

X . media . press



Tassilo Pellegrini
Andreas Blumauer (Hrsg.)

Semantic Web

Wege zur vernetzten Wissensgesellschaft

Mit 98 Abbildungen und 4 Tabellen

 Springer

Tassilo Pellegrini
Andreas Blumauer
Semantic Web School
Zentrum für Wissenstransfer
Lerchenfelder Gürtel 43
1160 Wien, Österreich
t.pellegrini@semantic-web.at

Bibliografische Information der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen
Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über
<http://dnb.ddb.de> abrufbar.

ISSN 1439-3107
ISBN-10 3-540-29324-8 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-29324-8 Springer Berlin Heidelberg New York

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland vom 9. September 1965 in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtsgesetzes.

Springer ist ein Unternehmen von Springer Science+Business Media
springer.de

© Springer-Verlag Berlin Heidelberg 2006
Printed in Germany

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften. Text und Abbildungen wurden mit größter Sorgfalt erarbeitet. Verlag und Autor können jedoch für eventuell verbliebene fehlerhafte Angaben und deren Folgen weder eine juristische Verantwortung noch irgendeine Haftung übernehmen.

Satz und Herstellung: LE-TeX, Jelonek, Schmidt & Vöckler GbR, Leipzig
Umschlaggestaltung: KünkelLopka Werbeagentur, Heidelberg
Gedruckt auf säurefreiem Papier 33/3100 YL - 5 4 3 2 1 0

Vorwort der Herausgeber

Tassilo Pellegrini, Andreas Blumauer

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

Anfang des Jahres 2005 ist der Springer Verlag an die Semantic Web School herangetreten, ein Lehrbuch über das Thema Semantic Web zu verfassen. Schnell war klar, dass ein Lehrbuch zu so einem umfangreichen Thema kaum aus der Perspektive einer einzelnen Person oder Institution zu erstellen ist, ohne dabei sträfliche Abstriche in der inhaltlichen Bandbreite in Kauf nehmen zu müssen. Auf die konzeptionelle Not folgte die gestalterische Tugend und die Idee, den ersten deutschsprachigen Sammelband über Semantic Web zu erstellen, mit dem Ziel, einen möglichst umfassenden – wenn auch nicht erschöpfenden – Einstieg in das Thema zu gewährleisten.

Nach knapp einem Jahr vereinigt der vorliegende Band 33 Beiträge von 57 Autoren aus 35 Institutionen zu einem virulenten und multidisziplinären Thema. Der Band richtet sich gleichermaßen an interessierte Laien und fachfremde Experten, die nicht nur aus der technischen Perspektive einen Zugang zum Thema suchen. Denn obwohl das Thema Semantic Web zu überwiegendem Maße ein technisches ist, sollen hier bewusst jene Aspekte angesprochen werden, die außerhalb einer ingenieurwissenschaftlichen Perspektive von Relevanz sind und vor allem die praktischen Aspekte semantischer Technologien adressieren. Dieser Anforderung wird durch die vielen Praxisbezüge und Anwendungsbeispiele innerhalb der einzelnen Beiträge Rechnung getragen.

Die Generalisten unter den Leserinnen und Lesern¹ werden ein spannendes Feld der inhaltlichen Auseinandersetzung entdecken, wenn es darum geht, die Umsetzung des Semantic Web von einer Vision zu einer lebensweltlichen Praxis zu beobachten. Gleichzeitig hoffen die Herausgeber, den unterschiedlichen Spezialisten in den Themenbereichen des Semantic Web

¹ Die Herausgeber sind bemüht den Richtlinien des geschlechtsneutralen Formulierens nachzukommen. Aus Gründen der Lesbarkeit bitten wir jedoch um Verständnis, wenn im Weiteren die männliche Schreibweise beibehalten wird.

die Vielfalt der inhaltlichen und konzeptionellen Zugänge zu eröffnen und so den Dialog zwischen den Disziplinen anzukurbeln.

Hierbei ist es den Herausgebern jedoch wichtig darauf hinzuweisen, das Semantic Web und semantische Technologien nicht als verheißungsvolles Allheilmittel der durch Informationstechnologien heraufbeschworenen Probleme und Herausforderungen zu betrachten. Ganz im Gegenteil plädieren die Herausgeber für eine verstärkte Auseinandersetzung mit dem Thema unter Einbeziehung einer großen Vielfalt an Experten aus den unterschiedlichsten Fachbereichen, die einen reflektierten und kritischen Beitrag zu den positiven *und* negativen Effekten semantischer Technologien beitragen sollen. Eine differenzierte Auseinandersetzung mit den Technikfolgen semantischer Technologien ist deshalb dringend geboten, hätte jedoch den Umfang dieses Bandes bei weitem gesprengt.

Wie eine Paneldiskussion auf der internationalen Konferenz „Semantics 2005“ in Wien gezeigt hat, ist die Debatte um die Regulierung semantischer Technologien bereits voll entfacht. Die Bandbreite erstreckt sich vom Wunsch einer expliziten Regulierung des Einsatzes semantischer Technologien und der damit generierten Inhalte („Web of Trust“) bis hin zur Forderung, durch das Semantic Web induzierte gesellschaftliche Veränderungen nicht explizit mit semantischen Technologien in Verbindung zu bringen, sondern als generelles IT-Thema wahrzunehmen. Die Entwicklung einer allgemein reflektierten und tragfähigen Position zur Regulierungsproblematik ist jedenfalls ein wichtiger und notwendiger Schritt für den gesellschaftlichen und kommerziellen Erfolg des Semantic Web.

Unabhängig von der Forderung nach mehr Technikreflexion identifizieren die Herausgeber drei große Trends, die einerseits den Einsatz semantischer Technologien erfordern und andererseits ihre Entwicklung beschleunigen und begünstigen (siehe Abbildung 1):

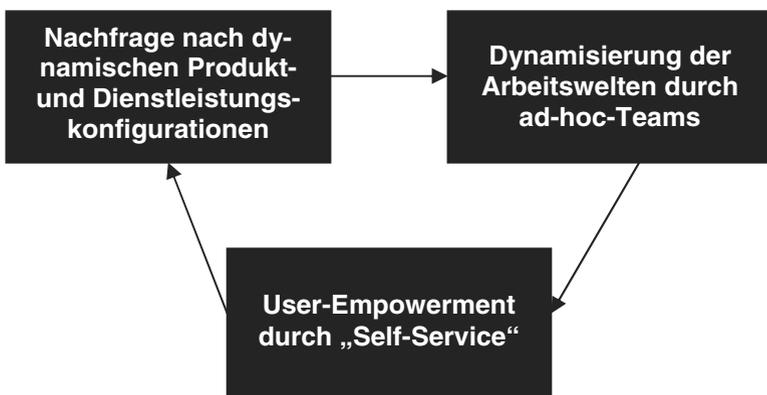


Abb. 1. Trends und arbeitstechnische Veränderungen, die mit der Entwicklung des Semantic Web in Wechselwirkung stehen

1. Die Nachfrage nach dynamischen Produkt- und Dienstleistungskonfigurationen steigt.
2. Die Dynamisierung der Arbeitswelten erfordert organisationsübergreifende Strukturen auf Basis von ad-hoc-Teambildungen.
3. Kunden und Anwender integrieren Informationsbestände zunehmend selbst unter Verwendung immer intelligenterer Dialogsysteme und Suchwerkzeuge („Self-Service“), um letztlich die in Punkt 1 angeführten Konfigurationen realisieren und nachfragen zu können.

Jeder dieser Trends manifestiert sich auf drei Ebenen:

1. Anwenderebene
2. Organisationsebene
3. Technikenebene

Diese drei Ebenen geben auch die Struktur vor, nach dem der Sammelband aufgebaut ist und damit auch versucht wird, der interdisziplinären Vielfalt des Themas gerecht zu werden.

Im *ersten Teil* wird neben der begrifflichen Klärung eine Reihe von Einstiegspunkten angeboten, ohne dass der Leser das Semantic Web in seiner Systematik und Funktionsweise kennen muss.

Im Beitrag von Andreas Blumauer und Tassilo Pellegrini werden die zentralen Begriffe rund um semantische Technologien vorgestellt und zentrale Konzepte überblicksartig dargestellt.

Die Arbeitsgruppe um Bernardi et al. leitet über in den Themenbereich der Arbeitsorganisation und diskutiert die Bedingungen für den Einsatz semantischer Technologien aus der Perspektive der Wissensarbeit.

Dem Thema Normen und Standards wurden sogar zwei Beiträge gewidmet. Während Christian Galinski die grundsätzliche Notwendigkeit von Normen zu Zwecken der Interoperabilität aus einer Top-Down-Perspektive beleuchtet, eröffnet Klaus Birkenbihl einen Einblick in die technischen Standards des Semantic Web aus der Bottom-Up-Perspektive des World Wide Web Consortiums (W3C).

Mit einem Beitrag zum Innovationsgrad semantischer Technologien in der ökonomischen Koordination betreten Michael Weber und Karl Fröschl weitgehend theoretisches Neuland und legen ein Fundament für weiterführende Auseinandersetzungen.

Agerundet wird der erste Teil noch mit einem Beitrag von Bernd Wohlkinger und Tassilo Pellegrini über die technologiepolitischen Dimensionen der Semantic Web Forschung in der europäischen Union.

Im *zweiten Teil* steht der Anwender des Semantic Web im Mittelpunkt, womit auch die erste Ebene der systematischen Auseinandersetzung mit semantischen Technologien angesprochen wird.

Nicola Henze zeigt auf, welchen Beitrag semantische Technologien für die Personalisierung von Informationssystemen leisten.

Stefanie Lindstaedt und Armin Ulbrich diskutieren die Möglichkeiten der Zusammenführung von Arbeiten und Lernen zu Zwecken der Kompetenzentwicklung in Arbeitsprozessen.

Leo Sauermaun stellt daraufhin mit der Metapher des „Semantic Desktop“ ein innovatives Konzept für den Arbeitsplatz der Zukunft vor und fragt – nicht ohne eine gewisse Ironie –, ob dieser Arbeitsplatz tatsächlich auf einen physischen Ort begrenzt ist.

Mark Buzinkay zeigt aus einer historischen Perspektive, wie semantische Strukturen die Navigation sowohl im Web als auch auf einzelnen Webseiten verändert haben und noch werden.

Michael Schuster und Dieter Rappold adressieren die Konvergenz von Social Software und Semantic Web entlang der persönlichen Aneignung von Informationstechnologien zu Zwecken der sozialen Vernetzung.

Remo Burkhard plädiert dafür, Wissensvisualisierung als Brückenfunktion zwischen technischer Infrastruktur und Nutzer wahrzunehmen und demonstriert das Potential der Wissensvisualisierung zur zielgruppengerechten Kommunikation komplexer Zusammenhänge.

Abschließend zeigt Gabriele Sauberer, welche Informationskompetenzen und Schlüsselqualifikationen in der modernen Informationsgesellschaft von Bedeutung sein werden, in der der Einsatz semantische Technologien zur täglichen Wissensarbeit gehören wird.

Der *dritte Teil* des Bandes thematisiert die organisationalen Dimensionen des Semantic Web und demonstriert unter dem Stichwort „Wissensmanagement“ eine Reihe von Konzepten und Anwendungen im betrieblichen und kollaborativen Umgang mit Information.

Der Beitrag von Andreas Blumauer und Thomas Fundneider bietet einen Überblick über den Einsatz semantischer Technologien am Beispiel eines integrierten Wissensmanagement-Systems.

Michael John und Jörg Drescher zeichnen den historischen Entwicklungsprozess des IT-Einsatzes für das Management von Informations- und Wissensprozessen im betrieblichen Kontext.

Vor dem Hintergrund der betrieblichen Veränderungen durch Globalisierung und angeheizten Wettbewerb zeigt Heiko Beier, welche Rollen, Prozesse und Instrumente in wissensbasierten Organisationen die effiziente Nutzung von Wissen unterstützen.

Mit dem Konzept des kollaborativen Wissensmanagement präsentiert das Autorenteam Schmitz et al. einen innovativen Wissensmanagement-Ansatz auf Peer-to-Peer-Basis mit dem Ziel der kollaborativen Einbindung und Pflege von dezentralisierten Wissensbasen.

York Sure und Christoph Tempich demonstrieren anhand der Modellierungsmethode DILIGENT, welchen Beitrag Ontologien bei der Wissensvernetzung in Organisationen leisten können.

Hannes Werthner und Michael Borovicka adressieren die Bedeutung semantischer Technologien für eCommerce und demonstrieren am Beispiel HARMONISE deren Einsatz im Bereich des eTourismus.

Erweitert wird diese Perspektive durch den Beitrag von Fill et al., in dem das Zusammenspiel zwischen Web-Services und Geschäftsprozessen aus der Perspektive der Wirtschaftsinformatik analysiert wird.

Abschließend präsentiert das Autorenteam Angele et al. eine Reihe von realisierten Anwendungen auf Basis semantischer Technologien und identifiziert kritische Faktoren für deren Einsatz.

Im *vierten Teil* des Bandes stehen die technischen und infrastrukturellen Aspekte im Mittelpunkt des Interesses, die für den Aufbau und Betrieb semantischer Systeme von Relevanz sind.

Wolfgang Kienreich und Markus Strohmaier identifizieren die Wissensmodellierung als Basis für den Einsatz semantischer Technologien für das Knowledge Engineering und stellen zwei grundlegende Modellierungsparadigmen vor.

Andreas Koller argumentiert, dass die strukturierte Ablage von Content in Content Management Systemen den Lift-Off des Semantic Web stützen wird und zeigt eine Reihe von einfachen Maßnahmen auf, wie CMS Semantic Web tauglich gemacht werden können.

Alois Reitbauer gibt einen leicht verständlichen Überblick über technische Fragestellungen der IT-Integration und demonstriert anhand von Beispielen die Vorteile semantischer Technologien gegenüber konventionellen Methoden.

Gerald Reif veranschaulicht die Einsatzgebiete und Leistungsfähigkeit der semantischen Annotation und stellt Tools vor, die den Nutzer bei der Dokumentenverschlagnwortung unterstützen.

Robert Baumgartner stellt die Funktionsweise von Wrappertechnologien zur Extraktion von Daten aus unstrukturierten Dokumenten vor und demonstriert den Nutzen am Beispiel eines B2B-Szenarios.

Michael Granitzer bietet einen Überblick über statistische Verfahren der Textanalyse und zeigt, welchen Beitrag diese zur Wartung von Ontologien leisten können.

Gerhard Budin geht auf die zentrale Rolle des Terminologiemanagements bei der Ordnung und Intersubjektivierung komplexer Wissensstrukturen ein und gibt Anleitung für die Entwicklung von terminologischen Metamodellen.

Marc Ehrig und Rudi Studer thematisieren Prinzipien und Herausforderungen der semantischen Integration von Ontologien zu Zwecken der Herstellung von Interoperabilität von Web Services.

Wolfgang May gibt eine Einführung in das Thema Reasoning im und für das Semantic Web und zeigt auf, welche Mechanismen und Konzepte in naher Zukunft für das Semantic Web relevant werden.

Abschließend führt die Autorengruppe um Polleres et al. in das junge Thema der semantischen Beschreibung von Web Services ein und adressiert Fragestellungen der Service Komposition und Automatisierung von Geschäftsprozessen.

In einem *Nachwort* widmet sich Rafael Capurro der Frage, wie es in Zeiten eines auftauchenden semantischen Web um die philosophische Hermeneutik bestellt ist. Und er kommt zu dem Schluss, dass das Semantic Web als ein weltpolitisches Projekt verstanden werden sollte, das zu wichtig ist, um es alleine den Technikern oder den Politikern zu überlassen.

Wien, am 25. Jänner 2006

Danksagung

Die Herausgeber des Bandes bedanken sich bei den Autorinnen und Autoren dieses Sammelbandes und ihrer Bereitschaft dieses Projekt durch ihre Expertise mit zu tragen. Besonderer Dank gilt zudem der kompetenten und unkomplizierten Unterstützung durch den Springer Verlag in allen produktionstechnischen Fragen und dem unermüdlichen Einsatz von Eva Pellegrini, die durch ihre Erfahrung und Kompetenz im Zuge des Lektorats einen wichtigen Beitrag zu der hohen formalen Qualität dieses Buches geleistet hat.

Inhaltsverzeichnis

Semantic Web – Geschichte und Ausblick einer Vision.....	1
<i>Klaus Tochtermann, Hermann Maurer</i>	

Einstiegspunkte

Semantic Web und semantische Technologien: Zentrale Begriffe und Unterscheidungen.....	9
<i>Andreas Blumauer, Tassilo Pellegrini</i>	

Komplexe Arbeitswelten in der Wissensgesellschaft	27
<i>Ansgar Bernardi, Harald Holz, Heiko Maus und Ludger van Elst</i>	

Wozu Normen? Wozu semantische Interoperabilität?	47
<i>Christian Galinski</i>	

Standards für das Semantic Web.....	73
<i>Klaus Birkenbihl</i>	

Das Semantic Web als Innovation in der ökonomischen Koordination.....	89
<i>Michael Weber, Karl Fröschl</i>	

Semantic Systems Technologiepolitik in der Europäischen Union	115
<i>Bernd Wohlkinger, Tassilo Pellegrini</i>	

Anwenderperspektive

Personalisierbare Informationssysteme im Semantic Web	135
<i>Nicola Henze</i>	

Integration von Arbeiten und Lernen – Kompetenzentwicklung in Arbeitsprozessen	147
<i>Stefanie N. Lindstaedt, Armin Ulbrich</i>	
Semantic Desktop – Der Arbeitsplatz der Zukunft	161
<i>Leo Saueremann</i>	
Finden und gefunden werden – Funneling im Semantic Web	177
<i>Mark Buzinkay</i>	
Social Semantic Software – was soziale Dynamik im Semantic Web auslöst	189
<i>Michael Schuster, Dieter Rappold</i>	
Knowledge Visualization: Die nächste Herausforderung für Semantic Web Forschende?	201
<i>Remo Burkhard</i>	
Informationskompetenz und Schlüsselqualifikationen in der Wissensarbeit	213
<i>Gabriele Sauberer</i>	

Organisationale Perspektive

Semantische Technologien in integrierten Wissensmanagement-Systemen	227
<i>Andreas Blumauer, Thomas Fundneider</i>	
Semantische Technologien im Informations- und Wissensmanagement: Geschichte, Anwendungen und Ausblick.....	241
<i>Michael John, Jörg Drescher</i>	
Betriebliches Wissensmanagement: Rollen, Prozesse, Instrumente	257
<i>Heiko Beier</i>	
Kollaboratives Wissensmanagement.....	273
<i>Christoph Schmitz, Andreas Hotho, Robert Jäschke und Gerd Stumme</i>	
Wissensvernetzung in Organisationen	291
<i>York Sure, Christoph Tempich</i>	

E-commerce und Semantic Web307
Hannes Werthner, Michael Borovicka

Web-Services und Geschäftsprozesse im Semantic Web321
Hans-Georg Fill, Dimitris Karagiannis und Johannes Lischka

Anwendungen und Good Practices Semantischer Technologien.....337
*Jürgen Angele, Eddie Mönch, Andreas Nierlich, Heiko Rudat
 und Hans-Peter Schnurr*

Methoden und Technische Infrastruktur

Wissensmodellierung – Basis für die Anwendung
 semantischer Technologien.....359
Wolfgang Kienreich, Markus Strohmaier

Web Content und Content Management Systeme: Ohne Struktur
 kein Semantic Web!373
Andreas Koller

IT Konsolidierung und Informationsintegration387
Alois Reitbauer

Semantische Annotation.....405
Gerald Reif

Methoden und Werkzeuge zur Webdatenextraktion419
Robert Baumgartner

Statistische Verfahren der Textanalyse437
Michael Granitzer

Kommunikation in Netzwerken – Terminologiemanagement453
Gerhard Budin

Wissensvernetzung durch Ontologien.....469
Marc Ehrig, Rudi Studer

Reasoning im und für das Semantic Web485
Wolfgang May

XIV Inhaltsverzeichnis

Semantische Beschreibung von Web Services.....505
Axel Polleres, Holger Lausen und Rubén Lara

Nachwort

Hermeneutik *revisited*527
Rafael Capurro

Semantic Web – Geschichte und Ausblick einer Vision

Klaus Tochtermann, Hermann Maurer

Know-Center und Technische Universität, Graz, Österreich;

ktochter@know-center.at

Technische Universität, Graz, Österreich;

hmaurer@iicm.tu-graz.ac.at

1 Am Anfang war der Wunsch

Stellen Sie sich folgendes Szenario vor. Frau Maler ist auf der Suche nach einem homöopathischen Arzt in ihrer Heimatstadt. Über ihre bevorzugte Suchmaschine gibt Frau Maler die Suchbegriffe *Arzt*, *Homöopathie* und *Stadt Graz* ein. Als Ergebnis wirft die Suchmaschine eine lange Liste an Links aus. Frau Maler sucht sich die Links aus, die direkt auf Ordinationen zeigen. Natürlich sind nun auch Ärzte in ihrer Liste, die sich *nicht* mit Homöopathie beschäftigen. Diese sortiert Frau Maler aus, genauso wie jene, die nicht mit den öffentlichen Verkehrsmitteln des örtlichen Verbunds erreichbar sind. Dazu muss sie natürlich zum Teil zuvor im Stadtplan nach der genauen Lage der Adressen suchen. Schließlich sortiert sie die Ärzte aus, die keine Bewertungen besitzen bzw. die nicht zumindest als gut bewertet wurden. Für die verbleibenden Ärzte sichtet Frau Maler noch die Ordinationszeiten und vergleicht diese mit ihrem Kalender. Nach 20 Minuten hat Frau Maler schließlich drei in Frage kommende Ärzte gefunden. Ob sie welche übersehen hat, weiß sie freilich nicht.

Geht das ganze nicht einfacher, schneller, besser?

Und jetzt versetzen Sie sich ein paar Jahre in die Zukunft und stellen sich einen Projektmanager in einem transnationalen Konzern vor, der zusammen mit zwei weiteren Großkonzernen, fünfzehn kleineren Firmen sowie einem breiten Netzwerk selbständiger Spezialisten an einem bahnbrechenden Weltraumtourismusprojekt arbeitet. Ein unerwartetes Gerichtsurteil hat soeben die gesamten Rahmenbedingungen des Projekts gehörig ins Wanken gebracht. Und in der *real time economy* erwarten seine Projektpartner,

Investoren und nicht zuletzt sein Chef schnelle und sichere Antworten – und das erfordert nicht nur Recherchen, die die Arztsuche von Frau Maler wie ein Kinderspiel aussehen lassen, sondern darüber hinaus auch noch enge virtuelle Zusammenarbeit über mehrere Zeitzonen und Fachdisziplinen hinweg ...

muss das ganze nicht einfacher, schneller, besser gehen als heute?

2 Es folgte die Tat

Die Antwort von Tim Berners-Lee, dem „Erfinder“ des World Wide Web, und einer recht großen und beständig wachsenden Gruppe von Forschern, Technikern und zunehmend auch Anwendern lautet: Ja. Und wir wissen auch, wie.

An dieser Stelle wird dann der Begriff des Semantic Web ins Spiel gebracht. Die grundlegende Idee besteht darin, Inhalte im Web so anzureichern, dass sie nicht nur für Menschen verständlich sind, sondern auch von Maschinen zumindest soweit erfasst werden können, dass Automatisierung auch auf der Ebene der *Bedeutung* möglich wird. Wie und wodurch das im Einzelnen geschieht, ist Gegenstand der Beiträge in diesem Band, die den Bogen von den prägenden Rahmenbedingungen – den Arbeitswelten in der Wissensgesellschaft – bis hin zu zukünftigen intelligenten Diensten – den Semantic Web Services – spannen.

Wodurch unterscheidet sich nun das Semantic Web als Konzept von bisherigen Ansätzen?

Der Versuch, Maschinen auf semantischer Ebene arbeiten zu lassen, ist definitiv nicht neu. Die Disziplin der künstlichen Intelligenz beschäftigt sich bereits seit mehreren Jahrzehnten mit Themen wie automatischem Schlussfolgern, und geht teilweise in ihrem Anspruch sogar noch weit über den des Semantic Web hinaus: Verfechter der starken künstlichen Intelligenz wollen die Leistungen des menschlichen Gehirns insgesamt nachbilden und übertreffen. Als Teil von Knowledge Engineering werden auch schon lange Methoden praktiziert, Wissen überhaupt erst einmal strukturiert aus Köpfen in elektronische Form zu überführen. Tatsächlich baut das Semantic Web wesentlich auf Wissensrepräsentationsformen dieser Herkunft auf. Die großen Erfolge der künstlichen Intelligenz beschränken sich jedoch bisher auf einzelne, relativ isolierte inhaltliche Domänen.

Die halb- oder vollautomatische Extraktion von Semantik aus gegebenen, nicht oder nur teilweise strukturierten Datenbeständen ist ebenfalls nicht wirklich neu. Auch das gab es bereits, bevor die Idee des Semantic Web überhaupt geboren wurde. Auf diesem Gebiet werden – vielleicht sogar noch mehr als auf dem der künstlichen Intelligenz – auch laufend

neue Fortschritte erzielt. Und selbstverständlich wird heute die automatische Informationsextraktion ebenfalls als wesentlicher Baustein für das Semantic Web gesehen. Mit ihrer Hilfe kann man innerhalb kürzester Zeit eine große Menge an Inhalten zumindest bis zu einem gewissen Grad semantisch anreichern, viel schneller als die manuelle Wissensmodellierung das vermag.

Auch die Verteiltheit von Information in großen Mengen an sich ist natürlich nicht neu – der Name Semantic Web selbst geht ja gerade auf das Paradebeispiel dafür, nämlich das World Wide Web, zurück.

Neu ist jedoch die Kombination dieser Elemente. Das Semantic Web ist eine effektivere Zusammenführung auf einheitlicher Basis mit breiterem Anspruch. Einflüsse, die Bedarfe widerspiegeln – wie eben große Mengen verteilter und heterogener Informationen, aber auch ständig zunehmende wirtschaftliche Anforderungen – treffen auf bereits aus früheren Anstrengungen vorliegende technologische Lösungselemente, die nun adaptiert und von einer bestimmten Vision getrieben gezielt weiterentwickelt werden. Die Basis dafür bilden die Entwicklung entsprechender Standards, aber auch Strategien, um den tatsächlichen Erfolg des gesamten Vorhabens sicherzustellen.

Eine wichtige Rolle spielt dabei Dezentralität, und zwar in zweierlei Hinsicht: Einerseits handelt es sich nicht um ein monolithisches, zentral gesteuertes Forschungsprogramm, sondern vielmehr um eine Bewegung, einen Trend, der von Forschern, Entwicklern und Anwendern eigenständig aufgegriffen wird. Andererseits beschreibt auch die Vision des Semantic Web selbst ein dezentrales Netz, das mit einem Minimum an zentral koordinierten Elementen (einer Handvoll technischer Standards) auskommt, wohingegen die Inhalte, und zwar einschließlich ihrer semantischen Dimension (Ontologien – explizite und maschinell verarbeitbare Beschreibungen von Domänenmodellen), verteilt entstehen und sich bloß der Standards bedienen, um Interoperabilität sicherzustellen.

3 Und irgendwann die Reflexion?

Obwohl seit dem populären Debüt der Vision des Semantic Web vor rund fünf Jahren durchaus beachtliche Fortschritte erzielt wurden, steckt das Semantic Web selbst – also die Realisierung der Vision – noch in den Kinderschuhen. Das heißt aber gleichzeitig: Sie ist noch nicht in Stein gemeißelt. Noch gestaltbar. Eine offene Chance. Und Verantwortung.

Das schreit natürlich geradezu nach einer kritischen Reflexion, die über die technische Ebene hinausgeht, möchte man meinen. Dieser Schrei wird jedoch erstaunlich selten wahrgenommen. Der vorliegende Band soll daher

dazu beitragen, eine solche Reflexion nicht erst irgendwann, sondern jetzt anzustoßen, und zwar auch auf ökonomischer, sozialer und kultureller Ebene. Einige Ansatzpunkte dafür seien gleich hier angerissen, damit sie bei der Lektüre der einzelnen Beiträge stets mitschwingen können.

Selbst wenn man sich der Reflexion von technischer Seite nähert, kommen schnell Fragen auf, deren Beantwortung einen weiteren Blickwinkel verlangt. Etwa im Kontext der vielen entstehenden Ontologien: Wie stellen wir eine vernünftige Wiederverwendung sicher? Das ist nicht nur eine Frage des Auffindens, sondern auch eine von subjektiver Qualitätsbeurteilung, Vertrauen und Feedbackmechanismen. Und unter welchen Umständen ist welche Art der semantischen Anreicherung von Inhalten am sinnvollsten – Ontologie-Entwicklung a priori oder Informationsextraktion a posteriori? Das Vorgeben von Strukturen mag eine gewisse Konsistenz sicherstellen, verschließt dabei aber inhaltlich immer auch so manche Tür. Die immer wieder geführte Debatte über den Umgang mit semi-strukturierten Daten legt auch eine andere Frage nahe: Welcher Grad an Formalisierung ist überhaupt letztendlich am sinnvollsten? Die Antwort wird natürlich nicht zuletzt von den im Einzelfall angepeilten Anwendungsszenarien abhängen (was soll mit den Inhalten geschehen – automatische Inferenz?), darf aber auch den psychologischen und sozialen Prozess der Entstehung dieser Inhalte nicht außer Acht lassen (ist eine weitgehende Formalisierung überhaupt möglich und auch unter Berücksichtigung nicht-technischer Rahmenbedingungen akzeptabel?). In eine ähnliche Kerbe schlägt die Bemerkung, dass die Benutzerschnittstellen der gängigen Werkzeuge im Semantic Web Umfeld noch lange nicht ein Niveau erreicht haben, das der gewünschten breiteren Akzeptanz wirklich dienlich wäre. Wie kann man von einer breiteren Bevölkerungsschicht erwarten, sich etwa mit Ontologie-Editoren auseinanderzusetzen, wenn selbst Profis, denen man wesentlich mehr Komplexität zumuten kann, heute immer noch unglücklich damit sind?

Bleibt man bei der Anwendungsperspektive, macht aber den technologischen Blickwinkel etwas weiter auf, so stellt sich schnell die Frage nach dem Zusammenhang zwischen dem Semantic Web und anderen aktuellen Trends, die unter Schlagworten wie Web 2.0 und Social Software zusammengefasst werden. Inwieweit kann das Semantic Web zu einer „Architektur der Partizipation“ beitragen? Erste Schritte sind mit semantischen Blogs und Wikis ja bereits unternommen. Und immerhin sind die wirklich auf breiter Front erfolgreichen Systeme am Web ihrer Natur nach als „bottom-up“ zu klassifizieren. Das Semantic Web ist zwar durchaus auch so gedacht, aber wie sehr wird die (notwendige?) Komplexität eines voll ausgeprägten Semantic Web seiner weiteren Verbreitung im Weg stehen? Was kann für das zukünftige Semantic Web aus Erfahrungen auf dem Peer-to-Peer Sektor gelernt werden? Und wie können (und sollen) umgekehrt semantische Technologien neue Arten von Peer-to-Peer Anwendungen ermöglichen, die

nicht nur technisch mächtiger sind, sondern in Folge womöglich auch gewisse soziale Machtverschiebungen nach sich ziehen? Außerdem lenkt der Blick auf andere Trends die Aufmerksamkeit auch auf sich ändernde Erwartungshaltungen der Benutzer – und zwar nicht nur in den Marktsegmenten, denen diese Trends üblicherweise zugeordnet werden. Wenn beispielsweise im Consumer Internet die oft ziemlich radikale Vereinfachung des Benutzer-Erlebnisses an der Tagesordnung steht (technisch momentan z. B. durch AJAX, bei vielen neuen Webdiensten aber auch funktional), wie wird sich das auf die Erwartungshaltungen von Wissensarbeitern in Unternehmen auswirken? Schließlich sind die beiden Gruppen alles andere als scharf getrennt: Jeder Wissensarbeiter ist in seiner Freizeit natürlich auch Konsument. Aber Unternehmensinfrastrukturen haben Planungshorizonte und Roadmaps – sie können mit den Entwicklungen im Consumer Bereich oft nicht Schritt halten.

Selbst wenn sich die Steine am Weg zur vollen technischen und inhaltlichen Entfaltung des Semantic Web als gar nicht so groß oder zumindest als einigermaßen leicht umgehbar entpuppen, bleiben dennoch grundsätzliche Fragen hinsichtlich dieser Vision selbst. Eine Menge an neuen und zweifellos ausgesprochen nützlichen Anwendungen wird sich auftun. Aber gerade das hat vielleicht auch eine Kehrseite: Kann der unbedarfte Benutzer die Kontrolle behalten oder wird er von der zunehmenden Automatisierung überrollt? Auf der technischen Ebene mag wohl von loser Kopplung (etwa von Web Services) die Rede sein, aber auf der Anwendungsebene ermöglicht das Semantic Web durchaus eine engere Kopplung von Abläufen. Gehen hier ökonomisch und sozial wichtige Pufferzonen verloren? Inwieweit können und wollen wir mit einer vom Semantic Web noch verschärften weiteren Beschleunigung unserer Lebenswelt Schritt halten? Und selbst wenn die Vision einer weitgehenden Automatisierung vieler Prozesse auf Bedeutungsebene perfekt aufgeht, laufen wir gerade dann nicht auch Gefahr, vernetztes Denken und Handeln zu verlernen?

Die von Rafael Capurro in seinem Nachwort zu diesem Band bereits angesprochene Wichtigkeit des beständigen Dialogs, der Vorläufigkeit jeglicher vorgeblichen Erkenntnis, des Offenhaltens von Perspektiven, wirft hier ebenfalls einen wesentlichen Fragenkomplex auf. An die Unbeständigkeit von Inhalten am World Wide Web haben wir uns ja bis zu einem gewissen Grad gewöhnt, aber wie weit ist das Semantic Web in der Lage, mit der aus dieser Perspektive nicht nur wünschenswerten, sondern geradezu notwendigen Fluidität von semantischen Strukturen (speziell: Ontologien) umzugehen? Die Idee des Semantic Web besteht ja gerade darin, immer mehr bisher essentiell menschliche Prozesse zu automatisieren. Wie weit wollen wir die Flexibilität unserer mentalen Modelle, und

damit letztendlich unserer gesamten Weltdeutung, der Automatisierung überantworten?

Schließlich muss natürlich auch die Kosten/Nutzen Frage gestellt werden dürfen. Und gleich anschließend: Wessen Kosten, und wessen Nutzen eigentlich? Ein breiter Blick ist dazu unerlässlich. Dieser Band soll dabei helfen.

Danksagung

Das Know-Center wird als Kompetenzzentrum innerhalb des Österreichischen Kompetenzzentrenprogramms Kplus (www.kplus.at) unter der Schirmherrschaft des Österreichischen Ministeriums für Verkehr, Innovation und Technologie gefördert. Schließlich möchten wir Dr. Gisela Dörsinger und Dr. Herwig Rollett für ihre Hilfe bei den Recherchen zu und der Erstellung von diesem Beitrag unseren Dank aussprechen.

Einstiegspunkte

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 eingenommen 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 dieselbe 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 Ausrufung 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 dabei 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 Prozess erfolgen.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“ beginnt⁷ [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. „aufzuhalsen“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

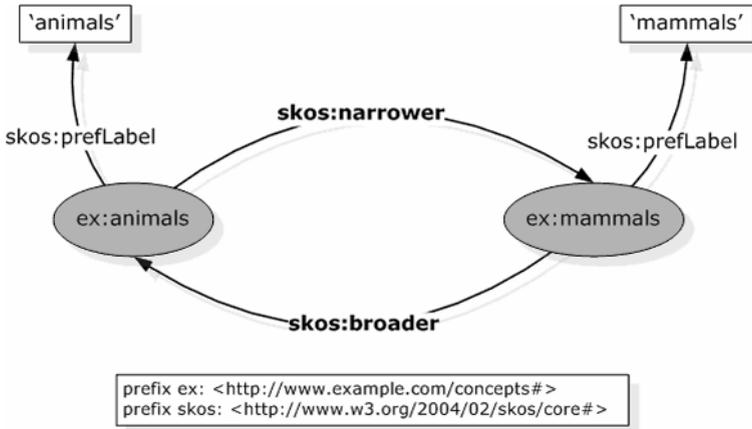


Abb. 1. Beispiel einer Visualisierung eines RDF-Dokuments

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-swbp-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://twoday.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:

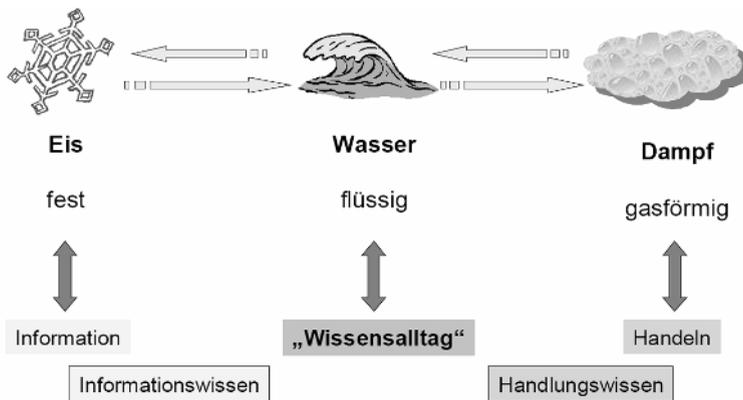


Abb. 2. „Wasseranalogie des Wissens“

Gasförmiges Wissen wird sozusagen laufend – 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 zufälligerweise 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:

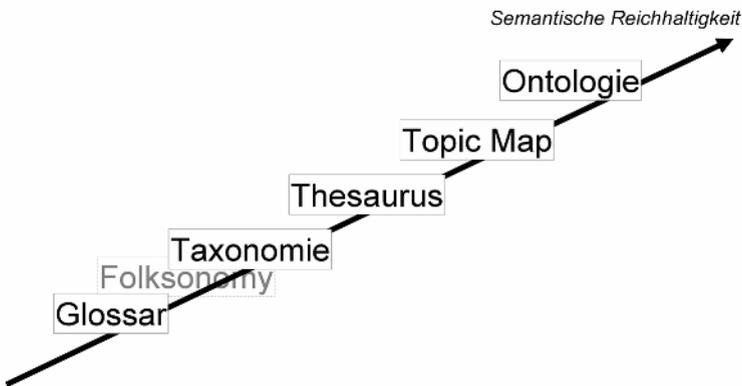


Abb. 3. „Semantische Treppe“

Von „Ontologien“ im eigentlichen Sinn sollte daher erst gesprochen 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 wiedererneuerbare 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 zu gestalten²² 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 bestehenden 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 Datenbestände 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.