

Semantic Web und semantische Technologien: Zentrale Begriffe und Unterscheidungen

Andreas Blumauer, Tassilo Pellegrini

Semantic Web School – Centre for Knowledge Transfer, Wien, Österreich;
E-Mail: {a.blumauer, t.pellegrini}@semantic-web.at

Zusammenfassung: Naturgemäß zieht jede interdisziplinäre Auseinandersetzung, wie hier zum Thema Semantic Web, Ungereimtheiten auf begrifflicher Ebene nach sich. „Wenn die Begriffe nicht klar sind, breitet sich Unordnung aus“ weiß man spätestens seit Konfuzius. Dieses Kapitel soll daher jene zentralen Begriffe der Semantic Web-Entwicklung einführen und weitgehend definieren, die dann in den einzelnen Kapiteln jeweils in unterschiedliche Kontexte gesetzt werden. Anhand des „A-O-I-Modells des Semantic Web“ werden schließlich jene drei Betrachtungsweisen zueinander in Beziehung gesetzt, denen in der derzeitigen Entwicklungsstufe des Semantic Web die jeweils höchste Aufmerksamkeit geschenkt wird: Anwenderkontext, Organisations-Kontext und (technische) Infrastruktur. Es eignet sich als Orientierungshilfe, um die unterschiedlichen Zielsetzungen in der Auseinandersetzung mit semantischen Technologien und dem Semantic Web zu einem „Big Picture“ verschmelzen zu lassen, und um damit etwaige Missverständnisse, die immer dann entstehen, wenn unterschiedliche Standpunkte einengenommen werden, zu vermeiden.

1 Zentrale Begriffe

1.1 Semantik

In einem Buch über das „semantische Web“ sollte zunächst geklärt sein, was an diesem Konzept eines Internet der nächsten Generation semantisch ist.

„Die Semantik (Bedeutungslehre) ist das Teilgebiet der Sprachwissenschaft (Linguistik), das sich mit Sinn und Bedeutung von Sprache beziehungsweise sprachlichen Zeichen befasst.“ [19]

In Neumüller wird folgende Definition verwendet: „... semantics is the study of meaning. As used by Charles Morris, that branch of semiotics devoted to studying the relationship between signs and their objects.“ ([15, p. 195])

Wesentlich ist also einerseits die Unterscheidung zwischen Sinn (engl. „meaning“) und Bedeutung (engl. „reference“). Neumüller [15] veranschaulicht dies anhand eines einfachen Beispiels: Morgenstern und Abendstern beziehen sich beide auf denselben Himmelskörper, nämlich die Venus, haben insofern die selbe Bedeutung („reference“), meinen („meaning“) jedoch unterschiedliche Dinge, genauso wie mit Sonnenaufgang und Sonnenuntergang Verschiedenes gemeint ist.

Diese in der Semiotik wichtige Unterscheidung nimmt in der zur Zeit eher aus Sicht der Informatik geführten Auseinandersetzung mit dem Semantic Web eine eher untergeordnete Rolle ein, umso wichtiger erscheint daher der Vergleich der Gebiete der Semiotik (Syntax, Semantik, Pragmatik), der anhand des semiotischen Dreiecks gezogen werden soll: Nach Charles S. Peirce ist jedes Zeichen untrennbar mit dem Bezeichneten (Gegenstand) und dem Interpret (Referent) in einer triadischen Struktur verbunden.¹

Die Syntax beschäftigt sich mit den Beziehungen zwischen den Zeichen untereinander², der semantische Aspekt (wie bereits im vorigen Absatz erwähnt) mit den Beziehungen zwischen Zeichen und Objekten bzw. Gegenständen der Außenwelt und der pragmatische Aspekt mit den Beziehungen der Zeichen gegenüber den Interpreten und Kontexten. [13]

Insofern ist der Begriff „Semantic Web“ eigentlich zu kurz gefasst, da in der gegenwärtigen Debatte alle drei Aspekte berücksichtigt werden. In Anlehnung an John F. Sowa [20] handelt es sich offensichtlich um das „Semiotic Web“³:

„The Internet is a giant semiotic system. It is a massive collection of Peirce's three kinds of signs: icons, which show the form of something; indices, which point to something; and symbols, which represent something according to some convention. But current proposals for ontologies and metadata have overlooked some of the most important features of signs. A sign has three aspects: it is (1) an entity that represents (2) another entity to (3) an agent. By looking only at the signs themselves, some metadata proposals have lost sight of the entities they represent and the agents – human, animal, or robot – which interpret them.“

Es scheint also, als hätte Tim Berners-Lee mit der Ausarbeitung einer Roadmap zum „Semantic Web“ [5] die Büchse der Pandora geöffnet, da die Auseinandersetzung mit Semantik im engeren Sinn jedenfalls seit über 100 Jahren geführt wird und die Bedeutungslehre weit über 2000 Jahre alt ist. Es ist daher kaum verwunderlich, dass an der Entwicklung eines globalen Projektes wie eben dem „Semantic Web“, das nun im Gegensatz zum

¹ Siehe dazu auch die Beiträge von Budin und Ehrig u. Studer in diesem Band.

² Paradoxerweise liegt der Schwerpunkt der Semantic Web – Forschung nach wie vor auf Ebene der Syntax, geht sie doch in vielen Fällen der Frage nach, wie technische Systeme durch geeignete Datenmodelle „interoperabel“ gestaltet werden können.

³ Für eine ergänzende Auseinandersetzung siehe auch den Beitrag von Weber u. Fröschl in diesem Band.

Internet der ersten Generation nicht einfach nur Daten übertragen soll, sondern auch Bedeutung vermitteln will, auch – und vor allem – Kultur- und Geisteswissenschaften mitwirken wollen und werden.

1.2 Metadaten

Web-Content (und ebenso Information im Intranet) wird zum gegenwärtigen Zeitpunkt vor allem für den Menschen aufbereitet und formatiert. Information, die mittels (X)HTML⁴ oder als Dokument im PDF-Format ausgedrückt wird, genügt dem menschlichen Auffassungsvermögen zumeist, damit verstanden werden kann, worum es sich da bei handelt, obwohl die dafür notwendige Meta-Information (z. B. Syntax- oder Kontext-Information auf semantischer Ebene) gar nicht explizit vorliegt.

Für Maschinen (Webcrawler, Softwareagenten etc.) ist z. B. zunächst nicht klar, dass es sich bei der Stringkette „Mo 21.12.2006, 9:30“ um eine Datumsangabe handelt.

Metadaten sind Daten über Daten. Mit Metadaten kann der Sinn („meaning“) von Daten ausgedrückt werden, bezieht sich also auf die Semantik im Semantic Web. [2]

Metadaten kommen prinzipiell seit Jahrhunderten in der bibliothekarischen Praxis zum Einsatz. Metadaten werden eingesetzt, um Informationsressourcen zu beschreiben und dadurch besser auffindbar zu machen und Beziehungen zwischen den Ressourcen herzustellen. Voraussetzung dafür ist Erschließung mit einem gewissen Standardisierungsgrad.

Metadaten im Semantic Web werden mittels RDF organisiert und strukturiert und bilden damit die Basis für den Einsatz von Domänenontologien:

„RDF is an infrastructure that enables the encoding, exchange, and reuse of structured metadata. Search engines, intelligent agents, information brokers, browsers and human users can make use of the semantic information. RDF is an XML application (i. e., its syntax is defined in XML) customized for adding meta-information to Web documents [...]“ [8, p. 19]

Wie im obigen Zitat angeschnitten, reicht XML alleine nicht aus, um jene Metadaten-Infrastruktur aufzubauen, die für die Realisation des Semantic Web Voraussetzung ist.⁵ Die Generierung von strukturierten Metadaten kann jedoch durch den Einsatz entsprechender Content Management Systeme, Portallösungen und Learning Management Systeme ohne zusätzlichen Aufwand im laufenden Prozesseffolgen.⁶ Damit zeigt sich auch, dass das Semantic Web nicht erst bei der Entwicklung „superintelligenter

⁴ Siehe dazu: <http://www.w3.org/MarkUp/>, zuletzt aufgerufen am 3.1.2006

⁵ Siehe dazu auch den Beitrag von Birkenbihl in diesem Band.

⁶ Siehe dazu den Beitrag von Koller in diesem Band.

Agenten“ be ginnt⁷ [2], sondern eben schon bei „basaleren“ Dingen wie der Verwendung geläufiger Metadaten-Schemata.

1.3 Ontologie

Zunächst muss festgestellt werden, dass jene philosophische Position, die unter „realistischer Ontologie“ [12] zusammengefasst wird, jedenfalls nicht mit jener Definition einer Ontologie in Einklang zu bringen ist, die innerhalb der Semantic Web-Community auf breitesten Konsens stößt und von Tom Gruber im Kontext der Wissensteilung und Wissenswiederverwendung 1993 vorgeschlagen wurde: „An ontology is an explicit specification of a conceptualization.“ [9].

Ontologien wurden im Umfeld der Künstlichen Intelligenz entwickelt und sind die zentralen Bausteine des Semantic Web: Mit ihnen kann Wissen einer Domäne formal repräsentiert und prinzipiell unabhängig von Programmen wieder verwendet werden. Sie beschreiben also Konzepte und ihre Beziehungen innerhalb einer Wissensdomäne und unterstützen Maschinen dabei, Inhalte im Web interpretieren zu können, anstatt sie einfach darzustellen und damit sämtliche Vernetzungstätigkeiten dem Menschen zu überlassen bzw. „aufzuheben“⁸. Sie bilden die höchste Stufe der semantischen Reichhaltigkeit (siehe Abbildung 3) und setzen auf semantische Netze auf: Der Begriff „Ontologie“ ist aber keineswegs eindeutig und wird auf vielfältige Weise gebraucht, da unterschiedliche Zielsetzungen in einem Spektrum adressiert werden, das sich von maschineller Lesbarkeit von Daten (Automatisierung) bis hin zur Unterstützung von Menschen bei der Erfüllung komplexer, wissensintensiver Arbeiten (Wissensmanagement) erstreckt.

Ontologien werden also entwickelt und eingesetzt, um

- den Datenaustausch zwischen Programmen zu ermöglichen
- die Vereinheitlichung und Übersetzung zwischen verschiedenen Wissensrepräsentationsformen zu ermöglichen
- Services zur Unterstützung von Wissensarbeitern zu entwickeln
- Theorien abzubilden
- die Semantik strukturierter und semi-strukturierter Information auszudrücken
- die Kommunikation zwischen Menschen zu unterstützen und zu erleichtern

⁷ Dies wird vor allem immer wieder von Kritikern, die gegenüber der Idee des Semantic Web, wie z.B. Clay Shirky (siehe: http://www.shirky.com/writings/semantic_syllogism.html), kontroversiell eingestellt sind, fälschlicherweise angenommen. *Semantic Web ist nicht gleich Künstliche Intelligenz!*

⁸ Siehe dazu <http://www.w3.org/TR/owl-guide/>, aufgerufen am 10.12.2005

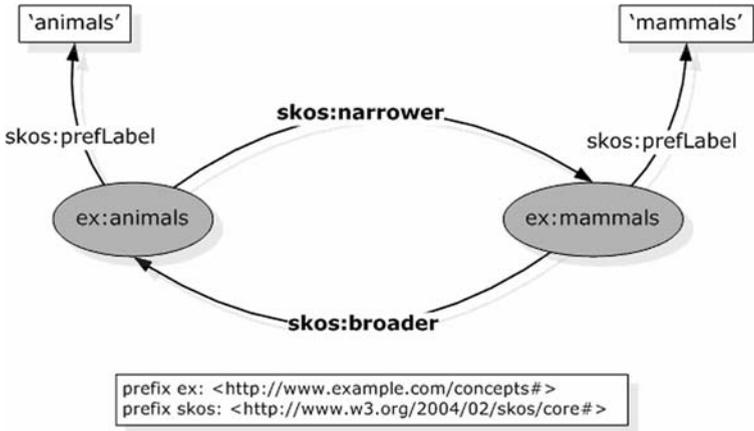


Abb. 1. Beispiel einer Visualisierung eines RDF-Dokuments

Anhand dieser Aufzählung wird auch klarer, dass weder mit Ontologien, noch mit anderen Formen der Wissensrepräsentation, wie semantischen Netzen, primär die Ebene der Darstellung und Visualisierung von Wissensnetzen – also ontologie-gestützte Navigationshilfen – gemeint ist, sondern zunächst einmal die zugrunde liegenden Modelle, die den Wissensraum formal beschreiben.⁹

Ein kleines Beispiel zeigt ein RDF-Dokument¹⁰ und eine korrespondierende Visualisierung¹¹:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">

  <skos:Concept
rdf:about="http://www.example.com/concepts#mammals">
    <skos:prefLabel>mammals</skos:prefLabel>
    <skos:broader
rdf:resource="http://www.example.com/concepts#animals"/>
  </skos:Concept>
```

⁹ Siehe dazu den Beitrag von Kienreich u. Strohmaier in diesem Band.

¹⁰ Das Beispiel wurde dem SKOS Core Guide des W3C entlehnt, siehe dazu: <http://www.w3.org/TR/2005/WD-swp-skos-core-guide-20051102/>, zuletzt aufgerufen am 11.12.2005

¹¹ Eine entsprechende grafische Darstellung kann mit dem RDF Validation Service des W3C generiert werden, siehe dazu: <http://www.w3.org/RDF/Validator/>, zuletzt aufgerufen am 11.12.2005

```

<skos:Concept
rdf:about="http://www.example.com/concepts#animals">
  <skos:prefLabel>animals</skos:prefLabel>
  <skos:narrower
rdf:resource="http://www.example.com/concepts#mammals"/>
  </skos:Concept>
</rdf:RDF>

```

1.3.1 *Ontologie als sozialer Verhandlungsprozess*

Die Ontologie als „gemeinsames Modell eines Konzeptes der Wirklichkeit“ [10] gemeint, das „in einem sozialen Verhandlungsprozess einem Konsens zugeführt wird“, stellt den Idealfall dar: Vorliegende Methoden zur Erstellung von Ontologien [1, 22] betrachten die soziale und kollaborative Komponente des Erstellungsprozesses als „Black Box“ und schlagen vor allem Top-down-Vorgehensmodelle vor. Auf das so genannte „Consensus-Building“ wird kaum eingegangen, vielmehr wird vom „Ontology-Engineering“ gesprochen, das auf der Annahme beruht, dass Modelle der Wirklichkeit, die die Basis kollaborativen Arbeitens bilden sollen, wie Automaten entworfen werden können. Dies mag für die Entwicklung von Expertensystemen eine erprobte Vorgehensweise sein, dürfte sich allerdings in der Entwicklung konsensfähiger Metamodelle in größeren Communities im WWW als ungeeignet herausstellen.

Inwiefern sich daher Entwicklungen im Bereich der Social Software wie „keyword tagging“, „social bookmarking“¹² und damit verbundenen „Folksonomies“¹³, also bottom-up-Strategien auf Basis von user feedback zur Organisation von Wissen dazu eignen¹⁴, einen Beitrag zur Erstellung von Ontologien leisten zu können, warum Folksonomies jedenfalls keine Alternative zu den teilweise mühevoll zu erstellenden Ontologien darstellen, und wie mit Hilfe von Ontologien wiederum das „tagging“ organisiert werden kann, darauf wird in [11] eingegangen.

Im gesamten Entwicklungsprozess zum Zwecke der Wissensorganisation überlagern sich Prozesse und Ordnungssysteme in Form einer Kombination aus „qualitätsgesicherten“ und standardisierten Ontologien, weitgehend akzeptierten Terminologiesystemen, wie z. B. Fachthesauri und schnelllebigen Folksonomies, die insgesamt einander ergänzen und jeweils die Basis für die weitere Entwicklung tragfähiger Wissensmodelle bilden. Bei Mika [14] wird auf diese Dynamik explizit eingegangen und der Begriff „Ontologie“ im Sinne einer „linguistischen Ontologie“, also einer Terminologie verwendet:

¹² Siehe dazu: <http://del.icio.us/>, aufgerufen am 10.12.2005

¹³ Ein guter Überblick zum Thema „Folksonomies“ wird in diesem Weblog geboten: <http://today.tuwien.ac.at/cheesy/>, aufgerufen am 10.12.2005. Siehe dazu auch den Beitrag von Schmitz et al. in diesem Band.

¹⁴ Siehe dazu den Beitrag von Schuster u. Rappold in diesem Band.

„Considering the dynamics of the community and the extent of neologism, the ontologies emerging from folksonomies such as del.icio.us also have a large potential for enriching established, but slowly evolving linguistic ontologies such as Wordnet.“ [23]

1.3.2 *Ontologie als Wissensmodell mit hoher semantischer Reichhaltigkeit*

Dieser Zusammenhang zwischen Ontologie, Terminologie und Folksonomie kann anhand des Münchener Modells des Wissensmanagements und dem damit verbundenen Konzept der „Wasseranalogie des Wissens“ [16] veranschaulicht werden:

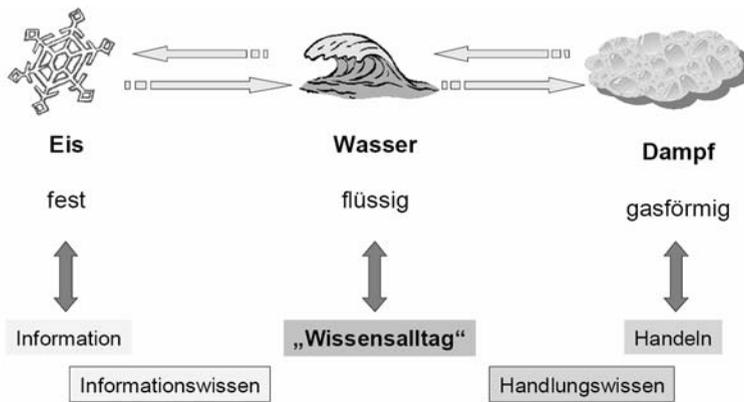


Abb. 2. „Wasseranalogie des Wissens“

Gasförmiges Wissen wird sozusagen lauffend – auch in zwischenmenschlicher Konversation oder in eMails – „produziert“ und wird dementsprechend, wenn überhaupt, unstrukturiert erfasst. Die Kommunikation über Blogs in der Blogosphäre stellt hier ein entsprechendes verteiltes Kommunikationssystem zur Verfügung, in dem laufend „gasförmiges Wissen“ produziert wird und daher auch nicht zu fälligerweise im Sinne eines laufenden „Trend-Scoutings“ als Struktur gebende Komponente und Basis für das laufende Crawlen von Webseiten der Suchmaschinenbetreiber herangezogen wird. Folksonomies und Social Tagging und die damit verbundene Bildung von Themen-Communities mündet schließlich in der nächsten Stufe in Form „flüssigen Wissens“ in der Bereitstellung und Systematisierung eines gemeinsamen, kontrollierten Vokabulars in Form von (multilingualen) Thesauri oder semantischen Netzen. Erst darauf aufbauend können verlässliche und tragfähige Ontologien im Sinne eines semantisch reichhaltigen Wissensmodells innerhalb einer Wissensdomäne geschaffen werden.

Zusammengefasst kann diese Evolution semantischer Modelle in Form der semantischen Treppe dargestellt werden:

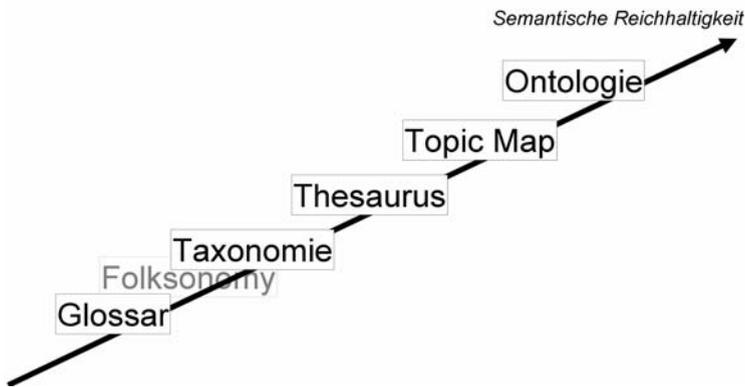


Abb. 3. „Semantische Treppe“

Von „Ontologien“ im eigentlichen Sinn so lte da her er st ge sprochen werden, sobald im semantischen Modell explizit Möglichkeiten vorgesehen sind, Schlüsse auf Ontologieebene (und nicht etwa auf Applikationsebene) zu ziehen, um beispielsweise festzustellen, ob eine bestimmte Konzeptdefinition überhaupt erfüllbar ist.¹⁵

1.3.3 Domänen-Ontologien, „Light-Weight-Ontologies“ und Datenschemata

Dementsprechend verwirrend ist daher auch die Verwendung des Begriffs „Ontologie“, um Metadatenschemata wie Dublin Core¹⁶, RSS¹⁷, FOAF¹⁸ zu bezeichnen. Eine Typisierung von Ontologien wird in Anlehnung an [8] vorgeschlagen und zeigt die Bandbreite der Verwendungsmöglichkeiten des Begriffs „Ontologie“:

- Domänen-Ontologien erfassen das Wissen innerhalb einer Domäne (z. B. Bioinformatik, Städtetourismus oder wiederverwendbare Energieformen).
- Metadaten-Ontologien dienen als Vokabular zur Beschreibung von Informationsquellen bzw. -typen.

¹⁵ Siehe dazu den Beitrag von May in diesem Band.

¹⁶ Siehe dazu: <http://dublincore.org/>, aufgerufen am 10.12.2005

¹⁷ Siehe dazu: <http://blogs.law.harvard.edu/tech/rss>, aufgerufen am 10.12.2005

¹⁸ Siehe dazu: <http://xmlns.com/foaf/0.1/>, aufgerufen am 10.12.2005

- Upper-Level-Ontologien (Generische Ontologien) bzw. Ontologien zur Beschreibung des Allgemeinwissens, um z. B. allgemeines Wissen über Teil-Ganzes-Beziehungen (Mereologie) auszudrücken, die die Basis für eine Vielzahl von spezifischeren Domänen-Ontologien bilden können. Ein Beispiel hierfür ist SUMO¹⁹.
- Light-Weight-Ontologies, die u. A. mittels SKOS²⁰, als Vorschlag einer Spezifikation zur Beschreibung von Thesauri oder durch den Topic Maps-Standard²¹ ausgedrückt werden können.

Die Vielzahl der Verwendungsarten des Begriffes Ontologie spiegelt sich auch in diesem Band wider, es wird an dieser Stelle jedoch vorgeschlagen, die zukünftige Entwicklung eines „semantischen“ Webs auch auf Basis entsprechend wohl definierter Terminologiesysteme zuzugeschlagen²² und gerade einen zentralen Begriff wie „Ontologie“ entsprechend genauer zu bezeichnen und nicht einfach als Homonym bzw. Oberbegriff zu verwenden.

1.4 Wissensmanagement

Wissensmanagement umfasst im Wesentlichen folgende vier Handlungsfelder [17]:

- *Inhalt / Kontext*: Transparenzierung von bestehendem Wissen
- *Mensch / Kompetenz*: Steuerung der Entwicklung von aktuell oder künftig benötigtem Wissen
- *Zusammenarbeit*: Förderung des Austausches von Wissen
- *Orientierung*: Information zielgerecht vernetzen und finden

Da Wissen stets personengebunden ist, können Informationssysteme also vor allem einen Beitrag zur Verbesserung der Informationslogistik und Steigerung der Informationsqualität beitragen, nicht aber das „Wissen speichern“ – diese Auseinandersetzung ist weitgehend überwunden.

Semantisch basiertes Wissensmanagement oder kurz „semantisches Wissensmanagement“ beruht auf der Annahme, dass Lern-, Innovations-, Kommunikations- und Entscheidungsprozesse durch semantisch angereicherte Informationsbestände – also durch das Verfügbarmachen von Kontextinformation – effizienter ablaufen, dazu Studer [21]:

¹⁹ Siehe dazu: <http://www.ontologyportal.org/>, aufgerufen am 10.12.2005

²⁰ Siehe dazu: <http://www.w3.org/2004/02/skos/>, aufgerufen am 10.12.2005

²¹ Siehe dazu: <http://www.topicmaps.org/xtm/1.0/>, aufgerufen am 10.12.2005

²² Siehe dazu: <http://www.semantic-web.at/thesaurus>, aufgerufen am 3.1.2006 bzw. <http://www.w3.org/2003/glossary/>, aufgerufen am 3.1.2006

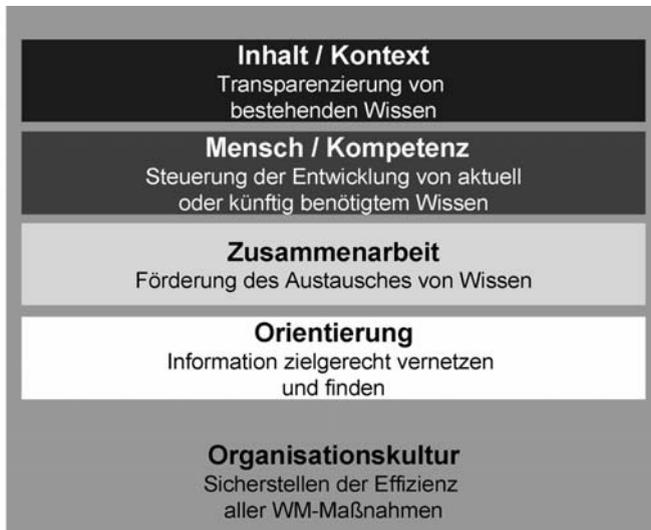


Abb. 4. Primäre Handlungsfelder und Ziele des Wissensmanagements

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

Zahlreiche Beiträge dieses Bandes beschäftigen sich mit der Rolle der Semantik in Wissensmanagement-Systemen.

1.5 Interoperabilität

Ein wichtiger Schritt für die Realisierung eines Semantic Web ist die Herstellung von Interoperabilität. Interoperabilität bezeichnet einen Zustand der Vereinheitlichung und Vereinfachung von Geschäftsprozessen und IT-Architekturen in der organisationsinternen und -übergreifenden Kommunikation. Das Ziel ist die Interaktion von dispersen Datenbeständen und Anwendungen auf technischer, organisationaler und semantischer Ebene²³ zu ermöglichen, ohne dass die Autonomie der einzelnen Teilsysteme aufgehoben wird.²⁴ Dies ist besonders dort von Relevanz, wo die interagierenden Akteure unterschiedliche Datenformate, Terminologien oder Definitionen verwenden und die Art und Intensität des Datenaustausches mit dem jeweiligen Kontext variiert. Man denke hier etwa an die staatenübergreifende

²³ Siehe dazu auch den Beitrag von Galinski in diesem Band.

²⁴ Siehe dazu den Beitrag von Reitbauer in diesem Band.

Kommunikation zwischen Behörden (z. B. Ministerien, Krankenkassen, Arbeitsämtern etc.) oder Konzernen, die aus historischen Gründen mit proprietären Datenformaten arbeiten oder kulturelle Eigenheiten pflegen, die außerhalb des organisationalen Kontextes nicht geläufig sind.

„The use of semantic technologies makes it possible to describe the logical nature and context of the information being exchanged, while allowing for maximum independence among communication parties. The results are greater transparency and more dynamic communication among information domains irrespective of business logic, processes and workflows.“ ([18, p. 37])

Seit Ende der 1990er Jahre hat sich eine Vielzahl an nationalen und supranationalen Initiativen gebildet, die das Thema Interoperabilität in unterschiedlichen Bereichen wie Landwirtschaft, Tourismus, Gesundheitswesen, Transportwesen u. v. a. m. vorantreiben.²⁵ Die Komplexität der hierbei zu bewältigenden Herausforderungen ergibt sich vor allem aus dem überbordenden Koordinationsbedarf zwischen den beteiligten Verhandlungspartnern und Interessensgruppen, wodurch das Thema Interoperabilität seit einigen Jahren als vorrangiges politisches Thema wahrgenommen wird.²⁶ Die politische Dimension ergibt sich vornehmlich aus dem Anspruch der branchen- und sektorenübergreifenden institutionellen Koordination von IT-Architekturen und dem Bestreben durch die Entwicklung von allgemein akzeptierten Normen und Standards bestehende Informations- und Kommunikationssysteme und Terminologien auf großtechnischer Ebene interoperabel zu machen.

2 Wesentliche Unterscheidungen

2.1 Semantic Web versus Semantische Technologien und andere Entwürfe des Internet der nächsten Generation

„Semantic Web“ wird im vorliegenden Band einerseits als Konzept für ein Internet der nächsten Generation verstanden und kann somit mit anderen ähnlich gelagerten Entwürfen, z. B. dem Web 2.0 verglichen werden, zeichnet sich demgegenüber jedoch dadurch aus, dass es bereits auf eine längere Entwicklungsgeschichte zurückblicken kann [4, 5] und daher auf entsprechend fundierten technischen Konzepten fußt. Insofern muss das

²⁵ Für eine umfangreiche Übersicht zu den Aktivitäten in der Europäischen Union siehe <http://europa.eu.int/idabc/en/chapter/550>, aufgerufen 15.12.2005. Für eine branchenspezifische Betrachtung im Bereich der USA automotiv Industrie siehe <http://www.mel.nist.gov/proj/mi.htm>, aufgerufen am 15.12.2005

²⁶ In diesem Zusammenhang ist insbesondere das European Interoperability Framework zu nennen. Siehe auch <http://europa.eu.int/idabc>, aufgerufen 15.12.2005

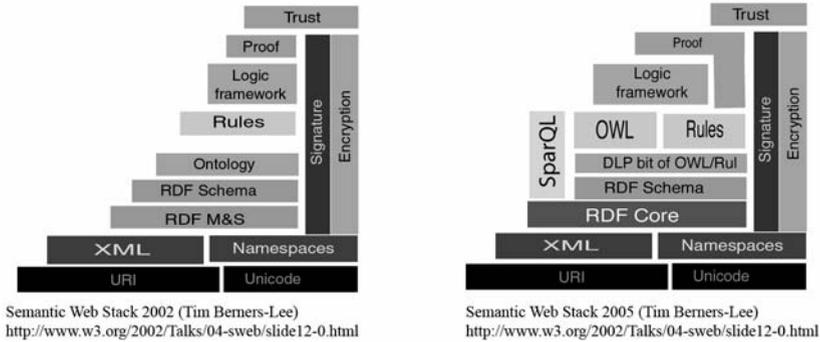


Abb. 5. Der Semantic Web Stack 2005 im Vergleich zu seinem Vorgänger aus dem Jahre 2002

Web 2.0 eher als Bemühung einzelner Konzerne verstanden werden, technische und soziale Entwicklungen im Internet rasch aufzugreifen und zu vermarkten, als ein durchdachtes Konzept für das Web der nächsten Entwicklungsstufe anzubieten, wie dies der Name „Web 2.0“ vielleicht verheißen mag.

Andererseits soll an dieser Stelle darauf hingewiesen werden, dass es nicht „ein“ Semantic Web gibt, beispielsweise im Sinne eines semantischen Netzes innerhalb einer Organisation, sondern lediglich „das“ Semantic Web als Erweiterung des bestehenden Internet.

Entsprechend wichtig ist auch die Unterscheidung zwischen Semantic Web und „semantischen Technologien“: Während das Semantic Web im Kern auf Standards zur Beschreibung von Prozessen, Dokumenten und Inhalten sowie entsprechenden Metadaten – vorwiegend vom W3C²⁷ vorgeschlagen – aufsetzt, und damit einen Entwurf für das Internet der nächsten Generation darstellt, adressieren semantische Technologien Herausforderungen zur Bewältigung komplexer Arbeitsprozesse, Informationsmengen bzw. Retrievalprozessen und Vernetzungs- oder Integrationsaktivitäten, die nicht nur im Internet, sondern auch innerhalb von Organisationsgrenzen in Angriff genommen werden.

Semantische Technologien, wie z. B. automatische Texterschließung, der Einsatz von semantischen Netzen, Ontologien oder Informationsextraktion mit Hilfe von Wrapper-Technologien, kommen in vielfältiger Weise bereits heute, vor allem im Kontext wissensintensiver, interdisziplinärer und kommunikationsintensiver Arbeitsabläufe als Mittel im „Kampf gegen die Informationsflut“²⁸ zum Einsatz.

²⁷ Siehe dazu: <http://www.w3.org/>, aufgerufen am 11.12.2005

²⁸ Wobei eigentlich nicht die Menge der verfügbaren Information zu bemängeln wäre, sondern deren niedrige Qualität.

Eine Studie der Fraunhofer Gesellschaft²⁹ [7] hat gezeigt, dass einerseits die Akzeptanz semantischer Technologien zur Unterstützung des innerbetrieblichen Wissensmanagements bereits weit fortgeschritten ist, eine Untersuchung der Semantic Web School hat andererseits aber auch vergleichsweise wenige konkrete Anwendungsfälle des „World Wide Semantic Webs“ ans Tageslicht gefördert [6].

2.2 Unterschiedliche Zielsetzungen: Automatisierung von Prozessen versus maschinelle Unterstützung bei der Bewältigung wissensintensiver Prozesse

Automatisierung als das maschinelle Verarbeiten von Information, um Prozessschritte zu automatisieren, ist zentrales Anliegen im Semantic Web. Semantik als Mittel zum Transport jener Information, die es Maschinen erlaubt, zu „verstehen“ und zu „entscheiden“ verfolgt allerdings grundlegend andere Ziele als die Anreicherung von Informationsobjekten mit semantischer Information, um den Wissensarbeiter bei der Bewältigung komplexer, wissensintensiver Prozesse zu unterstützen.

Während also im innerbetrieblichen Wissensmanagement semantische Technologien, wie semantische Netze beispielsweise Suchanfragen des Wissensarbeiters dahingehend unterstützen, dass auch Dokumente gefunden werden, in denen der Suchbegriff gar nicht vorkommt, vielleicht aber ein Synonym davon oder der englischsprachige Ausdruck dafür, so zielen jene Forschungsaktivitäten, in denen die Entwicklung von „Semantic Web Services“ im Zentrum steht, darauf ab, Prozesse zu automatisieren:

„Organizations that provide various services can tag those services with meaning; using Web-based software agents, you can dynamically find these services on the fly and use them to your benefit or in collaboration with other services.“ [3]

Mit der Automatisierung von Prozessen werden folgende Anwendungsfelder des Semantic Web adressiert: Web Commerce (B2C)³⁰, E-Business (B2B)³¹ und Enterprise Application Integration (EAI)³², wohingegen das Szenario „Wissensmanagement“ Technologien des Semantic Web mit anderen Schattierungen einzusetzen weiß.

²⁹ Siehe dazu: <http://www.fraunhofer-studie.de/>, aufgerufen am 11.12.2005

³⁰ Siehe dazu den Beitrag von Werthner u. Borovicka in diesem Band.

³¹ Siehe dazu den Beitrag von Fill et al. in diesem Band.

³² Siehe dazu den Beitrag von Reitbauer in diesem Band.

2.3 Unterschiedliche Perspektiven im Semantic Web: das A-O-I-Modell

Die Debatte um das „Semantic Web“ wurde zwar innerhalb akademischer Kreise der Informatik – vor allem im Umfeld um fassender EU-Projekte bzw. auf Initiative der DARPA³³ – gestartet, wird aber in der Zwischenzeit in zahlreichen Communities bzw. Berufsfeldern als Erfolg versprechendes Konzept zur Bewältigung der jeweils systeminhärenten Probleme mit unterschiedlich großer Euphorie diskutiert.

Diese zuhöchst transdisziplinäre Auseinandersetzung mit einem zentralen Thema der Wissensgesellschaft sollte daher auf Basis einer gemeinsamen Topographie geführt werden, aus dem Bewusstsein heraus, dass zwar unterschiedliche Perspektiven ein genommen und Subziele verfolgt werden, dass den noch im Grunde alle Proponenten vernetzter Wissensarbeit am selben Projekt arbeiten und ein Oberziel verfolgen, das in seinem Kern zusammengefasst lauten könnte:

Menschengerechte Computeranwendungen zu entwickeln, die auf Basis vernetzter Strukturen und effizienter Informationsflüsse die Verarbeitung und Veredelung von Informationsobjekten in unterschiedlichen Realitäten und Kontexten erlaubt, um schließlich handlungsrelevantes Wissen zu stimulieren und zu generieren.

In Folge werden typische Fragestellungen im Zusammenhang mit der Semantic Web-Debatte aufgezählt. Es wird so ersichtlich, dass die Entwicklung auf drei unterschiedlichen Ebenen im Gange ist und teilweise nur lose gekoppelt ist:

2.3.1 Anwenderperspektive³⁴

- Wie kann der Anwender fragmentierte, disperse Informationsbestände zusammenführen, um entscheidungsrelevante Informationen zur Verfügung zu haben?
- Wie können komplizierte Softwareanwendungen und damit verbundene Informationszusammenhänge mit Hilfe intelligenter, menschengerechter Benutzeroberflächen einfacher bedient und durchforstet werden?
- An welchen Stellen sollten vorerst noch fehlende Kontextinformationen zur effizienteren Suche nach Informationen und zur Steigerung der Merkfähigkeit angeboten werden?

³³ Siehe dazu: <http://www.swsi.org/>, aufgerufen am 3.1.2006

³⁴ Natürlich sind hier sowohl der Anwender als auch die Anwenderin angesprochen.

2.3.2 Organisationsperspektive

- Wie kann die Suche nach Ansprechpartnern zum Zwecke der effizienteren Vernetzung von Wissensträgern gestaltet werden?
- Wie kann das Innovationspotential eines Unternehmens erschlossen werden, indem bereichsübergreifende Informationsflüsse und die bessere Integration des Wissens der Kunden realisiert wird?
- Wie kann Doppelarbeit im Unternehmen vermieden werden?

2.3.3 Infrastrukturperspektive

- Wie können disperse Datenbestände (im Web verteilte Datenbanken und Dokumente) integriert werden?
- Wie kann mit Hilfe entsprechender Metadatenstrukturen die Integration verschiedener Applikationen erreicht werden?
- Wie können inkompatible Klassifikationen und Ontologien aufeinander „gemappt“ werden?

Die Foci einzelner Communities (Social Software, Künstliche Intelligenz, Knowledge Visualisation, Business Process Management) bzw. Berufsfelder (Wissensmanager, Terminologen, Bibliothekare, Dokumentare, Betriebswirte, Informationsmanager, ...) auf die unterschiedlichen Fragestellungen bzw. die verschiedenen Perspektiven zueinander in Beziehung zu setzen – das dürfte also der Schlüssel für eine erfolgreiche Entwicklung des globalen Projektes „Semantic Web“ sein.

3 Fazit

Das Semantic Web ist Vision, Programm und Konzept zugleich. Eine differenzierte Auseinandersetzung mit dem Thema ist ohne eine vorausgehende begriffliche Klärung kaum zu leisten. Am Beispiel von Begriffen wie Semantik und Ontologie wird deutlich, wie sehr die kontextspezifische Bedeutung der Begriffe im Vergleich zum ihrem ideen- und erkenntnisgeschichtlichen Erbe differieren kann. Deshalb ist es notwendig – wie in jeder neu entstehenden Disziplin – die neuen Begriffe und ihre Verwendungsweise einzuführen und den begrifflichen Rahmen im Groben abzustecken.

Dieser Beitrag erhebt keinesfalls den Anspruch dies erschöpfend geleistet zu haben. Vielmehr ging es darum anhand zentraler Begriffe einen ersten Einstieg in das Thema zu ermöglichen und gleichzeitig am Beispiel der unterschiedlichen Perspektiven auf das Semantic Web (A-O-I-Modell) ein grobes Orientierungsraster vorzuschlagen, das helfen kann, den Diskurs

um das Semantic Web zu ordnen und das gegenseitige Verständnis zu fördern, wenn in einem hoch interdisziplinären Feld über Semantic Web gesprochen wird. Die zahlreichen Verweise auf andere Beiträge in diesem Band sollen es dem Leser ermöglichen, Anknüpfungspunkte für die vertiefende Auseinandersetzung zu finden und auf diese Weise die Reichhaltigkeit und Transdisziplinarität des Themas weiter auszuloten.

Literatur

1. Apke S., Dittman L. (2003). Analyse von Vorgehensmodellen aus dem Software, Knowledge und Ontologies Engineering. Projektbericht 1/2003, Projekt KOWIEN, Institut für Produktion und Industrielles Informationsmanagement, Universität Essen, Essen 2003
2. Antoniou G., van Harmelen F. (2004). A Semantic Web Primer. Cambridge: MIT Press
3. Balani N. (2005). The Future of the Web is Semantic. In: <http://www-128.ibm.com/developerworks/web/library/wa-semweb/>, aufgerufen am 5.12.2005
4. Berners-Lee T., Hendler J., Lassila O. (2001). The Semantic Web. In: Scientific American, May 2001
5. Berners-Lee T. (1998). Semantic Web Roadmap. In: <http://www.w3.org/DesignIssues/Semantic.html>, aufgerufen am 15.12.2005
6. Blumauer A., Pellegrini T. (2005). Nutzen des Semantic Web für wissensbasierte Unternehmen und Netzwerke: Ausgewählte Praxisbeispiele. Präsentationsunterlagen I-Know 2005, http://www.semantic-web.at/file_upload/1001_tmpphpnEiTFA.pdf, aufgerufen am 16.12.2005
7. Decker B., Finke I., John M., Joisten M., Schnalzer K., Voigt S., Wesoly M., Will M. (2005). Wissen und Information 2005. Fraunhofer-Wissensmanagement Community: Fraunhofer IRB Verlag
8. Fensel D. (2004) Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce, 2. Auflage, Springer-Verlag Berlin Heidelberg
9. Gruber T. (1993). A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition, 5, (2):199–220
10. Gruber T. (2004). Every Ontology is a Treaty. Interview for Semantic Web and Information Systems SIG of the Association for Information Systems. SIGSEMIS Bulletin, Volume 1, Issue 3. October, 2004
11. Gruber T. (2005). Folksonomy of Ontology: A Mash-up of Apples and Oranges. First on-Line conference on Metadata and Semantics Research (MTR'05). <http://mtr.sigsemis.org/>. Paper available at <http://tomgruber.org/writing/mtr05-ontology-of-folksonomy.htm>, aufgerufen am 16.12.2005
12. Lauth B., Sareiter J. (2005). Wissenschaftliche Erkenntnis. Eine ideengeschichtliche Einführung in die Wissenschaftstheorie. Paderborn: mentis Verlag
13. Mey J. L. (2005). Pragmatics. An Introduction. Oxford: Blackwell Publishing
14. Mika P. (2005). Ontologies Are Us: A Unified Model of Social Networks and Semantics, in: The Semantic Web – ISWC 2005: 4th International Semantic Web Conference, ISWC 2005, Galway, Ireland, November 6–10, 2005. Proceedings, Springer Verlag, 2005

15. Neumüller M. (2001) Hypertext Semiotics in the Commercialized Internet. Dissertation, Wirtschaftsuniversität Wien
16. Reinmann-Rothmeier G. (2001). Wissen managen: Das Münchener Modell, Forschungsbericht Nr. 131. München: Ludwig-Maximilians-Universität, Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie.
17. Riempp G. (2004). Integrierte Wissensmanagement-Systeme - Architektur und praktische Anwendung. Berlin: Springer Verlag
18. Semantic Interoperability Community of Practice (2005). Introducing Semantic Technologies and the Vision of the Semantic Web. White Paper Series Module 1, 5.4. In:
<http://colab.cim3.net/file/work/SICoP/WhitePaper/SICoP.WhitePaper.Module1.v5.4.kf.021605.doc>, aufgerufen am 15.12.2005
19. Semantik. In: <http://de.wikipedia.org/wiki/Semantik>, aufgerufen am 5.12.2005
20. Sowa J. F. (2000). Ontology, Metadata and Semiotics. In:
<http://www.jfsowa.com/ontology/ontometa.htm>, aufgerufen am 5.12.2005
21. Studer R., Schnurr H., Nierlich A. (2005). Semantik für die nächste Generation von Wissensmanagement,
<http://www.community-of-knowledge.de/pdf/f05.pdf>,
aufgerufen am 16.12.2005
22. Uschold M. (1996). Building Ontologies. Towards A Unified Methodology. In: Proc. Expert Systems 96, Cambridge, December 1996
23. Wordnet. In: <http://wordnet.princeton.edu/>, aufgerufen am 10.12.2005