



Semantic Web Technologien und Strategien für Wissensportale – Ein Überblick

Autoren

Lars Zapf, Heinz-Dietrich Wuttke, Karsten Schmidt

Veröffentlichung

Tagungsband zum "8. Workshop "Multimedia für Bildung und Wirtschaft", Technische Universität Ilmenau, ISSN: 1436 – 4492, S.57-64

Semantic Web Technologien und Strategien für Wissensportale – Ein Überblick

Lars Zapf, Heinz-Dietrich Wuttke, Karsten Schmidt

Die Suche nach brauchbaren und nützlichen Informationen stößt im World Wide Web, in Firmennetzwerken und im privaten Bereich durch das rasante Daten- und Informationswachstum immer mehr an ihre Grenzen. Der informationsuchende Computernutzer verbringt inzwischen zu viel Zeit damit, aus dem Meer an Informationen die für ihn relevanten herauszufiltern. Aufbauend auf semantischen Modellen vernetzter Informationsressourcen bietet der Gedanke des Semantic Web neue Lösungswege zur Organisation von Informationen und einer effizienteren Informationsbereitstellung. Resource Description Framework (RDF), XML Topic Maps (XTM) und die Web Ontology Language (OWL) sind die auf XML basierende syntaktische und semantische Basis des Semantic Web. Mit ihnen entstehen eine Vielzahl neuer Anwendungen und Produktsuiten, die es ermöglichen, Wissensmanagement mit Semantic Web Technologie zu betreiben und darauf basierende Informationsportale zu erstellen. Speziell seitens der XML Topic Map Community gibt es mit dem TMRAP-Protokoll Bestrebungen, Seamless Knowledge und den Gedanken Virtueller Portale zu etablieren. Praktisch wurde der Einsatz von Semantic Web Technologien bereits anhand des Bildungsportals Thüringen skizziert und auf seine Anwendbarkeit validiert.

1. Das Semantic Web

„Das Web wird sein volles Potenzial erst dann erreichen, wenn es zu einer Umgebung wird, in der Daten nicht nur von Menschen, sondern auch von automatischen Werkzeugen getauscht und verarbeitet werden können.“ Tim Berners-Lee [BER03]

Semantische Netze sind ein Wissensrepräsentationsformat, bei dem Konzepte, die semantisch miteinander in Beziehung stehen, durch Strukturen repräsentiert werden, die durch sogenannte assoziative Beziehungen miteinander verbunden sind. Nach Bodendorf [BOD03] beruht der Gedanke, Wissen zu vernetzen, auf der Tatsache, dass auch im Gehirn Wissen nicht linear oder in hierarchisch definierten Strukturen, sondern durch Knoten (Konzepte) und Kanten (assoziative Beziehungen) abgebildet wird. Die Verwaltung von Wissen in Netzwerken kommt der Arbeitsweise des menschlichen Gehirns entgegen.

Tim Berners-Lee, der Erfinder des World Wide Web (WWW), griff den Gedanken der netzartigen Wissensrepräsentation mit semantischen Netzen auf und erwähnte 1998 erstmalig die Vision des Semantic Web. Er und das World-Wide-Web-Konsortium (W3C) gelten als die Vordenker des Semantic Web, dem WWW der Zukunft. Ziel der Evolutionsstufe des existierenden Webs ist es auf der einen Seite, die Bedeutung von Informationen im Web maschinenverständlich zu machen. Dabei sollen alle wesentlichen Informationen einer Quelle mit Angaben zu ihrer Bedeutung (Metadaten) versehen werden (vergleichbar mit der Systematik eines Bibliothekskataloges). Auf der anderen Seite sollen verschiedene Ressourcen (Dokumente, Bilder, Menschen, Konzepte, etc.) semantisch miteinander verknüpft und in Beziehung zueinander gestellt werden (istDort, arbeitetFür, istAutorVon, hatZumThema, istAbhängigVon, etc.). Vergleicht man dies mit der heutigen Situation des World Wide Web, stellt man fest, dass Ressourcen durch Links verbunden sind, die keinerlei Semantik implizieren und von Maschinen nicht verarbeitet werden können.

Das Semantic Web ist Ergebnis jahrelanger Forschungstätigkeiten in der Informatik zu den Themen Wissensrepräsentation und Künstliche Intelligenz (KI). Es ist keine Software, kein Produkt und kein Ersatz für das vorhandene WWW, sondern eine Evolutionsstufe. Das Semantic Web soll Rechner in die Lage versetzen, intelligenter mit dem Menschen und miteinander zu kommunizieren. Durch das Semantic Web werden die Nutzer Informationen in einer anderen Art und Weise aufnehmen, nicht mehr hierarchisch wie bisher.

Informationsdarstellung durch semantische Netze im Allgemeinen und das Semantic Web im Speziellen bieten dem User zu jedem gefundenen Begriff eine Reihe themenverwandter Begriffe an. Die Suche nach Informationen zieht sich wie ein roter Faden durch das Netz. Durch die Semantic Web Initiative kann begonnen werden, aus dem gegenwärtigen Web ein ausdrucksvolleres, semantisch reichhaltigeres Web zu entwickeln, die Qualität von Suchanfragen wesentlich zu verbessern und die Basis für eine Reihe neuer noch nicht vorstellbarer Anwendungen zu legen. Zunächst aber ist und bleibt das Semantic Web eine Vision. Semantik in das Web zu bringen, heißt noch nicht, Semantik aus dem Web herausziehen zu können. Höhere Konzepte und Anwendungen oberhalb des Semantic Web sind dafür notwendig. Das Semantic Web kommt dem Mensch hinsichtlich einer intuitiven Suche von Informationen entgegen, wie er sie alltäglich unter Beachtung ständig steigender Datenmengen bewältigen muss.

2. Semantic Web Standards

2.1. Resource Description Framework (RDF)¹

Das Resource Description Framework ist ein Standard des W3C, der den Austausch von Metadaten zwischen verschiedenen Anwendungen erlaubt. RDF ist ein System, das Relationen zwischen Objekten basierend auf gerichteten Graphen ausdrückt. In RDF wird das zu beschreibende Objekt mit einer eindeutigen Adresse, einer URI, identifiziert. Grundbaustein ist ein Subjekt-Prädikat-Objekt-Tripel P(S,O) (Subjekt S hat ein Prädikat (Eigenschaft) P mit dem Wert O). Das Tripel kann als beschriftete, gerichtete Kante zwischen zwei Knoten verstanden werden. Ein Subjekt eines Tripels kann auch die Rolle des Objekts in einem anderen übernehmen. Aufgabe des Resource Description Framework ist es, dieses Tripel in eine XML Beschreibung zu hüllen und die Integrität des Graphen zu erhalten.

Weiterhin definiert das so genannte RDF Schema (RDFS) Eigenschaften und Art von Ressourcen als Mitglieder von Klassenhierarchien. Im Unterschied zum klassischen objektorientierten Typsystem definiert RDFS nicht die Eigenschaften einer Klasse von Ressourcen, sondern beschreibt die Argumentrahmen der Eigenschaften.

2.2 XML Topic Maps (XTM)²

Im Jahr 2001 wurde durch die ISO der XML Topic Maps (XTM) Standard verabschiedet. Die hinter dem „GPS des Web“³ stehende Idee ist es, bestehende Webseiten und Dokumente nicht zu verändern, sondern vielmehr eine externe Sicht auf die Daten zu beschreiben. Topics können Gegenstände, Themen, Personen, Wörter u.a. sein. Jede dieser Topics existiert in einer Topic Map und verweist auf eine entsprechende Ressource. Topic Maps sind semantische Netzwerke, die von den referenzierten Dokumenten getrennt sind. Diese bleiben unangetastet.

Topics (Themen), Associations (Assoziationen) und Occurrences (Vorkommensangaben) sind die Grundbausteine einer jeden Topic Map oder anders gesagt – das TAO des Standards.

Unter einer *Topic* versteht man alles Beschreibbare, abhängig vom konkreten Anwendungsfall.

Occurrences verbinden Topics mit Ressourcen (Dokumente, Bilder, Webseiten, etc.), sofern sie durch eine URL adressierbar sind. Für nicht adressierbare Subjekte muss ein adressierbarer elektronischer Stellvertreter, der

¹ <http://www.w3.org/RDF/> (Seitenabruf: 14.09.04)

² <http://www.topicmaps.org/xtm/> (Seitenabruf: 14.09.04)

³ So lautet der Slogan der XTM-Arbeitsgruppe, geprägt durch Charles Goldfarb

Lars Zapf, Heinz-Dietrich Wuttke, Karsten Schmidt

so genannte *Subjekt-Indikator (eine URI)*, definiert werden. Eine Topic kann beliebig viele Occurrences aufweisen. Jede Occurrence kann eine bestimmte Rolle, die so genannte *Occurrence Role*, einnehmen. Diese Rolle weist der Occurrence eine bestimmte Semantik zu.

Associations schließlich beschreiben semantische Beziehungen zwischen den Topics. Eine Assoziation kann beliebig viele Themen in Beziehung zueinander stellen. Wie bei den Occurrences kann definiert werden, welche Rolle die Themen in der Beziehung übernehmen.

Durch das Scope-Konzept ist es möglich, alle Charakteristika einer Topic mit einem Gültigkeitsbereich zu versehen. Die Auswertung dieses Gültigkeitsbereiches bleibt der jeweiligen Anwendung überlassen. So sind beispielsweise die Verwendung mehrerer Sprachen, das Erstellen mehrerer Datensichten oder die Steuerung von Zugriffsrechten auf verschiedene Ressourcen umsetzbar. Der Standard ermöglicht ebenfalls das Merging, also das wohldefinierte Zusammenführen verschiedener Topic Maps zu einer, beziehungsweise verschiedener Topics zu einer Topic.

2.3. Web Ontology Language (OWL)⁴

Ontologien sind Definitionen von Objekten und deren Beziehungen bezogen auf einen Anwendungsbereich. Sie sind entscheidend für Anwendungen, die Informationen unterschiedlicher Herkunft suchen oder mischen müssen. Eine wichtige Anwendung von Ontologien ist, so Bodendorf [BOD03], die semantisch konsistente Kommunikation zwischen auf Wissensebene kooperierenden intelligenten Systemen zu ermöglichen. Gemeinsame Ontologien stellen sicher, dass eine Information vom Empfänger ebenso interpretiert wird, wie sie vom Absender gemeint war.

Die Web Ontology Language (OWL) ist eine auf XML basierende Sprache zur Beschreibung von Ontologien und Beziehungen zwischen Objekten. Ontologien können zum Beispiel durch Inferenzmaschinen oder intelligente Suchmaschinen interpretiert werden. Sie sagen auf einer höheren Ebenen, wie die Dinge auf der niedrigeren begrifflichen Ebene zusammenhängen, um die Semantik aus der Wissensbasis automatisiert für den Benutzer herauszuziehen.

2.4. Zusammenfassung

Hauptanwendungsgebiete für Semantic Web Standards und -Sprachen sind die Generierung von Wissensbasen, die Klassifikation und Strukturierung von Inhalten, das Ermöglichen intelligenter Navigation und Suche in semantischen Netzwerken, das Erstellen verschiedener Sichten auf gleiche Ressourcen, Informationsfilterung, Benutzerprofile und Rechte, sowie das Generieren von „virtuellen Informationen“ aus bereits bestehenden und die Wissensaggregation.

Die Möglichkeit der Suche nach Assoziationen und ihren Eigenschaften fördert auch die Entwicklung so genannter „intelligenter Suchmaschinen“, also semantisch gesteuerter Suchprozesse. Semantic Web Standards werden in den Bereichen Content-, Dokument- und Knowledge Management bereits praktisch eingesetzt. So arbeitet z.B. die Suchfunktion vom Bertelsmann Webportal *www.wissen.de* vollständig auf der Topic Map Ebene. Eine ausführlichere Betrachtung, Beispiele und Vergleiche findet man unter [ZAP04].

⁴ <http://www.w3.org/TR/owl-features/> (Seitenabruf: 14.09.04)

3. Semantic Web Tools und Produktsuiten

Ohne effiziente Software (Parser, API's, Frameworks, Editoren, Speicher- und Verarbeitungssysteme, etc.) ist die Verwendung von Semantic Web Standards relativ nutzlos. Ausgewählte Semantic Web Tools und Produktsuiten sollen nachfolgend kurz betrachtet werden.

Intelligent views hat die Plattform *K-Infinity* zum Aufbau und zur Nutzung von Wissensnetzen entwickelt. Sie bildet die Basis für intelligentes Wissensmanagement. K-Infinity ist modular aufgebaut und lässt sich einfach in bestehende Datenbank- und IT-Architekturen integrieren. Zum Aufbau und zur Pflege einer Wissensbasis dienen Editing-Tools (Knowledge-Builder), zur Ausnutzung des vernetzten Wissens die Usage-Tools (browserbasierend). Der Knowledge-Builder stellt sämtliche Operationen zur Verfügung, die zum Aufbau eines Wissensnetzes nötig sind. Mit Hilfe des Markup-Tools können Dokumente erstellt, erfasst und mit dem Wissensnetz verknüpft werden. Der Knowledge-Accelerator ermöglicht mit einem Web-Browser ohne Installationsaufwand Zugriff auf das Wissensnetz. Der Semantic-Finder ermöglicht eine semantische Suche. Mit dem Net-Navigator kann grafisch durch das semantische Netz navigiert werden. Die Ergebnisse von Anfragen werden durch die Layout-Engine automatisch in HTML-Seiten umgesetzt und individuell an ihr Design angepasst. Sämtliche K-Infinity-Komponenten lassen sich durch XML-Schnittstellen mit anderen Werkzeugen verbinden. Triggerkonzept und Workflowunterstützung sind weitere Features der Produktsuite. (<http://www.i-views.de/>)

Die norwegische Firma **Ontopia** hat bereits mehrjährige Erfahrung in der Entwicklung von Topic Map Software und ist in Fachkreisen bekannt für ihre Ontopia Knowledge Suite (OKS). Die OKS ist eine komplette Toolsammlung zum Entwerfen, Administrieren und Entwickeln Topic Map basierter Applikationen. Sie beinhaltet eine Topic Map Engine, die Topic Maps lädt, verarbeitet, speichert und eine Schnittstelle bietet, die andere Anwendungen auf Topic Maps zugreifen lässt. Weiterhin bietet sie webbasierte Editoren und einen frei erhältlichen Topic Map Browser (Omnigator), sowie eine eigene Abfragesprache. In seiner neuesten Version erlaubt der Omnigator das Browsen in RDF- und XTM-Dateien, sowie deren graphische Visualisierung. (<http://www.ontopia.net/>)

Als Teil der e:kms, der XML-basierten Enterprise Plattform für das Verwalten und Organisieren von Wissen und Inhalten, bietet **empolis** das Produkt e:km (empolis knowledge manager), den Nachfolger der Produktlinien k42 und orange, an. E:kms ist eine gesamtheitliche Knowledge Management Lösung mit Dokument- und Content Management, Workflow, Ontology Management, graphischer Visualisierung, intelligenter Suche basierend auf Ontologien, Intranet, Personalisierung und Community Support. (<http://www.empolis.de/>)

Der **USU KnowledgeMiner** extrahiert Metadaten aus bestehenden Datenquellen und vernetzt diese semantisch gemäß dem Topic Map Standard. Durch einen Fuzzy-Filter ist das System äußerst tolerant gegenüber Schreibfehlern. Der KnowledgeMiner dient der Themenstrukturierung und dem zentralen Informationszugriff, stellt die vernetzten Strukturen grafisch dar und bietet dem Anwender Navigations- und Suchfunktionen. Ebenfalls werden Topic Maps als flexibles Kategoriensystem dargestellt, so dass Themen und Begriffe im Zusammenhang aufgezeigt werden und der kategorienorientierte Anwender schnell zu den benötigten Informationen findet. Über Natural Language Processing stehen ihm erweiterte Suchmöglichkeiten zur Verfügung. Der User wird im Rahmen seiner Zugriffsberechtigung zu den erwünschten Informationen geleitet. Sensible Themen und Dokumente bleiben geschützt. (<http://www.usu.de/>)

Jena2 ist ein von Hewlett Packard entwickeltes Open Source Framework zum Erstellen von RDF basierten Semantic Web Applikationen. Es beinhaltet eine RDF-API zum Manipulieren von RDF-Datensätzen, unterstützt RDFS und OWL, bietet einen integrierten Java-basierten RDF-XML-Parser von ARP und eine Inferenzkomponente. Weiterhin erlaubt das Framework das persistente Speichern von Daten in einem relationalen Datenbanksystem (MySQL PostgreSQL, Oracle), das Hinzufügen von Daten und das Stellen von Abfragen an dieses System. (<http://www.hpl.hp.com/semweb/jena.htm>)

Das Ziel des **TM4J**-Projektes ist es, robuste Open Source Tools zum Erstellen, Manipulieren und Publizieren von Topic Maps anzubieten. Es besteht aus einer Topic Map Engine und einer Java-Applikation zum Browsen von Topic Maps mittels eines grafischen Nutzerinterfaces. (<http://tm4j.org/>)

Lars Zapf, Heinz-Dietrich Wuttke, Karsten Schmidt

Ohne die Verwendung spezieller Tools ist der professionelle und effiziente Einsatz von Semantic Web Technologie nicht möglich. Sie sind Basis für die Durchsetzung und den Erfolg der neuen Technologie. Alle hier vorgestellten Produkte eignen sich, das Wissensmanagement einer Institution wesentlich zu verbessern, einen schnelleren und qualitativ besseren Zugriff auf Informationen zu bieten, schnelleres Erstellen und Pflegen intelligenter Web Portale, wissensbasierter Intranets und Content Management Systeme zu ermöglichen. Eine ausführlichere Betrachtung, Beispiele und Vergleiche findet man unter [ZAP04].

4. Semantic Web basierende Informationsportale

Auf Semantic Web Technologie basierende Informationsportale bieten dem Nutzer eine Vielzahl an Möglichkeiten, die über diejenigen konventioneller Portale hinausgehen. Für das Bildungportal Thüringen wurde in [ZAP04] der Einsatz von Semantic Web Tools und Produktsuiten evaluiert. Nachfolgend sollen einige Anforderungen an Semantic Web basierende Informationsportale beschrieben werden, die sich aus den Ergebnissen der Evaluation herauskristallisiert haben.

Zu jeder gefundenen Information sollte der User die Möglichkeit besitzen, themenverwandte Inhalte weiterzuverfolgen und zugehörige Metainformationen zu erhalten. Sein Weg durch die Webseite gestaltet sich nicht hierarchisch, sondern individuell entlang eines Themenpfades. Die Navigation sollte einerseits über Links geschehen und andererseits alternativ über eine grafische Visualisierung des gesamten Themenbereiches erfolgen, ähnlich einem Hyperbolic-Tree-Ansatz. Weiterhin sollte automatisch und dynamisch ein Themenkatalog entsprechend der hinterlegten Taxonomie oder Ontologie generiert werden. Dem Nutzer sollten weiterhin intelligente semantische Such- und Abfragefunktion, wie z.B. Filtersuche, Suche über Metadaten, Volltextsuche, Thematische Suche, Semantische Query oder Suche in Synonymen zur Verfügung stehen. Ein Kontext-Filter dient zum Ausblenden von Inhalten, die z.B. mit einem Scope versehen sind oder vom Benutzer gewählt wurden (z.B. „Blende alle Topics zum Thema „TU Ilmenau“ aus“). Von der Benutzerverwaltung erhält jeder User ein individuelles personalisiertes Erscheinungsbild je nach zugehöriger Benutzergruppe (z.B. Hochschullehrer, Studenten, Private Bildungseinrichtungen, etc.). Suchergebnisse sollten mit zusätzlichen Informationen angereichert sein, so z.B. über die semantische Nähe zum Suchbegriff und die Häufigkeit der Aufrufe in den letzten Tagen. Weiterhin sollte ein Triggerkonzept implementiert sein, welches das Auslösen von Ereignissen beim Eintreten gewisser Ausgangsbedingungen ermöglicht. Schnittstellen zu bisherigen Systemen, sowie Im- und Exportfunktionen auf Basis von XML ermöglichen eine nahtlose Integration in andere Architekturen.

Abbildung 1 gibt einen ersten Eindruck der Visualisierung eines Hyperbolic-Tree zur grafischen Navigation im Informationsraum und zeigt darüberhinaus auf der linken Seite eine automatisch generierte Kategorieansicht, sowie im unteren Bereich die semantische Nähe und themenverwandte Bereiche zum aktuell gefundenen Suchbegriff.



Abbildung 1: Semantic BPT (Prototypischer Entwurf)

5. Seamless Knowledge und Virtuelle Portale

Seamless Knowledge, eine weitere Evolutionsstufe des Wissensmanagements für Informationsportale, bezeichnet den für den Benutzer nahtlosen Übergang von Wissen. Praktisches Anwendungsbeispiel für Seamless Knowledge sind Virtuelle Portale, d.h. mehrere physisch vorhandene (Wissens-)Portale, die zu einem großen virtuellen vernetzt werden. Bei der Umsetzung kann vor allem Semantic Web Technologie hilfreich sein. Gute Vorarbeiten auf diesem Gebiet stammen von Steve Pepper, dem Chief Strategy Officer der norwegischen Topic Map Firma Ontopia, auf dessen Ideen sich nachfolgend bezogen werden soll.

In der Praxis stehen die in Wissensportalen abgespeicherten Informations- und Wissensbestände oftmals getrennt voneinander zur Verfügung. Die Existenz des jeweils Anderen ist dem Nutzer nicht bekannt und kann folglich nicht abgerufen werden. Für den Nutzer besteht der Wunsch nach Knowledge Integration (sog. Seamless Knowledge). Seamless Knowledge ist jedoch nicht mit dem Begriff Semantic Web zu verwechseln, obwohl beide viele Gemeinsamkeiten haben. Das Semantic Web stellt lediglich eine mögliche Basistechnologie für eine nahtlose Informationsintegration zur Verfügung.

Norwegen ist Vorreiter, wenn es um auf Topic Map Technologie basierende Portale (sog. Topic Map driven portals) geht. Mit einer steigenden Anzahl von Informationsportalen wachsen natürlich auch die Überlappungen zwischen ihren Inhalten. Die drei Portale *forskning.no* (Research Council), *forbrukerportalen.no* (Norwegian Consumer Association) und *matportalen.no* (Biosecurity portal of the Department of Agriculture) sind die ersten, die unter dem Gesichtspunkt Virtueller Portale miteinander vernetzt wurden. Gleiche Inhalte werden in einer Datenquelle gespeichert und sind in allen drei Portalen sichtbar.

Aus technischer Sicht ist wenig nötig, um eine effektive Form nahtloser Wissensintegration zu erreichen. Zum einen ist ein Identifizierungsmechanismus notwendig, um die Identität eines so genannten *subjects* (kleinste modulare Informationseinheit) zu verifizieren. Portale, die miteinander kommunizieren, können somit sicher gehen, vom gleichen subject zu sprechen. Probleme, die durch die Verwendung von Synonymen, Homonymen,

Lars Zapf, Heinz-Dietrich Wuttke, Karsten Schmidt

etc. entstehen, werden damit beseitigt. Ein eindeutiger globaler Mechanismus zum Bezeichnen von subjects existiert zum Beispiel schon im Topic Map Standard und heißt dort Published Subjects. Er basiert technisch gesehen auf dem URI-Konzept. Published Subjects helfen nicht nur Maschinen, sich gegenseitig besser zu verstehen, sondern auch Menschen. Sie identifizieren ein subject via URI (sog. subject identifier). Diese URI ist gleichzeitig Adresse eines Dokumentes, das von einem User interpretiert werden kann (sog. subject indicator).

Zum anderen ist für die Umsetzung einer nahtlosen Wissensintegration ein Austauschprotokoll notwendig, um Information automatisch miteinander zu teilen. Hierfür gibt es seitens der Topic Map Community den Entwurf des TMRAP (Topic Maps Remote Access Protocol)⁵. Dies ist ein abstraktes Protokoll, um Informationen von Remote Repositories zu erhalten. Topic Map Applikationen können TMRAP sehr einfach unterstützen, für nicht-Topic Map Applikationen ist es ein nur relativ kleiner Anpassungsaufwand notwendig, um TMRAP verwenden zu können. Der Nutzen wäre enorm. Ähnliche Arbeiten findet man in der Semantic Web Gemeinde auch in der RDF Net API⁶ und bei SNAPI⁷.

Ontopia hat in seinem frei erhältlichen Topic Map Browser bereits eine erste Demonstrationsvariante des TMRAP eingebaut, die sogenannte *Omnigator Rap demo*. Diese basiert auf zwei räumlich getrennten Omnigator Browsern, auf denen jeweils unterschiedliche Topic Maps das vorhandene Wissen managen. Schaltet man die Unterstützung für TMRAP ein, werden in der jeweiligen Topic Map automatisch die Links zur Remote Topic Map des Anderen eingefügt. Über einen VISIT-Button kann man in die andere Map hineinbrowsen (vgl. Abbildung 2).



Abbildung 2: Omnigator Rap demo (<http://www.ontopia.net/> (Seitenabruf: 14.09.04))

Die hinter dem Topic Maps Remote Access Protocol stehende Technik mag auf den ersten Blick sehr einfach erscheinen. Ihr Potenzial jedoch ist sehr hoch, da auch andere Applikationen, die nicht den Topic Map Standard benutzen, sich des TMRAP bedienen können. Ontopia geht aber noch einen Schritt weiter, indem nicht nur

⁵ <http://www.jtc1sc34.org/repository/0507.htm> (Seitenabruf: 14.09.04)

⁶ <http://www.w3.org/Submission/2003/SUBM-rdf-netapi-20031002/> (Seitenabruf: 14.09.04)

⁷ <http://sourceforge.net/projects/snapi/> (Seitenabruf: 14.09.04)

einzelne Topic Page URI's ausgetauscht werden können, sondern auch Fragmente dieser, sog. Topic Maplets⁸. Dies können z.B. zusätzliche Metadaten, URL's, assoziierte Seiten mit Sichten und Rollen, etc. sein. Mitte 2005 wird die Topic Map Query Language (TMQL⁹) für Topic Maps durch die ISO standardisiert werden. Mit ihr wird es möglich sein, diese Topic Maplets als remote query results zu erhalten. Weiterhin ist seitens der Topic Map Community eine SOAP Anbindung für TMRAP geplant, um den Einsatz von Web Services zu unterstützen. Die dabei entstehenden Möglichkeiten wären immens.

Zusammenfassend kann man sagen, dass um Seamless Knowledge für Informationsportale erreichen zu wollen, folgenden Grundbausteine benötigt werden:

- Semantisch strukturierte Daten (Topic Maps, RDF, etc.)
- Global eindeutige Bezeichner (Published Subjects, etc.)
- Ein Austauschprotokoll (z.B. TMRAP)
- Eine Abfragesprache (z.B. TMQL) für präzisere Request-Anfragen

6. Zusammenfassung und Ausblick

Semantic Web Standards werden aufgrund der steigenden Anzahl praktischer Anwendungen und Produktsuiten populär. Vorreiter auf dem Gebiet der Topic Maps ist Norwegen. Hier gibt es sehr viele Webseiten und Portale, eLearning- und Wissensmanagementanwendungen, die sich bereits der Semantic Web Standards bedienen. Seamless Knowledge ist der nächste, auf dieser Technologie aufbauende Evolutionsschritt. Das Semantic Web wird für seine Durchsetzung und Akzeptanz sicher noch einige Jahre brauchen, Seamless Knowledge gibt es bereits heute.

	2003 in Tausend	% der Bevölkerung	2007 in Tausend	% der Bevölkerung
Deutschland	42.003	51,0	59.033	71,8

Tabelle 1: Webbenutzer 2003 und 2007 [EIT04]

Auch am praktischen Beispiel, konkret für das Bildungsportal Thüringen, ist der Einsatz von Semantic Web Technologie mittelfristig eine Alternative. Tabelle 1 verdeutlicht die zunehmende Anzahl von Webbenutzern bis zum Jahr 2007. Das damit verbundene zunehmende Daten- und Informationsaufkommen dürfte immens sein und exponentiell ansteigen. Neue Wege, Informationen zu organisieren, zu kategorisieren und dem Benutzer zu präsentieren werden benötigt, da sonst der Informationskollaps droht. Die in diesem Paper vorgeschlagenen Ansätze sind ein erster Schritt in die richtige Richtung.

Literatur

[BER03]

http://www.ercim.org/publication/Ercim_News/enw51/berners-lee.html (Seitenabruf: 14.09.04)

[BOD03]

Bodendorf, Freimut: Daten- und Wissensmanagement, Springer, 2003

[EIT04]

European Information Technology Observatory 2004, <http://www.eito.com/> (Seitenabruf: 14.09.04)

[ZAP04]

⁸ <http://www.ontopia.net/topicmaps/materials/xm-fragments.html> (Seitenabruf: 14.09.04)

⁹ <http://www.isotopicmaps.org/tmq/> (Seitenabruf: 14.09.04)

Lars Zapf, Heinz-Dietrich Wuttke, Karsten Schmidt

Zapf, Lars: Intelligentes Information Retrieval und Knowledge Management mit semantischen Netzen, Diplomarbeit an der Technischen Universität Ilmenau, 2004

Autorenangaben

Dipl.-Inf. Lars Zapf

Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH, Abteilung Wissensmanagement, Erwin-Schrödinger-Str. Geb. 57

D-67653 Kaiserslautern

E-Mail: zapf@dfki.uni-kl.de

www.dfki.de

Dr.-Ing. Heinz-Dietrich Wuttke

Technische Universität Ilmenau, Helmholtzplatz 1

98693 Ilmenau

E-Mail: dieter.wuttke@theoinf.tu-ilmenau.de

<http://www-ihs.theoinf.tu-ilmenau.de/forschung/projekte/sane/>

Dipl.-Kfm. Karsten Schmidt

bildungsportal-thueringen.de, Puschkinstraße 19

99084 Erfurt

E-Mail: Ka.schmidt@uni-jena.de

www.bildungsportal-thueringen.de

10. - 13. April 2005 in Kaiserslautern

3rd Conference Professional Knowledge Management

<http://wm-konferenz.de>
