

Bachelorarbeit

Konvergenzen von Web 2.0 und Semantic Web am Beispiel einer Tourismusplattform

ausgeführt von
Johannes Nagl
Dr.-Vogl-Gasse 37, 3400 Klosterneuburg

Begutachter/in: DI Helmut Gollner

Klosterneuburg, 06. August 2007

Ausgeführt an der Fachhochschule Technikum Wien
Bachelor-Studiengang Wirtschaftsinformatik



Kurzfassung

Die vorliegende Bachelorarbeit beschreibt die theoretischen Grundlagen von Web 2.0 und Semantic Web. Diese Technologien stellen konkrete Verbesserungen zum bisherigen World Wide Web dar. Anhand der Programmierung einer Konvergenzapplikation, die eine Verbindung der beiden Technologien darstellt, wird die Methodik der Erstellung eines semantischen Netzes erklärt. Dieses Netz wird automatisiert um Inhalte von Drittanbietern erweitert. Der Vergleich von MashUp Erstellungsdiensten wie Yahoo Pipes und Microsoft Popfly gibt Auskunft über den Einsatz von vorgefertigten Programmen zur Datenaggregation. Ein abschließender Überblick über ausgewählte Tourismusplattformen zeigt die Stärken und Schwächen sowie die wirtschaftliche Bedeutung von Konvergenzapplikationen im Gegensatz zu konventionellen Web-Angeboten.

Abstract

This Bachelor thesis describes the theoretical principles of the Semantic Web and Web 2.0. These technologies are definite improvements to World Wide Web in his current form. A convergence application, which is the combination of both technologies, shows the methodology of creating a semantic net. The net is automatically extended with third party data. The comparison of so called Mashup Creator Applications like Yahoo Pipes or Microsoft Popfly shows the benefits of programs for standardized data aggregation. Finally, an analysis as well as an economic inspection of selected tourist platforms shows the strengths and weakness of convergence applications in contrast to conventional web offerings.

Keywords

Semantic Web, Web 2.0, Konvergenzapplikation, Ontologie, MashUp, Yahoo Pipes, Microsoft Popfly, Topic Map, Data Web, Social Semantic Web, RDFa, MicroFormats, Tourismus, Webapplikation, Datenimport

Eidesstattliche Erklärung

„Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit selbständig angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder in gleicher noch in ähnlicher Form einer anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.“

Klosterneuburg, 06. August 2007

Danksagung

Die vorliegende Bachelorarbeit stellt den Abschluss des Bachelor-Studiums für den Autor dar. Daher möchte er die Chance wahrnehmen und allen Vortragenden für die vermittelten Inhalte während des Studiums danken. Vor allem waren es die vielen Diskussionen und spannenden Vorträge über neue Perspektiven im Web (eCommerce, Wissensmanagement, Kollaboration), die Lust auf *mehr* machten. Diese Arbeit stellt den Startschuss für *mehr* praktische Erfahrungen dar und vertieft die bereits im Unterricht aufgegriffen theoretischen Grundlagen.

Danke für eine solide Basis.

1. Inhaltsverzeichnis

1. Inhaltsverzeichnis	5
2. Problem- und Aufgabenstellung	6
3. Aufbau der Arbeit	8
4. Theoretische Grundlagen	9
4.1. Semantic Web	9
4.1.1. Das Top-down Design.....	10
4.1.2. Aufbau des Semantic Web.....	10
4.2. Web 2.0	11
4.2.1. Das Bottom-up Paradigma.....	12
4.2.2. Kombination von Datenbeständen durch MashUps.....	13
4.3. Theoretische Konvergenzen - Social Semantic Web.....	14
4.3.1. Potentielle Anwendungsbeispiele.....	15
5. Konzeptionierung einer Konvergenz-Applikation	16
5.1. Die Erstellung eines semantischen Netzes	17
5.1.1. Methodik der Erstellung	17
5.1.2. Beispiel-Netz für den Städte Tourismus	18
5.2. Automatisierter Datenimport mittels Mash Ups	19
5.2.1. Eigenständige Programmierung.....	20
5.2.2. Yahoo Pipes	20
5.2.3. Microsoft Popfly	21
5.2.4. Vergleich der Methodiken	22
5.3. Bereitstellung semantischer Daten.....	23
5.3.1. Microformats	24
5.3.2. RDFa	25
6. Kurzer Vergleich von Tourismus-Plattformen	27
6.1. Wahl geeigneter Kriterien.....	27
6.2. Anbieteranalyse	28
6.3. Ergebnis des Vergleichs	29
7. Zusammenfassung	30
8. Diskussion / Blick in die Zukunft	31
8.1. Wirtschaftliche Betrachtung	31
9. Glossar	34
10. Abbildungsverzeichnis	36
11. Abkürzungsverzeichnis	37
12. Literaturverzeichnis	38
13. Anhang	40
13.1. Struktur des semantischen Netzes (Auszug):.....	40

Die Bezeichnungen sollen immer geschlechtsneutral verstanden werden.

2. Problem- und Aufgabenstellung

Das Web verändert sich rasch. Ein sozialeres, intelligenteres Netz wird das bestehende bald ablösen. Hierfür werden Web 2.0 und Semantic Web (SW) sorgen. Abbildung 1 zeigt, dass der Begriff des Web 2.0 stetig an Bekanntheit zulegt, während dem Semantic Web kaum mediale Aufmerksamkeit geschenkt wird. Dennoch werden beide Technologien das Web nachhaltig verändern. Seit Entstehen der beiden Terminologien wurden die Bemühungen sehr gekapselt betrachtet.

„Ein neues Netzverständnis ist im Entstehen: Der Anwender steht immer mehr im Mittelpunkt; neue Technologien und Dienste emanzipieren ihn, seine Bedürfnisse in individuellerer Ausprägung zu befriedigen. Alle Bereiche der Entwicklung eines Produkts für und im Web werden von dieser Tendenz maßgeblich beeinflusst. Aus unternehmerischer Perspektive gilt es, neue Herausforderungen, aber auch Potenziale zu identifizieren und auf für Anwender und Unternehmen positive Weise zu nutzen.“ (Soergel & Munz, 2007, S. 1)

Die neuen Herausforderungen, die sich für Web-Entwickler stellen, sowie die vielen Potenziale, die sich durch den Einsatz von neuen Techniken bieten, werden bereits jetzt erkannt. Entwickler springen auf den Zug des Web 2.0 oder Semantic Web auf. Doch nur wenige erkennen, dass in Zukunft Anwender am meisten profitieren können, wenn beide Züge kombiniert werden. Das daraus entstehende Social Semantic Web, bzw. Web 3.0 steckt noch in den Kinderschuhen (vgl. Abbildung 1). Während das Web 3.0 ähnlich populär wie das Semantic Web ist, gibt es zu dem Begriff des Social Semantic Web zum Zeitpunkt der Veröffentlichung dieser Arbeit noch zu wenige Daten, um einen Trend abbilden zu können.

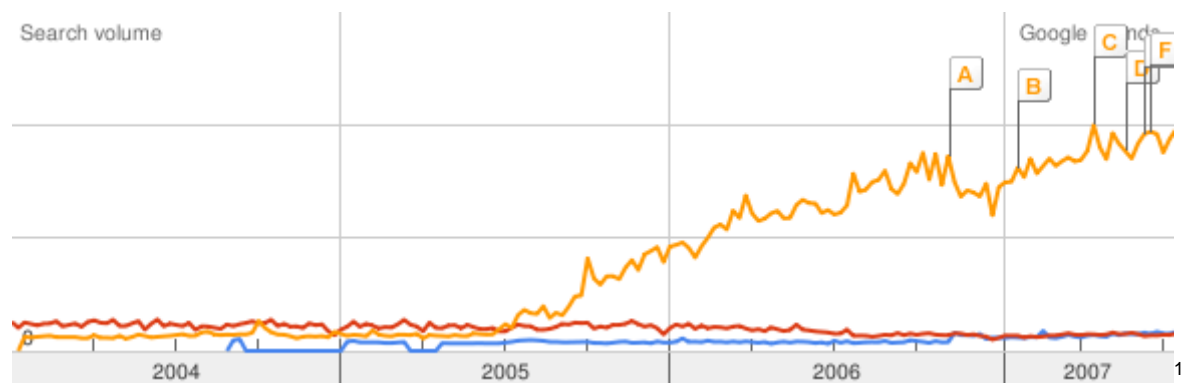


Abbildung 1: Popularität der Begriffe Web 2.0, Web 3.0, SW, Quelle: Google Trends

Die Kombination von Semantik und Kollektiver Intelligenz in einen Guss stellt die Hauptaufgabe der Arbeit dar. Um Aufschlüsse über zukünftige Entwicklungen zu liefern, beantwortet die vorliegende Arbeit daher folgende Fragen:

- Welche konkreten Stärken und Schwächen bietet eine Konvergenz-Applikationen im Tourismus-Sektor zu herkömmlichen Angeboten?
- Anhand welcher Kriterien kann ein qualitativer Vergleich der beiden Anbietergruppen Ergebnisse liefern?

¹ <http://www.google.de/trends?q=Web+3.0%2C+Semantic+Web%2C+Web+2.0&ctab=0&geo=all&date=all&sort=0> [26.07.2007]

- Welche Konzepte des Datenimports von bestehenden Quellsystemen in eine Applikation sind denkbar?
- Sind Konvergenzapplikation wirtschaftlich einsetzbar?

Web 2.0 und Semantic Web werden nicht gekapselt voneinander betrachtet, sondern werden miteinander verbunden. Eine Analyse der daraus resultierenden „*Konvergenzapplikationen*“ gibt Aufschluss über Stärken und Schwächen im Vergleich zu einer herkömmlichen Standardapplikation und stellt fest, ob es für Benutzer von Webapplikationen einen messbaren Mehrwert gibt, wenn Ansätze beider Technologien eingesetzt werden.

Da der Autor bereits über beide Themen im Rahmen seines Bachelor-Studiums geschrieben hat, ist die Kombination dieser Technologien in eine eigenständige Anwendung eine gelungene Abrundung des Bachelor-Studiums.

3. Aufbau der Arbeit

Die vorliegende Arbeit richtet sich an Technik-interessierte Personen, die bereits erste Erfahrungen mit der Programmierung im Web gemacht haben. Dabei stellt Kapitel 4 einen theoretischen Überblick über die aktuellen Trends im Internet dar. Anhand der Beschreibungen von Semantic Web und Web 2.0 wird der konträre Aufbau der beiden Trends analysiert und theoretische Konvergenzen aufgezeigt.

Anhand des Beispiels einer Tourismusplattform in Kapitel 5 liefert die Arbeit Vorschläge zur sinnvollen Kombination der technologischen Ansätze. Eine selbst modellierte Topic Map für den Themenraum des Städtetourismus wird automatisch um multimediale Web 2.0 Inhalte ergänzt. Die Arbeit liefert einen Überblick über ausgewählte Methoden der Erstellung von MashUp-Diensten während ein Vergleich zeigt, ob ein Best-Practice-Vorgehen möglich ist.

Der kurze Vergleich der erstellten Applikation zu herkömmlichen Anwendungen zeigt im Kapitel 6 die Stärken und Schwächen von Konvergenzapplikationen. Dabei wurde die Sichtweise des Benutzers gewählt, da bei Web-Portalen nicht die Technik ausschlaggebend sein soll, sondern der Nutzen für den Kunden.

Anhand der Erkenntnisse aus den vorherigen Kapiteln liefert Kapitel 7 einen Ausblick in die Zukunft und behandelt die Thematik der Wirtschaftlichkeit von Konvergenzapplikationen.

4. Theoretische Grundlagen

Das Kapitel über die theoretischen Grundlagen liefert einen Überblick über die zwei größten Trends im World Wide Web (WWW). Wie bereits in der Problemstellung erklärt werden die Themenkomplexe nicht gekapselt voneinander betrachtet sondern versucht zu verbinden. Obwohl die Funktionsweisen von Semantic Web und Web 2.0 auf den ersten Blick sehr verschieden wirken, können diese problemlos verbunden werden. Beispiele für Konvergenzen untermauern die theoretischen Inhalte mit ersten praktischen Beispielen, die für den eher praktischen Teil der Arbeit als Ausgangsbasis dienen sollen.

Das World Wide Web wurde bereits vor vielen Jahren in seinen Grundzügen entwickelt. Obwohl die Übertragungstechnik im Hintergrund annähernd dieselbe blieb, hat sich dennoch viel an der Art der angebotenen Inhalte verändert. Die wenigen, einfach gehaltenen Artikel von Universitätsprofessoren und Hobby-Technikern wurden bald von aufwändigen Marketing-Auftritten von Firmen in den Schatten gestellt. Internetverbindungen wurden schneller, die Datenmengen immer größer (vgl. ALBY, 2007, S. 3). Schon bald war das Internet nicht mehr nur starr text-basiert, sondern Bilder, Videos und Musik läuteten große Veränderungen ein. Das Netz wurde vielfältiger. Gleichzeitig wurde der Ruf nach mehr Partizipation im Netz laut. Benutzer wollten nicht nur Informationen finden, sondern auch eine Möglichkeit bekommen, wie sie selbst, ohne große technische Hürden bewältigen zu müssen, einfach Inhalte publizieren können. Dabei bleibt dennoch ein zentrales Problem bestehen. Inhalte im WWW sind für das menschliche Auge konzipiert – Maschinen können die angebotenen Inhalte nur anzeigen, nicht jedoch deren Sinn erkennen. Computer sind nicht in der Lage den Kontext von Wörtern und deren Bedeutung ohne zusätzliche Meta-Informationen zu verstehen. Während Menschen bei dem Satz „*Ich gehe zur Bank und hebe 20 EUR von meinem Konto ab.*“ aus dem Zusammenhang schließen können, dass es sich bei der Bank um ein Geldinstitut handelt, können Maschinen keine Aussage darüber treffen, ob es sich hierbei um eine Parkbank oder ein Geldinstitut handelt.

Um den Anforderungen an ein neues Web gerecht zu werden, bildeten sich zwei grundverschiedene Ansätze, die sowohl das Problem der Partizipation als auch das der maschinenlesbaren Informationen zu lösen versuchen. Die Modewörter „Web 2.0“ und „Semantic Web“ werden das World Wide Web in seiner bisherigen Form sehr rasch ablösen bzw. erweitern. Beide Trends bilden kein völlig neues Internet, sondern stellen Erweiterungen der bisherigen Technologien dar. Grundlegende Veränderungen wie „User generated content“ oder „intelligente Software Agenten“, die aktuelle Browser ablösen werden, erweitern das Netz der Netze nicht nur, sondern revolutionieren das Medium, das zukünftig eine noch größere Rolle für das tägliche Leben spielen wird.

Die wichtigsten Neuerungen dieser zwei Ansätze sollen in diesem Abschnitt näher erklärt werden, um eine theoretische Basis für die im Kapitel 5 erstellte Musterapplikation schaffen zu können.

4.1. Semantic Web

Seit den Anfängen des World Wide Web gibt es ein signifikantes Problem. HTML regelt nur die Darstellung von Webseiten, allerdings fehlt eine Möglichkeit die Aussage der darin enthaltenen Daten zu hinterlegen. Das heißt, dass das WWW für Menschen lesbar ist, für Maschinen jedoch nicht interpretierbar. Das Semantic Web stellt einen Ansatz dar, um dieses Problem zu lösen. Es wird versucht durch Hinterlegung von Metadaten Inhalte maschinenlesbar zu machen.

“The goal of the Semantic Web initiative is to create a universal medium for the exchange of data where data can be shared and processed by automated tools as well as by people. [...] We are talking about data here, not human documents.”
(Updegrave, 2005)

Ein Artikel von Tim Berners Lee, dem Erfinder des heutigen World Wide Web und Entwickler der Sprache HTML in *Scientific American* mit dem Titel „The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities“ war der Auslöser für die Veröffentlichung seiner Ideen an die breite Masse. (vgl. Berners-Lee et al, 2001, S. 34)

4.1.1. Das Top-down Design

Die Bemühungen des W3C (World Wide Web Consortium), das federführend an der Entwicklung des SW beteiligt ist, sowie der im nächsten Kapitel beschriebene technische Aufbau des SW verdeutlichen das klassische Top-down Design. Die Entwicklung wird von einer Instanz geleitet und forciert, während einige wenige Experten standardisierte Sprachen verabschieden und somit den so genannten Semantic Web Stack (siehe Glossar) formen. Diese wenig Massen-wirksame Vorgehensweise stellt den Grund dar, warum die Bekanntheit des SW im Vergleich zum Web 2.0, das von der Community selbst gepusht wird, sehr gering ist. Die technischen Spezifikationen müssen untereinander abgestimmt werden und in den hierarchischen Aufbau integriert werden. In der pyramidenartigen Aufbauweise finden sich Beispielapplikationen erst auf sehr tiefer Stufe. Die, aufgrund des komplexen technischen Aufbaus, sehr große Einlernphase verhindert ein *Rapid Development* von SW-Applikationen. Im Vergleich dazu kann im Web 2.0 jeder partizipieren und seinen Teil zur Popularität beitragen (konträrer Ansatz). Das vernetzte SW selbst ist momentan nur ein geistiges Konstrukt und ist primär für Menschen „unsichtbar“.

„Das W3C ist zwar bemüht die Entwicklungen zu lenken und Standards zu schaffen. Die Bemühungen scheitern aber meist am gegensätzlichen Interesse der Wirtschaft, die keine einheitliche Standardisierung zulässt. Einflussreiche Firmen entwickeln ihre eigenen proprietären de-facto Standards. Darüber hinaus stellt das Semantische Web eins der ersten globalen IT-Projekte überhaupt dar. Es bleibt also ungeklärt, ob die entstehenden Standards, Sprachen und deren Vokabulare in Zukunft ein interoperables „großes Ganzes“ bilden werden können.“ (Nagl, 2005, S. 12)

Die gegensätzlichen Interessen von W3C und den Industrievertretern, die rasch eigene Standards durchbringen wollen, führt zu einer Verlangsamung des Prozesses. Dennoch ist der Weg des W3C nachvollziehbar und wichtig – die unabhängige Entwicklung des SW durch Branchenexperten ist ein wichtiger Hoffnungsschimmer dafür, dass in naher Zukunft die Interoperabilität der Daten gegeben sein wird.

4.1.2. Aufbau des Semantic Web

Der Top-Down Ansatz des Semantic Web spiegelt sich in dessen Aufbau wieder. Es handelt sich um einen sehr technischen Ansatz, der viele Implementierungsansätze erlaubt bzw. notwendig macht. Das SW ist im Vergleich zum Web 2.0 ein sehr komplexer Ansatz und zum Zeitpunkt der Veröffentlichung dieser Arbeit sind viele Bereiche des Semantic Web Stacks noch als *Working Drafts* (=unfertig gestellte Entwürfe) des W3C anzusehen. Der Aufbau änderte sich in den letzten Jahren grundlegend und wird sich wohl auch noch bis zur tatsächlichen Fertigstellung weiter verändern. Zum Aufbau des SW befragt, meinte Tim Berners-Lee im Jahr 2007: *„The Semantic Web isn't inherently complex. The Semantic Web*

language, at its heart, is very, very simple. It's just about the relationships between things.
(King, 2007)

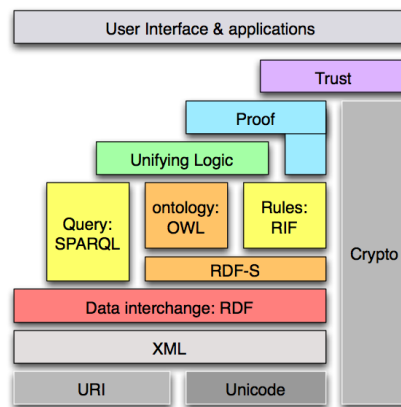


Abbildung 2: Semantic Web Stack (2005), Quelle: W3C

Abbildung 2 zeigt den aktuell gültigen Semantic Web Stack, der die stufenweise Integration der neuen Technologien in das World Wide Web spezifiziert. Basis des SW stellen vorhandene Konzepte wie URIs (Uniform Resource Indicator) und Unicode dar. Darauf aufbauend sind XML und RDF zentrale Austauschformate bei der Entwicklung von interdisziplinären Anwendungen. XML wurde zum einfachen Austausch von Daten entwickelt und ist die Grundlage für RDF, in dem jegliche Informationen des SW gespeichert werden. Auf die Wichtigkeit dieser Sprachen wird im Kapitel 5.3.2 näher eingegangen. Darauf aufbauend kann mit SPARQL bzw. OWL das semantische Netz abgefragt werden. Bisher gänzlich unbehandelt und daher für die Entwickler selbst zu lösen, sind die Ansätze für Proof & Trust. Das Vertrauen im SW ist eine der wichtigsten Maxime. Wenn sich Menschen auf die Entscheidungsfähigkeit der Maschinen verlassen wollen, muss gewährleistet sein, dass die Daten, auf deren Basis Software Agenten Entscheidungen treffen, 100%ig richtig sind.

4.2. Web 2.0

Die Begrifflichkeit des Web 2.0 wurde erstmals von Tim O'Reilly im Jahr 2004 geprägt. Er fasste viele Veränderungen zu einem Begriff zusammen, versäumte es jedoch den Begriff genau zu spezifizieren. Seiner Meinung nach werden unter dem Web 2.0 folgende Merkmale verstanden:

„Web 2.0 is the network as platform, spanning all connected devices; Web 2.0 applications are those that make the most of the intrinsic advantages of that platform: delivering software as a continually-updated service that gets better the more people use it, consuming and remixing data from multiple sources, including individual users, while providing their own data and services in a form that allows remixing by others, creating network effects through an "architecture of participation," and going beyond the page metaphor of Web 1.0 to deliver rich user experiences.“ (O'Reilly, 2005)

Durch die vage Formulierung des Begriffs in dem obigen Zitat entstanden viele Meinungen darüber, was unter dem neuen Buzzword zu verstehen sei - eine allgemein gültige Erklärung gibt es bis heute nicht. Tim Berners-Lee sagt über das Web 2.0 nüchtern: „... a piece of jargon, nobody even knows what it means.“ (BERNERS-LEE 2006). Durch die mediale Aufmerksamkeit, die dem Begriff ab 2006 zu Teil wurde, verbreitete sich der Begriff allerdings auch außerhalb der IT-Szene. Als die User des Web 2.0 zur Person des Jahres 2006 (vgl. GROSSMANN 2006) wurden, war der Begriff in aller Munde. Erstmals in der

Geschichte des amerikanischen Wahlkampfs stellten sich die Vertreter der demokratischen Partei Fragen von Usern, die auf der Videoplattform Youtube aufgezeichnet wurden. *„The lights and cameras were focused on the eight candidates, but it was the personal, heartfelt and, at times, comical nature of the user questions that stole the spotlight.“* (ANON, 2007) Journalisten waren sich einig: Nicht die Antworten waren die Highlights der Debatte, sondern die Fragen.

Die wichtigsten „Design Patterns“ des Web 2.0, die bereits in dem Zitat von Tim O’Reilly enthalten sind, zeigen, worauf es im *Mitmach-Web* ankommt:

Web als eine Plattform

Das Web stellt ein Medium dar, das von vielen, oftmals sehr verschiedenen Devices benutzt werden kann. Sei es der Zugriff auf Daten über ein Handy, einen Pocket PC, ein Notebook oder eine Großbildleinwand. Es ist zu einer Plattform für die Daten-Übertragung von modernen Applikationen geworden. Das Konzept des kompletten Neuladens von Seiten wird durch neuartige Mechanismen obsolet und viele Anwendungen stehen nur noch online zur Verfügung.

Die kollektive Intelligenz

Im neuen Web ist jeder in der Lage Inhalte zu erstellen. Wikis, Foto-Communities und das Verfassen von Blogbeiträgen bzw. Kommentaren in Foren zu jeglichen Themengebieten ermöglichen eine Demokratisierung der Inhalte. Anwendungen werden besser, je mehr Besucher diese aktiv benutzen. Das Kollektiv bestimmt den Erfolg oder Misserfolg von Portalen im Netz.

Social Tagging

Vordefinierte Navigationsstrukturen und starres Kategoriendenken wird durch das Beschlagworten von Ressourcen aller Art immer mehr in den Hintergrund verdrängt. Sogenannte Tag-Wolken liefern einen raschen Überblick über die angebotenen Inhalte, wichtige Begriffe werden größer dargestellt als weniger wichtige. Das Taggen von Informationen stößt jedoch rasch an seine Grenzen, da in den meisten Fällen keine inhaltliche/grammatikalische Überprüfung der Eingaben erfolgt.

4.2.1. Das Bottom-up Paradigma

Im Mittelpunkt des Web 2.0 stehen die Benutzer des WWW bzw. der einzelne Anwender. Das Phänomen stellt eine Demokratisierung des Internets dar. Benutzer erzeugen den Inhalt von Webapplikationen (vgl. Wikipedia, Youtube, Flickr, ...) und werden laufend über die Weiterentwicklung der angebotenen Features auf Webseiten befragt. Die Releasezyklen reduzieren sich auf die Ebene von wenigen Stunden, basierend auf dem Feedback der Benutzer werden Änderungen freigeschaltet oder verworfen. *„Coming of age for software developers means understanding that software is a cooperative effort, not something individuals do in isolation.“* Das Zitat von Tom DeMarco zeigt, dass bereits Branchenexperten die Wichtigkeit der gemeinsamen Entwicklung von Software begriffen haben. Der Ansatz des Web 2.0 ist daher bidirektionale Kommunikation. Während im *Web 1.0* keine Interaktionsmöglichkeiten bestand und Webangebote „read-only“ waren, findet man heutzutage auf gänzlich jeder Seite Formulare, über die Benutzer Kontakt mit anderen Benutzern oder den Seitenbetreibern aufnehmen können.

Im Vergleich zum Semantic Web gibt es keinen starren technologischen Aufbau, Web 2.0 ist die Menge aller Ideen das Web sozialer zu machen. Es ist daher jedem Benutzer möglich, selbst Teil des Web 2.0 zu werden, in dem er einen Blog startet, seine Favoriten öffentlich zugänglich macht oder selbst einen MashUp-Dienst (vgl. 5.2) entwickelt. Dadurch, dass keine allgemeingültige Definition vorhanden ist, kann jeder Benutzer das Web 2.0 täglich durch neue Ideen neu erfinden und zum Erfolg beitragen.

4.2.2. Kombination von Datenbeständen durch MashUps

Unter dem Begriff Mashup versteht man die Kombination von zwei oder mehreren Datenbeständen von Drittanbietern zu einem neuen Dienst. Diese Lösung ist in den letzten Monaten durch den Ausbau von Schnittstellen und angebotenen Daten ein großes Thema geworden. Viele Webplattformen stellen Teile ihrer Inhalte als standardisierte XML-Dateien (zum Beispiel in der Form von RSS-Feeds) zur Verfügung. Dies ermöglicht eine einfache Aggregation von Datenbeständen. Die Kombination dieser ist Basis für viele innovative Dienstleistungen im Web. Gute Beispiele für Mashup-Quellen sind Google Maps und Flickr. Google stellt mit seinen Maps eine Möglichkeit zur Verfügung Landkarten, Straßenkarten und Satellitenbilder mit zusätzlichen Informationen anzureichern. Es können auf den Karten eigene „Pins“ markiert werden, die zum Beispiel Fotos von den Stationen der eigenen letzten Weltreise zeigen können. Fügt ein Reisender seine Fotos auf Flickr ein und versieht diese mit Geoinformationen, wo das Foto aufgenommen wurde, lässt sich durch die Kombination von beiden Diensten auf der eigenen Webseite eine Karte einbinden, die zusätzlich zu den Pins der Aufenthaltsorte ebenso Fotos von den jeweiligen Orten/Ländern anzeigt.

Viele klassische Applikationen, die sich auf den Ansatz konzentrieren, sich auf ein Themengebiet zu fokussieren, können so leicht ihr Angebot um weitere sinnvolle Inhalte ergänzen. Der technische Aufwand ist hierbei minimal und durch die dynamische Verknüpfung mit anderen Webangeboten kann viel administrativer Aufwand entfallen.

In den letzten Monaten haben sich nicht nur eigentliche Mashups gebildet, sondern auch sogenannte *Mashup Application Creators* (vgl. Cubrilovic, 2007). Diese Anwendungen ermöglichen das einfache Erstellen von Mashups, ohne das Code hierzu geschrieben werden muss. Die hieraus resultierenden Applikationen können einfach in das eigene Web-Angebot inkludiert werden und stellen somit kosten- und zeitgünstige Alternativen zur eigenständigen Programmierung dar. Auf zwei ausgewählte Anbieter wird im Kapitel 5.2 näher eingegangen. Die in Kapitel 5 erstellte Konvergenzapplikation bedient sich ebenfalls der Datenintegration von mehreren Mashup-Datenquellen.

4.3. Theoretische Konvergenzen - Social Semantic Web

Obwohl beide Ansätze konträrer nicht sein könnten (vgl. Kapitel 4.1.1 und 4.2.1) geht der Trend eindeutig in eine gemeinsame Richtung. Beide Technologien können zu einem einzigen Zukunftsszenario zusammengefasst werden. Das Verschmelzen der beiden Technologien zu einem Produkt bezeichnet der Autor als „**Konvergenzapplikation**“. Experten streiten wie üblich über eine Namensgebung des Begriffs. Zum Zeitpunkt der Veröffentlichung der Arbeit sind die Begriffe „Social Semantic Web“ oder „Web 3.0“ aktuell. (vgl. Trendhistorie, Kapitel 2) Folgende Grafik veranschaulicht mögliche Konvergenzszenarien, die in Zukunft zu einem gemeinsamen Web verschmelzen könnten:

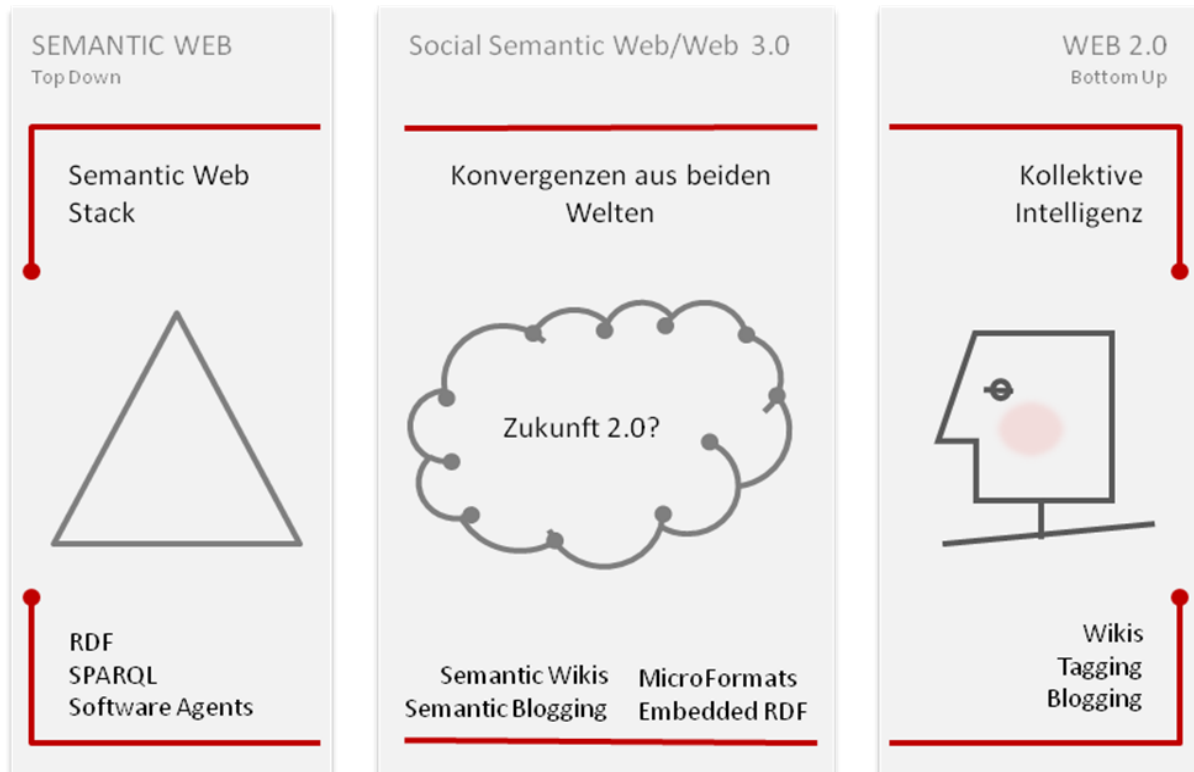


Abbildung 3: Konvergenzen zwischen SW und Web 2.0, Quelle: Eigenentwicklung

Was genau kann man sich jedoch unter dem „*Social Semantic Web*“ vorstellen?

„Das „Social Semantic Web“ vereint Social Software, wie sie im Web 2.0 zu finden ist, mit dem Semantic Web, das sich als nächste große Entwicklung des Web positioniert. Einerseits kann Social Software erheblich von Fortschritten im Bereich des Semantic Web profitieren, aber andererseits ist es eine rein technologische Entwicklung im Bereich des Semantic Web, die die sozialen Aspekte außer Acht lässt, aus unserer Sicht nicht zielführend.“ (Pelligrini, 2007)

Bei der Kombination beider Technologien gibt es viele mögliche Konvergenzszenarien, die in zwei große Themenblöcke geteilt werden können.

Web 2.0 Anwendungen werden um Semantik erweitert

Bei diesem Szenario werden bestehende Applikationen von der Semantic Web Gemeinde um Funktionen erweitert. Als Beispiel sei hier ein semantisches Wiki genannt. Die bisherige Software wird um die Möglichkeit der Entitäten ersetzt, die miteinander in Beziehung gesetzt werden können. Mit Hilfe dieser semantischen Techniken können Datenstände einerseits

dynamisiert werden. Übersichtsseiten sind in aktuellen Wikis händisch zu editieren, oftmals ergeben sich so lückenhafte Seiten. Durch den erweiterten Ansatz können so genannte *Inline-Queries* statische Inhalte ersetzen. Andererseits bietet das Entitäten-Modell die Möglichkeit der gezielten Suche. Abfragen wie „*Nenne mir alle Personen, die Herr Mustermann kennt*“ liefern genau die richtigen Personen als Ergebnis zurück. In einem konventionellen Wiki wären solche Suchabfragen nicht möglich. Dieser Ansatz wird in Zukunft sehr oft angewendet werden, weil bereits jetzt klassische Web 2.0 Angebote um Alleinstellungsmerkmale kämpfen müssen und der Einsatz von semantischen Technologien einen enormen Mehrwert für die Benutzer bringen kann.

Semantic Web Anwendungen werden um User generated content erweitert

Der Zweck von Ontologien bzw. Topic Maps ist in den meisten Fällen der Einsatz im Backoffice-Bereich. Das Netz als solches wird in heutigen Applikationen nur sehr selten dem Benutzer einer Applikation präsentiert bzw. als Navigationsübersicht zur Verfügung gestellt. Daher sind Semantic Web Applikationen meist für deren Benutzer nicht sichtbar. Dennoch ist eine Ontologie prädestiniert dafür, um mit Inhalten von Benutzern erweitert zu werden. Ein gutes Beispiel für dieses Szenario ist die Plattform *healthline.com*, die es ermöglicht Informationen zu medizinischen Themen durch den Einsatz einer Topic Map (so genannten Healthmaps) sehr schnell zu finden. Darüber hinaus können registrierte Mitglieder der Seite Inhalte in Form von eigenen Notizen und Tags erweitern. Dieser Ansatz ist im Vergleich zu oben genanntem Szenario weitaus seltener, da es zum heutigen Zeitpunkt noch sehr wenige reine Semantic Web Anwendungen am Markt gibt.

Es ist darauf hinzuweisen, dass der Schritt von einer Semantic Web Anwendung hin zu einer Konvergenzapplikation einfacher ist, da der Bottom-up Design des Web 2.0 einfacher in Applikationen zu integrieren ist. Im Gegensatz dazu ist ein enormes Know How für den Einsatz von semantischen Technologien erforderlich.

4.3.1. Potentielle Anwendungsbeispiele

Nicht immer müssen beide Ansätze komplett in eine Applikation übernommen werden. Öfter werden einzelne Teile von Webseiten (Navigation, Inhaltsbereiche, ...) um Social Features oder Topic Maps ergänzt. Dies hat den Vorteil, dass einerseits der Implementierungsaufwand deutlich niedriger ist und andererseits die Benutzer bereits in einer frühen Phase Feedback geben können.

Beispiele für solche Anwendungsgebiete sind semantische Tag Clouds. Durch einen Thesaurus oder einer Topic Map können gezielte Vorschläge für passende Tags abgeleitet werden – es können so etwaige (Grammatik-)Fehler minimiert werden und ähnliche Begriffe wie „Notebook“ und „Laptop“ zu einem Begriff zusammengefasst werden. Die daraus entstehende Wolke an Schlüsselwörtern ist weitaus repräsentativer als eine Standardwolke.

Das Speichern und Vergleichen von Surf-Pfaden ist ein Projekt von G. Beydoun, R. Kultchitsky und G. Manasseh, das zeigen soll, dass man mit Hilfe von ähnlichen Pfaden eine semantische Struktur bilden kann, die bei der Suche nach geeigneten Informationsquellen behilflich sein kann. „... *individuals in a given interest group produce trails that are of interest to the whole group. Moreover, special interest groups will be higher motivated than casual users to rate utility of pages they browse...*“ (Beydoun et al, 2006)

Die Verknüpfung von Elementen eines semantischen Netzes mit Kommentaren und automatischen Multimedia-Inhalten wird im Kapitel 5 näher beschrieben.

5. Konzeptionierung einer Konvergenz-Applikation

Um die theoretischen Grundlagen praktisch untermauern zu können, wurde im Rahmen dieser Bachelorarbeit eine eigene Applikation entwickelt. Die Vorgangsweise der Implementierung sowie eine Analyse von MashUp Creator Applications werden in folgendem Kapitel näher erläutert. Dabei wird nicht nur der automatische Datenimport näher erklärt, sondern auch Lösungsansätze für einen semantisch korrekten Export der Daten aufgezeigt.

Die auf den Markt befindlichen Semantic Web Applikationen sind für Endanwender noch nicht ausgereift genug und konzentrieren sich häufig sehr stark auf den technologischen Ansatz. Web 2.0 Plattformen sind genauso wie Semantic Web Applikationen sehr jung und ebenfalls erst im Aufbau. Es bleibt daher für eine Kombination der beiden Technologien in den meisten Fällen für Entwickler keine Zeit, um die Ansätze zu verbinden. Die eigens erstellte Konvergenzapplikation soll als „*proof of concept*“ dienen und zeigen, ob die Kombination der beiden Ansätze sinnvoll ist. Folgende Features werden in Phase 1 der Applikation eingebaut:

Auf Basis einer Ontologie für den Städtetourismus, die sowohl die Strukturinformationen und die wichtigsten Beziehungen der Entitäten untereinander beinhaltet, sollen Testdatensätze für die Stadt Wien angelegt werden. Das so erstellte semantische Netz ist Grundlage für die Navigationsstruktur und den Aufbau der Webseite. Zusätzlich zu den statischen Entitäten, die nur durch eine Administrationsoberfläche gewartet werden können, werden weitere Inhalte von Drittanbietern automatisiert in die Anwendung geladen. Fotos und Veranstaltungsinformationen zu wichtigen Gebäuden, sowie eine Kommentarmöglichkeit runden die Applikation ab.

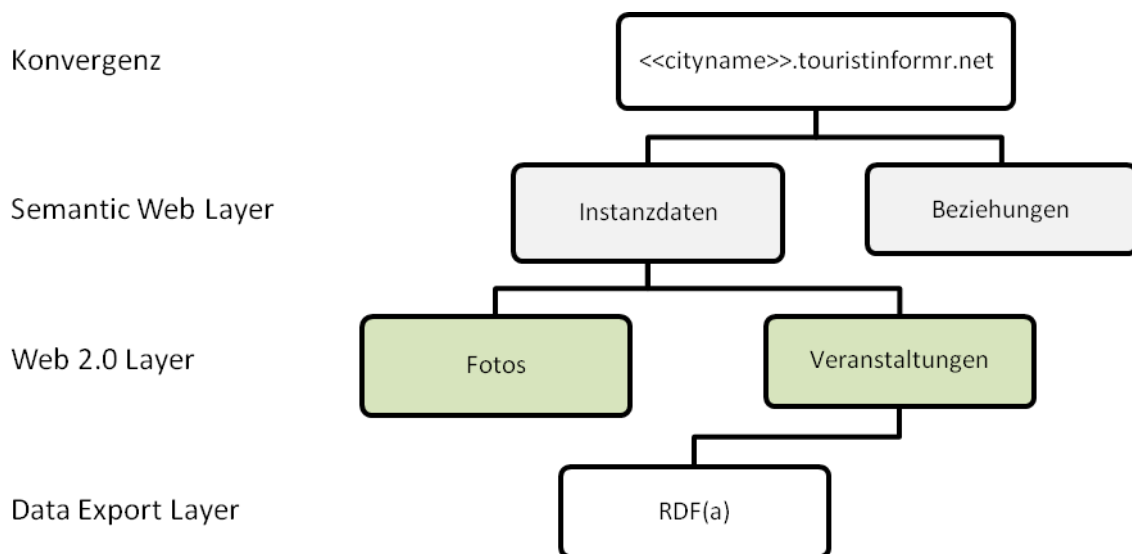


Abbildung 4: Aufbau der Beispielanwendung touristinformr.net, Quelle: Eigenentwicklung

Abbildung 4 zeigt den schematischen Aufbau der Applikation. Auf die technische Implementierung wird in dieser Arbeit nicht näher eingegangen, da diese den Rahmen der Arbeit deutlich übertreffen würde. Es wird versucht alle Teile der Applikation selbst zu programmieren. Die Speicherung der Ontologie erfolgt nicht in reinen RDF Dateien, wie es dem aktuellen Semantic Web Stack entsprechen würde, sondern in einem adaptierten Tripplstore in Datenbankform. Dieser Ansatz hat zur Folge, dass das semantische Netz

nicht über SPARQL (Abfragesprache für RDF-Graphen) abfragbar ist, ermöglicht aber eine einfachere Implementierung, ohne auf Drittanbieter-Produkte zurückgreifen zu müssen. Dem Autor war eine komplette eigenständige Entwicklung wichtig, da nur in diesem Fall ein aussagekräftiges Resümee möglich ist.

5.1. Die Erstellung eines semantischen Netzes

Die folgenden Unterkapitel sollen den Einsatz und die Modellierung einer Ontologie verdeutlichen. Die sehr abstrakte Definition „*An ontology is a specification of a conceptualization.*“ (Gruber, 1993) des Begriffs zeigt bereits, dass der Aufbau eines semantischen Netzes nicht trivial ist. Es erfordert Know How und Abstraktionsfähigkeit.

Eine Ontologie im Semantic Web ist ein Datenmodell für einen bestimmten Zweck (=Domäne), die Informationen über Entitäten und deren Beziehungen zueinander definiert. Dieses Netz dient als Meta-Schnittstelle zur Speicherung und Aufbereitung von Datenbeständen.

“The result is an information structure that breaks out of the traditional hierarchical straightjacket that we have gotten used to squeezing our information into. A topic map usually contains several overlapping hierarchies which are rich with semantic cross-links like “Part X is critical to procedure V.” This makes information much easier to find because you no longer act as the designers expected you to; there are multiple redundant navigation paths that will lead you to the same answer. You can even use searches to jump to a good starting point for navigation. The most common use for topic maps right now is to build web sites that are entirely driven by the topic map, in order to fully realize their information-finding benefits. The topic map provides the site structure, and the page content is taken partly from the topic map itself, and partly from the occurrences. This solution is perfect for all sorts of portals, catalogs, site indexes, and so on.” (Garshol, 2002)

Die erstellte Topic Map (Ontologie) stellt, wie L. Garshol bereits 2002 feststellte, eine ideale Lösung zur Aufbereitung von Navigationsstrukturen dar. Dennoch blieb der Einsatz der semantischen Netze in den letzten 5 Jahren eine Seltenheit und die semantischen Netze konnten sich im Vergleich zu hierarchischen Katalogstrukturen noch nicht am Markt etablieren.

5.1.1. Methodik der Erstellung

Bei der Erstellung einer Ontologie sollte man immer die gleiche Vorgehensweise einhalten. Dabei ist zu beachten, dass das Rad nicht immer neu erfunden werden muss. Unter der Adresse www.schemaweb.info gibt es eine Ansammlung an vorgefertigten Ontologien, die man beliebig um seine eigenen Wünsche und Anforderungen erweitern kann. Die Methodik der Erstellung sollte dem Wasserfallprinzip entsprechen. Alle Schritte müssen für einen Modellierzyklus durchlaufen werden. Wird die Reihenfolge nicht eingehalten, kann dies zu inkonsistenten Datenmodellen führen.

Entitätenmodell

Der erste Schritt beim Erstellen des Modells ist die sinnvolle Abgrenzung der Domäne. Die wichtigsten generischen Objekte müssen definiert werden und stellen die Grundlage für das weitere Vorgehen dar. Die Erstellung des Entitätenmodells ist der wichtigste Schritt bei der Einführung eines semantischen Netzes, da die abstrakten Objekte Ausgangspunkt für die Instanzdaten darstellen. Wenn in der Modellierung der Entitäten ein Fehler unterlaufen ist,

und Instanzdaten bereits angelegt wurden, kann der Fehler nur noch schwer rückgängig gemacht werden.

Beziehungen

Wenn die Entitäten fertig modelliert wurden, müssen auf Meta-Ebene die Beziehungen definiert werden. Entwickler sollten darauf achten, dass immer von einer Entität ausgehend alle möglichen Beziehungen definiert werden. Durch das Einbinden von bestehenden Ontologien und Namensräumen kann hierbei auf bereits bestehende Beziehungstypen Rücksicht genommen werden, bzw. ein Mapping der eigenen Beziehungen zu standardisierten Beziehungen vorgenommen werden. Beliebte Beispiele sind hierfür die Namensräume von Friend of a Friend, Dublin Core und den Creative Commons. Das Berücksichtigen von bestehenden Beziehungen in dieser Phase der Entwicklung hilft Zeit in der Phase des Exports bzw. des Datenabgleichs zu minimieren.

Instanzdaten (händisch/automatisch)

Nachdem auf Meta-Ebene Entitäten und deren Beziehungen geregelt wurden kann die Erstellung von Instanzdaten durch Redakteure erfolgen. Die Eingabe der Daten sollte möglichst simpel anwendbar sein, da die Generierung von Instanzdaten den meisten und zeitintensivsten Aufwand der Arbeit darstellt. Ein automatisches Erstellen von Instanzdaten bestehend aus Vordaten ist ebenfalls denkbar. Hierzu unbedingt notwendig ist das richtige Mapping von bestehenden Daten auf das erstellte semantische Entitätenmodell. Ein beliebtes Tool zur einfachen Erstellung von Datenbank basierten Beständen in RDF-Triple Form ist D2RQ. (vgl. <http://sourceforge.net/projects/d2rq-map/>)

5.1.2. Beispiel-Netz für den Städte Tourismus

Das zugrunde liegende semantische Netz für die Plattform *touristinformr.net* beinhaltet Instanzdaten zu 5 Haupt-Kategorien, die ebenfalls auch als Navigationselemente heranzuziehen sind. Bei der Entwicklung des Netzes wurde die oben genannte Methodik befolgt. Es wurde ein allgemeingültiges Objektmodell angelegt und mit Instanzdaten befüllt. Erst im zweiten Schritt wurden spezifische Objekte für die Stadt Wien ergänzt. Die exemplarische Verknüpfung der einzelnen Entitäten wird am Beispiel der Instanz „*Wiener Rathaus*“ näher erklärt. Die vollständige Verknüpfung der Datensätze ist online ersichtlich und ein Auszug des Aufbaus befindet sich im Anhang.

Das Wiener Rathaus ist als Typ Sehenswürdigkeit definiert und besitzt die Beziehungen:

- Ist Austragungsort von Lokales Event:
 - Life Ball
 - Wiener Christkindlmarkt
- Ist in der Nähe von Haltestelle
 - Rathaus Haltestelle

Möchte ein Besucher der Seite Informationen über weitere Sehenswürdigkeiten in der Nähe von Rathaus Haltestelle, kann er durch einen Klick feststellen, dass es sich hierbei um eine U-Bahnhaltestelle der U2 handelt, und sich in der Nähe der Haltestelle ebenfalls das Österreichische Parlament sowie das Theater an der Josefsstadt befinden. Die exemplarischen Verknüpfungen der Daten zeigen, wie rasch man über das Umfeld eines Objekts Informationen darstellen kann und diese für den Besucher strukturiert angezeigt werden. Weitere Navigationspfade könnten nun zu einer Liste an Stationen der U-Bahnlinie U2 führen, oder zu dem Österreichischen Parlament, das durch die Volkstheater Haltestelle ebenfalls in der Nähe des Museumsquartiers liegt.

Eine Messung der User Trails (vgl. den gleichnamigen Ansatz von Beydoun in Kapitel 4.3.1) und das kontinuierliche Anreichern der Daten mit weiteren Verknüpfungen würden mögliche

Erweiterungsszenarien des semantischen Netzes darstellen. Vor allem die händische Verknüpfung der in diesem Fall wenigen Entitäten (rund 300) stellt allerdings einen enormen Aufwand dar. Abbildung 5 zeigt einen exemplarischen Ausschnitt aus den Entitäten und deren Verbindungen als ein graphisch-aufbereitetes Netz. Zum Zeitpunkt des Screenshots wurde auf eine andere Technologie gesetzt, die in der finalen Version nicht mehr in der Applikation inkludiert ist. Die Visualisierung der Informationen zeigt wie eng die Daten miteinander verbunden sind und dass bereits bei kleinen Datenmengen die kontext-relevanten Informationen enorm hoch sein können.



Abbildung 5: Grafische Darstellung des semantischen Netzes, Quelle: Screenshot

Durch den Einsatz von weiteren semantischen Technologien wie *Reasoning* könnten anhand von Testdaten Zusammenhänge automatisch abgeleitet werden und Daten automatisch um weitere Beziehungen angereichert werden.

5.2. Automatisierter Datenimport mittels Mash Ups

Wie bereits in Kapitel 4.2.2 beschrieben, läuft der Trend in die Richtung der Kombination von Datenbeständen. Durch die einfache technische Abwicklung der Integration bzw. Vermischung von Daten stellt sich die Frage, ob Web-Plattformen gänzlich automatisch befüllt werden können, oder weiterhin redaktionelle Arbeit erforderlich ist.

Die Vermischung digitaler Quellen ist auf die Musikindustrie zurückzuführen. In den Anfängen des 21. Jahrhunderts kreieren Produzenten durch das Übereinanderlegen von mehreren, meist sehr verschiedenen Musikrichtungen, neue Lieder. Diese so genannten „Bootleg Remixes“ haben bereits eine große Fangemeinde gefunden und stellen eine wichtige Basis für die Content Revolution im Internet dar. (vgl. Rötgers, 2007)

Bevor auf die verschiedenen Möglichkeiten zur Erstellung von Mashups im Internet eingegangen wird, möchte der Autor auf die Copyright Thematik hinweisen. Die Befugnis auf

Datenbestände zugreifen zu dürfen, stellt den wichtigsten Teil der Konzeptionierung eines Mashups dar. Ohne die Erlaubnis befindet man sich auf rechtlich grauem Niveau und man läuft Gefahr viele Stunden der Programmierung/Integration zu verlieren, wenn der Inhaber der Daten die Vermischung mit den eigenen nachträglich verbietet.

5.2.1. Eigenständige Programmierung

Die Kombination von Datenbeständen ist für Firmen bereits immer eines der zentralen Themen der internen IT gewesen. Die vielen dezentralen Datenbanken (ERP-System, Customer Relationship Management, Lagerbestände) müssen seit den Anfängen der Datenverarbeitung miteinander abgeglichen werden. Es bestanden daher in der Vergangenheit bereits viele Standardschnittstellen (zB Edifact), die den Austausch von standardisierten Datentypen regelten. Neu am Ansatz des Web 2.0 allerdings ist das Kombinieren von Daten, die am ersten Blick nicht miteinander in Verbindung stehen. Es erforderte daher in der Vergangenheit viel Aufwand um Datenbestände miteinander verbinden zu können.

Durch die Entwicklung von XML und darauf aufbauenden Formaten wie zum Beispiel RSS war es jedoch erstmals möglich mit wenigen Programmierkenntnissen Datenbestände verarbeiten zu können. Anfangs wurden RSS Feeds zur Aufbereitung von Blog-Inhalten entwickelt, doch rasch entwickelte sich das Format zu einem Standard für dynamische Listen im Internet. Heutzutage werden Musikgewohnheiten, Eventinformationen, aktuelle Auktionsartikel und Fotos generisch über RSS Feeds angeboten und meist um eigene Attribute erweitert.

Die eigenständige Programmierung von Importschnittstellen stellt daher keine unüberwindbare Hürde dar. Dennoch muss ein Entwickler einige Dinge berücksichtigen. Der Abgleich von Dubletten sowie das Überprüfen auf gültige Daten in einzelnen Feldern bleibt weiterhin in der Obhut des Programmierers. Als Beispiel seien hier Datumsfelder genannt, die in jedem System unterschiedlich behandelt werden. Es müssen daher für alle Datenquellen eigene Umwandlungsmethoden eingeführt werden.

Aus einer eigenen Implementierung können viele Vorteile entstehen. Das Persistieren der Datenbestände sowie die Möglichkeit einer internen generischen Datenstruktur sind nur zwei Beispiele, die für eine Eigenentwicklung sprechen könnten. Der Zeit- und Know How Aufwand ist jedoch für viele Hobby-Programmierer abschreckend. Daher wird in den nächsten zwei Kapiteln auf Produkte eingegangen, die die Erstellung von Mashups für die Masse vereinfachen.

5.2.2. Yahoo Pipes

Yahoo Pipes gilt als der erste Vertreter von MashUp Creator Applikationen und wurde im Februar 2007 veröffentlicht. Ziel der Pipes (=Röhren) ist es, Daten miteinander zu verbinden und diese aufbereitet wiederzugeben („Feed Mashup and Manipulation Engine.“). Erstellte MashUps werden als XML-Dateien zur Verfügung gestellt. Es erfordert daher eine eigene Applikation, die den Output der Pipes weiterverwenden kann. Die vordefinierten Module richten sich an technisch versierte Personen, die sehr detaillierte Pläne der Datenaggregation verfolgen. Die angebotenen Beispiele sind sehr umfangreich und nicht immer für Anfänger geeignet, die oftmals keinen technischen Background haben. Zum Benutzen des Dienstes ist eine Yahoo ID erforderlich. Öffentliche Pipes können von anderen Mitgliedern geklont werden, um diese für eigene Zwecke zu verändern, bzw. auch als Quelle für deren MashUps dienen. Die Abbildung 6 zeigt einen Transformationsschritt und die im Vergleich zu Popfly sehr technische Oberfläche von Yahoo Pipes.

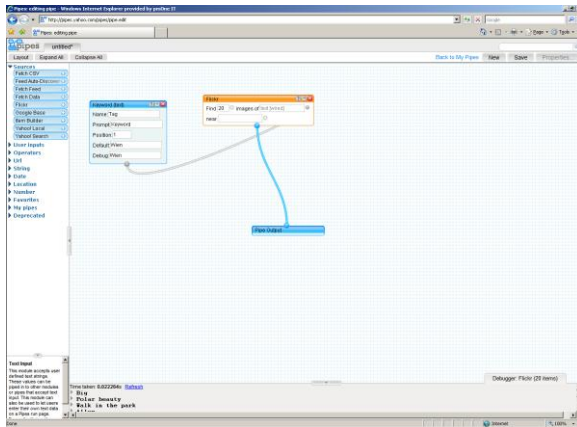


Abbildung 6: Entwicklungsumgebung Yahoo Pipes, Quelle: Screenshot

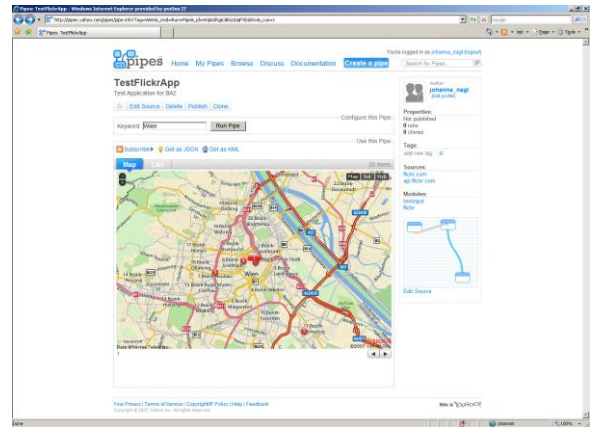


Abbildung 7: Möglicher Output Yahoo Pipes, Quelle: Screenshot

Die gefertigten Pipes können einerseits über die Yahoo Plattform aufgerufen werden und andererseits als RSS/JSON-Feeds aufgerufen werden. Yahoo Pipes berücksichtigt keine Module für die Anzeige der aggregierten Daten, es wird jedoch automatisch ein Yahoo Maps Kartendialog eingeblendet, sobald Geo-Informationen (siehe Abb. 7) in den aggregierten Daten enthalten sind. Für die Verwendung der Daten in einer bestehenden Web-Applikation muss daher ein Import-Szenario vorhanden sein, dass die Daten zumindest durch eine XSLT-Transformation visualisieren kann.

5.2.3. Microsoft Popfly

Microsoft Popfly wurde im Mai 2007 der Öffentlichkeit präsentiert und befindet sich seither in einem Alpha Stadium. Der Trend, Software im Beta Stadium zu veröffentlichen, wird hierbei noch weiter vor-verlagert. Der Dienst steht ausschließlich einem eingeschränkten Kreis zur Verfügung. Um Popfly benutzen zu können benötigen Besucher eine Einladung. Um den vollen Funktionsumfang (wie die Erstellung *und* Anzeige eigener Mashups) nutzen zu können, muss Microsoft Silverlight (vgl. Adobe Flash) installiert werden. Microsoft zwingt Benutzer seine eigenen Programme zu installieren und so bekannter zu machen.

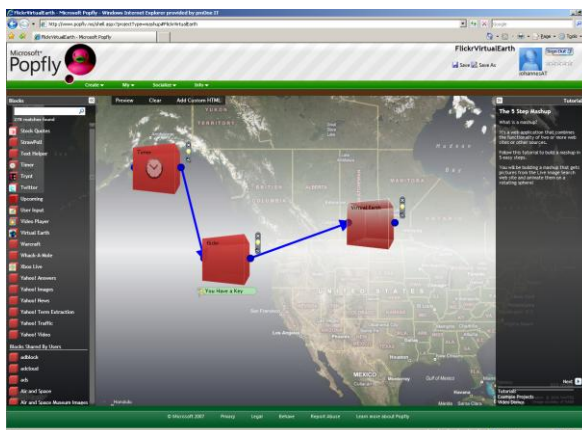


Abbildung 8: Entwicklungsumgebung MS Popfly, Quelle: Screenshot



Abbildung 9: Eigenschaften eines Blocks MS Popfly, Quelle: Screenshot

Im Gegensatz zu Yahoo Pipes ist das Konzept von Popfly stark vereinfacht. Module werden „Blocks“ (vgl. Abbildung 8) genannt, die immer gleich aussehen/anwendbar sind. Es können

wie bei Yahoo Pipes für jeden Block beliebige Input und Output-Verbindungen definiert werden. Das System erkennt jedoch automatisch mögliche Übereinstimmungen der Daten. Zu den vorgefertigten Blocks können beliebige Module erweitert und der Community zur Verfügung gestellt werden. Popfly beschränkt sich nicht nur auf die Generierung von XML-Dateien, sondern stellt ebenfalls Möglichkeiten zur Anzeige der Daten zur Verfügung. Eigene Blocks für Virtual Earth (Kartografie), eine Slideshow zur Bilderdarstellung, sowie ein Modul für die Erstellung von Balkengrafiken sind nur wenige Beispiele für das reichhaltige Angebot an Standardmodulen.

Die standardisierten Abläufe/Module können durch eigenen HTML/JS Code ergänzt werden und Microsoft bietet neben dem *Popfly Creator* (Online-Tool zur Erstellung von Mashups) zusätzlich den *Popfly Explorer* an, der ein Addin für Visual Studio (Programmierumgebung) darstellt. Popfly kann somit das gesamte Spektrum der Entwicklung abdecken. Der Fokus der Applikation liegt dennoch bei Benutzern, die keinen Code schreiben und rasch zum Ziel kommen möchten. Die Einbindung des erstellten Mashups ist zum Zeitpunkt der Analyse leider nur über IFrames möglich. Dieser Ansatz ist für Produktivsysteme nicht praktikabel. Da ebenfalls keine Möglichkeit zum Export der Daten als reine XML-Feeds besteht ist das System nur in den wenigsten Fällen einsetzbar.

5.2.4. Vergleich der Methodiken

Der nachfolgende Vergleich, den der Autor auf Grund einer methodischen Analyse erstellt hat, stellt die 3 verschiedenen Möglichkeiten zur Aggregation von Daten gegenüber und ermittelt einen Gesamtsieger. In den folgenden Vergleich fließen verschiedenste Kriterien ein, die für die Erstellung von MashUps notwendig sind. Die Gewichtungsskala stellt eine subjektive Bewertung der Bedeutung jedes einzelnen Kriteriums dar.

- **Entwicklungsaufwand in Stunden**
Wie lange dauert eine Musterimplementierung eines einfachen Mashups? Je weniger Zeit notwendig ist, desto besser. Der Aufwand sollte gering ausfallen, da sonst eine Ankoppelung von Drittanbieter-Daten sehr teuer werden kann und die wirtschaftliche Sinnhaftigkeit in den Raum gestellt werden muss.
- **Anzahl von ausgelieferten Standardmodulen**
Bietet der Technologieansatz bereits fertige Module, die eine Implementierung vereinfachen? Je mehr Module vorhanden sind, desto besser.
- **Verfügbarkeit von Dokumentation und Beispielcode**
Gibt es für die Entwicklung eine vorliegende Dokumentation und Beispielmuster? Je mehr, desto besser. (Bei der eigenständigen Entwicklung hängt dieser Punkt sehr stark von der verwendeten Programmiersprache ab. Es wird jedoch davon ausgegangen, dass Dokumentationen zu der exemplarischen Struktur von MashUp-Diensten im Internet vorhanden sind.)
- **Leichtigkeit der Integration in bestehende Applikationen**
Wie einfach ist die Anpassung/Einbindung der MashUp-Daten in bestehende Applikationen? Welche technologischen Ansätze sind möglich? Je einfacher, desto besser. Dieses Kriterium ist aus Sicht des Autors das wichtigste, da eine fließende Integration in bestehende Applikationen das Erstellen von MashUps noch attraktiver macht.
- **Notwendiges, vorhandenes Know-How**
Wie viel Vorwissen ist notwendig, um den Technologieansatz einsetzen zu können? Ist der Einsatz für die breite Masse geeignet oder stellt die Benutzung einen enormen

Lernaufwand dar. Je weniger, desto besser. Dieses Kriterium verhält sich ähnlich wie der Entwicklungsaufwand, denn eine zu große Einarbeitungsphase sollte in einem Outsourcing-Prozess enden.

Ergebnis:

	Gewichtung	Entwicklung	Popfly	Pipes
Entwicklungsaufwand*	20	0	20	12
Standardmodule	15	0	15	8
Dokumentation/Beispiele	10	5	5	10
Integration	35	35	5	30
Know-How*	20	0	20	10
Summe	100	40	65	70

Das Ergebnis zeigt, dass der Platzhirsch Pipes im Endergebnis die Nase vorne hat. Der Microsoft Dienst Popfly schlägt den Primus zwar im Entwicklungsaufwand, angebotenen Modulen und dem aufzubringenden Know-How, landet bei der Integration der erstellten MashUps und deren Daten jedoch an letzter Stelle. Sollte Microsoft hier bis zur Fertigstellung des Produkts nachbessern, stellt Popfly ein sehr mächtiges Tool zur Erstellung von MashUps dar. Das Konzept der generischen Blocks und die Möglichkeit eigene Blocks zu erstellen/von der Community zu beziehen, ist im Gegensatz zu Yahoo Pipes weitaus fortgeschrittener.

Die Eigenentwicklung verliert in allen Kategorien, da sie zeitaufwändig ist und viel Know-How voraussetzt, bei der Implementierung in (bestehende) Applikationen ist die Flexibilität jedoch nicht zu schlagen. Das Resultat des Vergleichs ist daher: Je tiefer die Integration eines MashUps in die eigene Anwendung ist, desto mehr sollte man eine Eigenentwicklung forcieren und an die gewünschten Bedürfnisse anpassen. Möchte man rasch zu einem vorzeigbaren Ergebnis kommen, dann ist der Einsatz von Yahoo Pipes oder Microsoft Popfly die klügere Möglichkeit.

5.3. Bereitstellung semantischer Daten

Die in Kapitel 5.1 erstellte Ontologie dient in der Konvergenzapplikation dazu, um die Navigation für Besucher der Tourismusplattform zu vereinfachen. Die kontextsensitiven Daten und die Beziehung der Objekte untereinander führen zu Navigationselementen, die konventionelle Applikationen nicht (oder nur bedingt) bieten können. Damit die angebotenen Inhalte allerdings auch für Maschinen lesbar gemacht werden können, muss die Anzeige der Daten adaptiert werden.

Die theoretischen Grundlagen über das Semantic Web haben gezeigt, dass Daten über einfache Triples (Subjekt, Prädikat, Objekt) abgebildet werden.

Grundsätzlich unterscheidet man bei der Bereitstellung bzw. dem Export von semantischen Daten zwei Ansätze:

- externe RDF Dateien, die nur für Computer lesbare Informationen enthalten, sowie
- adaptierte (X)HTML-Seiten, die die gewohnten Inhalte für den Menschen enthalten, zusätzlich jedoch mit erweitertem Markup ausgestattet werden.

* Für mit * gekennzeichnete Kriterien gelten umgekehrte Werteskalen, die in der Punktevergabe bereits berücksichtigt wurden.

Wenn bereits konkrete Exportszenarien vorliegen bzw. Content-Partnerschaften geschlossen werden, ist der Einsatz von externen RDF Dateien vorzuziehen. Wenn, wie in dem Fall der Tourismusplattform, der Fokus auf den Besuchern der Seite liegt und das Bereitstellen von Maschinen-interpretierbaren Informationen ein „nice-to-have“ Feature ist, empfiehlt sich der Ansatz der Adaptierung der bestehenden Webseiten.

Das Integrieren von RDFa bzw. Mikroformaten wird auf Grund der aktuell größeren Verbreitung in den folgenden Unterkapiteln näher analysiert.

5.3.1. Microformats

Der Ansatz von Mikroformaten in Web-Applikationen ist „as simple as possible“. Es wird versucht für spezifische Probleme Lösungsansätze zu definieren, die in erster Linie für den Menschen designt werden – erst im zweiten Schritt wird an die Maschinen gedacht. (vgl. ANON, 2007)

„Microformats are an implementation of the REST (Representation State Transfer) architectural style. REST is a method for sending domain-specific data over HTTP without using a messaging layer like the XML-based SOAP (Simple Object Access Protocol). Microformats are one way to implement the REST style by mixing metadata with HTML or XHTML. [...] The big difference between the microformats version and a comparable XML encoding is that the microformats can be displayed without an XSLT transformation file. With XML data files, the machine gets preference. But with microformats, the human-readable form is augmented so software can extract useful information.“ (Elecdesign, 2007)

Vereinfacht gesagt, wird bestehender HTML-Code um zusätzliche, für spezielle Domänen standardisierte Attribute ergänzt. Mikroformate stellen offene Standards dar, die von Personen eingereicht werden können. Es steht jedoch keine namhafte Organisation hinter diesen *Vorschlägen für Standards*. Nachfolgendes Code-Listing zeigt den Einsatz des hCalendar-Formats, der die Struktur für Kalenderdaten regelt.

```
<div class="vevent">
<a class="url" href="http://www.technikum-wien.at/">http://www.technikum-wien.at/</a>
<span class="summary">Präsentation der Bachelorarbeit im Rahem der Bachelorprüfung</span>:
<abbr class="dtstart" title="2007-09-07">7. September</abbr>-
<abbr class="dtend" title="2007-09-11">11. September</abbr>,
im <span class="location">Seminarraum B2.07</span>
</div>
```

Die Hinterlegung der *class*-Attribute mit den Werten „vevent“, „url“, „summary“, „dtstart“, „dtend“, „location“ ermöglicht einem Browser, der Mikroformate unterstützt, die einfache Extraktion der Daten als Kalenderobjekte. Wenn die Anzeigetexte nicht mit den tatsächlichen Werten (wie in diesem Beispiel bei den Datumfeldern der Fall) übereinstimmen, werden diese in dem *title*-Attribute näher spezifiziert. Der HTML-Code bleibt in den meisten Fällen sehr schlank und Mikroformate können schnell in bestehende Applikationen integriert werden, ohne große Veränderungen am Code vornehmen zu müssen. *„The big difference between the microformat version and a comparable XML encoding is that the microformat can be displayed without an XSLT transformation file. [...] the human-readable form is augmented so software can extract useful information.“ (Wong, 2007).*

Das Design des simplen Integrierens wurde von den Entwicklern wie in Abbildung 10 sichtbar umgesetzt. Die Basis der Microformats stellt XHTML dar, das um die spezifizierten Formate ergänzt wird.

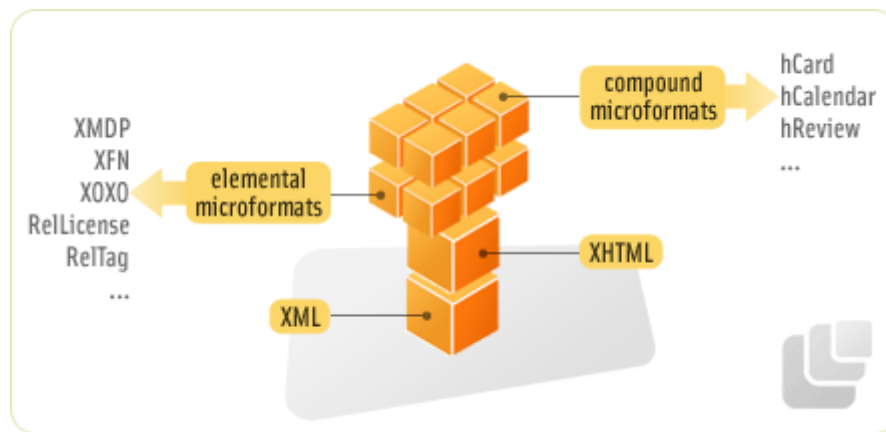


Abbildung 10: Aufbau der Microformats, Quelle: microformats.org

Die domänenspezifische Umsetzung der Formate macht den Einsatz von Mikroformaten allerdings unflexibel. Sie sind zwar sehr einfach für Insellösungen zu benutzen, wenn es aber um die Verknüpfung von verschiedenen Entitäten geht, schlägt der Ansatz fehl. Im zweistufigen Ansatz werden „elemental microformats“ und „compound microformats“ definiert, es fehlt jedoch eine weitere, höhere Stufe, die unterschiedliche Formate und deren Attribute miteinander verbinden könnte.

5.3.2. RDFa

RDFa ist in seiner jetzigen Form ein so genanntes Working Draft des W3C. Es dient zum Einbetten von RDF direkt in (X)HTML-Code. In erster Linie ist die Idee eine ähnliche wie bei den Mikroformaten, da der Einsatz ebenfalls über den REST Architekturstil erfolgt. Die RDF Daten werden also über denselben HTTP-Kanal wie die für den Benutzer sichtbaren Daten gesendet.

“RDFa is a syntax that expresses this structured data using a set of elements and attributes that embed RDF in HTML. An important goal of RDFa is to achieve this RDF embedding without repeating existing HTML content when that content is the structured data. RDFa is designed to work with different XML dialects, e.g. XHTML1, SVG, etc. given proper schema additions.” (Adida & Berdeck, 2007)

Der Ansatz von RDFa ist jedoch im Gegensatz zu den Microformats deutlich ausgereifter. Durch den mehrstufigen Semantic Web Stack lassen sich die Attribute untereinander kombinieren und somit einfacher wiederverwenden. Der Einsatz von Namensräumen regelt die Beziehung der Attribute untereinander.

```
<html xmlns:cal="http://www.w3.org/2002/12/cal/ical#"
      xmlns:contact="http://www.w3.org/2001/vcard-rdf/3.0#"
  ...
  <p class="cal:Vevent" about="#bachelorexamen">
  Ich trete am
  <span property="cal:dtstart" content="20070907T1000+0100"> 7. September um 10 Uhr</span> zur
  <span property="cal:summary">Bachelorprüfung am Technikum Wien</span> an.
  </p>
  <p class="contactinfo" about="http://johannes.nagl.name/">
  Mein Name ist <span property="contact:fn">Johannes Nagl</span>.
  Ich bin <span property="contact:title">Student</span> an der <a rel="contact:org"
  href="http://www.technikum-wien.at/">FH Technikum Wien</a>. Sie können mich unter der Adresse
  <a rel="contact:email" href="mailto:johannes.nagl@technikum-wien.at">johannes.nagl@technikum-
  wien.at</a> erreichen.
  </p>
  ...
  </html>
```

Durch die Angabe der XML-Namensräumen werden Attribute aus Ontologien referenziert, die sowohl für strukturierte Ansätze wie der des Kalendereintrags (vgl. Code-Listing des Microformats) angewendet werden können. Darüber hinaus können Attribute aus anderen Ontologien ebenfalls mitbenutzt/referenziert werden. Ein übergreifender Einsatz der Attribute ist hierdurch gewährleistet und bietet Vorteile, wie zum Beispiel das Suchen über verschiedene Domänen hinweg.

Der Implementierungsaufwand ist im Vergleich zu den Mikroformaten höher, da eine Angabe der Namensräume zwingend erforderlich ist. RDFa-fähige Browser können allerdings aus den RDFa Daten gültige Trippel extrahieren, die leicht für Maschinen verständlich sind. *„At the moment, it is very difficult for software like web browsers and search engines to make use of this implicit data. We need a standard mechanism to explicitly express it. This is precisely where RDFa comes in.”* (ebenda, 2007)

Die aus obigem Code-Listing entstandenen Trippel würden wie folgt lauten:

```
<http://johannes.blogt.nursoals.info/?p=Bachelorarbeit#bachelorexamen>
  Rdf:type cal:Vevent;
  Cal:summary „Bachelorprüfung am Technikum Wien“^^XMLLiteral;
  Cal:dtstart „20070907T1000+0100“ .

<http://johannes.nagl.name/>
  Contact:fn „Johannes Nagl“^^XMLLiteral;
  Contact:title "Student"^^XMLLiteral;
  Contact:org <http://www.technikum-wien.at/>;
  Contact:email <mailto:johannes.nagl@technikum-wien.at> .
```

Ein Software Agent des Betreuers von Johannes Nagl wäre nun in der Lage unter der Berücksichtigung obiger Trippel alle relevanten Informationen auszulesen, automatisch zu überprüfen, ob bereits andere Termine zu dem vorgeschlagenen Termin eingetragen sind und könnte diese bei Bedarf automatisch verschieben und E-Mails an alle Betroffenen mit Vorschlägen für neue Termine verschicken.

6. Kurzer Vergleich von Tourismus-Plattformen

Um einen Überblick über die Stärken und Schwächen der entwickelten Applikation zu bekommen, ist ein Vergleich mit anderen Anbietern unerlässlich. Der Vergleich in diesem Kapitel soll zeigen, ob der exemplarische Ansatz des Social Semantic Web für den Tourismussektor neue Impulse bringen kann, oder ob bereits vorhandene Angebote die Vorteile der sozialen Interaktion und die vernetzte Strukturierung von Daten implementiert haben.

Dazu wurden die angebotenen Inhalte von *touristinformr.net* mit drei weiteren Anbietern, die Inhalte über die Stadt Wien präsentieren, verglichen. Es handelt sich dabei um den Webauftritt des Wien Tourismus, der eine Messlatte für die inhaltliche Komponente darstellen sollte. Die Seite *wikitravel.org* stellt einen Vertreter der Web 2.0 Nische dar und soll zeigen, wie stark das Web 2.0 bereits Einzug in den Tourismussektor gehalten hat. Als dritten Vergleichsauftritt wurden die Wiener Webseiten der Firma *tiscover.com* analysiert, die vor allem im deutschsprachigen Raum eine Vorreiterrolle im virtuellen Tourismus einnimmt. Im Jahr 2001 erlangte die Firma große mediale Aufmerksamkeit, als zum Start der Registrierungsphase der TLD „.info“ 5000 Städte-Domainnamen durch *tiscover* registriert wurden.

„Weil der Name von Orten und Städten eben nicht als Warenzeichen eingetragen werden könne, habe man in geschäftlichem Eigeninteresse handeln müssen – und zwar auch schon in der Sunrise-Periode, in der nach Vorstellungen der Afilias nur Inhaber von eingetragenen Warenzeichen .info-Domains vormerken sollten. (Anon, 2001)“

Der damalige Erwerb hat zu großen Streitigkeiten mit einzelnen Städten geführt, dennoch sind die info-Domains immer noch im Eigentum von *tiscover*, die die Domains ihren Kunden gratis zur Verfügung stellt.

6.1. Wahl geeigneter Kriterien

Damit die Plattformen untereinander verglichen werden können, müssen geeignete Kriterien definiert werden. Diese wurden aus Sicht eines potentiellen Touristen gewählt und stellen daher alltägliche Kriterien dar, auf den Vergleich von technischen Ansätzen wird bei dem Vergleich verzichtet, da dieser für Benutzer der Angebote keine Relevanz darstellt. Auf folgende drei Kriterien wird bei dem Test eingegangen. Die Gewichtung der einzelnen Kriterien (wie im Vergleich von den MashUp Anwendungsfällen in Kapitel 5.2.4) entfällt, da der kurze Vergleich lediglich die Unterschiede der einzelnen Anbieter darstellen soll. Jede der verglichenen Tourismusportale hat eine Existenzberechtigung und somit ist es unmöglich einen einzigen Gewinner zu küren.

Informationsgehalt für den Benutzer

Die Fülle an (leicht zu findenden) Inhalten stellt die wichtigste Komponente für eine Plattform dar. Dabei wird beim Vergleich auf die Tiefe der Informationen eingegangen und analysiert, wie rasch man verwandte Informationen zu einem Thema finden kann. Speziell für die Auswahl von Reisezielen interessant sind multimediale Inhalte wie Fotos oder Videos zu einzelnen Sehenswürdigkeiten. Daher wird ebenfalls bewertet, wie informativ die angebotenen Informationen für Touristen sind.

Zielgruppenspezifizierung innerhalb der Applikation

Für welche Zielgruppen wurde die Seite entwickelt? Sind Themenkomplexe wie Wochenendurlaube, Urlaubsziele für die Generation 55+ oder für Familien gesondert gekennzeichnet? Sind eigene Informationen für verschiedene Zielgruppen gekennzeichnet, bzw. gesondert abrufbar?

Vorhandene Interaktionsmöglichkeiten

Können die angebotenen Informationen kommentiert/bewertet werden? Kann ein interessierter Besucher bereits vorab Hotel- sowie Kartenreservierungen vornehmen? Besteht die Möglichkeit Reisepläne zusammenzustellen und Informationen in verschiedenen Kontexten (Geographisch, Zeitlich, nach Eintrittspreisen sortiert, ...) darzustellen?

6.2. Anbieteranalyse**Travelwiki.org (<http://www.travelwiki.org/de/Wien>)**

Alle Informationen zur Stadt Wien werden auf einer Webseite präsentiert. Die Unterseiten für die jeweiligen Bezirke sind nach dem gleichen Schema aufgebaut und beinhalten sehr wenige Informationen. Sehenswürdigkeiten werden genannt, allerdings gibt es keine multimedialen Inhalte. Travelwiki stellt den kleinsten Informationsgehalt im Testumfeld dar. Die angebotenen Inhalte sind auf Deutsch verfügbar, die englischsprachige Seite bietet komplett verschiedene Informationen.

Eine Segmentierung der Informationen für Zielgruppen ist nicht erkennbar. Es erfolgt lediglich eine preisliche Unterteilung bei Gastronomie und Unterkünften. Der klassische Web 2.0 Vertreter ist durch seine Offenheit bei Datenänderungen der einzige Vertreter im Feld, der nicht nur Inhalte der wichtigsten Sehenswürdigkeiten anbietet, sondern auch kleine wienerische „Schmankerln“ anbietet. (vgl. Konzept des *Long Tails*)

Auf Grund der Tatsache, dass die Webseite als Wiki konzipiert wurde, können alle Informationen von jedem Besucher geändert/erweitert werden. Eine Kommentarmöglichkeit ist ebenso wie eine Möglichkeit zur Online Reservierung nicht vorhanden. Die Inhalte können nicht nach Kontexten verschieden dargestellt werden.

Tiscover.com (<http://wien.tiscover.com/>)

Die angebotene Informationsvielfalt ist beeindruckend. Es gibt zu allen wichtigen Sehenswürdigkeiten überdurchschnittlich viele Informationen. Neben historischen Informationen gibt es Eintrittspreise, Fotos und selbst virtuelle Rundgänge. Die angebotenen Inhalte sind in übersichtlichen Kategorien hinterlegt, eine Übersicht über die Struktur fehlt jedoch. Als Besucher verirrt man sich rasch in der Tiefe an Unterseiten und ist gezwungen vom Start weg neu anzufangen. Die Seite ist in 2 Sprachen verfügbar.

Es gibt keine Zielgruppenspezifizierung, jedoch bietet tiscovers als einziger Anbieter getrennte Informationen für Winter und Sommertourismus.

Die angebotenen Informationen können nicht kommentiert oder zu Reiseplänen zusammengestellt werden. Eine Online-Reservierung ist sowohl via Weboberfläche als auch Telefon möglich.

Wien.info (<http://www.wien.info/>)

Wie bereits erwartet beinhaltet die Seite des Wien Tourismus die vielseitigste Breite an Informationen rund um die Stadt Wien. Die aktuellen Inhalte werden um Stadtpläne, Fotos und Eventdaten ergänzt. Das Navigationsschema ist im Vergleich zu tiscovers.com weitaus intuitiver, dennoch fällt der Überblick schwer. Die Informationen zu Sehenswürdigkeiten und Events sind stark von den restlichen Inhalten auf der Seite entkoppelt. Ein rasches Navigieren durch verschiedene Bereiche der Webseite ist nicht möglich.

Als einzige Plattform im Vergleich bietet der Wien Tourismus eigene Bereiche für besondere Themenkomplexe. Es finden sich Informationen für Familien, Personen mit Handicap, Hochzeitsgesellschaften und vielem mehr. Dabei fehlt jedoch wie bei den Standardinformationen eine Vernetzung der Inhalte untereinander. Benutzer der Seite müssen also gezielt zwischen allgemeinen und speziellen Informationen wechseln.

Es besteht sowohl die Möglichkeit der Online Reservierung als auch der Erstellung eines eigenen Reiseplanes mit Kommentarmöglichkeit zu den einzelnen Sehenswürdigkeiten, der komfortabel zusammengestellt werden kann um ihn im Anschluss an Freunde zu verschicken. Eine öffentliche Kommentarmöglichkeit fehlt jedoch.

Touristinformr.net (<http://vienna.touristinformr.net/>)

Als Neuling am Markt kann die Eigenentwicklung dennoch wichtige Akzente setzen. Der Mehraufwand bei der Konzeptionierung des semantischen Netzes bringt spürbare Verbesserungen für das Finden von Informationen. touristinformr.net ist der einzige Teilnehmer im Feld, wo kontext-relevante Informationen leicht verständlich dargestellt werden können. Die angebotenen Inhalte hinken der Konkurrenz allerdings weit hinterher. Bis zu einem kommerziellen Marktstart müsste noch sehr viel an redaktioneller Arbeit getan werden, um mit der Konkurrenz mithalten zu können. Bei diesem Thema scheitert die Konvergenzapplikation also. Besucher können die einfache Navigation zwar als hilfreich empfinden, Nutzen wird ihnen allerdings nur die gewünschte Information.

Durch die Hinterlegung der Beziehungen von einzelnen Objekten wäre es in einem weiteren Ausbauschritt sehr einfach möglich Themenkomplexe besonders hervor/abzuheben. Die Auswahl der Darstellungsmerkmale (geographisch, nach Preis, ...) könnte allerdings ebenfalls durch eine konventionelles Plattform angeboten werden.

Durch die Kombination mit Web 2.0 Komponenten wie dem Kommentieren von einzelnen Sehenswürdigkeiten und dem Bereitstellen von RDFa/Mikroformaten hebt sich die Plattform allerdings deutlich von der Konkurrenz ab. Die angebotenen Inhalte können wiederum als Mashup-Quelle dienen und einfach mit anderen Applikationen kombiniert werden.

6.3. Ergebnis des Vergleichs

Die verglichenen Plattformen schneiden sehr unterschiedlich ab und eine genaue Betrachtung zeigt bei allen Teilnehmern deutliche Schwächen. Keine der analysierten Plattformen kann sich vollkommen an die Bedürfnisse der Benutzer anpassen. Bei allen Anbietern ist ein Verändern der Darstellungsoptionen/Kontexte nicht möglich. Obwohl tiscover.com und Wien.info über immenses Datenmaterial verfügen und daher für Touristen die erste Anlaufstelle für Informationen zur Stadt darstellen sollten, gelingt es beiden Anbietern nicht die Inhalte strukturiert und einfach auffindbar darzustellen.

Der Test zeigt, dass vor allem bei der Aufbereitung von Daten der Einsatz von semantischen Technologien große Synergieeffekte bieten kann. Anbieter wie tiscover werden deshalb in Zukunft nicht um den Einsatz von semantischen Technologien herumkommen.

Touristen, die bereits jetzt ihre Urlaube im Internet planen und Informationen vorab sammeln, können also getrost in die Zukunft blicken. Urlaube zu planen wird im Social Semantic Web definitiv einfacher.

7. Zusammenfassung

Das World Wide Web verändert sich. Durch die vielen Web 2.0 Angebote und dem medialen Echo strömen immer mehr (junge) Menschen (vgl. Schwarz, 2006) in Communities und interagieren. Das Web wird durch die vielen Veränderungen, die das „Mitmach-Web“ mit sich bringt, multimedialer und sozialer. Durch die vielen neuen Inhalte, die täglich publiziert werden steigt jedoch ebenfalls das Chaos. Die richtigen Informationen werden immer schwerer auffindbar und daher ist die Entwicklung des Semantic Web positiv zu sehen. Die Anreicherung der Daten und die maschinelle Verarbeitung und Aufbereitung von Informationen wird definitiv das Internet der nächsten Jahrzehnte prägen. Die Kombination dieser beiden Welten ist daher bereits jetzt eine gute Möglichkeit um ein Alleinstellungsmerkmal für Web-Angebote zu schaffen. Dabei sollte jedoch nicht die Technik im Vordergrund stehen, sondern der Mehrwert für die Benutzer. Das Verbinden von Ontologien mit Social Tagging, oder das Verknüpfen von Informationen stellen typische Anwendungsfälle dar, in denen Benutzer Vorteile aus beiden Welten nutzen können. Die Anreicherung von SW-Applikationen um soziale Features wie das Kommentieren bzw. Taggen von Ressourcen stellt einen weit aus geringeren Aufwand dar, als die Implementierung einer eigenen Ontologie, da Web 2.0 Komponenten keinerlei aufwendige technische Rahmenbedingungen unterliegen. In Zukunft werden soziale Dienste jedoch nicht mehr ausreichen, um aus der Fülle von Anbietern hervorzustechen.

Der Vergleich von klassischen Tourismusplattformen zu *touristinformr.net* hat gezeigt, dass semantische Web-Angebote einen klaren Vorteil bei der Vernetzung und Visualisierung von Daten aufweisen können. Eine vollständige Automatisierung des Datenimports ist zum heutigen Zeitpunkt ohne finanzielle Mittel jedoch nicht möglich. Hier scheitert eine Konvergenzapplikation deutlich. Daten, die per XML-Schnittstelle für den Import heranziehbar wären, sind selten und es fehlen notwendige Metainformationen, die nur per RDF(a) abbildbar wären.

Der Import von Daten über XML Schnittstellen (RSS, Atom, RDF, SOAP, ...) stellt jedoch ein Best Practice Szenario dar. Für Benutzer, denen diese Schnittstellen nicht zur Verfügung stehen, bzw. sich nicht mit diesen technischen Feinheiten beschäftigen möchten, stellen Mashup Creator Applications wie Microsoft Popfly oder Yahoo Pipes ebenfalls eine mögliche Alternative dar. Diese Web 2.0 Tools bieten einen allgemeingültigen Ansatz für die Erstellung von Mashups. Der Vergleich in Kapitel 5.2.4 hat gezeigt, dass eine Eigenentwicklung oder der Einsatz von Yahoo Pipes vorzuziehen ist, wenn auf eine Integration in bestehende Applikationen Wert gelegt wird.

8. Diskussion / Blick in die Zukunft

Die zugrunde liegende Konvergenzapplikation stellt ein einfaches, aber dennoch wirkungsvolles Beispiel der Verbindung von Semantic Web und Web 2.0 Inhalten dar. Das Vergleichskapitel hat gezeigt, dass der Einsatz von neuen Technologien deutliche Vorteile, sowohl für die Entwicklung als auch für die Benutzung der Applikation, bringt und der Mehraufwand an Konzeption und Implementierung gerechtfertigt ist. Am Ende der Arbeit blickt der Autor daher über den Tellerrand und hinterfragt, wie das zukünftige Web aussehen könnte.

Bevor nicht große Anbieter den Schritt in das SW wagen und mehr standardisierte Ontologien verfügbar sind, stellen Musteranwendungen Insellösungen dar. Ein automatisierter Abgleich von Datenbeständen ist zwar technologisch möglich, jedoch für professionelle Produktivsysteme ungeeignet, da mit heutigem Datum noch zu wenig semantische Datenbestände vorliegen. Es wird daher auch in naher Zukunft notwendig sein Inhalte redaktionell zu bearbeiten und zu filtern.

Durch die wenigen bisher standardisierten Technologien und Sprachen des Semantic Web sowie den wenigen vernetzten Muster-Applikationen ist der Traum von maschinenlesbaren Informationen schon in greifbarer Nähe – dennoch wird noch viel Arbeit notwendig sein, um den Sprung in ein Web 3.0 oder Social Semantic Web zu schaffen, in dem maschinelle Dienste, Software Agenten und soziale Communities koexistieren. Konvergenzapplikationen zeigen jedoch bereits heute, wie die Welt von morgen aussehen könnte.

8.1. Wirtschaftliche Betrachtung

Bis neue technische Entwicklungen von Firmen tatsächlich akzeptiert und angenommen werden, vergeht oft eine sehr lange Zeit. Zu groß ist die Skepsis und die Gefahr, dass Geld in fadenscheinige Technologietrends investiert wird. Die Frage, die nach dem Vergleich der Konvergenz- mit der konventionellen Applikation übrig bleibt, ist die der gesamtheitliche, wirtschaftlichen Betrachtung der Thematik. Informatiker tendieren oft „*der Technik wegen*“ Projekte zu starten, ohne auf den finanziellen (und oft auch sozialen) Hintergrund zu blicken.

Durch die hohe Partizipation von potentiellen Kunden in Blogs, Foren und Bewertungsplattformen sind bereits jetzt viele Firmen ohne deren Zustimmung im Internet vertreten. Es ist also nicht die Frage, ob Firmen den Einstieg in das Mitmach-Web wagen sollen, sondern viel mehr, wie sie planen, diese Plattform für sich zu nutzen.

„In der Vergangenheit (Web 1.0) konnte sich der Gast nur nach den Inhalten des Anbieters informieren. Heute versorgen sich Hotelsuchende gegenseitig mit Hinweisen, Empfehlungen aber auch Kritik. [...] Mehr als die Hälfte der Internet-User entscheiden sich eher für den Kauf, wenn sie positive Rückmeldungen über ein Produkt bzw. Dienstleistung im Web gelesen haben. Ca. 34% der Internet-User entscheiden sich gegen einen Kauf, wenn sie negative Kommentare über das Produkt/Dienstleistung gelesen haben.[...]“ (Reisenzahn, 2007)

Eine Beteiligung am Web 2.0 ist daher für jedes Unternehmen sinnvoll und notwendig, um in Zukunft gegen die Konkurrenz bestehen zu können. Die Transparenz, die durch die vielen Rückmeldungen der Kunden entsteht, kann jedoch zu Problemen für Firmen führen. (Gefälschte) negative Nachrichten über Produkte oder Dienstleistungen führen zu Umsatzeinbußen und Imageverlust. Es muss daher genau darauf geachtet werden, wie sich die Reputation des eigenen Unternehmens entwickelt.

„[...] Diese Zahlen zeigen aber auch die mit dieser Entwicklung verbundenen Konsequenzen und Gefahren auf. Die meisten Einträge auf Websites mit Hotel- oder Restaurantbewertungen sind zwar ernsthafte Erfahrungsberichte. Wer erkennt und bereinigt jedoch irreführende negative Einträge, welche nicht zwingend auf Erfahrungswerten beruhen? Oftmals missbrauchen User dieses Tool, um absichtlich negative Mundpropaganda zu verbreiten und die Kaufentscheidung interessierter Urlauber negativ zu beeinflussen.“ (ebenda)

Die Präsenz auf verschiedenen Web-Plattformen ist für Unternehmen eine große Chance, benötigt allerdings sehr viel Zeit und Aufmerksamkeit.

Da der Einsatz von semantischen Technologien zum Zeitpunkt des Verfassens dieser Arbeit noch nicht für die breite Masse zur Verfügung steht, können diese Technologien daher nur für das eigens gewählte Beispiel der Tourismus-Plattform beurteilt werden. Die Analyse fällt jedoch nicht einfach aus. Vor allem die Kombination von Ontologien in Kombination mit automatisierten Datenimporten muss unter mehreren Betrachtungswinkeln analysiert werden.

Kundensicht

In erster Linie soll der Einsatz von neuen Technologien für den Kunden (=Besucher der Webseite) Vorteile bringen. Zum aktuellen Zeitpunkt kann dieser Vorteil nur durch veränderte Navigationsstrukturen und dem Anbieten von vernetzten Informationen erfolgen. Das Ziel der Semantic Web Initiative, die Daten über mehrere Web-Portale zu verbinden ist noch in weiter Ferne. Aktuelle Browser bieten momentan weder Support für Microformate und RDFa, noch können Sie externe RDF Daten plausibel verarbeiten. Intelligente Software Agenten für das Semantic Web müssen erst entwickelt werden und der breiten Masse zur Verfügung stehen, ehe der Durchbruch der semantischen Idee stattfinden kann. Bis dahin muss analysiert werden, ob einfachere Navigationsstrukturen und beziehungsrelevante Informationen zu einzelnen Objekten der Seiten ausreichen, um Kunden eines Web-Angebots langfristig zufriedenzustellen bzw. neue Kunden anzulocken.

Die Aggregation der verschiedenen Datenanbieter (Fotomaterial, Veranstaltungshinweise, Kartografie, ...) stellt definitiv einen enormen Mehrwert für Kunden dar. Konventionelle Angebote, die sich auf ein Themengebiet wie Veranstaltungshinweise fokussieren, erreichen naturgemäß eine viel kleinere Zielgruppe. Je mehr sinnvollen Inhalt eine Webseite hat, desto mehr Kunden können dadurch angelockt werden. Die einzige Gefahr hierbei ist, dass die Datenflut zu groß und unübersichtlich wird. Da alle besprochenen Technologien auf keine proprietären Formate setzen und teilweise auf der REST Architektur (vgl. Kapitel 5.3) basieren, stellen heutige Konvergenzapplikation keine zusätzlichen Barrieren für Benutzer dar. Aus Sicht der Kunden eines Webangebots können geplante Adaptierungen in Richtung Semantic Web und Web 2.0 nur positiv gesehen werden.

Anbietersicht

Das Betreiben von semantischen Applikationen stellt einen erheblichen initialen Mehraufwand dar. Es muss für das Themengebiet eine Ontologie entwickelt, bzw. auf die eigenen Bedürfnisse adaptiert werden. Dieser Vorgang erfordert Know-How der Mitarbeiter im Bereich Modellierung und Wissensmanagement. Es gibt nur wenige Anbieter von Standardsoftware in diesem Bereich. Eigenentwicklungen setzen ein hohes Maß an Programmierkenntnissen voraus. Ist das semantische Netz jedoch fertig, rechtfertigt der Einsatz der Technik die entstandenen Kosten. Die zugrunde liegenden Daten können rasch miteinander in Beziehung gebracht werden. Durch die Verknüpfung der Daten kann aus implizitem Wissen sehr rasch explizites Wissen gewonnen werden. Zusammenhänge die in (un)strukturierten Datenbeständen nicht abbildbar waren, können durch simple Transformierungen dargestellt werden.

Sobald eine größere Verbreitung von semantischen Applikationen gegeben ist und der Semantic Web Stack in seiner endgültigen Form verabschiedet wird, ist eine Vereinfachung beim Einsatz von semantischen Technologien prognostizierbar. Der Know-How Vorsprung, der bereits jetzt erlangt werden kann, wird allerdings nur schwer aufzuholen sein.

Der Einsatz von automatisierten Web 2.0 Inhalten ermöglicht die rasche Anreicherung der Daten mit zusätzlichen Informationen. Hierbei ist kaum mehr manueller Eingriff notwendig. Die Applikation ist laufend am aktuellsten Stand. Am Beispiel der Tourismusplattform ist erkennbar, dass für den tatsächlichen Betrieb keine Redakteure notwendig wären, wenn genügend verwendbare Datenquellen vorhanden wären. Momentan ist der Einsatz von gänzlich automatisiertem Datenimport jedoch nicht ratsam.

Die Vorgehensweise der Content-Partnerschaften stellt allerdings eines der größten Problemfelder zukünftiger Applikationen dar. Die Abhängigkeit zu anderen Anbietern ist enorm hoch. Betreiber müssen sich daher überlegen, wie externe Daten behandelt werden. Müssen die angebotenen Inhalte zwischengespeichert werden, um etwaige Schnittstellenprobleme temporär abfangen zu können? Was passiert, wenn Anbieter die lose Kooperation aufheben? Für ausfallssichere Lösungen muss ein erhöhter (finanzieller) Aufwand betrieben werden, um automatisierte Verfahren garantiert aufrechterhalten zu können. Das gezeigte Szenario der Datenaggregation am Beispiel von Kapitel 5.2 ist daher in der Praxis nicht 1:1 wiederverwendbar.

9. Glossar

Data Web	Siehe Semantic Web
MashUp	Ein MashUp ist die Kombination von zwei oder mehreren Datenquellen zu einer neuen Applikation, die oftmals eine völlig neue Betrachtungsweise von bestehenden Daten erlaubt. Oftmals werden Inhalte um Geodaten (zB Google Maps) angereichert.
Ontologie	Eine Ontologie ist eine Spezifizierung von Elementen und deren Beziehung in einer bestimmten Domäne (Anwendungsgebiet).
Semantic Web	Das Semantic Web (SW) stellt eine Erweiterung des bisherigen World Wide Web dar, in dem mit Hilfe von Metadaten Inhalte für Maschinen interpretierbar gemacht werden. Ein Synonym dafür ist „Data Web“.
Semantic Web Stack	Der Semantic Web Stack stellt das technische Rahmenwerk des SW dar. Auf abstrakter Ebene wird das Zusammenspiel der verwendeten Techniken dokumentiert.
Social Semantic Web	Siehe Web 3.0
SPARQL	SPARQL ist aus aktueller Sicht die Abfragesprache von RDF-Graphen im SW. Die Syntax ist ähnlich zu der Sprache SQL, die Daten in relationalen Systemen abfragt.
Unicode	Unicode stellt einen Standard für Schriftzeichen dar, der in seiner aktuellen Fassung 99.000 Zeichen beinhaltet. Der Standard ist Grundlage für die interoperable Kommunikation von Systemen.
URI	Uniform Resource Indicator sind eindeutige Namen im World Wide Web. Die Adresse http://johannes.nagl.name/ stellt einen eindeutigen Pfad für die Entität „Johannes Nagl“ dar, der überall im Internet benutzt werden kann.
W3C	Das World Wide Web Consortium (kurz: W3C) stellt die Normierungsbehörde des Internets dar. Geleitet wird sie von Tim

	Berners-Lee, dem „Erfinder“ des heutigen Webs und dem Vordenker des → SW.
Web 2.0	Das Web 2.0 positionierte sich in den letzten Jahren als ein Sinnbild für die Demokratisierung des bestehenden Web. Die Benutzer von Diensten stehen im Vordergrund und durch das gemeinsame Partizipieren auf Webplattformen (zB Wikipedia oder YouTube) entstehen neue Dienste. Ein Synonym dafür ist „Mitmachweb“.
Web 3.0	Das Web 3.0 ist die Verbindung von Semantic Web und Web 2.0. Der Begriff ist nicht allgemeingültig und die Bedeutung ändert sich sehr häufig.

10. Abbildungsverzeichnis

Abbildung 1: Popularität der Begriffe Web 2.0, Web 3.0, SW	6
Abbildung 2: Semantic Web Stack (2005).....	11
Abbildung 3: Konvergenzen zwischen SW und Web 2.0.....	14
Abbildung 4: Aufbau der Beispielapplikation touristinformr.net.....	16
Abbildung 5: Grafische Darstellung des semantischen Netzes (Screenshot)	19
Abbildung 6: Entwicklungsumgebung Yahoo Pipes	21
Abbildung 7: Möglicher Output Yahoo Pipes	21
Abbildung 8: Entwicklungsumgebung MS Popfly	21
Abbildung 9: Eigenschaften eines Blocks MS Popfly.....	21
Abbildung 10: Aufbau der Microformats	25

11. Abkürzungsverzeichnis

ERP	Enterprise Resource Planning
HTML	Hypertext Markup Language
JSON	JavaScript Object Notation
RDF	Resource Description Framework
RDFa	<i>Keine Langschreibweise auffindbar.</i>
REST	Representation State Transfer
RSS	Really Simple Syndication
SOAP	Simple Object Access Protocol
SW	Semantic Web
TLD	Top Level Domain
URI	Uniform Resource Indicator
W3C	World Wide Web Consortium
XLST	Extensible Stylesheet Language Transformation
XML	Extensible Markup Language

12. Literaturverzeichnis

- ADIDA, B. & BIRBECK, M., 2007. RDFa Primer 1.0 [Online]., Verfügbar bei: <http://www.w3.org/TR/xhtml-rdfa-primer/> [17.07.2007]
- ALBY, T. 2006. Web 2.0 Konzepte, Anwendungen, Technologien. München: Carl Hanser Verlag
- ANON, 2001. Tourismuskonzern reserviert rund 5.000 .info-Domains [Online]., Verfügbar bei: <http://www.heise.de/newsticker/meldung/21652> [25.07.2007]
- ANON, 2007. About Microformats [Online]. Verfügbar bei: <http://microformats.org/about> [17.07.2007]
- CUBRILOVIC, N. Microsoft Launches Popfly: Mashup App Creator Built On Silverlight [Online]., Verfügbar bei: <http://www.techcrunch.com/2007/05/18/microsoft-launches-popfly-mashup-app-creator-built-on-silverlight/> [17.07.2007]
- BERNERS-LEE, T. & HENDLER, J. & LASSILA, O. 2001. The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, 284 (5)
- BERNERS-LEE, T. 2006. developerWorks Interviews: Tim Berners-Lee / Transcript. [Online]. Verfügbar bei: <http://www-128.ibm.com/developerworks/podcast/dwi/cm-int082206.txt> [11.01.2007]
- GARSHOL, L. 2002. What are Topic Maps? [Online]., Verfügbar bei: <http://www.xml.com/pub/a/2002/09/11/topicmaps.html> [23.07.2007]
- GROSSMANN, L. Time's Person of the Year: YOU [Online]. Verfügbar bei: <http://www.time.com/time/magazine/article/0,9171,1569514,00.html> [20.07.2007]
- KING, R. 2007. Q&A with Tim Berners-Lee [Online]., Verfügbar bei: http://www.businessweek.com/technology/content/apr2007/tc20070409_961951.htm [22.07.2007]
- NAGL, J. 2005. Das Semantische Web - Aktuelle Umsetzungen. Seminararbeit. FH Technikum Wien, unveröffentlicht.
- O'REILLY, T. 2005a. Web 2.0: Compact Definition? [Online]., *O'Reilly Radar*. Verfügbar bei: http://radar.oreilly.com/archives/2005/10/web_20_compact_definition.html [11.01.2007]
- PELLIGRINI, T. 2007. Sebastian Schaffert: "Ein Semantic Web, das soziale Aspekte außer Acht lässt, ist aus unserer Sicht nicht zielführend." [Online]., Verfügbar bei: <http://www.semantic-web.at/10.36.188.article.sebastian-schaffert-ein-semantic-web-das-soziale-aspekte-ausser-acht-laesst-ist-aus-unsere.htm> [10.07.2007]
- REISENZAHN, T. 2007. Mit Web 2.0 in ein neues touristisches Zeitalter. [Online]. Verfügbar bei: http://www.ots.at/presseaussendung.php?schluessel=OTS_20070129_OTS0191&ch=technologie [15.06.2007]

RÖTGERS, J. 2007. Interview mit dem Chefentwickler von Yahoo Pipes. In: Ö1 Matrix. On Air am: 20. Mai 2007

SCHWARZ, T. 2006. Die beliebtesten Web 2.0 Portale nach Alter [Online]., Verfügbar bei: <http://www.marketing-boerse.de/Fachartikel/details/Web20-Nutzer> [12.07.2007]

SOERGEL, J & MUNZ, S. 2007. Agile Produktentwicklung im Neuen Web. Diplomarbeit (Diplom-Ingenieur/in für Druck- und Medientechnik). Fachhochschule Berlin.

UPDEGROVE, A. 2005. The Semantic Web: An interview with Tim Berners-Lee [Online]., Verfügbar bei: <http://www.consortiuminfo.org/bulletins/semanticweb.php> [12.07.2007]

WONG, W. 2007. Microformats: People First, Machines Second. [Online]. Verfügbar bei: <http://www.elecdesign.com/Articles/Index.cfm?AD=1&ArticleID=15742> [17.07.2007]

13. Anhang

13.1. Struktur des semantischen Netzes (Auszug):

- Berühmte Persönlichkeit
 - Historie
 - Kunst
 - Malerei
 - Musik
 - Schauspiel
 - Literatur
 - Politik
 - Sport
 - Wissenschaft
- Freizeit & Kultur
 - Lokales Event
 - Sehenswürdigkeit
 - Denkmal
 - Kirche
 - Museum
 - Sonstiges
 - Theater
 - Vergnügen
 - Nachtleben
 - Bar
 - Casino
 - Club
 - Kino
 - Wellness
- Infrastruktur
 - Unterkunft
 - Camping Platz
 - Ferienhaus
 - Herberge
 - Hotel
 - 2 * Hotel
 - 3 * Hotel
 - 4 * Hotel
 - 5 * Hotel
 - 6 * Hotel
 - Familienhotel
 - Pension
 - Verkehr
 - Flughafen
 - Individualverkehr
 - Öffentlicher Verkehr
 - Wichtige Gebäude
 - Botschaft
 - Bankinstitut
 - Post
- Gastronomie
 - Kaffeehaus
 - Restaurant
- Shopping
 - Einkaufsstraße
 - Einkaufszentrum
 - Geschäft