

Semantic Web: Anwendungen

Pascal Hitzler Markus Krötzsch Sebastian Rudolph

Institut AIFB · Universität Karlsruhe

Semantic Web Technologies 1 (WS07/08)

13. Februar 2008

<http://semantic-web-grundlagen.de>

Die nichtkommerzielle Vervielfältigung, Verbreitung und Bearbeitung dieser Folien ist zulässig (→ Lizenzbestimmungen CC-BY-NC).



Die Vision der Semantic Web

[Berners-Lee, Hendler, Lassila, 2001]:

“A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities”

“Computers will find the meaning of semantic data by following hyperlinks to definitions of key terms and rules for reasoning about them logically.”

“Naturally, you want to check this, so your computer asks the service for a proof of its answer, which it promptly provides by translating its internal reasoning into the Semantic Web’s unifying language.”

Und wie sieht es in der Realität aus?



Semantic Web Technologies 1

- 1 Einleitung und Ausblick
- 2 XML und URIs
- 3 Einführung in RDF
- 4 RDF Schema
- 5 Logik – Grundlagen
- 6 Semantik von RDF(S)
- 7 OWL – Syntax und Intuition
- 8 OWL – Semantik und Reasoning
- 9 SPARQL – Syntax und Intuition
- 10 Semantik von SPARQL
- 11 Konjunktive Anfragen und Regelsprachen
- 12 OWL 1.1 – Syntax und Semantik
- 13 Bericht aus der Praxis
- 14 **Semantic Web – Anwendungen** (→ Webseite)

Literatur und Verweise zu dieser Vorlesung online siehe
→ Vorlesungswebseite



Herausforderungen der Praxis

In der Praxis geht es um mehr als um Technik!

- Wer möchte semantische Daten veröffentlichen (und wer nicht)?
- Wie passen semantische Technologien in bestehende Nutzungsprozesse?
- Welchen Gewinn bringt der Umstieg auf semantische Lösungen (und was kostet er)?
- Wer bietet Dienste und Produkte für solche Technologien an (und wer kauft sie)?

The Semantic Web Chicken-And-Egg-Problem

Was kommt zuerst: formalisierte semantische Daten oder Anwendungen und Dienste zu ihrer produktiven Nutzung?



Was ist eine „Anwendung“?

Eine „Anwendung“ semantischer Technologie könnte vieles sein ...

- ... eine erste Implementierung?
- ... eine einsetzbare Implementierung?
- ... das Produkt einer Firma?
- ... der Einsatz in einem konkreten Anwendungsfall?

↪ Echte Anwendungen müssen sich in der echten Welt durchsetzen!



RSS – RDF Site Summary

Das am weitesten verbreitete Semantic-Web-Vokabular: ➔ **RSS**

- „RDF Site Summary“ (RSS 1.0)
≠ „Really Simple Syndication“ (RSS 2.0)
- Herausgegeben von der *RSS-DEV Working Group* 2000
- Übertragung von Listen aus (Hyper-)Texteinträgen
- Relevante Metadaten: Autor, Datum, Inhalt, ...
- Erzeugung durch Blogs zur Übertragung von Newsfeeds
- Auswertung/Aggregation durch RSS-Feed-Reader



Vokabulare im Semantic Web

Grundidee semantischer Wissensrepräsentation:

- Eindeutige Bezeichnung relevanter Elemente eines Anwendungsfelds (↔ URI)
- Formale Darstellung der Beziehung zwischen diesen Elementen (↔ RDF(S), OWL, ...)

Woher kommen URIs?

Vokabular

Sammlung eindeutiger Bezeichner (URIs), deren pragmatische Bedeutung (informell) spezifiziert ist, eventuell ergänzt durch einfache ontologische Beziehungen.

Meist als Schema zur Übertragung/Verknüpfung semantischer Daten in bestimmten Anwendungsfeldern.



Daten über Personen: FOAF

➔ FOAF: „Friend of a Friend“

- Vokabular für Personendaten und Social Networking
- Herausgegeben vom FOAF-Projekt (Libby Miller, Dan Brickley) seit 2000
- Heute (2008) etwa 900,000 FOAF-Beschreibungen im Netz, größtenteils automatisch erstellt
- Verschiedene Anzeigetools: ➔ FOAFexplorer, ➔ FOAFnaut, ...
- Beschreibung des Vokabulars mittels OWL



Beispiele für Elemente in FOAF:

- `foaf:Person`: Klasse aller Personen
- `foaf:mbox`: DatatypeProperty für Email-Adressen; invers-funktional
- `foaf:knows`: ObjectProperty zur Herstellung von Beziehungen zwischen Personen
- `foaf:mbox_sha1sum`: DatatypeProperty für Hashs von Email-Adressen; invers-funktional

~> `mbox` und `mbox_sha1sum` zur Zuordnung verschiedener FOAF-Files zu einer Person

Weitere Elemente und Eigenschaften für Personen, Organisationen, Gruppen, Bilder, ...



Woher kommen FOAF-Files?

FOAF wird größtenteils automatisch generiert, z.B. aus Nutzerprofilen.

Beispiel: ➔ AIFB-Portal

Einige Tools zur direkten Erstellung.

Beispiel: ➔ FOAF Creator, ➔ FOAF-a-matic



Auch Vokabularkonstruktion ist nicht einfach!

Wie stellt man Namen dar?

- International unterschiedliche Namensgebungen
- FOAF hat `firstName`, `surname`, `givenname`, `family_name`, `name`

Wie stellt man Geburtstage dar?

- Problem: Angabe von Tag oder Jahr eventuell nicht erwünscht
- XML-Schema kennt Typen für Datum und Zeitpunkt, nicht für Tage ohne Jahr
- Zusammenhang zwischen XSD-Datum und Jahreszahl nicht in OWL darstellbar



Weitere Vokabulare im Semantic Web

Andere weit verbreitete Vokabulare:

- ➔ GEO: geographische Positionen
- ➔ DOAP (Description of a Project): Metadaten zu Softwareprojekten
- ➔ CC (Creative Commons): Lizenzinformationen in RDF
- ➔ SIOC (Semantically Interlinjked Online Communities): Online-Gemeinschaften



Ontologie

Formale Beschreibungen für ein Anwendungsgebiet: Dokument einer Ontologiesprache

- Zielstellung von Ontologien nicht nur Datenaustausch
- Axiomatisierung relevanter Zusammenhänge zwecks „intelligenter“ Verarbeitung
- Nutzung ausdrucksstarker Ontologiesprachen

(„Ontologie“ und „Vokabular“ keine allgemein definierten Begriffe, Überschneidungen möglich)



Medizinische Terminologien

Anwendungen komplexer Ontologien besonders in Medizin/Biologie

Gene Ontology

- Domäne: Gene in allen Organismen
- Ontologiesprache: „OBO“, Abbildung auf OWL teilweise möglich
- Anwendung: Kombination von Gen-Daten

SNOMED – Systematized Nomenclature of Medicine (Clinical Terms)

- Domäne: Krankheiten, Diagnosen, Medikamente, ...
- Ontologiesprache: Beschreibungslogik \mathcal{EL}^{++}
- Anwendung: Informationsaustausch in Medizinanwendungen

GALEN

- Domäne: Krankheiten (medizinische Klassifikation)
- Ontologiesprache „GRAIL“, Abbildung auf OWL teilweise möglich
- Anwendung?: „basis for teaching, training and services“ (Zitat)

↪ Formale Semantik hilft bei Ontologieerstellung



Gibt es domänenunabhängige Ontologien?

↪ **Upper-Level Ontologien** (auch: *Foundational Ontologies*)

Beispiele:

- DOLCE: Descriptive Ontology for Linguistic and Cognitive Engineering
Klassen wie „Perdurant“ und „Non-Physical Endurant“
- SUMO: Suggested Upper Merged Ontology
Klassen wie „Abstract“ und „Physical“
- GFO: General Formal Ontology
- ...

Verwendung vor allem zur Modellierung und Vereinigung unabhängiger Ontologien ... *eher kompliziert* ...



Wie entstehen Ontologien?

Erstellung von Ontologien kompliziert:

- genaue Kenntnis der Domäne erforderlich
- gutes Verständnis der Ontologiesprache erforderlich

↪ Problem: Domänenexperten sind häufig keine Informatiker ...

Weitere Herausforderungen:

- Ontologien können sehr groß werden – Aufteilung?
- Viele Freiheiten bei Modellierung – wie entscheiden?
- Gemeinschaftliche Erstellung von Ontologien? Prozesse und Methoden?
- Entwicklung und Versionierung von Ontologien (Ontologie-Evolution)
- Beurteilung der Qualität von Ontologien (Ontologie-Evaluation)



Ontologien werden nicht manuell in XML geschrieben!

Freie und kommerzielle Ontologieeditoren:

- ➔ Protégé: freier Ontologieeditor, Plugins/Erweiterungen
- ➔ Topbraid Composer: kommerzieller Editor
- ➔ NeOn-Toolkit: freier Editor mit kommerziellen Erweiterungen, in Entwicklung

↔ mehr dazu in ➔ Semantic Web Technologies 2, Sommer 2008



Semantic MediaWiki

⟨ *Kurze Einleitung zu Wikis, semantischen Wikis und SMW* ⟩

Homepage: ➔ Semantic MediaWiki



Noch einmal *SmartWeb* (siehe auch ➔ Praxisvortrag Ontoprise ...)

- Fallstudie zum Einsatz semantischer Technologien in mobilen Umgebungen
- komplexe Aufgabenstellung, Verknüpfung vieler Technologien
- funktionsfähiger Prototyp
- Projekt gefördert durch BMBF, viele deutsche Partner aus Forschung und Industrie

↔ Video



Semantic MediaWiki (SMW)

- freie Erweiterung von MediaWiki (Software unter Wikipedia)
- seit Anfang 2006 verfügbar
- große weltweite Nutzergemeinschaft
- zahlreiche Anwendungen in WWW und Industrie (700 Downloads in Jan 2008)
- Übersetzung in über 10 Sprachen (darunter Chinesisch und Hebräisch)
- Entwicklung auch in verschiedenen Projekten am AIFB

Homepage: ➔ Semantic MediaWiki

↔ mehr dazu in ➔ Semantic Web Technologies 2, Sommer 2008



Beispielanwendungen von Semantic MediaWiki:

- ➔ Ontoworld.org – Semantic Web Community Wiki
- ➔ DiscourseDB – Politische Berichterstattung in den Medien
- ➔ SpieleWiki – Aktivitäten für Jugendgruppen
- ➔ Sydney Directory – Alles über Sydney
- ➔ Creative Commons Wiki – Alles über freie Inhalte und CC
- Innovationsmanagement bei Volkswagen (intern)
- Projekt- und Arbeitsgruppenwikis in verschiedenen Organisationen
- Die ➔ Homepage dieser Vorlesung
- und selbst ➔ *das*

↔ sehr viele neue OWL/RDF Daten im Netz

↔ sehr viele neue „Vokabulare“ für neue Anwendungen



Fallstudien aus Industrie- und Forschungsprojekten

Viele weitere fallstudien zum Einsatz semantischer Technologien in Karlsruhe:

- ➔ NeOn-Projekt:
Wissensverwaltung bei der FAO (UN Food and Agriculture Organisation)
Informationsintegration in der pharmazeutischen Industrie
- ➔ SEKT (inzwischen ausgelaufen):
z.B. Anwendungsstudie zur semantischen Suche in digitalen Bibliotheken
- ACTIVE (Folgeprojekt von SEKT, Start März 2008)
Einsatz semantischer Technologien in Firmenkontexten, SMW
- ➔ Nepomuk
Umsetzung des *semantischen Desktops*
- ➔ XMedia
Semantische Daten und Multimedia



Freie Positionen im Umfeld von SMW

- Studentische Hilfskräfte
 - Entwicklung
 - Kommunikation/Öffentlichkeitsarbeit
 - ...
- Diplom- und Projektarbeiten
- Stellen für Doktoranden

↔ Mail an Markus Krötzsch – mak@aifb.uni-karlsruhe.de



Ontologien speichern und verarbeiten

Systeme zum Umgang mit Ontologien/RDF-Daten:

- RDF-Datenbanken (RDF-Stores): Jena, Oracle 10g*, RAP, Redland, Sesame, Virtuoso*, ...
- OWL-Reasoner: FACT++, HermiT, KAON2 (OntoBroker-OWL*), Pellet, RacerPro*, ...
- Leicht-gewichtige OWL-Unterstützung: CEL, Oracle 11g*, QuOnto*, ...

Siehe auch: ➔ Online-Liste der DL-Reasoner

* Kommerzielle Systeme



Einbettung semantischer Metadaten in existierende Dateiformate:

- Adobe XMP – Extensible Metadata Platform
RDF-basiertes Metadatenformat in Dateien
Verfügbar in Photoshop 7.0, Acrobat 5.0, Illustrator 10, ...
- SVG (z.B. in ↪ Inkscape)
RDF-Daten in SVG (Scalable Vector Graphics, XML-basiert)
Einbettung von Metadaten Grafiken

↪ Semantik fest an existierende Dateiinhalte gekoppelt

↪ einfache Integration in bestehende IT-Prozesse

Anwendungen mit „Semantic Web inside“

Nicht alle Anwendungen semantischer Technologien exportieren Daten!

↪ Yahoo! Food

- Yahoo!-Portal zu Ernährung und Kochen
- interne Datenhaltung mit RDF

Vodafone Live! Mobile Portal

- Vodafone-Portal für handy-Downloads
- interne Datenverwaltung mit RDF
- Verbesserung in der Suche: 50% weniger Seitenaufrufe pro Download

OWL/RDF als Metadatenformat in der Medienverwaltung:

↪ Microsoft „Interactive Media Manager“

Semantische Websuche

Websuche auf Basis von semantischen Dokumenten in Netz

↪ Yahoo! Creative Commons Search

- Suche nach Webinhalten auf Basis ihrer Lizenzbestimmungen
- Angabe von gewünschten Rechten (kommerzielle Nutzung, Bearbeitung)
- Basierend auf Annotationen in Creative-Commons-RDF-Vokabular (z.B. von Flickr)

Aktuelle Forschung: Semantic-Web-Suchmaschinen

- Aggregation von semantischen Daten, crawling
- Suche nach technischen/strukturellen Merkmalen von Metadaten
- Beispiele: ↪ Swoogle, ↪ Sindice
- Meist relativ unvollständige Crawler, unvollständige Datensätze

Webdienste als wichtiger Architekturteil moderner IT-Systeme (SOA)

➔ Microsoft Connected Services Framework

- Interne Verwaltung von Nutzerprofilen in RDF
- Zugriff via SPARQL möglich

Aktuelle Forschung: Semantic Webservices

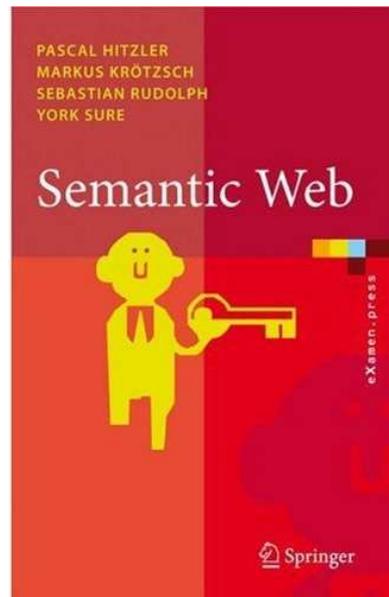
- automatische Suche, Konfiguration, und Invokation von Diensten aufgrund von semantischen Beschreibungen
- Neu in Karlsruhe: ➔ *Karlsruhe Service Research Institute (KSRI)*

Literatur

Pascal Hitzler
Markus Krötzsch
Sebastian Rudolph
York Sure

Semantic Web
Grundlagen

Springer 2008, 277 S., Softcover
ISBN: 978-3-540-33993-9
Aktuelle Literaturhinweise online



Semantische Technologien sind in der Praxis angekommen

- Industrielle Produkte, Dienstleistungen und Anwendungen verfügbar
- Vokabulare, Ontologien und „Best practices“ für verschiedene Domänen
- Fokus auf „Web of Data“ oder „Web 3.0“
- RDF stärker verbreitet als OWL

Was bringt die Zukunft?

- weiter verstärkte Annahme durch Industrie
- neue Geschäftsmodelle
- Weiterentwicklung der Grundlagentechnologien
- Übergang von RDF-Vokabularen zu OWL-Ontologien?

Wer bringt die Zukunft?