

OWL und das Semantische Web

OWL ist ein wichtiger Schritt in Richtung semantischem Web

71desa1bif

Hochschule für Technik Stuttgart

5. November 2008

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen und Entwicklung	2
2.1	Problem	2
2.2	Begriffe	2
2.3	Tim Berners-Lee	3
2.4	Bestrebungen und Ziele	4
2.5	Lösung	4
3	OWL	5
3.1	Allgemein	5
3.2	Herkunft	5
3.3	Untersprachen	6
4	Anwendungen	7
4.1	Semantische Suche und Suchmaschinen	7
4.2	Semantic MediaWiki	8
4.3	Semantische Webservices mit OWL-S	10
4.4	Anwendungen und Prototypen	11
4.5	Editoren und Repositorien	11
5	Zusammenfassung	12

1 Einleitung

Internet, 2008, Suchanfrage: [eine noch zu definierende Suchanfrage].

Resultate:

1. ...
2. ...
3. ...
4. ...
5. ...

Bei komplexen Suchanfragen kommt es oft zu unbrauchbaren Suchergebnissen. Durch raffiniertes Setzen von Schlagworten oder durch Einsatz von Geldmitteln (z.B. Google Adwords) enthalten die Ergebnisse oft unangeforderte Werbebotschaften, die ohne Relevanz für die Suchenden sind. Die Suchenden müssen Teilergebnisse durch Eingrenzungen erzielen und diese dann zusammensetzen. Es ist wünschenswert die Qualität der Suchergebnisse zu verbessern.

Um bei komplexeren Suchanfragen - wie dem einleitende Beispiel - eine sinnvolle Ergebnismenge zu erhalten sind für den Menschen mehrere Arbeitsschritte nötig. Suchmaschinenbetreiber investieren zwar viel Zeit und Geld unterschiedliche Methoden zu implementieren, die dem Suchenden ein hohes Maß an Ergebnisqualität liefern, jedoch sind die Suchmaschinen nicht in der Lage die Zusammenhänge zwischen Informationen zu erkennen.

Damit eine rechnergestützte Informationsbeschaffung erfolgen kann müssen einerseits die Fragen der Suchenden, andererseits die vorhandenen Informationen und deren Zusammenhänge maschinell auswertbar sein. So kann eine Ergebnisqualität erreicht werden die der Zusammenstellung von Teilergebnissen durch einen Menschen entsprechen. Dies wird erreicht indem die Informationen mit zusätzlichen semantischen Informationen versehen werden.

Durch die zusätzliche Auswertung dieser semantischen Informationen bei einer Suchanfrage wird nicht nur die Ergebnisqualität verbessert, sondern es ist auch möglich - ähnlich logischer Programmiersprachen - über Inferenzen implizites Wissen zu extrahieren und zu nutzen.

Im Rahmen dieser Ausarbeitung wird gezeigt, dass es sich bei dem Einsatz der Web Ontology Language (OWL) zur Beschreibung der semantischen Beziehungen zwischen Informationen um einen vielversprechenden Schritt in Richtung semantisches Web handelt.

Hierfür werden in Kapitel 2 die vorhandenen Probleme aufgezeigt und wenige zum Verständnis wichtige Begriffe erklärt. Die Vision des semantischen Webs stammt vom Erfinder des Internets Tim Berners-Lee selbst. Es werden die Ziele und die aktuellen Bestrebungen zur Erreichung dieser Ziele aufgezeigt. Im letzten Teil von Kapitel 2 werden Lösungsansätze aufgezeigt. Kapitel 3 widmet sich OWL, deren Herkunft und verfügbaren Untersprachen. Kapitel 4 zeigt einige

semantische Anwendungen auf, speziell semantische Suchmaschinen und exemplarisch den Ansatz des Semantic MediaWiki. Im letzten Kapitel werden einige bereits implementierte semantische Web Anwendungen angesprochen. Besonders spannend ist der Ansatz einer semantischen Webservice Landschaft, also einem Netz von Rechnern, die sich beginnen zu verstehen. Zuletzt wird eine Auswahl von Ontologie Editoren und Ontologie Repositorien aufgelistet.

2 Grundlagen und Entwicklung

2.1 Problem

Die Suche im Internet hat in viele Lebensbereiche Einzug gehalten. Aktuelle Börsenkurse, Fußballergebnisse oder die Suche nach einem günstigen Preis für ein Produkt. Über 90 Prozent der Deutschen nutzten 2007 das Internet um sich zu informieren¹.

Bei einer Suche ist die gewünschte Information zwar meist verfügbar, in der Menge der zu durchsuchenden Information ist sie jedoch oft nur schwer zu finden.

Das liegt daran, dass die Speicherung der Informationen im Internet auf den Menschen als Konsumenten ausgerichtet ist. Somit ist die Information garnicht oder nur schwer maschinell verarbeitbar. Ein weiterer Grund ist, dass die Suchmaschinen zwar die Suchbegriffe in den vorhandenen Informationen lokalisieren, jedoch nicht in der Lage sind den Sinn einer Anfrage zu erfassen. Leider werden die Suchergebnisse oft noch durch die bezahlte Platzierung von Werbebotschaften oder durch die raffinierte Ausstattung von Werbeseiten mit Suchbegriffen verfälscht².

Ein weiteres Problem ist dem Internet inhärent. Die "dezentrale Struktur und Organisation des Web [bringt] zwangsläufig eine Heterogenität der vorhandenen Informationen auf ganz verschiedenen Ebenen mit sich: [von] [...] unterschiedlichen Dateiformaten [...] Kodierungstechniken [...] verschiedenen [...] natürlichen Sprachen bis hin zum ganz unterschiedlichen Aufbau privater Homepages"³.

Zunächst sollen jedoch wenige Begriffe die zum Verständnis von OWL und dem semantischen Web notwendig sind erläutert werden.

2.2 Begriffe

Eine Hierarchie ist ein System von Elementen, bei dem eine Ordnungsrelation entweder einen Baum (Monohierarchie) oder einen gerichteten azyklischen Graphen (Polyhierarchie) definiert. Diese Einteilung in eine Hierarchie (ähnlich dem umgangssprachlichen Gebrauch) impliziert meist schon eine gewisse Wertigkeit der Elemente untereinander⁴.

¹vgl. [EIFR2007], Seite 367

²vgl. [HKR+2008], Seite 10

³[HKR+2008], Seite 10

⁴vgl. [WIKIa]

Eine Taxonomie ist eine „Begriffshierarchie mit festgelegter Syntax. [Die] Semantik beschränkt sich auf Generalisierungs-Spezialisierungs-Hierarchie[n]“⁵. Taxonomie in der Informatik wird beschrieben als „[...] Bezug auf Dokumente bzw. Inhalte [...] für ein Klassifikationssystem, eine Systematik oder den Vorgang des Klassifizierens [...]“⁶.

Eine Ontologie hingegen ist eine „[...] Taxonomie mit erweiterter Semantik (verschiedene Relationstypen, erhöhter Vernetzungsgrad der Daten zu Modell)“⁷.

Das Informatiklexikon der Gesellschaft für Informatik beschreibt „[...] eine Ontologie also einen Wissensbereich (knowledge domain) mit Hilfe einer standardisierenden Terminologie sowie Beziehungen und ggf. Ableitungsregeln zwischen den dort definierten Begriffen. Das gemeinsame Vokabular ist in der Regel in Form einer Taxonomie gegeben, die als Ausgangselemente (modelling primitives) Klassen, Relationen, Funktionen und Axiome enthält“⁸. Man kann sagen eine Ontologie beschreibt Objekte und die bestehenden Beziehungen zwischen den Objekten.

Das maschinelle erkennen und auswerten von Beziehungen zwischen Informationen ist der Grundgedanke des Semantischen Webs. Tim Berners-Lee, der Erfinder des Internets, gilt als Vordenker auf dem Gebiet des semantischen Webs.

2.3 Tim Berners-Lee

Sir Timothy John Berners-Lee ist der Erfinder des Internets am CERN. Er gründete 1994 das World Wide Web Consortium und ist dessen Direktor. Er ist Professor für Computer Science am MIT und trägt diverse Ehrendoktorwürden.

Bereits 1998 veröffentlichte Tim Berners-Lee die Idee vom Semantic Web. Seither werden Bestrebungen in diese Richtung von vielen Forschern und Industriepartnern durchgeführt.

Seit der Empfehlung des W3C von SPARQL steht nun auch die formalisierte Möglichkeit zur Formulierung von semantischen Suchanfragen zur Verfügung. Tim Berners-Lee begrüßt diese Empfehlung ausdrücklich. Es ist möglich über mehrere Datenquellen eine Suchanfrage zu starten. Diese Zukunft der Suche wird zu keiner Ablösung der klassischen Textsuchmaschinen führen, eher zu einer Koexistenz zwischen textuell- und datenorientierter Suche.

Die Suche im semantischen Web wird nichtmehr auf Dokumente ausgeführt, sondern auf Daten. Wie das Internet mittelfristig aussieht ist noch unklar und eine sehr spannende Frage. Im Wesentlichen müssen Standards definiert werden auf deren Basis Software-Entwicklungen durchgeführt werden können.

Wie fast immer im Bereich des Internets kümmert sich bei den Bestrebungen zur Verwirklichung des semantischen Webs das World Wide Web Consortium um die Harmonisierung der einzelnen Entwicklungen und der Empfehlung von einheitlichen Standards.

⁵[SCHO200X], Folie 6

⁶[WIKIb], #Taxonomie_in_der_IT

⁷[SCHO200X], Folie 6

⁸[HESS2002]

XXX

Siehe bzw. Quelledetails:

http://server02.is.uni-sb.de/courses/wiki/Semantic_Web

http://de.wikipedia.org/wiki/Tim_Berners-Lee

<http://dig.csail.mit.edu/breadcrumbs/blog/4> (tim bl blog)

<http://www.w3.org/TR/rdf-sparql-query/>

XXX

2.4 Bestrebungen und Ziele

Die W3C Semantic Web Activity will ein Programmiergerüst schaffen, mit dem es möglich ist Informationen gemeinsam über Grenzen (Firmen, Anwendungen) hinweg zu nutzen⁹. Es geht darum Informationen zu verbinden (linked data).

Für den Nutzer wünschenswert sind gute Suchergebnisse. Gut heisst, das zu finden was gesucht wird. Hierfür müssen die Informationen maschinell verarbeitet und in Bezug zueinander gesetzt werden können. Nicht nur auf Basis einer Lokalisierung von Teilstrings sondern auf der erweiterten Basis von Informationen die miteinander verknüpft sind.

Durch diese Verknüpfungen der Informationen ist es auch möglich implizites Wissen nutzbar zu machen.

Die Ziele sind:

- Gute Suchergebnisse
- Semantisches Web (d.h. Informationen und deren Zusammenhänge können von Computern ausgewertet werden)
- glückliche Entdeckungen (Serendipität)
- Maschinenverarbeitung
- Erweiterung Key/Value Konzept zu S/P/O („Key:=Value“ soll zu „Subjekt_in_bestimmter_Relation_zu_Objekt“ werden)

Was wird nun genau benötigt um diese Ziele zu erreichen. Auf welche bestehenden Standards kann man sich zurückziehen und an welchen Stellen gibt es zur Abbildung der Anforderungen noch Defizite. Dazu im folgenden Unterkapitel Lösungen.

2.5 Lösung

Um eine weltweite Interoperabilität zu gewährleisten ist es wichtig offene Standards zu nutzen. Verkomplizierungen der Abläufe durch notwendige Lizenzierungen bzw. der Nutzung unklarer proprietärer Formate verzögern oder verhindern die kontinuierliche Entstehung eines semantischen Webs.

⁹vgl. [W3SW2008]

Ein wichtiger offener Standard ist XML. XML ermöglicht einen durch Computer verarbeitbaren Austausch strukturierter Informationen. XML liefert auch den syntaktischen Rahmen der Ontologiesprachen RDF und OWL¹⁰.

Bereits mit RDF ist es möglich Verknüpfungen zwischen Informationen zu hinterlegen. Jedoch sind die beschränkten Ausdrucksmittel nicht ausreichend um komplexere Zusammenhänge darzustellen. Hierfür kann die Ontologiesprache OWL genutzt werden, die auf der Prädikatenlogik erster Stufe basiert (d.h. es wird „nur über Individuen quantifiziert, in der Prädikatenlogik höherer Stufen auch über Prädikate.“¹¹).

Im Folgenden soll nun die Ontologiesprache OWL genauer betrachtet werden.

3 OWL

3.1 Allgemein

Die Web Ontology Language (OWL) ist eine der Empfehlungen des W3C zur Verwirklichung des Semantischen Webs¹².

Mit OWL ist es möglich Elemente und die Beziehungen der Elemente untereinander eines bestimmten Wissensbereichs mit formalen Mitteln zu beschreiben. Man nutzt OWL um Ontologien zu formulieren. Es werden zusätzlich „Sprachkonstrukte eingeführt, die es erlauben, Ausdrücke ähnlich der Prädikatenlogik zu formulieren“¹³.

3.2 Herkunft

Die nun folgende Beschreibung der Herkunft von OWL und der weiteren genutzten Technologien orientiert sich am OWL Overview des W3C¹⁴.

Durch die Nutzung von XML ist es möglich wohlgeformte und strukturierte Dokumente zu erzeugen und diese ohne Verlust von Information auszutauschen. XML Dokumente enthalten per se jedoch keine Informationen über die Bedeutung der im XML Dokument enthaltenen Informationen.

XML Schemata ermöglichen es die Struktur eines XML Dokumentes zu erzwingen. Desweiteren können Datentypen gefordert werden. Die Festlegung eigener Datentypen ist ebenfalls möglich.

Mit RDF entstand ein Datenmodell für Elemente und für die Beziehungen der Elemente untereinander. RDF ermöglicht einfache semantische Informationen. Es ist möglich RDF Datenmodelle in XML darzustellen.

RDF Schemata legen ein Vokabular für einen speziellen Wissensbereich fest. Auch können die in diesem Wissensbereich vorkommenden Elemente, deren Eigenschaften und Beziehungen untereinander formalisiert dargestellt werden. Die

¹⁰vgl. [HKR+2008], Seite 31

¹¹[WIKIc], #System_der_Pr.C3.A4dikatenlogik

¹²vgl. [W3OW2004a], Kapitel 1.2

¹³[WIKId]

¹⁴vgl. [W3OW2004a]

semantischen Möglichkeiten eignen sich für Generalisierungen dieser Eigenschaften und Beziehungen¹⁵.

Bei DAML+OIL handelt es sich um einen Ansatz - aufbauend auf XML und RDF - eine Beschreibungssprache für Ontologien zu erhalten. Dieser Ansatz wurde seit etwa 2001 nichtmehr weiter verfolgt. Die Web Ontology Working Group des W3C empfiehlt OWL (als direkter Nachfolger von DAML+OIL) als die Ontologiesprache fuer das Semantische Web. OWL enthält die 'lessons learned' des DAML+OIL Ansatzes. So finden sich viele Namen der DAML+OIL Autoren auch auf der Autorenliste von OWL¹⁶.

OWL erweitert das Vokabular mit dem die Eigenschaften und Beziehungen zwischen Elementen eines Wissensbereiches beschrieben werden können. Zusätzliche Möglichkeiten sind die Nutzung von z.B. Kardinalitäten, Gleichheit und Merkmalen von Eigenschaften. Es können Ausdrücke ähnlich der Prädikatenlogik formuliert werden.

3.3 Untersprachen

Die OWL existiert in drei Untersprachen und damit verbundenen Ausdrucksmöglichkeiten. Die folgende Zusammenfassung stützt sich auf die in der Fußnote genannten Quellen¹⁷.

Diese Untersprachen sind OWL Full, OWL DL und OWL Lite.

OWL Lite kann genutzt werden um einfach Beschränkungen und eine Hierarchie darzustellen. OWL Lite wurde hauptsächlich mit dem Ziel geschaffen die Implementierung einfach zu gestalten. OWL Lite ist eine Teilsprache von OWL DL bzw. OWL Full. Sie hat die kleinste Ausdrucksstärke der drei Untersprachen, ist aber immer entscheidbar. Die Komplexität im schlimmsten Fall ist ExpTime. Man kann OWL Lite als eine Erweiterung eines eingeschränkten RDF verstehen.

OWL DL hat die maximal mögliche Ausdrucksstärke bei der zum einen die Berechenbarkeit der Anfragen und zum anderen die Berechenbarkeit der Anfragen in endlicher Zeit garantiert werden kann. OWL DL enthält OWL Lite und ist eine Teilsprache von OWL Full. OWL DL enthält alle OWL Sprachkonstrukte, jedoch teilweise mit Einschränkungen (beispielsweise kann eine Klasse eine Unterklasse verschiedener anderer Klassen sein, eine Klasse kann jedoch keine Instanz einer anderen Klasse sein). Auch OWL DL ist immer entscheidbar, jedoch ist hier die Komplexität im schlimmsten Falle »NExtTime«. OWL DL wird von den momentan gängigen Softwarewerkzeugen fast vollständig unterstützt. DL steht für »description logic« (Beschreibungslogik) »SHOIN(D)«. Diese ist „zu einer entscheidbaren Untermenge der Prädikatenlogik erster Stufe äquivalent“¹⁸. Auch OWL DL kann man als eine Erweiterung eines eingeschränkten RDF verstehen.

¹⁵vgl. [WIKIe], [W3RD2004]

¹⁶vgl. [DAML2002]

¹⁷vgl. [WIKId], [W3OW2004a], [HHL+2008a]

¹⁸[WIKId], #OWL_DL

OWL Full ist die ausdrucksstärkste Ontologiesprache. Sie enthält sowohl OWL Lite als auch OWL DL. Sie enthält als einzige OWL Untersprache ganz RDF(S). Die Berechenbarkeit und Entscheidbarkeit ist nicht gewährleistet. Auch enthält die Semantik einige Aspekte, die aus logischem Blickwinkel problematisch sind¹⁹.

Aktuelle Softwarewerkzeuge unterstützen OWL Full nur bedingt. OWL Full kann man als eine Erweiterung von RDF sehen.

Jedes OWL Dokument ist ein RDF Dokument. Jedes RDF Dokument ist ein gültiges OWL Full Dokument, jedoch nicht immer ein gültiges OWL Lite bzw. OWL DL Dokument.

Für die Auswahl der OWL Untersprache für die eigene Nutzung ist Ziel und die damit verbundene benötigte Ausdrucksstärke der Ontologiesprache der geplanten Anwendung entscheidend. Zunächst ist die gewünschte Ausdrucksstärke wichtig. Die Ausdrucksstärke von OWL Lite wird in den wenigsten Fällen ausreichen. Für die Wahl zwischen OWL DL und OWL Full hängt es davon ab, inwieweit die Möglichkeiten der Meta-Modellierung von RDF-S für die Anwendung genutzt werden sollen (z.B. an Klassen weitere Eigenschaften anzufügen, Klassen von Klassen definieren). Sind diese Möglichkeiten nicht notwendig reicht OWL DL aus. Ein Vorteil bei der Nutzung von OWL DL ist noch, dass Folgerungen/Entscheidbarkeiten in endlicher Zeit garantiert sind²⁰.

4 Anwendungen

4.1 Semantische Suche und Suchmaschinen

Um die Qualität der Suchergebnisse zu verbessern nutzt die semantische Suche statt statistischen Mitteln oder der Anzahl von Querverweisen oder Nennungen bei der Auswahl der Ergebnisse die durch OWL/RDF hinterlegten Beziehungen zwischen den vorhandenen Informationen.

Die mögliche Doppeldeutigkeit von Begriffen (ist ein Schimmel ein weißes Pferd oder eine Struktur an einem unverschlossenen Marmeladenglas) macht es notwendig Bedeutungen in einzelne Wissensbereiche unterschiedlich hinterlegen zu können. Auch kann ohne Kontext eine Aussage nicht ausgewertet werden (wenn man eine Frau mit einem Fernglas sieht - wer hat nun das Fernglas in der Hand ?).

Die W3C Empfehlung zur Durchführung von Anfragen an RDF Daten ist SPARQL. Mit SPARQL sind Anfragen an unterschiedliche Datenquellen möglich. Hierbei können Wertabfragen, Mengenoperationen und Einschränkungen abgebildet werden²¹

Die folgende Beispielabfrage sucht alle Länder und deren Hauptstädte in Afrika²²:

¹⁹XXX QUELLE XXX

²⁰[HHL+2008a]

²¹vgl. [W3SP2008]

²²vgl. [WIKI]

```

PREFIX xxx: <http://example.com/exampleOntology#>
SELECT ?capital ?country
WHERE {
  ?x xxx:cityname ?capital.
  ?y xxx:countryname ?country.
  ?x xxx:isCapitalOf ?y.
  ?y xxx:isInContinent xxx:africa.
}

```

Als Beispiele für semantische Suchmaschinen sind zu nennen²³:

- <http://sindice.com/>
Entwicklung der DERI, umfangreiche Suche möglich
- <http://swoogle.com/>
Suchmaschine die nach RDF Dokumenten oder HTML Dateien mit enthaltenen RDF Daten sucht, es existiert eine API
- <http://www.ontosearch.org/>
Universität von Aberdeen, direkte Eingabe von SPARQL möglich

4.2 Semantic MediaWiki

Semantic MediaWiki (SMW) ist eine Erweiterung des vor allem durch die Nutzung durch Wikipedia bekannten MediaWiki. Mit SMW kann man zusätzliche semantische Anmerkungen (Annotationen) in die Texte die für den Menschen bestimmt sind einbringen, sodass die Bedeutung der Inhalte und die Beziehung der Inhalte untereinander von Computern verarbeitet werden können. Die Redakteure können direkt in den Inhalten Relationen und Attribute platzieren. Relationen können beispielsweise Zusammenhänge wie „ist Vater von“ oder „wurde erfunden von“ darstellen. Attribute stellen besondere Werte wie z.B. „hat integrierte Schaltkreise (Anzahl)“ oder „befindet sich auf Breitengrad (Ortsangabe)“ dar. Je nach Datentyp des Attributes werden auch automatische Methoden angeboten, die beispielsweise Umrechnungen von Maßeinheiten vornehmen können.

Man kann nun z.B. (vorausgesetzt die Informationen liegen vor) Suchen nach „wer ist Vater von Horst Zuse“ oder „welche Städte liegen auf dem 52. Breitengrad“ durchführen. Es ist möglich diese semantischen Informationen als OWL Ontologien im RDF Format zu erzeugen. Diese semantischen Daten können in weiteren Anwendungen genutzt werden oder in Bearbeitungsprogrammen für Ontologien weiterverwendet werden. Auch ist es möglich auf diese Weise Informationen aus unterschiedlichen Wikis zusammenzuführen.

Ein Beispiel auf der MediaWiki-Seite ist die Suche nach den „100 größten Städten die eine Frau als Bürgermeisterin“ haben. Die notwendigen Informationen (die größten Städte, aktuelle Informationen zur politischen Lage) sind in Wikipedia enthalten. Jedoch kann diese Suche momentan noch nicht durchgeführt

²³vgl. [PKR2008], Folie 30ff.

werden. Es müssten von einem Menschen so lange die Liste der größten Städte durchforstet werden, bis die 100. Bürgermeisterin gefunden ist. Dies ist nicht nur sehr zeitaufwendig, sondern die erlangte Information ist auch womöglich nur für den Moment gültig und könnte am nächsten Tag schon wieder überholt sein. Es handelt sich eindeutig um eine Aufgabe die ein Computer erledigen sollte. Mit einem Semantic MediaWiki ist eine solche Suche möglich.

Besonders hilfreich ist eine solche Ergänzung der Informationen durch Annotationen bei:

- der Erstellung solcher oder anderer Auflistungen
- bei einer Suche mit Einschränkungen (z.B. alle 1.lebenden 2.weiblichen 3.Ärzte)
- bei der Eindämmung der inflationären Verwendung von Schlagworten (Tagging) und Kategorien
- bei der Überbrückung von Sprachgrenzen (z.B. die Abfrage der aktuellen Einwohnerzahlen der Stadtteile von Taipeh/Taiwan direkt von wikipedia.tw)
- bei der Verwendung dieser Informationen in weiteren Applikationen (durch den Export/Wiederverwendung der Ontologien)
- bei der Aufdeckung von Inkonsistenzen von Informationen über Wiki-Grenzen hinweg

Es existiert ein Mapping zwischen OWL und SMW²⁴.

Eine Abfrage zur Erstellung einer Auflistung aller Länder in Afrika, die zwischen 1 und 10 Millionen Einwohner haben sieht folgendermaßen aus (als Ergebnismenge wird neben dem Name des Landes die Anzahl der Einwohner, die Fläche in Quadratkilometer und die Liste der angrenzenden Länder gewünscht):

```
{{#ask:
[[Category:Country]]
[[located in::Africa]]
[[population:=>1,000,000]]
[[population:=<10,000,000]]
|?population
|?area=km^2
|?borders
}}
```

XXX

Siehe bzw. Quellendetails:

http://de.wikipedia.org/wiki/Semantic_MediaWiki

²⁴[HHL+2008b], Folie 7

http://semantic-mediawiki.org/wiki/Semantic_MediaWiki

http://semanticweb.org/wiki/Main_Page

XXX

4.3 Semantische Webservices mit OWL-S

Service Orientierte Architekturen (SOA) stellen die Zwischenschicht zur Verfügung mit der Anwendungen über Betriebssystem-, Geräte- und Netzwerkgrenzen hinweg kommunizieren können (Interoperabilität). Hervorzuheben sind die Möglichkeiten der Einbindung von externen Service- und Datenanbietern sowie die einfache Integration von Alt-Systemen die nicht ohne weiteres in die neue Architektur überführt werden können oder sollen²⁵.

Ein Web Service (eine Einheit die innerhalb einer SOA identifizierbar ist) modelliert ein Geschäftsprozess. Dieser ist über Standard Web-Protokolle erreichbar. Die Kommunikation findet über einheitliche plattformunabhängige Protokolle statt (XML, SOAP). Die Hauptvorteile sind neben der Kapselung ganzer Geschäftsprozesse und deren Komplexität (grobgranular) die Möglichkeit zeitnaher Bindung zwischen Service Bereitsteller und Service Konsument die Integration der Alt-Systeme und deren Funktionalität.

Diese Interoperabilität beschränkt sich auf ein einheitliches Format und Syntax. Es gibt jedoch keine Möglichkeit die Funktionalität, d.h. die Bedeutung des Geschäftsprozesses/Web Services zu beschreiben. Die Herausforderung ist unabhängige und meist heterogene Datenquellen, Anwendungen und Geschäftsprozesse so miteinander zu verbinden, dass Flexibilität und Performance beibehalten und semantisch gleiche Systeme/Konzepte als solche erkannt werden können.

Ein idealisierter Ablauf der Nutzung von Webservices in einer SOA sieht folgendermaßen aus:

- die genaue Beschreibung der Funktionalität eines Webservice muss veröffentlicht werden (Publication)
- sodass im besten Fall unterschiedliche/mehrere Webservices entdeckt werden können für eine Aufgabe (Discovery)
- damit der Webservice ausgewählt werden kann, der am besten geeignet ist (Selection)
- ggf. werden unterschiedliche Webservices kombiniert um die Aufgabe zu erfüllen (Composition)
- ggf. Harmonisierung bezüglich Datendarstellung und genutztem Protokoll der zusammengestellten Webservices (Mediation)
- Ausführung der Webservices (Execution)
- Beobachtung und Kontrolle der Ausführung (Monitoring)

²⁵vgl. [HHL+2008c]

- Bereitstellung von Transaktionskontrolle, Undo (Compensation)
- ggf. Austausch/Ersatz von nicht verfügbaren WebServices (Fehler/Netzwerkprobleme) durch entsprechende weitere WebServices (Replacement)

Jedoch können bei der Discovery eines WebService Serviceanbieter und -konsument unterschiedliche Sichten auf eine Problemstellung haben. Es muss möglich sein Teilergebnisse zu erkennen und weiterzuverarbeiten. Bei der Composition von WebServices müssen unterschiedliche (Zwischen-)Ergebnisse zusammengefasst werden können.

Die Probleme sind hier die Koordination der WebServices und die eigentliche Zusammenführung der Teilergebnisse. Gewünscht ist eine automatisierte Entdeckung, Komposition und Ausführung von WebServices.

OWL-S ist eine auf OWL basierende Ontologie für das semantische Web, die Web Services beschreibt. Damit wird eine automatisierte Entdeckung, Komposition, Nutzung und Überwachung von Web Services zur Modellierung von Geschäftsprozessen möglich²⁶.

4.4 Anwendungen und Prototypen

Viele semantische Anwendungen befinden sich im Prototypenstadium oder sind Insel-Entwicklungen für eine bestimmte Domain. Manche Anwendungen sind nur als Machbarkeitsbeweis zu sehen.

- <http://food.yahoo.com>
Yahoo!-Portal zu Ernährung und Kochen
- Vodaphone Live! Mobile Portal
Vodaphone-Portal für Handy-Downloads, 50% Reduzierung Downloads
- <http://www.ontogame.com>
Unter Ausklammerung der sozialen Aspekte von Web 2.0 ein spielerischer Ansatz bei dem semantische Informationen generiert werden
- [http://www.dfki.uni-kl.de/\[userDir\]grimnes/2006/03/RDFRoom/](http://www.dfki.uni-kl.de/[userDir]grimnes/2006/03/RDFRoom/)
isometrische RDF Viewer - eine wirklich neue Perspektive

4.5 Editoren und Repositorien

Die Bearbeitung von Ontologien ist ein aufwändiges Unterfangen welches mit Softwareunterstützung erfolgen sollte.

Editoren:

- <http://protege.stanford.edu/>
Freier Ontologie Editor

²⁶vgl. [W3OW2004b]

- <http://www.co-ode.org/>
CO-ODE Projekt Universität Manchester in Zusammenarbeit mit Stanford/Protege
Online OWL-Browser
- [http://www.uvm.edu/\[userDir\]skrivov/growl/](http://www.uvm.edu/[userDir]skrivov/growl/)
Visualisierungs- und Bearbeitungswerkzeug für OWL-Ontologien

Repositorien:

- http://protegewiki.stanford.edu/index.php/Protege_Ontology_Library
Repositorium wiederverwendbarer Ontologien
- <http://www.schemaweb.info/>
Repositorium wiederverwendbarer Ontologien in RDFS, OWL und DAML+OIL

5 Zusammenfassung

Die zunehmende Menge an verfügbaren Informationen im Internet macht es unumgänglich die Art der Informationsspeicherung und des Informationsabrufs zu überdenken. Filtersysteme, statistische Methoden, Vorauswahl und Methoden aus der künstlichen Intelligenz stoßen dabei - wie im einleitenden Beispiel - an ihre Grenzen. Die Auswertung der verfügbaren Informationen und deren Zusammenhänge muss automatisiert geschehen, da sonst die Ergebnisqualität bei komplexeren Suchanfragen nicht gewährleistet ist. Das zusätzliche Versehen von Informationen mit Schlagworten und der Einsatz von ad hoc Lösungen wie z.B. Microformaten sind Zeichen dieser Notwendigkeit. Um dies zu verwirklichen ist OWL als semantische Sprache momentan eine sehr gute Möglichkeit.

OWL ist eine sehr ausdrucksstarke Sprache. Bei einer Suchanfrage im semantischen Web werden die Informationen gefunden, die gesucht werden. Die Informationen und die Beziehungen der Informationen untereinander werden rechnergestützt auswertbar. Um eine über Grenzen von Informationsquellen hinausgehende automatisierte Suche zu ermöglichen ist es wichtig, daß die semantische Auszeichnung der Informationen einheitlich ist. Bei der einheitlichen Verknüpfung von Informationen greift man auf sogenannte Ontologien zurück.

Für diese Ontologien ist wichtig, dass es sich um einen offenen Standard handelt und die Möglichkeit der Erweiterbarkeit besteht. Dies erhöht die Bereitschaft der Entwickler diese Art Technologien in ihre Anwendungen zu integrieren und zu nutzen. Die bereits bestehenden Anwendungen und Prototypen sind ein Beweis für die Erfolgsfähigkeit des Konzeptes.

[XXX Eventuell: 'Die Erhebung von SPARQL als Abfragesprache zur W3C Empfehlung 2008 ist ...' XXX]

Es bleibt jedoch abzuwarten, wann eine sinnvolle semantische Suche über große Teile der Informationen im Internet möglich sein wird. Die jeweiligen Betreiber werden ihre bestehenden Inhalte mit viel Aufwand nachbearbeiten müssen. Eine große Gefahr für den tatsächlichen Nutzen dieser Technologie für

die Suchenden bleiben die Werbenden. Diese werden einen Weg suchen sich auch im semantischen Web zu platzieren.

Literatur

- [DAML2002] DAML Committee 2002, *DAML+OIL* (<http://www.daml.org/2001/03/daml+oil-index.html>) Abgerufen am 02.11.2008
- [EIFR2007] ARD/ZDF-Online-Studie 2007, *Internetnutzung zwischen Pragmatismus und YouTube-Euphorie* Birgit van Eimeren und Beate Frees Media-Perspektiven 8/2007 Seite 362-378 (http://www.media-perspektiven.de/uploads/tx_mppublications/08-2007_Eimeren.Frees.pdf) Abgerufen am 29.10.2008
- [HESS2002] Informatik Spektrum, Springer-Verlag Berlin Heidelberg 2002, *Ontologie* Wolfgang Hesse. Universität Marburg (<http://www.gi-ev.de/service/informatiklexikon/informatiklexikon-detailansicht/meldung/57/>) Abgerufen 02.11.2008
- [HHL+2008a] , Karlsruhe Institute of Technology 2008, *Semantic Web Technologies II - OWL Syntax* Peter Haase, Pascal Hitzler, Steffen Lamparter, Denny Vrandecic (http://semantic-web-grundlagen.de/w/images/5/55/2008-04-23_SWT_02.pdf) Abgerufen am 02.11.2008
- [HHL+2008b] , Karlsruhe Institute of Technology 2008, *Semantic Web Technologies II - Semantic Web 2.0 III* Peter Haase, Pascal Hitzler, Steffen Lamparter, Denny Vrandecic (http://semantic-web-grundlagen.de/w/images/2/25/2008-07-09_SWT_13.pdf) Abgerufen am 02.11.2008
- [HHL+2008c] , Karlsruhe Institute of Technology 2008, *Semantic Web Technologies II - Semantic Web Services* Peter Haase, Pascal Hitzler, Steffen Lamparter, Denny Vrandecic (http://semantic-web-grundlagen.de/w/images/2/22/2008-06-09_SWT_08.pdf) Abgerufen am 02.11.2008
- [HKR+2008] Springer-Verlag Berlin Heidelberg 2008, *Semantic Web Grundlagen* Pascal Hitzler and Markus Krtzsch and Sebastian Rudolph and York Sure, ISBN 978-3-540-33993-9
- [PKR2008] AIFB Karlsruhe 2008, *Semantic Web: Anwendungen* (http://semantic-web-grundlagen.de/w/images/8/8f/14-anwendungen_4up.pdf) Abgerufen am 02.11.2008
- [SCHO200X] Vortrag MDC-Berlin *Ontologien - Wissensrepräsentation für die Biowissenschaften* Daniel Schober (<http://www.bioinf.mdc-berlin.de/schober/OntologieVortrag.ppt>) Abgerufen am 02.11.2008
- [W3SW2008] World Wide Web Consortium Webseite 2008, *W3C Semantic Web Activity* W3C (<http://www.w3.org/2001/sw/>) Abgerufen am 02.11.2008

- [W3OW2004a] W3C Recommendation 2004, *OWL Web Ontology Language Overview* W3C (<http://www.w3.org/TR/owl-features/>) Abgerufen 02.11.2008
- [W3OW2004b] W3C Member Submission 2004, *OWL-S: Semantic Markup for Web Services* W3C Members (<http://www.w3.org/Submission/OWL-S/>) Abgerufen 02.11.2008
- [W3RD2004] W3C Recommendation 2004, *RDF Vocabulary Description Language 1.0: RDF Schema* W3C (<http://www.w3.org/TR/rdf-schema/>) Abgerufen am 02.11.2008
- [W3SP2008] W3C Recommendation 2008, *SPARQL Query Language for RDF* (<http://www.w3.org/TR/rdf-sparql-query/>) Abgerufen am 02.11.2008
- [WIKIa] Wikipedia *Hierarchie* (<http://de.wikipedia.org/wiki/Hierarchie>) Abgerufen am 22.10.2008
- [WIKIb] Wikipedia *Taxonomie* (<http://de.wikipedia.org/wiki/Taxonomie>) Abgerufen am 22.10.2008
- [WIKIc] Wikipedia *Prädikatenlogik* (<http://de.wikipedia.org/wiki/Pr%C3%A4dikatenlogik>) Abgerufen am 02.11.2008
- [WIKId] Wikipedia *Web Ontology Language* (http://de.wikipedia.org/wiki/Web_Ontology_Language) Abgerufen am 02.11.2008
- [WIKIe] Wikipedia *RDFS* (<http://de.wikipedia.org/wiki/RDFS>) Abgerufen am 02.11.2008
- [WIKIf] Wikipedia *SPARQL* (<http://de.wikipedia.org/wiki/SPARQL>) Abgerufen am 02.11.2008