

WIRTSCHAFTSUNIVERSITÄT WIEN

DIPLOMARBEIT

Titel der Diplomarbeit:

Semantic Web in der Praxis

Nutzungspotentiale und Einsatzmöglichkeiten von Semantic-Web-Technologien am Beispiel der Musikbranche

Verfasserin/Verfasser: Isabella Hidalgo

Matrikel-Nr.: 0353788

Studienrichtung: Internationale Betriebswirtschaft

Beurteilerin/Beurteiler: Prof.Dr. Wolfgang Panny

Ich versichere:

dass ich die Diplomarbeit selbstständig verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubten Hilfe bedient habe.

dass ich dieses Diplomarbeitsthema bisher weder im In- noch im Ausland (einer Beurteilerin/ einem Beurteiler zur Begutachtung) in irgendeiner Form als Prüfungsarbeit vorgelegt habe.

dass diese Arbeit mit der vom Begutachter beurteilten Arbeit übereinstimmt.

Datum

Unterschrift

Wirtschaftsuniversität Wien



Institut für Informationswirtschaft

Diplomarbeit

„Semantic Web in der Praxis“

Nutzungspotentiale und Einsatzmöglichkeiten von Semantic-Web-Technologien am Beispiel der Musikbranche

Betreuender Assistent: Dipl.-Ing. Mag. Dr. Albert Weichselbraun

Verfasser: Isabella Hidalgo

Matrikel-Nr.: 0353788

Wien, 21. Oktober 2008

Inhaltsverzeichnis

Inhaltsverzeichnis	1
Abbildungs- und Tabellenverzeichnis	4
Abkürzungsverzeichnis	5
Einleitung	6
1 Semantic Web	8
1.1 Semantic Web - Definition und Ziele	8
1.2 Semantic-Web-Technologien	9
1.2.1 Resource Description Framework (RDF)	10
1.2.2 RDF-Schema (RDFS)	14
1.2.3 Ontologien und Web Ontology Language (OWL)	14
1.3 Abgrenzung des Begriffs Semantic Web	17
2 Nutzenpotentiale des Semantic Web	18
2.1 Anwendungen der Semantic-Web-Technologie im E-Commerce	18
2.1.1 Semantische Empfehlungssysteme	19
2.1.2 Mobile Dienste	22
2.1.3 Werbung	25
2.2 Anwendungen der Semantic-Web-Technologie im Wissensmanagement	29
2.2.1 Textbasierte Suchmaschinen	31
2.2.2 Wikis	35
2.2.3 Social Software	39
2.2.3.1 Social Networks	40
2.2.3.2 Blogs	42
2.2.3.3 Social Bookmarking	44
2.3 SWOT-Analyse des Semantic Web	46
2.3.1 Stärken	46
2.3.2 Schwächen	50
2.3.3 Chancen	53
2.3.4 Gefahren	56
2.4 Bewertung der Nutzenpotentiale und Ausblick anhand von Experten	59
2.4.1 Fragebogen	60
2.4.2 Ergebnisse	61

3	Fallstudie mit Universal Music Deutschland	64
3.1	Universal Music Deutschland	64
3.2	Die Newsletter	65
3.3	Problemstellung.....	67
3.3.1	Hohe Informationsdichte	68
3.3.2	Unterschiedliche Bedürfnisse der Zielgruppen.....	69
3.3.3	Problem: Relevanz der Inhalte	70
3.4	Lösung: Einsatz von Semantic-Web-Technologien zur Personalisierung der Inhalte der Genre-Newsletter	71
3.4.1	Aufbau der semantischen Infrastruktur	72
3.4.1.1	Entwicklung oder Verwendung von bestehenden Ontologien	72
3.4.1.2	Semantische Annotation der Daten	76
3.4.2	Erstellung von Benutzerprofilen	83
3.4.3	Anwendungen für die Newsletter von Universal Music	86
3.4.3.1	Exkurs: Bestehende Online-Musikdienste	87
3.4.3.2	Semantische Empfehlungssysteme.....	93
3.4.3.3	Applikationen zur Personalisierung der Inhalte der Newsletter.....	95
3.4.3.4	Vorteile für Universal Music	100
3.4.3.5	Ausblick	102
4	Fazit.....	104
5	Literaturverzeichnis	106
6	Anhang.....	120
6.1	Ontologien von Music Ontology.....	120
6.1.1	Event-Ontologie	120
6.1.2	Zeitleisten-Ontologie	121
6.1.3	Ontologie Musikalischer Ausdruck und Event-Schema.....	122
6.1.4	Musikalisches-Werk-Schema	122
6.1.5	Veröffentlichungs-Schema	123
6.1.6	Musikobjekte-Schema.....	124
6.1.7	Musikünstler-Ontologie, Musikgruppen-Ontologie	124
6.1.8	Veröffentlichungsstatus-Ontologie	124
6.1.9	Veröffentlichungsart-Ontologie.....	125
6.1.10	Gestaltungs-Ontologie	121
6.2	Interviews	125

6.2.1	Interview mit Felix Schramm	125
6.2.2	Interview mit Mario Szpuszta	129
6.2.3	Interview mit Christian Boser.....	134
6.2.4	Interview mit Thomas Hochhaltinger	138
6.2.5	Interview mit Markus Linder	141

Abbildungs- und Tabellenverzeichnis

Abb. 1:	RDF-Triples	11
Abb. 2:	RDF-Triples mit mehreren Daten.....	11
Abb. 3:	Dateneingabe und -ergebnis in der Applikation RDFmaker	13
Abb. 4:	Gastronomie-Ontologie.....	16
Abb. 5:	Screenshot des digitalen Empfehlungssystems Esolda.....	21
Abb. 6:	Mobile Anfragen durch Kombination von Handykamera und Spracheingabe.....	24
Abb. 7:	Navigationsdialog durch die Kombination von mehreren Webdiensten ...	25
Abb. 8:	Kontextbasierte Werbe-Initiative von ThoughtTrail und MSN	28
Abb. 9:	Kontextbasierte Brillenwerbung in Facebook.....	29
Abb. 10:	Stichwortbasierte Suche versus Fragebeantwortung.....	33
Abb. 11:	Zugriff auf Informationen durch den semantischen Mediator	34
Tab. 1:	Beispielfragen und -antworten von AskWiki.....	39
Abb. 12:	Verlinkungen bei der Wikipedia	38
Abb. 13:	Verlinkungen bei der Semantic Wiki	38
Abb. 14:	Darstellung der semantischen Beziehungen beim ExpertFinder-System	41
Abb. 15:	Screenshot eines semantischen Blogeintrags	43
Abb. 16:	Screenshot der Gruppe Jazz Legends bei der GroupMe!-Seite	45
Abb. 17:	Eintrittsbarrieren des Semantic Webs	51
Abb. 18:	Analyse der Inhalte der Genre-Newsletter	67
Abb. 19:	Sinus-Milieus in Deutschland 2007	69
Abb. 20:	Ontologie mit redaktioneller Information von Music-Ontology.....	74
Abb. 21:	Ausschnitt aus einer Ontologie über Radioprogramme	75
Abb. 22:	Grafische Repräsentation der MusicBrainz!-RDF-Daten	78
Abb. 23:	Metadaten-Architektur und Retrieval Architektur	81
Abb. 24:	Screenshot von der Musicoverly-Applikation.....	90
Abb. 25:	Architektur des Empfehlungssystems Strands Recommender	92
Abb. 26:	Screenshot vom Idiomag-Tool, um Lieder zu bewerten.....	97
Abb. 27:	Screenshot aus dem Musicoverly-Tool.....	98
Abb. 28:	Beispielprofil der Community von Jimi Blue	100

Abkürzungsverzeichnis

API	Application Programming Interface
APML	Attention Profiling Markup Language
B2B	Business to Business
B2C	Business to Consumer
DAML+OIL	DARPA Agent Markup Language + Ontology Inference Layer
E-Commerce	Electronic Commerce
HTML	Hypertext Markup Language
M-Commerce	Mobiles Commerce
MDA	Mobile Digital Assistant
RDF	Resource Description Framework
RDFa	Resource Description Framework attribu- tes
RDFS	RDF-Schema
SWEO	Semantic Web Education and Outreach
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
VoIP	Voice over Internet Protocol
W3C	World Wide Web Consortium
WWW	World Wide Web
XML	Extensible Markup Language

Einleitung

Das Semantic Web ist eine Vision von Tim Berners-Lee, dem Schöpfer des World Wide Web (WWW). Aufgrund des Erfolges des WWW wurden unzählige Informationen, Angebote und Daten unkontrolliert im Web publiziert. Laut einer Studie aus dem Jahre 2005 [Gulli und Alessio, 2005] gibt es mehr als 11,5 Milliarden indizierbare Webseiten. Durch diese Datenflut ist es für die Internetbenutzer zunehmend schwierig, die richtigen Informationen und Angebote zu finden.

Die große Herausforderung ist derzeit, diese riesige Datenmenge semantisch anzureichern, damit sie von Maschinen interpretiert werden kann. Dadurch können intelligentere und effizientere Applikationen geschaffen werden, die die Internetnutzer bei der Abwicklung ihrer täglichen Internet-Aktivitäten unterstützen. Zum Beispiel suchen die heutigen Suchmaschinen wie Google oder Yahoo „nur“ nach Zeichenketten und einzelnen Wörtern. Die Suchergebnisse enthalten oft eine sehr lange Liste an Links mit vielen irrelevanten Inhalten. Das Semantic Web ermöglicht, das Wissen im Web besser zu ordnen, und somit zum Beispiel gezieltere Ergebnisse auf Suchanfragen zu liefern [Berners-Lee, 1998]. Das ist aber nur eine der vielen Anwendungsmöglichkeiten dieser Technologie. Weitere Anwendungsgebiete findet man im E-Commerce-Bereich, Wissensmanagement oder in Web 2.0-Applikationen. Diese werden in dieser Arbeit näher beschrieben.

Im Gegensatz zum Web 2.0 hat das Semantic Web noch keine starke Verbreitung in den Medien gewonnen und ist der Masse noch ziemlich unbekannt. Seit Jahren wird viel im Semantic-Web-Bereich geforscht, aber in den meisten Publikationen wird dieses Thema hauptsächlich von der technischen Seite betrachtet. Damit das Semantic Web sich durchsetzt und eine breite Anwendbarkeit findet, müssen die Internetbenutzer dieses Konzept verstehen und den Mehrwert wahrnehmen, den diese Technologien ihnen bieten.

Vor diesem Hintergrund liegt das primäre Ziel dieser Arbeit darin, die Anwendungspotentiale und den Praxisbezug der semantischen Technologien für jeden Menschen verständlich zu machen und den Mehrwert dieser Technologie für Benutzer und Unternehmen darzustellen. Darauf aufbauend wird eine Fallstudie erstellt, die demonstriert, wie der Prozess des Einsatzes von Semantic-Web-Technologien abläuft

und was das Semantic Web für Möglichkeiten zur Personalisierung der Inhalte bietet. Für Unternehmen in der Musikindustrie ist die Fallstudie besonders wertvoll, weil bestehende Ontologien und Initiativen zur Gewinnung von Metadaten erklärt werden, die für diese Domäne bereits entwickelt wurden.

Die vorliegende Arbeit ist wie folgt strukturiert:

Kapitel eins gibt einen Überblick über den Begriff Semantic Web und die damit verbundenen Technologien. Darüber hinaus wird in diesem Kapitel das Themengebiet abgegrenzt.

Im Kapitel zwei stehen die Nutzungspotentiale der Semantic-Web-Technologie im Vordergrund. Zuerst werden die Anwendungen im E-Commerce erläutert, wie zum Beispiel semantische Empfehlungssysteme, mobile Services oder intelligente Werbung. Danach werden die Anwendungen im Wissensmanagement dargestellt. Anhand von Beispielen für Suchmaschinen, semantische Wikis und Social Software, werden diese näher erläutert. Anschließend werden in einer SWOT-Analyse die Stärken und Schwächen, Chancen und Risiken des Semantic Web präsentiert. Darüber hinaus wird ein Ausblick gegeben, in welchen Branchen das Semantic Web Zukunft hat und welche Anwendungsgebiete laut Experten am meisten Sinn machen.

Im letzten Kapitel wird eine Fallstudie für Universal Music Deutschland dargestellt. Dieses Projekt befasst sich mit dem Einsatz von semantischen Technologien im E-Commerce-Bereich und fokussiert sich auf die Personalisierung und Optimierung der Newsletter von Universal Music Deutschland. In diesem Kapitel wird zuerst ein Überblick über das Unternehmen und die Problemstellung gegeben. Anschließend werden Handlungsempfehlungen präsentiert.

Das Fazit gibt eine Zusammenfassung über die behandelten Themen und ein Ausblick über die Entwicklung und Zukunft des Semantic Web.

1 Semantic Web

Im ersten Kapitel wird der Begriff Semantic Web definiert und die damit verbundene Technologie erklärt. Danach wird das Themengebiet abgegrenzt und die die Anwendungsgebiete des Semantic Web werden in E-Commerce und Wissensmanagement aufgeteilt.

1.1 Semantic Web - Definition und Ziele

Tim Berners-Lee, Pionier des World Wide Web (WWW) und Direktor des World Wide Web Consortiums, definiert das Semantic Web folgendermaßen:

„The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation.“ [Berners-Lee et al., 2001]

Im Semantic Web werden die Informationen im World Wide Web maschinenlesbar und maschinenverarbeitbar so dargestellt, dass der Computer in gewisser Hinsicht die „Bedeutung“ der Wörter versteht und den Sinnzusammenhang zwischen verschiedenen Quellen erkennt. Durch die Bedeutungszuweisung profitieren einerseits die Maschinen, weil sie die Inhalte im Internet besser verstehen, und somit schlussendlich die Benutzer, weil sie durch verschiedene semantische Applikationen im Web unterstützt werden können [Henze, 2006, S. 140].

Durch die Trennung der Inhalte und deren Bedeutung kann man Inhalte neu zusammenstellen und in verschiedenen Kontexten, Anwendungsszenarien und Sichten anwenden. Webinhalte werden optimal ausgewählt, zusammengestellt und angezeigt. Das Ergebnis sind Informationen oder Informationssysteme, die für den Benutzer personalisiert und strukturiert dargestellt werden. Ein Beispiel dafür sind Empfehlungssysteme, die im E-Commerce erfolgreich eingesetzt werden [Henze, 2006, S. 137, ff.].

Das World Wide Web Consortium [W3C FAQs] identifiziert folgende Ziele des Semantic Web:

- **Datenintegration:** Es gibt Daten, die man jeden Tag nutzt und die in Fotos, Büchern, Kalendern, Blogs, PDF-Dokumenten und Wikis individuell vorhanden sind. Viele dieser Daten sind nicht Teil des Webs oder werden von verschiedenen Applikationen genutzt. Ein Ziel des Semantic Web ist, ein Web zu schaffen, auf dem Daten mit verschiedenen Formaten und aus unterschiedlichen Orten in einer Applikation integriert werden. Man kann zum Beispiel in einem Online-Kalender Fotos anschauen und sehen, was man gemacht hat, als man sie geschossen hat, und auch, wie viel Geld man damals auf dem Girokonto hatte.
- **Wissensteilung und -austausch:** Intelligente Software-Agenten ermöglichen im Semantic Web eine bessere Wissensteilung und besseren -austausch.
- **Quellen entdecken und klassifizieren:** Die Informationen in World Wide Web können durch semantische Technologien besser gefunden und klassifiziert werden. Dadurch kann man die Ergebnisse der Suchmaschinen optimieren und besser an die Anforderungen der Benutzer anpassen.

Die Vision des Semantic Web ist, das gegenwärtige Internet, das überwiegend aus textlichen Informationen besteht, mit Daten und Metadaten zu erweitern, damit intelligente Applikationen oder Systeme einen besseren Zugriff auf die Informationen haben. Ein gemeinsames System soll geschaffen werden, das erlaubt, dass Daten geteilt und wiederverwendet werden. Die Daten können einerseits automatisch durch verschiedene Applikationen wiederverwendet werden und andererseits durch manuelle Einträge [W3C FAQs].

1.2 Semantic-Web-Technologien

Zur Verwendung des Semantic Web muss man zunächst die Daten mit semantischen Annotationen in einem einheitlichen Format versehen, damit diese auch von Maschinen „verstanden“ werden können. Danach erfolgt die Klassifizierung der Daten aus verschiedenen Domains anhand ihrer Eigenschaften und Beziehungen zu anderen Daten. Dafür werden Semantic-Web-Technologien wie RDF, RDFS und OWL verwendet [SemWeb, 2008].

1.2.1 Resource Description Framework (RDF)

Das RDF (Resource Description Framework) ist ein XML-basierter Standard zur Beschreibung von Ressourcen im Internet, Intranet und Extranet [SemWeb, 2008]. Das RDF wird benutzt, um Wissen in kleine Teile aufzusplittern. Diese kleinen Teile enthalten Regeln über die Bedeutung (Semantic) des Wissens und werden auch oft als Triples bezeichnet [RDF, 2006].

Die wichtigsten Bestandteile von RDF sind Ressourcen, Merkmale und Statements [Antoniou und van Harmelen, 2004, S. 64]:

- **Ressourcen:** Eine Ressource ist ein Objekt, eine „Sache“, über die man reden will. Ressourcen können Autoren, Bücher, Verlage, Orte, Hotels, Zimmer usw. sein. Jede Ressource hat einen URI (Universal Resource Identifier), die sie eindeutig identifiziert. Dieser URI kann zum Beispiel ein URL (Unified Resource Locator oder Webseiten-Adresse) sein.
- **Merkmale (Properties):** Merkmale sind Attribute oder Relationen, um eine Ressource zu beschreiben. Beispiele dafür sind die Attribute: „geschrieben von“, „Alter“, „Titel“ usw. Die Merkmale in RDF werden auch mit URIs versehen. Somit haben sie weltweit ein eindeutiges Namensschema, das sie identifiziert.
- **Statements (Triples):** Statements stellen die Aussagen über die Ressourcen dar. Ein Statement ist ein Subjekt-Prädikat-Objekt-Triple. Das Subjekt besteht aus einer Ressource, das Prädikat aus einer Eigenschaft und das Objekt entweder aus einer Ressource (dargestellt durch eine Ellipse) oder einem Literal (Eigenschaftswert, der nicht weiter annotiert werden kann; dargestellt durch ein Rechteck).

Ein Beispiel für ein einfaches RDF-Triple ist [SemWeb, 2008]:

[Ressource]	[Eigenschaft]	[Wert] .
Das Hotel	heißt	Linderhof.
[Subjekt]	[Prädikat]	[Objekt]

Zur Veranschaulichung kann man RDF-Triples mittels XML/RDF schematisieren oder in einem Graphen darstellen (siehe Abbildung 1).

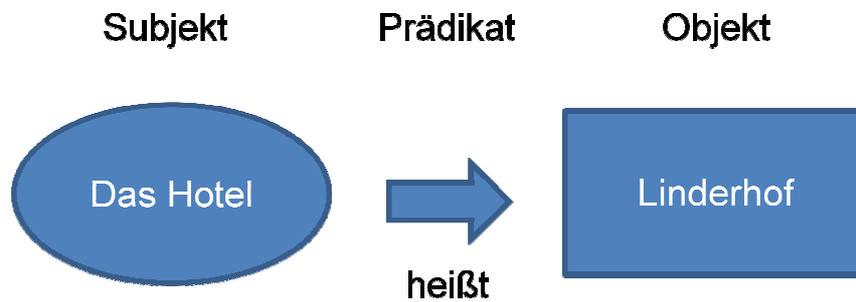


Abb. 1: RDF-Triples [Eigene Darstellung unter Verwendung der Daten von [SemWeb, 2008]]

Durch Kombination von Einzelaussagen sind komplexe Konstrukte möglich. Man kann zum Beispiel die in Abbildung 1 dargestellten Triples mit weiteren Ressourcen erweitern.

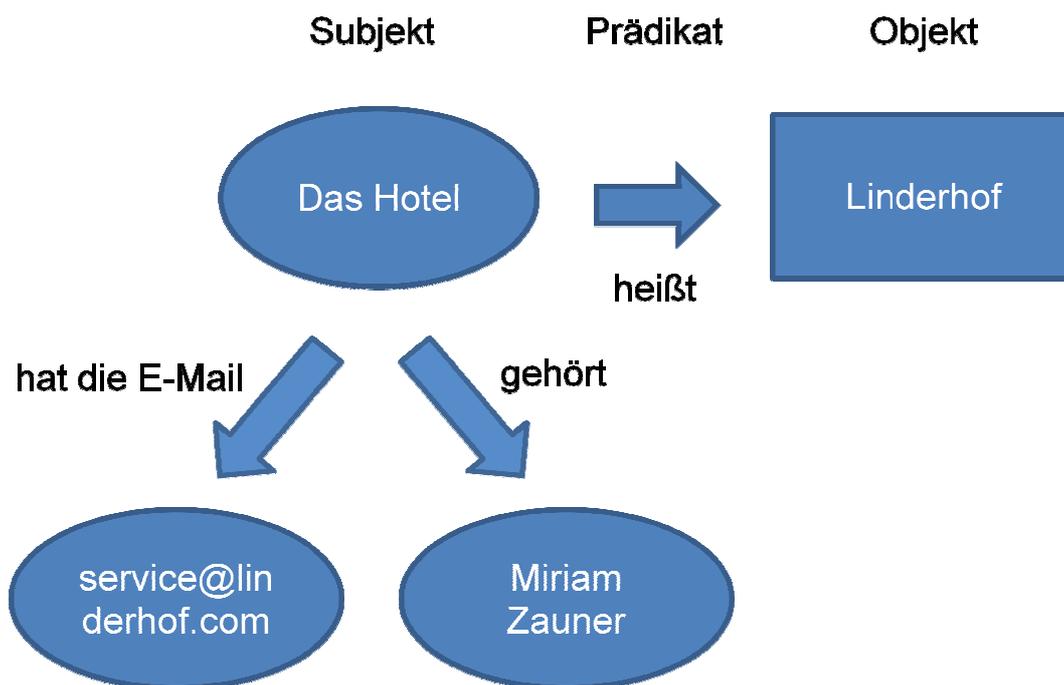


Abb. 2: RDF-Triples mit mehreren Daten [Eigene Darstellung unter Verwendung der Daten von [SemWeb, 2008]]

Damit die Triples von den Maschinen aufgerufen werden können, werden sie entweder in RDF/XML- oder n-Triples-Formaten kodiert. Die Unterscheidung zwischen Subjekt, Prädikat und Objekt ermöglicht den Maschinen, Verknüpfungen zwischen Subjekt und Objekt zu machen. Wichtig ist dabei, dass sämtliche Elemente wie die

Ressourcen, Klassen und Eigenschaften eindeutig identifizierbar sind. Um dieses zu erreichen, verwendet man Namensräume. Die Klassen und Eigenschaften erhalten durch Kombination mit dem Namensraum-Präfix eine eindeutige URI [SemWeb, 2008].

Das abstrakte RDF-Modell kann mit vier einfachen Regeln beschrieben werden [RDF, 2006]:

1. Ein Statement wird von einem Subjekt-Prädikat-Objekt-Triple beschrieben.
2. Subjekte, Prädikate und Objekte stellen Namen für Entitäten (Ressourcen) dar. Entitäten entsprechen entweder einer Person, Webseite oder etwas Abstraktem wie Beziehungen und Zustände.
3. Namen sind URIs, die einen globalen Anwendungsbereich haben und sich bei jedem RDF-Dokument auf dieselbe Entität beziehen.
4. Objekte können „text values“ enthalten - auch genannt „literal values“, die in den XML-Schema-Datentypen geschrieben werden können.

Das RDF hat drei wesentliche Einsatzgebiete. Erstens benutzt man das RDF zum Finden von Daten und Ressourcen für die Bereitstellung von besseren Suchmaschinenkapazitäten. Zweitens kann man die Inhalte und Beziehungen von Objekten in Webseiten oder digitalen Bibliotheken beschreiben und drittens kann man das vorhandene Wissen im Web besser und automatisiert teilen und austauschen [Andersen, 2003].

Damit das Semantic Web sich durchsetzt, müssen so viele Daten wie möglich im RDF-Format beschrieben sein. Es gibt verschiedene Applikationen, die es ermöglichen, sehr einfach Daten mit RDF-Annotationen zu versehen. Ein Beispiel ist der RDFmaker¹, der derzeit noch in Entwicklung ist und mit dem man kostenlos Daten und Dokumente im RDF-Format beschreiben kann [RDFmaker, 2007].

Das Ziel des RDFmakers ist, beliebige Sachverhalte (Angebote, Produktbeschreibungen) maschineninterpretierbar zu beschreiben und für Semantic-Web-Applikationen kompatibel zu machen. Somit können E-Commerce-Anbieter zum

¹ <http://myontology.deri.at/rdfmaker>

Beispiel die Eigenschaften eines Mobiltelefons beschreiben (Marke, Gewicht, Frequenzen, Preis, Verfügbarkeit, Versandkosten usw.) [RDFmaker, 2007].

The image displays two screenshots of the RDFmaker application interface. The left screenshot, titled 'Schritt 3: Dateneingabe', shows a form for entering data. It includes fields for 'Preis' (299,50), 'Währung' (Euro), 'Verfügbarkeit (in Tagen)' (1), 'Nachnameversand' (radio buttons for Ja and Nein, with Nein selected), and 'Versandkosten'. The right screenshot, titled 'Ergebnis: RDF Datensatz', shows the resulting RDF XML code in a text area, with a 'Download' button and a 'Speichern unter...' field above it. The XML code is as follows:

```
<?xml version="1.0" encoding="UTF-8" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://smart-
infosys.com/ebsemantics/offerontology/">
  <rdf:Description rdf:about="http://smart-
infosys.com/ebsemantics/offerontology/offer">
    <dc:title>Product Offer</dc:title>
    <dc:publisher>RDFmaker Online Instancinf Tool
  </dc:publisher>
  </rdf:Description>
  <dc:Product>
  <.../></dc:Product>
</rdf:RDF>
```

Abb. 3: Dateneingabe und -ergebnis in der Applikation RDFmaker [RDFmaker, 2007]

Diese Information wird von dem RDFmaker in RDF annotiert; dadurch maschineninterpretierbar und somit für intelligenten Semantic-Web-Applikationen nutzbar [RDFmaker, 2007].

RDFmaker bietet zwei Möglichkeiten, einzelne Instanzen von gewünschten Gegenständen (z.B. Produkte wie Mobiltelefone oder Digitalkameras) zu generieren: manuell, d.h. durch das Ausfüllen der Felder einer HTML-Form und semiautomatisch, basierend auf einem Datenbankauszug (CSV-Datei). Bei der CSV-basierten Instanzgenerierung können mehrere Produkt-Instanzen in einem Vorgang erstellt werden [RDFmaker, 2007].

RDF beschreibt Ressourcen nach einem bestimmten Modell mit einer bestimmten Syntax, ohne Angaben über die Semantik (Bedeutung) der Ressourcen zu machen. Die Semantik wird erst mit dem RDFS und OWL erfasst [SemWeb, 2008].

1.2.2 RDF-Schema (RDFS)

In einem ausgesprochenen Satz zwischen zwei Menschen hat jedes einzelne Wort eine Bedeutung, die für das Verständnis der Aussage wichtig ist. Das heißt, der Sender und der Empfänger müssen der Aussage dieselbe Bedeutung zuschreiben, damit die Kommunikation reibungslos verläuft [Andersen, 2003].

Ähnlich verhält es sich bei RDFS-Statements, die vorwiegend zur Interkommunikation zwischen heterogenen Software-Systemen dienen und so einheitlich und präzise wie möglich sein sollen [Andersen, 2003].

Mit RDFS kann man Vokabulare erstellen, die die Beziehungen und Bedeutung von RDF-Ressourcen beschreiben. Mit dem RDFS-Vokabular kann man die RDF-Ressourcen an bestimmte Domains zuweisen und Ressourcenklassen festlegen, die ähnliche Eigenschaften haben [SemWeb, 2008].

RDFS wird mittels RDF-Statements dargestellt und ist hierarchisch aufgebaut. Die Ressourcen werden als Instanzen von Klassen definiert, wobei jede Klasse eine Unterklasse von einer anderen sein kann. Diese Hierarchie erlaubt den Maschinen, die Bedeutung der Informationen auf Basis der Eigenschaften und Klassen von Ressourcen zu erfassen [SemWeb, 2008].

Das Web Ontologie Language (OWL) baut auf RDFS auf und verwendet ein umfangreiches Vokabular zur Definition von Semantic-Web-Ontologie.

1.2.3 Ontologien und Web Ontology Language (OWL)

Ontologien sind zentrale Bestandteile des Semantic Web und bilden die höchste Stufe semantischer Reichhaltigkeit. Sie sind formale Strukturen, die einen Teil der Realität beschreiben. Sie stellen Konzepte und ihre Beziehungen innerhalb einer Wissensdomäne dar und unterstützen die Maschinen bei der Interpretation der Information [Blumauer und Pellegrini, 2006, S. 12]. Das OWL wurde vom W3C als Standard-Beschreibungssprache von Ontologien vorgeschlagen. Diese definiert die Arten von Beziehungen, die in RDF (im XML-Vokabular) ausgedrückt werden, um die

Hierarchien und Beziehungen zwischen den verschiedenen Daten darzustellen [SemWeb, 2008].

Laut Blumauer und Pellegrini [2006, S. 12] haben Ontologien folgende Einsatzbereiche:

- Vereinheitlichung und Übersetzung zwischen verschiedenen Wissensrepräsentationsformen.
- Entwicklung von Applikationen, die zur Unterstützung der Wissensarbeiten dienen.
- Datenaustausch zwischen Programmen.
- Erleichterung und Unterstützung der menschlichen Kommunikation.
- Definition von neuen Termen durch die Kombination von bestehenden.
- Bedeutung der Termen wird „formal“ spezifizierbar.

Darüber hinaus können Ontologien genutzt werden, um Produkt- und Dienstleistungsstandards zu spezifizieren. Diese können von Unternehmen benutzt werden, um Produkte zu beschreiben und zu klassifizieren [Blumauer und Pellegrini, 2006, S. 12].

In der Abbildung 4 wird ein Teil einer Gastronomie-Ontologie dargestellt, die zur Beschreibung von Restaurants dient². Das Restaurant wird in der Ontologie mit dem Namen „Inn“ dargestellt. Von diesem Restaurant werden unter anderem das Menü und die Öffnungszeiten dargestellt. Diese Ontologien stellen einen Inhaltsstandard dar, der von der Gastronomie-Branche genutzt werden kann, um die Restaurants einheitlicher zu beschreiben. Je mehr Restaurants mit dieser Ontologie beschrieben werden, desto besser können sie mit semantischen Applikationen gefunden werden.

² <http://www.ebsemantics.net/doc/gastro.jpg>

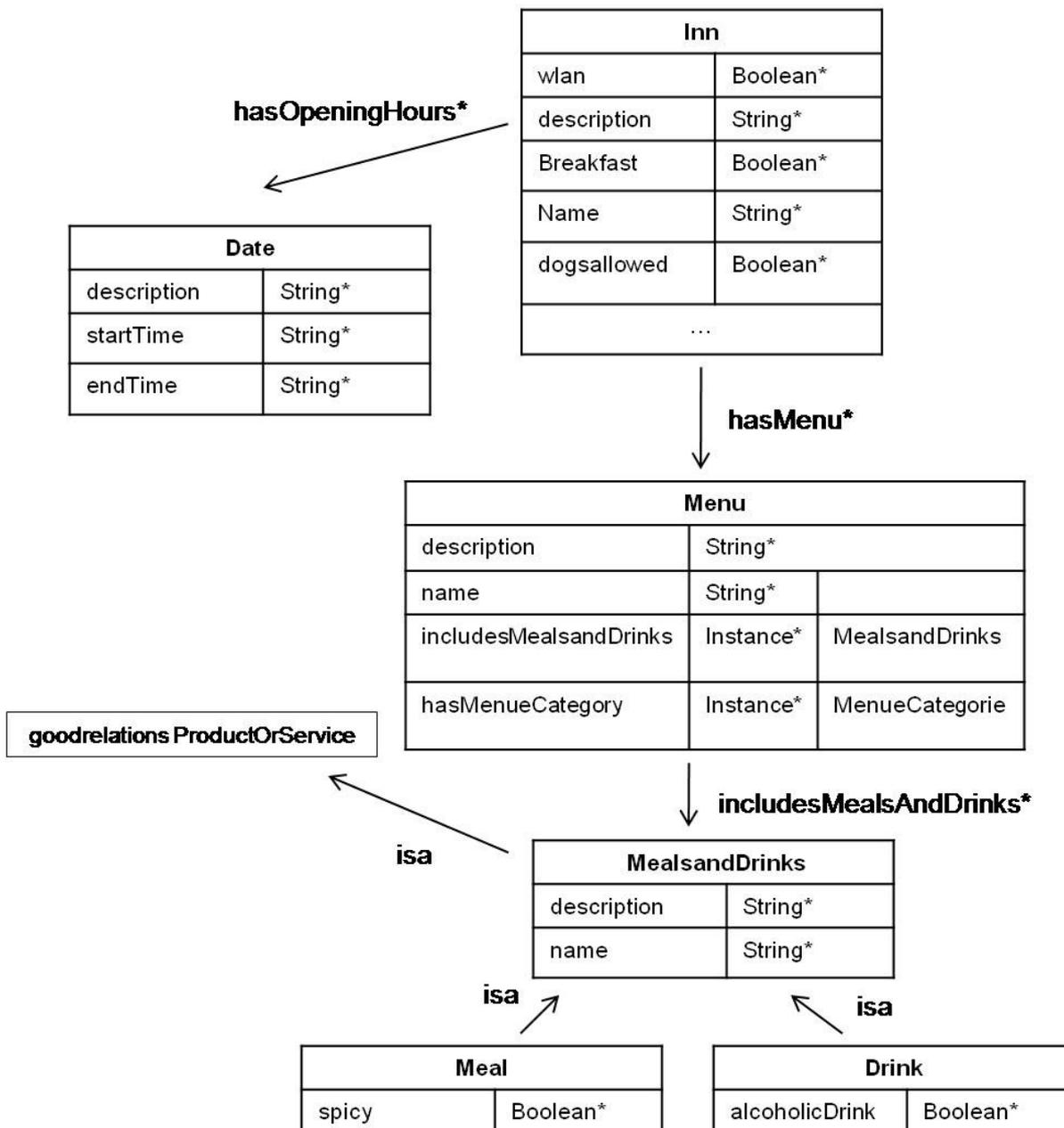


Abb. 4: Ausschnitt einer Gastronomie-Ontologie [Eigene Darstellung unter Verwendung der Daten von [ebSemantics Ont.]]

Im Wissensmanagement und im E-Commerce stellen Inhaltsstandards einen hohen Stellenwert dar. Ein Vorteil der Ontologien gegenüber anderen Inhaltstandards ist die Verwaltung der Versionsänderungen. Die Branchen-Standards durchlaufen pro Jahr hunderte von Aktualisierungen, die pro Aktualisierung zwischen 50 und 600 Änderungen enthalten. Ontologievergleichs-Techniken ermöglichen einen Vergleich der unterschiedlichen Ontologie-Versionen und heben die Unterschiede hervor [Fensel et al., 2002]. Somit können Restaurant-Betreiber über die Änderungen in den Ontolo-

gien informiert werden und ihre Daten entsprechend aktualisieren.

Damit sich das Semantic Web durchsetzt, ist es wichtig, gute Ontologien zu entwickeln, die Inhaltstandards schaffen. Es gibt verschiedene Initiativen, die Ontologien für bestimmte Einsatzbereiche designen oder Werkzeuge für die Entwicklung von Ontologien zur Verfügung stellen. Ein Beispiel ist die kollaborative Ontologie-Entwicklungsplattform, myOntology [myOnt]. Diese basiert auf Wiki-Technologien und ermöglicht eine dynamische und zeitnahe Weiterentwicklung von Produktontologien. Durch die enge Kooperation mit internationalen Standardisierungsorganisationen und Experten aus verschiedenen Branchen kann man Ontologien schaffen, die die nötigen Informationen über die Produkte, Produkteigenschaften und Anbieter darstellen können [ebSem AUSTRIAPRO, 2007, S. 40].

Eine weitere Initiative, um Ontologien zu entwickeln, ist OntoGame³, eine Webseite mit verschiedenen Spielen, in denen Menschen spielend Ontologien entwickeln. Durch kostenlose Spiele kann man viele Personen besser motivieren und dafür begeistern mitzuarbeiten [Siorpaes und Hepp, 2007].

1.3 Abgrenzung des Begriffs Semantic Web

Es gibt verschiedene Sichtweisen und Einsatzmöglichkeiten von Semantic-Web-Technologien. In Anlehnung an den Sammelband von Pellegrini und Blumauer [2006] und Fensel et al. [2002] wird in dieser Arbeit das Semantic Web von zwei Sichtweisen betrachtet: einerseits der Einsatz von Semantic Web im E-Commerce-Bereich und andererseits im Wissensmanagement.

Das Electronic Commerce ist Teil des E-Business, in dem elektronische Geschäftsprozesse angebahnt, vereinbart und abgewickelt werden. Das heißt, zur Erzielung einer Wertschöpfung findet ein Leistungsaustausch im Internet statt [Meier und Stormer, 2005, S. 2].

Mit Wissensmanagement wird die Informationsbeschaffung, -aufbereitung und -bewahrung bezeichnet [ITWissen, 2008]. Dieses Wissen kann sich auf das Wissen im Internet oder in einem Unternehmen beziehen.

³ <http://www.ontogame.org>

2 Nutzenpotentiale des Semantic Web

Der folgende Abschnitt führt in die Anwendungsgebiete der Semantic-Web-Technologien ein, indem zunächst die Anwendungen im E-Commerce (Electronic Commerce) und im Wissensmanagement erläutert werden. Ein besonderer Schwerpunkt des Wissensmanagements wird auf die Verknüpfung des Semantic Web mit Web 2.0-Applikationen gelegt. Anschließend wird anhand wissenschaftlicher Literatur und aufgrund von Expertengesprächen eine SWOT-Analyse des Semantic Web dargestellt. Am Ende des Kapitels werden die Ergebnisse von einer vom Verfasser geführten Experten-Befragung erläutert, die über die Anwendbarkeit, den Nutzen und die Zukunft der semantischen Technologien Auskunft gibt.

2.1 Anwendungen der Semantic-Web-Technologie im E-Commerce

Das E-Commerce zeichnet sich durch eine Wachstumstendenz aus. Immer mehr Nutzer lesen die Nachrichten, kaufen Theaterkarten, suchen nach elektronischen Geräten und planen ihren Urlaub im Internet. Der Stellenwert des Internet als Verkaufskanal wird immer wichtiger und die Anzahl an Händlern und Anbietern, die dieses Medium nutzen, um ihre Produkte zu verkaufen, wächst ständig [Dustdar et al., 2006].

Dadurch, dass fast jeder Informationen und Angebote ins Netz stellen kann, gibt es eine Überflut an Informationen und Angeboten. Diese Überflut, gemeinsam mit der Fülle an unstrukturierten Daten im Web, erschwert es den Benutzern, die richtigen Produkte oder Dienstleistungen zu finden [Dustdar et al., 2006].

Ein wichtiges Ziel des Semantic Web ist, die E-Commerce-Interaktionen flexibler und automatisierter zu gestalten. Dafür ist eine Standardisierung von Ontologien, Nachrichten-Inhalten und Nachrichten-Protokollen notwendig [Lei und Horrocks, 2003]. Durch die Semantic-Web-Technologie werden die Inhalte des Internet mit einer Bedeutungsebene versehen. Dadurch können Computer den Sinnzusammenhang zwischen verschiedenen Quellen erkennen und somit Suchprozesse optimieren [Web Inhalte, 2008]. Übertragen auf das E-Commerce heißt dies, dass die Produkte und deren Eigenschaften durch die Semantic-Web-Technologie so beschrieben

werden, dass diese für Maschinen und Menschen auswertbar werden [Dustdar et al., 2006].

In den folgenden Kapiteln werden exemplarisch die verschiedenen Anwendungen des Semantic Web im E-Commerce erläutert. Als Erstes werden die digitalen semantischen Empfehlungssysteme dargestellt. Diese unterstützen die Benutzer beim Kauf von Produkten und bieten ihnen bedürfnisorientierte Angebote an. Zweitens werden die Einsatzmöglichkeiten des Semantic Web im mobilen Bereich erläutert. Anschließend wird die Verknüpfung von Semantic Web und Werbung dargestellt.

2.1.1 Semantische Empfehlungssysteme

Digitale B2C (Business to Consumer)-Empfehlungssysteme unterstützen die Benutzer bei der Selektion von Informationen über Produkte und Dienstleistungen und beim Treffen von Entscheidungen [Mayr, 2006, S. 2]. Das Ziel der digitalen Empfehlungssysteme ist, die Entscheidungsqualität der Benutzer bei der Produktauswahl zu optimieren und zu verbessern [West et al., 1999].

Es gibt zwei Arten von digitalen Empfehlungssystemen, die personalisierten und nicht-personalisierten Empfehlungssysteme. Die nicht-personalisierten Empfehlungssysteme basieren auf den Empfehlungen und Ratings von anderen Benutzern. Die personalisierten Empfehlungssysteme beziehen sich auf die Eigenschaften oder Informationen über die Benutzer und geben dementsprechend Empfehlungen ab, die am besten zu der Person passen [Spitzer, 2005, S.6, ff.].

Die meisten personalisierten Empfehlungssysteme werden bei Online-Shops wie zum Beispiel Amazon⁴ oder Reel⁵ eingesetzt. Diese Systeme haben meistens einen engen Fokus bei der Erfassung von transaktionellen Daten und empfehlen ähnliche Produkte, die die Benutzer in der Vergangenheit gekauft haben. Zum Beispiel, ein Online-Shop weiß, dass ein bestimmter Kunde eine Bluse gekauft hat oder nach dieser gesucht hat. Dieser Shop weiß vielleicht auch, welche weiteren Kleidungsstücke diese Person in der Vergangenheit erworben hat. Er kennt die Artikelnummer, den gezahlten Preis, die Farbe, die Marke und Tag und Zeit, an dem der Artikel

⁴ <http://www.amazon.com>

⁵ <http://www.reel.com>

gekauft wurde. Diese Daten sind nützlich, aber es gibt viel Information, die vom System nicht erfasst wird. Dieses weiß nicht, ob die Bluse eher konservativ oder auffällig ist, ob sie modisch ist, ob sie eher casual oder formell ist. Wie wichtig es für den Benutzer ist, dass sie von einer bekannten Marke ist, ist ihm auch nicht bekannt [Ghani und Fano, 2002].

Semantic-Web-Technologien ermöglichen, dass digitale Empfehlungssysteme Produkte im Internet finden, die den spezifischen Bedürfnissen von den Benutzern entsprechen. Heutzutage sind die Benutzer bei der Produktsuche auf die kleine Anzahl von Angeboten angewiesen, die zentral in E-Commerce-Plattformen verfügbar sind. Zukünftige Applikationen mit Semantic-Web-Technologien werden den Benutzern einen Suchprozess ermöglichen, der auf Produktattributen basiert, die alle Produkte erfassen, die in einer solchen Form im Internet publiziert werden. Darüber hinaus werden im nächsten Schritt digitale Empfehlungssysteme sogar fähig sein, Geschäftsprozesse wie Kauf und Verkauf oder Verhandlungen automatisch zu bewältigen [Dustdar et al., 2006].

Ein Beispiel für ein solches semantisches Empfehlungssystem wurde von Ghani und Fano [2002] entwickelt und analysiert die Produkte und die Produkteigenschaften, die durch semantische Attribute beschrieben werden. Mittels Text-Lerntechniken extrahiert das System semantische Eigenschaften, die den angestrebten Bereich, in diesem Fall Kleidung, beschreiben. Als Ergebnis entsteht eine Wissensbasis mit den Produkten und Produkteigenschaften in semantischer Form. Dieses System analysiert die Beschreibungen der Produkte, die der Benutzer im Internet angeschaut oder gekauft hat. Diese Daten werden in semantische Attribute übersetzt und automatisch gespeichert, um den Benutzer zu profilieren und seinen „Geschmack“ zu verstehen. Diese Abstraktion erlaubt es dem System, Produkte von der gleichen Produktkategorie vorzuschlagen und gleichzeitig Produkte aus anderen Kategorien zu empfehlen, die zu dem Benutzerprofil passen. Der Ansatz dieses Systems ist, die Empfehlungsergebnisse in qualitativen Eigenschaften zu „erklären“, um das Vertrauen der Kunden in die Produktempfehlungen zu stärken.

In anderen Branchen, wie zum Beispiel der Unterhaltungselektronik-Branche, kann man digitale semantische Empfehlungssysteme ebenfalls einsetzen. Gerade bei beratungsintensiven Produkten, wie zum Beispiel Digitalkameras oder Fernsehern,

ist das Bedürfnis nach einer guten Beratung hoch. Laut einer internationalen Umfrage der Wiener Wirtschaftsuniversität⁶, die vom Institut „International Marketing und Management“ in Zusammenarbeit mit dem Unternehmen Smart Information Systems⁷ durchgeführt wurde [Versandhaus, 2008; Interview Linder; Presseportal, 2007], fühlt sich mehr als die Hälfte der Benutzer bei der Auswahl von technischen Produkten im Online-Bereich überfordert. Besonders sind die technisch wenig versierten Nutzer betroffen. Des Weiteren besagt die Studie, dass 70% der Online-Shopper sich einen Produktberater direkt im Online-Shop wünschen.

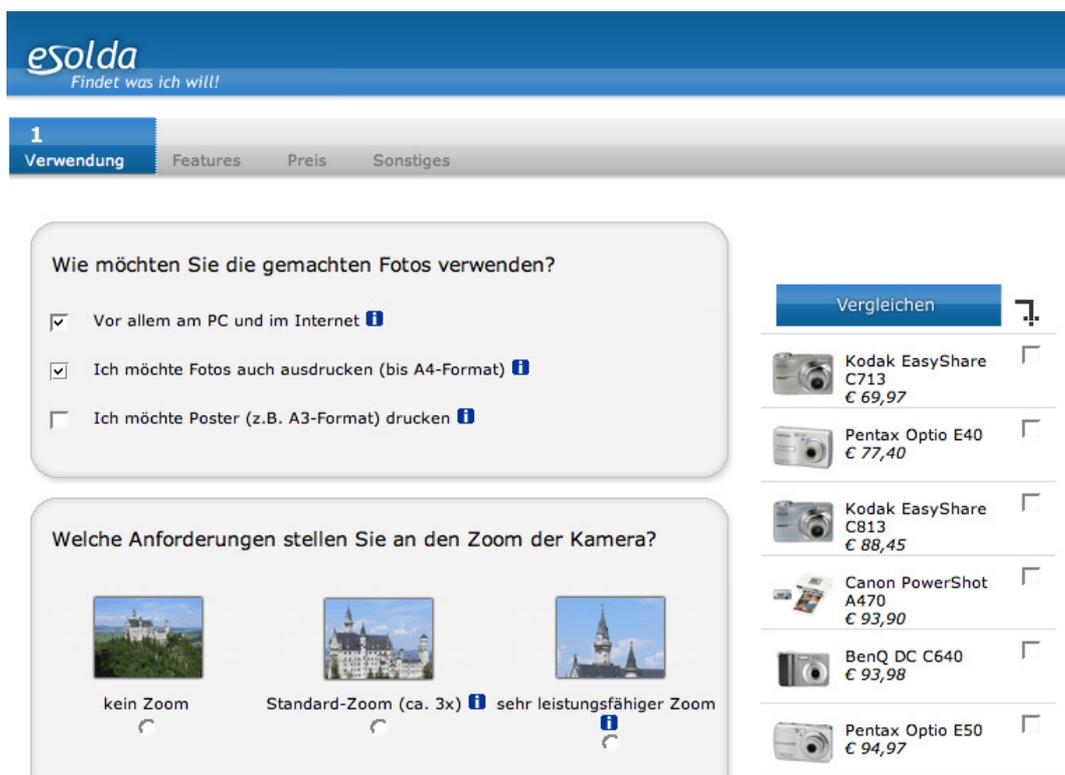


Abb. 5: Screenshot des digitalen Empfehlungssystems Esolda [Esolda, 2008]

Ein Beispiel für ein digitales semantisches Empfehlungssystem ist die Plattform Esolda⁸ (siehe Abbildung 5). Dieses System berät die Benutzer beim Kauf von elektronischen Produkten wie zum Beispiel Fernsehern oder Digitalkameras. Die Kunden werden durch einen dynamischen Dialogprozess geführt, in dem sie ihre Bedürfnisse erklären und am Ende die Beschreibung der Produkte aus verschiedenen Online-Shops erhalten, die am besten zu ihren Anforderungen passen. Wenn die

⁶ <http://www.wu-wien.ac.at>

⁷ <http://www.smart-infosys.at>

⁸ <http://www.esolda.at>

Kunden sich für ein Produkt entscheiden, dann werden sie auf die Webseite des jeweiligen Online-Shops weitergeleitet [Esolda, 2008].

Semantische Empfehlungssysteme bieten Vorteile für die Benutzer und Anbieter. Laut Markus Linder [Interv. Linder] profitieren die Benutzer dadurch, dass sie durch semantische Empfehlungssysteme einen raschen Überblick über die Produktpalette bekommen und bedürfnisorientiert zum Produkt hingeführt werden. Darüber hinaus erhalten sie eine Produktempfehlung, die auf ihre Bedürfnisse zugeschnitten ist. Somit sparen sie Zeit und haben einen geringeren Aufwand bei der Suche.

Weitere Vorteile liegen darin, dass die Empfehlungssysteme nicht nur einzelne Produktkategorien wie Fernseher, Digitalkameras oder Hotels anbieten, sondern auch Produktbündelberatung. Mithilfe der Semantic-Web-Technologien ist es möglich, die Produktberatung intelligent über verschiedene Produktbereiche durchzuführen [Interv. Linder]. Wenn man in einem Portal nach einem Urlaub sucht, kann das Empfehlungssystem zum Beispiel ein komplettes Urlaubspaket mit Hotel, Auto, Kulturevents und Konzerttickets anbieten.

Für die Anbieter bietet das Semantic Web auch sehr viele Möglichkeiten. Durch diese Technologie können zum Beispiel Hersteller und Anbieter diejenigen Kunden erreichen, die an Nischenprodukten interessiert sind. Das ist möglich, weil die semantischen Empfehlungssysteme Angebote aussuchen, die am besten den Kundenanforderungen entsprechen und nicht die, die am meisten gekauft wurden. Allerdings steigt durch diese Technologie auch die Markttransparenz und damit der Wettbewerb zwischen den verschiedenen Produkthanbietern [Dustdar et al., 2006].

2.1.2 Mobile Dienste

Die mobilen Dienste sind Teil des Mobile-Commerce (M-Commerce) und bieten viele Verwendungszwecke, die die Benutzer und Unternehmen bei personalisierten und zeitkritischen Aufgaben unterstützen [Sadeh, 2002]. Heutzutage bestehen die mobilen Dienste überwiegend aus einfachen Dienstleistungen. Man kann zum Beispiel Parktickets und Kinotickets kaufen, Nachrichten abonnieren, den TAN-Code für Online-Banking beziehen oder nach einer bestimmten Route oder einem Restaurant in der Nähe suchen. Gerade bei den mobilen Anwendungssituationen wie zum

Beispiel im Auto oder auf dem Motorrad ist der Benutzer unter Zeitdruck und will bei den Suchanfragen nicht eine lange Liste an Links bekommen (wie bei den traditionellen Suchmaschinen), die er durchsuchen muss, um die richtige Antwort zu finden. Stattdessen will er eine konkrete Antwort auf seine Anfrage bekommen, die am besten seine Bedürfnisse befriedigt [Wahlster, 2007a].

Laut Sheshagiri et al. [2004] ist der nächste Schritt der mobilen Dienste, den Kontext, in dem sich die Benutzer bewegen, zu verstehen und somit mobile Dienste mit einem höheren Automatisierungsgrad anzubieten. Sie sind der Meinung, dass es mit semantischen Technologien möglich ist, mobile Dienste mit intelligenten Agenten zu erweitern, die automatisch Aufgaben für den Benutzer lösen. Dafür sind verschiedene Daten über den Kontext des Benutzers notwendig, wie zum Beispiel, wo er sich befindet, wer seine Freunde und Kollegen sind und weitere kontextbezogene Attribute und Präferenzen. Diese Daten, die sich in verschiedenen Applikationen wie dem Terminkalender, Adressenverzeichnis usw. des mobilen Endgerätes befinden, sollten semantisch aufbereitet sein, damit sie von den intelligenten Agenten erfasst werden können. Darüber hinaus kann dieses System auch auf öffentliche semantische Daten, Webdienste, öffentliche Ontologien oder andere Ressourcen zugreifen, wie zum Beispiel Webdienste von naheliegenden Restaurants oder Wetterberichte usw. Die semantische Technologie ermöglicht, dass die Agenten diese Informationen aus dem Semantic Web „wahrnehmen“, Zusammenhänge erkennen und dieses Wissen nutzen, um automatisierte und personalisierte Lösungen für die täglichen Aufgaben des Benutzers zu liefern. Die Anwendungsmöglichkeiten sind vielfältig, wie zum Beispiel bei Vereinbarung von Treffen, Dokumenten-Austausch, Organisation von Veranstaltungsbesuchen, Filterung von eingehenden Nachrichten und dem Bestellen von Essen.

Nach Wahlster [2007a] können mithilfe der semantischen Technologien mobile Suchanfragen getätigt werden, die konkrete Antworten auf die jeweiligen Fragen in Alltagssprache liefern und somit einen konkreten Vorteil gegenüber anderer Dienste bieten. Diese Dienste können mit Bild- und Spracherkennungsfunktionen erweitert werden. Ein Anwendungsbeispiel dieser intelligenten mobilen Endgeräte ist folgendes: Man will ein Restaurant oder einen bestimmten Veranstaltungsort in einer Stadt finden, dann kann man ein Foto oder eine Abbildung dieses Ortes zum Beispiel aus einem Reiseführer fotografieren. Das Bilderkennungssystem im Handy ordnet das

Foto einem Bild einer Bilddatenbank zu. Die Datenbank erkennt das Foto aufgrund charakteristischer Bildpunkte auch bei geänderten Aufnahmewinkel und modifizierten Beleuchtungsparametern. Die Bildeingabe kann mit einer Spracheingabe kombiniert werden. Zum Beispiel, eine Person macht ein Foto von einem Modell des Brandenburger Tors und fragt, wie sie dorthin kommt (siehe Abbildung 6) [Wahlster, 2007a].

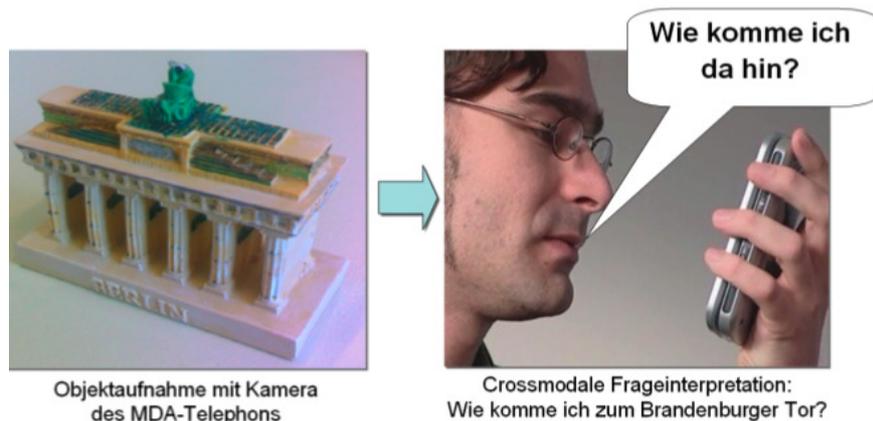


Abb. 6: Mobile Anfragen durch Kombination von Handykamera und Spracheingabe [Wahlster, 2007a]

Bei diesen Systemen findet die Hauptverarbeitungsleistung nicht im Handy oder MDA (Mobile Digital Assistant) statt, sondern auf einem Server, der über eine breitbandige Funkverbindung erreicht wird. Dieser enthält eine leistungsfähige Spracherkennung mit einer semantischen Anfrageanalyse. Weitere Werkzeuge wie eine Bilddatenbank können hinzugefügt werden.

Eine weitere Anwendung des Semantic Web ist die Verknüpfung mit Internet- und Fahrzeug-Diensten. Wie in Abbildung 7 ersichtlich, kann ein Fahrer eines PKW die Frage im Handy eingeben „Wo gibt es ein italienisches Restaurant hier in der Nähe?“ Darauf folgend werden die jeweiligen Restaurants auf einer digitalen Karte im Handy angezeigt. Nachdem er ein Restaurant ausgesucht hat, fragt er das Handy „Wie komme ich dahin? Vorher muss ich tanken.“ Als Ergebnis wird im Handy eine Karte mit dem Weg zum Restaurant und einem Tankstopp angezeigt [Wahlster, 2007a].

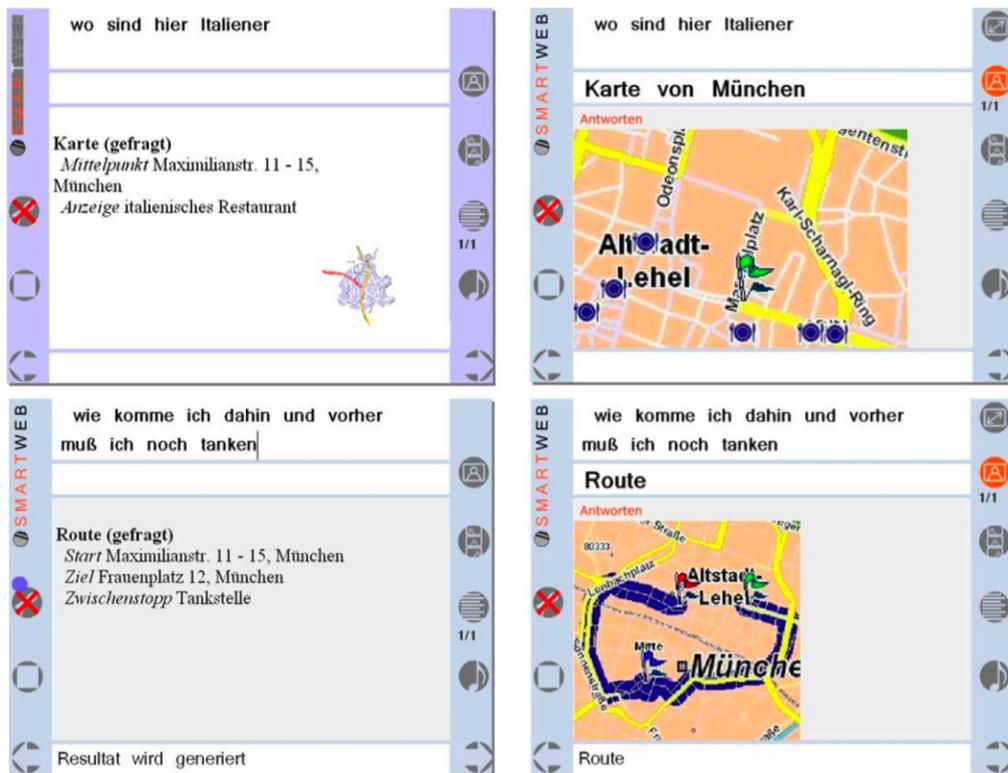


Abb. 7: Navigationsdialog durch die Kombination von mehreren Webdiensten [Wahlster, 2007a]

Die Möglichkeiten in diesem Bereich sind sehr umfassend. Man kann Sprache, Schrift, Gestik, Haptik oder Video im Mobiltelefon kombinieren und als Ergebnis wiederum Sprache, Haptik, Graphik, Video oder Ton bekommen [Wahlster, 2007a].

2.1.3 Werbung

Grundbestandteile der neuen Online-Werbeformen sind kontextbezogenen Inhalte, das Tracking (Folgen) und Targeting (Zielgruppenansprache) der Kunden [TrendONE, 2008]. Bei der kontextbezogenen Werbung geht es darum, die kommerzielle Werbung innerhalb des Inhalts von Webseiten zu platzieren. Bei der konventionellen kontextbezogenen Werbung wird die Webseite, die ein Benutzer gerade anschaut, nach Keywords (Schlüsselwörter) durchsucht und passende Werbung wird geliefert. Das Ziel ist, relevante Werbung zu den Inhalten der Webseite zu liefern und dadurch eine höhere Klickrate und somit höheren Umsatz zu erreichen [Broder et al., 2007]. Darüber hinaus haben viele Studien gezeigt, dass kontextbezogene Werbung den Return on Investment (ROI) erhöht [Knight, 2008].

Die konventionelle kontextbezogene Werbung ist manchmal nicht in der Lage, den echten Kontext der Webseite zu identifizieren und liefert öfters Werbung, die fehlplatziert ist. Als Ergebnis werden Werbekosten verschwendet und es wird nicht die gewünschte Zielgruppe erreicht. Die Ursache dafür ist, dass die kontextbasierten Systeme wenig Verständnis über die Bedeutung und den Zusammenhang der Wörter haben. Das führt zu falscher Platzierung der Werbung, die zu Imageschäden für die verschiedenen Marken führen kann und auch für die Webseiten-Betreiber keinen Mehrwert bietet, da die Werbung somit ihre Wirkung verliert. Ein Beispiel dafür ist folgendes: Eine Werbeanzeige von einem Anbieter für Küchenmesser wird durch die kontextbezogene Werbung neben einem Bericht über eine Messerstecherei platziert. Rein technisch passt die Werbung zum Artikel, bei beiden taucht das Wort Messer auf, aber vom Sinnzusammenhang und von der Bedeutung passen beide Botschaften nicht zusammen [iSense, 2008]. Darüber hinaus kann die Vielseitigkeit der Sprache auch manchmal zu unpassenden Werbebotschaften führen. Wenn zum Beispiel der Artikel über Tiger Woods handelt, könnte es passieren, dass die Suchmaschine das Wort „Tiger“ findet und eine Werbung zu der Salbe „Tiger Balm“ schaltet.

Durch semantische Technologien werden die Inhalte der Webseiten für Maschinen verständlich gemacht. Die Computer erfassen daher die Bedeutung der Inhalte der Webseiten und können auch Sinnzusammenhänge zwischen den Wörtern ableiten [Henze, 2006, S. 140]. Somit hätte die Maschine im vorherigen Beispiel über Tiger Woods verstanden, dass es in dieser Nachricht um Golf geht und nicht um das einzelne Wort „Tiger“.

Ein Beispiel für eine semantisch angereicherte Werbemethode ist die iSense-Applikation⁹, die mithilfe von semantischen Technologien alle Wörter von Webseiten durchsucht, analysiert, Wortbedeutungen erfasst, Redewendungen berücksichtigt und die lexikalische Thematik des Inhalts in wenigen Sekunden erfasst. So wird Werbung für das jeweilige Themenumfeld geschaltet und die Ergebnisse sind zuverlässiger und zielgerichteter als bei konventionellen Online-Werbemethoden. Hauptbestandteil der Applikation ist die „Sense-Engine“, die den lexikalischen Inhalt von Webseiten in Echtzeit erfasst und kategorisiert. Damit dies funktioniert, müssen

⁹ <http://www.isense.net>

die Webseiten semantisch klassifiziert werden, sodass der Ad-Server automatisch die Werbung auf diejenigen Werbeplätze schaltet, die für die werbetreibenden Unternehmen relevant sind [iSense, 2008].

Eine weitere Art von kontextbezogener Werbung wird in Social-Networking-Seiten betrieben. Diese Plattformen speichern viele Daten über die Benutzer, wie zum Beispiel Name, Wohnort, Fotos, Freunde und unterschiedliche Interessen. Man kann sehen, wo sich die Benutzer befinden, mit wem sie sich aus dem Social Network unterhalten und welche Informationen und Produkte sie brauchen [TrendONE, 2008]. Es gab schon einige Initiativen wie die von Liu und Maes [2005], um die Information der Community-Mitglieder durch Ontologien zu strukturieren und „Interest Maps“ zu bilden, die die Interessen der Benutzer darstellen. Mit diesen Daten kann gezielt auf die Bedürfnisse der Benutzer eingegangen und passende Werbung geschaltet werden. Die Werbung ist dann keine Werbung im klassischen Sinne, sondern eine Dienstleistung, die den Benutzern das bietet, wonach sie suchen [TrendONE, 2008].

Es gibt verschiedene Applikationen, in denen neue Werbeformen mit semantischen Technologien innerhalb von Social-Networking-Plattformen entwickelt wurden. Die Firma ThoughtTrail¹⁰ hat zusammen mit dem Microsoft Instant Messenger (MSN)¹¹ eine Initiative gestartet, bei der die Benutzer anhand ihres Gesprächs kontextrelevante Werbung in Form von Videos und Bildern bekommen. In Abbildung 8 wird dies verdeutlicht: Zwei Benutzer unterhalten sich im MSN über den neuesten Film von Will Smith, „I am a Legend“ und einer sagt, dass er sich diesen gerne anschauen will. Parallel bekommen die Benutzer in einer Nebenleiste relevante Filme, Bilder und Links über Will Smith und seinen Film.

¹⁰ <http://www.thoughttrail.com>

¹¹ <http://www.msn.com>



Abb. 8: Kontextbasierte Werbe-Initiative von ThoughtTrail und MSN [TrendONE, 2008]

Weitere semantische Werbearten können in Social-Networking-Seiten mit VoIP (Voice over Internet Protocol) platziert werden. Die Firma Pudding Media¹² ermöglicht Echtzeitanalysen der Stimme in IP-Telefonaten, um auf Basis der kontextrelevanten Informationen zu Produkten und Dienstleistungen zu liefern. Ein Beispiel dafür ist: Zwei Freunde unterhalten sich, wo sie heute Abend in Berlin asiatisch essen können und bekommen dann im Bildschirm verschiedene Anzeigen und Angebote über asiatische Restaurants in Berlin [TrendONE, 2008].

Kontextbasierte Werbung kann auch mit semantischen Bilderkennungs-Methoden verknüpft werden. Mittels einer visuellen Suchmaschine wie Like.com¹³ kann man Bilder analysieren und passende Werbung schalten. Die Fotos aus den Social-Networking-Plattformen werden genutzt, um angepasste Werbung zu platzieren. Zum Beispiel, eine Person trägt auf ihrem Profilfoto eine Sonnenbrille, die vom System als solche erkannt wird (siehe Abbildung 9). Mit diesem Wissen wird in der

¹² <http://puddingmedia.com>

¹³ <http://www.like.com>

Social-Networking-Seite automatisch Werbung über ähnliche Sonnenbrillen geschaltet [TrendONE, 2008].



Abb. 9: Kontextbasierte Brillenwerbung in Facebook [TrendONE, 2008]

Die Webseitenbetreiber und werbetreibende Unternehmen profitieren von der semantischen Anreicherung der Werbung. Für den Werbetreibenden erreicht seine Werbung nur diejenigen Personen, die potentiell Interesse an einem Produkt oder einer Marke haben. Für die Webseiten-Betreiber bietet sich auch ein Mehrwert, weil die Benutzer der Webseite durch die Werbeeinblendungen nicht gestört werden, da diese für sie relevant sind und sie eventuell einen Nutzen daraus ziehen können [iSense, 2008].

2.2 Anwendungen der Semantic-Web-Technologie im Wissensmanagement

Das Wissensmanagement beschäftigt sich mit der Akquise, dem Zugriff und der Wartung von Information [Grigoris und Van Harmelen, 2004]. Traditionelles Wissensmanagement wird mit neuen Problemen im Web-Bereich konfrontiert wie zum Beispiel Informationsüberlastung, ineffizienter „Keyword“-Suche, heterogener Integration der Information und geografisch-verteilten Intranet-Problemen [Fensel et al.,

2002]. Darüber hinaus gibt es nach Grigoris und Van Harmelen [2004] noch weitere Schwächen, einerseits bei der Wartung der Information, da es derzeit Widersprüche bei den Terminologien gibt und veraltete Information, die nicht gelöscht wird. Ein weiteres Defizit findet bei der Auswahl und Suche nach Information statt. Bei unstrukturierten Daten ist es schwer, mit herkömmlichen Systemen die richtige Information zu finden, und wenn man ein Dokument findet, dann müssen die Benutzer dieses lange durchsuchen, um die relevanten Daten zu finden.

Ein effizientes Wissensmanagement ist im Zeitalter der Informationsflut entscheidend und wurde als Schlüsselfaktor für das Beibehalten der Wettbewerbsvorteile in Organisationen identifiziert [Fensel et al., 2002]. Das Ziel des Semantic Web ist, fortgeschrittene Wissensmanagement-Systeme zu schaffen, die Folgendes ermöglichen [Grigoris und Van Harmelen, 2004]:

- Wissen wird in konzeptuelle Räume gemäß deren Bedeutung eingeteilt.
- Automatisierte Werkzeuge unterstützen die Wartung der Daten und suchen nach Widersprüchen und nach neuem Wissen.
- Keyword-basierte Suche wird durch Antwortmaschinen ersetzt. Das heißt, das gesuchte Wissen wird abgefragt, ausgewählt und benutzerfreundlich dargestellt.
- Die Fragenbeantwortung über viele Dokumente hinaus wird unterstützt.

Durch das Semantic Web kann das vorhandene Wissen im Web entwickelt, lokalisiert, verteilt und genutzt werden. Neues Wissen kann durch die Kombination von bestehendem Wissen abgeleitet werden. Studer et al. [2005, S.2] erklären die Entwicklungen des Wissensmanagement folgendermaßen:

„Die zentrale Anforderung an die nächste Generation von Wissensmanagement-Systemen ist die Möglichkeit, Informationen geeignet zu kombinieren, um damit implizites Wissen ableiten und somit neues Wissen generieren zu können. Semantik kann diese Anforderungen erfüllen und bildet somit die Grundlage für eine neue Landschaft an Anwendungen, welche die Informationstechnologie in eine Wissenstechnologie transformiert.“

Im Kapitel 2.2 werden verschiedene Anwendungen dargestellt, in denen man das Semantic Web sinnvoll einsetzen kann. Ein besonderer Schwerpunkt wird auf Web 2.0-Applikationen gelegt, die laut einer Studie von McKinsey [2008] verstärkt in Unternehmen eingesetzt werden.

Das Web 2.0 (user-generated-content) ist ein wichtiger Teil des Wissens im Internet, der schwer zu durchsuchen und zu erfassen ist. Die im Web 2.0 generierten Inhalte stellen die „kollektive Intelligenz“ der beteiligten Personen dar. Die Information ist verteilt gespeichert im Online-Kalender, bei den getaggtten Bildern auf Social-Networking-Seiten oder bei den gespeicherten Links in einer Webseite. Web 2.0-Applikationen entstehen aus der Kombination verschiedener Datenarten und sind im Web verstreut. Die Rolle des Semantic Web ist, diese Daten zu strukturieren und zu beschreiben, damit diese im Web auffindbar sind und von allen genutzt werden können. Web 2.0 und das Semantic Web fokussieren sich auf den intelligenten Datenaustausch und stehen eher in einer ergänzenden als konkurrierenden Beziehung [W3C FAQs, 2008].

2.2.1 Textbasierte Suchmaschinen

Um sich mit der Informationsflut im Internet zurechtzufinden, sind leistungsfähige Online-Suchmaschinen wie Google¹⁴, Yahoo¹⁵ oder Altavista¹⁶ unabdingbar. Diese weisen aber nach Wahlster [2007b] einige Defizite auf:

- Die Anfragen werden als Zeichenkette verstanden, die Bedeutung und Inhalte der Wörter werden vernachlässigt.
- Die Suchanfragen werden voneinander unabhängig betrachtet, der Inhalt und Zusammenhang von vorherigen Suchanfragen wird nicht beachtet. Daher ist auch die Liste an Ergebnissen sehr umfangreich.
- Die Anfragen, die komplexe Restriktionen haben oder sich auf mehrere Relationen beziehen, bringen keine zufriedenstellenden Resultate und können sogar in eine ganz falsche Richtung gehen.

¹⁴ <http://www.google.com>

¹⁵ <http://www.yahoo.com>

¹⁶ <http://www.altavista.com>

- „Versteckte Inhalte“, die nur über PDF-Dokumente und Webdienste verfügbar sind, werden nicht gefunden („Deep Web“).

Duke et al. [2007] unterstützen die Aussagen von Wahlster [2007b] und sind der Meinung, dass es bei herkömmlichen Suchmaschinen öfter der Fall ist, dass die Benutzer bei Suchanfragen eine sehr lange Liste an Links erhalten, die unstrukturiert ist und viele irrelevante Ergebnisse anzeigt. Ein Grund dafür ist, dass die Mehrdeutigkeit und Synonymie kaum analysiert wird und die Bedeutung (Semantik) der Wörter selten erfasst wird. Es gibt zwar Algorithmen wie den PageRank (Rangordnung) Algorithmen von Google, die benutzt werden, um die Suchergebnisse zu optimieren, aber mit diesem Ansatz bekommen neue Webseiten über populäre Themen eine niedrige Rangordnung, die auch niedrig bleibt, im Speziellen wenn die Personen die Seite nicht finden und diese nicht verlinken.

Das Semantic Web will das bestehende Web erweitern und bietet ein integriertes Gesamtsystem mit den technischen Voraussetzungen, um die Defizite von klassischen Suchmaschinen zu überwinden [Wahlster, 2007b]. Die Daten bekommen dafür eine für Maschinen verarbeitbare Semantik (Bedeutung), die es ermöglicht, die Informationen der einzelnen Wörter zu verstehen und Zusammenhänge zwischen verschiedenen Konzepten zu schließen, um die Suchergebnisse gemäß den Anforderungen des Benutzers zu verarbeiten [Andersen, 2003]. Nach Berners-Lee [1998] sind die Suchmaschinen der Zukunft eine Kombination zwischen Suchmaschinen und logischen Maschinen, die es ermöglichen, „Proofs“ zu bilden. Diese Maschinen können umfassende Indexe durchsuchen und mithilfe von Logik die wichtigsten Informationen filtern und somit eine Lösung zur Suchanfrage bieten.

Nach Wahlster [2007a] werden semantische Suchmaschinen in Zukunft nicht mehr nach Schlüsselwörtern suchen und eine lange Liste an Links mit Dokumentenverweisen liefern, sondern auch komplexe, natürlich-sprachlich formulierte Anfragen zulassen, die eine konkrete Antwort auf die Anfrage liefern. Somit wandelt sich die klassische Suchmaschine in eine Antwortmaschine, die sich nicht mehr auf den Suchprozess konzentriert, sondern auf die Lieferung von gezielten Antworten zur Beantwortung der Suchanfragen.

Heute: Google

• **Eingabe:** Umsatz von Siemens im Jahr 2005 in Deutschland

Web Bilder Groups News Froogle Mehr »

Suche: Das Web Seiten auf Deutsch Seiten

• **Ausgabe:**

Frühaufsteher - Investor - FAZ NET - Medienschau Porsche ...
Der Umsatz legte um sechs Prozent auf 12,54 Milliarden Dollar zu ... die aus Siemens-Kassen rund 200 Millionen Euro abgezweigt und die Summe als ...
[www.faz.net/.../](#)
Doc-E580BA22D797A43A3996599E450FE112A~ATpl-Ecommon-Scontent.html - 62k -
[Im Cache](#) - [Ähnliche Seiten](#)

Siemens Financial Services: "Leasingfabrik" als neue ...
Siemens Financial Services präsentiert aktuelle Studie zu ... für die USA) hatten dafür zwischen Februar und April 2005 in Deutschland, Frankreich, ...
[www.silicon.de/enid/it_services/14407 - 32k - Im Cache](#) - [Ähnliche Seiten](#)

Morgen: SmartWeb

• **Eingabe:** Umsatz von Siemens im Jahr 2005 in Deutschland?



• **Ausgabe:** 15,685
Millarden
Euro

	2005		2004	
Sales by % of customer	2005	2004	2005	2004
Germany	15,685	16,223	24,798	25,872
Europe (other than Germany)	24,429	22,787	23,684	20,869
U.S.	14,686	13,378	14,806	13,377
Americas other than U.S.	4,414	3,522	3,651	2,899
Asia-Pacific	10,057	8,990	6,612	5,586
Africa, Middle East, C.I.S.	6,174	5,337	1,894	1,634
Siemens	75,445	70,237	75,445	70,237

Abb. 10: Stichwortbasierte Suche versus Fragebeantwortung [Wahlster, 2007a]

Wie in Abbildung 10 ersichtlich, suchen die semantischen Suchmaschinen einerseits Informationen aus den Passagen von Dokumenten im offenen Web, wie bei den klassischen Suchmaschinen, andererseits haben sie auch Zugriffsmöglichkeiten auf verschiedene Webdienste (zum Beispiel Routenplanung, Wetterinformation, Veranstaltungskalender, Verkehrsinformation, Webcam-Suche) und auf PDFs enthalten, die mittels Bildextraktion durchsucht werden. Verschiedene Lösungswege werden von der semantischen Suchmaschine parallel vorgehalten und je nach Aufgabenstellung wird der beste Ansatz für die Antwort der Suchanfrage geliefert (siehe Abbildung 11) [Wahlster, 2007a].

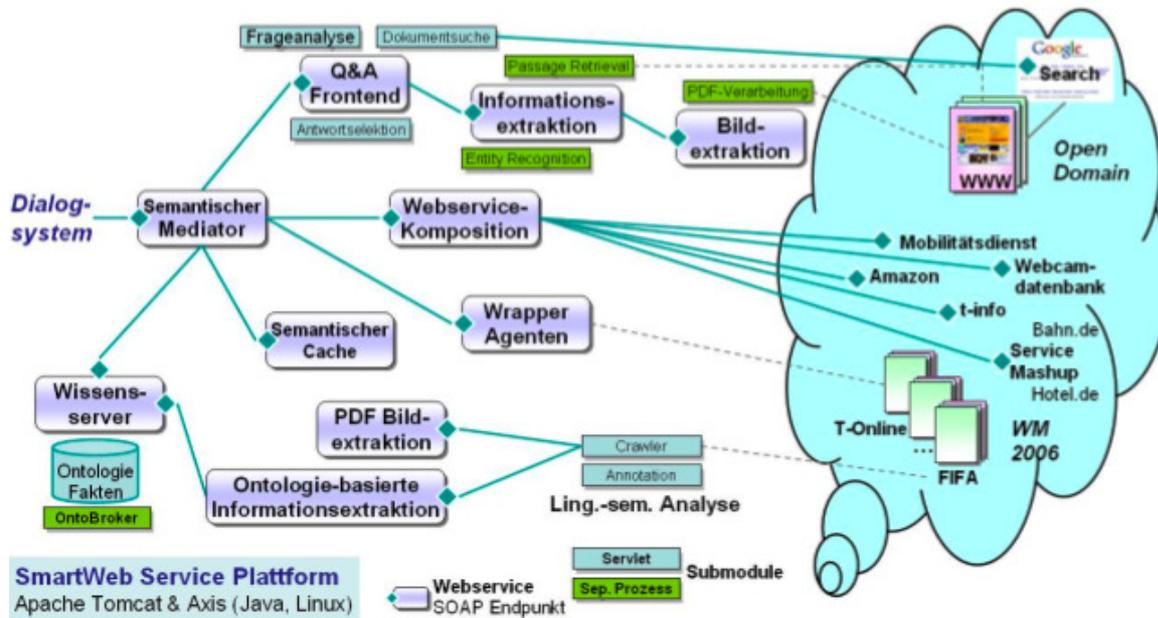


Abb. 11: Zugriff auf Informationen durch den semantischen Mediator [Wahlster, 2007a]

Ein Beispiel für eine semantische Suchmaschine wurde von Yahoo entwickelt und heißt SearchMonkey¹⁷. Diese Suchmaschine bietet ein Framework, um kleine Applikationen zu schaffen, welche die Yahoo!-Suchergebnisse mit zusätzlichen Metadaten und Strukturen wie zum Beispiel Bildern und Links bereichern. Dies ermöglicht Entwicklern und Eigentümern von Inhalten, ihre Daten visuell attraktiver zu gestalten und mehr Traffic auf ihren Webseiten zu generieren. Das SearchMonkey-Tool hilft Entwicklern, bestehende Datendienstleistungen zu finden und eigene Datendienstleistungen zu entwickeln. Sie können außerdem festlegen, wann und wie diese zusätzlichen Metadaten in den Suchergebnissen dargestellt werden. Wenn man eine Applikation geschaffen hat, kann man sie selber nutzen, mit anderen teilen oder in einer Galerie der Öffentlichkeit verfügbar machen [SearchMonkey, 2008].

Eine weitere semantische Suchmaschine wurde von Duke et al. [2007] entwickelt und heißt Squirrel. Das Ziel von Squirrel ist, die Keyword-basierte Suche mit der semantischen Suche zu kombinieren. Der Benutzer kann einen Text oder verschiedene Wörter in der Suchmaschine eingeben und erhält Suchergebnisse, die mithilfe von ontologischer Unterstützung verfeinert werden können. Zum Beispiel, man kann zwischen verschiedenen Themen und passenden Entitäten aussuchen, um genauere

¹⁷ <http://developer.yahoo.com/searchmonkey>

Ergebnisse zu bekommen. Squirrel will damit die Schnelligkeit und Einfachheit der herkömmlichen Suchmaschinen beibehalten und diese mit den Fähigkeiten der semantischen Technologien erweitern.

Die Ergebnisse der semantischen Suchmaschinen können durch die Nutzung von Benutzerprofilen optimiert werden. Wenn der Benutzer nach einem Wort sucht, das zu zwei Entitäten passt, dann muss sich das System für eine der beiden Entitäten entscheiden. Eine Lösung ist, die Suchergebnisse mit dem Profil des Benutzers zu vergleichen und nach passenden Themen zu suchen. Zum Beispiel, wenn ein Benutzer nach Dokumenten sucht, die der Autor Müller geschrieben hat (der verschiedene Instanzen hat), dann kann das System alle Dokumente durchsuchen, die die verschiedenen Autoren namens Müller geschrieben haben. Diese werden mit den Themengebieten im Benutzerprofil verglichen und der Autor, der am besten zu dem Benutzerprofil passt, wird angezeigt [Duke et al., 2007].

2.2.2 Wikis

Bei Wikis handelt es sich um eine Web 2.0-Technologie, welche eine Ansammlung von Seiten und Artikeln zu verschiedenen Themen und Begriffen enthält. Diese Informationen sind in einer Webseite verfügbar und deren Inhalt kann von den Benutzern gelesen und kooperativ bearbeitet werden. Bei semantischen Wikis werden große Wissensbasen miteinander vernetzt, die gemeinsam bearbeitet werden können [Meißner et al., 2007].

Nach Völkel et al. [2006] und Tazzoli et al. [2004] ist eine Schwäche von normalen Wikis, wie zum Beispiel der Wikipedia¹⁸, der Mangel an durch Maschinen verarbeitbaren Daten und die fehlende Verknüpfung der Konzepte. Völkel et al. [2006] meinen, dass die Information und das Wissen in der Wikipedia sehr wertvoll sind, aber derzeit in vielen Gebieten nicht angewendet werden können, weil die Inhalte der Wikipedia von Maschinen nicht „verstanden“ werden. Es ist derzeit nicht möglich, automatisch verknüpfte Information über verschiedene Artikel zu sammeln und Anfragen wie zum Beispiel „Zeige alle Filme aus den 1960-ern mit italienischen Filmregisseuren“ zu beantworten. Obwohl die Daten in der Wikipedia ziemlich struk-

¹⁸ <http://www.wikipedia.org>

turiert sind (jeder Film hat einen eigenen Artikel und Links zu Darstellern und Regisseuren), ist die Bedeutung (Semantik) der Wörter für die Maschinen unklar. Das strukturelle Wissen über die Verknüpfung der verschiedenen Konzepte kann derzeit nicht formell erfasst oder automatisch bearbeitet werden.

Tazzoli et al. [2004] sind auch der Meinung, dass Schwächen von herkömmlichen Wikis an der fehlenden Daten-Verknüpfung und der Isolation zwischen den verschiedenen Wikis liegen. Durch das rasante Wachstum von Wissen im Web gewinnt das Bedürfnis der Strukturierung dieses Wissens immer mehr an Bedeutung. Verschiedene Lösungen wurden von den Wiki-Communities vorgeschlagen, wie zum Beispiel die Ansätze von InterWiki¹⁹ oder SisterSites²⁰, die die Verlinkung zwischen verschiedenen Wikis ermöglichen.

Semantisch basierte Wikis gehen aber einen Schritt weiter. Die Inhalte der Wikis werden mit Metadaten angereichert und Ontologien werden entwickelt, um die Daten zu strukturieren. Dafür sind W3C-Standards wie RDF, RDFS und OWL notwendig [Tazzoli et al., 2004; Meißner et al., 2007]. Diese Methode ist flexibler und ermöglicht, Inhalte und Metadaten von unterschiedlichen Wikis zu verschachteln. Den verschiedenen Themen oder Begriffen innerhalb des Wiki-Systems werden Attribute bzw. Relationen angefügt. Somit ist es möglich, Beziehungen zu anderen Semantic-Wiki-Seiten oder -Begriffen herzustellen [Meißner et al., 2007]. Darüber hinaus wird eine spezifische Suche ermöglicht. Etwas komplexere Fragen, wie zum Beispiel die oben genannte Frage über die italienischen Filmregisseure der 1960er Jahre, können mit semantischen Wikis beantwortet werden [Völkel et al., 2006].

Semantische Wikis sind für Computer sinnvoll, weil diese ermöglichen, dass Webinhalte für Maschinen lesbar gemacht werden. Andererseits sind sie auch für Menschen vorteilhaft, weil durch benutzerfreundliche Anwendungen auch technisch nicht versierte Benutzer Zugang zu den semantischen Technologien haben [Meißner et al., 2007]. Darüber hinaus können die angereicherten Daten durch die RDF-Export-Funktion für das gesamte semantische Web verfügbar gemacht werden und von verschiedenen externen Applikationen genutzt werden [Völkel et al., 2006].

¹⁹ <http://c2.com/cgi/wiki?InterWiki>

²⁰ <http://c2.com/cgi/wiki?AboutSisterSites>

Ein Beispiel für eine semantische Wiki ist die Platypus-Wiki. Mit der Platypus-Wiki können die Benutzer die Wiki-Artikel mit RDF-Klassen und -Eigenschaften anreichern. Für jede Seite wird eine Liste von URLs spezifiziert, aus der man Inhalte sammelt und eine andere Liste von URLs, aus der man Metadaten sammelt. Diese URLs können lokal- oder ferngesteuert sein und es ist irrelevant, von wo die Information aggregiert wird. Die Qualität der Daten und Metadaten hängt, wie bei den normalen Wikis, von der Kollaboration der Benutzer ab. Die Platypus-Wiki kann als ein persönliches Wissensmanagementsystem gesehen werden und als ein Wiki-System. Unternehmen können es auch nutzen, um Wissen der Mitarbeiter zu sammeln und mit anderen Abteilungen zu teilen. Platypus kann daher als ein Lernsystem gesehen werden, das benutzt wird, um anderen Wissen weiterzugeben und vom Wissen anderer zu lernen [Tazzoli et al., 2004].

Völkel et al. [2006] haben in einer weiteren Initiative das bekannteste Wiki, die Online-Enzyklopädie Wikipedia, mit semantischen Technologien angereichert und somit die Semantic-Wikipedia geschaffen. Diese eröffnet neue Möglichkeiten für die Suche, die Aggregation oder den Export von Wissen. Die Benutzer haben in der Semantic-Wiki die Möglichkeit, Links zwischen den verschiedenen Artikeln zu setzen und Daten innerhalb der Artikel semantisch zu spezifizieren. Dafür werden semantische Links und Attribute verwendet. Die semantischen Links sind eine Erweiterung der bestehenden Links von der Wikipedia. In den Abbildungen 12 und 13 werden die Verlinkungen der aktuellen Wikipedia (Abbildung 12) und die Verlinkungen der Semantic-Wikipedia (Abbildung 13) gegenübergestellt. Bei der Wikipedia (Abbildung 12) werden einzelne Wörter zu anderen Artikeln verlinkt während bei der Semantic-Wikipedia (Abbildung 13) zusätzlich die Relationen zwischen den Artikeln dargestellt (semantische Links) werden. Diese Relationen werden in Abbildung 13 dargestellt. Im Artikel über London wird das Wort „London“ mit einem weiteren Artikel über „England“ verlinkt. Dieser ist nicht willkürlich verlinkt, sondern mit der Beziehung „ist die Hauptstadt von“. Somit weiß das System ganz genau, dass London die Hauptstadt von England ist und kann auf spezifische Fragen wie „Wie heißt die Hauptstadt von England?“ eine konkrete Antwort liefern.

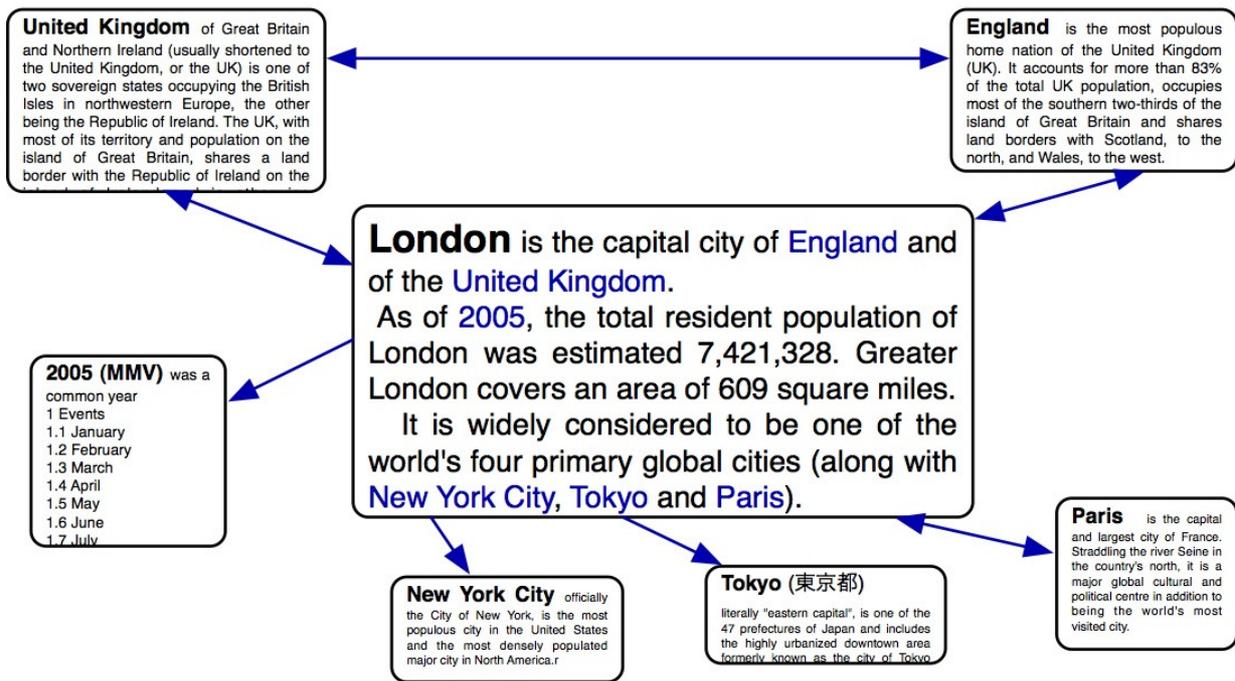


Abb. 12: Verlinkungen bei der Wikipedia [Völkel et al., 2006]

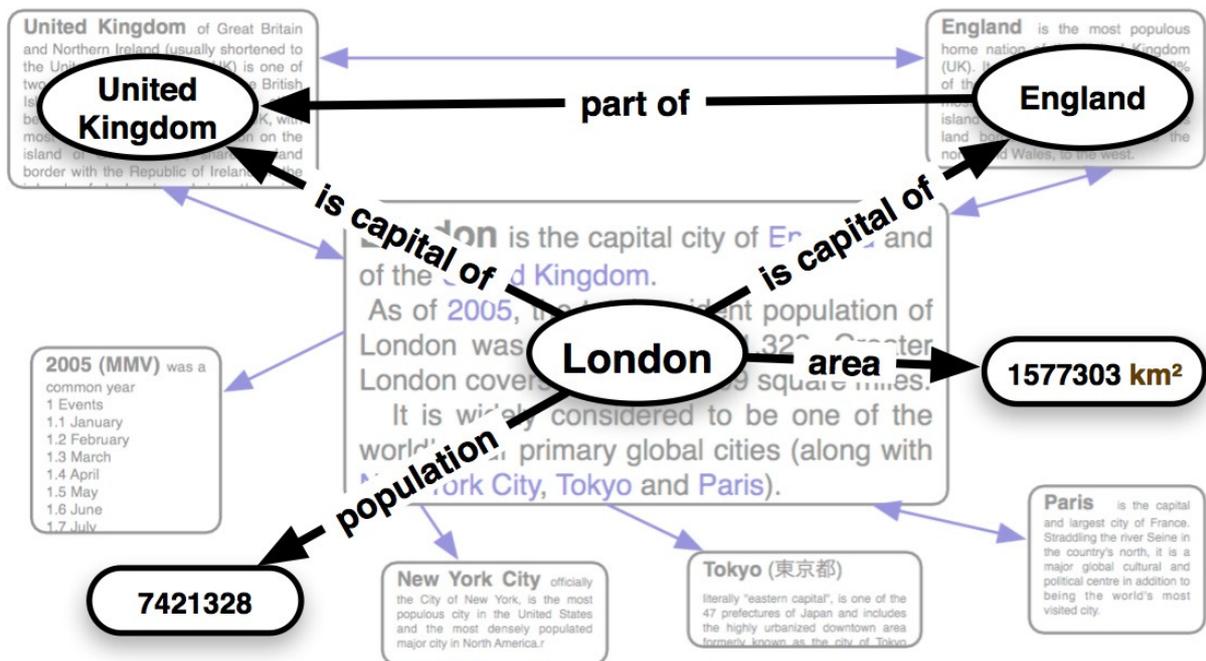


Abb. 13: Verlinkungen bei der Semantic-Wiki [Völkel et al., 2006]

In einer weiteren Initiative zur Erweiterung der Wikipedia mit semantischen Technologien wurde die Suchmaschine AskWiki²¹ entwickelt. Diese ist derzeit nur auf Englisch verfügbar und in der Entwicklungsphase. Sie agiert überwiegend als Ant-

²¹ <http://www.askwiki.com>

wortsuchmaschine und liefert konkrete Antworten auf spezifische Fragen. Die Genauigkeit der Suchergebnisse der Maschine wurde von Finin [2007] geprüft und einige Ergebnisse sind in Tabelle 1 dargestellt.

Frage	Antwort
1. Wie heißt die Hauptstadt von Belgien?	Brüssel
2. Wie viele Einwohner hat Brüssel?	1,975,000 Einwohner
3. Wie viele Einwohner hat die Hauptstadt von Belgien?	1,975,000 Einwohner
4. Was ist die Wurzel von 16?	Jahr 16 war ein Schaltjahr im Julianischen Kalender, der am Mittwoch begonnen hat.

Tab. 1: Beispielfragen und -antworten von AskWiki [Finin, 2007]

In Tabelle 1 wird erkennbar, dass die Suchmaschine einfache Fragen wie die Fragen Nr. 1 und 2 richtig beantwortet, sowie auch komplexere Fragen wie die Nr. 3 „Wie viele Einwohner hat die *Hauptstadt* von Belgien“. Bei der Frage Nr.4 „Was ist die Wurzel von 16“ hat sie eine falsche Antwort gegeben. Die AskWiki ist noch in der Entwicklungsphase und muss optimiert werden. Der Ansatz und die Erweiterung der Wikipedia mit Metadaten ermöglichen komplexere Suchanfragen und eine bessere Strukturierung der Inhalte [Finin, 2007].

Abschließend wird deutlich, dass die semantischen Wikis die bestehenden Wikis erweitern und unterstützen. Sie verbessern das laterale Denken und die Verknüpfungen zwischen den verschiedenen Wikis. Dadurch wird der Austausch von wertvoller Information zwischen den Benutzern ermöglicht [Tazzoli et al., 2004].

2.2.3 Social Software

Mit Social Software bezeichnet man „*all die Anwendungen, die Kommunikation, Interaktion und Zusammenarbeit im Internet unterstützen*“ [Brückmann, 2007, S. 1].

Social Software kann innerhalb von Unternehmen oder für private Zwecke genutzt werden. Unternehmen profitieren von Social Software dadurch, dass neue Möglichkeiten für das Knowledge Management und die Kollaboration zwischen den Mitarbeitern aus verschiedenen Unternehmensabteilungen entstehen. Durch Social

Software kann die interne und externe Kommunikation der Unternehmen durch Blogs, Wikis, Tagging und die Integration von externen RSS Feeds erweitert werden [Passant, 2008].

Semantic Web Technologien können des Weiteren als Mediator zwischen den verschiedenen Web 2.0-Applikationen und den Benutzern agieren [Passant, 2008]. In den folgenden Unterkapiteln werden drei Applikationen von Social Software erläutert, die mithilfe der semantischen Technologien optimiert werden können. Diese sind Social Networks, Blogs und Social-Bookmarking-Systeme.

2.2.3.1 Social Networks

Social Networks sind Beziehungsgeflechte aus Social-Network-Dienstleistungen, die von Teilnehmern mit gleichen Interessen genutzt werden. Die Benutzer kommunizieren untereinander und tauschen ihre persönlichen Daten aus [ITWissen, 2008].

Derzeit gibt es viele Online-Communities, die verschiedene Applikationen nutzen und untereinander nicht verlinkt sind. Die meisten Social Networks bearbeiten und teilen die Information zwischen den Mitgliedern durch einen zentralen Punkt. Wenn man eine bestimmte Information aus einem Netzwerk finden will, tut man sich schwer, weil die Suchanfragen normalerweise Keyword-basiert sind. Das heißt, jedes einzelne Wort wird gesucht, ohne auf den Zusammenhang und die Bedeutung der Wörter zu achten. Die Benutzer haben dadurch nur beschränkte Möglichkeiten um Informationen zu finden [Harth et al., 2004].

Durch Semantic-Web-Technologien kann man Social Networks bereichern und zugrundeliegende Information für Menschen und Maschinen verfügbar machen [Harth et al., 2004].

Ein Anwendungsbeispiel ist das ExpertFinder²²-System, mit dem man Experten aus einem bestimmten Gebiet finden kann. Ziel des ExpertFinders ist, RDF-Vokabulare und Regelerweiterungen zu vereinen, um persönliche oder Firmen-Homepages sowie auch Webseiten mit Publikationen oder akademische Konferenzen mit Metadaten zu versehen [Aleman-Meza et al., 2007].

²² <http://expertfinder.info>

Um das Projekt technisch umzusetzen, wurde ein Framework erstellt, welches auf existierenden Vokabularen (Industriestandards, bestehenden Klassifikationen und Taxonomien) basiert und diese kombiniert. Um dieses Framework zu erstellen, muss man bei allen Bereichen überprüfen, wie die Information in der Regel formalisiert wird, welche Ontologien es bereits gibt und ob Überlappungen ersichtlich sind. Auf Basis dieser Informationen kann man Richtlinien und Vorschläge aufbereiten, wie diese Vokabulare wiederverwendet und miteinander verbunden werden können, um eine optimale Suche nach Personen und Experten zu ermöglichen [Aleman-Meza et al., 2007].

Die Experten werden durch verschiedene Elemente beschrieben, wie zum Beispiel allgemeine Daten, Ausbildungsdaten, Qualifikationen, vergangene Projekte usw. (siehe Abbildung 14). Darüber hinaus werden auch die Beziehungen zwischen den Personen, Communities und Organisationen erfasst. Durch die semantisch angereicherten Daten wird eine genauere und komplexere Suchanfrage möglich [Aleman-Meza et al., 2007]. Wenn man ein Team für ein komplexes Projekt sucht, kann man nach bestimmten Qualifikationen suchen, nach vergangenen Projekten, die in diesem Bereich durchgeführt wurden und auch nach Personen, die bereits miteinander gearbeitet haben.

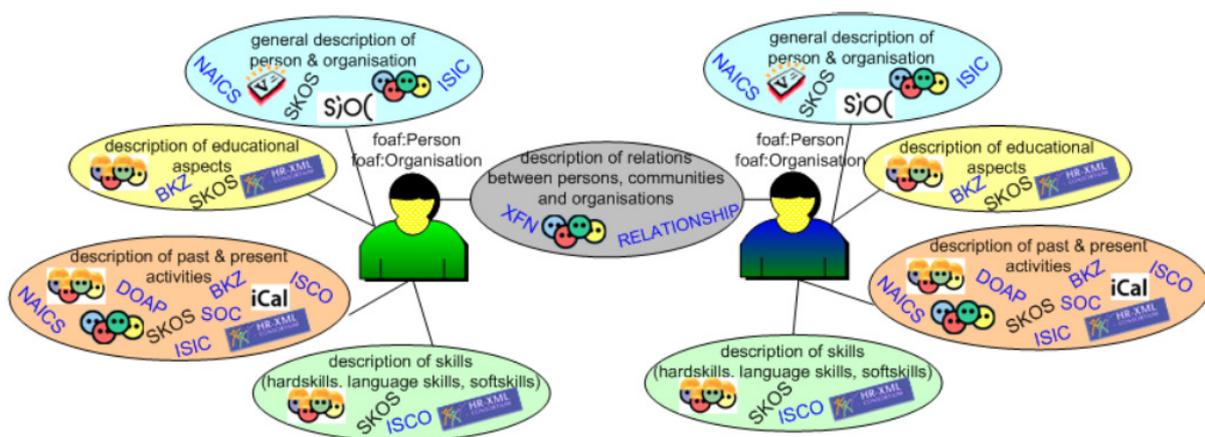


Abb. 14: Darstellung der semantischen Beziehungen beim ExpertFinder-System [Aleman-Meza et al., 2007]

Das System kann im Bereich Personalwesen eingesetzt werden. Mithilfe geeigneter RDF-Vokabulare können die Jobsuchenden ihr Profil in maschinenlesbarer Form im Internet publizieren. Es können auch Mitarbeiter einer Firma ihre Daten und Fähig-

keiten im Intranet der Firma speichern. Somit wird für Firmen und Jobvermittlungsagenturen die Suche nach geeignetem Personal erleichtert und die Ergebnisse können besser den Anforderungen der Suchenden entsprechen [Aleman-Meza et al., 2007].

2.2.3.2 Blogs

Ein Blog ist eine Art Web-Tagebuch, das im Internet geführt wird. Blogs enthalten Einträge, Bilder, Kommentare und Notizen, die chronologisch geordnet sind [ITWissen, 2008]. Die meisten Blogs werden derzeit manuell erstellt. Die Blog-Autoren (Blogger) schreiben den Text und suchen dann nach passenden Bildern und Links.

Der Unterschied zwischen den traditionellen und semantischen Methoden, um Blogeinträge zu schreiben, wurde von Ankolekar et al. [2007] anhand eines Beispiels dargestellt. In diesem Beispiel geht es um einen Blogger, der in seinem Blog über die Filme berichtet, die er im Kino angeschaut hat. Bei der traditionellen Methode fügt der Blogger zuerst den Titel des Films ein, schreibt dann den Text und addiert eventuell noch einige Tags dazu, die zum Beispiel das Genre des Filmes beschreiben.

Wenn der Blogger ein Plug-In mit semantischer Technologie nutzt, dann ändert sich der Prozess des Blogeintrags. Am Anfang muss er explizit festlegen, dass es sich um einen Filmeintrag handelt. Danach kann er den Titel des Films spezifizieren oder nach dem Filmmamen, durch Angabe der Darsteller oder des Regisseurs, suchen. Der zweite Schritt ist, den Bericht über den Film zu schreiben. Dafür wird vom Benutzer oder automatisch von der Maschine ein Sidebar erstellt (siehe Abbildung 15), in dem Felder wie Regie, Länge, Produzent, Erscheinungsjahr, Darsteller oder die URL zu der offiziellen Homepage dargestellt werden. Diese Sidebar wird einmal erstellt und kann in weiteren Blogeinträgen benutzt werden. Die Information, die in der Sidebar angezeigt wird, kann aus verschiedenen Quellen wie zum Beispiel semantische Wikis abgerufen werden (siehe Kapitel 2.2.2). Darüber hinaus kann der Blogger dynamische Information in seinen Blog integrieren, wie zum Beispiel die Filme, die derzeit in den Kinos gespielt werden, oder die Verfügbarkeit des Films in den Videotheken. Dies ist möglich, wenn verschiedene Online-Dienste wie die Kino-Webseiten

oder Veranstaltungs-Webseiten diese Information in RDF publizieren [Ankolekar et al., 2007].

Der Blogger erzielt durch den Einsatz von Semantic-Web-Technologien einen großen Nutzen. Er muss nur festlegen, aus welchen Webdiensten Daten abgerufen werden sollen und im Gegenzug bekommt er automatisch zusätzliche Daten, Bilder und Informationen, mit denen er seine Blogbeiträge anreichern kann [Ankolekar et al., 2007].

Archive

[June 2007](#)
[May 2007](#)
[April 2007](#)
[March 2007](#)

[About me](#)
[RSS-Feed](#)

Blogroll

[nutkidz](#)
[nakit-arts](#)
[Blog of the rings](#)
[Matrix Reblogged](#)

Links

[Gloria Cinema](#)
[Ecoshop](#)
[Legolas fanzine](#)

Pirates of the Caribbean 3

June 21st, 2007

I just went with [Till](#) into the last part of the Pirates of the Caribbean, where our heroes (the adoringly cute [Orlando Bloom](#) and [Keira Knightley](#) reprise their roles) go to the end of the world to save the one and only Captain Jack Sparrow ([Johnny Depp!](#) xOxOx!) from the claws of the Kraken. And guess what - Jack Sparrows daddy has a special appearance, played by old Rolling Stone [Keith Richards!](#) Weeeeha!

Best movie of the year, until know, without a question! Tons of fun, and colorful action.



Director [George Verbinski](#)
Running time 126 minutes
Starring [Johnny Depp](#), [Keira Knightley](#), [Bill Nighy](#), [Orlando Bloom](#), [Geoffrey Rush](#)
Info from [Wikipedia](#)

See [Pirates of the Caribbean 3](#) in the [Gloria](#)
Today 16:00, 18:30, 21:00
Tomorrow 16:00, 18:30, 21:00
[Reserve tickets now](#)

no comments yet – [post your comment](#) - [backtrack](#)

Abb. 15: Screenshot eines semantischen Blogbeitrags [Ankolekar et al., 2007]

Aus diesem Beispiel wird erkennbar, dass semantische Applikationen die Blogger bei der Suche nach passendem Kontext unterstützen. Semantische Technologien sollen es ermöglichen, relevante Bilder, Links und Artikel zu dem eigenen Blogbeitrag schneller und einfacher zu finden [Kilpatrick, 2008].

Ein Beispiel für eine semantische Applikation für Blogs ist Zemanta²³. Diese wird durch Natural-Language-Algorithmen unterstützt und beliefert die Blogger während des Schreibens mit wertvollen Inhalten wie Links, Tags, Dokumenten und Fotos [Zemanta, 2008]. Diese Applikation ist für Firefox²⁴ und Internet Explorer²⁵ erhältlich

²³ <http://www.zemanta.com>

²⁴ <http://www.mozilla.com/firefox>

²⁵ http://www.microsoft.com/internet_explorer

und arbeitet mit vielen Blogmaschinen wie WordPress²⁶, MovableType²⁷ and TypePad²⁸, Blogger²⁹ und Live Journal³⁰ [Kilpatrick, 2008]. Eine weitere semantische Applikation, das Toolkit Calais³¹, ermöglicht den Bloggern, Publishern und Entwicklern, ihre Texte oder Anwendungen automatisiert mit Metadaten anzureichern. Blogs und andere Anwendungen verarbeiten die Texte und beliefern diese mit semantischen Ergänzungen. Über APIs (Application Programming Interface) von Calais Web Services können Entwickler auf die Funktionen von Calais zugreifen. Sie schicken unstrukturierten Text und bekommen den Text annotiert mit RDF, Microformaten u.a. [Spath, 2008].

2.2.3.3 Social Bookmarking

Social Bookmarks sind „*Internet-Lesezeichen, die weltweit von vielen Benutzern indexiert, kommentiert und standortunabhängig abgefragt werden können*“ [ITWissen, 2008]. Es gibt Social-Bookmarking-Plattformen wie del.icio.us³², in denen die Benutzer Internetadressen (URLs) speichern und mit verschiedenen Lesezeichen wie Tags oder Kommentaren versehen können. Diese Social Bookmarks können für andere freigegeben werden [ITWissen, 2008]. Somit können Benutzer des Verzeichnisses Internetadressen mit Kommentaren von anderen Benutzern lesen und von deren Wissen profitieren.

Wie bei den Blogs oder Social Networks, herrscht bei den Social-Bookmarking-Plattformen auch das Problem, dass zu wenig Information und Wissen sinnvoll strukturiert ist und daher der Wissensaustausch zwischen den Plattformen noch sehr limitiert ist [Abel et al., 2007]. Um dieses Problem zu bewältigen, können semantische Technologien eingesetzt werden. Ein Beispiel, das dies bestätigt, wird in den folgenden Absätzen erklärt.

²⁶ <http://www.wordpress.org>

²⁷ <http://www.movabletype.org>

²⁸ <http://www.typepad.com>

²⁹ <http://www.blogger.com>

³⁰ <http://www.livejournal.com>

³¹ <http://www.opencalais.com>

³² <http://www.del.icio.us>

Die Initiative GroupMe!³³ erweitert die Idee des Social-Bookmarking-Systems mit einem Gruppensystem und ermöglicht es den Benutzern, eine Sammlung von Web-Ressourcen zu erstellen (wie zum Beispiel Fotos, Videos oder Webseiten), die ihren Interessen entsprechen. Diese Web-Ressourcen und die Gruppen-Aktivitäten werden mithilfe des Semantic Web in Beziehung gesetzt. GroupMe! bietet folgende Funktionalitäten: Erstens können die Benutzer Gruppen bilden und Webseiten zu diesen Gruppen addieren, bearbeiten und mit Stichwörtern taggen. Zweitens wird die Suche nach Web-Ressourcen innerhalb der Gruppen optimiert, indem das System nicht nur nach Tags sucht, sondern auch nach kontextueller Information der Web-Ressourcen. Drittens werden die Inhalte, die durch den Prozess der Gruppenbildung und Ressourcen-Organisation in den Gruppen gebildet werden, mit RDF-Annotationen versehen und für die Öffentlichkeit zugänglich gemacht [Abel et al., 2007].

In der Abbildung 16 werden die Funktionalitäten von GroupMe! bildlich dargestellt. Die angezeigte Gruppe wurde mit Jazz Legends getaggt und bei der Beschreibung werden die Künstler aufgelistet, die auf Videos aus YouTube³⁴ oder Fotos aus Flickr³⁵ angezeigt werden [Abel et al., 2007].

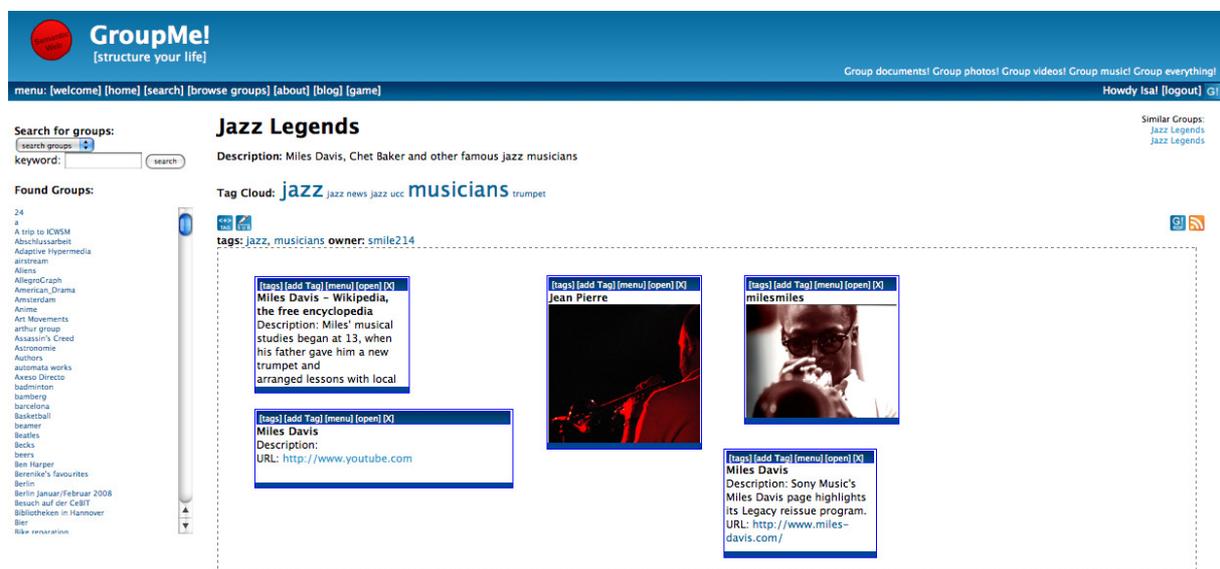


Abb. 16: Screenshot der Gruppe Jazz Legends bei der GroupMe!-Seite [GroupMe!, 2008]

³³ <http://www.groupme.com>

³⁴ <http://www.youtube.com>

³⁵ <http://www.flickr.com>

2.3 SWOT-Analyse des Semantic Web

Im folgenden Kapitel wird eine SWOT-Analyse dargestellt. Diese stellt die Stärken und Schwächen den Chancen und Risiken des Semantic Web gegenüber und wurde aufgrund von Literatur und Expertengesprächen erstellt, die im Anhang 6.2 vorliegen. Folgende Experten wurden vom Verfasser persönlich interviewt:

- Christian Siegfried Boser, Leiter des Fachbereichs „Semantic Web“ bei AUSTRIAPRO, Wirtschaftskammer Österreich³⁶ [Interv. Boser].
- Thomas Hochhaltinger, technischer Projektleiter bei Smart Information Systems³⁷ [Interv. Hochhaltinger].
- Markus Linder, Gründer und Managing Director bei Smart Information Systems [Interv. Linder].
- Felix Schramm, Concept Developer bei Universal Music Deutschland³⁸ [Interv. Schramm].
- Mario Szpuszta, Strategic Architect bei Microsoft Österreich³⁹ [Interv. Szpuszta].

2.3.1 Stärken

„Die Notwendigkeiten zum Einsatz von semantischen Technologien in der betrieblichen Anwendung sind offensichtlich.“ [John und Drescher, 2006, S. 247]

Durch die Steigerung der Komplexität der Dienstleistungs- und Warenwelt steigt auch der Bedarf nach einer besseren Beschreibung der Prozesse und Produkte. Semantische Technologien dienen zur Prozessunterstützung in Organisationen und erleichtern die Mensch-Maschine-Kommunikation. Dieser Informationsprozess der Verarbeitung der Information in maschinenlesbarer Form wird „Information Retrieval“ genannt [John und Drescher, 2006, S. 247].

³⁶ <http://www.austriapro.at>

³⁷ <http://www.smart-infosys.at>

³⁸ <http://www.universal-music.de>

³⁹ <http://www.microsoft.com/austria>

Das Semantic Web zeichnet sich durch folgende Stärken aus:

- **Offen zugängliche Daten**

Das „World Wide Web Consortium“⁴⁰ (W3C) und „Semantic Web Education and Outreach“⁴¹ (SWEO) unterstützen das „Open Data Movement“, das als Ziel hat, Daten für jeden frei verfügbar zu machen. Das Web soll durch das Semantic Web erweitert und die Daten im RDF-Format sollen für alle Plattformen zugänglich gemacht werden. Die RDF-Daten werden somit von beliebigen semantischen Suchmaschinen oder weiteren Applikationen erfasst [LinkingOpenData, 2008].

- **Integration von verschiedenen Formaten und Datenquellen**

RDF-Vokabulare sind deskriptive Elemente, wie zum Beispiel Dienstleistung, Buch, Bild, Titel, Beschreibung, Rechte usw., die nützlich sind, um Informationen umzucodieren. Damit kann man verschiedene Daten effektiv nutzen, integrieren und aggregieren [Berners-Lee und Miller, 2002].

Wie im Kapitel 2.2.1 beschrieben, kann man mittels Semantic-Web-Annotationen nicht nur aus den Passagen von Dokumenten im offenen Web suchen, sondern auch aus „versteckten Inhalten“, wie zum Beispiel Webdiensten (Routenplanung, Wetterinformation, Veranstaltungskalender, Verkehrsinformation, Webcam-Suche) und aus PDFs, die mittels Bildextraktion durchsucht werden [Wahlster, 2007a]. Ebenso kann man in RDF viele unterschiedliche Datenressourcen repräsentieren. Der Browser Mozilla⁴² benutzt für die Darstellung der Bookmarks, History, Suchergebnisse, E-Mail usw. ein gemeinsames RDF-Format [Andersen, 2003].

Das Semantic Web ermöglicht es, die Daten so darzustellen, dass ein Programm die Formatierung, Bilder oder Werbung einer Webseite nicht entfernen muss, um die relevanten Teile einer Information zu finden. Andererseits kann man im Semantic Web Dateien generieren, die die Beziehung zwischen verschiedenen Datenbeständen erklären. Die RDF-Daten können verlinkt werden

⁴⁰ <http://www.w3.org>

⁴¹ <http://www.w3.org/2001/sw/sweo>

⁴² <http://www.mozilla.org>

und ermöglichen dem Internet-Benutzer, im Semantic-Web-Browser von einem Datenelement in einer Datenquelle zu verwandten Datenelementen in anderen Datenquellen zu navigieren. Zum Beispiel kann ein „semantischer Link“ zwischen einer Datenbank mit einer Spalte mit Postleitzahlen und einem Online-Formular darstellen, dass sie dasselbe beschreiben. Dadurch wird es Maschinen erleichtert, die Beziehungen zwischen den Datenquellen zu erfassen und Daten aus verschiedenen Quellen zu integrieren [Berners-Lee und Miller, 2002].

- **Zusammenhänge werden verstanden**

Semantische Technologien ermöglichen ein schnelles „Verständnis“ von komplexen Zusammenhängen. Die Merkfähigkeit der Zusammenhänge wird gesteigert und individuelle Lernpfade sind möglich. Durch die Offenlegung der bereichsübergreifenden Zusammenhänge wird die interdisziplinäre Kommunikation in Unternehmen unterstützt und das Ideenmanagement gefördert [Blumauer und Fundneider, 2006, S. 232].

- **Optimierung des betrieblichen Wissensmanagement**

Semantische Technologien können Wissens- und Geschäftsprozesse innerhalb von Unternehmen optimieren [John und Drescher, 2006]. Durch die Wissensnetze, die im Semantic Web entstehen, wird das Komplexitätsmanagement optimiert und die Informationsflüsse werden effektiver [Blumauer und Fundneider, 2006, S. 232]. Nach John und Drescher [2006] ermöglicht diese Technologie die Analyse und Extraktion von Daten aus unterschiedlichen Dokumenten und Datenbanken der Unternehmen. Die Bedeutung der Daten wird durch die Maschinen erfasst und Zusammenhänge und Schlussfolgerungen können abgeleitet werden.

In der Boeing Corporation⁴³ wurden zum Beispiel die Datenbanken „Flugzeugentwurf“ und „Flugzeug-Wartung“ in RDF beschrieben. Mit dem RDF-Schema kann man Zielfragen formulieren, die semantisch übersetzt werden und zu besseren Suchergebnissen führen [Andersen, 2003].

⁴³ <http://www.boeing.com>

- **Intelligente Anwendungen**

Das Semantic Web wird Struktur für die bedeutungsvollen Inhalte der Webseiten liefern, um ein Umfeld zu schaffen, in dem intelligente Agenten verschiedene Webseiten durchsuchen und komplexe Aufgaben für die Benutzer schnell lösen. Zum Beispiel: Wenn jemand krank ist und einen Arzt braucht, dann sucht der intelligente Agent nach Krankenhäusern in der Nähe des Wohnortes des Benutzers, die seine Krankheit behandeln können. Dieser Agent sucht in der Webseite des Krankenhauses nicht nur nach den Schlüsselwörtern "Behandlung, Medizin, physische, Therapie", wie es bei den aktuellen Agenten der Fall ist, sondern er weiß auch, dass Dr. Hartman in diesem „Krankenhaus“ am „Montag, Mittwoch und Freitag“ „arbeitet“ und dass das Skript einen Datumsbereich im „Format tt.mm.jjjj“ annimmt und „Arzttermine“ zurückliefert [Berners-Lee et al., 2001].

- **Personalisierung der Suchanfragen**

Durch semantische Suchmaschinen kann man Suchanfragen den Bedürfnissen der Anwender anpassen und höhere Trefferquoten erreichen. Man kann auch bei den Suchanfragen Fragen stellen und erhält eine konkrete Antwort anstatt einer langen Link-Liste. Diese Suchanfragen kann man in den Unternehmensbereichen Personalmanagement, Teambildung und Kompetenzmanagement einsetzen. Durch die Wissensnetze und Ontologien des Semantic Web werden die passenden Ansprechpartner und Wissenslücken in Unternehmen identifiziert [Blumauer und Fundneider, 2006, S. 232].

Weitere Stärken des Semantic Web sind im E-Commerce-Bereich zu sehen, wie zum Beispiel:

- **Effiziente Platzierung der Angebote**

Im Bereich des E-Commerce bietet das Semantic Web Vorteile für Benutzer und Unternehmen. Die Unternehmen können ihre Angebote einfach und effizient platzieren. Die Benutzer bekommen Suchergebnisse, die an ihre Wünsche und Bedürfnisse angepasst sind [Interv. Linder].

- **Mechanische Unterstützung bei der Suche nach Verkäufern und Angeboten**

Derzeit wird die Suche nach E-Commerce-Plattformen oder nach Angeboten von den Internet-Benutzern manuell getätigt. Dadurch wird die Skalierbarkeit der Suchprozesse beschränkt. Mit Hilfe semantischer Technologien kann man diesen Prozess automatisieren [Fensel et al., 2002]. Nach Andersen [2003] können semantische Technologien verschiedene Prozesse der B2B (Business to Business)-Plattformen, wie zum Beispiel die Suche nach passenden Anbietern, das Abschließen von Geschäften oder die Durchführung von Transaktionen, automatisieren und beschleunigen. Dieses wird durch semantische Beschreibungssprachen wie DAML+OIL⁴⁴ (DARPA Agent Markup Language + Ontology Inference Layer) oder RDF erreicht, welche die komplexen Beschreibungen über die Angebote der Verkäufer und die Anforderungen der Käufer für Maschinen verständlich machen.

- **Einheitliche Standards**

Im Web gibt es derzeit eine große Anzahl verschiedener Standards, um Produkte, Dienstleistungen, Produktkataloge und Geschäftsdokumente zu beschreiben. Mit dem Semantic Web kann man die große Anzahl an heterogenen Formaten, die im Web sind, besser koordinieren. Ontologien ermöglichen eine bessere Definition und Abbildung der Standards. Durch diese Technologie können die verschiedenen Terminologien effizienter verarbeitet werden und dadurch die Skalierbarkeit erhöht werden [Fensel et al., 2002].

2.3.2 Schwächen

Das Semantic Web weist für Benutzer und Entwickler noch geringe Attraktivität auf, da das Angebot an semantisch aufbereiteten Inhalten nicht sehr umfassend ist (siehe Abbildung 17). Dadurch sinkt der Anreiz der Softwarehersteller, benutzerfreundliche Tools zu entwickeln, um Ontologien und Annotationen zu erstellen, die die semantischen Daten strukturiert darstellen. Wenn es keine Ontologien gibt, dann werden die

⁴⁴ <http://www.w3.org/TR/daml+oil-reference>

semantischen Inhalte im Web auch nicht zunehmen. Diese Tatsache kann man durch das Phänomen der Netzwerkexternalität erklären: Bei positiven Netzwerkexternalitäten wächst der Nutzen an einem Standard oder Netzwerk, wenn dessen Nutzerzahl größer wird. Durch diesen erhöhten Nutzen wird das Netzwerk für noch mehr Personen interessant. Dadurch wächst die Nutzerzahl weiter an und somit steigt wiederum der Nutzen für alle. Übertragen auf das Semantic Web heißt dies, dass je häufiger Inhalte und Services im Web mittels akzeptierter Ontologien beschrieben und verfügbar gemacht werden, desto sinnvoller wird es, sich selbst an diese Standards zu halten und entsprechend Inhalte zu publizieren. Nur dann können Suchmaschinen, Software-Agenten und verschiedene Applikationen geschaffen werden, die die Wiederverwendbarkeit der Inhalte erhöhen und den Nutzen für alle Benutzer erweitern. Hiermit wird ersichtlich, dass der Erstellungsaufwand von Ontologien sich erst dann lohnt, wenn es auch ein ausreichendes Angebot an für Semantic Web tauglichen Inhalten gibt [Blumauer und Fundneider, 2006, S. 227-229].



Abb. 17: Eintrittsbarrieren des Semantic Web [Blumauer und Fundneider, 2006, S. 228]

Des Weiteren weist das Semantic Web folgende Schwächen auf:

- **Sinkende Anonymität und Vertraulichkeit**

Durch die eindeutige Referenzierung der Inhalte können Semantik-basierte Suchmaschinen zuverlässig die gesuchten Inhalte auffinden und liefern. Das

bedeutet aber, dass diese Maschinen den Anwender und sein Verhalten kennen und dadurch die Anonymität der Benutzer verloren geht [Dostal et al., 2004]. Dazu meint Boser [Interv. Boser]: *„Die Speicherung meines detaillierten Suchprofils beinhaltet unter Umständen viele Informationen zu meiner Person, wie zum Beispiel, „Christian Boser macht Urlaub in Kärnten, möchte vegetarisch essen, und hat zwei Kinder und einen Hund mit“. Auch wenn die Verfügbarkeit dieser Information nicht wirklich bedrohlich klingt, muss das Thema der Privatsphäre und der Schutz personenbezogener Daten stärker thematisiert werden. Hier können klare Regeln für den Umgang mit Daten aus strukturierten Suchabfragen Hilfe bieten.“*

- **Sinkende Verlässlichkeit und Missbrauch der Semantikbeschreibung**

Bei vielen semantischen Applikationen wird der Entscheidungsprozess an Maschinen übergeben. Diese entscheiden, welche Inhalte „relevant“ und „interessant“ sind. Wenn man diesen Prozess kennt, dann ist das System auch angreifbar. Man könnte somit Inhalte mit falschen Beschreibungen versehen, damit sie die semantischen Filterprozesse bestehen und öfter gefunden werden [Dostal et al., 2004].

Schramm [Interv. Schramm] und Szpuszta [Interv. Szpuszta] unterstützen die Meinung von Dostal et al. [2004] und sehen im Semantic Web die Gefahr, dass Webseitenbetreiber oder Hersteller die semantische Beschreibung missbrauchen können, um ihre Inhalte und Angebote im Web prominenter zu platzieren oder um besser gefunden zu werden.

Schramm fügt hinzu: *„Die Entscheidungskompetenz, welche Information auf welche Frage zutrifft, wird im Semantic Web an die Maschine ausgelagert. Wenn die Personen den Interpretationsprozess der Maschinen verstehen, wird dieses System angreifbar und kann missbraucht werden. Es entsteht daher die Gefahr, dass Inhalte mit falschen Semantik-Beschreibungen/Metadaten beschrieben werden und dadurch den semantischen Filterprozess umgehen. Das könnte zu Formen des Spamming im semantischen Netz führen.“* [Interv. Schramm]

- Sinkende Ergebnisbreite

Viele Ergebnisse einer Web-Recherche werden durch Zufall im Web gefunden. Man sucht nach einem Thema und findet bei der Suche nebenbei auch interessante, relevante Daten, nach denen man nicht direkt gesucht hat. Durch die automatisierte Suche des Semantic Web werden diese „nebenbei angefundene Daten“ eliminiert und der filternden menschlichen Interpretation entzogen [Dostal et al., 2004]. Zu diesem Thema fügt Schramm [Interv. Schramm] hinzu: *„Dadurch, dass Informationen so spezifisch und punktgenau ausgeliefert werden, werden andere Angebote unter Umständen der menschlichen Interpretation entzogen. Bei Google gibt man die Wörter „Semantic Search“ ein und man bekommt 100 Suchergebnisse. Manchmal sind auch Sachen dabei, die gar nicht mit der Frage zu tun haben. Aber trotzdem können diese Links Personen mehr inspirieren, als wenn eine spezifische Antwort angezeigt wird. Das sehe ich als ein generelles Problem des Semantic Web.“*

2.3.3 Chancen

Unternehmen produzieren, teilen, kombinieren und verarbeiten eine enorme Menge an Daten und Dokumenten. Die Anwendungsgebiete dieser Daten sind vielfältig und können von einfachen Transaktions-Daten bis zu komplexen Geschäfts-Prozessdefinitionen reichen. Da Unternehmen sich in unterschiedlichen Teilen der Wertschöpfungskette befinden, werden auch die Daten in verschiedenen Kontexten und Prozessen genutzt. Dadurch können widersprüchliche Anforderungen für die Repräsentation solcher Daten entstehen [Hepp et al., 2007].

Eine Strukturierung und Standardisierung von Daten in der Geschäftswelt wird derzeit dringend gebraucht. Erstens laufen die meisten Unternehmensprozesse kollaborativ und dezentralisiert ab. Darüber hinaus steigen der Arbeitskraftanteil und die spezifischen Teile und Materialien, die von Unternehmen genutzt werden. Weiters sind die Geschäftspartner in der heutigen globalen Welt geographisch verstreut und man will Prozesse optimieren, damit die Ambiguitäten und Widersprüche, die bei einer schlechten Daten-Strukturierung entstehen können, vermieden werden [Hepp et al., 2007]. Christian Boser [Interv. Boser] meint, dass die Daten aus unterschiedli-

chen Plattformen, die im Internet verstreut vorliegen, durch das Semantic Web in einer Anwendung aggregiert werden. Zum Beispiel, wenn man einen Urlaub plant, sollte es mit semantischen Technologien nicht mehr notwendig sein, auf verschiedenen Webseiten einen Flug, eine Unterkunft und Kulturangebote vor Ort zu suchen, weil die semantische Suchmaschine die Inhalte aus verschiedenen Anwendungen aggregiert.

Schramm [Interv. Schramm] sieht im Semantic Web eine Chance, um *„den Datenüberfluss im Web zu bewältigen. In einer Zeit, in der sich Nutzerbedürfnisse immer mehr ausdifferenzieren, stellen zielgenaue Angebote eine große Chance dar. Durch die konstante Weiterentwicklung von Nutzerbedürfnissen lösen sich auch die Personen aus den klassischen Milieus oder Subkulturen und springen zwischen den verschiedenen Bereichen hin und her.“*

Wie im Kapitel 2.3.2 dargestellt, liegt eine Schwäche des Semantic Web darin, dass es derzeit wenig Ontologien gibt, die Produkte und Dienstleistungen beschreiben. Um dieses Problem zu bewältigen, muss man Werkzeuge und Systeme schaffen, die den Prozess der Erstellung von Ontologien vereinfachen [Hepp, 2005].

Ein Beispiel für die Entwicklung von solchen Werkzeugen ist die kollaborative Ontologie-Entwicklungsplattform myOntology [myOnt]. Diese basiert auf Wiki-Technologien und ermöglicht eine dynamische und zeitnahe Weiterentwicklung von Produktontologien. Durch die enge Kooperation mit internationalen Standardisierungsorganisationen und Experten aus verschiedenen Branchen kann man Ontologien schaffen, die die nötigen Informationen über die Produkte, Produkteigenschaften und Anbieter darstellen können [ebSem AUSTRIAPRO, 2007]. Markus Linder von Smart Information Systems meint, dass solche Initiativen nötig sind, damit sich das Semantic Web durchsetzt. *„Denn letztendlich braucht man Werkzeuge, Mittel und Wege, die es erleichtern, gemeinsam Ontologien zu entwickeln“*, meint Linder [Interv. Linder].

Die Risiken des Semantic Web, wie zum Beispiel die Authentifizierungsprobleme der Datenquellen, können durch Trustworthy-Systeme [Thuraisingham, 2007] überwältigt werden. Eine Initiative, um solche Systeme zu erreichen, ist die Erweiterung des Firmenverzeichnisses von der Wirtschaftskammer Österreich [Interv. Boser]. Durch

diese Initiative haben Unternehmen die Möglichkeit, zusätzlich zu den Firmenstammdaten semantisch annotierte Informationen (RDF) zu registrieren, um diese für semantische Suchmaschinen auffindbar zu machen. Wo derzeit schon alle österreichischen Unternehmen mit ihren Stammdaten gelistet sind, wird in Zukunft zum Beispiel jeder einzelne Tourismus- und Freizeitbetrieb sein Angebot zusätzlich auf sehr einfache, aber strukturierte Weise beschreiben können [WKÖ, 2008]. Das Trustworthy-System wird dadurch gewährleistet, dass nur die im Firmenbuch registrierten Unternehmen ihre Daten und Angebote eintragen und ändern können [Interv. Boser].

Dieses Verzeichnis [WKÖ 2008] ermöglicht unter anderem die:

- Eingabe von Links zu RDF-Daten der Unternehmen.
- Ausgabe von gesicherten Informationen über Unternehmen in RDF.
- Bereitstellung der Unternehmensinformationen in Maschinen-interpretierbarer Form für alle Anwendungen, die auf Semantic-Web-Technologien basieren.

Mario Szpuszta [Interv. Szpuszta], Strategic Architect bei Microsoft Österreich, meint, dass die Begriffe des Semantic Web, wie zum Beispiel Ontologien und RDF, für die breite Masse verständlich gemacht werden sollen, damit das Semantic Web in der Industrie und in der Praxis eine stärkere Chance bekommt. Des Weiteren meint er: *„Man muss das Ganze auf bodenständige Probleme herunterbrechen, die die Leute im Business draußen haben. Man muss denen zeigen, wie diese Probleme oder Teile dieser Probleme über das Semantic Web leichter gelöst werden können. Das ist ein Problem, mit dem nicht nur Semantic-Web-Leute zu kämpfen haben, mit dem haben auch Plattformen und Hersteller wie Microsoft, IBM oder Yahoo zu kämpfen.“* [Interv. Szpuszta]

Dr. Juhnyoung Lee vom IBM Thomas J. Watson Research Center meint, dass die Industrie eine Vorreiterrolle bei der Forschung über Daten-Standardisierung hat und dass Partnerschaften mit der Industrie, Forschungsinstituten und staatlichen Organisationen entscheidend sein werden, damit die Dienstleistungs-Wirtschaft in der Zukunft vorankommt [Hepp et al., 2007]. Durch die oben genannten Initiativen zur Erstellung von Ontologien und Authentifizierung der Daten wird ersichtlich, dass es

Bewegungen gibt, die das Semantic Web unterstützen und damit die Erfolgschancen sichern.

2.3.4 Gefahren

Christian Boser [Interv. Boser] und Felix Schramm [Interv. Schramm] sind sich darüber einig, dass die Datenschutzproblematik im Semantic Web vermehrt beachtet werden soll. Schramm: *„Es ist zwar interessant, den Nutzer zu profilieren und Benutzerprofile zu erstellen, aber in dem Moment, wo die Anonymität des Nutzers nicht mehr gewährleistet ist, halte ich es doch für sehr problematisch. Da muss man aufpassen. Gerade Deutschland ist im Thema Datenschutz glücklicherweise relativ weit vorne - im Gegensatz zu einigen Ländern im asiatischen Raum. Die Datenschutzproblematik sollte man jedenfalls nicht aus dem Auge verlieren.“* [Interv. Schramm]

Die deutsche Regierung hat bereits Gesetzesentwürfe gestartet, um den Datenschutz zu verschärfen. In Zukunft werden die Daten nur mit ausdrücklicher Einwilligung der Betroffenen weitergegeben. Eine Kennzeichnungspflicht in Bezug auf die Herkunft der Daten soll eingeführt werden, damit die Benutzer Bescheid wissen, wie und woher Unternehmen ihre Daten bezogen haben [Focus Online, 2008].

Laut Mario Szpuszta, Microsoft Österreich, liegt das Risiko des Semantic Web darin, *„dass die Ontologien zu vielfältig und zu umfangreich werden; Informationen können nicht mehr zielgerichtet gefiltert werden. Damit genau das nicht passiert und die Ontologien in einem gesunden Rahmen gehalten werden, müssen sich alle Hersteller von zum Beispiel Digitalkameras auf allgemeine Charakteristika von Kameras einigen, mit denen man Informationen sinnvoll verknüpfen kann. Das ist extrem aufwendig. Und wenn man alle Märkte der Welt in Betracht zieht, dann bin ich mir absolut sicher, dass es Märkte und Branchen gibt, die man damit nicht erfassen kann.“* [Interv. Szpuszta]

Im technischen Bereich hat das Semantic Web folgende Herausforderungen zu bewältigen:

- **Standardisierung von Ontologien [Lu et al., 2002]**

Integration und Interoperabilität sind bei jedem offenen System wichtig, somit auch beim Semantic Web. Es gibt viele Vorschläge für Ontologie-Repräsentations-Sprachen, aber es gibt keine Einigung, welches die Standard-Ontologie-Repräsentationssprache sein soll. Solange diese nicht festgelegt wird, stellt dies eine große Hürde für die Durchsetzung des Semantic Web dar.

- **Formelle Semantic-Web-Sprachen**

RDF wurde anfänglich genutzt, um Metadaten zu definieren und zu benutzen. Danach wurde RDF mit RDFS ergänzt. Forscher haben erkannt, dass es RDFS an formeller Semantic mangelt, und entwickelten Vorschläge, die RDFS-Architektur mithilfe von UML⁴⁵ (Unified Modeling Language) zu definieren. Um die HTML-Webseiten einfacher in semantischer Sprache zu annotieren, wurde der RDFa⁴⁶ (Resource Description Framework attributes) entwickelt [Adida und Birbeck, 2008].

Je mehr Webseiten ihre Inhalte in formellen Semantic-Web-Sprachen annotieren, desto größer ist der Nutzen des Semantic Web. Daraus wird ersichtlich, dass eine Standard-Sprache des Semantic Web unbedingt notwendig ist [Lu et al., 2002].

- **Entwicklung von „Trust“- und „Proof“-Modellen**

Als ein offenes und dezentralisiertes System unterstützt das Semantic Web die Idee, dass „jeder alles mögliche über jeden sagen kann“. Daher können Personen auf der ganzen Welt Statements veröffentlichen. Manche dieser Statements können eventuell zu Konflikten führen. Daher muss jemand sicherstellen, dass das Statement von der Originalquelle stimmt (Proof) und dass diese Quelle vertrauenswürdig ist (Trust) [Lu et al., 2002].

Beim Proof ist das Ziel nachzuvollziehen, wie es zu einem Ergebnis gekommen ist. Im Semantic Web sollten „Proofs“ ausgetauscht werden, die in „unifying

⁴⁵ <http://www.uml.org>

⁴⁶ <http://www.w3.org/2006/07/SWD/RDFa>

language“ beschrieben sind. Diese Programmiersprache drückt die logischen Rückschlüsse aus, die mittels verschiedener Regeln ausgedrückt werden. Zum Beispiel, die Daten von Frau Cook wurden von einer Online-Dienstleistung gefunden und haben sie in Johannesburg platziert. Wenn der Benutzer glaubt, dass diese Information richtig ist, schickt der Computer dem Dienstleister eine Anforderung, diese Information zu prüfen. Das „proof“ wird geliefert, indem die interne Logik in die „unifying language“ des Semantic Web übersetzt wird. Eine Inferenz-Maschine verifiziert die Information, dass Frau Cook in Johannesburg lebt, und zeigt auch weitere Webseiten, die das bestätigen [Berners-Lee et al., 2001]. Des Weiteren spielt die digitale Signatur eine wichtige Rolle im „Proof“. Die Quelle muss ihre Statements codieren, so dass die semantischen Agenten überprüfen können, dass die Information von dieser spezifischen, vertraulichen Quelle stammt [Lu et al., 2002, Berners-Lee, 2001]. Außerdem kann man auch weitere Sicherheits-Technologien, wie zum Beispiel Verschlüsselungen und Zugriffskontrolle, nutzen, um die Vertraulichkeit der Information zu garantieren [Lu et al., 2002].

Bezüglich des „Trusts“ ist es für jeden Menschen möglich, ein eigenes Vertrauensmodell aufzubauen und festzulegen, wie viel Vertrauen man in jede einzelne Quelle des Semantic Web hat. Da es unrealistisch ist, das Vertrauensmaß für jede Quelle festzulegen, braucht man einen Mechanismus, der den Grad an „Trust“ ableitet. Eine Lösung ist das „Web of Trust“, das heißt, wenn jemand der Quelle A vertraut, dann vertraut er auch anderen Quellen, denen von Quelle A vertraut wird, aber in einem kleineren Maße. Somit bildet man ein großes, hierarchisches Netzwerk, das den Agenten hilft, Information zu finden, die von vielen Personen als vertrauenswürdig angesehen wird [Lu et al., 2002].

Der Gedanke von „Proof“ und „Trust“ ist noch nicht formalisiert und muss für das Semantic Web entwickelt werden, da diese Technologie die Grundlage für „gute“ und vertrauenswürdige E-Commerce-Applikationen ist [Lu et al., 2002].

2.4 Bewertung der Nutzenpotentiale und Ausblick anhand von Experten

Im Rahmen der vorliegenden Arbeit wurden zehn Experten anhand eines Fragebogens (siehe Kapitel 2.4.1) nach dem Nutzen und der Zukunft der semantischen Technologien befragt.

Die zu beantwortenden Fragen beziehen sich auf die im Kapitel 2.1 und 2.2 genannten Anwendungspotentiale der Semantic-Web-Technologie und auch auf die Branchen, in denen das Semantic Web sinnvoll eingesetzt werden kann. Des Weiteren werden die Experten nach dem Aufwand und Nutzen der semantischen Technologie befragt und ob sie denken, dass sich diese in den nächsten 5 Jahren durchsetzen wird.

Das Ziel dieser Befragung ist es nicht, eine repräsentative Aussage für die ganze Gesellschaft zu machen, sondern Tendenzen anhand von Experten zu erfassen, die sich mit diesem Thema beschäftigen.

Zusätzlich zu den im Kapitel 2.3 genannten Experten Boser [Interv. Boser], Hochhaltinger [Interv. Hochhaltinger], Linder [Interv. Linder], Schramm [Interv. Schramm] und Szpuszta [Interv. Szpuszta] wurden fünf weitere Experten vom Verfasser telefonisch befragt. Diese sind Bortecin Ege [Interv. Ege], Moderator der Xing-Community „Semantic Web“⁴⁷; Christoph Möller [Interv. Möller], Trend Analyst bei TrendONE⁴⁸, ein deutsches Trendforschungsunternehmen, welches im Juli 2008 eine Studie über Semantic Web publiziert hat [TrendONE, 2008]; Alexander Oelling [Interv. Oelling], Gründer der Firma sones⁴⁹, die verschiedene Produkte anbietet, die auf semantischen Technologien basieren; Julian Rothkamp [Interv. Rothkamp], Junior Manager New Business bei Universal Music Group Deutschland⁵⁰, und Jan Voss [Interv. Voss], Leiter der New Business-Abteilung bei Universal Music Group Deutschland. Die Befragung wurde zusätzlich zum Fragebogen mündlich geführt, damit die Begründungen für die jeweilige Auswahl erfasst werden können.

⁴⁷ <https://www.xing.com/net/nextgenerationweb>

⁴⁸ <http://www.trendone.de>

⁴⁹ <http://www.sones.de>

⁵⁰ <http://www.universal-music.de>

2.4.1 Fragebogen

Der vorliegende Fragebogen wurde für die Experten-Interviews verwendet.

Wie schätzen Sie den Nutzen von Semantic-Web-Technologien ein?				
Für Unternehmen				
<input type="radio"/> 1 (sehr hoch)	<input type="radio"/> 2 (niedrig)	<input type="radio"/> 3 (mittel)	<input type="radio"/> 4 (niedrig)	<input type="radio"/> 5 (sehr niedrig)
Für Benutzer				
<input type="radio"/> 1 (sehr hoch)	<input type="radio"/> 2 (niedrig)	<input type="radio"/> 3 (mittel)	<input type="radio"/> 4 (niedrig)	<input type="radio"/> 5 (sehr niedrig)
Wie hoch schätzen Sie den Aufwand einer Implementierung von Semantic-Web-Technologien?				
<input type="radio"/> 1 (sehr hoch)	<input type="radio"/> 2 (niedrig)	<input type="radio"/> 3 (mittel)	<input type="radio"/> 4 (niedrig)	<input type="radio"/> 5 (sehr niedrig)
Lohnt sich der Aufwand in Anbetracht des Nutzens? (Kreuzen Sie bitte die richtige Antwort an.)				
<input type="radio"/> JA	<input type="radio"/> NEIN			
Von welchen Faktoren hängt es ab, dass das Semantic Web sich durchsetzt?				
Denken Sie, dass sich das Semantic Web in den nächsten 5 Jahren durchsetzen wird? (Kreuzen Sie bitte die richtige Antwort an.)				
<input type="radio"/> JA	<input type="radio"/> NEIN			
In welchen Branchen macht das Semantic Web Sinn bzw. lohnt es sich, Semantic-Web-Technologien einzusetzen?				
- Bildungswesen:				
<input type="radio"/> 1 (sehr hoch)	<input type="radio"/> 2 (niedrig)	<input type="radio"/> 3 (mittel)	<input type="radio"/> 4 (niedrig)	<input type="radio"/> 5 (sehr niedrig)
- Musikindustrie:				
<input type="radio"/> 1 (sehr hoch)	<input type="radio"/> 2 (niedrig)	<input type="radio"/> 3 (mittel)	<input type="radio"/> 4 (niedrig)	<input type="radio"/> 5 (sehr niedrig)
- Öffentlicher Bereich/Behörden/Regierung:				
<input type="radio"/> 1 (sehr hoch)	<input type="radio"/> 2 (niedrig)	<input type="radio"/> 3 (mittel)	<input type="radio"/> 4 (niedrig)	<input type="radio"/> 5 (sehr niedrig)
- Finanzbereich:				
<input type="radio"/> 1 (sehr hoch)	<input type="radio"/> 2 (niedrig)	<input type="radio"/> 3 (mittel)	<input type="radio"/> 4 (niedrig)	<input type="radio"/> 5 (sehr niedrig)
- Tourismus:				
<input type="radio"/> 1 (sehr hoch)	<input type="radio"/> 2 (niedrig)	<input type="radio"/> 3 (mittel)	<input type="radio"/> 4 (niedrig)	<input type="radio"/> 5 (sehr niedrig)
- Unterhaltungselektronik:				
<input type="radio"/> 1 (sehr hoch)	<input type="radio"/> 2 (niedrig)	<input type="radio"/> 3 (mittel)	<input type="radio"/> 4 (niedrig)	<input type="radio"/> 5 (sehr niedrig)
- Verlage/Medien:				
<input type="radio"/> 1 (sehr hoch)	<input type="radio"/> 2 (niedrig)	<input type="radio"/> 3 (mittel)	<input type="radio"/> 4 (niedrig)	<input type="radio"/> 5 (sehr niedrig)

Wie hoch schätzen Sie die Auswirkung des Semantic Web auf Wissensmanagement und E-Commerce?				
- E-Commerce:				
O 1 (sehr hoch)	O 2 (niedrig)	O 3 (mittel)	O 4 (niedrig)	O 3 (sehr niedrig)
- Wissensmanagement:				
O 1 (sehr hoch)	O 2 (niedrig)	O 3 (mittel)	O 4 (niedrig)	O 3 (sehr niedrig)
Wie bewerten Sie den Nutzen des Semantic Web (in Anbetracht des Aufwands) in folgenden Anwendungsgebieten?				
- Empfehlungssysteme/Recommender-Systeme				
O 1 (sehr hoch)	O 2 (niedrig)	O 3 (mittel)	O 4 (niedrig)	O 3 (sehr niedrig)
- Mobile Services mit Semantic-Web-Applikationen				
O 1 (sehr hoch)	O 2 (niedrig)	O 3 (mittel)	O 4 (niedrig)	O 3 (sehr niedrig)
- Semantic Web im Werbebereich				
O 1 (sehr hoch)	O 2 (niedrig)	O 3 (mittel)	O 4 (niedrig)	O 3 (sehr niedrig)
- Semantische Suchmaschinen				
- Textbasierte Suche				
O 1 (sehr hoch)	O 2 (niedrig)	O 3 (mittel)	O 4 (niedrig)	O 3 (sehr niedrig)
- Angebotssuchmaschinen (spezialisierte Suche nach Produkten)				
O 1 (sehr hoch)	O 2 (niedrig)	O 3 (mittel)	O 4 (niedrig)	O 3 (sehr niedrig)
- Semantische WIKIs				
O 1 (sehr hoch)	O 2 (niedrig)	O 3 (mittel)	O 4 (niedrig)	O 3 (sehr niedrig)
- Semantische Corporate Social Software (Blogs und Social Networks)				
O 1 (sehr hoch)	O 2 (niedrig)	O 3 (mittel)	O 4 (niedrig)	O 3 (sehr niedrig)

2.4.2 Ergebnisse

Nach der Auswertung der Fragebögen können folgende Ergebnisse vorgewiesen werden.

Die meisten Befragten finden, dass das Semantic Web einen hohen bis sehr hohen Nutzen hat, wobei der Nutzen für den Benutzer höher bewertet wurde als für Unternehmen.

Bezüglich des Aufwands für den Einsatz von semantischen Technologien gibt es unterschiedliche Meinungen. Viele sind der Meinung, dass momentan der Aufwand sehr groß ist, dass aber das Ziel ist, Applikationen zu entwickeln, die es ermöglichen,

diesen niedrig zu halten. Nach Hochhaltinger und Schramm [Interv. Hochhaltinger,; Interv. Schramm] hängt der Aufwand von der Komplexität der Daten ab. Wenn die Unternehmen eine einfache Datenstruktur haben und es schon bestehende Ontologien in dieser Branche gibt, dann ist auch der Aufwand niedrig, wie dies bereits in der Tourismusbranche der Fall ist. Wenn jedoch die Datenstruktur sehr komplex ist, dann braucht man dementsprechend mehr Zeit und Geld, um semantische Technologien einzusetzen.

Die Tourismusbranche hat den Experten zufolge das größte Potential, gefolgt von Verlagswesen, Bildungswesen, Unterhaltungselektronik und Finanzen.

Bei den Applikationen wurden die semantischen Empfehlungssysteme von allen Befragten als sehr hoch oder hoch eingestuft. Des Weiteren wurde die mobile Nutzung der semantischen Technologien ebenfalls als sehr vielversprechend bewertet, und laut Hochhaltinger [Interv. Hochhaltinger] ist sie die Zukunft des Web. Bei den Suchmaschinen wurde das Potential für Produktsuchmaschinen höher als die textbasierte Suchmaschinen eingestuft. Ein Grund dafür wurde von Szpuszta [Interv. Szpuszta] genannt, und zwar, dass die menschliche Sprache zu komplex ist, um die textbasierte Suche vom Semantic Web vollständig ersetzen zu lassen. Sie kann aber auf jeden Fall unterstützend wirken.

Szpuszta [Interv. Szpuszta] sieht Potentiale bei der Verknüpfung der Werbung mit dem Semantic Web, da hier der Bezug zu einem Geschäftsmodell höher ist. Die Werbebranche kann semantische Technologien nutzen, um effizienter Werbung zu platzieren. Er sieht aber auch eine Gefahr, dass dann die Ontologien nicht im Sinne der Benutzer, sondern im Sinne des Unternehmens gestaltet werden.

Bezüglich der Verknüpfung des Semantic Web mit dem Web 2.0 meint Hochhaltinger [Interv. Hochhaltinger], dass semantisch angereicherten Blogs viel Sinn machen, weil man damit das Wissen innerhalb der Blogs mit anderem Wissen intelligent verknüpfen und nutzen kann. Unternehmen können zum Beispiel die Produktbewertungen oder Kommentare, die in Blogs publiziert werden, erfassen und als Zusatzinformation in semantische Empfehlungssysteme einfließen lassen. Szpuszta [Interv. Szpuszta] ist diesbezüglich einer anderen Meinung. Für ihn stehen bei den Blogs oder Social Networks die Personen oder Nachrichten im Vordergrund und nicht die Verknüpfungen, die daraus entstehen können.

Auf die Frage, ob sich der Aufwand, semantische Technologien einzusetzen, in Anbetracht des Nutzens lohnt, gab es unterschiedliche Meinungen. Obwohl die meisten zugestimmt haben, sagten viele, dass die Technologie derzeit noch unreif ist, aber dass es sich in Zukunft auf jeden Fall lohnen wird. Daraus folgt die letzte Frage, ob sich die Semantic-Web-Technologie in den nächsten fünf Jahren durchsetzen wird. Bis auf einen Experten haben alle zugestimmt.

Aus diesen Interviews wird ersichtlich, dass das Semantic Web einen hohen Nutzen und großes Potential hat. Das bestehende Web kann mit den semantisch angereicherten Applikationen Prozesse verbessern und beschleunigen. Darüber hinaus können Benutzer davon profitieren, dass die Inhalte an deren Bedürfnisse angepasst werden und sie von den Online-Diensten besser unterstützt werden. Die große Herausforderung liegt darin, die nötigen Werkzeuge zu schaffen, damit der Einsatz von semantischen Technologien einfacher und schneller getätigt werden kann und mehr Leute davon profitieren können.

3 Fallstudie mit Universal Music Deutschland

In diesem Teil der Arbeit wird ein Projekt für das Unternehmen Universal Music Deutschland vorgestellt. Dieses Projekt befasst sich mit dem Einsatz von Semantic-Web-Applikationen im E-Commerce-Bereich und fokussiert sich auf die Personalisierung und Optimierung der Newsletter von Universal Music Deutschland. Die Handlungsempfehlungen erfolgen anhand der im Kapitel 2.1 beschriebenen Anwendungsgebiete im E-Commerce-Bereich und auf Grundlagen von Experten-Gesprächen. Des Weiteren werden in diesem Kapitel eigene Annahmen getroffen, die aufgrund der Zusammenarbeit mit Universal Music entstanden sind.

3.1 Universal Music Deutschland

Die Universal Music International Group⁵¹ entstand im Jahre 1996 aus der Übernahme der Music Corporation of America (MCA). Durch den Zukauf von PolyGram im Jahr 1998 entstand das größte Musikunternehmen der Welt. Die weitere Fusion mit Vivendi, Seagram und Canal+ führte zum weltweit zweitgrößten Medien- und Kommunikationsunternehmen Vivendi mit Hauptsitz in Paris [UMD, 2008].

Universal Music Deutschland⁵² ist Teil der Universal Music International Group, vermarktet sowohl deutsche als auch internationale Künstler und hat als Ziel, sie im deutschen Markt zu etablieren. Künstler und Stars von Universal Music Deutschland kann man in allen möglichen Genres finden. Von Pop über Dance, Rock, Volksmusik, Schlager und Jazz bis zur Klassik reicht der größte und vielfältigste Musikkatalog Deutschlands und der Welt. Bei Universal Music stehen unter anderen André Rieu, Die Ärzte, Christina Stürmer, Juli, Paul van Dyk, Rammstein, Tokio Hotel, 50 Cent, Black Eyed Peas, Bon Jovi, Elton John, Eminem, Jack Johnson, Juanes, Metallica, Mariah Carey und U2 unter Vertrag [UMD, 2008].

Das Frontline-Geschäft von Universal Music Deutschland wird von sechs Abteilungen ausgeführt, der Domestic Division, International Division, Koch, Classics & Jazz

⁵¹ <http://www.umusic.com>

⁵² <http://www.universal-music.de>

sowie Strategic Marketing und Family Entertainment. Diese Abteilungen entdecken und vermarkten Künstler aus verschiedenen Genres [UMD, 2008].

Die Abteilung Strategic Marketing vermarktet medienbeworbene Compilation-Tonträger, limitierte Sammlerauflagen, Midprice- und Budget-Serien. Außerdem schließt sie Kooperationen zwischen Künstlern und Unternehmen, Marken, Produkten und Medien. Im Bereich Business Development werden neue Geschäftsmodelle und Kooperationen im digitalen und physischen Markt entwickelt. Die Aktivitäten umfassen die Bereiche Direct to Consumer, Customer Relationship Management und Ausbau der Wertschöpfungsstufen [UMD, 2008].

Der Bereich Business Development ist verantwortlich für die Suche neuer Geschäftsmodelle und innovativer Konzepte - daher auch das Interesse für den Einsatz von Semantic-Web-Technologien zur Datenstrukturierung, Personalisierung der Inhalte und Wissensmanagement [Interv. Schramm].

In dieser Arbeit wird der mögliche Einsatz des Semantic Web im Unternehmen Universal Music anhand des Fokusszenarios Newsletter behandelt. Die Herausforderung für Universal Music Deutschland ist es, die Newsletter auf die Bedürfnisse und das veränderte Informationsverhalten der Abonnenten anzupassen.

In den folgenden Unterkapiteln wird das Unternehmen „Universal Music Deutschland“ mit der abgekürzten Version „Universal Music“ bezeichnet.

3.2 Die Newsletter

Universal Music bietet Genre-Newsletter und Künstler-Newsletter an. Die Ziele der Newsletter sind, für die Künstler und deren Repertoires zu werben und den Abverkauf an CDs, Singles, Ringtones und Merchandising zu fördern [Interv. Schramm].

Die Newsletter können auf den jeweiligen Portalen der Künstler wie z.B. Lang Lang⁵³ und auf der Unternehmensseite von Universal Music abonniert werden. Auf diesen Unterseiten kann man sein eigenes Profil im Bereich „MyProfile“ anlegen. Persönliche Daten, Lieblingskünstler und Genres können dort angegeben werden. Darüber

⁵³ <http://www.langlang.com>

hinaus werden die Benutzer gefragt, ob sie an Aktionen und Umfragen teilnehmen wollen [UMD, 2008].

In dieser Fallstudie wird der Schwerpunkt auf die Genre-Newsletter gelegt. Diese Newsletter können auf den jeweiligen Genre-Portalen abgerufen werden. Es gibt sieben verschiedene Genre-Portale von Universal Music [UMD, 2008]:

1. Vertigo mit Rock-Musik (<http://www.vertigo.fm>)
2. Urban mit Black-Musik (<http://www.urban.de>)
3. Compilations (<http://www.polystar.de/>)
4. POP24 mit Pop-Musik (<http://www.pop24.de>)
5. UK-Sounds mit Brit-Rock und -Pop (<http://www.uk-sounds.de>)
6. Klassik Akzente mit klassischer Musik (<http://www.klassikakzente.de>)
7. JAZZECHO mit Jazz-Musik (<http://www.jazzecho.de>)

Der Inhalt der sieben Genre-Newsletter von Universal Music [UMD, 2008] wird in Abbildung 18 verglichen. Die Struktur, Farbkombination und redaktionellen Inhalte der verschiedenen Newsletter sind sehr unterschiedlich und es kann kein einheitliches Format festgestellt werden. Bei den Klassik- und Jazz-Newslettern wird man gesiezt, bei den Compilations wird man geduzt und bei den restlichen werden die Abonnenten gar nicht persönlich angesprochen. Einige Newsletter enthalten Videos, andere Gewinnspiele oder Empfehlungen. Was aber alle gemeinsam anzeigen, sind Informationen über neue Veröffentlichungen.

	Vertigo	Urban	Compilations	Pop24	UK-Sounds	Klassik	JazzEcho
Persönliche Ansprache mit Namen			X (Man wird geduzt)			X (Man wird gesiezt)	X (Man wird gesiezt)
Inhaltsangabe						X	X
Nachrichten	X	X			X	X	X
Rezensionen						X	X
Neue Releases	X	X	X	X	X	X	X
- Information	X	X	X	X	X	X	X
- CD kaufen	X		X	X		X	X
- Lied runterladen	X			X	X		
- Lied hören			X				X
Angebot Konzerte	X	X				X	X
- Tickets kaufen		X					
Angebot TV						X	X
Angebot Radio							X
Videos	X			X	X		
Gewinnspiele				X	X	X	
Empfehlungen		X	X				
- Downloads		X	X				
- Ringtones			X				
- Newsletter			X				
- Werbung	X	X	X	X			
Extras				X (Merchandising und Spiel)		X (Spiel)	X (Seite weiterempfehlen)

Abb. 18: Analyse der Inhalte der Genre-Newsletter [Eigene Darstellung]

Die Abbildung 18 verdeutlicht die Tatsache, dass die Genre-Newsletter sehr unterschiedlich sind und viel Information enthalten. Da die Abonnenten auch sehr verschieden sind, werden mit den Genre-Newslettern von Universal Music nicht alle Abonnenten zielgerichtet angesprochen. Zum Beispiel: Ein Abonnent des Newsletters Pop24 interessiert sich für internationale Pop-Musik und weniger für deutsche Pop-Musik. Ein anderer hat wiederum eine besondere Vorliebe für deutsche Pop-Musik [Interv. Schramm].

3.3 Problemstellung

Um die aktuellen Herausforderungen der Newsletter von Universal Music zu verstehen, werden in diesem Unterkapitel zwei Phänomene dargestellt, mit denen sich

Universal Music bei der Gestaltung der Newsletter auseinandersetzen muss und die zum Problem der relevanten Inhalte führen. Diese sind die hohe Informationsdichte und die unterschiedlichen Bedürfnisse der Zielgruppen.

3.3.1 Hohe Informationsdichte

Durch die Revolution der digitalen Medien hat sich im Internet viel Information angesammelt. Eine enorme Menge an Daten wurde generiert und wächst mit einer exponentiellen Rate weiter. Diese Datenmenge kann ohne Unterstützung von Computer unmöglich bewältigt werden [Garcia und Celma, 2005].

Die Internet-Benutzer werden durch die zunehmende Kommerzialisierung des Internet im E-Commerce-Bereich mit Informationen und Angeboten überflutet [Schackmann, 2003, S. 53]. Daher versuchen die Benutzer, die Information zu filtern und nur diejenige Information zu lesen, die für sie relevant ist [Interv. Schramm]. Das führt dazu, dass Werbetreibende neue Methoden und Technologien nutzen müssen, um die Aufmerksamkeit der Nutzer durch bessere zielgruppen- und umfeldorientierte Online-Werbung zu gewinnen [Fittkau & Maaß, 2003].

Wenn das nicht geschieht, dann werden die Leseraten der Newsletter weiter fallen. Zu diesem Thema äußert sich Schramm [Interv. Schramm]: *„Die Problematik von Newslettern generell ist, dass ihre Leseraten tendenziell zurückgehen. Ein Grund dafür ist, dass sich relevante Informationen mit einer unglaublich hohen Zahl an sonstiger Information vermengen und man sie in dieser Informationsdichte nur schwer findet“*.

Folgende Daten verdeutlichen diese Tatsache: Die meisten deutschen Internet-Benutzer (86%) haben einen E-Mail-Newsletter abonniert oder eingewilligt, von Unternehmen via E-Mail informiert zu werden [Schwarz, 2004]. Laut Fittkau & Maaß Consulting [Fittkau & Maaß, 2003], die in einer Studie 117.000 Abonnenten von Newslettern befragt haben, lesen nur 21 Prozent der Abonnenten ihre Newsletter, 51 Prozent überfliegen sie und 12 Prozent löschen die Newsletter sofort. Die niedrige Öffnungs-Rate ist dreimal so hoch (ca. 65 Prozent), wenn die Benutzer eine aktive E-Mail-Beziehung mit den Unternehmen führen. Im Durchschnitt stehen Internet-Benutzer mit 16 Unternehmen in einem engeren Kontakt. Für Unternehmen ist es

sehr wichtig, diese „loyalen“ Kunden zu pflegen. Die Personen, die seit zwei Jahren im Verteiler sind, kaufen neunmal jährlich Produkte, und Abonnenten, die länger im Verteiler sind, dreizehn Mal im Jahr [Schwarz, 2004, S. 8, ff.].

3.3.2 Unterschiedliche Bedürfnisse der Zielgruppen

Anhand regelmäßiger, ethnologischer Einzelfallstudien und repräsentativer Umfragen ermittelt das Heidelberger Forschungsinstitut Sinus Sociovision seit Mitte der 1980er Jahre die Schichtzugehörigkeit und die Wertorientierung der deutschen Bevölkerung. Die Sinus-Milieus-Landkarte (siehe Abbildung 19) teilt die deutsche Bevölkerung in zehn verschiedene Milieus ein. Personen mit ähnlichen Lebenseinstellungen, Lebensweisen und Wertorientierungen werden dargestellt [Sinus-Socio, 2007].

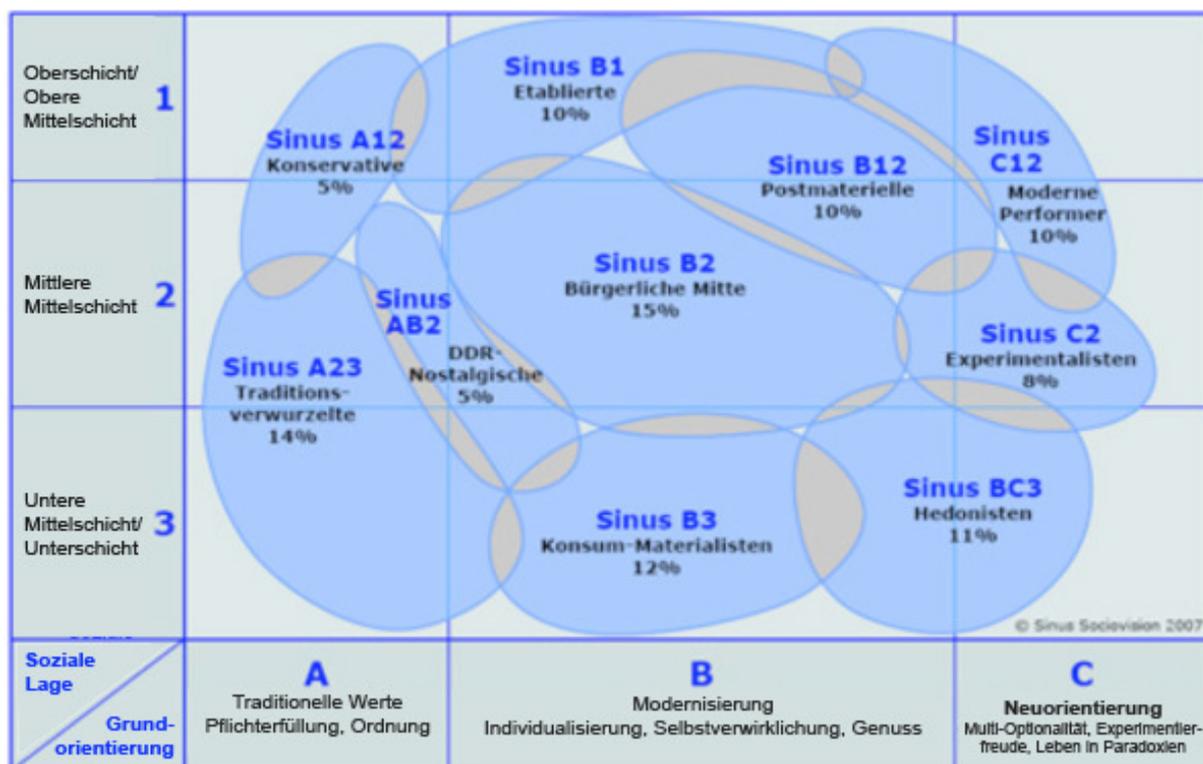


Abb. 19: Sinus-Milieus in Deutschland 2007 [Sinus-Socio, 2007]

Im Modell der Sinus-Milieus kann man vier Milieugruppen unterscheiden, die gesellschaftlichen Leitmilieus (Postmaterielle, Moderne Performer, Etablierte), die traditionellen Milieus (Konservative, DDR-Nostalgiker, Traditionsverwurzelte), die Mainstream-Milieus (Konsum-Materialisten, bürgerliche Mitte) und die hedonistischen Milieus (Experimentalisten, Hedonisten). Das Ziel des Sinus-Milieus ist, die komplexe

Wirklichkeit zu reduzieren und die Zielgruppen zu clustern, um eine vereinfachte Ansprache zu ermöglichen [Gimpel und Dommisch, 2006].

Die Schwachpunkte dieses Modells liegen in der Verallgemeinerung der Konsumenten, da der heutige Konsument nicht in einer einzelnen Zielgruppe abgebildet werden kann, sondern zwischen den Gruppen wechselt. Der Konsument ist unberechenbar und seine Entscheidungen sind wechselhaft. Vor allem bei Massenkonsumgütern haben situative Faktoren einen stärkeren Einfluss auf den Kauf als die Grundorientierung des Konsumenten. Die klassischen Ansätze zur Einteilung der Gesellschaft in stabile Zielgruppen werden durch die Individualisierungstendenzen der modernen Gesellschaft immer mehr in Frage gestellt. Die Konsumenten können nicht mehr in eine Schublade gesteckt und zielgerecht angesprochen werden, da sie in verschiedenen Situationen zu unterschiedlichen Zielgruppen gehören [Gimpel und Dommisch, 2006].

Basierend auf den oben genannten Erkenntnissen von Gimpel und Dommisch [2006] und dem Interview mit Felix Schramm [Interv. Schramm], Concept Developer bei Universal Music, können alle Zielgruppen des Sinus-Milieus über die Universal Music-Newsletter erreicht werden. Laut Schramm [Interv. Schramm] kann man in der heutigen postmodernen Gesellschaft kaum die Abonnenten der Genre-Newsletter von Universal Music auf bestimmte Milieus oder Subkulturen ausrichten, wie es einst der Fall war. Ein Abonnent des Klassik-Newsletter könnte sowohl ein „Konservativer“ wie auch ein „Etablierter“ sein und sogar auch Rock-Musik mögen. Ein Rock-Fan kann sowohl zu der Gruppe der „Experimentalisten“ als auch zu der Gruppe „Moderne Performer“ gehören. Somit wird ersichtlich, dass die Abonnenten der Genre-Newsletter in keine eindeutige Zielgruppe eingeordnet werden können. Ein einheitlicher Newsletter ist daher nicht effektiv genug und eine Personalisierung des Newsletter ist notwendig, um Konsumenten zielgerichtet anzusprechen.

3.3.3 Problem: Relevanz der Inhalte

Die Abonnenten der Newsletter von Universal Music bekommen durch die hohe Informationsdichte (siehe Kapitel 3.3.1) und die unterschiedlichen Bedürfnisse der Abonnenten (siehe Kapitel 3.3.2) viel Information, die teilweise nicht spezifisch genug

und für sie nicht relevant ist.

Universal Music hat ein Künstlerportfolio mit verschiedenen Künstlern pro Genre. Zum Beispiel ist der Pop24-Newsletter sehr breit gefächert und enthält Informationen und Angebote von Künstlern wie Fräulein Wunder, Tokio Hotel, Klee und andererseits Songs von Lang Lang oder Norah Jones [UMD, 2008]. Um eine breite Masse an Abonnenten anzusprechen, wird im Newsletter so viel Information wie möglich publiziert [Interv. Schramm]. Der Abonnent kann nicht festlegen, welche Künstler ihn aus einem bestimmten Genre interessieren oder über welche Künstler er keine Informationen bekommen will. Er kann auch nicht aussuchen, welche Rubriken angezeigt werden sollen oder wie die Struktur des Newsletter aussehen soll. Aus diesem Grund muss er lange suchen, bis er einen Künstler oder eine Nachricht findet, die seinem Interesse entspricht.

Durch den Einsatz von semantischen Technologien können Informationen besser strukturiert und verarbeitet werden [Garcia und Celma, 2005]. Darüber hinaus ermöglichen sie die Erstellung von personalisierte Angeboten, die den Bedürfnissen der Benutzer entsprechen [Dustdar et al., 2006].

3.4 Lösung: Einsatz von Semantic-Web-Technologien zur Personalisierung der Inhalte der Genre-Newsletter

„Customers, whether consumers or businesses, do not want more choices. They want exactly what they want - when, where, and how they want it - and technology now makes it possible for companies to give it to them. Interactive and database technology permits companies to amass huge amounts of data on individual customers' needs and preferences.“ [Pine et al., 1995, S.103]

Schackmann [2003, S.53] bestätigt dieses Zitat von Pine et al. [1995, S.103] mit dem Gedanken, dass ein einheitliches Angebot nach dem Motto „one-fits-it-all“ nicht ausreicht, um die Kunden zufrieden zu stellen. Die technologischen Entwicklungen ermöglichen den Unternehmen, diese Bedürfnisse wahrzunehmen und zugeschnittene Angebote und Inhalte an die Benutzer weiterzuleiten [Pine et al., 1995, S.103]. Eine Technologie, um dies zu erreichen, ist das Semantic Web.

Personalisierung bedeutet „*Inhalte zu liefern, die spezielle Informationen für einen Empfänger oder eine Gruppe von Empfängern enthalten*“ [Schwarz, 2004, S.156]. Personalisierte Newsletter sind ein wichtiger Bestandteil für die Stärkung einer langfristigen Kundenbeziehung [Schwarz, 2004, S.156] und stellen eine Lösung für das im Kapitel 3.3.3 genannte Problem der Relevanz der Inhalte dar.

Um die Inhalte der Newsletter mithilfe von Semantic-Web-Technologien zu personalisieren, müssen bei Universal Music erstens eine semantische Infrastruktur und zweitens Benutzerprofile geschaffen werden. Diese Punkte werden in den Kapiteln 3.4.1 und 3.4.2 behandelt.

Im Kapitel 3.4.3 werden verschiedene Einsatzmöglichkeiten von semantischen Technologien dargestellt, die Universal Music in Anspruch nehmen kann, um die Inhalte der Newsletter zu personalisieren und angepasste Angebote an die Abonnenten zu liefern.

3.4.1 Aufbau der semantischen Infrastruktur

In dem vorliegenden Abschnitt werden die verschiedenen Schritte erklärt, die zum Einsatz von Semantic-Web-Technologien in Universal Music nötig sind.

Erstens müssen Ontologien geschaffen oder bestehende Ontologien genutzt werden. Zweitens werden die relevanten Daten mit Metadaten und RDF-Annotationen versehen [Interv. Hochhaltinger].

3.4.1.1 Entwicklung oder Verwendung von bestehenden Ontologien

Wenn man das Semantic Web in einem Unternehmen einsetzen will, muss man entscheiden, ob man es intern einsetzt (zum Beispiel im Wissensmanagement), oder ob man seine Daten und Ontologien für das ganze Semantic Web verfügbar macht. Wenn man sich für die zweite Alternative entscheidet, ist es wichtig, dass die Ontologien kompatibel mit den bestehenden Systemen sind, und im optimalen Fall auf bestehenden Standards aufsetzen [Interv. Hochhaltinger].

Der zweite Schritt ist die Entwicklung von Ontologien. Man kann entweder neue Ontologien schaffen oder bestehende Ontologien verwenden, die sich mit der eigenen Branche und dem eigenen Themengebiet beschäftigen. Die bestehenden Ontologien muss man dennoch an die eigenen Daten und Anforderungen anpassen. Die zweite Alternative ist, dass ein Unternehmen eine Ontologie für sich selber entwickelt [Interv. Hochhaltinger].

Folgende Ontologien wurden für die Musik-Branche entwickelt:

- **Music-Ontology**

Die Music-Ontology-Spezifikation [MusicOnt] ist ein formelles Framework mit semantischen, musikbezogenen Informationen, wie zum Beispiel redaktionelle, kulturelle und akustische Information. Diese Informationen stehen in Beziehung zueinander und können im Semantic Web verlinkt werden [Raimond et al., 2007].

Das Ziel des Music-Ontology-Projektes ist, ein umfassendes, maschinenverarbeitbares kulturelles Umfeld zu schaffen, das sich mit verschiedenen Konzepten im Musikbereich beschäftigt und in vielen Bereichen angewendet werden kann. Diese Bereiche können von der Musikproduktion über Musikkonsum bis zur Musikempfehlung gehen. Blogger werden damit ein mit dem Handy gefilmtes Konzert auf ihren Blogs publizieren und Musikologen komplexe Tonaltitäten eines Jazzstückes in semantischer Sprache beschreiben können [Raimond et al., 2007].

Um diese Vielseitigkeit zu ermöglichen, wurden im Rahmen des Projektes Music-Ontology verschiedene Ontologien entwickelt (Abbildungen von diesen Ontologien findet man im Anhang 6.1). Um die Integration von feinkörnigen Zusammenhängen zu ermöglichen, wurde die Ontologie in verschiedenen Expressivitätsniveaus aufgeteilt [Raimond et al., 2007]:

Die erste Ebene befasst sich nur mit redaktioneller Information, mit der man zum Beispiel Folgendes ausdrücken kann: „*Dieser* Künstler gehört *dieser* bestimmten Musikgruppe und *diese* Organisation gehört *diesem* Label (siehe Abbildung 20).“

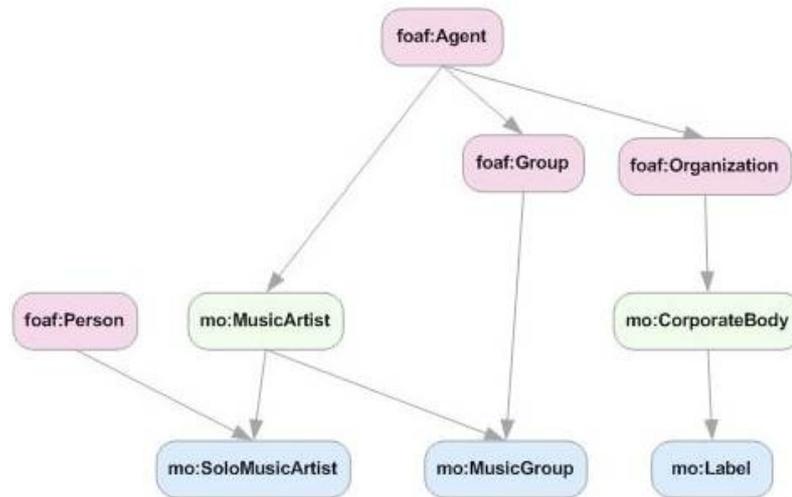


Abb. 20: Ontologie mit redaktioneller Information von Music-Ontology [Wiki MusicOnt, 2008]

Die zweite Ebene bezieht das Konzept „Event“ ein und beschäftigt sich mit der Gestaltung des Werkes, der Aufführung der Lieder und der Aufnahme der Aufführungen. Zum Beispiel: „Ich war letzten Abend bei *dem* Konzert, das ich mit meinem Handy aufgenommen habe, und *hier* ist der dazugehörige Audio-Stream.“

Die dritte Ebene bezieht das Element „Zerlegung“ ein. Dieses ermöglicht, die Musikstücke zu zerlegen und zu beschreiben, wie zum Beispiel: „In *dieser* Aufführung wurde diese Tonart in *diesem* bestimmten Moment von *dieser* Person auf dem Klavier gespielt.“

Die Ontologien von Music-Ontology können mit weiteren Ontologien oder RDF-Daten aus dem Internet bereichert werden. Zum Beispiel ist es möglich, Information aus der Wiki DBpedia⁵⁴ zu entnehmen, um Musik-Genres zu beschreiben und Zusammenhänge zu schließen [Raimond et al., 2007].

- **Music Ontologie – Universität Ulm [UniUlm, 2006]**

Eine weitere Ontologie im Musikbereich wurde von der Universität Ulm für ein Radioportal erstellt (siehe Abbildung 21). Diese Ontologie unterteilt das Radioprogramm in drei Kategorien: Bildung, Unterhaltung und Information. Im Unterhaltungsbereich werden die verschiedenen Musikrichtungen in weitere

⁵⁴ <http://dbpedia.org>

Unterkategorien aufgeteilt. Jazz kann zum Beispiel in Bigband, Dixieland, Latin-Jazz usw. eingeteilt werden. Die verschiedenen Informationsthemen der Radio-Nachrichten können auch in der Ontologie beschrieben werden und in verschiedene Unterkategorien wie Wirtschaft, Sport, Wetter und Reisen unterteilt werden.

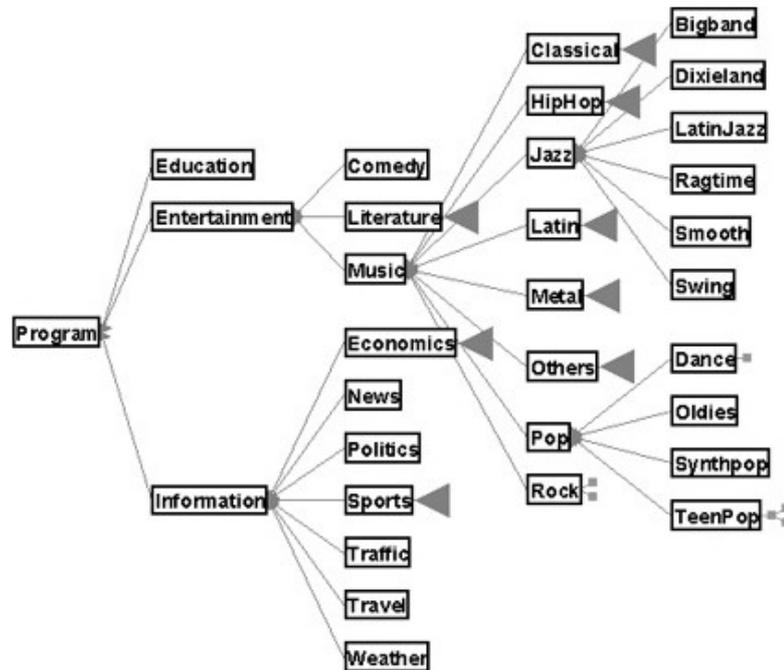


Abb. 21: Ausschnitt aus einer Ontologie über Radioprogramme [UniUlm, 2006]

Neben der Verwendung von bestehenden Ontologien können diese von Benutzern und Unternehmen selbst erstellt werden. Wie im Kapitel 1.2.3 erwähnt, gibt es kollaborative Ontologie-Entwicklungsplattformen wie myOntology [myOnt], die auf Wiki-Technologien basieren und eine dynamische und zeitnahe Entwicklung von Ontologien ermöglichen. Die Ontologien sind für alle zugänglich und können durch Anregung, Kritik oder Verbesserungsvorschläge von anderen Entwicklern oder Fachexperten optimiert werden [ebSem AUSTRIAPRO, 2007]. Des Weiteren empfehlen Grigoris und Van Harmelen [2004, S. 219] folgende Werkzeuge, um Ontologien zu erstellen: OILed⁵⁵, Protégé⁵⁶ und OntoEdit⁵⁷, wobei sie OILed bevorzugen. Chebotko et al. [2004] ziehen wiederum OntoEdit vor.

Da sich Universal Music mit dem Konzept und Einsatz von semantischen Technolo-

⁵⁵ <http://oiled.man.ac.uk>

⁵⁶ <http://protege.stanford.edu>

⁵⁷ <http://ontoprise.de>

gien in der Orientierungsphase befindet, ist es sinnvoll, bestehende Ontologien wie zum Beispiel solche aus MusicOntology [MusicOnt] zu verwenden (siehe Anhang 6.1) und diese dann an die internen Anforderungen zu adaptieren.

Für die Newsletter von Universal Music sind folgende Ontologien notwendig:

- Die Event-Ontologie (Anhang 6.1.1), um die Konzerte der Künstler zu beschreiben.
- Die Künstler- und Musikgruppen-Ontologie (Anhang 6.1.7), um die Künstler, Gruppen und Labels zu beschreiben.
- Die Veröffentlichungsart-Ontologie, um die Art der Veröffentlichung zu beschreiben, wie zum Beispiel: Soundtrack, Remix, Live, Album oder Compilation.
- Die Genre-Ontologie, um die verschiedenen Genres und ihre dazugehörigen Künstler darzustellen.
- Benutzerprofil-Ontologien, um den Musikgeschmack, Interessen und Vorlieben der Abonnenten der Newsletter kennen zu lernen (siehe Kapitel 3.4.2).

3.4.1.2 Semantische Annotation der Daten

„Die Existenz von RDF-Metadaten ist eine Grundvoraussetzung für die Realisierung des Semantic Web.“ [Reif, 2006]

Das RDF-Modell und die Syntaxspezifikation wurden von dem World Wide Web Consortium (W3C) geprüft und als eine Empfehlung für die Nutzung und Repräsentation von Metadaten (Daten über Daten) im Internet befürwortet (siehe Kapitel 1.2.1) [W3C RDF, 1999]. Wenn die Dokumente mit semantischen Metadaten versehen werden, dann spricht man von einer semantischen Annotation der Daten [Reif, 2006].

Nachdem Unternehmen ihre Daten mittels Ontologien (Kapitel 3.4.1.1) strukturiert haben, ist nach Hochhaltinger [Interv. Hochhaltinger] der nächste Schritt, ihre Daten mit RDF-Annotationen zu beschreiben, welche sich auf die verwendeten Ontologien beziehen:

„Die meisten Unternehmen haben schon klassische Datenbanken oder MySQL-Datenbanken oder haben das Ganze über XML beschrieben. Um diese Daten im Semantic Web zu nutzen, müssen diese auf Basis der Ontologien in RDF beschrieben werden. Der Aufwand der Daten-Beschreibung in RDF hängt hauptsächlich nicht von der Anzahl an Daten ab, sondern von der Komplexität der Daten, wie breit die Datenbank ist. Wenn die Datenbank viele Felder und Relationen hat, dann ist es schwieriger, die Daten in RDF zu beschreiben. Die große Herausforderung ist, dass so viele Unternehmen wie möglich ihre Daten in RDF umwandeln sollen, damit semantische Daten generiert werden können und sich dieses Konzept durchsetzt.“
[Interv. Hochhaltinger]

Im Musikbereich können die Künstler, Lieder usw. mit Metadaten angereichert werden. Die ganze Information wird der Musik hinzugefügt, ist aber nicht Teil von ihr, sie beschreibt diese nur [Swartz, 2002]. Der Vorteil der semantischen Annotation mit Metadaten besteht darin, dass diese getrennt von den eigentlichen Nutzdaten sind und sehr schnell und ökonomisch aufgerufen werden können [ITWissen, 2008]. Andererseits meinen Garcia und Celma [2005] und Turnbull et al. [2007], dass die manuelle Gewinnung von semantischen Musik-Annotationen mit viel Aufwand und Zeit verbunden ist und sehr kostenintensiv sein kann. Wenn man die Verfügbarkeit von semantischen Multimedia-Metadaten erhöhen will, dann braucht man auch Methoden, die produktiver sind. Die direkteste Lösung ist, von dem großen Umfang an Metadaten zu profitieren, die bereits im Internet sind [Garcia und Celma, 2005].

Es gibt verschiedene Methoden, um Musik semantisch zu annotieren. Turnbull et al. [2007] teilen diese in vier auf: die Musik kann manuell [MusicBr], durch Befragungen, Web-Mining-Methoden oder Online-Spiele [ListGame] semantisch annotiert werden.

Einige Initiativen zur semantischen Aufbereitung der Daten im Musikbereich sind folgende:

- **Online-Games [Turnbull et al., 2007]**

Listen Game [ListGame] ist ein Online-Multiplayer-Spiel, in dem Menschen spielerisch Musikstücke semantisch annotieren. Das heißt, die Internet-Benutzer bekommen im „normalen“ Modus eine Liste mit semantisch verwandten Wörtern zu Instrumenten, Gefühlen oder Verwendungszwecken und

müssen das am besten und das am wenigsten passende Wort für dieses Lied aussuchen. Im „freestyle“-Modus sollen sie ein neues Wort aussuchen, um diese Musik zu beschreiben. Jeder Spieler bekommt Echtzeit-Feedback über die Übereinstimmung mit allen Spielern. Laut Turnbull et al. [2007] kann man durch diese „Spiele für einen guten Zweck“ viele Menschen gewinnen, die durch ein Spiel bereit sind zu kooperieren.

- **MusicBrainz! [Swartz, 2002]**

Der Musikindustrie mangelt es an Standardisierungen in Bezug auf Metadaten-Formate sowie auch an öffentlichen Metadaten. Das MusicBrainz!-Projekt [MusicBr] versucht diese Situation zu ändern, indem es eine umfassende Datenbank von Musik-Metadaten zur Verfügung stellt. Die Information in MusicBrainz! ist benutzerbasiert.

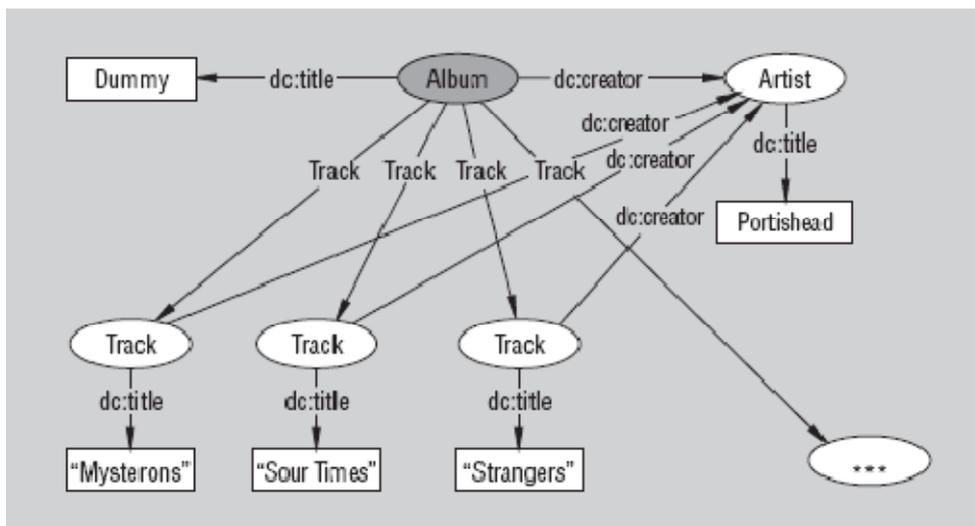


Abb. 22: Grafische Repräsentation der MusicBrainz!-RDF-Daten [Swartz, 2002]

Je mehr Benutzer Teil der MusicBrainz!-Community werden, desto wertvoller ist das System. Wenn zum Beispiel ein Benutzer eine neue Musik-CD kauft und die auf ihr enthaltenen Daten auf seinen Computer spielt, dann wird die Musik-CD wahrscheinlich einen generischen Namen haben wie „Audio CD47“ und die Lieder als „Track1“, „Track2“ usw. bezeichnet werden. Wenn man jedoch MusicBrainz! verwendet, dann greift es im Audiosystem ein und prüft, ob es Metadaten für diese CD gibt. Wenn das Lied im System gespeichert ist, dann wird das Audiosystem die CD und die Lieder mit dem jeweiligen Namen benen-

nen, zum Beispiel CD „Dummy“ und Lied „Sour Times“ oder „Strangers“ (siehe Abbildung 22). Im Falle, dass Metadaten nicht verfügbar sind, können die Daten von den Benutzern manuell eingefügt werden.

Das Semantic Web hat zum Ziel, Inhalte im Web mit maschinell verarbeitbaren Informationen zu bereichern und damit auch Prozesse zu optimieren. MusicBrainz! tut beides. Es liefert Inhalte in HTML und im RDF-Format und beschleunigt die Musikbeschreibung. Alle Objekte, wie zum Beispiel Künstler, Alben und Lieder, bekommen eine eindeutige URI zugeordnet, das heißt eine Internetadresse in der Datenbank, die von anderen Applikationen genutzt werden kann. Die Datenbank ist für die Öffentlichkeit unter der OpenContent-Lizenz verfügbar. Die offenen Protokolle ermöglichen den Benutzern, auf die Information zuzugreifen. Sie können auch aussuchen, von welchem Server sie die Daten bekommen wollen und diejenigen Server bevorzugen, die höhere qualitative Information bieten. Alle Server sind untereinander kompatibel, so dass Benutzer von einem zum anderen wechseln können.

Weil das MusicBrainz!-Format offen ist und mit RDF den Industriestandard für Metadaten nutzt, kann es in vielen Applikationen weiter verwendet werden. Zum Beispiel können andere Benutzer auf die Daten von MusicBrainz! zugreifen und diese weiterentwickeln. Musikanbieter können die Metadaten von MusicBrainz! nutzen, um die Information in deren Webseiten anzureichern und um darauf hinzuweisen, wo man die Musik legal kaufen kann.

- **MPEG-7**

MPEG-7 ist die größte Multimedia-Metadaten-Initiative, die das Ziel verfolgt, einen Standard für die Beschreibung von Multimedia-Inhalten zu schaffen [Garcia und Celma, 2005]. MPEG-7 kann audiovisuelle Information beschreiben, egal, welcher Speicher, Codierung, Getriebe, Display, Medium oder Technologie verwendet wird. Darüber hinaus umfasst er verschiedene Medientypen wie zum Beispiel Standbildübertragung, Grafiken, 3D-Modelle, Audio, Sprache, Video und Kombinationen von diesen [Tsinaraki, 2007].

Der Nachteil von MPEG-7 ist, dass er auf XML (Extensible Markup Language)-Schema basiert und keine formelle Semantik enthält. Durch semantische Tech-

nologien können jedoch MPEG-7-Multimedia-Daten besser integriert, erweitert und verwaltet werden [Garcia und Celma, 2005]. Daher gab es bereits verschiedene Initiativen [Garcia und Celma, 2005; Hunter, 2003; Tsinaraki et al., 2004; Tsinaraki, 2007], die das Ziel verfolgt haben, MPEG-7 semantisch anzureichern.

Der Ansatz von Garcia und Celma [2005] ermöglicht die automatische Abbildung der MPEG-7-Metadaten in OWL, einer Standard-Ontologiesprache (siehe Kapitel 1.2.3). Um MPEG-7-Daten mit Semantik anzureichern, schlagen sie vor, die XML-Instanzen in semantische RDF-Metadaten zu konvertieren.

In diesem semantischen Framework ist es einfacher, Metadaten, die aus verschiedenen Quellen kommen, zu integrieren. Die Maschinen können somit die Bedeutung der Multimedia-Daten interpretieren und die Genauigkeit des Information-Retrievals, also der Informationswiedergewinnung, dramatisch erhöhen. Dann sind auch komplexe Dienstleistungen über komplexe Multimedia-Datenbehälter möglich [Garcia und Celma, 2005].

Die Architektur der Initiative wird in Abbildung 23 dargestellt. Die MPEG-7 OWL-Ontologie bildet die Grundlage für das ontologische Framework, das die Multimedia-daten im Internet integriert. Weitere externe Ontologien können durch das XSD2OWL-Modul integriert werden. Semantische Metadaten können direkt im System mit XML-Metadaten integriert werden. Garcia und Celma [2005] haben drei Metadaten-Schemen im Rahmen ihres Forschungsprojektes integriert: das MusicBrainz-Schema (siehe Kapitel 3.4.1.2), die Simac-Music-Ontology [Herrera et al., 2005] und ein Musik-Vokabular, um Konzerte zu beschreiben. Die verschiedenen Schemen verwenden das RDF-Format zur Beschreibung der Lieder, Künstler oder Alben [Garcia und Celma, 2005].

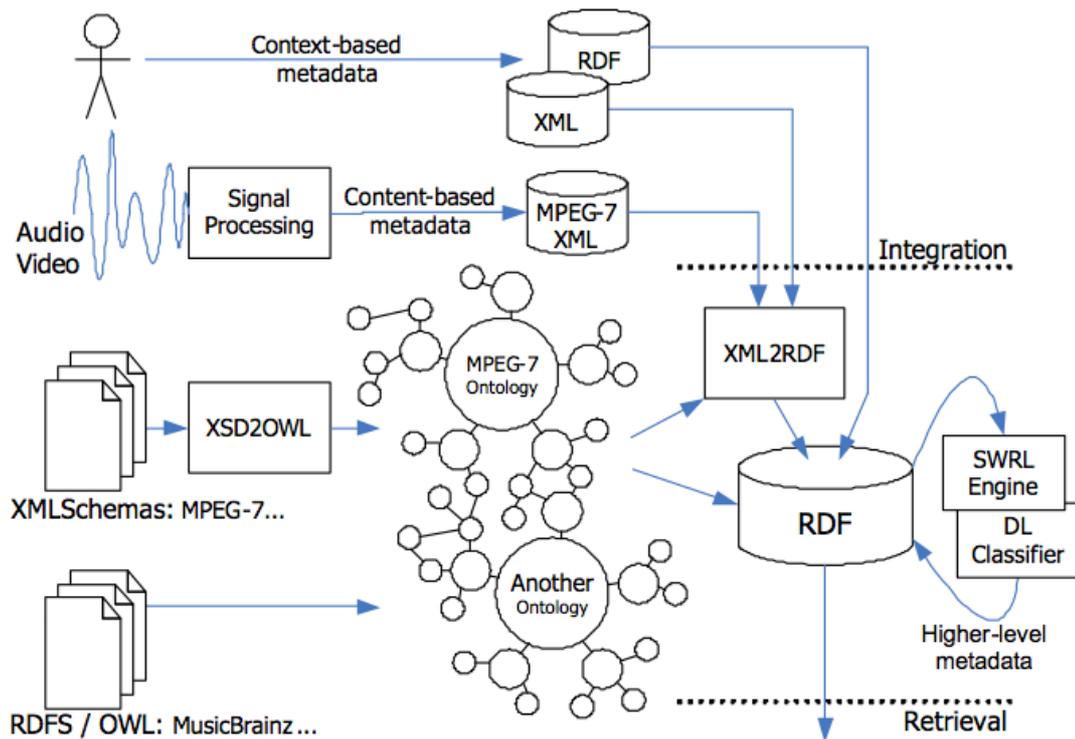


Abb. 23: Metadaten-Architektur und Retrieval Architektur [Garcia und Celma, 2005]

Dieses Framework wird von einer RDF-Ablage unterstützt, die Metadaten und Ontologien beinhaltet. Wenn all die Metadaten zusammengesetzt werden, dann findet die semantische Integration statt und die Metadaten können von verschiedenen semantischen Applikationen abgerufen werden [Garcia und Celma, 2005].

Des Weiteren können Unternehmen und Benutzer die Daten mit Annotations-Tools semantisch beschreiben. Die Tools, die im folgenden Absatz erklärt werden, sind im Rahmen von Forschungsprojekten erstellt worden. Ein Beispiel ist der RDFmaker [RDFmaker, 2007], der bereits im Kapitel 1.2.1 dieser Arbeit erwähnt wird. Mit diesem kann man beliebige Sachverhalte maschineninterpretierbar beschreiben und er ist besonders für E-Commerce-Anbieter geeignet, da man ohne viel Programmierungkenntnisse die Eigenschaften von Produkten wie zum Beispiel von Musikstücken beschreiben kann (Genre, Dauer, Künstler, Musikgruppe, Art der Veröffentlichung, Preis, Verfügbarkeit usw.) [RDFmaker, 2007]. Weitere Initiativen zur manuellen Daten-Annotation sind der SHOE Knowledge Annotator [Heflin et al., 2003], der CREAM OntoMat-Annotizer [Handschuh und Staab, 2003] und SMORE [Kalyanpur et al., 2003]. Ein weiteres Tool ist der S-cream [Handschuh et al., 2002],

der ein semiautomatisches Tool ist und damit den Aufwand der Annotierung wesentlich verringert.

Turnbull et al. [2007] und Reif [2006] sind sich einig, dass manuelle Methoden zur semantischen Datenannotation einen sehr zeitaufwendigen Prozess darstellen. Darüber hinaus meinen Turnbull et al. [2007], dass der Einsatz von Web-Mining-Techniken in einigen Fällen auch keine sehr positiven Ergebnisse gezeigt hat, da sie nicht immer gute Beschreibungen von Liedern liefern. Die Befragungen haben bessere Resultate ergeben, dafür sind sie aber langwierig und zeitaufwendig. Sie bevorzugen daher die Gewinnung von Daten durch Online-Spiele, weil diese durch verschiedene Anreize innerhalb des Spiels (wie zum Beispiel das Feedback der anderen Benutzer), die Menschen anspornen, gute Annotationen zu liefern [Turnbull et al., 2007].

Wenn aber die manuelle Annotation durch kollaborative Ansätze wie MusicBrainz! geschaffen wird, dann vermindert sich der von Turnbull et al. [2007] genannte Aufwand erheblich. Die Musicbrainz!-Datenbank hat schon einige Use-Cases und wurde von verschiedenen semantischen Applikationen benutzt, wie zum Beispiel in der semantischen Musik-Suchmaschine Squiggle [Celino et al., 2006], die der Indexierung und dem Retrieval von Multimedia-Inhalten dient.

Die Metadaten können nicht direkt als solche verwendet werden, sondern sie müssen der Struktur und den Erfordernissen der Ontologien entsprechen [Reif, 2006]. Wenn die im Kapitel 3.4.1.1 genannten Ontologien von Universal Music geschaffen werden, dann muss man die Daten, die dafür benötigt werden, mit semantischen Annotationen versehen. Das heißt, die Lieder der jeweiligen Genres, Künstler, Musikgruppen und Veröffentlichungsart müssen mit RDF versehen werden. Informationen über die Lieder können aus der MusicBrainz!-Datenbank entnommen werden. Weitere Metadaten können durch semiautomatische Methoden wie S-cream [Handschuh et al., 2002] oder durch Online-Games wie Listen Game [Turnbull et al., 2007] erworben werden.

3.4.2 Erstellung von Benutzerprofilen

Eine wichtige Voraussetzung, um erfolgreich Personalisierungsstrategien einzusetzen, ist, Information über die Kunden und deren Bedürfnisse zu besitzen. Zutreffende Empfehlungen können nur geliefert werden, wenn ausführliche Kundenprofile vorhanden sind [Tusek, 2006, S. 38, ff.].

Bei der Datengewinnung ist es wichtig, dass die Interessenten nicht gleich bei der Anmeldung der Newsletter mit vielen Fragen nach persönlichen Daten „überrumpelt“ werden. Wenn dies geschieht, werden die Interessenten verschreckt und geben tendenziell weniger Daten von sich preis. Wenn nur die E-Mail-Adresse als Pflichtfeld angezeigt wird, verdoppelt sich oft die Anzahl der Neuanmeldungen. Immerhin können die Unternehmen den Benutzern die Option anbieten, mehr Information anzugeben. Die Benutzer geben lieber im Nachhinein mehr Daten an, wenn eine Vertrauensbasis mit dem Unternehmen aufgebaut wurde und sie einen Mehrwert dabei erkennen [Schwarz, 2004].

Daten über die Abonnenten der Newsletter können auf verschiedenste Weise gewonnen werden:

- **Selbsteingabe**

Wie bereits im vorletzten Absatz erwähnt, kann man den Benutzern die Option geben, verschiedene Informationen über sich anzugeben.

- **Klickverhalten**

Anhand der bisher angeklickten und nicht angeklickten Webseiten kann man von den Benutzern ein Profil erstellen [Schackmann, 2003] und ihre Interessensgebiete und bisheriges Kaufverhalten beschreiben. Dieser Prozess kann mithilfe von semantischen Technologien optimiert werden [Roberto und Silva, 2007].

Roberto und Silva [2007] haben einen Algorithmus entwickelt, der das semantische Wissen und die Navigationsdaten der Benutzer integriert, um die Interessen der Benutzer besser zu identifizieren. In der Praxis sieht es folgen-

dermaßen aus: Ein Benutzer ruft eine semantische Internetseite auf und navigiert durch die Seite. Der Navigationsverlauf wird in einem semantischen Log-File gespeichert. Diese Daten werden in semantischer Form aufbereitet und ein Benutzerprofil wird erstellt, das linguistische und kognitive Aspekte beinhaltet.

Der Ansatz basiert auf der Theorie, dass das menschliche Gedächtnis in einer semantischen, netzwerkartigen Form organisiert ist. Wenn ein Konzept der Fokus unserer Aufmerksamkeit wird, dann werden auch alle Konzepte, die damit verbunden sind, aktiviert. Im Vergleich zu anderen Ansätzen wie der frequenzbasierten Klassifikation und dem bayes'schen Modell hat der Ansatz von Roberto und Silva [2007] sehr positive Ergebnisse gezeigt.

- **Communities**

Viele Künstler von Universal haben eine Community auf ihrer Webseite. Die zusätzlichen Informationen der Benutzer, wie zum Beispiel Motto, Selbstbeschreibung, Herkunft, Lieblingskünstler, -filme, -spiele, -bücher, können von Universal Music genutzt werden, um umfassende Benutzerprofile der Benutzer zu erstellen.

Liu und Maes [2005] haben in einer Studie eine Methode entwickelt, die Daten über die Benutzer von Online-Communities strukturiert, um diese für semantische Produkt-Empfehlungen zu nutzen. In Online-Communities sieht man nicht nur die Freunde der Mitglieder, sondern auch deren Interessen und Leidenschaften. Da diese Angaben in natürlicher Sprache angegeben wurden, können sie derzeit von Maschinen schwer verstanden werden. Liu und Maes [2005] haben 100.000 Profile von Online-Communities durchsucht und versucht, die Information über die Benutzer zu strukturieren und semantisch aufzubereiten. Für diesen Zweck wurden Ontologien geschaffen, um die verschiedenen Interessens kategorien wie Bücher, Musik und Filme zu strukturieren, damit sie von Maschinen verstanden und verarbeitet werden können. Durch die maschinelle Analyse der Korrelationsmuster zwischen den verschiedenen Interessen und der kulturellen Identitäten, wie zum Beispiel „Raver“, „Hundeliebhaber“ und „intellektuell“, wird eine „InterestMap“ (Interessenskarte) der Benutzer generiert.

Dieses „InterestMap“ liefert sehr genaue und verlässliche Empfehlungen an die Benutzer.

Die „InterestMap“-Ansatz kann bei den Communities von Universal Music eingesetzt werden. Der Prozess, diese „InterestMaps“ zu bilden, erfolgt nach Liu und Maes [Liu und Maes, 2005] folgendermaßen:

1. Community-Profile werden durchsucht.
2. Die Information über die Interessen und Identität der Benutzer wird durch Ontologien vereinheitlicht. Die unstrukturierte Information über die Interessen der Benutzer wird in Kategorien eingeteilt (zum Beispiel Musik, Filme, Sport, Essen) und die Information über die Nutzer wird in einem von Maschinen lesbaren und verarbeitbaren Format dargestellt.
3. Das strukturierte und vereinheitlichte Profil der Benutzer wird mit Metadaten angereichert, um Zusammenhänge festzustellen (zum Beispiel „Madonna“, „Like a Virgin“, „Pop“, „Evita“).
4. Eine maschinelle Lerntechnik wird eingesetzt, um die semantischen Beziehungen zwischen zwei Beschreibungsobjekten zu gewichten.

Das Ergebnis ist ein riesiges semantisches Netzwerk mit Informationsknoten, das die Identität und Interessen der Benutzer beschreibt und die semantischen Zusammenhänge zwischen den verschiedenen Kategorien und Objekten gewichtet.

Wie bereits bei der SWOT-Analyse des Semantic Web beschrieben (Kapitel 2.3), spielt der Datenschutz bei der Erstellung von Benutzerprofilen eine wichtige Rolle.

Unternehmen haben die Pflicht, den Empfänger zu informieren, welche Daten sie von ihm speichern. Als besonders problematisch wird die Speicherung des Klickverhaltens angesehen. Die Benutzer müssen explizit damit einverstanden sein, dass die Unternehmen ihr Klickverhalten gemeinsam mit den persönlichen Daten speichern dürfen [Schwarz, 2004].

Aus der oben genannten Literatur können folgende Empfehlungen für Universal Music abgeleitet werden:

Bei den meisten Newslettern von Universal Music muss man bei der Anmeldung nur die E-Mail-Adresse angeben. Basierend auf Schwarz [2004], sollte Universal nicht mehr Daten bei der Anmeldung abfragen. Viele zukünftige Abonnenten könnten somit abgeschreckt werden und die Bestellung abbrechen. Universal Music könnte bei der Bestätigungs-E-Mail (Opt-In oder Double-Opt-In) einen Link integrieren, den man bestätigen muss, um die Newsletter zu abonnieren. Wenn der Benutzer diesen anklickt, wird sein Abonnement bestätigt und er kann auf der Landing-Page sein Profil mit seinen Musikpräferenzen und weitere beliebige Informationen anlegen. Diese können direkt im RDF-Format gespeichert werden.

Persönliche Daten wie Vorname, Name, Adresse, Telefonnummer, Lieblings-CD und T-Shirt-Größe können über Gewinnspiele in den Newsletter von Universal Music einfließen. Die Gewinnspiele könnten auch im Betreff des Newsletter erwähnt werden und somit als Anreiz für die Abonnenten fungieren.

Nach Liu und Maes [2005] enthalten Online-Communities wertvolle Information für die Profilierung der Internet-Benutzer. Universal sollte versuchen, für die Künstler-Communities intensiver zu werben. Eine Möglichkeit wäre, in den Newslettern einen Link zu integrieren, der auf diejenige Online-Communities verweist, die für den Benutzer relevant sind. Die Relevanz für den Benutzer kann dabei durch Empfehlungssysteme erforscht werden (siehe nächstes Unterkapitel). Dadurch werden einerseits mehr Mitglieder für die Communities gewonnen und Universal kann andererseits mehr Daten und Information über die Abonnenten sammeln.

3.4.3 Anwendungen für die Newsletter von Universal Music

Durch das Semantic Web ergeben sich neue Perspektiven für die Personalisierung von Webinteraktionen. Diese Technologie ermöglicht, dass Web-Inhalte von Maschinen gelesen und verstanden werden und dadurch in verschiedenen Kontexten und Anwendungsszenarien angewendet werden [Henze, 2006].

Universal Music kann mit den semantisch aufbereiteten Daten (siehe Kapitel 3.4.1) und den semantischen Benutzerprofilen (siehe Kapitel 3.4.2) die Inhalte der Newsletter personalisieren und individualisierte Inhalte an jeden Abonnenten schicken. In den folgenden Unterkapiteln werden Anwendungsmöglichkeiten für Universal Music dargestellt.

3.4.3.1 Exkurs: Bestehende Online-Musikdienste

In der Musikindustrie ist der Trend des personalisierten Inhalts in einem kontinuierlichen Wachstum. Verschiedene Initiativen wie Pandora⁵⁸, Last.Fm⁵⁹, Idiomag⁶⁰, iLike⁶¹, Mog⁶² und MyStrands⁶³ verdeutlichen dies. Nach einem Artikel im Spiegel [Patalong, 2007] gilt Last.Fm als die „größte und erfolgreichste europäische Web-2.0-Community überhaupt“.

In den folgenden Absätzen werden einige Beispiele für Online-Musikdienste aufgezählt. Diese arbeiten mit verschiedenen Technologien und Applikationen, um den Benutzern Empfehlungen zu liefern. Im nächsten Unterkapiteln 3.4.3.2 und 3.4.3.3 wird auf die Eigenschaften und Technologien dieser Musikdienste hingewiesen. Daher ist ein Exkurs mit einer Aufzählung dieser Dienste sinnvoll:

- **Last.Fm: ein interaktives Webradio**

Last.Fm ist ein Webradio, das die Benutzer mit Liedern versorgt, die ihrem Geschmack entsprechen. Es verfügt über rund eine Million Musik-Lizenzen und gilt als eine der größten Web-2.0-Communities im Internet [Patalong, 2007]. Dieses kollaborative Empfehlungssystem [Knees et al., 2007] funktioniert folgendermaßen: Man gibt einen Künstler in der Suchmaske ein und das System liefert dazu ein Lied. Weitere Lieder-Empfehlungen werden anhand des angegebenen Lieblings-Künstlers, Empfehlungen von anderen Benutzern und des Feedbacks des Benutzers, der angeben kann, ob ihm die Lieder gefallen oder nicht, generiert [Patalong, 2007]. Das dahinterliegende Empfehlungssystem heißt

⁵⁸ <http://www.pandora.com>

⁵⁹ <http://www.lastfm.com>

⁶⁰ <http://www.idiomag.com>

⁶¹ <http://www.ilike.com>

⁶² <http://www.mog.com>

⁶³ <http://www.mystrands.com>

Audioscrobber und ist im Prinzip eine riesige Musik-Datenbank. Diese verfolgt die Hörgewohnheiten der Benutzer und erstellt Zusammenhänge und Empfehlungen auf Basis der Musik, die sich die Benutzer anhören. Die vom Benutzer angehörten Lieder werden vom Audioscrobber in einem Musikprofil gespeichert und durch statistische Methoden können basierend auf diesen, passende Musik-Empfehlungen geliefert werden [Audioscrobber].

Darüber hinaus können die Benutzer in Last.Fm die Lieder, die sie hören, bewerten [Stein, 2007] und mit Tags oder Labels versehen, um sie zu beschreiben. Diese zusätzlichen Metadaten können genutzt werden, um Suchmechanismen zu verbessern, Daten besser zu strukturieren oder um personalisierte Empfehlungen zu liefern, die an die Interessen der Benutzer angepasst sind [Tso-Sutter et al., 2008]. Andere beliebte Internetdienste wie del.icio.us, Gmail⁶⁴ oder Flickr⁶⁵ verwenden auch die Tagging-Methode, um Bookmarks, E-Mails oder Fotos zu beschreiben.

- **Idiomag: Eine personalisierte Online-Zeitung**

Idiomag⁶⁶ ist eine personalisierte Online-Zeitung über Musik, die die Kategorisierung von Inhalten mit einem innovativen Tagging- und Gewichtungssystem kombiniert [IdiomagPress]. Bei der Anmeldung muss der Benutzer zwei Lieblingskünstler angeben. Falls vorhanden, erfasst das System die Hörgewohnheiten der Benutzer aus verschiedenen Online-Musikdienstleistungen wie Last.fm, MyStrands oder Pandora und liefert dem Benutzer dementsprechend passende Inhalte. Mittels Benutzer-Feedback kann der Benutzer dem System mitteilen, ob ihm das Lied gefällt oder nicht. Falls ihm ein Lied nicht gefällt, kann er die Option ankreuzen, dass dieses Lied nie wieder gespielt werden soll. Die Inhalte der Online-Zeitung Idiomag können mittels eines Widget in eigene Blogs oder in verschiedene Social-Networking-Applikationen integriert werden, wie zum Beispiel in Facebook⁶⁷ [IdioWidget].

⁶⁴ <http://mail.google.com>

⁶⁵ <http://www.flickr.com>

⁶⁶ <http://www.idiomag.com>

⁶⁷ <http://www.facebook.com/apps/application.php?id=4098878913>

Idiomag besitzt auch ein zusätzliches Tool, das APML⁶⁸ (Attention Profiling Markup Language)-Import. Diese Funktion ermöglicht, die Aktivitäten der Benutzer auf anderen Seiten zu erfassen, die die APML-Export-Funktion unterstützen. Mit diesen zusätzlichen Informationen kann man das Verhalten der Benutzer besser kennenlernen, wodurch genauere Empfehlungen möglich werden [Kirkpatrick, 2008].

- **Die Hype-Machine: ein Empfehlungssystem, das auf Blogpublikationen basiert [HypeM]**

Die Hype-Machine⁶⁹ verfolgt Diskussionen über Musik in Blogs, die MP3-Links enthalten, sammelt diese Links in einer Datenbank und zeigt sie auf der Titelseite an. Die Lieder, die am häufigsten in Blogs erwähnt werden, werden vom Server der Hype-Machine abgefangen und von ihr für Empfehlungen verwendet.

Die eingeloggtten Besucher der Webseite können ein Benutzerprofil anlegen, in dem ihre Lieblingslieder als Playlist oder auch Nachrichten aus ihren Lieblingsblogs angezeigt werden. Des Weiteren können die Benutzer Beurteilungen zu den Liedern abgeben und sie als beliebt oder nicht beliebt markieren. Die Lieder aus der Hype-Machine sind nicht zum Download verfügbar. Neben den Liedern gibt es jedoch Links zu den Webseiten Amazon oder iTunes⁷⁰, wo Lieder zum Kauf angeboten werden. Darüber hinaus ist bei jedem Lied auch ein Link zu dem Blog integriert, der das Lied publiziert hat. Somit hat man die Möglichkeit, im Blog weitere Informationen über das Lied zu erhalten und auch weitere Artikel darüber zu schreiben.

- **Musicoverly: ein Empfehlungssystem mit Stimmungsmatrix**

Musicoverly⁷¹ ist ein interaktives und personalisiertes Webradio, in dem die Benutzer mit ein paar Klicks ein Radioprogramm erstellen können, das zu verschiedenen Hörsituationen und den Präferenzen der Benutzer passt. Wie in der Abbildung 24 ersichtlich, können folgende Angaben festgelegt werden, um

⁶⁸ <http://www.apml.org>

⁶⁹ <http://www.hypem.com>

⁷⁰ <http://www.apple.com/itunes>

⁷¹ <http://www.musicoverly.com>

die Musik zu personalisieren: Stimmung (positiv, düster, ruhig oder energiege- laden), zeitliche Einordnung der Musik (1960er, -70er, -80er usw.), Genre, Tempo, Musikrichtung. Das Besondere an dieser Applikation ist die Stim- mungsmatrix, eine durchgehende Fläche, die sich in zwei Dimensionen teilt (positiv-düster und ruhig-energiegeladen) und in der man mit sich mit der Maus positionieren kann. Somit können die Benutzer passend zu deren Stimmung oder zur gesuchten Atmosphäre die Musik ändern [Musicoverypress]. Des Wei- teren kann die Personalisierung durch das Markieren von beliebten und unbeliebten Liedern verfeinert werden [Scholl, 2008].

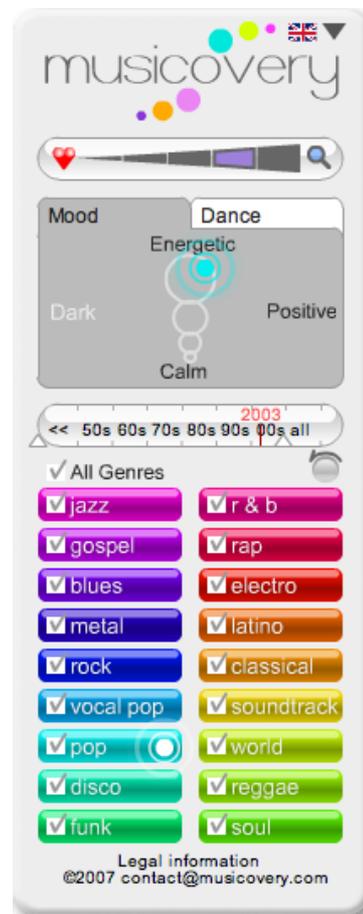


Abb. 24: Screenshot von der Musicoverypress-Applikation [Musicoverypress]

Je mehr Musik mit dem System gehört wird, desto personalisierter wird das Ra- dio. Diese Dienstleistung kann sowohl im Computer als auch mobil genutzt werden. Die Technologie von Musicoverypress kann durch eine Lizenz erworben (in White Label oder Co-Branding) und auf verschiedenen Gebieten angewendet werden, wie zum Beispiel Internetseiten, Spiele, mobile Dienstleistungen, Ra- dios, Marken [Musicoverypress].

- **MyStrands**

Eine Möglichkeit zur Personalisierung der Webinhalte ist der Strands-Recommendier [Strands, 2008], ein Empfehlungssystem, das von Firmen genutzt wird, um Real-time-Produkte und Inhaltsempfehlungen zu liefern, die an die Benutzer angepasst sind. Die Strands-Technologie besteht aus einer sich gegenseitig ergänzenden Kombination aus statistischer Analyse, künstlicher Intelligenz und komplexen Filter-Algorithmen [Strands BS, 2008].

Wie in Abbildung 25 ersichtlich, erfolgt der Prozess des Strands-Recommendier folgendermaßen [Strands BS, 2008]: Die Unternehmen laden ihre Produkt-Kataloge in die Strands-Datenbank hoch. Diese Daten werden mit Produkteigenschaften und Tags versehen, die als Metadaten agieren. Darüber hinaus erfasst dieses System die von Menschen generierten Zusammenhänge zwischen den Objekten, das implizite und explizite Benutzerverhalten (vorangegangene Käufe und das Klickverhalten in Echtzeit) und die Analyse von bestimmten intrinsischen Eigenschaften. Diese Zusammenhänge stellen in aggregierter Form dar, wie die Gesellschaft diese Produkte wahrnimmt, welche Inhalte und Informationen sie erstellt und konsumiert. Zusätzlich können die Benutzer die Empfehlungsfiler kontrollieren und selbst bestimmen, welche Benutzerdaten das System verfolgen soll. Auf Basis dieser Daten liefert der Strands-Recommendier persönliche Empfehlungen für die Benutzer, die kontextsensitiv sind und sich im Laufe der Zeit an Änderungen anpassen. Mit diesem Werkzeug können Empfehlungen in der eigenen Webseite aufgebaut und die Inhalte von Webseiten personalisiert werden.

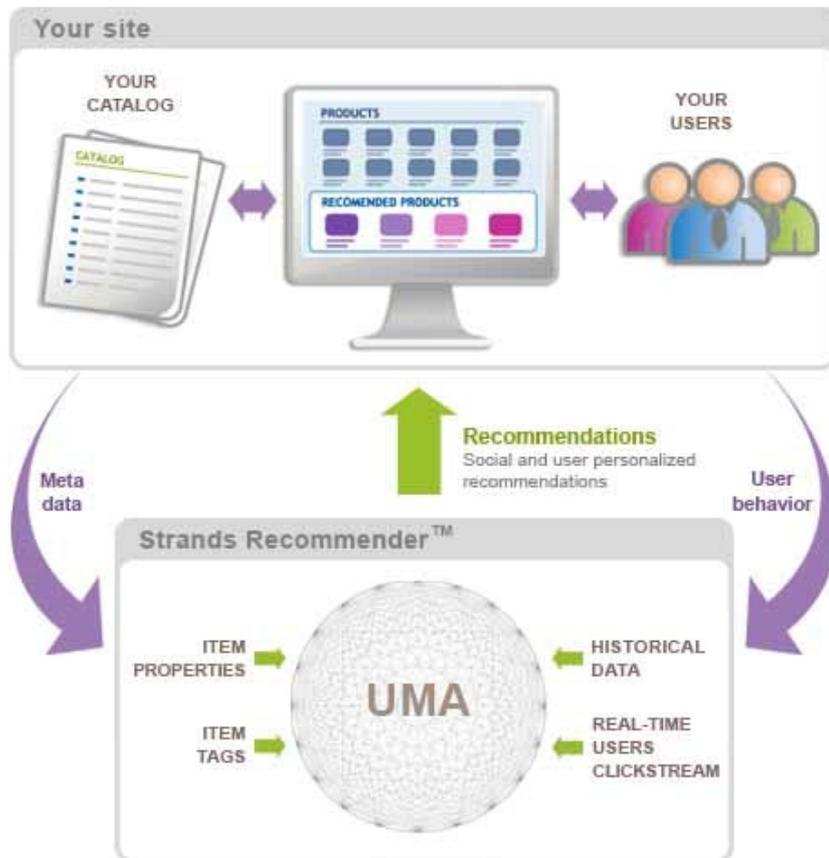


Abb. 25: Architektur des Empfehlungssystems Strands Recommender [Strands BS, 2008]

Diese Strands-Technologie wurde bei der Musik-Community MyStrands.com angewendet. In dieser Plattform können Personen den Musikgeschmack mit anderen Personen teilen und kostenlos Musik-Empfehlungen bekommen. Wie bei Last.Fm kann man auch die Lieder mit Tags versehen, um diese zu beschreiben und schneller wiederzufinden [López Coronado, 2005]. Dieses System profitiert von der „Weisheit“ der Massen. Erstens analysiert es, wie die Benutzer ihre Musik hören und ordnen und bildet daraus Hörprofile, um die Benutzer zu vergleichen und Benutzergruppen mit ähnlichem Musikgeschmack zu finden [Music Ally, 2008]. Durch den Einsatz von verschiedenen Technologien, wie statistisches Lernen, bayes'scher Wahrscheinlichkeitsbegriff, probabilistisches Denken und Visualisierungs-Techniken, ist es dem System möglich, die richtigen Lieder den richtigen Benutzern in Echtzeit zu empfehlen [López Coronado, 2005]. Darüber hinaus können die Benutzer Hörmuster von Freunden ansehen, die laut dem Empfehlungssystem als gleich gesinnt eingestuft wurden [Music Ally, 2008].

- **Strands-Recommendier und Case based Reasoning**

Baccigalupo und Plaza [2006] haben in einer Studie die Strands-Technologie für die Erstellung von sinnvollen Playlists verwendet. Diese Playlists werden anhand von Liedern ausgesucht, die in einer sinnvollen Sequenz angeordnet sind. Dieses Musik-Empfehlungssystem schlägt nicht nur die Lieder vor, die der Benutzer bereits gehört oder gekauft hat, sondern führt eine tiefere Analyse der Lieder durch und empfiehlt ihm Lieder mit einer ähnlichen sequenziellen Struktur. Die empfohlenen Lieder sind akustisch und im Metadaten-Bereich den Lieblingsliedern ähnlich, die der Benutzer in seinem Profil eingetragen hat. Dadurch, dass eine ähnliche sequenzielle Struktur der Lieder im Vordergrund der Empfehlungen steht, kann es vorkommen, dass den Benutzern Lieder empfohlen werden, die sie zwar nicht kennen, aber ihrem Musikgeschmack entsprechen. Daher bietet dieses System die Möglichkeit, neue Musik auszuprobieren.

3.4.3.2 Semantische Empfehlungssysteme

Viele Inhalte der Newsletter von Universal Music können durch semantische Empfehlungssysteme an den Abonnenten angepasst werden. Besonders der Teil der Newsletter mit Empfehlungen und neuen Releases kann gezielt auf die Bedürfnisse der Abonnenten zugeschnitten werden. Durch die Personalisierung der Inhalte kann auch mehr Umsatz generiert werden [Roth und Voss, 2002].

An den verschiedenen Beispielen von bestehenden Online-Musikdiensten (Kapitel 3.4.3.1) wird ersichtlich, dass Empfehlungen anhand verschiedener Methoden getätigt werden können. Diese Filterverfahren können entweder kollaborativ (collaborative) oder inhaltsbasiert (content-based) sein. Mit den Collaborative-Filtering-Methoden können neuen Benutzern Empfehlungen anhand der Präferenzen ähnlicher Benutzer geliefert werden. Zum Beispiel haben Person X und Y einen ähnlichen Musikgeschmack. Das System empfiehlt Person Y dementsprechend die Band „Oasis“, weil X die Beatles mag [Cohen und Fan, 2000]. Andererseits liefern die Content-based-Filtering-Systeme ihre Empfehlungen aufgrund der Korrelation zwischen dem Inhalt der Objekte und den Benutzerpräferenzen [Loizou und

Dasmahapatra, 2006]. Wie bereits im Kapitel 3.4.2 erklärt, kann man Informationen über den Benutzer durch die Informationen, die er selber angibt, sein Klickverhalten in Webseiten und seine Community-Profile sammeln.

Nach Loizou und Dasmahapatra [2006] weisen die Collaborative-Filtering- und die Content-based-Filtering-Methoden einige Defizite auf. Um diese zu überwinden, können die beiden genannten Methoden mit semantischen Technologien angereichert werden. Es gab verschiedene Initiativen [Middleton et al., 2004; Anand und Shapcot, 2007], die sich damit beschäftigt haben, Benutzerprofile mit Ontologien anzureichern. Wenn eine Domain-Ontologie besteht, dann können die Benutzerprofile die Domain-Struktur wiedergeben und eine komplexere Repräsentation schaffen, die es den Maschinen ermöglicht, die Empfehlungen zu „durchdenken“ und dem Benutzer besser zu erklären, wieso sie diese Empfehlungen bekommen haben. Durch die erweiterte „Intelligenz“ der Maschine werden effektivere Empfehlungen ermöglicht [Mobasher, 2007].

Eine weitere Methode, die sehr oft bei Empfehlungssystemen genutzt wird, ist das Social Tagging. Diese Methode wird von Applikationen wie Last.Fm, Idiomag oder dem Strands-Recommendier benutzt. Damit kann man Metadaten in Form von Keywords angeben [Stein, 2007]. Wenn aber diese wertvolle Meta-Information nicht semantisch aufbereitet wird, dann kann sie nur begrenzt sinnvoll genutzt werden [Stein, 2007], da die Bedeutung der Tags vom Computer nicht verstanden wird. Wenn ein Lied mit dem Tag „Queen“ versehen wird, weiß das System nicht, ob es sich um den Sänger Queen handelt oder ob die Königin von England gemeint ist. Damit die Maschinen dies unterscheiden können und auch Zusammenhänge finden, wie zum Beispiel „Queen heißt Freddy Mercury“ und „gehört ins Genre Rock“, benötigt man semantische Technologien. Darüber hinaus kann man die Information zu den Liedern, die durch Tags gesammelt wurden (siehe Kapitel 3.4.3.1), mithilfe von semantischen Technologien strukturieren und damit ermöglichen, dass treffsichere Suchergebnisse oder Empfehlungen von den Maschinen getätigt werden [Stein, 2007].

Universal Music kann mithilfe der semantisch aufbereiteten Daten über das Musikangebot (siehe Kapitel 3.4.1) und die semantische Profilierung der Abonnenten (siehe Kapitel 3.4.2) die Inhalte der Newsletter personalisieren und jedem Abonnen-

ten individualisierte Inhalte schicken. Außerdem können sie verschiedene Tools einsetzen, um die Empfehlungssysteme zielgerichteter zu machen.

3.4.3.3 Applikationen zur Personalisierung der Inhalte der Newsletter

In den folgenden Absätzen werden verschiedene Applikationen vorgestellt, die Universal Music in die Newsletter integrieren kann, um die Personalisierung noch weiter zu optimieren. Diese Vorschläge basieren auf verschiedenen Beispielen der Musikdienste, die im Kapitel 3.4.3.1 genannt wurden.

- Applikation: Anpassung der Struktur der Newsletter

Es wurde ermittelt, dass die Struktur der Webseiten die Auswahlmöglichkeiten der Benutzer einschränkt und indirekt den Ausdruck der Interessen beeinflusst [Roberto und Silva, 2007].

Laut Schackmann [2003] können die Inhalte von Webseiten personalisiert werden, indem man dem Benutzer erlaubt, die Struktur der Webseite zu konfigurieren und das Informationsangebot nach seinen Bedürfnissen zu strukturieren. Darüber hinaus kann dieser Prozess vollautomatisiert erfolgen, indem Fachbereichsmitarbeiter contextual-inference-Regeln definieren, um dem Benutzer auf Basis ihrer Profile Inhalte anzubieten [Schackmann 2003]. Daraus wird ersichtlich, dass Universal Music den Abonnenten der Newsletter einerseits die Option geben sollte, die Struktur der Newsletter nach den eigenen Wünschen zu gestalten. Andererseits kann mit den Daten der semantischen Benutzerprofile (Kapitel 3.4.2) die contextual-inference-Methode genutzt werden, um die Inhalte und Struktur der Newsletter automatisch an die Bedürfnisse der Abonnenten anzupassen. Wenn zum Beispiel ein Abonnent oft die Konzertangebote anklickt, wird er in den nächsten Newslettern die Rubrik Konzerte in einer prominenteren Position sehen.

Die Grundstruktur der Genre-Newsletter von Universal Music besteht in der Regel aus fünf Teilen:

- Persönliche Ansprache
- News aus dem Genre
- Neue Releases: CD bestellen, Lied runterladen, Lied anhören, weitere Infos
- Empfehlungen: CDs, Ringtones, Communities, Newsletter

Wenn möglich, sollten auch folgende Informationen im Newsletter vorkommen:

- Konzertangebote
- Videos
- Gewinnspiele

Nach Schwarz [2004] sollte das Layout von Newslettern einfach und übersichtlich sein. Damit die Genre-Newsletter von Universal Music übersichtlicher werden, sollten nicht mehr als drei Releases, News und Empfehlungen dargestellt werden. Die Konzerte sollten nur für die nächsten 15 Tage angezeigt werden. Mit einem Link „weitere Info“ kann man dann weitere Konzerte anschauen. Maximal ein Video sollte im Newsletter eingebettet werden. Diese Zahlen sind Anhaltspunkte, die je nach Darstellung der Inhalte variieren können.

- **Applikation: Inhalte bewerten**

Erfolgreiche Online-Musikdienste wie Pandora, Last.Fm und Idiomag haben Tools in ihren Plattformen eingesetzt, mit denen die Benutzer festlegen können, welche Songs oder Artikel sie interessieren, nicht interessieren oder welche sie nie wieder hören beziehungsweise sehen wollen (siehe Abbildung 26) [Kirkpatrick, 2008]. Pine et al. [1995] haben ein Software-System namens SMART (for System for Manipulation and Retrieval of Text) entwickelt, mit dem verschiedene Arten von Artikeln empfohlen werden. Die Benutzer konnten dann diese Artikel jede Woche bewerten. Nach vier Wochen hat das System so viel über die Benutzer gelernt, dass diese 80 bis 90 Prozent der empfohlenen Artikel als für sie relevant empfunden haben, im Gegensatz zu den 40 bis 60 Prozent in der ersten Woche.



Abb. 26: Screenshot vom Idiomag-Tool, um Lieder zu bewerten [Idiomag]

Basierend auf den oben genannten Gründen, könnte Universal im Newsletter folgende Buttons bei den Nachrichten, Releases oder Empfehlungen einbinden, mit denen der Abonnent seinen Newsletter nach seinen Bedürfnissen anpassen kann:

- „Interessiert mich“
 - „Interessiert mich nicht“
 - „Dieses Lied überspringen“ (dann noch ein extra Kästchen mit der Option „Immer“)
 - „Will mehr Informationen über diesen Künstler bekommen“
- **Applikation: Emotionen der Kunden erfassen**

Nach Scherer und Zentner [2001] kann Musik jederzeit Emotionen in den Hörern hervorrufen. Aufgrund des Wachstums der Musik-Datenbanken gewinnt die Datenabfrage mit Berücksichtigung der Emotionen der Benutzer immer mehr an Bedeutung [Trohidis et al., 2008]. Es gab viele Initiativen zur Beschreibung der Emotionen, die die Musik in den Personen bewirkt [Trohidis et al., 2008; Oliveira und Cardoso, 2008]. Darüber hinaus gibt es auch Applikationen, die Emotionen des Benutzers berücksichtigen, wie zum Beispiel in Musikempfehlungssystemen [Rui et al., 2007], TV, Radioprogrammen [Musicoverly] oder Musik innerhalb von Spielen [Proud Music].

Die in Kapitel 3.4.3.1 vorgestellte Plattform Musicoverly ist ein solches Beispiel, wo die Lieder mit zusätzlichen Metadaten, wie Stimmungszustand und Musikrichtungen, angereichert werden. Den Stimmungszustand kann man in der Stimmungsmatrix einstellen (Abbildung 27). Durch diese Metadaten, die auch semantisch angereichert sind, kann das System mit hoher Wahrscheinlichkeit

erstaunlich gut passende Musik für den entsprechenden Moment liefern [Stein, 2007].



Abb. 27: Screenshot aus dem Musiccovery-Tool [Idiomag]

Laut oben genannter Literatur wäre es für Universal Music empfehlenswert, Musikstücke mit Metadaten zu versehen, die beschreiben, welche Emotionen oder Stimmungen sie erzeugen. Das sollte beim Prozess der semantischen Annotation der Daten (Kapitel 3.4.1.2) geschehen.

Da es nicht so einfach ist, die Stimmung der Musik zu beschreiben, könnte Universal seine Kunden dazu bringen, dies mittels Online-Spielen oder Fragen zu tun (siehe Kapitel 3.4.1.2). Darüber hinaus kann auch die entsprechende Technologie von Musiccovery durch eine Lizenz erworben werden (in White Label oder Co-Branding).

Schlussfolgernd kann man sagen, dass diese Applikation es ermöglicht, eine Liederauswahl näher an die Bedürfnisse und Gemütszustände der Benutzer anzupassen und dadurch bessere Empfehlungen zu liefern.

- **Applikation: personalisierte Werbung**

Die Online-Zeitung Idiomag liefert ihren Lesern zielgerichtete und relevante Inhalte und gibt Werbetreibenden daher die Möglichkeit, die Zielgenauigkeit und Effektivität der Kampagnen zu erhöhen. Je mehr Informationen sie über die Benutzer haben, desto höher ist die Wahrscheinlichkeit, dass die Werbung ihren Interessen entspricht. Neue Geschäftsmodelle in der Werbeindustrie werden in

Tausend-Kontakt-Preis-Basis (TKP) kalkuliert plus einem zusätzlichen Umsatzanteil, den der Inhaltsanbieter aufgrund der Popularität des angebotenen Inhalts bekommt [IdiomagPress]. Diese neuen Werbearten bieten ein großes Werbepotential für personalisierte Webdienste, wie zum Beispiel die Newsletter von Universal Music.

Wie bereits im Kapitel 2.1.3. erwähnt, ist es das Ziel der intelligenten, kontextbezogenen Werbung, zum Inhalt von Webseiten relevante Werbung zu schalten. Das Ziel dieser Art von Werbung ist, die Werbebotschaften so auszusuchen, dass sie optimal zum Inhalt der Webseite passen. Dies führt zu einer Steigerung der Klickrate und der Umsätze [Broder et al., 2007].

Viele Anwendungen suchen auf den Webseiten nach Schlüsselwörtern, um personalisierte Werbebotschaften zu liefern. Das Semantic Web ermöglicht die Erfassung der Bedeutung und der Zusammenhänge zwischen diesen Schlüsselwörtern und somit zielgerichtete Werbung [Broder et al., 2007].

Universal plant zusammen mit der Media-Agentur Hi-Media Deutschland AG, Genre-Webseiten und Genre-Newsletter zu vermarkten. Derzeit werden nur Standard-Werbeformate angeboten, jedoch ist eine Erweiterung auf individuelle Video-Ad-Formate geplant [Adzine, 2008]. Mithilfe des Semantic Web kann man in den Genre-Newslettern intelligente, kontextbezogene Werbung schalten und dadurch die Konsum- und Kaufbereitschaft erhöhen. Verschiedene semantisch angereicherte Werbeformen kann man dem Kapitel 2.1.3 dieser Arbeit entnehmen.

Besonders interessant für die Personalisierung der Werbung ist die Information in den Social Networks, weil dort ausführliche Profile der Internet-Nutzer vorhanden sind [Broder et al., 2007]. Universal Music verfügt über verschiedene Online-Communities. Viele Mitglieder dieser Communities bekommen auch Genre-Newsletter zugeschickt. Die Information aus diesen Communities kann von Universal Music benutzt werden, um intelligente Werbung zu schalten. Zum Beispiel werden in der Community von Jimi Blue (Abbildung 28) Lieblings-Band, -Song, -Buch und -Film der Mitglieder abgefragt. Anhand des Profils wird dann beispielsweise ersichtlich, dass ein Mitglied gerne Nintendo DS und Nintendo

Wii spielt. Universal Music und Hi-Media Deutschland haben mit diesen Angaben dann die Möglichkeit, Werbung für die neuen Nintendo-Spiele oder Konsolen mit einer hohen Erfolgsquote zu schalten.

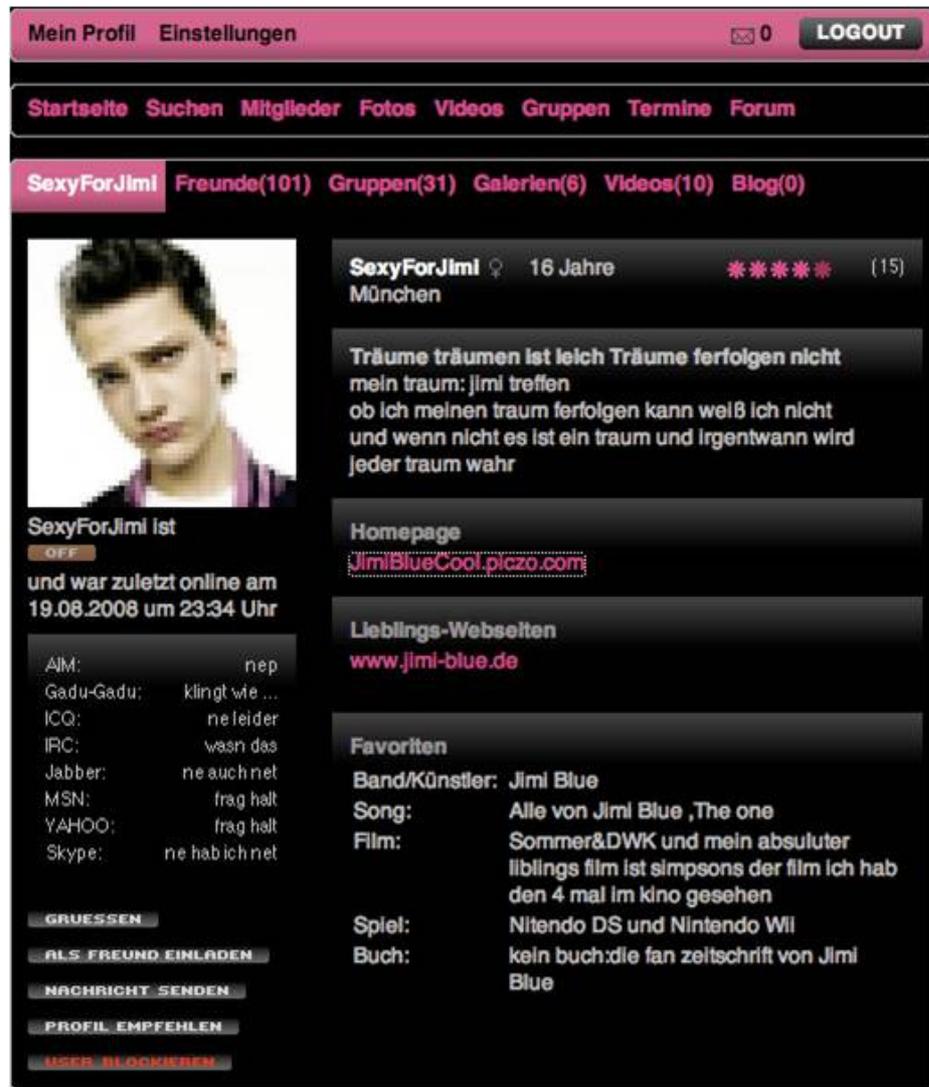


Abb. 28: Beispielprofil der Community von Jimi Blue [JimiBlue]

3.4.3.4 Vorteile für Universal Music

Durch die Schaffung einer semantischen Infrastruktur und die semantische Aufbereitung der Inhalte kann Universal Music effektivere Inhalte und Empfehlungen in den Newslettern platzieren. Nach Loizou und Dasmahapatra [2006] kann man durch semantische Empfehlungssysteme die Änderungen der Benutzer erfassen, was bei den herkömmlichen Systemen nicht der Fall ist. Darüber hinaus lernen diese Systeme den Benutzer besser kennen und verbessern dadurch weiter seine Ergebnisse.

Im Semantic Web können die Benutzerbedürfnisse durch Ontologien besser klassifiziert und Nachfrage-Trends überprüft werden [Gaul und Schmidt-Thieme, 2002].

Personalisierte Empfehlungssysteme haben folgende Vorteile für Universal Music und für die Abonnenten:

- **Benutzerfreundlichkeit**

Durch Empfehlungssysteme werden lange Newsletter mit irrelevanten Inhalten vermieden. Langes Suchen nach relevanten Informationen über Künstler und Lieder wird damit vermieden [Strands BS, 2008].

- **Traffic**

Durch die relevanten Inhalte werden mehr Kunden aufmerksam, mehr Newsletter werden bestellt und auch tatsächlich gelesen [Strands BS, 2008].

- **Umsatz**

Durch die Personalisierung der Angebote erhöht sich die Tendenz, die Webseiten-Besucher zu Käufern zu machen und die Kunden-Loyalität verfestigt sich [Gaul und Schmidt-Thieme, 2002]. Je genauer die Empfehlungen die Bedürfnisse der Benutzer befriedigen können, umso höher ist die Wahrscheinlichkeit des Kaufes [Tusek, 2006].

- **Nutzung des kompletten Musikkatalogs**

Unbekannte Lieder, die aber genau den Musikgeschmack der Abonnenten treffen, werden von den Empfehlungssystemen empfohlen und verkauft. Dieses Phänomen entspricht dem „Long Tail“-Effekt, das heißt Nischenprodukte mit einer höheren Marge werden angepriesen und können Umsatz generieren [Strands BS, 2008].

- **Cross-Selling-Möglichkeiten**

Die Daten über die Benutzerbedürfnisse können auch genutzt werden, um Angebote in anderen Bereichen zu schalten [Gaul et und Schmidt-Thieme, 2002]. Universal kann bei den Empfehlungen weitere Dienstleistungen anbieten, die

nicht mit deren Kerngeschäft zu tun haben, aber den Bedürfnissen der Abonnenten entsprechen, und somit Kooperationen mit anderen Unternehmen schließen.

- **Customer Relationship Management**

Semantische Empfehlungssysteme können als Customer-Relationship-Management-Instrument gesehen werden. Durch die personalisierten Empfehlungen erhöht sich der Zufriedenheitsgrad des Benutzers. Dieses hat wiederum eine Auswirkung auf die Motivation des Benutzers, der die Webseite oder Dienste des Unternehmens öfter beziehen wird [Tusek, 2006]. Somit können gelegentliche Leser der Newsletter von Universal Music zu treuen Kunden gemacht werden, die auch öfter Produkte kaufen, die in den Newslettern angeboten werden.

3.4.3.5 Ausblick

Um die Newsletter von Universal Music wirksamer und effektiver zu gestalten, ist es erforderlich, die Bedürfnisse der einzelnen Abonnenten zu verstehen und zu befriedigen.

Der Aufwand, die semantische Infrastruktur in Universal Music zu schaffen, ist zwar nicht gering, aber lohnenswert. Denn sie muss nur einmal aufgebaut werden und kann dann in vielen verschiedenen Gebieten eingesetzt werden. Die Newsletter sind ein Beispiel, jedoch liegen weitere Möglichkeiten im Einsatz von semantischen Musik-Suchmaschinen oder Empfehlungssystemen in den Online-Shops oder Künstler- und Genre-Webseiten. Das Konzept der intelligenten Werbung kann auch in den Künstler-Communities und -Webseiten eingesetzt werden.

Des Weiteren hat der Einsatz von semantischen Technologien im Wissensmanagement sehr viel Potential. Diesbezüglich sieht Felix Schramm, Concept Developer bei Universal Music [Interv. Schramm], Einsatzmöglichkeiten des Semantic Web bei der Personalsuche und beim innerbetrieblichen Wissensaustausch. Laut Schramm wird der Informationsaustausch durch semantische Technologien effizienter und könnte innerhalb des Unternehmens und in einem erweiterten Netzwerk über den Vivendi Konzern eingesetzt werden. Bezüglich der Personalsuche findet er, dass die Kompe-

tenzen und die Informationen über die Personen durch semantische Technologien besser strukturiert sind und sich daher besser verorten lassen. Er fügt hinzu: *„Wenn Leute in Netzwerken aktiv sind, dann kann man besser einschätzen, wie sie sind, was sie machen, und wo ihre Kompetenzen liegen, aber auch, welche Informationen sie als Person ansammeln und möglicherweise bisher aus verschiedensten Gründen nicht unbedingt weitergeben.“* [Interv. Schramm]

4 Fazit

Die Semantic-Web-Technologie ist nach Berners-Lee et al. [2001] eine Erweiterung des bestehenden Web, indem die Information des Web maschineninterpretierbar dargestellt wird, so dass die „Bedeutung“ der Wörter und Zusammenhänge zwischen verschiedenen Begriffen erkannt werden können.

Diese Technologie hat viele verschiedene Anwendungsgebiete und dient dazu, die Anwendungen intelligenter und effizienter zu gestalten. Durch das Semantic Web kann man das vorhandene Wissen im Web besser strukturieren, entwickeln, verteilen, finden und nutzen. Im E-Commerce-Bereich, besonders im Tourismus, bietet das Semantic Web wesentliche Vorteile: einerseits für Unternehmen, um ihre Angebote effizienter zu platzieren und andererseits für die Benutzer, die umfassende Produktangebote oder -bündel suchen, die genau ihren Bedürfnissen entsprechen sollen. Darüber hinaus hat die Kombination von Web 2.0 und Semantic Web ein großes Potential, um die kollektive Intelligenz des Web zu strukturieren und Zusammenhänge zu erfassen.

Unternehmen können mittels semantischer Technologien ihr betriebliches Wissensmanagement besser strukturieren und dieses nutzen, um komplexe Fragestellungen zu klären, die ein bestimmtes „Verständnis“ und „Intelligenz“ erfordern. Durch semantische Technologien können nicht nur Dokumente schneller gefunden und miteinander verknüpft werden, sondern auch die richtigen Mitarbeiter, die verschiedene Anforderungen erfüllen müssen.

Einige Herausforderungen müssen für den umfassenden Einsatz des Semantic Web noch überwunden werden. Der Einsatz von semantischen Technologien ist derzeit noch mit viel Aufwand verbunden. Im technischen Bereich gibt es noch keine Einigung über eine Standard-Ontologie-Repräsentations-Sprache oder darüber, wie die verschiedenen Ontologien miteinander integriert und angepasst werden können. Darüber hinaus ist der Datenschutz auch ein wichtiges Thema, das stärker thematisiert werden muss. Die Benutzer wollen einerseits, dass ihre Daten geschützt werden, andererseits publizieren sie viele persönliche Daten im Web, sei es durch Communities, Webseiten usw. Eines steht fest: Um zielgerichtete Angebote oder Inhalte zu bekommen, ist auch ein umfassendes Benutzerprofil notwendig. Wenn die

Benutzer also den Mehrwert dieser Dienstleistungen nutzen wollen, müssen sie auch bereit sein, Daten freizugeben.

Am Fallbeispiel Universal Music kann erkannt werden, dass das Semantic Web verschiedene Möglichkeiten bietet, Inhalte zu personalisieren und effizienter zu gestalten. Diese können Unternehmen einen hohen Nutzen bieten, wie zum Beispiel effizienteres Customer-Relationship-Management, erhöhtes Traffic, Cross-Selling-Möglichkeiten, und daher eine Profiterhöhung. Darüber hinaus kann der Aufwand des Einsatzes von Semantic-Web-Technologien durch die Nutzung von bestehenden Ontologien und Metadaten verringert werden.

Die Experten, die im Rahmen der vorliegenden Arbeit interviewt wurden, stehen der semantischen Technologie positiv gegenüber. Das bestehende Web kann mithilfe von semantisch angereicherten Applikationen Prozesse verbessern und beschleunigen. Anbieter und Nachfrager von Informationen profitieren davon. Die Benutzer können einerseits Inhalte beziehen, die an deren Bedürfnisse angepasst werden und die Unternehmen können ihre Daten und Angebote damit besser strukturieren. Die große Herausforderung liegt darin, die nötigen Werkzeuge zu schaffen, damit der Einsatz von semantischen Technologien einfacher und schneller getätigt werden kann und mehr Leute davon profitieren können.

5 Literaturverzeichnis

[Abel et al., 2007]

Abel, Fabian; Mischa, Frank; Henze, Nicola; Krause Daniel; Plappert, Daniel; Siehndel, Patrick (2007): „Group Me! - Where Semantic Web meets Web 2.0“. In Proceedings of the 6th International Semantic Web Conference (ISW 2007), S. 871-878.

[Adida und Birbeck, 2008]

Adida, Ben; Birbeck, Mark (2008): „RDFa Primer - Bridging the Human and Data Webs“. In W3C Working Draft, 20 Juni 2008.
<http://www.w3.org/TR/xhtml-rdfa-primer/#id84448>, abgerufen am 15.09.2008.

[Adzine, 2008]

Adzine news (2008): „Hi-Media vermarktet Online-Angebote von Universal Music“. In Adzine News, 1. Juli 2008.
<http://www.adzine.de/de/site/contentfcmsv44869f6163262b/newsArtikel.xml>, abgerufen am 19.09.2008.

[Aleman-Meza et al., 2007]

Aleman-Meza, Boanerges; Bojars, Uldis; Boley, Harold; Breslin, John G. ; Mochol, Malgorzata ; Nixon, Lyndon JB ; Polleres, Axel; Zhdanova, Anna V. (2007): „Combining RDF Vocabularies for Expert Finding“. In Proceedings of the 4th European Semantic Web Conference (ESWC2007), S. 235–250.

[Anand und Shapcot, 2007]

Anand, Sarabjot Singh; Kearney, Patricia; Shapcott, Mary (2007): „Generating semantically enriched user profiles for web personalization“. ACM Transactions on Internet Technology, 7(4) 22, S. 1533-5399.

[Andersen, 2003]

Andersen, Hans-Christoph (2003): „Service-Orientierung und das Semantic Web“. In Paar, Alexander ; Szeder, Gábor; Gelhausen, Tom; Schanne, Marc: „Grundlagen des Autonomen Rechnens“. Universität Karlsruhe, Institut für Programmstrukturen und Datenorganisation, S. 36-68.

[Audioscrobbler]

Audioscrobbler (2008): „Webseite von Audioscrobbler – The Music Technology Playground from Last.Fm“. <http://www.audioscrobbler.net>, abgerufen am 02.10.2008.

[Baccigalupo und Plaza, 2006]

Baccigalupo, Claudio; Plaza, Enric (2006): „Case-Based Sequential Ordering of Songs for Playlist Recommendation“. In Proceedings of the ECCBR 2006 Conference, S. 286-300.

[Berners-Lee, 1998]

Berners-Lee, Tim (1998) „Semantic Web Road Map“. In W3C Design Issues.
<http://www.w3.org/DesignIssues>, abgerufen am 07.10.2008.

- [Berners-Lee et al., 2001]
Berners-Lee, Tim; Hendler, James; Lassila, Ora (2001): „The Semantic Web“. *Scientific American*, 284 (5), S. 10.
- [Berners-Lee und Miller, 2002]
Berners-Lee, Tim; Miller, Eric (2002): „The Semantic Web lifts off“. In *ERCIM News*, (51), S. 9-10.
- [Blumauer und Fundneider, 2006]
Blumauer, Andreas; Fundneider, Thomas (2006): „Semantische Technologien in integrierten Wissensmanagement-Systemen“. In *Blumauer, Andreas; Pellegrini, Tassilo*: „Semantic Web, Wege zur vernetzen Wissensgesellschaft“. Springer Verlag, Berlin Heidelberg, S. 227-239.
- [Blumauer und Pellegrini, 2006]
Blumauer, Andreas; Pellegrini, Tassilo (2006): „Semantic Web und semantische Technologien: Zentrale Begriffe und Unterscheidungen“. In *Blumauer, Andreas; Pellegrini, Tassilo*: „Semantic Web, Wege zur vernetzen Wissensgesellschaft“. Springer Verlag, Berlin Heidelberg, S. 9-25.
- [Broder et al., 2007]
Broder, Andrei; Fontoura, Marcus; Josifovski, Vanja; Riedel, Lance (2007): „A semantic approach to contextual advertising“. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, S. 559-566.
- [Brückmann, 2007]
Brückmann, Bert (2007): „Web 2.0 - Social Software der neuen Generation“. In *Science Garden, Magazin für junge Forschung*, 1. Februar 2007.
<http://www.sciencegarden.de/content/2007-02/web-20%E2%80%93social-software-der-neuen-generation>, abgerufen am 18.09.2008.
- [Celino et al., 2006]
Celino, Irene; Della Valle, Emanuele; Cerizza, Dario; Turati, Andrea (2006): „Squiggle: a semantic search engine for indexing and retrieval of multimedia content“. In *Proceedings of the 1st International Workshop on Semantic-enhanced Multimedia Presentation Systems*.
- [Chebotko et al., 2004]
Chebotko, Artem; Shiyong, Lu; Fotouhi, Farshad (2004): „Challenges for Information Systems Towards the Semantic Web“. *AIS SIGSEMIS Bulletin* 1(1), S. 24-28.
- [Cohen und Fan, 2000]
Cohen, William W.; Fan, Wie (2000): „Web-collaborative filtering: recommending music by crawling the web“. In *WWW9/ Computer Networks*, 33 (1-6), S. 685-698.
- [Dostal et al., 2004]
Dostal, Wolfgang; Jeckle, Mario; Melzer, Ingo; Zengler, Barbara (2004): „Semantic Web“. In *OBJEKTspektrum*, Mai 2004.

http://www.sigs.de/publications/os/2004/05/dostal_melzer_OS_05_04.pdf, abgerufen am 02.10.2008.

[Duke et al., 2007]

Duke, Alistair; Glover, Tim; Davies, John (2007): „Squirrel: An Advanced Semantic Search and Browse Facility“. In Proceedings of the 4th European Semantic Web Conference (ESWC 2007), S. 341–355.

[Dustdar et al., 2006]

Dustdar, Schahram; Fensel, Dieter; Linder, Markus; Schliefnig, Martin; Otruba, Heinrich; Pellegrini, Tassilo (2006): „The realization of Semantic Web based E-Commerce and its impact on Business, Consumers and the Economy“. In Proceedings of Semantics 2006.

http://www.ocg.at/ak/semantics/files/060426_artikel.pdf, abgerufen am 10.07.2008.

[ebSem AUSTRIAPRO, 2007]

Linder, Markus (2007): „ebSemantic AUSTRIA PRO“. AK Semantic Web, 26. September 2007, Wirtschaftskammer Österreich, Wien.

<http://portal.wko.at/?375008>, abgerufen am 10.07.2008.

[ebSemantics Ont., 2008]

ebSemantics (2008): „ebSemantics II - Ontologien“.

<http://www.ebsemantics.net/doc>, abgerufen am 07.09.2008.

[Esolda, 2008]

Smart Information Systems (2008): „Esolda - Findet was ich will“.

<http://www.esolda.at>, abgerufen am 09.07.2008.

[Fensel et al., 2002]

Fensel, Dieter; Bussler, Christoph; Ding, Ying; Kartseva, Vera; Korotkiy, Maksym; Omelayenko, Borys; Siebes, Ronald (2002): „Semantic Web Application Areas“. In Proceedings of the 7th International Workshop on Applications of Natural Language to Information Systems, S. 27-28.

[Finin, 2007]

Finin, Tim (2007): „AskWiki uses Wikipedia for semantic search“. 3. November 2007. <http://ebiquity.umbc.edu/blogger/2007/11/03/askwiki-uses-wikipedia-for-semantic-search>, abgerufen am 29.09.2008.

[Fittkau & Maaß, 2003]

Fittkau & Maaß Consulting (2003): „WWW-Benutzer-Analyse W3B“. Oktober/November 2003, Fittkau & Maaß Consulting.

<http://www.fittkaumaass.de/reports>, abgerufen am 24.09.2008.

[Focus Online, 2008]

Focus Online (2008): „Schäuble kündigt Verschärfung des Datenschutzes an“. In Focus Online, 5. September 2008.

http://www.focus.de/politik/deutschland/datenschutz-schaeuble-kuendigt-verschaerfung-des-datenschutzes-an_aid_331134.html, abgerufen am 12.09.2008.

- [Garcia und Celma, 2005]
Garcia, Roberto; Celma, Óscar (2005): „Semantic Integration and Retrieval of Multimedia Metadata“. In 5th International Workshop on Knowledge Markup and Semantic Annotation (SemAnnot'05), S. 69-80.
- [Gaul und Schmidt-Thieme, 2002]
Gaul, Wolfgang; Schmidt-Thieme, Lars (2002): „Web Controlling und Recommendersysteme“. In Hippner, Hajo; Merzenich, Melanie; Wilde, Klaus: „Handbuch Web Mining im Marketing - Konzepte, Systeme, Fallstudien“. S. 234-265.
- [Ghani und Fano, 2002]
Ghani, Rayid; Fano, Andrew (2002): „Building Recommender Systems using a Knowledge Base of Product Semantics“. In Proceedings of the Workshop on Recommendation and Personalization in E-Commerce, at the 2nd International Conference on Adaptive Hypermedia and Adaptive Web Based Systems.
- [Giasson und Raimond, 2007]
Giasson, Frédéric; Raimond, Yves (2007): „Music Ontology Specification Specification Document“. 5 Februar 2007. <http://musicontology.com>, abgerufen am 08.09.2008.
- [Gimpel und Dommisch, 2006]
Gimpel, Lea; Dommisch, Katharina (2006): „Die Sinus-Milieus - Konzeption, Kritik und Zukunft“. In Reader zum Seminar Zielgruppenkonzeptionen, Gesellschafts- und Wirtschaftskommunikation an der Universität der Künste Berlin, S. 118-148.
- [Grigoris und Van Harmelen, 2004]
Antoniou, Grigoris; Van Harmelen, Frank (2004): „A Semantic Web Primer“. The MIT Press, Cambridge, Massachusetts.
- [GroupMe!, 2008]
GroupMe! (2008): „GroupMe! - Structure your life“. <http://groupme.org>, abgerufen am 18.09.2008.
- [Gulli und Alessio, 2005]
Gulli, Antonio; Signorini, Alessio (2005): „The indexable Web is more than 11.5 billion pages“. In Proceedings of the 14th International Conference on World Wide Web, S. 902-903.
- [Handschuh et al., 2002]
Handschuh, Siegfried; Staab, Steffen; Ciravegna, Fabio (2002): „S-cream - semiautomatic creation of metadata“. In Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EAKAW 2002), S. 358-372.
- [Handschuh und Staab, 2003]
Handschuh, Siegfried; Staab, Steffen (2003): „Annotation of the shallow and the deep web“. In Handschuh, Siegfried; Staab, Steffen (Ed.), Annotation for the

Semantic Web, Vol. 96 of Frontiers in Artificial Intelligence and Applications, IOS Press, Amsterdam, S. 25-45.

[Harth et al., 2004]

Harth, Andreas; Breslin, John G.; O'Murchu, Ina; Decker, Stefan (2004): „Linking Semantically-Enabled Online Community Sites“. Proceedings of the 1st Workshop on Friend of a Friend, Social Networking and the (Semantic) Web (FOAF Galway), S. 19-29.

[Heflin et al., 2003]

Heflin, Jeff; Hendler, James A.; Luke, Sean (2003): „Shoe: A blueprint for the semantic web“. In Fensel, Dieter; Hendler, James; Liebermann, Henry; Wahlster, Wolfgang : „Spinning the Semantic Web“, S. 29–63. The MIT Press, Cambridge, Massachusetts.

[Henze, 2006]

Henze, Nicola (2006): „Personalisierbare Informationssysteme im Semantic Web“. In Blumauer, Andreas; Pellegrini, Tassilo: „Semantic Web, Wege zur vernetzen Wissensgesellschaft“. Springer Verlag, Berlin Heidelberg, S. 135-146.

[Hepp, 2005]

Hepp, Martin (2005): „eClassOWL: A fully-fledged products and services ontology in OWL“. In Proceedings of the International Semantic Web Conference (ISWC 2005).

[Hepp et al., 2007]

Hepp, Martin; Sayal, Mehmet; Lee, Sang-goo; Lee, Juhnyoung; Shim, Junho (2007): „Data engineering issues in e-commerce and services: DEECS 2007 workshop summary“. In Proceedings of the 3rd International Workshop on Data Engineering Issues in E-Commerce and Services (DEECS 2007), ACM International Conference Proceeding Series, Vol. 236, ACM Press, S. 1-3.

[Herrera et al., 2005]

Herrera, P.; Bello, J.; Widmer, G.; Sandler, M.; Celma, O.; Vignoli, F.; Pampalk, E.; Cano, P.; Pauws, S.; Serra, X. (2005): „SIMAC: semantic interaction with music audio contents“. In Proceedings of the 2nd European Workshop on the Integration of Knowledge, Semantic and Digital Media Technologies, S. 399-406.

[Hunter, 2003]

Hunter, Jane (2003): „Enhancing the Semantic Interoperability of Multimedia through a Core Ontology“. IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Conceptual and Dynamical Aspects of Multimedia Content Description, (13) 1, S. 49 - 58.

[HypeM]

The Hype Machine (2008): „Webseite von The Hype Maschine“. <http://hypem.com/about>, abgerufen am 03.10.2008.

[Idiomag]

Idiomag (2008): „Webseite von Idiomag - personalized music magazine“. <http://www.idiomag.com>, abgerufen am 30.09.2008.

[IdiomagPress]

Idiomag (2006): „Idio launches personalized digital magazine“. 1. November 2006. In Idiomag - Pressemitteilung. <http://www.idiomag.com/files/releases/011006-Launch.pdf>, abgerufen am 20.09.2008.

[IdioWidget]

Idiomag (2008): „Webseite über die Widgets auf der Online-Zeitung Idiomag“. <http://www.idiomag.com/widget>, abgerufen am 03.10.2008.

[Interv. Boser]

Boser, Siegfried Christian (2008): „Persönliches Interview, geführt vom Verfasser“. 10. September 2008, Wien.

[Interv. Ege]

Ege, Börtecin (2008): „Telefonisches Interview, geführt vom Verfasser“. 20. September 2008.

[Interv. Hochhaltinger]

Hochhaltinger, Thomas (2008): „Persönliches Interview, geführt vom Verfasser“. 10. September 2008, Wien.

[Interv. Linder]

Linder, Markus (2008): „Persönliches Interview, geführt vom Verfasser“. 10. September 2008, Wien.

[Interv. Möller]

Möller, Christoph (2008): „Telefonisches Interview, geführt vom Verfasser“. 18. September 2008.

[Interv. Oelling]

Oelling, Alexander (2008): „Telefonisches Interview, geführt vom Verfasser“. 19. September 2008.

[Interv. Rothkamp]

Rothkamp, Julian (2008): „Persönliches Interview, geführt vom Verfasser“. 18. August 2008, Berlin.

[Interv. Schramm]

Schramm, Felix (2008): „Persönliches Interview, geführt vom Verfasser“. 15. September 2008, Berlin.

[Interv. Szpuszta]

Szpuszta, Mario (2008): „Persönliches Interview, geführt vom Verfasser“. 9. September 2008, Wien.

[Interv. Voss]

Voss, Jan (2008): „Telefonisches Interview, geführt vom Verfasser“. 24. Sep-

tember 2008, Berlin.

[iSense, 2008]

iSense (2008): „iSense: Semantisches Targeting“. <http://www.isense.net>, abgerufen am 12.10.2008.

[ITWissen, 2008]

ITWissen (2008). „Das große Online-Lexikon für die Informationstechnologie“. <http://www.itwissen.info>, abgerufen am 30.09.2008.

[JimiBlue]

Universal Music (2008): „Community von Jimi Blue“. <http://community.jimi-blue.de>, abgerufen am 21.08.2008.

[John und Drescher, 2006]

John, Michael; Drescher, Jörg (2006): „Semantische Technologien im Informations- und Wissensmanagement“. In Blumauer, Andreas; Pellegrini, Tassilo: „Semantic Web, Wege zur vernetzen Wissensgesellschaft“. Springer Verlag, Berlin Heidelberg, S. 241-255.

[Kalyanpur et al., 2003]

Kalyanpur, Aditya; Hendler, James; Parsia, Bijan; Golbeck, Jennifer (2003): „SMORE - semantic markup, ontology, and RDF editor“. Technical report, University of Maryland.

[Kilpatrick, 2008]

Kilpatrick, Kristi (2008): „ZEMANTA HELPS BLOGGERS UP THEIR GAME BY DELIVERING CONTENT FROM ACROSS THE WEB IN REAL-TIME“. In Zemanta - Pressemitteilung. 3. Juni 2008, San Francisco und London. <http://www.zemanta.com/press>, abgerufen am 19.08.2008.

[Kirkpatrick, 2008]

Kirkpatrick, Marshall (2008): Idiomag: „Sweet Online Music Magazine Now With Attention Data Import“. In ReadWriteWeb, 7. Januar, 2008. http://www.readwriteweb.com/archives/idiomag_apml.php, abgerufen am 02.10.2008.

[Knees, 2007]

Knees, Peter; Pohle, Tim; Schedl, Markus; Widmer, Gerhard (2008): „A music search engine built upon audio-based and web-based similarity measures“. In Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '07), S. 447-454.

[Knight, 2008]

Knight, Kristina (2008): „Twitter, Summarize deal could give Social Marketers more push“. In BizReport, 21. Juli 2008. http://www.bizreport.com/2008/07/twitter_summarize_deal_could_give_social_marketers_more_push.html, abgerufen am 20.09.2008.

[Lei und Horrocks, 2003]

Li, Lei; Horrocks, Ian (2003): "A software framework for matchmaking based on semantic web Technology". In Proceedings of the Twelfth International World Wide Web Conference (WWW 2003), S. 331–339.

[LinkingOpenData, 2008]

Linking Open Data (2008): „Project Description“. In Linking Open Data - W3C SWEO Community Project.
<http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>, abgerufen am 20.09.2008.

[ListGame]

Listen Game (2008): „Webseite von Listen Game“. <http://www.listengame.org>, abgerufen am 30.09.2008.

[Liu und Maes, 2005]

Liu, Hugo; Maes, Pattie (2005): „InterestMap: Harvesting social network profiles for recommendations“. In Beyond Personalization - IUI 2005, S. 54-59.

[Loizou und Dasmahapatra, 2006]

Loizou, Antonis; Dasmahapatra, Srinandan (2006): "Recommender Systems for the Semantic Web". In Proceedings of the ECAI 2006, S. 76-81.

[López Coronado, 2005]

López Coronado, Octavi (2005): „MusicStrands uses artificial intelligence to recommend music to site visitors“. In Innovations Report - Forum of Science, Industry and Business, 13. April 2005. http://www.innovations-report.com/html/reports/communication_media/report-42941.html, abgerufen am 10.08.2008.

[Lu et al., 2002]

Lu, Shiyong; Dong, Ming and Fotouhi, Farshad (2002): "The Semantic Web: opportunities and challenges for next-generation Web applications". Information Research 7(4).

[Mayr, 2006]

Mayr, Thomas; Bauernfeind, Ulrike; Zins, Andreas (2006): "A Conceptual Model for Quality Dimensions for B2C Recommender Systems". In Proceedings of the 14th European Conference on Information Systems (ECIS).

[McKinsey, 2008]

McKinsey (2008): „Building the Web 2.0 Enterprise“. McKinsey Global Survey Results.

[Meier und Stormer, 2005]

Meier, Andreas; Stormer, Henrik (2005): „eBusiness & eCommerce“. Springer, Berlin Heidelberg.

[Meißner et al., 2007]

Meißner Ronny; Müller, Andreas; Schieder, Miriam (2007): „Semantic Wikis“. Technische Universität Dresden, Fakultät Informatik, Lehrstuhl Multimedia-technik. <http://www-mmt.inf.tu->

dresden.de/Lehre/Sommersemester_07/Hauptseminar/hs-proceedings.pdf,
abgerufen am 29.09.2008.

[Middleton et al., 2004]

Middleton, Stuart E.; Shadbolt, Nigel R. ; De Roure, David C. (2004); "Ontological user profiling in recommender systems". *ACM Transactions on Information Systems*, 22 (1), S. 54-88.

[Mobasher, 2007]

Mobasher, Bamshad (2007): „Recommender Systems“. In *Künstliche Intelligenz, Special Issue on Web Mining*, Nr. 3, BottcherIT Verlag, Bremen, S. 41-43.

[Music Ally, 2008]

Music Ally (2008): „Nokia plotting Apple showdown“. In *The REPORT*, Issue 157, 14. Dezember 2006. www.mystrands.com/corp/files/musically_dec06.pdf, abgerufen am 19.09.2008.

[MusicOnt]

Music Ontology (2008): "Webseite von Music Ontology Specification". <http://musicontology.com>, abgerufen am 28.09.2008.

[Musicoverly]

Musicoverly (2008): "Webseite von Musicoverly". <http://www.musicoverly.com>, abgerufen am 30.09.2008.

[MusicoverlyPress]

Musicoverly (2008): „Pressemitteilung von Musicoverly“. In *Musicoverly – Pressemitteilung*. <http://www.musicoverly.com/pressRelease/PressKitMUSICOVERY.doc>, abgerufen am 02.10.2008.

[myOnt]

MyOntology (2008): "Webseite von MyOntology - Open Ontology environment for Semantic Web-based E-Commerce". <http://www.myontology.org>, abgerufen am 09.07.2008.

[MyStrands]

MyStrands (2008): "Webseite von MyStrands". <http://www.mystrands.com>, abgerufen am 30.09.2008.

[Oliveira und Cardoso, 2008]

Oliveira, António; Cardoso, Amílcar (2008): "Controlling Music Affective Content: A Symbolic Approach". In *Conference on Interdisciplinary Musicology*, Thessaloniki.

[Passant, 2008]

Passant, Alexandre (2008): „Case Study: Enhancement and Integration of Corporate Social Software Using the Semantic Web“. In *WWW3.com*, Juni 2008, Electricité de France R&D and LaLIC, Université Paris-Sorbonne, Paris. <http://www.w3.org/2001/sw/sweo/public/UseCases/EDF>, abgerufen am 19.09.2008.

[Patalong, 2007]

Patalong, Frank (2007): „Das Geschmacks-verstärkte Radio“. In Spiegel Online, 30. Mai 2007. <http://www.spiegel.de/netzwelt/web/0,1518,485632,00.html>, abgerufen am 01.10.2008.

[Pellegrini und Blumauer, 2006]

Pellegrini, Tassilo; Blumauer, Andreas (2006): „Semantic Web, Wege zur vernetzen Wissensgesellschaft“. Springer Verlag, Berlin Heidelberg.

[Pine et al., 1995]

Pine, Joseph B. II; Peppers, Don; Rogers, Martha (1995): „Do you want to keep your customers forever?“. In Harvard Business Review, Jg. 73, Nr. 2, S. 103-114.

[Presseportal, 2007]

Presseportal (2007): „Online-Shopper beim Kauf von Fernsehern, Digi-Cams & Co häufig überfordert“. In Presseportal.de, 11. Oktober 2007. http://www.presseportal.de/pm/68459/1063409/smart_information_systems_gmbh, abgerufen am 20.10.2008

[Proud Music]

Proud Music (2008): „Webseite von Proud Music Library“. <http://www.proudmusiclibrary.com>, abgerufen am 06.10.2008.

[Raimond et al., 2007]

Raimond, Yves; Abdallah, Samer; Sandler, Mark; Giasson, Frederick (2007): „The music ontology“. In Proceedings of the International Conference on Music Information Retrieval, September 2007, S. 417–422.

[RDF, 2006]

Tauberer, Joshua (2006): „What Is RDF“. In XML.com, 26. Juli 2006. <http://www.xml.com/lpt/a/2001/01/24/rdf.html>, abgerufen am 27.06.2008.

[Reif, 2006]

Reif, Gerald (2006): „Semantische Annotation“. In Blumauer, Andreas; Pellegrini, Tassilo: „Semantic Web, Wege zur vernetzen Wissensgesellschaft“. Springer Verlag, Berlin Heidelberg, S. 405-346.

[Roberto und Silva, 2007]

Roberto, Rafael Liberato; Silva, Sérgio Roberto P. (2007): „An Approach for Identification User's Intentions During the Navigation in Semantic Webseites“. In Proceedings of the 4th European Semantic Web Conference (ESWC2007), S. 371-383.

[Roth und Voss, 2002]

Roth, Michael; Voss, Jan-Martin (2002): „Web Mining Application Services Providing - Erfahrungen und Erfolgsfaktoren“. In Hippner, Hajo; Merzenich, Melanie; Wilde, Klaus: „Handbuch Web Mining im Marketing - Konzepte, Systeme, Fallstudien“. S.472-485.

[Rui et al., 2007]

Cai, Rui; Zhang, Chao; Wang, Chong; Zhang, Lei; Ma, Wei-Ying (2007): „Mu-

sicsense: contextual music recommendation using emotional allocation modeling. In Proceedings of the 15th international conference on Multimedia. S. 553-556.

[Sadeh, 2002]

Sadeh, Norman M. (2002): "M-Commerce: Technologies, Services, and Business Models". John Wiley & Sons, Inc., New York.

[Schackmann, 2003]

Schackmann, Jürgen (2003): „Ökonomisch vorteilhafte Individualisierung und Personalisierung - Eine Analyse unter besonderer Berücksichtigung der Informationstechnologie und des Electronic Commerce“. Verlag Dr. Kovac.

[Scherer und Zentner, 2001]

Scherer, Klaus R.; Zentner, Marcel R. (2001): „Emotional effects of music: Production rules“. In Juslin, Patrick N.; Sloboda, John A.: "Music and emotion: Theory and research". Oxford, Oxford University Press, S. 361–392.

[Scholl, 2008]

Scholl, Christoph (2008): "Musiccovery im Test: Webradio mal anders". In netzwelt.de, <http://www.netzwelt.de/news/75400-musiccovery-im-test-webradio-mal.html>, abgerufen am 02.10.2008.

[Schwarz, 2004]

Schwarz, Torsten (2004): „Leitfaden eMail Marketing und Newsletter-Gestaltung“, Waghäusel.

[Seagran, 2007]

Segaran, Toby (2007): "Programming Collective Intelligence: Building Smart Web 2.0 Applications". O'Reilly Media.

[SearchMonkey, 2008]

SearchMonkey (2008): „SearchMonkey Guide: A Manual for SearchMonkey Developers and Publishers“. Yahoo! Inc., <http://developer.yahoo.com/searchmonkey>, abgerufen am 19.09.2008.

[SemWeb, 2008]

Altova (2008): „Was ist das Semantic Web?“. http://www.altova.com/de/semantic_web.html, abgerufen am 04.07.2008.

[Sheshagiri et al., 2004]

Sheshagiri, Mithun; Sadeh, Norman; Gandon, Fabien (2004): "Using Semantic Web Services for Context-Aware Mobile Applications". In MobiSys 2004, Workshop on Context Awareness, Boston.

[Sinus-Socio, 2007]

Sinus-Sociovision (2007): „Sinus Milieus- Deutschland 2007“. <http://www.sinus-sociovision.de>, abgerufen am 20.08.2008.

[Siorpaes und Hepp, 2007]

Siorpaes, Katharina; Hepp, Martin (2007): „OntoGame: Towards Overcoming the Incentive Bottleneck in Ontology Building“. In Proceedings of the 3rd Inter-

national IFIP Workshop On Semantic Web & Web Semantics (SWWS '07) co-located with OTM Federated Conferences, S. 1222-1232.

[Spath, 2008]

Spath, Markus (2008): „Calais 2.0: Semantischer Baukasten für Blogger und Verleger“. In *Netzwerkartig.com*, 20. Mai 2008.

<http://netzwerkartig.com/2008/05/20/calais-20-semantischer-baukasten-fuer-blogger-und-verleger/#more-768>, abgerufen am 19.09.2008.

[Spitzer, 2005]

Spitzer, Ursula (2005): „Recommender Systeme im E-Commerce“. In Abteilung für Informationswirtschaft, Wirtschaftsuniversität Wien.

[Stein, 2007]

Stein, Armin (2007): „*Semantic Web vs. Web 2.0*“. In Grob, H. L. und Vossen, G.: „Entwicklungen im Web 2.0 aus technischer, ökonomischer und sozialer Sicht“. Arbeitsberichte des Kompetenzzentrums Internetökonomie und Hybridität, Nr. 51, S. 13-22.

[Strands, 2008]

Strands (2008): „Strands Recommender“.

<http://www.strands.com/corp/recommender.vm>, abgerufen am 22.08.2008.

[Strands BS, 2008]

Strands (2008): „Strands Business Solutions“. <http://bizsolutions.strands.com>, abgerufen am 22.08.2008.

[Studer et al., 2005]

Studer, Rudi; Schnurr, Hans-Peter; Nierlich, Andreas (2005): „Semantik für die nächste Generation Wissensmanagement“. Institut AIFB, Universität Karlsruhe. <http://www.community-of-knowledge.de/pdf/f05.pdf>, abgerufen am 08.07.2008.

[Swartz, 2002]

Swartz, Aaron (2002): „MusicBrainz: A Semantic Web Service“, In *IEEE Intelligent Systems*, Intelligent Web Services, Januar/Februar 2002.

[Tazzoli et al., 2004]

Tazzoli, Roberto; Castagna, Paolo; Campanini, Stefano Emilio (2004): „Towards a semantic wiki wiki web“. In *Demo Session at ISWC2004*.

[Thuraisingham, 2007]

Thuraisingham, Bhavani (2007): „Building Trustworthy Semantic Webs“. Auerbach Publications.

[TrendONE, 2008]

TrendONE (2008): „Schlüssel-trend Smartwebte digital intellingee expansion“. TrendONE, Berlin.

[Trohidis et al., 2008]

Trohidis, Konstantinos; Tsoumakas, Grigorios; Kalliris, George; Vlahavas, Ioannis (2008): "Multilabel Classification of Music into Emotions". In *Proceedings of*

the International Conference on Music Information Retrieval (ISMIR 2008), S. 325-330.

[Tsinaraki, 2007]

Tsinaraki, Chrisa (2007): "Ontology-Driven Interoperability for MPEG-7". In Proceedings of the DELOS-Multimatch Interoperability Workshop.

[Tsinaraki et al., 2004]

Tsinaraki, Chrisa; Polydoros, Panagiotis; Christodoulakis, Stravros (2004): "Integration of OWL Ontologies in MPEG-7 and TV-Anytime Compliant Semantic Indexing". In Proceedings of the 16th Int'l Conf. Advanced Information Systems Eng. (CAiSE), S. 398-413.

[Tso-Sutter et al., 2008]

Tso-Sutter, Karen H.L., Balby Marinho, Leandro; Schmidt-Thieme, Lars (2008): "Tag-Aware Recommender Systems by Fusion of Collaborative Filtering Algorithms". In Proceedings of 23rd Annual ACM Symposium on Applied Computing (SAC'08), S. 16-20.

[Turnbull et al., 2007]

Turnbull, Douglas; Liu, Ruoran; Barrington, Luke; Lanckriet Gert (2007): "A game-based approach for collecting semantic annotations of music". In Proceedings Intl. Symp. Music Information Retrieval (ISMIR 2007), S. 535-538.

[Tusek, 2006]

Tusek, Jasna (2006): „Semantic Web - Einführung, wirtschaftliche Bedeutung, Perspektiven“. VDM, Müller, Saarbrücken.

[UMD, 2008]

Universal Music Deutschland (2008): „Unternehmen - Universal Music Deutschland“. <http://www.universal-music.de>, abgerufen am 28.08.2008.

[UniUlm, 2006]

Universität Ulm (2006): „Mobile Informationsdienste“. Universität Ulm, Abteilung Künstliche Intelligenz, Juni 2006.
http://www.uni-ulm.de/aktuelles/aktuelles_thema/aktuell0606/index.html, abgerufen am 10.08.2008.

[Versandhaus, 2008]

Versandhaus (2008): „Wie Online-Beratung die Umsätze von Online-Shops erhöht“. In Der Versandhausberater Spezial, 14. Januar 2008.
<http://smartassistant.eu/presse.html>, abgerufen am 08.07.2008

[Völkel et al., 2006]

Völkel, Max; Krötzsch, Markus; Vrandečić, Denny; Haller, Heiko; Studer, Rudi (2006): "Semantic Wikipedia". In Proceedings of the 15th international conference on World Wide Web, WWW 2006, Edinburgh.

[W3C FAQs]

World Wide Web Consortium (2008): „W3C Semantic Web Frequently Asked Questions“. <http://www.w3.org/RDF/FAQ>, abgerufen am 04.07.2008.

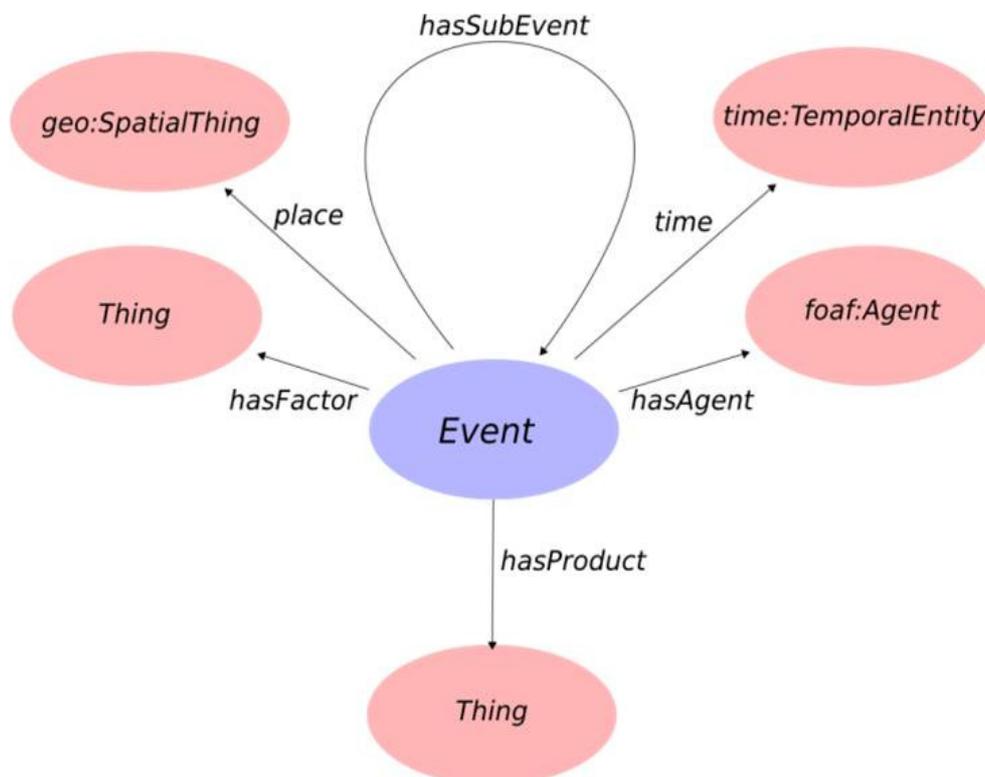
- [W3C RDF, 1999]
W3C (1999): „Pressemitteilung: W3C gibt Recommendation für Resource Description Framework (RDF) frei“. 24. Februar 1999, <http://www.w3c.at/RDF-REC.html>, abgerufen am 25.09.2008.
- [Wahlster, 2007a]
Wahlster, Wolfgang (2007): „Von Suchmaschinen zu Antwortmaschinen: Semantische Technologien und Benutzerpartizipation im Web 3.0“. In Mattern, F.: „Mehr als Stichworte: Wie Arbeiten die Suchmaschinen von morgen?“. Acatech-Schriftenreihe, Berlin, München.
- [Wahlster, 2007b]
Wahlster, Wolfgang (2007): SmartWeb: „Ein multimodales Dialogsystem für das semantische Web“. In Reuse, B., Vollmar, R.: „Informatikforschung in Deutschland“, Heidelberg, Berlin, Springer, 2007.
- [Web Inhalte, 2008]
Telekom (2008): „Semantic Web: Interpretation von Web-Inhalten“. In Telekom - Pressemitteilung, 22. April 2008. http://www.telekom-presse.at/channel_internet/news_32755.html, Bassersdorf, abgerufen am 06.07.2008.
- [West et al., 1999]
West, Patricia; Ariely, Dan; Bellman, Steve; Bradlow, Eric; Huber, Joel; Johnson, Eric; Kahn, Barbara; Little, John; Schkade, David (1999): „Agents to the Rescue“. Marketing Letters, 10 (30). Kluwer Academic Publishers.
- [Weihs, 2005]
Weihs, Eric (2005): „Zur Anwendung "intelligenter" Suchmaschinen zur Vermittlung von Umweltdaten“. In Knetsch, Gerlinde; Lehmann, Angela: „Umweltdatenbanken und Netzwerke“, Umweltbundesamt.
- [Wiki MusicOnt, 2008]
Music Ontology (2008): „Class Schemas – Music Ontology Wiki“. http://wiki.musicontology.com/index.php/Classes_Schemas, abgerufen am 09.09.2008.
- [WKÖ, 2008]
Wirtschaftskammer Österreich (2008): „Startschuss eines zukunftsweisenden Internet-Projekts für die Tourismusbranche“. 26. März 2008. http://www.ots.at/presseaussendung.php?schluessel=OTS_20080326_OTSO101&ch=technologie, abgerufen am 16.08.2008.
- [Zemanta, 2008]
Zemanta (2008): „Zemanta service launches worldwide“. 26. März 2008, Ljubljana. <http://www.zemanta.com/press>, abgerufen am 20.08.2008.

6 Anhang

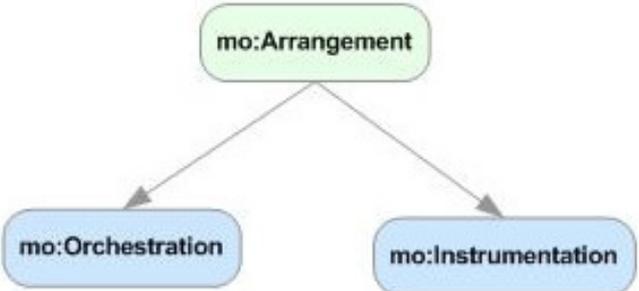
6.1 Ontologien von Music Ontology

Die Music-Ontology-Spezifikation wurde von Giasson und Raimond [2007] entwickelt. Folgende Klassen-Schemas zeigen die Verknüpfungen zwischen Klassen von externen Ontologien und die Musikontologie. Die rosaroten Ellipsen sind Super-Klassen externer Ontologien, die grünen Ellipsen sind Super-Klassen der Music-Ontology-Ontologie, die blauen Ellipsen sind Subklassen der Musik-Ontology-Ontologie und die gelben Ellipsen sind individuelle Instanzen der Music-Ontology [Wiki MusicOnt, 2008].

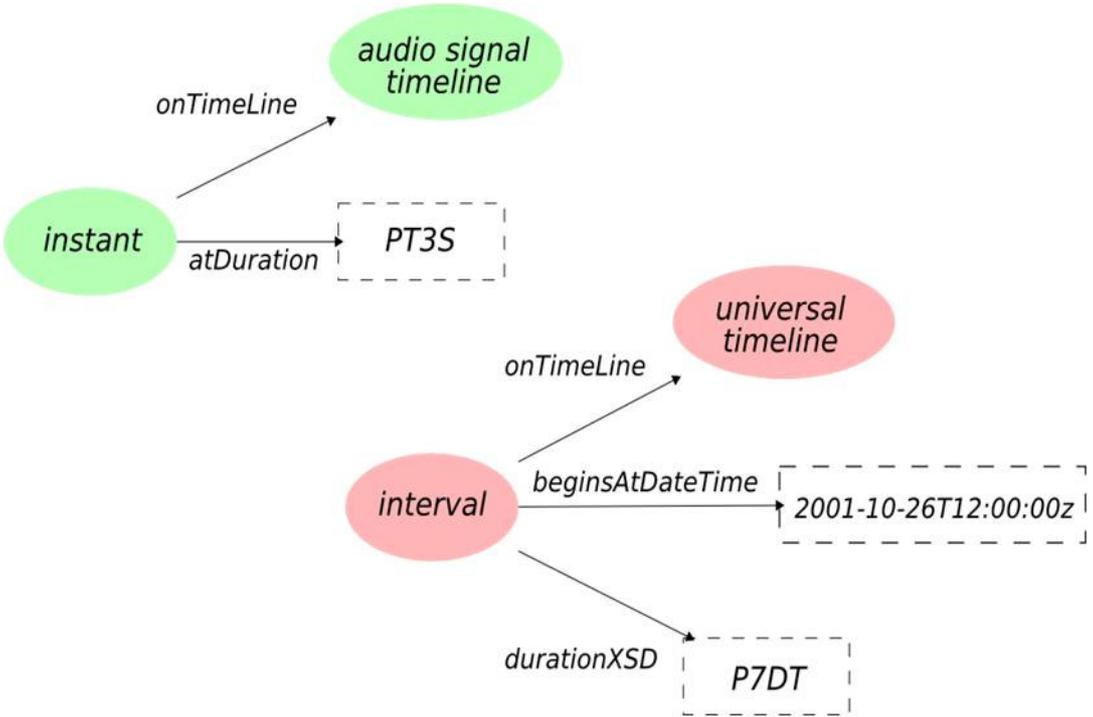
6.1.1 Event-Ontologie



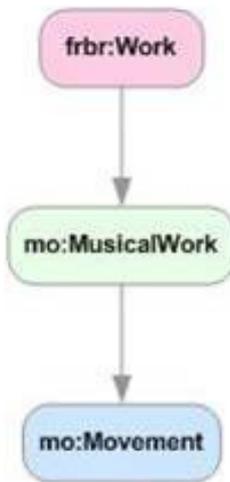
6.1.2 Gestaltungs-Ontologie



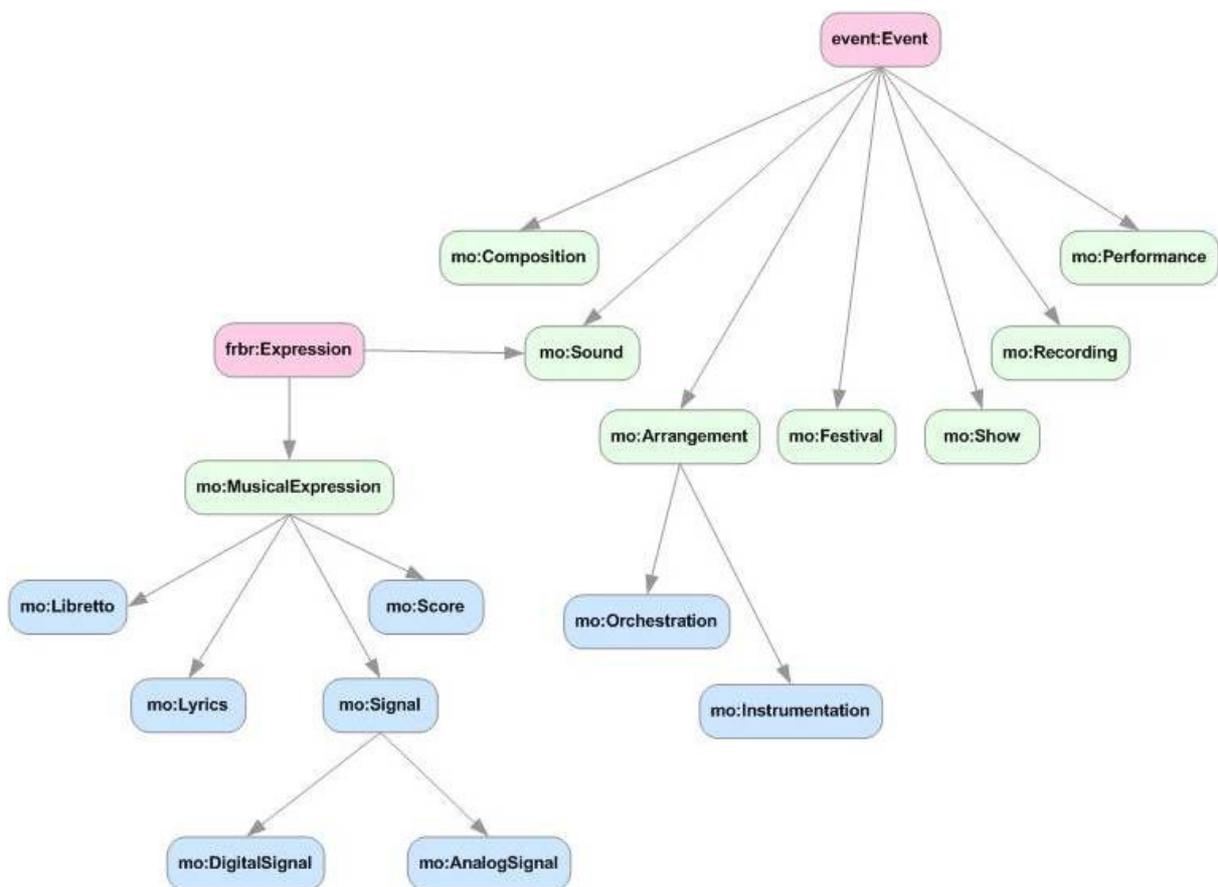
6.1.3 Zeitleisten-Ontologie



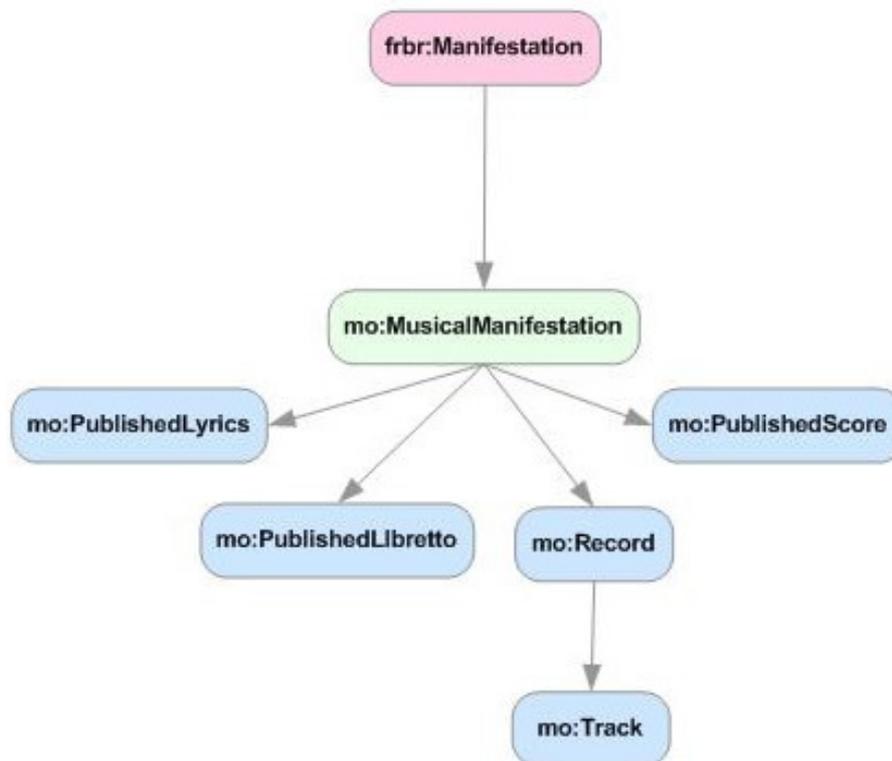
6.1.4 Musikalisches-Werk-Schema



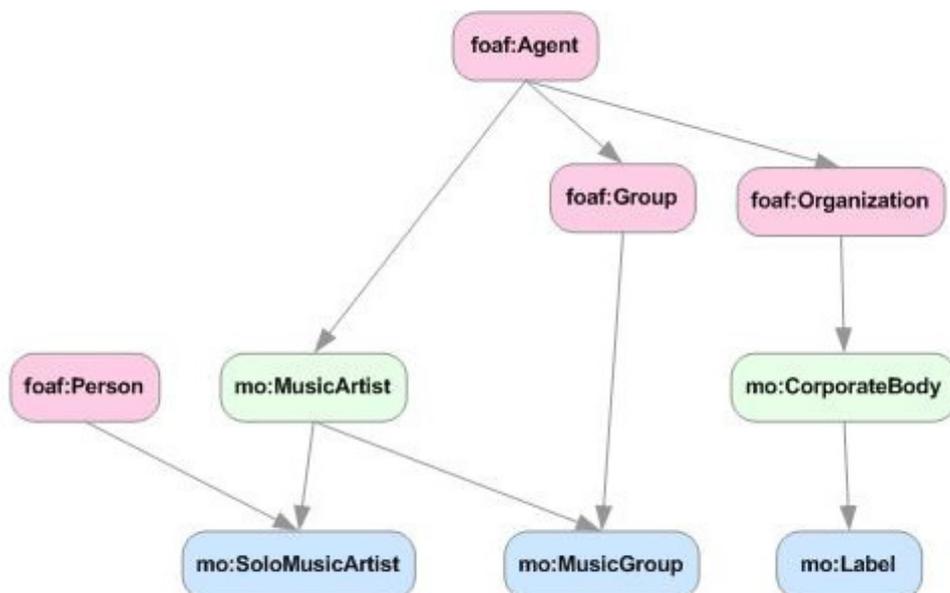
6.1.5 Ontologie Musikalischer Ausdruck und Event-Schema



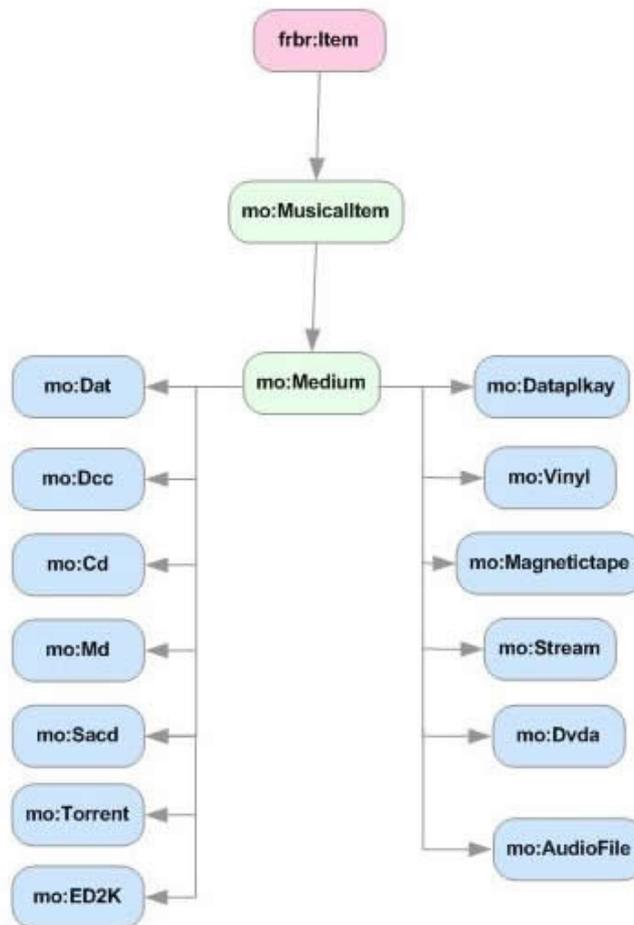
6.1.6 Veröffentlichungs-Schema



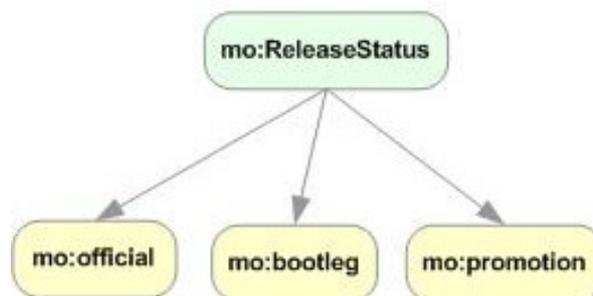
6.1.7 Musikkünstler-Ontologie, Musikgruppen-Ontologie und Organisations-Schema



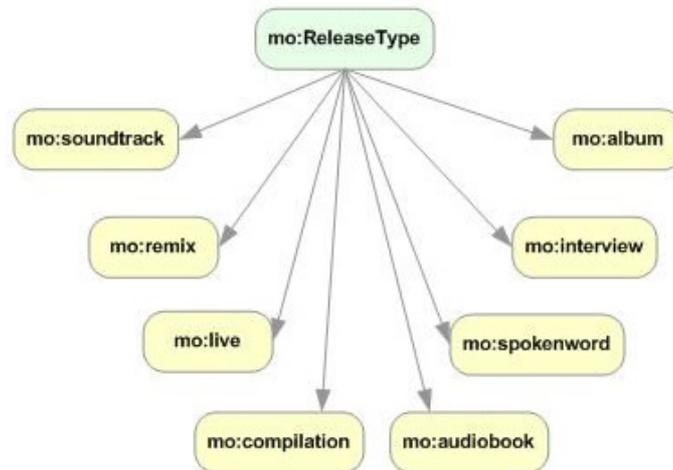
6.1.8 Musikobjekte-Schema



6.1.9 Veröffentlichungsstatus-Ontologie



6.1.10 Veröffentlichungsart-Ontologie



6.2 Interviews

Im folgenden Abschnitt werden fünf Interviews, die im Rahmen dieser Arbeit vom Verfasser durchgeführt wurden, dargestellt. Ziel dieser Interviews war, Informationen für verschiedene Teile dieser Arbeit zu gewinnen, wie zum Beispiel für die SWOT-Analyse und die Fallstudie über Universal. Weitere informelle Interviews wurden mit den Mitarbeitern der Abteilung Business Development von Universal Music geführt, um die Richtung und Inhalte der Fallstudie festzulegen und Information über die Tätigkeiten von Universal Music zu bekommen.

6.2.1 Interview mit Felix Schramm, Concept Developer bei Universal Music Deutschland (vom 15.09.2008, Berlin)

Wie ist die Einstellung von Universal Music zum Semantic Web?

Unsere Einstellung zum Semantic Web ist im Grunde sehr positiv. Wir befinden uns in der Orientierungsphase. Obwohl das Thema Semantic Web nicht unsere Priorität ist, glaube ich, dass ein generelles Bewusstsein innerhalb des Unternehmens herrscht, dass man diesen Trend nicht verschlafen darf. Die Musikindustrie aber auch andere Akteure haben den mit dem Web 2.0 einhergehenden Wandel weitgehend verpasst. Die überwiegende Zahl der Betreiber von Communities, im Bereich Musik sind beispielsweise keine zuvor etablierten Unternehmen. Erst langsam ziehen Konzerne wie Universal Music nach und haben es dementsprechend schwerer sich auf diesem Gebiet zu behaupten. Einmal mehr hat sich hier erwiesen, dass Innovation selten aus dem Zentrum kommt, sondern eher vom Rand und seinen häufig flexibleren und mutigeren Akteuren ausgeht. Diese neuen Akteure haben mit dem Erfolg des Web 2.0 eine Position gewonnen mit der man heutzutage rechnen muss.

Nun geht es darum Konstellationen zu finden um sich professionell behaupten zu können, also z.B. Kontexte zu schaffen, die unter diesen veränderten Bedingungen funktionieren, aber auch Partnerschaften einzugehen. Für uns eine Herausforderung die vor dem Einsatz semantischer Technologien steht. Um das semantische Web aktiv zu gestalten, braucht man Ressourcen, die wir momentan nicht haben. Dennoch existiert ein vitales Interesse, weil wir sehen wie wesentlich das Ganze für uns wird.

Wie viel beschäftigt sich Universal Music mit dem Thema Semantic Web?

Zur Zeit haben wir informell ein vergleichsweise kleines Team von etwa 6 Personen, die vor allem zu Think Tanks und zu Veranstaltungen zum Thema Semantic Web fahren. Das ist momentan die einzige Alternative die wir haben, um uns dem Thema anzunähern. Es geht darum fehlende Strukturen wie zum Beispiel eine Entwicklungsabteilung durch gezielte Expertise zumindest teilweise auszugleichen. Des Weiteren könnten wir nur schwer Agenturen für ein Thema verpflichten, das derzeit noch relativ schwer zu greifen ist. Für uns gilt es daher in erster Linie eine Orientierung innerhalb eines so komplexen Themas und der unzähligen Anwendungsgebiete und -möglichkeiten zu gewinnen. Dazu zählt auch die Suche nach potentiellen Partnern z.B. Technologieträgern, mit denen wir zusammenarbeiten wollen um Applikationen zu schaffen.

Welcher Antrieb steht dahinter Semantic Web Technologien bei Universal Music einzusetzen wie zum Beispiel bei den Newslettern?

Wir haben zahlreiche Newsletter, die Genre-Newsletter, die Künstler-Newsletter, usw. Die Problematik von Newslettern generell ist, dass ihre Leseraten tendenziell zurückgehen. Ein Grund dafür ist, dass sich relevante Information mit einer unglaublichen hohen Zahl an sonstiger Information vermengen und man sie in dieser Informationsdichte nur schwer findet. Gerade unsere Genre-Newsletter sind auf bestimmte Milieus oder Subkulturen ausgerichtet die in einer postmodernern Gesellschaft wie unserer kaum noch so trennscharf zu definieren sind, wie das einst der Fall war. Solche einheitlichen und dementsprechend umfangreichen Newsletter werden den Abonnenten zugesendet. Unter Umständen hat dieser zusätzlich mehrere Newsletter abonniert so dass er eine Fülle an zum Beispiel Produkt oder Artist-Informationen erhält. Ganz zu schweigen davon, dass wir nur eines von diversen Unternehmen sind die den Kunden tagtäglich mit Informationen bombardieren. An dem Punkt AN DEM DER Nutzer unseren Newsletter nicht mehr als Mehrwert für sein Informationsbedürfnis empfindet, sollte man sich überlegen ob man ihn derartig zuspammt. Das bringt uns nichts, weil der Nutzer eher ein negatives Grundgefühl gegenüber Universal Music bekommen wird und sich davor hüten wird weitere Angebote zu abonnieren. Einen tatsächlichen Mehrwert können wir nur durch die Steigerung der Relevanz von Informationen im Hinblick auf die individuellen Bedürfnisse des Nutzers erreichen.

Im Bezug auf den Newsletter beabsichtigen wir den Kunden noch spezifischere Angebote zu machen. Angebote die eine Relevanz für ihn haben, und über Verortungssysteme wie Lifestyle-Segmente hinaus seine individuellen Interessen berücksichtigen. Bislang beruht die inhaltliche Zusammensetzung der Newsletter vor allem auf Erfahrungswerten welche die Musikindustrie zweifelsohne über viele Jahre lang gesammelt hat. Durch angedeutete Entwicklungen werden die Milieus und

Subkulturen jedoch sehr stark aufgebrochen und diffundieren mehr und mehr ineinander. Für wirkungsvolle Alternativen die den Nutzer als Individuum fokussieren sind semantischen Technologien sind eine entscheidende Erneuerung.

Welche Vorteile hat Universal Music und Ihre Kunden wenn Semantic-Web-Technologien eingesetzt werden?

Ein Vorteil liegt wahrscheinlich darin, dass wir mithilfe des Semantic Webs unsere Promotion und Marketing Tools effektivieren können. Bezogen auf die Reichweite, also die Anzahl der Abonnenten, sind wir relativ schlagkräftig. Nun müsste es darum gehen unsere Instrumente mittels semantischer Technologien effizienter zu machen.

Der andere Vorteil ist, dass wir mit semantischen Technologien zum Beispiel im Bereich Smart Advertising, Kontexte ganz genau verorten und dadurch Ressourcen sparen könnten. Bei der Werbung, die wir schalten, müssen nicht mehr auf die Masse gehen, wenn wir wissen, es klickt eh nur ein geringer Teil an. Genau um diesen kleinen Personenkreis geht es jedoch und auf eben diesen kann man sich nun konzentrieren. Darüber hinaus, können Werbetreibende in unseren Newslettern, Communities und Web-Angeboten ihre Werbung schalten. Durch semantische Technologien wie Empfehlungssysteme können wir auch unseren Werbekunden einen besseren Service bieten.

Welche Nachteile entstehen für Universal Music und für Ihre Kunden wenn Semantic-Web-Technologien eingesetzt werden?

Dadurch dass Informationen so spezifisch und punktgenau, ausgeliefert werden, werden andere Angebote unter Umständen der menschlichen Interpretation entzogen. Bei Google gibt man die Wörter „Semantic Search“ ein und man bekommt 100 Suchergebnisse. Manchmal sind auch Sachen dabei, die gar nicht mit der Frage zu tun haben. Aber trotzdem können diese Links die Personen mehr inspirieren, als wenn eine spezifische Antwort angezeigt wird. Das sehe ich als ein generelles Problem des Semantic Webs.

Eine andere Geschichte ist die Analyse von Nutzerverhalten. Wenn man den Nutzer, nur noch spezifische Angebote macht, dann bedeutet dies gleichzeitig, dass die Vielfalt, von Angeboten in der wir das Nutzerverhalten messen können, verkleinert wird. Wenn man zum Beispiel eine komplexe „Generell interest Site“ hat, ist die Navigation des Kunden durch diese Informationen letztlich eine breite und wertvolle Informationsquelle über dessen Nutzerverhalten. Mit Lösungen wie sie die amerikanische Firma Semanticator anbieten wird diese Informationsquelle möglicherweise eingeschränkt.

Ein anderer Punkt den ich persönlich am wesentlichsten finde ist der, dass man die Entscheidungskompetenz, zum Beispiel welche Information auf welche Frage zutrifft, auf die Maschine ausgelagert. Diese Auslagerung, des Interpretationsprozesses an die Maschine, ist letztlich auch immer analysier- und dadurch angreifbar. In dem Moment, in dem man also versteht wie er funktioniert, kann man ihn letztlich auch umgehen. Wenn du deine Inhalte also mit falschen Semantikbeschreibungen/ Metadaten bestückst, umgehst du den semantischen Filterprozess und kannst dein Angebot sehr prominent platzieren. Auf diese Art und Weise wird auch das semantische Netz durch Formen des Spaming ermöglichen

Wo liegen die Chancen des Semantic Webs?

Global gesehen sind die Einsatzbereiche vielfältig. Semantic Search ist ein Bereich, Smart Advertising ein anderer.

Eine Chance liegt daran den Datenüberfluss im Web zu bewältigen. In einer Zeit in der sich Nutzerbedürfnisse immer mehr ausdifferenzieren, stellen zielgenaue Angebote eine große Chance dar. Durch die konstante Weiterentwicklung von Nutzerbedürfnisse lösen sich auch die Personen aus den klassischen Milieus oder Subkulturen und springen zwischen den verschiedenen Bereichen hin und her.

Wir sehen im Semantic Web Chancen, die Information mit denen wir unsere Kunden versorgen, beispielsweise wie zum Beispiel in Newsletter, zu optimieren. Wir wollen aber auch die Kontexte, die wir selber erstellen und vermarkten, wie zum Beispiel Communities oder Webseiten optimieren.

Eine weitere essentielle Anwendungsmöglichkeit sehen wir im Bereich Wissensmanagement. Ein sehr wichtiger und spannender Bereich. Vor allem, um die interne Kommunikation zu verbessern und die Komplexität eines Unternehmens wie Universal Music bewältigen zu können, werden mit semantischen Technologien ganz klare Vorteile erwachsen.

Wo liegen die Risiken des Semantic Webs?

Der Datenschutz könnte Probleme bringen. Es ist zwar interessant den Nutzer zu profilieren und Benutzerprofile zu erstellen, aber in dem Moment wo die Anonymität des Nutzers nicht mehr gewährleistet ist, halte ich es doch für sehr problematisch. Da muss man aufpassen. Gerade Deutschland ist im Thema Datenschutz glücklicherweise relativ weit vorne, im Gegensatz zu einigen Ländern im asiatischen Raum. Die Datenschutzproblematik kann man jedenfalls nicht aus dem Auge verlieren.

Welche Semantic-Web-Applikationen oder -Anwendungen könnten Sie sich vorstellen, bei Universal Music einzusetzen?

Das Semantic Web könnte, wie erwähnt, bei Universal beispielsweise eingesetzt werden um einen effizienteren Informationsaustausch innerhalb des Unternehmens aber auch in einem erweiterten Netzwerk wie z.B. anderen zum Vivendi Konzern gehörenden Unternehmen zu gewährleisten

Ein anderes Szenario ist die Suche nach Personal innerhalb und außerhalb des Unternehmens. Personalmanagement und Entwicklung könnten mithilfe semantischer Technologien wesentlich gefördert werden, Kompetenzen lassen sich besser verorten und entwickeln. Auch auf aktuelle Entwicklungen könnte ein Unternehmen dadurch schneller reagieren. Wenn Leute in Netzwerken aktiv sind, dann kann man besser einschätzen wie sie sind, was sie machen, und wo ihre Kompetenzen liegen aber auch welche Informationen die sie als Person ansammeln und möglicherweise bisher aus verschiedensten Gründen nicht unbedingt weitergeben.

Wie viele andere Unternehmen haben wir in diesem Bereich jedoch zuerst weiterhin Basis-Arbeit zu leisten, um die Akzeptanz und das Bewusstsein für den Mehrwert eines derartigen Wissenstransfers zu schaffen. Die Instrumente der Stunde sind also Wikis, Projektplattformen, oder Diskussionsforen. Erst auf der Basis solcher Instru-

mente ist die Idealkonstellation, also ein semantisches Wissensmanagement überhaupt denkbar.

Inwiefern verdrängt das Semantic Web bestehende Geschäftsmodelle und wird durch neue ersetzt?

Wenn Suchergebnisse so genau werden, dann fällt beispielsweise das Suchmaschinenmarketing unter dem Tisch. Also ein grundsätzlicher Teil des Geschäftsmodells von Google. Sucht man etwas erhält man Suchergebnisse und Werbung. Auch wenn ich gestehen muss, dass eine Ablösung der Google Suche durch semantische Suchen mittelfristig nicht vorstellbar ist.

Für uns ist es erstmals nicht so dramatisch. Ich denke, dass das Semantic Web in unserem Fall aber auch allgemein eher unterstützend wirkt. Grundlegend ist jedoch ein Umdenken erforderlich. Informationsangebote die semantische Technologien ausklammern, werden zunehmend unattraktiv weil sie bezogen auf die Suche nach relevanten Informationen nicht länger konkurrenzfähig sind.

6.2.2 Interview mit Mario Szpuszta, Strategic Architect bei Microsoft Österreich (vom 09.09.2008, Wien)

Wie ist der Bezug von Microsoft Österreich zum Semantic Web?

Microsoft hat generell Interesse und befasst sich mit den ganzen modernen Entwicklungen im Software- und Services-Umfeld, vor allem in Zusammenhang mit Web 2.0 und Economies of Scale. In manchen Bereichen wie Software und Services investieren wir viel und sind „Thought Leader“. In anderen Bereichen sind wir das eher nicht. Semantic Web zählt zu diesen Bereichen wo Microsoft Österreich sich nicht als „Thought Leader“ sieht. Wir sehen die Sache da sehr pragmatisch, dort wo es Sinn macht Semantic-Web-Konzepte einzusetzen, dort spricht nichts dagegen. Da unterstützen wir technologisch auch entsprechend die Standards rund um RDF und unsere Produkte und Technologien sind auch so aufgebaut, dass man sie leicht mit Semantic Web Standards (und anderen Standard) integrieren kann. Dort wo es nicht sinnvoll ist, werden wir auch nichts in diesem Bereich investieren. Wir sehen dieses Thema ganz allgemein und sehen es in der Plattform weniger als eine Top-Priorität. Wir unterstützen es und sehen es als eine interessante Entwicklung im Markt. Wir haben aber selber keine Semantic-Web-Produkte, die direkt auf diese Konzepte aufbauen und werden in absehbarer Zeit auch keine derartigen Produkte oder Technologien am Markt anbieten. Dennoch stellen wir sicher, dass unsere Produkte und Technologien so offen gestaltet sind, dass sich Aspekte wie etwa Semantic Web Formate dort wo erforderlich und sinnvoll, leicht integrieren lassen.

Vom Standpunkt Microsoft Österreichs aus betrachtet sehen wir durchaus einige Bereiche, wo es sinnvoll erscheint, auf die Konzepte und Ideen, die man im Semantic Web bearbeitet, zu setzen. Aber vielleicht gar nicht so unmittelbar auf das Web beschränkt, sondern viel mehr aus einem Paradigma, welches man im Umfeld der „User-Experience“ immer stärker wahrnimmt. Was ich damit meine, ist das Konzept der Kontext-sensitiven Anwendungen. Egal ob es sich dabei um Client Applikationen,

Server Applikationen oder Web-Browser basierte Applikationen handelt – es geht immer darum, dass sich Anwendungen entsprechend der Situation in der sich der Benutzer befindet auch die Bedürfnisse und Anforderungen eines Benutzers automatisch anpassen und darauf reagieren. Microsoft sieht semantische Technologien durchaus als eine Möglichkeit, derartiges Verhalten in Applikationen umzusetzen. Dort betrachten wir es auch als durchaus sinnvoll, die Ideen und auch die Technologiestandards die dahinter verbergen, einzusetzen.

Welche semantische Applikationen hat Microsoft Österreich umgesetzt?

Ein Beispiel wo wir gemeinsam mit Smart Information Systems und AUSTRIAPRO Wirtschaftskammer gearbeitet haben, ist die Schaffung einer Wirtschaftsplattform für Tourismusapplikationen, die sich Kontext-sensitiv Verhalten. Das spannende dabei ist, dass man damit zum Beispiel Applikationen entwickeln kann, die sich anhand von Benutzerprofilen, von Benutzerverhalten und von anderen Eigenschaften der Endbenutzer gezielte, auf den Kontext abgestimmte Offerings, machen können. Diese Anpassungen können teilautomatisiert auf Basis der Informationen, die über Semantic Web im Hintergrund erfasst wurden, basieren.

Voraussetzung für dieses Projekt war, dass Tourismus-Ontologien vorhanden sind. Mit myOntology, eine Web 2.0 basierte Plattform, in der man solche Ontologien mit einer Community aufbauen kann, war diese Voraussetzung gegeben. Mit dieser Voraussetzung haben wir konkret mit Smart Information Systems und AUSTRIAPRO an einen Prototyp gearbeitet, einem Plug-In für Microsoft Office Produkte, über welches die Hotels die Möglichkeit haben, Ihre Offerings auf Basis der Charakteristika der Ontologien von myOntology zu beschreiben und publizieren. Die Frage dahinter war, wie kann ich das Semantic Web in Tools integrieren, in der alle Benutzer in den Branchen jeden Tag bearbeiten. Jemand beschreibt einen Text in Word oder in einer E-Mail in Outlook. Der bestehenden Text, welchen der Benutzer mit Word oder einem anderen Office-Tool schon für andere Zwecke (Broschüre, Briefe, Werbung etc.) erfasst hat, wird durch das Microsoft Plug-In mit den Charakteristika der Ontologien von myOntology verknüpft. Daraus entsteht ein Maschinen-lesbares Format im Dokument im Hintergrund, welches damit von einer Anwendung automatisiert verarbeitet werden kann, wenn das Dokument auf einer Tourismusplattform veröffentlicht wird. Wir haben als Prototyp das Plug-In mal für Microsoft Word entwickelt. Das heißt, ich kann jedes beliebige Word-Dokument hernehmen und kann es mit Ontologien und Attributen aus diesen Ontologien, die zum Beispiel in myOntology definiert sind, verknüpfen. Der Prototyp nützt das neue Office Open XML File Format (ein ISO Standard für Office Dokumente) und bettet ein RDF-XML ein. Die Inhalte des RDF-XML werden mit den Inhalten des Dokuments selbst verknüpft. Das hat für den Benutzer den Vorteil, dass er das Dokument einmal schreibt und die Charakteristika seines Offerings entsprechend den Ontologien verknüpfen kann. Und somit entsteht ein RDF-Dokument, der die Verknüpfung mit dem Semantic Web darstellt.

Was sind die Voraussetzungen um personalisierte Inhalte an die Benutzer anzubieten?

Um kontext-sensitives Verhalten von Applikationen zu erreichen, die sich aus bestimmten Aktionen des Benutzer heraus oder aus bestimmten Profilinformationen heraus ableiten lassen kann, braucht man ein allgemeines Verständnis, das auch technisch interpretierbar ist, über das Umfeld in dem sich der Benutzer bewegt. Am Beispiel Tourismus brauche ich irgendeine Form von allgemeiner Infrastruktur, die

mir es ermöglichen generelle Eigenschaften über Tourismusangebote, Eigenschaften auf den Kontext der Benutzer mappen lassen, um ihm Kontext-sensitiv-Angebote anbieten zu können. Dafür brauche ich allgemeines Verständnis auf konzeptioneller Ebene über Angebote die es gibt, über Charakteristika von Angeboten die es gibt, die dann zur gezielten Filterung von Informationen und Angeboten herangezogen werden können. Und die muss man mal schaffen - im Vorfeld. Im Semantic Web Bereich werden derartige Informationen nach meinem Verständnis in Ontologien gefasst, die versuchen Begrifflichkeiten zu formulieren und mit Informationen verbinden.

Ich glaube, dass man nicht ein umfassendes Benutzerprofil braucht, um Kontext-sensitive Angebote an die Benutzer anzubieten. Zum Beispiel, der Benutzer navigiert ohne sich anzumelden in einer Tourismusplattform 20 Minuten lang. Selbst wenn er das Anonym macht, dann kann ich eventuell, wenn ich ein intelligentes System habe, unter der Art und Weise wie der Benutzer sich durch navigiert, Rückschlüsse ziehen, wonach er sucht. Ohne dass ich persönlich Daten über ihn erfasst habe.

Worin liegen die Vorteile des Semantic Webs?

Für mich ist der Vorteil ganz klar, dass ich Applikationen entwickeln kann, die aufgrund ihres Kontext-sensitiven Verhaltens genauer auf meine Bedürfnisse als Endbenutzer reagieren und mir nicht so wie die typischen Suchmaschinen Millionen von Ergebnissen aus einer Suche anzeigen, die mich im Moment überhaupt nicht interessieren. Ich glaube, dass der klassische Ansatz der Websuche langfristig aufgrund der Vielfalt von Informationen nicht erfolgreich sein kann. Da sehe ich ganz andere Konzepte und mit dem Semantic Web kann ich strukturiert Beziehungen zwischen Informationen aufbauen, um gezielter den Benutzer aus seiner Situation heraus Ergebnisse präsentieren zu können, die eher das treffen was er will und wonach er sucht. Das ist für mich der Hauptvorteil. Aber – Semantic Web ist da aber aus meiner Sicht nur ein Mittel zum Zweck.

Ein anderes Mittel zum Zweck, dass man da ganz klar sieht, ist die ganze Thematik rund um Social Networking, Web 2.0. Der Grund weshalb sich die großen Key Player in der Industrie wie Google, Yahoo oder Microsoft sehr stark an diese Social Networks á la Xing, Facebook, etc. interessiert, liegt klar darin, dass sich viele Aspekte der Informationssuche dorthin verlagern werden. Die Suche in einem sozialen Netzwerk mit Peers, die ähnliche Interessen haben, kann extrem schnell zu den passenden Ergebnissen führen. Spannend ist es nun, wenn man das Semantic Web und das Web 2.0 verknüpft. Wenn ich in meinem Social Network einen Begriff mit einem Attribut aus einer Ontologie verknüpfen kann und wenn dann der nächste meinen Beitrag liest und der Browser intelligent genug ist, um diese Information, die damit mit einem Attribut aus einer Ontologie verknüpft sind, wieder über Querverbindungen mit anderen Dokumenten oder Inhalten aufzulösen. Das macht es ganz spannend.

Semantic Web ist ein Thema welches im Hintergrund agiert. Dem Benutzer interessiert Semantic Web überhaupt nicht. Wenn der Benutzer eine Information hat, interessiert es ihm nur, Querverbindungen zu den Informationen zu erhalten, die damit in Verbindung stehen, die für ihm auch in seinem aktuellen Kontext interessant sind. Der Mehrwert für den Benutzer ist das Kontext-sensitive-Verhalten der Applikation. Was da im Hintergrund teilautomatisch über Verknüpfungen, über Ontologien

oder zielgerichtete Indizierung oder Suchmaschinen passiert, das sind technische Aspekte, die dem Benutzer eigentlich nicht interessieren.

Worin liegen die Nachteile des Semantic Webs?

Das schwierigste ist die Definition der Ontologie. Wenn ich eine Infrastruktur schaffen will, die den Benutzer zielgerichtet Informationen aus seinem Kontext heraus anbieten möchte. Nehmen wir mal an, er sucht nach einer Digitalkamera, die bestimmte Kriterien erfüllt. Das ist eine Kontext-sensitive Suche. Für Kameras wird man sich dafür auf verschiedene Basis-Kriterien einigen müssen, nach welchen Kameras gesucht und gefunden werden können. In diesen Basiskriterien müssen sich alle Hersteller von Canon über Nikon über Casio gleichermaßen einig sein. Sonst habe ich wieder den Wildwuchs an Informationen und werde bestimmte Dinge einfach wieder nicht finden, weil irgendein Hersteller sich einen Sonderattribut einfallen hat lassen, welches durch keines der Backend Systeme zielgerichtet mit anderen Informationen in Verbindung gesetzt werden kann. Oder man findet nur wiederum Angebote von nur einem Hersteller, weil er gerade einen Attribut ausgewählt hat, welches nur von einem Hersteller unterstützt und akzeptiert wird aber nachdem viele Leute einfach in anderem Kontext suchen. Das würde das Ganze wieder ad absurdum führen, wenn solche Spezialattribute, auf die sich die Hersteller nicht einigen hat können, ins Spiel kommen. Ich habe die Befürchtung oder sehe das Risiko, dass genau so was passieren wird. Dass sich gerade eine Firma, irgendeine Spezialontologie für seine Kameras einfallen lassen wird, die sie zielgerichtet mit Informationen im Web verknüpft, wo andere nicht mitspielen können. Damit ist für den Benutzer die Information wieder gefiltert und er kriegt wieder nicht alles oder wieder zu viel, weil die Basis an Informationen wieder zu breit ist.

Worin liegen die Risiken des Semantic Webs?

Das Risiko für die Benutzer ist dass die Ontologien zu vielfältig werden, zu umfangreich, so dass Informationen nicht mehr zielgerichtet gefiltert werden können. Damit genau das nicht passiert, damit die Ontologien in einem gesunden Rahmen gehalten werden, dann müssten sich alle Hersteller von zum Beispiel Digitalkameras auf allgemeine Charakteristika von Kameras einigen, mit denen man Informationen sinnvoll verknüpfen kann. Das ist extrem aufwendig. Und wenn man alle Märkte der Welt in Betracht nimmt, dann bin ich mir absolut sicher, dass es Märkte und Branchen gibt, die man damit nicht erfassen kann.

Insgesamt ist die Gefahr sehr groß, dass viele Unternehmen so getrieben sein werden, dass sie Ontologien zu Ihrem Vorteil definieren, damit sie den Benutzer eher auf ihre Produkte lenken.

Worin liegen die Chancen des Semantic Webs?

Das Thema Semantic Web müsste rasch aus dem rein akademischen Umfeld in ein pragmatisches praxisnahes Umfeld kommen. Dieser Sprung ist aus meiner Sicht kritisch. Den hat man noch nicht gemacht. Wenn der Sprung nicht bald passiert, ist das Risiko groß, dass es durch andere Trends am Markt, die sich schneller entwickeln, überrumpelt wird.

Ich nehme jetzt die Web 2.0 und das Social Networking her. Die Leute werden schnell merken, dass im Social Network Leute mit gleichen Interessen agieren, die vielleicht Dinge doch noch ein bisschen schneller wahrnehmen als Systeme. Denn

viele dieser Systeme bauen darauf auf, dass man die richtigen Modelle erstmals erkennt und gestaltet. Wenn man über solche Mittel und Wege Informationen schneller findet, dann kann es passieren, dass solche Plattformen schneller Momentum erlangen und das Semantic Web überrumpeln.

Außerdem muss man die Begriffe des Semantic Webs, wie zum Beispiel Ontologien und RDF, für die breite Masse verständlich machen. Das ist der Schritt der notwendig ist, damit Semantic Web in der Industrie und in der Praxis eine stärkere Chance bekommt. Man muss das Ganze auf bodenständige Probleme runterbrechen, die die Leute im Business draußen haben. Man muss denen zeigen wie diese Probleme oder Teile dieser Probleme über das Semantic Web leichter gelöst werden können. Das ist ein Problem mit dem nicht nur Semantic-Web-Leute zu kämpfen haben, mit dem haben auch Plattformen und Hersteller wie Microsoft oder IBM, Yahoo zu kämpfen.

Semantic Web alleine kann nicht alle Probleme der Unternehmen lösen. Damit sich das Semantic Web durchsetzt, sollte man das kreativ mit anderen Dingen verbinden. Wie in vielen anderen Bereichen auch, ist man im Semantic Web leider sehr fokussiert auf sein eigenes Thema (Semantic Web) und versucht alle Probleme damit zu lösen. So wird es aber nicht funktionieren. Man muss versuchen es mit anderen Dingen zu kombinieren. Die Wahrheit liegt überall, in allen Bereichen irgendwo in der Mitte. Bis jetzt habe ich noch kein Extrem gesehen, dass als Extrem wirklich überlebt hat. Sondern es ist immer irgendwo ein Kompromiss mit anderen Dingen geworden.

Dieser Schritt, die Dinge in die Praxis zu bringen und aufzuzeigen, welche alltäglichen Probleme sich damit lösen lassen, ist ein extrem schwieriger. Microsoft hat auch viel Erfahrung damit, denn wir kämpfen bei jeder Weiterentwicklung mit dieser Tatsache. Manchmal sind gute Ideen da, aber diese muss man mit dem lokalen Markt verbinden. Man muss erstmals herausfiltern welche Aspekte relevant für den lokalen Markt sind und welche konkreten Probleme der Leute damit gelöst werden. Auch bei der Microsoft Plattform gibt es nicht eine Technologie, die die Lösung für alle Probleme bietet. Deshalb ist die Plattform auch so vielseitig. Aber genau eben diesen Schritt sehe ich als Chance für Semantic Web um sich besser im Markt zu platzieren und um schlussendlich wie viele andere Web 2.0-Technologien endlich mal abheben zu können.

Lohnt sich der Aufwand in Anbetracht des Nutzens?

Vom ROI betrachtet eher nicht.

Wenn ich priorisiere und das Semantic Web mit anderen Dingen, die sich am Markt entwickeln vergleiche und mir den Return on Investment anschau, dann gibt es Dinge, die wesentlich wichtiger sind, mehr bringen und mehr Potential haben. Es kommt aber darauf an, von welchem Kontext man das betrachten soll.

Bei welchen Applikationen lohnt sich der ROI?

Ein Beispiel, was Firmen ziemlich intensiv einsetzen, sind die themenbezogene Wikis und Blogs. Es funktioniert sehr gut, wenn man diese Dinge richtig aufzieht. Beispiel größerer Softwareentwicklungskonzern, die haben größere Projekte laufen, die in bestimmte Kategorien einzuteilen sind, wo sie bestimmte allgemeine Komponenten wieder verwenden. Diese allgemeinen Komponenten sind themenbezogen in Wikis

dokumentiert. Zum Beispiel, es gibt einen Wiki mit Dokumentationen im Bereich Finance, im Bereich Foreign Exchange und im Bereich Accounting. Wenn ich jetzt Entwickler im Konzern bin und eine Accounting-Komponente für meine Applikation entwickeln muss, dann suche ich nicht ins Blaue, sondern ich schaue erst im Accounting-Wiki nach und durchsuche gezielt dieses Wiki. Dort innerhalb dieses Wikis kann ich nach Standardindizierungsmechanismen zurückgreifen, weil die Informationsvielfalt von vornherein schon eingeschränkt ist auf Accounting. Somit würde ich auch nicht überrumpelt von Informationen sein. Das sind pragmatischere Ansätze, die sich schnell umsetzen lassen und von denen man schnell einen Nutzen kriegt, ohne dass im Vorfeld aufwendigst Ontologien für meine Softwareentwickler zu definieren sind.

Denken Sie, dass sich das Semantic Web in den nächsten 5 Jahren durchsetzen wird?

Im Vergleich zu den anderen Trends ist es ein schwieriges Spiel. Auf der einen Seite ja, es ist irrsinnig viel Potential da wie zum Beispiel die personalisiert dargestellten Inhalte. Aber es müsste im Semantic Web klarer hervorgehoben werden, dass Semantic Web extrem viel beitragen kann. Aber ich glaube es ist alles viel zu akademisch und aufwendig.

Im Moment ist es eher schwierig, dass es sich durchsetzt, weil es im Moment viele weniger akademisch aufgehängte Trends gibt, die am Markt sind, und sich rasant entwickeln. Daher hat das Semantic Web eher ein schwieriges Spiel.

Wie sehen Sie die Entwicklung des Semantic Webs in der Musikbranche?

Die Musikbranche ist für mich zu einfach gestrickt, da finde ich, dass der Einsatz des Semantic Web nicht so viel Sinn macht. Es kommt aber natürlich darauf an. Man könnte jetzt natürlich sagen, ich verknüpfe „Every you and every me“ von Placebo mit Gefühlslagen und mit Stilrichtungen und kann da kann man wiederum Querverbindungen herstellen. Aber das erscheint mir fast zu komplex. Weil das ist so ein emotionaler Aspekt, dass ich glaube, dass sich der Aufwand dafür nicht lohnt.

6.2.3 Interview mit Christian Boser, Leiter des Fachbereiches Semantic Web bei AUSTRIAPRO, Wirtschaftskammer Österreich (vom 10.09.2008, Wien)

Wie ist der Bezug von der Wirtschaftskammer Österreich zum Semantic Web?

AUSTRIAPRO ist ein gemeinnützig tätiger Verein in Kooperation mit der Wirtschaftskammer Österreich, welcher sich mit technologischen und betriebsorganisatorischen Innovationen im E-Business befasst. Der Jungunternehmer Markus Linder (Smart Information Systems) ist auf uns zugekommen und hat uns den Bedarf für Semantic Web zur effizienteren Darstellung von Angebotsdaten aufgezeigt. AUSTRIAPRO gründete hierauf einen eigenen Arbeitskreis, der dieses Themenfeld seit etwa 2 Jahren bearbeitet und der von Herrn Linder, zwischenzeitlich zum Vorstandsmitglied des Vereines avanciert, geleitet wird. Mit dem Projekt ebSemantic leistet AUSTRIAPRO Vorschubarbeit zu Etablierung des Semantic Web basierten E-

Commerce in Österreich. Unsere Rolle im Projekt ist das Projektmanagement und die Koordination innerhalb der Wirtschaftskammer Österreich.

Wir sind einerseits ein Forschungsverein und sind so in der Lage, rasch und flexibel auf Bedarf nach E-Business Standards zu reagieren. Andererseits haben wir als Vorfeldorganisation der Wirtschaftskammer guten Zugang zu den österreichischen Unternehmen, und können unserer Forschungsergebnisse und Pilotanwendung einer breiten wirtschaftsnahen Fachöffentlichkeit kommunizieren. Der Mehrwert von AUSTRIAPRO für die WKÖ ist unser technisches Know How und die Best Practice, die wir mit diesem schaffen. Der Mehrwert der WKÖ für uns hingegen ist der Multiplikator-Effekt und eine Teilfinanzierung unserer Projekte. Es handelt sich somit um eine private-public-partnership im Gebiet des Business Innovationen.

Um was geht es im Projekt ebSemantics?

Im Projekt ebSemantics geht darum, in bestimmten Branchen die Vorzüge des Semantic Webs zu zeigen. In dem aktuellen Demonstrations-Projekt ebSemantics II haben wir aufgrund der großen wirtschaftlichen Bedeutung die Tourismus-Branche ausgewählt. Das Vorhaben zeigt neue Wege auf, wie ich als touristischer Anbieter (Gastronomiebetrieb, Hotel, Eventorganizer) meine Angebote beschreiben kann und diese dann maschinenlesbar und damit automatisch über eine größere Anzahl von Internet-Plattformen hinaus exportieren kann.

Unsere Vorgangsweise und die Semantic-Web-Technologien sind für alle zugänglich. Ich kann einen Baustoffkatalog genauso semantisch gestalten, wie ich Tourismusangebote gestalten kann. Langfristig haben wir auch große Hoffnungen in Sachen Dienstleistungen beschreibungen das Gesundheitswesen. Mit semantischen Konzepten kann ich jede Dienstleistung beschreiben und auch jedes Produkt. Und bei ebSemantic I haben wir mit den Produkten angefangen. Die Herkunft von dem Ganzen war, so Sachen wie Digitalkameras und Mobiltelefone zu beschreiben. Also Unterhaltungs-Elektronik. Wir sind aber zu der Überzeugung gelangt, dass es bei den Dienstleistungen einfach noch mehr Nutzen hat, die man das auch deutlicher zeigen kann. Speziell bei solchen Dienstleistungen, die auch verstreut auf Informationen zurückgreifen. Und beim Tourismus ist das klar gegeben. In Österreich ist der Tourismus ein sehr wichtiger wirtschaftlicher Zweig, in dem die Informationen auf viele Plattformen verstreut sind. Darum haben wir dieses Modell gemacht, aber im Grunde ist es für alle Branchen anwendbar.

Worin liegen die Vorteile für die Benutzer und Unternehmen im E-Commerce-Bereich, falls die Konzepte des Semantic Web im E-Commerce-Bereich Realität werden?

Für Unternehmen ist es einfach der, ich kann besseres Marketing machen; ich kann meine Angebote darstellen und komme auf verschiedenen Plattformen raus. Ich müsste mich einmal selbst als Anbieter und mein Angebot anmelden, wie zum Beispiel beim Firmen A-Z von der Wirtschaftskammer Österreich und komme auf den Plattformen raus. Durch das Semantic Web versuche ich Mehrfacharbeit zu reduzieren und es ist ein effizientes Marketing für meine Angebote.

Und für den Suchenden liegt der Vorteil darin, dass mehr Informationen zur Verfügung stehen. Ich habe eine größere Grundgesamtheit. Wenn ich eine intelligente Suchmaschine habe, dann kriege ich ein besseres für mich spezifisches Ergebnis.

Worin liegen die Nachteile und Gefahren für die Nutzer und Unternehmen, falls die Konzepte des Semantic Web im E-Commerce Bereich Realität werden?

Aus dem Stand heraus gesagt, kann es für mich als Suchenden gibt es zwei mögliche Nachteile geben:

- i) Die Speicherung meines detaillierten Suchprofils enthält unter Umständen viel Information zu meiner Person, wie zum Beispiel, „Christian Boser macht Urlaub in Kärnten, möchte vegetarisch essen, und hat zwei Kinder und einen Hund mit“. Auch wenn die Verfügbarkeit dieser Information nicht wirklich bedrohlich klingt, muss das Thema der Privatsphäre und der Schutz personenbezogener Daten stärker thematisiert werden. Hier können klare Regeln für den Umgang mit Daten aus strukturierten Suchabfragen Hilfe bieten.
- ii) Wenn eine Suchmaschine für mich Daten aus verschiedenen Quellen im WWW für mich aggregiert, habe ich wenig Information über die Datenherkunft. Es könnten auch „fakes“ dabei sein – nicht ernst gemeinte oder unseriöse Angebote, oder solche von Anbietern ohne Gewerbeberechtigung („Pfuscher“ im weiteren Sinne). Hier schafft. z.B. der „Trust Faktor“ von Daten aus dem FAZ Abhilfe. Die Anbieter aus dem FAZ sind Firmen mit aufrechter Gewerbeberechtigung, und somit auch im Falle von Problemen mit dem Angebot (unrichtig, unseriös, unklar, inaktuell...) identifizierbar und „auffindbar“.

Für die Anbieter ist klar: die Vergleichbarkeit des Anbieters wird gesteigert. Wer sich vor dem Wettbewerb fürchtet, muss sich vorm Semantic Web „fürchten“ – dies kann aber sinnvollerweise die Haltung eines Unternehmers sein. Jedenfalls sollte dieser „Nachteil“ vom Vorteil der besseren Auffindbarkeit mehr als aufgewogen werden.

Der Einsatz von Semantic Web hängt auch mit einem gewissen Aufwand zusammen. Sehen Sie das in gewisser Hinsicht als ein Nachteil?

Das wird von uns selbst abhängen. Wenn wir Systeme bauen, die einfach zu bedienen sind, dann nicht. Wir haben uns die interne Vorgabe gemacht, es soll kein Fachwissen nötig sein, um meine Daten in RDF zu übersetzen. Das soll ich als Benutzer gar nicht merken. Ich gebe es ein, in ein Formular, drücke unten drauf und sage, übersetze es jetzt auf RDF. Das muss ich auch nicht selber machen. Das ist auch die Idee dahinter. Damit soll der Benutzer, damit meine ich den Firmenanwender, nicht belästigt werden, dass er RDF lernen muss.

Worin liegen die Chancen des Semantic Webs?

Für mich liegt die Chance darin, dass es möglich wird, Daten aus unterschiedlichen Plattformen, die aus unterschiedlichen Stellen im Internet verstreut vorliegen, zusammen zu führen. Das heißt, ich habe als Anwender mehr Nutzen, ich muss nicht zuerst auf die eine Anwendung, dann auf die andere und andere. Ich habe dann eine

semantische Suchmaschine und erhalte aus verschiedensten Anwendungen Daten aggregiert. Das ist auch das Neue im Vergleich zu bestehenden Suchmaschinen.

Worin liegen die Risiken des Semantic Webs?

Die Risiken liegen einerseits darin, dass genug Daten vorhanden sein müssen. Die Güte des gesamten Systems sowie die Güte der einzelnen Datenquellen müssen vorhanden sein. Dann habe ich ein gewisses Authentifizierungsproblem. Diesen versuchen wir mit dem Firmen A-Z zu begegnen. Ich muss wissen von wem die Daten wirklich stammen. Kann da jeder für jeden anderen auch eintragen? Weil es ist der Unterschied, dass ich sie eben nicht von einer einzelnen Webseite kriege. Ich kriege sie über eine Maschine zusammengesucht. Ich weiß grundsätzlich nicht automatisch woher diese Daten kommen. Daher wird die Verlässlichkeit der Quelle ein Thema sein im Semantic Web und eine mögliche Schwachstelle. Daher die Idee, ich nehme so Trustworthy-Systeme wie das Firmen A-Z. Man weißt, dass dort nur der Unternehmer Daten eintragen darf.

Wie sehen Sie das Problem des Datenschutzes im Semantic Web?

Anbieterseitig stellt sich kein Datenschutzproblem Beim Semantic Web muss ich keine Daten über eine bestimmte anbietende Person sammeln, ich sammle über das Angebot. Zum Beispiel ich suche eine Unterkunft bei der Plattform „Kärnten Urlaub“ und finde eine ganz bestimmte Pension. Das sind Daten der Firma, die die Firma über ihr Marketing der Öffentlichkeit zur Verfügung stellt. Die Anbieterdaten sind ja öffentliches Gut bis zu einem gewissen Grad, es ist nur das was auf der Webseite steht, nichts anderes.

Der Datenschutz ist ein Thema im Bezug auf die Suchmaschinen (siehe oben).

Wie hoch schätzen sie den Aufwand einer Implementierung von Semantic-Web-Technologien in Ihrem Unternehmen?

Der Aufwand sollte für die Unternehmen, die zum Beispiel Angebote maschinenlesbar beschreiben, so gering wie möglich sein. Wenn wir diese Tools fertig gebaut haben, ist der Aufwand auch tendenziell gering. Der Aufwand liegt in der Entwicklung der Tools. Das heißt, es ist technologisch anspruchsvoll für alle Entwickler. Es ist auch eine Herausforderung für die Early Adpoters, die das am Anfang eintragen müssen. Aber langfristig ist es ein Szenario, der möglich ist. Ich trage mein Angebot Im grunde in einen Formular ein und exportiere es dann in RDF. Der Aufwand ist daher langfristig niedrig.

Von was hängt es ab, dass sich das Semantic Web in den nächsten 5 Jahre durchsetzt? Was sind die Kriterien?

Der Aufwand des Einsatzes von Semantic Web lohnt sich unter den Prämissen, dass wir die aufwendige Anfangshürden wie Datenschutz und das Entwicklungsrisiko überwinden. Eine erfolgreiche Entwicklungsarbeit muss geschaffen werden. Des Weiteren muss auch eine gute Menge an ersten Anwendungen vorhanden sein und größere Plattformen, die damit einsteigen. Das heißt, Early Adpoters, was den Plattformen angeht und Early Adopters, was den Firmen angeht, die eintragen. Der

Datenschutz ist insofern wichtig, dass ich auf einfacher Weise aktiv zustimmen kann, dass meine Daten im RDF-Format weiterverarbeitet werden. Der andere wichtige Aspekt betrifft die Suchenden, die wissen wollen, welche Daten von denen verwendet werden und was die Plattformen mit den Daten machen.

6.2.4 Interview mit Thomas Hochhaltinger, technischer Projektleiter bei Smart Information Systems (Wien, 10.09.2008)

Was sind die Eigenschaften und Besonderheiten des digitalen Empfehlungssystems, „Smart Assistant“, der Smart Information Systems anbietet?

E-Commerce-Plattformen bieten heute eine große Zahl von Angeboten und Produkten. Durch die Fülle von Produktinformationen oder Angebotsinformationen fühlt sich der Benutzer oft überfordert. Er kann nicht genau entscheiden, welche Informationen für ihn relevant sind und tut sich somit schwer die Angebote untereinander zu vergleichen. Oft fehlt im dazu aber auch das Know-How über das Produkt. Denn wer weiß schon genau, welche Standards ein 3 Stern Hotel bietet oder wie groß ein 72 Zoll Fernseher wirklich ist?

Mit unseren Empfehlungssystem „Smart Assistant“ bieten wir den Kunden von Webshops eine dynamische Produktberatung an, die auf semantischen Technologien basiert. Der Kunde bekommt eine Empfehlung des digitalen Assistenten, die optimal zu seinen individuellen Wünschen passt. Dadurch erhöht sich die Sicherheit sich für das richtige Produkt entscheiden zu können, was sich schlussendlich auch auf die Zufriedenheit des Konsumenten mit dem Produkt auswirkt..

Der Prozess des digitalen Assistenten erfolgt folgendermaßen: Ohne den Benutzer zu kennen, stellen wir ihn Fragen. Durch diese Fragen leiten wir seine Bedürfnisse ab und diese Bedürfnisse übersetzen wir in technischen Anforderungen. Diese Informationen werden zueinander in Zusammenhang gesetzt. Auch Erfahrungswerte von Verkäufern und Experten des Produkts werden dabei miteinbezogen. Daraus kann man schlussendlich Produkt- bzw. Angebotsempfehlungen aussprechen und diese nach den Wünschen des Konsumenten ranken. Nach jeder Frage werden dem Benutzer die Empfehlungen laut seinen bisher geäußerten Bedürfnissen in einer Nebenleiste angezeigt. Somit kann er jederzeit selbst entscheiden, wann er sich gut genug beraten fühlt und den Beratungsprozess beenden.

Welcher Vorteil hat der „Smart Assistant“ in Vergleich zu herkömmlichen Empfehlungssystemen?

Das Semantic Web ermöglicht uns Eigenschaften der Produkte zu gewichten und Prioritäten zu setzen, um so ein optimales Ranking zu ermöglichen und das bestmögliche Produkt anzuzeigen. Und zwar auch dann, wenn die Ansprüche so hoch sind, dass es kein Produkt gibt, das alle Ansprüche erfüllt. Hier haben wir klare Vorteile im Vergleich zu herkömmlichen „Filter-Tools“, die dann eine Nulllösung ausgeben würden. Gerade im Tourismusbereich kennen wir das alle, dass wir nach einem 3-Stern Hotel in einer bestimmten Region und zu einer bestimmten Zeit suchen und dann kein Ergebnis bekommen. In so einem Fall würde der Smart

Assistant einfach das nächst bessere, verfügbare Hotel empfehlen. Mit unseren digitalen Assistenten kann man Gewichtungen vornehmen und somit zwischen weichen Kriterien, die nicht so ausschlaggebend für die Kaufentscheidung sind und sogenannten Killerkriterien unterscheiden. Wenn ein Attribut ein Killerkriterium ist, wie zum Beispiel, ein behindertengerechtes Hotel, dann wird unser digitaler Assistent auch nur diejenigen Angebote anzeigen, die das haben. Wir können also definieren, welche Kriterien nach unseren Erfahrungen und Statistiken entscheidend sind. Sie können auch selber die Gewichtung vorgeben und sagen wie wichtig ihnen die Erfüllung des Kriteriums ist. Um dieses mit der Technik zu vereinen, sind gewisse Zusammenhänge notwendig. Somit können wir mit Stolz behaupten, dass unser Smart Assistant ständig durch Benutzerverhalten dazu lernt.

Einen zusätzlicher Vorteil ist, dass Angebote und komplexe Zusammenhänge, bei denen zwar der Mensch einen Konnex herstellen könnte, aber diese für Computer nicht erkennbar sind, durch Semantic-Web plötzlich auch für die Maschine verständlich werden. Zum Beispiel erkennt der Smart Assistant, dass ein Hotelangebot zu einer bestimmten Zeit verfügbar ist und kann Konnex zu einem Event herstellen, das ebenfalls zu diesem Zeitpunkt in der Region angeboten wird. Mit herkömmlichen Suchtechnologien würde dies auf automatisierter Ebene kaum realisierbar sein.

Wie ist der Prozessablauf um Semantic Web in Unternehmen einzusetzen?

Im wesentlichen geht es darum, dass man seine Inhalte, die schon in verschiedenen Datenbanken dargestellt sind, in gewisse Relationen setzt und diese somit semantisch beschreibt.

Erstens muss man entscheiden ob man das Semantic Web intern einsetzt, oder man seine Daten und Ontologien für das ganze Semantic Web verfügbar macht. Im letzten Bereich ist es wichtig, dass man eine große Kompatibilität mit bestehenden Systemen hat und im Optimalfall sogar auf bestehende Standards setzt. Im ersten Bereich, kann man semantische Technologien, was derzeit üblicher ist, auf sein eigenes System anwenden. Ein Paradebeispiel dafür ist der Einsatz von Semantic Web im Wissensmanagement.

Der zweite Schritt ist die Entwicklung von Ontologien. Hier ist es empfehlenswert, dass man erstmals schaut, ob es bereits Ontologien gibt, die sich mit der eigenen Branche und Themengebiete beschäftigen. Die bestehenden Ontologien muss man dann mit den internen Daten in Relation setzen. Es gibt bereits einige Initiativen, die Ontologien entwickeln und mit der Menschheit teilen. Ein Beispiel ist myOntology (www.myontology.org). Die zweite Alternative ist, dass ein Unternehmen eine Ontologie für sich selber entwickelt. Die Unternehmen müssen ihre Ontologie ja nicht zwangsläufig für das gesamte Semantic Web erstellen. Sie können das auch intern machen und ihren Nutzen daraus ziehen, in dem sie mit intelligenten Suchtechnologien intern einsetzen.

Der dritte Schritt für die Unternehmen ist, die Daten des Unternehmens, die mittels Ontologien strukturiert worden sind, schlussendlich auf deren Basis mit RDF-Annotationen zu beschreiben. Die meisten Unternehmen haben schon klassische Datenbanken oder MySQL-Datenbanken oder haben das Ganze über XML beschrieben. Um diese Daten im Semantic Web zu nutzen, müssen sie diese auf Basis der Ontologien in RDF beschrieben werden. Der Aufwand der Daten-Beschreibung in

RDF hängt hauptsächlich nicht von der Anzahl an Daten ab, sondern von der Komplexität der Daten, wie breit die Datenbank ist. Wenn die Datenbank viele Felder und Relationen hat, dann ist es schwieriger, die Daten in RDF zu beschreiben. Die große Herausforderung ist, dass so viele Unternehmen wie möglich, ihre Daten in RDF umwandeln sollen, damit semantische Daten generiert werden können und sich dieses Konzept durchsetzt.

Macht es mehr Sinn selber Ontologien zu erstellen oder bestehende Ontologien einzusetzen?

Das kommt darauf an, in welchen Bereich sie eingesetzt werden sollen und was das Ziel des ganzen ist. Ist der Ziel der Datenaustausch bzw. sein Angebot möglichst intelligent und zielgruppenorientiert einem breiten Publikum zur Verfügung zu stellen? Oder stehen interne Verwendungsszenarien von Semantic-Web-Technologien im Vordergrund?

Unser Ziel ist eine möglichst große Kompatibilität zu schaffen und unsere Ontologien auf kollaborative wikibasierte Ontologie-Entwicklungsplattformen wie myOntology zu veröffentlichen, damit die Leute die Know How haben, die Ontologien verbessern und für Ihre Ansprüche adaptieren können. Wir haben gemerkt, dass für die Entwicklung von einer Ontologie sehr viel dazugehört. Denn um Ontologien erfolgreich zu entwickeln, braucht es sehr viel Wissen über das Fach, die Zielgruppe, die Technologie und sprachliches Wissen. Es ist ein komplexer Prozess, nicht nur technisch, sondern auch im fachlichen und sprachlichen Bereich. Deswegen versuchen wir bei unseren Ontologien möglichst im Rahmen von Expertenforen, Touristiker mit ins Boot zu holen, die fachliches Know-How haben. Wir wollen auch deren Wording verwenden um Ontologien zu erstellen. Es ist nämlich sehr wichtig, das alles in deren Sprache abläuft und nicht in technischer Sprache. Genauso von Bedeutung ist aber, dass alles technisch einwandfrei ist. Die große Herausforderung ist alles optimal zu verknüpfen.

Bei der kollaborativen Ontologie-Entwicklungsplattform myOntology hat man den Vorteil, dass das Ganze in einem kollaborativen Prozess voranschreitet. Da kommt viel Input von draußen, an den man sonst als Unternehmen gar nicht denken würde. Da stehen auch gewisse Leute dahinter, die sagen, „so kann man das eigentlich nicht darstellen“. Wenn myOntology weit gewachsen ist, dann wird es sicherlich so sein, dass für viele Themen Ontologien existieren. Umso mehr Ontologien auf myOntology sind, desto größer wird die Community sein.

Wir schätzen das Potential für Initiativen wie myOntology als sehr hoch ein, weil man sich eine Ontologie oder Teile von bestehen Ontologien entnehmen kann und das Rad nicht neu erfinden muss. Man nimmt sich dann einen großen Teil der Ontologie heraus, lädt sich das herunter und macht auf Basis dessen eine eigene Ontologie. Diese Vorgangsweise eignet sich natürlich nur, dann, wenn man mit den Zusammenhängen nicht unternehmensinterne Prozesse beschreiben will. Dafür wird MyOntology nicht das geeignete Tool sein, um das Semantic Web in Unternehmen einzusetzen, sondern eher als Beispiel für schöne Modellierungen dienen.

Eine weitere Initiative zur Benutzung von bestehenden Ontologien ist die „Good relations“-Ontologie vom renommierten Semantic-Web Experten Univ.-Prof. Dr. Martin Hepp. Hier geht es darum, eine Grund-Ontologie zur Beschreibung von E-

Commerce-Angebote zu schaffen. Verschiedene Teile von dieser Grund-Ontologie können genutzt werden, um neue Ontologien zu schaffen. Zum Beispiel baut unsere Gastronomie-Ontologie auch auf Good Relations auf, um Angebote von Menüs zu beschreiben. Wir haben in diesem Fall einige Teile der Ontologie genommen und sie in unserer Ontologie hineinpasst. Verschiedene Spezifikationen wie zum Beispiel Preise und Verfügbarkeit, die ein klassisches Angebot auszeichnen, mussten wir damit nicht neu definieren und sind somit automatisch mit vielen anderen Strukturen kompatibel.

Welche Daten über den Benutzer sind wichtig um zielgruppenorientierte Angebote mit dem „Smart Assistant“ zu erstellen?

Wir schauen, wie sich die Benutzer innerhalb von unseren Empfehlungssystemen verhalten. Wir haben ein eigenes intelligentes Analysesystem entwickelt, wo wir sehen, welche Fragen die Bedürfnisse der Kunden gut abdecken und welche Fragen wiederum den Konsumenten nicht weiter bringen. Damit versuchen wir den Konsumentenverhalten zu erlernen. Wichtige Indikatoren sind da z.B. unter anderem, wie häufig und in welchem Konnex eine Frage beantwortet wird Ein weiterer wichtiger Aspekt, den man wissen soll, ist, wann der Punkt erreicht wird, an dem sich der Kunde gut beraten fühlt In diesem Moment klickt er auf das Endergebnis. Hier ist es wiederum sehr interessant welche Produkte in einem bestimmten Bereich die beliebtesten sind. All diese Informationen sind nicht nur für unsere Beratungsoptimierung, sondern auch vor allem für unsere Kunden sehr vorteilhaft.

Die Information in Web 2.0-Applikationen wie zum Beispiel Blogs wird diesbezüglich in Zukunft auch eine wichtige Rolle spielen. Wir planen, dass wir Information in Blogs und in anderen Web 2.0-Applikationen, für die Verbesserung von unserem Empfehlungssystem berücksichtigen. Wir wissen, dass die Information in Blogs über verschiedene Anwendungen und Themen ein sehr wichtiger Entscheidungspunkt ist. Diese haben einen sehr großen Einfluss auf die Kaufentscheidung. Wenn eine Person eine vernichtende Kritik über ein Produkt liest, da kann unser Empfehlungssystem noch so sehr sagen, dass dieses das optimale Produkt für diese Person ist, er wird es wahrscheinlich nicht kaufen. Weil es so einen großen Einfluss hat, wollen wir vor allem im Tourismusbereich ein Augenmerk über die Hotel- und Gastronomie-Bewertungen in Blogs und weitere Web 2.0 Applikationen legen.

6.2.5 Interview mit Markus Linder, Smart Information Systems (vom 10.09.2008, Wien)

Was sind die Voraussetzungen, um Semantic Web in Unternehmen einzusetzen?

Zuerst muss man unterscheiden ob man es im E-Commerce oder im Wissensmanagement einsetzen will. Im letzteren Fall gibt es fertige Tools, die auf Basis von Semantic-Web-Technologien es dem Unternehmen erleichtern, Wissen zu verarbeiten und zu verwalten.

In Bezug auf den Semantic-Web-basierten E-Commerce muss ein Unternehmen Tools nutzen, um Produkte und Dienstleistungen strukturiert beschreiben zu können. Das reicht von einer einfachen Web-Anwendung, die man nutzen kann, bis zu Anwendungen, die man installiert.

Welche Lösungen bietet das Semantic Web im E-Commerce-Bereich den Unternehmen und Benutzern?

Im Bereich des Semantic-Web-basierten E-Commerce bietet das Semantic Web den Benutzern und Unternehmen, Angebote einfach und effizient zu platzieren und intelligent auffindbar zu machen.

Den Unternehmen ermöglicht es einerseits ihre Produkte für die Zielgruppe optimal auffindbar zu machen. Die Benutzer andererseits profitieren von den Suchergebnissen, die sie bekommen, die optimal zu diesen Menschen passen.

Worin liegen die Vorteile der semantischen Empfehlungssysteme für die Benutzer und für die Unternehmen?

Semantische Empfehlungssysteme eingesetzt in Online-Shops oder auf Hersteller-Portalen bieten den Benutzern einen großen Vorteil. Erstens bekommen sie einen raschen Überblick über die Produktpalette. Zweitens können sie sehr schnell und bedürfnisorientiert zu den Produkten hingeführt werden. Drittens passt die Produktempfehlung optimal zu ihren Bedürfnissen. Somit spart der Benutzer ein Zeit und hat weniger Aufwand bei der Suche.

Auf der anderen Seite kann durch unsere Empfehlungssysteme, wie zum Beispiel Esolda.at, eine höhere Entscheidungssicherheit herbeigeführt werden, weil Konsumenten, die sich gut beraten fühlen, sich eher trauen die Kaufentscheidung zu treffen. Ein weiterer großer Vorteil ist, dass die Empfehlungssysteme nicht nur einzelne Produktkategorien wie Fernsehen, Digitalkameras oder Hotels anbieten, sondern auch eine Produktbündelberatung. Mithilfe der Semantic-Web-Technologien ist es wesentlich einfacher, die Produktberatung intelligent über verschiedene Produktbereiche verknüpft durchzuführen.

Welche möglichen Applikationen zur Personalisierung der Inhalte im Internet fallen Ihnen ein?

Personalisierte Inhalte sind gegeben, wenn ein Portal gewisse Informationen über einen Benutzer hat. Eine Information ist zum Beispiel die Postleitzahl des Wohnortes des Benutzers. Diese Information kann in verschiedenster Weise von unterschiedlichen Portalen genutzt werden.

Wenn sich ein Benutzer im Portal www.ichkoche.at einloggt und eingibt, dass er aus Graz kommt, dann könnte das Portal zielgruppenorientierte Veranstaltungen wie laufende gastronomische Events oder Kochmessen aus dem Semantic Web ziehen und für den Benutzer individualisiert darstellen. Das gleiche gilt für das Portal www.bikemap.net. Wenn den Benutzer Hotels einer bestimmten Kategorie interessieren oder wenn er Vegetarier isst, kann ihm das System die Hotels der jeweiligen Kategorie und vegetarische Restaurants auf der Strecke anzeigen, die offen haben, wenn er die Radtour durchführt. Da gibt es viele Anwendungsmöglichkeiten. Der

große Vorteil, den Semantic-Web-Technologien bieten ist, dass die Anbieter von Semantic-Web-Inhalten und die Benutzer sich in Bezug auf den Nutzen nicht abstimmen müssen.

Nach dem Hersteller ihr Produkt möglichst umfassend und strukturiert auf der Basis von Semantic-Web-basiertem E-Commerce beschrieben haben, besteht die Möglichkeit, dass damit auch kleine Portale ihre Inhalte, die für ihre individuellen Benutzer oder Zielgruppen interessant sind, aufgreifen bzw. miteinbeziehen. Und das geschieht im Semantic Web, ohne dass dazwischen bilateral irgendwelche Daten speziell formatiert werden müssen.

Wie sehen sie die Verknüpfung der Werbung mit dem Semantic Web?

Durch die immer größere Individualisierbarkeit von Inhalten und Web-Auftritten ist es natürlich auch möglich, Werbung immer gezielter zu platzieren. Durch das Wachstum strukturierter Informationen über die Benutzer, Produkte und Dienstleistungen, sehe ich die Möglichkeit, dass man diese Information nutzt, um sie intelligenter mit Werbeeinhalten zu verknüpfen. Daher sehe ich es als realistisch an, dass neue Methoden in der Zukunft verwendet werden, um zielgruppenorientierte Werbung platzieren zu können.

Ein Anwendungsbeispiel im Musikbereich wäre Folgendes: Die Plattenlabels könnten Kurzversionen von Songs im Web semantisch annotiert zur Verfügung stellen. Diese kurzen Songs könnten automatisch in einer Vielzahl von Community-Portalen eingebunden werden. Man könnte das in Kombination mit Werbung, Affiliate-Programmen und Social Networks nutzen. Allgemein könnte man in der Musikindustrie neue Applikationen schaffen, die Personen Musik zielgruppenorientiert anbietet. Diesbezüglich sind Collaborative-Filtering-Methoden und Ähnliches sinnvoll.

Lohnt sich der Aufwand in Anbetracht des Nutzens? (Kreuzen Sie bitte die richtige Antwort an)

Ja. Im Bereich des Semantic-Web-basierten E-Commerce ist es ähnlich wie bei der Einführung des Faxes oder des World Wide Web. Es sind Netzwerkexternalitäten da. Es lohnt sich auf jeden Fall, allerdings gibt es am Anfang eine gewisse Phase, in der man nicht sofort den Vorteil im vollen Ausmaß sehen wird. Dafür hat man sich sehr früh mit der neuen Technologie vertraut gemacht, und man kann sehr schnell lernen, wie man damit umgehen muss, um von ihr zu profitieren. Unternehmen, die sich früh umstellen, werden davon dann sicher den größten Vorteil haben.

Von welchen Faktoren hängt es ab ob sich das Semantic Web durchsetzt?

In Bezug auf die Technologie muss eine Infrastruktur geschaffen werden. Wir arbeiten sehr aktiv daran, Ontologien zu entwickeln. Unser Weg ist eine kollaborative Entwicklungsplattform, myOntology. Denn letztendlich braucht man Werkzeuge, Mittel und Wege, die es erleichtern, gemeinsam Ontologien zu entwickeln. Natürlich braucht man auch Werkzeuge, um Produkte, Dienstleistungen und Ontologien zu beschreiben und konkrete Instanzen zu bilden. Content-Management-Systeme sind nötig, um diesen Prozess zu unterstützen und den RDF-Export zu tätigen. Des Weiteren brauchen bestehende Portale Unternehmensstandard-Software-

Anwendungen, um ihre Inhalte in RDF-Form zu exportieren. Das sind lauter Meilensteine, die die Integration von RDF-Daten erleichtern.

Denken Sie, dass sich das Semantic Web in den nächsten 5 Jahren durchsetzen wird?

In 5 Jahren ja. Das World Wide Web hat sich auch schrittweise durchgesetzt. Am Anfang war es auch schwer. Bei allen neuen Technologien ist es so, dass die Sachen schrittweise kommen und dann erst boomen. Genauso wird es hier sein. Erst wird es einzelne Teilbereiche geben, an denen man sieht, dass es funktioniert. Mit diesem Ziel versuchen wir, unser Projekt im Tourismusbereich voranzutreiben, damit andere Branchen sich ein Beispiel nehmen können und das Semantic Web auch in anderen Bereichen Fuß fassen kann.