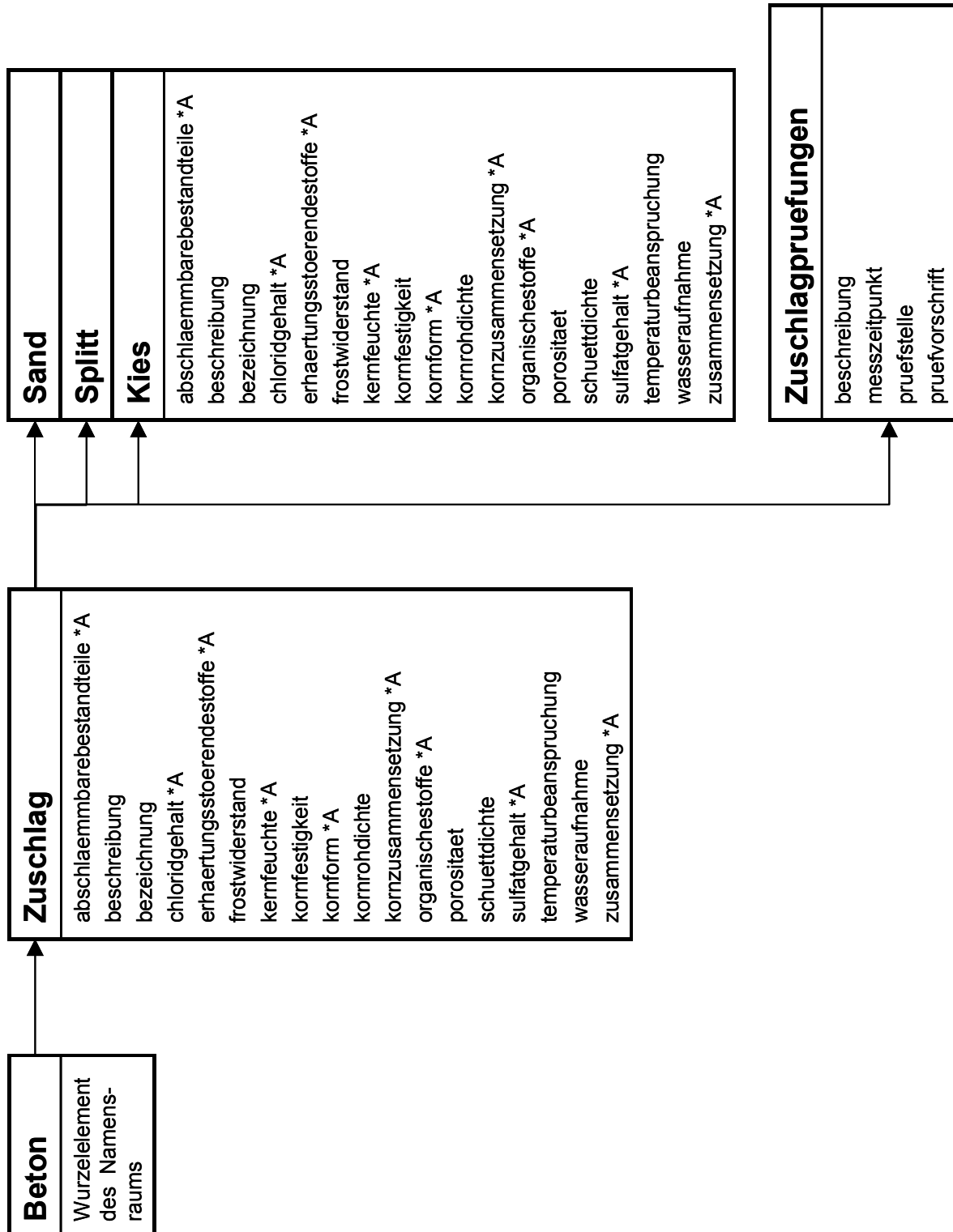
Anhang A2 Schema Festbeton



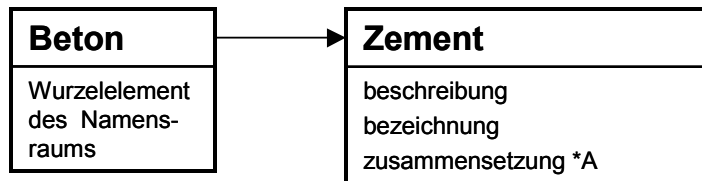
Anhang A3 Schema Frischbeton



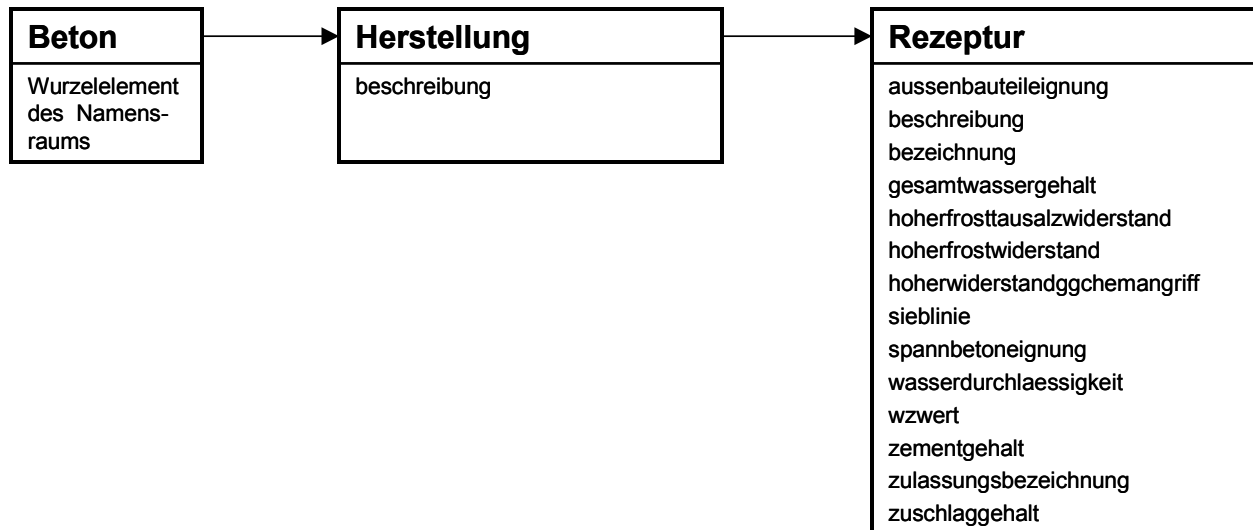
Anhang A4 Schema Betonzuschlag



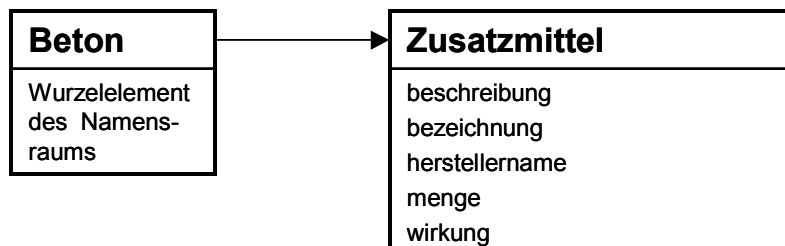
Anhang A5 Schema Zement



Anhang A6 Schema Herstellung



Anhang A7 Schema Zusatzmittel



Anhang B XML-Serialisierung des met@BiM-RDF-Schemas

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!-- ISO-8859-1 Angabe für korrekte Interpretation der Umlaute -->
<!--

RDF-Schema declaration for the building materials data element set 1.0
2002/02/01
comments etc. to <schreyer@iwb.uni-stuttgart.de>

This Version: http://www.b-i-m.de/2001/01/metabim#
Latest Version: http://www.b-i-m.de/2001/01/metabim#
Previous Version: -

Editors:
Marcus Schreyer, University of Stuttgart

-->
<!DOCTYPE rdf:RDF>
<rdf:RDF xml:lang="de"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:dc="http://purl.org/dc/elements/1.1/"
>

<!-- Description of the Schema -->

<rdf:Description rdf:about="http://www.b-i-m.de/2001/01/metabim#">
  <dc:title>The met@BiM (metabim) Element Set v1.0</dc:title>
  <dc:creator>Marcus Schreyer</dc:creator>
  <dc:publisher>Institute of Construction Materials (IWB), University of
    Stuttgart (Germany)</dc:publisher>
  <dc:description>The met@bim metadata vocabulary is a simple vocabulary
    intended to facilitate discovery of resources concerning
    construction materials in the WWW. It's part of my work
    on a semantic description of the contents of BiM-Online
    (http://www.b-i-m.de)</dc:description>
  <dc:identifier>http://www.b-i-m.de/2001/01/metabim#</dc:identifier>
  <dc:language>DE</dc:language>
  <dc:relation>http://www.b-i-m.de</dc:relation>
  <dc:date>2002-02-15</dc:date>
</rdf:Description>

<!-- Beginn der metabim RDFS Deklaration -->

<!-- Deklaration der Klassen in BiM-Online -->

<rdf:Class ID="DOKUMENT">
  <rdfs:subClassOf
    rdf:resource="http://www.w3.org/2000/01/rdf-schema#Resource"/>
  <rdfs:label>Dokument</rdfs:label>
  <rdfs:comment>Webdokumente mit betontechnologischen Informationen, z.B.
    Forschungsberichte, Fachveröffentlichungen, Projektbe-
    schreibungen, Dissertationen etc.</rdfs:comment>
</rdf:Class>

<rdf:Class ID="KURS">
  <rdfs:subClassOf
    rdf:resource="http://www.w3.org/2000/01/rdf-schema#Resource"/>
```

```

    <rdfs:label>Online-Kurs</rdfs:label>
    <rdfs:comment>Interaktive Online Kurse, die auch aus mehreren, zusammen-
                    hängenden Dokumenten bestehen können. Z.B. interaktive
                    Flash Animationen, Online Präsentationen, BiM-Online
                    Lektionen etc.</rdfs:comment>
</rdf:Class>

<rdf:Class ID="MATHEMATISCHEFORMEL">
    <rdfs:subClassOf
        rdf:resource="http://www.w3.org/2000/01/rdf-schema#Resource"/>
    <rdfs:label>Mathematische Formel</rdfs:label>
    <rdfs:comment>Mathematische Gleichung</rdfs:comment>
</rdf:Class>

<rdf:Class ID="CHEMISCHEFORMEL">
    <rdfs:subClassOf
        rdf:resource="http://www.w3.org/2000/01/rdf-schema#Resource"/>
    <rdfs:label>Chemische Formel</rdfs:label>
    <rdfs:comment>Chemische Gleichung</rdfs:comment>
</rdf:Class>

<rdf:Class ID="BETON">
    <rdfs:subClassOf
        rdf:resource="http://www.w3.org/2000/01/rdf-schema#Resource"/>
    <rdfs:label>Betontechnologische Daten</rdfs:label>
    <rdfs:comment>In dieser Klasse werden Ressourcen und Properties zur
                    Modellierung von Daten aus der Betonherstellung sowie
                    Materialprüfungen zu Beton definiert</rdfs:comment>
</rdf:Class>

<!-- Deklaration der betontechnologischen Klassen -->

<rdf:Class ID="FESTBETON">
    <rdfs:subClassOf
        rdf:resource="http://www.b-i-m.de/2001/01/metabim#BETON"/>
    <rdfs:label>Festbeton</rdfs:label>
    <rdfs:comment>Daten zu Beton im erstarrtem Zustand, z.B. ein
                    Probekörper</rdfs:comment>
</rdf:Class>

<rdf:Class ID="FESTBETONPRUEFUNGEN">
    <rdfs:subClassOf
        rdf:resource="http://www.b-i-m.de/2001/01/metabim#FESTBETON"/>
    <rdfs:label>Festbetonprüfungen</rdfs:label>
    <rdfs:comment>Daten zu Materialprüfungen an Beton im erstarrtem Zustand
    </rdfs:comment>
</rdf:Class>

<rdf:Class ID="FRISCHBETON">
    <rdfs:subClassOf
        rdf:resource="http://www.b-i-m.de/2001/01/metabim#BETON"/>
    <rdfs:label>Frischbeton</rdfs:label>
    <rdfs:comment>Daten zu Beton im noch nicht erstarrtem Zustand
    </rdfs:comment>
</rdf:Class>

<rdf:Class ID="FRISCHBETONPRUEFUNGEN">
    <rdfs:subClassOf
        rdf:resource="http://www.b-i-m.de/2001/01/metabim#FRISCHBETON"/>
    <rdfs:label>Frischbetonprüfungen</rdfs:label>
    <rdfs:comment>Daten zu Materialprüfungen an Beton im noch nicht
                    erstarrtem Zustand</rdfs:comment>
</rdf:Class>

```

```

<rdf:Class ID="ZUSCHLAG">
  <rdfs:subClassOf
    rdf:resource="http://www.b-i-m.de/2001/01/metabim#BETON"/>
  <rdfs:label>Betonzuschlag</rdfs:label>
  <rdfs:comment>Daten zu Zuschlagstoffen für Beton</rdfs:comment>
</rdf:Class>

<rdf:Class ID="SAND">
  <rdfs:subClassOf
    rdf:resource="http://www.b-i-m.de/2001/01/metabim#ZUSCHLAG"/>
  <rdfs:label>Sand</rdfs:label>
  <rdfs:comment>Daten zu Sand als Zuschlagfraktion 0/2 mm zur Verwendung
    als Betonzuschlag</rdfs:comment>
</rdf:Class>

<rdf:Class ID="KIES">
  <rdfs:subClassOf
    rdf:resource="http://www.b-i-m.de/2001/01/metabim#ZUSCHLAG"/>
  <rdfs:label>Kies</rdfs:label>
  <rdfs:comment>Daten zu Kies als auf natürliche Weise entstandene
    Zuschlagfraktion größer 2 mm, z.B. Flussskies
  </rdfs:comment>
</rdf:Class>

<rdf:Class ID="SPLITT">
  <rdfs:subClassOf
    rdf:resource="http://www.b-i-m.de/2001/01/metabim#ZUSCHLAG"/>
  <rdfs:label>Splitt</rdfs:label>
  <rdfs:comment>Daten zu Spitt als z.B. im Brecher maschinell zerkleinerte
    Zuschlagfraktion grösser 2 mm</rdfs:comment>
</rdf:Class>

<rdf:Class ID="ZUSCHLAGPRUEFUNG">
  <rdfs:subClassOf
    rdf:resource="http://www.b-i-m.de/2001/01/metabim#ZUSCHLAG"/>
  <rdfs:label>Zuschlagprüfungen</rdfs:label>
  <rdfs:comment>Daten zu Materialprüfungen an Betonzuschlag</rdfs:comment>
</rdf:Class>

<rdf:Class ID="ZEMENT">
  <rdfs:subClassOf
    rdf:resource="http://www.b-i-m.de/2001/01/metabim#BETON"/>
  <rdfs:label>Zement</rdfs:label>
  <rdfs:comment>Daten zu Zement</rdfs:comment>
</rdf:Class>

<rdf:Class ID="ZUSATZMITTEL">
  <rdfs:subClassOf
    rdf:resource="http://www.b-i-m.de/2001/01/metabim#BETON"/>
  <rdfs:label>Zusatzmittel</rdfs:label>
  <rdfs:comment>Daten zu Zusatzmitteln zur Beeinflussung der Betoneigen-
    schaften</rdfs:comment>
</rdf:Class>

<rdf:Class ID="HERSTELLUNG">
  <rdfs:subClassOf
    rdf:resource="http://www.b-i-m.de/2001/01/metabim#BETON"/>
  <rdfs:label>Herstellung</rdfs:label>
  <rdfs:comment>Daten zur Betonherstellung</rdfs:comment>
</rdf:Class>

<rdf:Class ID="REZEPTUR">

```

```

    <rdfs:subClassOf
      rdf:resource="http://www.b-i-m.de/2001/01/metabim#BETON"/>
    <rdfs:label>REZEPTUR</rdfs:label>
    <rdfs:comment>Daten zur Betonrezeptur</rdfs:comment>
  </rdf:Class>

<rdf:Class ID="EINHEITEN">
  <rdfs:label>EINHEITEN</rdfs:label>
  <rdfs:comment>Einheiten für Werte der met@BiM Properties</rdfs:comment>
  <EINHEIT rdf:ID="mm"/>
  <EINHEIT rdf:ID="cm"/>
  <EINHEIT rdf:ID="m"/>
  <EINHEIT rdf:ID="mm2"/>
  <EINHEIT rdf:ID="cm2"/>
  <EINHEIT rdf:ID="m2"/>
  <EINHEIT rdf:ID="cm3"/>
  <EINHEIT rdf:ID="m3"/>
  <EINHEIT rdf:ID="s"/>
  <EINHEIT rdf:ID="min"/>
  <EINHEIT rdf:ID="h"/>
  <EINHEIT rdf:ID="d"/>
  <EINHEIT rdf:ID="a"/>
  <EINHEIT rdf:ID="N"/>
  <EINHEIT rdf:ID="KN"/>
  <EINHEIT rdf:ID="MN"/>
  <EINHEIT rdf:ID="Nmm2"/>
  <EINHEIT rdf:ID="KNcm2"/>
  <EINHEIT rdf:ID="g"/>
  <EINHEIT rdf:ID="kg"/>
  <EINHEIT rdf:ID="gcm3"/>
  <EINHEIT rdf:ID="kgcm3"/>
  <EINHEIT rdf:ID="kgm3"/>
  <EINHEIT rdf:ID="°C"/>
  <EINHEIT rdf:ID="K"/>
  <EINHEIT rdf:ID="°"/>
  <EINHEIT rdf:ID="M.-%"/>
  <EINHEIT rdf:ID="Vol.-%"/>
  <EINHEIT rdf:ID="FTW"/>
  <EINHEIT rdf:ID="mmm"/>
  <EINHEIT rdf:ID="lm3"/>
</rdf:Class>

<!-- Deklaration der Properties -->

<rdf:Property ID="abschlaemmbarbestandteile">
  <rdfs:domain rdf:resource="#ZUSCHLAG"/>
  <rdfs:domain rdf:resource="#KIES"/>
  <rdfs:domain rdf:resource="#SAND"/>
  <rdfs:domain rdf:resource="#SPLITT"/>
  <rdfs:comment>Abschlämbbare Bestandteile (vgl. DIN 4226, Teil 3)
  </rdfs:comment>
</rdf:Property>

<rdf:Property ID="alter">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:comment>Alter der Probe</rdfs:comment>
</rdf:Property>

<rdf:Property ID="aussenbauteileignung">
  <rdfs:domain rdf:resource="#REZEPTUR"/>
  <rdfs:comment>Eignung der Betonmischung für Außenbauteile (vgl. DAfStb
    Richtlinie "Beton mit rezykliertem Zuschlag")
  </rdfs:comment>

```

```
</rdf:Property>

<rdf:Property ID="beschreibung">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:domain rdf:resource="#FESTBETONPRUEFUNGEN"/>
  <rdfs:domain rdf:resource="#ZUSCHLAG"/>
  <rdfs:domain rdf:resource="#KIES"/>
  <rdfs:domain rdf:resource="#SAND"/>
  <rdfs:domain rdf:resource="#SPLITT"/>
  <rdfs:domain rdf:resource="#ZUSCHLAGPRUEFUNGEN"/>
  <rdfs:domain rdf:resource="#FRISCHBETON"/>
  <rdfs:domain rdf:resource="#FRISCHBETONPRUEFUNGEN"/>
  <rdfs:domain rdf:resource="#HERSTELLUNG"/>
  <rdfs:domain rdf:resource="#REZEPTUR"/>
  <rdfs:domain rdf:resource="#ZEMENT"/>
  <rdfs:domain rdf:resource="#ZUSATZMITTEL"/>
  <rdfs:comment>Kurze Erläuterung der Besonderheiten</rdfs:comment>
</rdf:Property>

<rdf:Property ID="biegezugfestigkeit">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:comment>Zugfestigkeit des Probekörpers ermittelt im Biegezug-
    versuch</rdfs:comment>
</rdf:Property>

<rdf:Property ID="bluten">
  <rdfs:domain rdf:resource="#FRISCHBETON"/>
  <rdfs:comment>Angaben zum Bluten des Betons beim Verdichten
</rdfs:comment>
</rdf:Property>

<rdf:Property ID="carbonatisierung">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:comment>Carbonatisierungstiefe</rdfs:comment>
</rdf:Property>

<rdf:Property ID="chloridgehalt">
  <rdfs:domain rdf:resource="#ZUSCHLAG"/>
  <rdfs:domain rdf:resource="#KIES"/>
  <rdfs:domain rdf:resource="#SAND"/>
  <rdfs:domain rdf:resource="#SPLITT"/>
  <rdfs:comment>Angaben zum Chloridgehalt</rdfs:comment>
</rdf:Property>

<rdf:Property ID="druckfestigkeit">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:comment>Druckfestigkeit des Probekörpers</rdfs:comment>
</rdf:Property>

<rdf:Property ID="druckspannung">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:comment>Angaben zur Druckspannung im Probekörper während eines
    Versuchs</rdfs:comment>
</rdf:Property>

<rdf:Property ID="emodul">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:comment>Angaben zum E-Modul der Probe</rdfs:comment>
</rdf:Property>

<rdf:Property ID="erhaertungsstoerendestoffe">
  <rdfs:domain rdf:resource="#ZUSCHLAG"/>
  <rdfs:domain rdf:resource="#KIES"/>
```

```

    <rdfs:domain rdf:resource="#SAND"/>
    <rdfs:domain rdf:resource="#SPLITT"/>
    <rdfs:comment>Angaben zum Gehalt erhärtungsstörender Stoffe (vgl. DIN
        4226, Teil 3)</rdfs:comment>
</rdf:Property>

<rdf:Property ID="erstarren">
    <rdfs:domain rdf:resource="#FRISCHBETON"/>
    <rdfs:comment>Angaben, die das Erstarren des Frischbetons beschreiben
    </rdfs:comment>
</rdf:Property>

<rdf:Property ID="farbe">
    <rdfs:domain rdf:resource="#FESTBETON"/>
    <rdfs:comment>Farbe des Probekörpers</rdfs:comment>
</rdf:Property>

<rdf:Property ID="festigkeitsklasse">
    <rdfs:domain rdf:resource="#FESTBETON"/>
    <rdfs:comment>Einstufung der Probe in eine Festigkeitsklasse. Klassen-
        angabe ohne Leerzeichen (z.B. C20/25, B25)</rdfs:comment>
</rdf:Property>

<rdf:Property ID="frostdausalzwiderstand">
    <rdfs:domain rdf:resource="#FESTBETON"/>
    <rdfs:comment>Frost-Tausalz Widerstand der Probe (vgl. DIN 4226, Teil 1
        bzw. DAfStb Heft 422)</rdfs:comment>
</rdf:Property>

<rdf:Property ID="frostdwiderstand">
    <rdfs:domain rdf:resource="#FESTBETON"/>
    <rdfs:domain rdf:resource="#ZUSCHLAG"/>
    <rdfs:domain rdf:resource="#KIES"/>
    <rdfs:domain rdf:resource="#SAND"/>
    <rdfs:domain rdf:resource="#SPLITT"/>
    <rdfs:comment>Frostdwiderstand der Probe (vgl. DIN 4226, Teil 1 bzw.
        DAfStb Heft 422)</rdfs:comment>
</rdf:Property>

<rdf:Property ID="gesamtwassergehalt">
    <rdfs:domain rdf:resource="#REZEPTUR"/>
    <rdfs:comment>Gesamtwassergehalt der Betonmischung, z.B. Zugabewasser
        und Kernfeuchte des Zuschlags</rdfs:comment>
</rdf:Property>

<rdf:Property ID="herstellername">
    <rdfs:domain rdf:resource="#HERSTELLUNG"/>
    <rdfs:comment>Name des Herstellers, z.B. Betonwerk, Zuschlaghersteller
        etc.</rdfs:comment>
</rdf:Property>

<rdf:Property ID="hoherfrostdausalzwiderstand">
    <rdfs:domain rdf:resource="#REZEPTUR"/>
    <rdfs:comment>Eignung der Betonmischung für die erhöhte Beanspruchung
        durch Frost-Tausalz (vgl. DIN 4226, Teil 1 bzw. DAfStb
        Heft 422, DAfStb-Richtlinie "Beton mit rezykliertem
        Zuschlag")</rdfs:comment>
</rdf:Property>

<rdf:Property ID="hoherfrostdwiderstand">
    <rdfs:domain rdf:resource="#REZEPTUR"/>
    <rdfs:comment>Eignung der Betonmischung für die erhöhte Beanspruchung
        durch Frost (vgl. DIN 4226, Teil 1 bzw. DAfStb Heft 422,

```

```

        DAFStb-Richtlinie "Beton mit rezykliertem
        Zuschlag")</rdfs:comment>
</rdf:Property>

<rdf:Property ID="hoherwiderstandggchemangriff">
  <rdfs:domain rdf:resource="#REZEPTUR"/>
  <rdfs:comment>Eignung der Betonmischung für die erhöhte Beanspruchung
    durch schwachen chemischen Angriff (vgl. DIN 1045, DAFStb
    Richtlinie "Beton mit rezykliertem Zuschlag")
  </rdfs:comment>
</rdf:Property>

<rdf:Property ID="kernfeuchte">
  <rdfs:domain rdf:resource="#ZUSCHLAG"/>
  <rdfs:domain rdf:resource="#KIES"/>
  <rdfs:domain rdf:resource="#SAND"/>
  <rdfs:domain rdf:resource="#SPLITT"/>
  <rdfs:comment>Angaben zur Kernfeuchte der Zuschlagkörner</rdfs:comment>
</rdf:Property>

<rdf:Property ID="kornfestigkeit">
  <rdfs:domain rdf:resource="#ZUSCHLAG"/>
  <rdfs:domain rdf:resource="#KIES"/>
  <rdfs:domain rdf:resource="#SAND"/>
  <rdfs:domain rdf:resource="#SPLITT"/>
  <rdfs:comment>Angaben zur Druckfestigkeit der Zuschlagkörner
  </rdfs:comment>
</rdf:Property>

<rdf:Property ID="kornform">
  <rdfs:domain rdf:resource="#ZUSCHLAG"/>
  <rdfs:domain rdf:resource="#KIES"/>
  <rdfs:domain rdf:resource="#SAND"/>
  <rdfs:domain rdf:resource="#SPLITT"/>
  <rdfs:comment>Angaben zur Kornform (vgl. DIN 4226, Teil 3)
  </rdfs:comment>
</rdf:Property>

<rdf:Property ID="kornrohddichte">
  <rdfs:domain rdf:resource="#ZUSCHLAG"/>
  <rdfs:domain rdf:resource="#KIES"/>
  <rdfs:domain rdf:resource="#SAND"/>
  <rdfs:domain rdf:resource="#SPLITT"/>
  <rdfs:comment>Angaben zur mittlere Rohddichte der Zuschlagkörner
  </rdfs:comment>
</rdf:Property>

<rdf:Property ID="kornzusammensetzung">
  <rdfs:domain rdf:resource="#ZUSCHLAG"/>
  <rdfs:domain rdf:resource="#KIES"/>
  <rdfs:domain rdf:resource="#SAND"/>
  <rdfs:domain rdf:resource="#SPLITT"/>
  <rdfs:comment>Angaben zu den einzelnen Komponenten des Zuschlags
    innerhalb einer Fraktion</rdfs:comment>
</rdf:Property>

<rdf:Property ID="kriechen">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:comment>Verformungen der Probe durch Kriechen</rdfs:comment>
</rdf:Property>

<rdf:Property ID="lastverformungen">
  <rdfs:domain rdf:resource="#FESTBETON"/>

```

```
<rdfs:comment>Verformungen unter aufgebrachtter Last</rdfs:comment>
</rdf:Property>

<rdf:Property ID="luftgehalt">
  <rdfs:domain rdf:resource="#FRISCHBETON"/>
  <rdfs:comment>Angaben zum Luftgehalt von Frischbeton</rdfs:comment>
</rdf:Property>

<rdf:Property ID="menge">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:domain rdf:resource="#ZUSCHLAG"/>
  <rdfs:domain rdf:resource="#KIES"/>
  <rdfs:domain rdf:resource="#SAND"/>
  <rdfs:domain rdf:resource="#SPLITT"/>
  <rdfs:domain rdf:resource="#FRISCHBETON"/>
  <rdfs:domain rdf:resource="#ZUSATZMITTEL"/>
  <rdfs:comment>Absolute Zugabemenge</rdfs:comment>
</rdf:Property>

<rdf:Property ID="messzeitpunkt">
  <rdfs:domain rdf:resource="#FESTBETONPRUEFUNGEN"/>
  <rdfs:domain rdf:resource="#FRISCHBETONPRUEFUNGEN"/>
  <rdfs:domain rdf:resource="#ZUSCHLAGPRUEFUNGEN"/>
  <rdfs:comment>Messzeitpunkte der Wertermittlung</rdfs:comment>
</rdf:Property>

<rdf:Property ID="organischestoffe">
  <rdfs:domain rdf:resource="#ZUSCHLAG"/>
  <rdfs:domain rdf:resource="#KIES"/>
  <rdfs:domain rdf:resource="#SAND"/>
  <rdfs:domain rdf:resource="#SPLITT"/>
  <rdfs:comment>Angaben zum Gehalt an organischen Stoffen</rdfs:comment>
</rdf:Property>

<rdf:Property ID="probekoerpergeometrie">
  <rdfs:domain rdf:resource="#FESTBETONPRUEFUNGEN"/>
  <rdfs:comment>Kurze Beschreibung der Probekörpergeometrie z.B. Form,
    Maße</rdfs:comment>
</rdf:Property>

<rdf:Property ID="pruefstelle">
  <rdfs:domain rdf:resource="#FESTBETONPRUEFUNGEN"/>
  <rdfs:domain rdf:resource="#FRISCHBETONPRUEFUNGEN"/>
  <rdfs:domain rdf:resource="#ZUSCHLAGPRUEFUNGEN"/>
  <rdfs:comment>Name und Angaben zur prüfenden Stelle</rdfs:comment>
</rdf:Property>

<rdf:Property ID="pruefvorschrift">
  <rdfs:domain rdf:resource="#FESTBETONPRUEFUNGEN"/>
  <rdfs:domain rdf:resource="#FRISCHBETONPRUEFUNGEN"/>
  <rdfs:domain rdf:resource="#ZUSCHLAGPRUEFUNGEN"/>
  <rdfs:comment>Der Prüfung zugrundegelegte Vorschrift oder Richtlinie
  </rdfs:comment>
</rdf:Property>

<rdf:Property ID="quellen">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:comment>Verformungen durch Quellen</rdfs:comment>
</rdf:Property>

<rdf:Property ID="rohdichte">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:domain rdf:resource="#FRISCHBETON"/>
```

```

    <rdfs:comment>Rohdichte des Probekörpers</rdfs:comment>
</rdf:Property>

<rdf:Property ID="schuettdichte">
    <rdfs:domain rdf:resource="#ZUSCHLAG"/>
    <rdfs:domain rdf:resource="#KIES"/>
    <rdfs:domain rdf:resource="#SAND"/>
    <rdfs:domain rdf:resource="#SPLITT"/>
    <rdfs:comment>Angaben zur Schüttdichte</rdfs:comment>
</rdf:Property>

<rdf:Property ID="schwinden">
    <rdfs:domain rdf:resource="#FESTBETON"/>
    <rdfs:comment>Verformungen durch Schwinden</rdfs:comment>
</rdf:Property>

<rdf:Property ID="sieblinie">
    <rdfs:domain rdf:resource="#REZEPTUR"/>
    <rdfs:comment>Gewählte Sieblinienklassifizierung</rdfs:comment>
</rdf:Property>

<rdf:Property ID="spaltzugfestigkeit">
    <rdfs:domain rdf:resource="#FESTBETON"/>
    <rdfs:comment>Zugfestigkeit des Probekörpers ermittelt im Spaltzug-
        versuch</rdfs:comment>
</rdf:Property>

<rdf:Property ID="spannbetoneignung">
    <rdfs:domain rdf:resource="#REZEPTUR"/>
    <rdfs:comment>Eignung der Betonmischung für die Verwendung in
        Spannbetonbauteilen (vgl. DIN 4227 T.1)</rdfs:comment>
</rdf:Property>

<rdf:Property ID="sulfatgehalt">
    <rdfs:domain rdf:resource="#ZUSCHLAG"/>
    <rdfs:domain rdf:resource="#KIES"/>
    <rdfs:domain rdf:resource="#SAND"/>
    <rdfs:domain rdf:resource="#SPLITT"/>
    <rdfs:comment>Angaben zum Sulfatgehalt</rdfs:comment>
</rdf:Property>

<rdf:Property ID="temperaturbeanspruchung">
    <rdfs:domain rdf:resource="#FESTBETON"/>
    <rdfs:domain rdf:resource="#ZUSCHLAG"/>
    <rdfs:domain rdf:resource="#KIES"/>
    <rdfs:domain rdf:resource="#SAND"/>
    <rdfs:domain rdf:resource="#SPLITT"/>
    <rdfs:comment>Temperaturbeanspruchung der Probe, z.B. vor Lastversuchen
    </rdfs:comment>
</rdf:Property>

<rdf:Property ID="verarbeitbarkeit">
    <rdfs:domain rdf:resource="#FRISCHBETON"/>
    <rdfs:comment>Angaben zur Verarbeitbarkeit des Frischbetons. Dies können
        sowohl die numerischen Werte aus Ausbreitversuchen sein,
        als auch daraus abgeleitete Angaben, wie z.B. die
        Konsistenzklasse</rdfs:comment>
</rdf:Property>

<rdf:Property ID="wasseraufnahme">
    <rdfs:domain rdf:resource="#ZUSCHLAG"/>
    <rdfs:domain rdf:resource="#KIES"/>
    <rdfs:domain rdf:resource="#SAND"/>

```

```

    <rdfs:domain rdf:resource="#SPLITTT"/>
    <rdfs:comment>Angaben zur Wasseraufnahme des Zuschlaggemisches
  </rdfs:comment>
</rdf:Property>

<rdf:Property ID="wasserundurchlaessigkeit">
  <rdfs:domain rdf:resource="#REZEPTUR"/>
  <rdfs:comment>Eignung der Betonmischung für den erhöhten Anspruch an die
    Wasserundurchlässigkeit (vgl. DIN 1045-2 bzw. 1045, DAFStb
    Richtlinie "Beton mit rezykliertem Zuschlag")
  </rdfs:comment>
</rdf:Property>

<rdf:Property ID="wirkung">
  <rdfs:domain rdf:resource="#ZUSATZMITTEL"/>
  <rdfs:comment>Kurze Erläuterung der Wirkung, die durch die Zugabe auf
    den Beton erzielt wird</rdfs:comment>
</rdf:Property>

<rdf:Property ID="wzwert">
  <rdfs:domain rdf:resource="#REZEPTUR"/>
  <rdfs:comment>Wasser-Zement Wert (w/z-Wert) der Mischung</rdfs:comment>
</rdf:Property>

<rdf:Property ID="zementgehalt">
  <rdfs:domain rdf:resource="#REZEPTUR"/>
  <rdfs:comment>Zementgehalt der Mischung</rdfs:comment>
</rdf:Property>

<rdf:Property ID="zugspannung">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:comment>Angaben zur Zugspannung im Probekörper während eines
    Versuchs</rdfs:comment>
</rdf:Property>

<rdf:Property ID="zulassungsbezeichnung">
  <rdfs:domain rdf:resource="#REZEPTUR"/>
  <rdfs:comment>Bezeichnung der Zulassung für die Mischung</rdfs:comment>
</rdf:Property>

<rdf:Property ID="zuschlaggehalt">
  <rdfs:domain rdf:resource="#REZEPTUR"/>
  <rdfs:comment>Anteil des Zuschlags in der Mischung</rdfs:comment>
</rdf:Property>

<rdf:Property ID="zusammensetzung">
  <rdfs:domain rdf:resource="#FESTBETON"/>
  <rdfs:comment>Beschreibung der unterschiedlichen Komponenten des
    Zuschlaggemisches</rdfs:comment>
</rdf:Property>

<!-- Deklaration der Werteeigenschaften -->

<rdf:Property ID="einheit">
  <rdfs:range rdf:resource="#EINHEITEN"/>
  <rdfs:comment>Ausdruck der physikalischen Kenngröße und mathematischen
    Dimension, die dem Wert zugewiesen wird</rdfs:comment>
</rdf:Property>

<rdf:Property ID="anteil">
  <rdfs:comment>Relativer Wert der beschriebenen Eigenschaft
  </rdfs:comment>
</rdf:Property>

```

```
<rdf:Property ID="wertebereich">
  <rdfs:comment>Wertebereich der beschriebenen Eigenschaft</rdfs:comment>
</rdf:Property>

<rdf:Property ID="obergrenze">
  <rdfs:subPropertyOf rdf:resource="#wertebereich"/>
  <rdfs:comment>Obergrenze des Wertebereichs</rdfs:comment>
</rdf:Property>

<rdf:Property ID="untergrenze">
  <rdfs:subPropertyOf rdf:resource="#wertebereich"/>
  <rdfs:comment>Untergrenze des Wertebereichs</rdfs:comment>
</rdf:Property>

</rdf:RDF>
```

Anhang C XML-Serialisierung des Dublin Core-Metadaten-RDF-Schemas

```

<?xml version="1.0" ?>
<!--
  RDF-Schema declaration for the Dublin Core Element Set 1.1
  2001/08/14
  comments, etc. to webteam <dcmi-feedback@dublincore.org>
-->

<!DOCTYPE rdf:RDF (View Source for full doctype...)>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:eor="http://dublincore.org/2000/03/13/eor#">

  <!-- Description of Schema -->
  <eor:Schema rdf:about="http://purl.org/dc/elements/1.1/">
    <rdf:value>The Dublin Core Element Set v1.1</rdf:value>
    <dc:title>The Dublin Core Element Set v1.1</dc:title>
    <dc:publisher>The Dublin Core Metadata Initiative</dc:publisher>
    <dc:description>The Dublin Core metadata vocabulary is a simple
      vocabulary intended to facilitate discovery of resources.
    </dc:description>
    <dc:language>English</dc:language>
    <dc:relation rdf:resource="http://dublincore.org/documents/dces/" />
    <dc:date>2000-07-02</dc:date>
  </eor:Schema>

  <!-- Begin: Title Declaration -->
  <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/title">
    <rdfs:label>Title</rdfs:label>
    <rdfs:comment>A name given to the resource.</rdfs:comment>
    <eor:comment>Typically, a Title will be a name by which the resource is
      formally known.</eor:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
  </rdf:Property>

  <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/contributor">
    <rdfs:label>Contributor</rdfs:label>
    <rdfs:comment>An entity responsible for making contributions to the
      content of the resource.</rdfs:comment>
    <eor:comment>Examples of a Contributor include a person, an organisation,
      or a service. Typically, the name of a Contributor should be
      used to indicate the entity.</eor:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
  </rdf:Property>

  <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/creator">
    <rdfs:label>Author/Creator</rdfs:label>
    <rdfs:comment>An entity primarily responsible for making the content of
      the resource.</rdfs:comment>
    <eor:comment>Examples of a Creator include a person, an organisation, or
      a service. Typically, the name of a Creator should be used
      to indicate the entity.</eor:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
  </rdf:Property>

  <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/publisher">
    <rdfs:label>Publisher</rdfs:label>
    <rdfs:comment>An entity responsible for making the resource available.
  </rdfs:comment>

```

```

    <eor:comment>Examples of a Publisher include a person, an organisation,
        or a service. Typically, the name of a Publisher should be
        used to indicate the entity.</eor:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
</rdf:Property>

<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/subject">
    <rdfs:label>Subject</rdfs:label>
    <rdfs:comment>The topic of the content of the resource.</rdfs:comment>
    <eor:comment>Typically, a Subject will be expressed as keywords, key
        phrases or classification codes that describe a topic of the
        resource. Recommended best practice is to select a value
        from a controlled vocabulary or formal classification
        scheme.</eor:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
</rdf:Property>

<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/description">
    <rdfs:label>Description</rdfs:label>
    <rdfs:comment>An account of the content of the resource.</rdfs:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
</rdf:Property>

<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/date">
    <rdfs:label>Date</rdfs:label>
    <rdfs:comment>A date associated with an event in the life cycle of the
        resource.</rdfs:comment>
    <eor:comment>Typically, Date will be associated with the creation or
        availability of the resource. Recommended best practice for
        encoding the date value is defined in a profile of ISO 8601
        [W3CDTF] and follows the YYYY-MM-DD format.</eor:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
</rdf:Property>

<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/type">
    <rdfs:label>Resource Type</rdfs:label>
    <rdfs:comment>The nature or genre of the content of the
        resource.</rdfs:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
</rdf:Property>

<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/format">
    <rdfs:label>Format</rdfs:label>
    <rdfs:comment>The physical or digital manifestation of the
        resource.</rdfs:comment>
    <eor:comment>Typically, Format may include the media-type or dimensions
        of the resource. Format may be used to determine the
        software, hardware or other equipment needed to display or
        operate the resource. Examples of dimensions include size
        and duration. Recommended best practice is to select a value
        from a controlled vocabulary (for example, the list of
        Internet Media Types [MIME] defining computer media
        formats).</eor:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
</rdf:Property>

<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/identifier">
    <rdfs:label>Resource Identifier</rdfs:label>
    <rdfs:comment>An unambiguous reference to the resource within a given
        context.</rdfs:comment>
    <eor:comment>Recommended best practice is to identify the resource by
        means of a string or number conforming to a formal
        identification system. Example formal identification systems

```

```

        include the Uniform Resource Identifier (URI) (including the
        Uniform Resource Locator (URL)), the Digital Object
        Identifier (DOI) and the International Standard Book Number
        (ISBN).</eor:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
</rdf:Property>

<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/language">
    <rdfs:label>Language</rdfs:label>
    <rdfs:comment>A language of the intellectual content of the
        resource.</rdfs:comment>
    <eor:comment>Recommended best practice for the values of the Language
        element is defined by RFC 1766 [RFC1766] which includes a
        two-letter Language Code (taken from the ISO 639 standard
        [ISO639]), followed optionally, by a two-letter Country Code
        (taken from the ISO 3166 standard [ISO3166]). For example,
        'en' for English, 'fr' for French, or 'en-uk' for English
        used in the United Kingdom.</eor:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
</rdf:Property>

<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/relation">
    <rdfs:label>Relation</rdfs:label>
    <rdfs:comment>A reference to a related resource.</rdfs:comment>
    <eor:comment>Recommended best practice is to reference the resource by
        means of a string or number conforming to a formal
        identification system.</eor:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
</rdf:Property>

<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/source">
    <rdfs:label>Source</rdfs:label>
    <rdfs:comment>A Reference to a resource from which the present resource
        is derived.</rdfs:comment>
    <eor:comment>The present resource may be derived from the Source resource
        in whole or in part. Recommended best practice is to
        reference the resource by means of a string or number
        conforming to a formal identification system.</eor:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
</rdf:Property>

<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/coverage">
    <rdfs:label>Coverage</rdfs:label>
    <rdfs:comment>The extent or scope of the content of the
        resource.</rdfs:comment>
    <eor:comment>Coverage will typically include spatial location (a place
        name or geographic coordinates), temporal period (a period
        label, date, or date range) or jurisdiction (such as a named
        administrative entity). Recommended best practice is to
        select a value from a controlled vocabulary (for example,
        the Thesaurus of Geographic Names [TGN]) and that, where
        appropriate, named places or time periods be used in
        preference to numeric identifiers such as sets of
        coordinates or date ranges.</eor:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
</rdf:Property>

<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/rights">
    <rdfs:label>Rights</rdfs:label>
    <rdfs:comment>Information about rights held in and over the
        resource.</rdfs:comment>
    <eor:comment>Typically, a Rights element will contain a rights management
        statement for the resource, or reference a service providing

```

```
such information. Rights information often encompasses
Intellectual Property Rights (IPR), Copyright, and various
Property Rights. If the Rights element is absent, no
assumptions can be made about the status of these and other
rights with respect to the resource.</eor:comment>
<rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
</rdf:Property>

</rdf:RDF>
```

Anhang D Beispiele für RDF-Beschreibungen

Anhang D1 Einfluss der Rohdichte rezyklierter Zuschläge auf die Betondruckfestigkeit (Tab. 6.1)

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:metabim="http://www.b-i-m.de/2001/01/metabim#">
  <rdf:Description about="http://www.b-i-m.de/berichte/d03/d03z0998.htm#4">

    <dc:type>http://www.b-i-m.de/2001/01/metabim#BETON</dc:type>
    <dc:format>image/gif</dc:format>
    <dc:identifizier>http://www.b-i-m.de/berichte/D03/d03z0998.htm#14
    </dc:identifizier>
    <dc:date>1999-10-26</dc:date>
    <dc:title>Einfluss der Rohdichte rezyklierter Zuschläge auf die
      Betondruckfestigkeit
    </dc:title>
    <dc:creator>Dipl.-Ing. Ch. Müller, Dipl.-Ing. U. Wiens</dc:creator>
    <dc:publisher>Rheinisch-Westfälische Technische Hochschule Aachen,
      Institut für Bauforschung (ibac)</dc:publisher>
    <dc:source>http://www.b-i-m.de/berichte/d03/d03z0998.htm</dc:source>
    <dc:relation>http://www.b-i-m.de/berichte/d03/d03z0998.htm#TabelleA1
    </dc:relation>
    <dc:relation>http://www.b-i-m.de/berichte/d03/d03z0998.htm#TabelleA2
    </dc:relation>
    <dc:relation>http://www.b-i-m.de/berichte/d03/d03z0998.htm#TabelleA3
    </dc:relation>
    <dc:relation>http://www.b-i-m.de/berichte/e04/e04z0998.htm#Tabelle1
    </dc:relation>
    <dc:description>Ziel ist die Bewertung der bei der Aufbereitung von
      Bauschutt anfallenden Recyclingzuschläge hinsichtlich
      der Eignung als Betonzuschlag.</dc:description>
  <dc:subject parseType="Resource">

    <metabim:pruefvorschrift parseType="Resource">
      <rdf:type resource=
        "http://www.b-i-m.de/2001/01/metabim#FESTBETONPRUEFUNGEN"/>
      <rdf:li>DIN 1048-5</rdf:li>
      <rdf:li>DIN 4226-3</rdf:li>
    </metabim:pruefvorschrift>

    <metabim:probekoerpergeometrie parseType="Resource">
      <rdf:type resource=
        "http://www.b-i-m.de/2001/01/metabim#FESTBETONPRUEFUNGEN"/>
      <rdf:value>Würfel</rdf:value>
    </metabim:probekoerpergeometrie>

    <metabim:wzwert parseType="Resource">
      <rdf:type resource=
        "http://www.b-i-m.de/2001/01/metabim#REZEPTUR"/>
      <rdf:li>0.33</rdf:li>
      <rdf:li>0.4</rdf:li>
      <rdf:li>0.55</rdf:li>
      <rdf:li>0.6</rdf:li>
      <rdf:li>0.7</rdf:li>
      <rdf:li>0.75</rdf:li>
    </metabim:wzwert>

  </rdf:Description>
</rdf:RDF>
```

```

<metabim:bezeichnung parseType="Resource">
  <rdf:type resource=
    "http://www.b-i-m.de/2001/01/metabim#ZEMENT"/>
  <rdf:li>CEM I 32,5R</rdf:li>
  <rdf:li>CEM I 42,5R</rdf:li>
</metabim:bezeichnung>

<metabim:messzeitpunkt parseType="Resource">
  <rdf:type resource=
    "http://www.b-i-m.de/2001/01/metabim#FESTBETONPRUEFUNGEN"/>
  <metabim:einheit>d</metabim:einheit>
  <rdf:value>28</rdf:value>
</metabim:messzeitpunkt>

<metabim:druckfestigkeit parseType="Resource">
  <rdf:type resource=
    "http://www.b-i-m.de/2001/01/metabim#FESTBETON"/>
  <metabim:einheit>nmm2</metabim:einheit>
  <rdf:li>59.0</rdf:li>
  <rdf:li>57.1</rdf:li>
  <rdf:li>42.6</rdf:li>
  <rdf:li>32.8</rdf:li>
  <rdf:li>37.0</rdf:li>
  <rdf:li>35.1</rdf:li>
  <rdf:li>35.6</rdf:li>
  <rdf:li>29.8</rdf:li>
  <rdf:li>23.3</rdf:li>
  <rdf:li>30.1</rdf:li>
  <rdf:li>30.0</rdf:li>
  <rdf:li>28.2</rdf:li>
  <rdf:li>20.8</rdf:li>
  <rdf:li>20.3</rdf:li>
  <rdf:li>82.5</rdf:li>
  <rdf:li>65.1</rdf:li>
  <rdf:li>64.2</rdf:li>
  <rdf:li>65.6</rdf:li>
  <rdf:li>67.1</rdf:li>
  <rdf:li>57.2</rdf:li>
</metabim:druckfestigkeit>

<metabim:kornrohddichte parseType="Resource">
  <rdf:type resource=
    "http://www.b-i-m.de/2001/01/metabim#ZUSCHLAG"/>
  <metabim:einheit>kgm3</metabim:einheit>
  <rdf:li>2620</rdf:li>
  <rdf:li>2480</rdf:li>
  <rdf:li>2430</rdf:li>
  <rdf:li>2180</rdf:li>
  <rdf:li>2310</rdf:li>
  <rdf:li>2400</rdf:li>
  <rdf:li>2000</rdf:li>
  <rdf:li>2300</rdf:li>
  <rdf:li>2490</rdf:li>
</metabim:kornrohddichte>

</dc:subject>

</rdf:Description>
</rdf:RDF>

```

Anhang D2 Carbonatisierung von Beton mit unterschiedlichen Brechsandanteilen im Zuschlag (Abb. 6.7)

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:metabim="http://www.b-i-m.de/2001/01/metabim#">
  <rdf:Description about="http://www.b-i-m.de/berichte/e02/e02z0998.htm#4">

    <dc:type>http://www.b-i-m.de/2001/01/metabim#BETON</dc:type>
    <dc:format>image/gif</dc:format>
    <dc:identifizier>http://www.b-i-m.de/berichte/e02/e02z0998.htm#4
    </dc:identifizier>
    <dc:date>1999-10-19</dc:date>
    <dc:title>Carbonatisierung von Beton mit rezyklierter Gesteinskörnung
    </dc:title>
    <dc:creator>Dr.-Ing. E. Siebel, Dipl.-Ing. B. Kerkhoff</dc:creator>
    <dc:publisher>Verein Deutscher Zementwerke e.V. (VDI), Forschungs-
      institut der Zementindustrie (FIZ)</dc:publisher>
    <dc:source>http://www.b-i-m.de/berichte/e02/e02z0998.htm</dc:source>
    <dc:relation>http://www.b-i-m.de/berichte/e02/e02z0998.htm#8
    </dc:relation>
    <dc:description>Einfluss von Recyclingzuschlägen aus Altbeton im Sand-
      anteil auf den Carbonatisierungsfortschritt.
    </dc:description>
    <dc:subject parseType="Resource">

      <metabim:carbonatisierung parseType="Resource">
        <rdf:type resource=
          "http://www.b-i-m.de/2001/01/metabim#FESTBETON"/>
        <metabim:einheit>mm</metabim:einheit>
        <metabim:wertebereich parseType="Resource">
          <metabim:untergrenze>0</metabim:untergrenze>
          <metabim:obergrenze>4.1</metabim:obergrenze>
        </metabim:wertebereich>
      </metabim:carbonatisierung>

      <metabim:messzeitpunkt parseType="Resource">
        <rdf:type resource=
          "http://www.b-i-m.de/2001/01/metabim#FESTBETONPRUEFUNGEN"/>
        <metabim:einheit>d</metabim:einheit>
        <metabim:wertebereich parseType="Resource">
          <metabim:untergrenze>0</metabim:untergrenze>
          <metabim:obergrenze>180</metabim:obergrenze>
        </metabim:wertebereich>
      </metabim:messzeitpunkt>

      <metabim:zusammensetzung parseType="Resource">
        <rdf:type resource="http://www.b-i-m.de/2001/01/metabim#ZUSCHLAG"/>
        <metabim:einheit>%</metabim:einheit>
        <rdf:li rdf:value="Brechsand aus B15" metabim:anteil="100"/>
        <rdf:li rdf:value="Brechsand aus B45" metabim:anteil="0"/>
        <rdf:li rdf:value="Brechsand aus B45" metabim:anteil="50"/>
        <rdf:li rdf:value="Brechsand aus B45" metabim:anteil="100"/>
        <rdf:li rdf:value="Rheinkiessand" metabim:anteil="100"/>
      </metabim:zusammensetzung>

    </dc:subject>
  </rdf:Description>
</rdf:RDF>
```

Lebenslauf

Persönliche Angaben

Marcus Schreyer

geboren am 03. Januar 1971 in Backnang

Ausbildung

1977 - 1981 Stöckach Grundschule in Winnenden

1981 - 1990 Georg-Büchner-Gymnasium in Winnenden, Abitur

1990 - 1996 Studium des Bauingenieurwesens an der Universität Stuttgart, Diplom

Berufliche Tätigkeit

1988 - 1996 Nebentätigkeiten als EDV-Berater

1989 - 1995 Verschiedene Baustellen- und Büropraktika

1993 - 1995 Wissenschaftliche Hilfskraft am Institut für Tragwerksentwurf und -konstruktion bei Prof. Dr.-Ing. Drs.h.c. J. Schlaich

1996 - 1997 Wissenschaftlicher Mitarbeiter der Abteilung Innovationsforschung an der FMPA Baden-Württemberg (Otto-Graf-Institut)

1997 - 2002 Wissenschaftlicher Mitarbeiter am Institut für Werkstoffe im Bauwesen der Universität Stuttgart bei Prof. Dr.-Ing. H.-W. Reinhardt