

X . media . press



Andreas Blumauer · Tassilo Pellegrini

Social Semantic Web

Web 2.0 – Was nun?

Andreas Blumauer
Semantic Web Company
Lerchenfelder Gürtel 43
1160 Wien
Österreich
a.blumauer@semantic-web.at

Tassilo Pellegrini
Semantic Web Company
Lerchenfelder Gürtel 43
1160 Wien
Österreich
t.pellegrini@semantic-web.at

ISBN 978-3-540-72215-1

e-ISBN 978-3-540-72216-8

DOI 10.1007/978-3-540-72216-8

X.media.press ISSN 1439-3107

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© 2009 Springer-Verlag Berlin Heidelberg

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funk-sendung, 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. Zuwider-handlungen unterliegen den Strafbestimmungen des Urheberrechtsgesetzes.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk be-rechtigt 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.

Einbandgestaltung: KinkelLopka, Heidelberg

Gedruckt auf säurefreiem Papier

9 8 7 6 5 4 3 2 1

springer.de

Vorwort der Herausgeber

Tassilo Pellegrini und Andreas Blumauer

Semantic Web Company, Wien, Austria
{t.pellegrini; a.blumauer}@semantic-web.at

Knapp zwei Jahre nach der ersten Publikation „Semantic Web – Wege zur vernetzten Wissensgesellschaft“¹ liefern wir mit diesem Buch ein „Update“ zur aktuellen Debatte über die Entwicklung und Perspektiven des Internets im Allgemeinen und des World Wide Web im Speziellen.

Entgegen dem vorherrschenden Trend mittels Modebegriffen und Versions-Metaphern das vermeintlich „Neue“ möglichst plakativ und oftmals undifferenziert den interessierten Beobachtern und Aktivisten zu präsentieren, soll mit dem aktuellen Band eine differenzierte Diskussion über ein Phänomen angeregt werden, das mit dem Zusammenwachsen von Web 2.0 und Semantic Web beschrieben und unter dem Begriff „Social Semantic Web“ zusammengefasst werden kann. Hierbei ist es uns wichtig darauf hinzuweisen, dass der Begriff nicht von Dauer sein kann, zumal – so die Meinung der Herausgeber – in absehbarer Zeit viele der in diesem Band vorgestellten Konzepte und Technologien als Selbstverständlichkeit wahrgenommen werden und sich damit die Einführung neuer Kunstbegriffe weitgehend erübrigt. Unser Buchtitel soll vielmehr als Paraphrase auf einen aktuellen Transformations- und Konvergenzprozess verstanden werden, der keinen Zustand, sondern einen evolutionären Pfad beschreibt, welcher das Web in seinen Funktionen sowohl als technische Infrastruktur als auch als Ort der sozialen Interaktion in den kommenden Jahren prägen, verändern und modernisieren wird.

Das Buch ist dem interessierten Laien und themenfremden Experten gewidmet. Es dient als Einstiegswerk in die breite Fülle an technologischen und methodischen Trends der Web-Entwicklung und soll den LeserInnen in Ansätzen Antworten auf die Frage geben: „Web 2.0 – Was nun?“

¹ Pellegrini, Tassilo; Blumauer, Andreas (2006). Semantic Web. Wege zur vernetzten Wissensgesellschaft. Berlin: Springer Verlag

Hierbei setzen wir voraus, dass die LeserInnen bereits ein grundlegendes Vorwissen in Bezug auf Web 2.0 und Semantic Web mitbringen, weshalb wir insbesondere in Bezug auf die (technischen) Grundlagen des Semantic Web auf eine ausführliche Einführung verzichtet haben, da dies nicht zuletzt mit unserem Vorgängerband abgedeckt wurde.² Das Buch soll vielmehr unterschiedliche Perspektiven eröffnen, worauf bei der Konzeption, Entwicklung und Anwendung der nächsten Generation von Social Web-Anwendungen zu achten ist, ohne sich zu sehr mit den technischen Feinheiten und Detailproblemen zu belasten. Die umfangreichen Quellen, auf denen die einzelnen Beiträge basieren, sind jedoch für eine vertiefende Auseinandersetzung bestens geeignet.

Bei der Zusammenstellung der Beiträge wurde auf hohe Praxisrelevanz geachtet, was dazu geführt hat, dass ausschließlich alle der vorgestellten Technologien, Methoden und Applikationen sich bereits in Anwendung befinden und live getestet werden können. Außerdem haben wir am Ende des Bandes einen umfangreichen Index angefügt, der die Orientierung in diesem neuen und umfangreichen Themenfeld erleichtern soll, um nicht zuletzt die Reichhaltigkeit und Tiefe der einzelnen Beiträge besser zu erschließen. Zusätzlich haben wir den Band in vier Abschnitte eingeteilt, um auch einen besseren Überblick über die Foki der einzelnen Beiträge geben zu können.

Abschnitt 1 bietet einen niedrigschwelligen Einstieg in das Thema Web 2.0 und seine Entwicklungsperspektiven.

Andreas Blumauer und Tassilo Pellegrini beginnen mit einer grundsätzlichen Verortung des Phänomens „Social Semantic Web“ vor dem Hintergrund der Konvergenz von Web 2.0 und Semantic Web.

Sonja Bettel wirft einen eloquenten Blick auf die Entstehungsgeschichte des Phänomens Web 2.0 und fragt anhand plakativer Beispiele nach den Konsequenzen für Gesellschaft, Demokratie und Medien.

Alexander Raabe führt den Begriff der Social Software ein und zeigt am Beispiel unterschiedlicher Einsatzmöglichkeiten, wie sich vor dem Hintergrund der zunehmenden Mobilisierung, aber auch vor dem Hintergrund der wachsenden Metadatenbewirtschaftung – z. B. mittels Semantic Web Technologien – das organisationale, wirtschaftliche und soziale Gefüge verändern könnte.

Alexander Stocker und Klaus Tochtermann geben einen Überblick über die aktuellen Technologien und Anwendungen des Web 2.0 und legen damit die Basis für die weiterführende technologische Diskussion über die Transformation des Social Web zu einem Social Semantic Web.

² Pellegrini, Tassilo; Blumauer, Andreas (2006). Semantic Web. Wege zur vernetzten Wissensgesellschaft. Berlin: Springer Verlag

Aufbauend auf der technologischen Betrachtung stellt **Jörg Linder** die Frage nach der Usability von Rich Internet Applications und thematisiert gleichzeitig die Veränderungen in der hohen Schule des Web-Designs vor dem Hintergrund von AJAX und den darauf basierenden Gestaltungsmöglichkeiten.

Anupriya Ankolekar, Markus Krötzsch, Than Tran und Denny Vrandeic schließen den ersten Abschnitt mit drei Hypothesen über die Konvergenz von Web 2.0 und Semantic Web vor dem Hintergrund der zunehmenden Datenverfügbarkeit am Netz, der Möglichkeit darauf neue Mehrwertdienste zu konzipieren und den unzureichenden Möglichkeiten, diese mit „reinen“ Web 2.0 Technologien zu realisieren.

Abschnitt 2 bietet einen Überblick über die zentralen Technologien und Methoden des Social Semantic Web.

Dazu stellt **Barbara Geyer-Hayden** unterschiedliche Ansätze der Wissensmodellierung vor, und zeigt auf, welche Rolle diese für die Beschreibung von Web-Ressourcen und die Veröffentlichung von strukturierten Daten am Netz spielen.

Michael Hausenblas zeigt an den Beispielen Mikroformate und RDFa unterschiedliche Strategien, um bestehende Web-Ressourcen mit Metadaten zu annotieren, in HTML zu integrieren und damit die maschinelle Erschließbarkeit bestehender Web-Ressourcen zu verbessern.

Armin Ulbrich und Patrick Höfler konzentrieren sich in ihrer Betrachtung auf die Modellierung von Anwenderverhalten im Web und zeigen anhand unterschiedlicher Beispiele, wie eine strukturierte Aufbereitung der gewonnenen Daten zu einer qualitativen Verbesserung von Applikationen und der Nutzer-zentrierten Knowledge Discovery führen kann.

Axel Polleres und Malgorzata Mochol zeigen in ihrem Beitrag unterschiedliche Methoden der Expertensuche im Web. Am Beispiel des Projektes „ExpertFinder“ diskutieren sie die Rahmenbedingungen zur Kombination, Wiederverwendung und Erweiterung bestehender RDF-Vokabulare zur Beschreibung von Personen, Organisationen und Expertise im Web und deren Konsequenz für die Jobsuche der Zukunft.

Erich Gams und Daniel Mitterdorfer diskutieren in ihrem Beitrag die Anforderungen an die nächste Generation von Content Management Systemen und stellen eine generische Architektur eines Semantic Web-tauglichen Systems vor.

Ergänzend dazu vergleichen **Andreas Blumauer und Martin Hochmeister** unterschiedliche Tag-Recommender-Systeme zur gemeinschaftlichen Verschlagwortung und stellen ein ontologiegestütztes Tag-Recommender-System vor, auf Basis dessen sich neue Anwendungen und Mehrwertdienste realisieren lassen.

Sebastian Schaffert, François Bry, Joachim Baumeister und Malte Kiesel führen in die Grundlagen von semantischen Wikis ein und zeigen,

wie durch die Verknüpfung von „sozialer Intelligenz“ und „künstlicher Intelligenz“ das Potential von Wikis weiter erschlossen werden kann.

Sören Auer, Jens Lehmann und Christian Bizer bieten einen Einstieg in die spannende Welt von Open Data und zeigen anhand semantischer Mashups, wie offene Datenquellen am Netz mittels semantischer Technologien und Reasoning (teilautomatisiert) vernetzt und zum weiteren Gebrauch veröffentlicht werden können.

Abschnitt 3 thematisiert bestehende Anwendungen und deren Perspektiven in einem Social Semantic Web.

Am Beispiel Plattform Wissensmanagement geben **Stefanie N. Lindstaedt und Claudia Thurner** einen praxisnahen Einblick in das gelebte Management einer Community of Interest und zeigen, wie soziale Prozesse in der Online- und Offline-Welt mittels semantischer Technologien unterstützt werden können.

Gernot Tscherteu und Christian Langreiter stellen unterschiedliche Methoden und darauf basierende Visualisierungsapplikationen vor, die dazu geeignet sind Online-Communities visuell zu erschließen und so die Nutzer bei der sozialen Interaktion am Netz unterstützen.

Mit dem Semantic Desktop präsentieren **Leo Sauermann, Malte Kiesel, Kinga Schumacher und Ansgar Bernardi** die nächste Generation von Desktop-Anwendungen. Mittels des Konzeptes „persönliches Wissensmanagement“ werden Ansätze aus dem Web 2.0 und dem Semantic Web verknüpft und in eine generische Architektur überführt.

Andreas Hotho, Robert Jäschke, Dominik Benz, Miranda Grahl, Beate Krause, Christoph Schmitz und Gerd Stumme zeigen am Beispiel BibSonomy ein kooperatives Verschlagwortungssystem für Web-Links und Publikationen. Sie gehen auf die bestehenden Funktionalitäten und die dahinter stehende Architektur ein, und zeigen, wie die gewonnenen Daten für weitere Zwecke ausgewertet werden können.

Markus Krötzsch und Denny Vrandečić thematisieren am Beispiel Semantic Wikipedia die Entwicklungsoptionen und verfügbaren Technologien für die nächste Generation der beliebten Grass-Roots-Enzyklopädie und zeigen, wie das gemeinschaftlich generierte Wissen mittels entsprechender Exportmöglichkeiten in anderen Anwendungen weiterverwendet werden kann.

Christoph Wieser und Sebastian Schaffert stellen die Fragen nach der Zeitung der Zukunft, und wie sich das Zeitunglesen auf Basis semantisch gestützter Redaktionsprozesse verändern wird.

Holger Stenzhorn und Matthias Samwald zeigen zum Abschluss die Relevanz des Semantic Web für den Bereich der biomedizinischen Forschung, insbesondere um aus der unfassbaren Datenfülle jene relevanten Erkenntnisse zu extrahieren, die zu Durchbrüchen bei neuen Medikamenten und Heilungsmethoden führen sollen.

Abschnitt 4 versucht eine Reflexion des Phänomens „Social Semantic Web“ aus einer sozialwissenschaftlichen Perspektive³.

Jan Schmidt und Tassilo Pellegrini nähern sich dem Social Semantic Web aus kommunikationssoziologischer Perspektive und identifizieren sowohl unterschiedliche Wissenstypen der sozialen Organisation als auch gängige „Bemächtigungsstrategien“ zur Nutzung und Aneignung von User-Generated Content.

Im einzigen englischsprachigen Aufsatz dieses Bandes fragen **Narayan Kulathuramaiyer und Hermann Maurer** nach den Implikationen des zunehmenden Data Minings im Spannungsfeld zwischen technischer Notwendigkeit für gesteigerte Convenience und Endnutzertauglichkeit von Anwendungen und sozialer Bedrohung durch Informationsasymmetrien und Machtmissbrauch.

Michael Nagenborg beschließt diesen Abschnitt mit einer fundierten Diskussion von Privacy im Social Semantic Web. Aus der Perspektive der Informationsethik stellt er die Forderung das Private als ein normatives, handlungsleitendes Konzept zu verstehen, dem die technische Infrastruktur entsprechen muss um sozial verträglich zu sein.

Wien, im Juni 2008

³ Leider sind zwei wichtige Themenkomplexe unterrepräsentiert: Intellectual Property Rights und Governance. Wir hoffen u. a. mittels dieses Buches den Diskurs zu diesen Themen anzuregen und den Wissenstransfer zwischen den Disziplinen zu fördern.

Danksagung

Bücher schreiben ist Knochenarbeit, sowohl für die Autoren als auch die Herausgeber. Gleich vorweg gilt daher unser Dank allen Autorinnen und Autoren, die mit Eifer und Energie an diesem anspruchsvollen Projekt mitgewirkt haben. Weiters bedanken wir uns bei Eva Pellegrini, die wie schon im letzten Band durch ihr Lektorat einen wertvollen Beitrag zur formalen Qualität beigesteuert hat, ebenso bei Jana Herwig und Jan Krone, die das Buch mit wertvollen Kommentaren bereichert haben. Und last but not least gehört unser Dank dem Springer-Verlag, welcher uns in jeder Phase des Projektes unterstützt hat und uns mit Rat und Tat bei der Erstellung der Druckvorlage zur Seite stand.

Inhaltsverzeichnis

Einstiegspunkte

1. **Semantic Web Revisited – Eine kurze Einführung
in das Social Semantic Web.....** 3
Andreas Blumauer und Tassilo Pellegrini
2. **Warum Web 2.0? Oder: Was vom Web 2.0 wirklich bleiben wird** 23
Sonja Bettel
3. **Entwicklungsperspektiven von Social Software und dem Web 2.0** 43
Alexander Raabe
4. **Anwendungen und Technologien des Web 2.0: Ein Überblick** 63
Alexander Stocker und Klaus Tochtermann
5. **Die Usability von Rich Internet Applications** 83
Jörg Linder
6. **Die zwei Kulturen.....** 99
*Anupriya Ankolekar, Markus Krötzsch, Than Tran
und Denny Vrandečić*

Technologien und Methoden

7. **Wissensmodellierung im Semantic Web.....** 127
Barbara Geyer-Hayden
8. **Anreicherung von Webinhalten mit Semantik –
Microformats und RDFa** 147
Michael Hausenblas

9. Modellierung von Anwenderverhalten im Social Semantic Web.....	159
<i>Armin Ulbrich, Patrick Höfler und Stefanie Lindstaedt</i>	
10. Expertise bewerben und finden im Social Semantic Web.....	175
<i>Axel Polleres und Malgorzata Mochol</i>	
11. Semantische Content Management Systeme.....	207
<i>Erich Gams und Daniel Mitterdorfer</i>	
12. Tag-Recommendier gestützte Annotation von Web-Dokumenten.....	227
<i>Andreas Blumauer und Martin Hochmeister</i>	
13. Semantische Wikis.....	245
<i>Sebastian Schaffert, François Bry, Joachim Baumeister und Malte Kiesel</i>	
14. Semantische Mashups auf Basis Vernetzter Daten	259
<i>Sören Auer, Jens Lehmann und Christian Bizer</i>	

Anwendungen und Perspektiven

15. Web-gestütztes Social Networking am Beispiel der „Plattform Wissensmanagement“.....	289
<i>Stefanie N. Lindstaedt und Claudia Thurner</i>	
16. Explorative Netzwerkanalyse im Living Web.....	313
<i>Gernot Tscherteu und Christian Langreiter</i>	
17. Semantic Desktop	337
<i>Leo Sauermann, Malte Kiesel, Kinga Schumacher und Ansgar Bernardi</i>	
18. Social Bookmarking am Beispiel BibSonomy	363
<i>Andreas Hotho, Robert Jäschke, Dominik Benz, Miranda Grahl, Beate Krause, Christoph Schmitz und Gerd Stumme</i>	
19. Semantic Wikipedia	393
<i>Markus Krötzsch und Denny Vrandečić</i>	
20. Die Zeitung der Zukunft.....	423
<i>Christoph Wieser und Sebastian Schaffert</i>	
21. Das Semantic Web als Werkzeug in der biomedizinischen Forschung	435
<i>Holger Stenzhorn und Matthias Samwald</i>	

Reflexion

22. Das Social Semantic Web aus kommunikationssoziologischer Perspektive.....	453
<i>Jan Schmidt und Tassilo Pellegrini</i>	
23. Implications of Emerging Data Mining.....	469
<i>Narayanan Kulathuramaiyer and Hermann Maurer</i>	
24. Privacy im Social Semantic Web.....	485
<i>Michael Nagenborg</i>	
Index	507

I Einstiegspunkte

1. Semantic Web Revisited – Eine kurze Einführung in das Social Semantic Web

Andreas Blumauer und Tassilo Pellegrini

Semantic Web Company, Wien, Österreich
{a.blumauer; t.pellegrini}@semantic-web.at

Zusammenfassung: Während in den vergangenen Monaten Themen wie Web 2.0 und Social Software ein erstaunliches Konjunkturhoch erlebt haben, vollzieht sich weitgehend abseits der öffentlichen Wahrnehmung eine technologische Komplexitätsinnovation. Die wachsende Adaption semantischer Technologien zu Zwecken der strukturierten Erschließung von „Web 2.0 Content“, aber auch der Einsatz von Social Software zur kollaborativen Anreicherung von Web Content mit maschinenlesbaren Metadaten sind Ausdruck eines Trends in Richtung „Social Semantic Web“. Bezeichnendes Merkmal dieser Entwicklung ist die voranschreitende Konvergenz zwischen Social Software und Semantic Web Technologien. Dieser Beitrag hat das Ziel ein allgemeines Bewusstsein und Verständnis dieser Entwicklung zu schaffen und nähert sich dem Phänomen aus einer nicht-technischen Perspektive.

Ein Blick zurück nach vorne

Noch im Jahr 2005, als der Begriff Web 2.0 erstmals breitenwirksam in Erscheinung trat, hätte sich niemand vorstellen können, welche weitreichenden Auswirkungen der Trend zum Social Web auf das Internet-Business haben würde. Nicht nur, dass die großen Service-Anbieter in atemberaubender Geschwindigkeit mit eigenen Web 2.0-Angeboten in einen scharfen Konkurrenzkampf um die Gunst der Nutzer traten, auch spektakuläre Firmenübernahmen wie etwa Yahoo!'s Kauf der Social Bookmarking Plattform del.icio.us [21] und der Foto-Sharing-Site Flickr [10], Rupert Murdochs Übernahme des Social Networking-Anbieters MySpace [39], sowie eBays Übernahme von Skype für knapp 2,6 Milliarden Dollar [26] markierten den Anfang einer Reihe von Akquisitionen, die dem Dot-Com-Hype der späten 1990er Jahre nicht nachzustehen

schienen.¹ Und noch ist kein Ende der Akquisitionswelle in Sicht, denn seit 2006 bemüht sich Yahoo! sich gegen ein Übernahmeangebot von Microsoft [31] zu behaupten, das im April 2008 bei knapp 45 Milliarden Dollar [24] lag.² Kurz gesagt: Der Internet-Markt ist in Bewegung.

Die Frage nach den wirtschaftlichen Motiven dieser Transaktionen lässt sich ganz trocken auf zwei Aspekte zurückführen: Zum einen geht es um die strategische Besetzung neuer Kunden- und Technologie-Märkte, die sich in Network-Industries [26] herausbilden und jeden Player unter Druck setzen, entsprechende Windows of Opportunity zu nutzen – auch auf die Gefahr hin, Fehlinvestitionen zu machen³; zum anderen geht es um die Erschließung eines der lukrativsten Geschäftsbereiche der Zukunft, nämlich der Online-Werbung, die seit mehreren Jahren mit zweistelligen Wachstumsraten⁴ nicht nur hohe Profite verspricht, sondern mit der voranschreitenden Konvergenz aus Medien, Telekommunikation, Unterhaltung und Internet auch neue Geschäftsmöglichkeiten und Dienstleistungen eröffnet.

Die strategische Sicherung von Marktpositionen, auch ohne ein revolutionäres Geschäftsmodell in der Tasche, ist eine Grundvoraussetzung für den wirtschaftlichen Erfolg von morgen. Und dieser Erfolg ist auf Gedeih und Verderb von der Verfügungsgewalt über die wichtigste strategische Ressource einer vernetzten, auf „Social Production“ basierenden Informationsökonomie [3] abhängig: Daten.⁵

¹ Weitere erwähnenswerte Übernahmen waren z.B. die YouTube-Akquisition durch Google um 1,65 Milliarden Dollar Ende 2006 [8], die Übernahme der deutschen Social Networking Plattform StudiVZ durch die Holtzbrinck-Gruppe um 100 Millionen Euro Anfang 2007 [28] und Microsofts Beteiligung von 1,6% an Facebook für 240 Millionen Dollar im Oktober 2007 [40]. Kleinere Akquisitionen, wie etwa die Übernahme des Technologieanbieters ClearForest durch Reuters für 25 Millionen Dollar [23] oder eBays Kauf des Bookmarking-Dienstes StumbleUpon für 75 Millionen Dollar [7] im Jahr 2007, nehmen sich dabei etwas bescheidener aus.

² Microsoft zog im Monat darauf (Mai 2008) sein Angebot wieder zurück, was mit verheerenden Kurseinbrüchen der Yahoo!-Aktie einherging. Es bleibt zu spekulieren, ob es sich dabei um eine strategische Entscheidung Microsofts im weiteren Übernahmeprozess handelt.

³ Wie etwa am Beispiel eBay-Skype kolportiert wird [34].

⁴ Im Vergleich zur Print- und TV-Werbung befindet sich die Online-Werbung jedoch noch auf relativ geringem Niveau, konkurriert jedoch anteilig mit Radio- und Kinowerbung [42].

⁵ Die aktuelle Debatte über Datenschutz und Privacy in Zusammenhang mit Googles Informationsgebarung [16], das gescheiterte „Beacon-Experiment“ von facebook [15], die peinlichen Änderungen der AGBs von StudiVZ [11], der Überwachungsskandal bei deutschen Lebensmitteldiskontern wie Schlecker [33] und Debatten wie Bundestrojaner [13], Vorratsdatenspeicherung [12],

Von Daten-Webs und Intel-Metaphern

Daten sind der Rohstoff und die Währung in Netzwerk-Industrien. Und die Möglichkeit, diese kostengünstig zu erschließen, ist eine der strategischen Kernkompetenzen der Zukunft. Nicht ganz ohne Zufall verwendet Tim Berners-Lee gerne den Ausdruck „Web of Data“ [32] als Synonym für die etwas glücklose Wahl des Begriffes „Semantic Web“, und auch Tim O'Reilly postuliert in seinen Web 2.0 Design Patterns plakativ: „Data is the Next Intel Inside“ [18].

Die Frage nach der Erschließbarkeit und Logistik von Daten im Web steht im Zentrum der Diskussion um Nutzen und Funktion von Semantic Web Technologien. Und diese Diskussion wird kontroversiell geführt.

Skeptiker des Semantic Web nennen die vom World Wide Web Consortium entwickelten Standards und Methoden als zu kompliziert und Top-Down-getrieben, um im vom Grass-Roots-Ethos gesteuerten Web erfolgreich zu sein.

Im Gegenzug werfen die Vertreter der Semantic Web Community die berechtigte Frage auf, welche technologische und methodische Alternative zur Verfügung steht, um die durch Bottom-Up-Prozesse rasch wachsende Datenflut am Netz erfolgreich in den Griff zu bekommen und gleichzeitig Applikationen zu entwickeln, die in der Lage sind, den Informationsreichtum im weitläufigen, dezentralen Internet und kleinteiligen Intranets zu erschließen.

Beide Fraktionen können einen Gültigkeitsanspruch auf ihre Argumente erheben. Und langsam wächst das Verständnis, dass nur eine kombinierte Form von Top-Down- und Bottom-Up-Ansätzen der Datenorganisation eine nachhaltige Lösung darstellt, die jedoch nicht nur an der technologischen Machbarkeit, sondern auch der sozialen Akzeptanz und Verträglichkeit zu messen ist. Denn die Bewirtschaftung von (Meta-)Daten – ob durch gemeinnützige Communities, profitorientierte Unternehmen oder staatliche Institutionen – geht Hand in Hand mit der Frage über die Verfügungsgehalt der existierenden Datenbestände und fügt dem Technologiediskurs eine Dimension über geistiges Eigentum, Datenschutz und Macht hinzu.

Aus technologischer Perspektive erleben wir eine voranschreitende gegenseitige Annäherung der Semantic Web und Web 2.0 Glaubensgemeinschaften mit dem Effekt, dass parallel zum Web 2.0-Hype eine atemberaubende Dynamik bei der Konvergenz von Web 2.0 und Semantic Web zu

Flugdatenaustausch [9] u. v. a. m. sind Symptome für die gehobene mediale und öffentliche Aufmerksamkeit gegenüber der Datengebarung von Unternehmen und Behörden.

beobachten ist – ein Phänomen, das unter dem Begriff „Social Semantic Web“ Einzug in die akademische Diskussion [2] gefunden hat.

Parallel dazu formiert sich rund um Tim Berners-Lee seit etwa 2005 eine interdisziplinäre Gruppe von Technikern, Ökonomen, Psychologen und Soziologen, die mit einer „Science of the Web“ [36] den komplexen Wechselwirkungen zwischen Technologie, Wirtschaft und Gesellschaft auf den Grund gehen wollen, um u. a. im Sinne einer Technikfolgenabschätzung den disruptiven Effekten, die durch den Einsatz und die Kombination unterschiedlicher Methoden und Technologien des Webs entstehen, nachzuforschen.

Neben der akademischen Diskussion manifestiert sich das Social Semantic Web für den aufmerksamen Beobachter in einer Vielzahl an Ereignissen, die als praktische Beispiele für die voranschreitende Konvergenz aus Web 2.0 und Semantic Web herangezogen werden können.

So hat sich parallel zu Open Source ein weltweites Phänomen etabliert, das unter dem Begriff „Open Data“ firmiert und eine enorme Menge an hoch-strukturierten, frei verfügbaren Daten am Web bezeichnet, die aufeinander referenzieren und daher auch wie eine riesige, verteilte Datenbank abgefragt werden können. Das Open Dataset⁶ speist sich aus Quellen wie der Wikipedia⁷, dem US Census⁸, dem CIA World Factbook⁹, der GeoNames Datenbank¹⁰, der MusicBrainz Datenbank¹¹, Flickr¹² und einer Vielzahl an kleinteiligen Quellen und stellt mit mehr als 2 Milliarden Fakten das derzeit größte offene Daten-Repository der Welt dar [20].

Im Jänner 2008 traten Google, Facebook und Plaxo der Data Portability Group¹³ bei, der u. a. Unternehmen wie Microsoft, Yahoo!, Mozilla, LinkedIn und Digg angehören. Die Data Portability Group und ihre Tochterorganisation Open Data Definition¹⁴ verstehen sich als offene Plattform für alle Stakeholder zum Thema Social Web-Interoperabilität und verfolgen das Ziel, die technischen, sozialen und rechtlichen Implikationen des Datenaustausches über Social Web-Dienste und -Applikationen hinweg zu verstehen und in nachhaltige, sozial verträgliche Geschäftsmodelle umzumünzen.

⁶ Siehe <http://richard.cyganiak.de/2007/10/lod/>

⁷ Siehe <http://dbpedia.org/About>

⁸ Siehe <http://www.rdfabout.com/demo/census/>

⁹ Siehe <http://www4.wiwiss.fu-berlin.de/factbook/>

¹⁰ Siehe <http://www.geonames.org/ontology/>

¹¹ Siehe <http://fgiasson.com/blog/index.php/2007/05/22/browsing-musicbrainzs-dataset-via-uri-dereferencing/>

¹² Siehe <http://apassant.net/blog/2007/12/18/rdf-export-of-flickr-profiles-with-foaf-and-sioc/>

¹³ Siehe <http://dataportability.org/>

¹⁴ Siehe <http://www.opendd.net/>

Yahoo! verlautbarte im Februar 2008, dass es im Rahmen seiner Open Search Strategie [37] verstärkt auf die Einbindung von Mikroformaten und RDF-Daten aufbauen werde, und bekennt sich zu einem „Semantic Web Light“. Die verstärkte Berücksichtigung von Kontextdaten im Rahmen des Dienstes oneSearch 2.0 soll eine bisher unerreichte Qualität der Suchergebnisse auf mobilen Endgeräten gewährleisten.

Im März 2008 ließ Microsoft Research auf der Open Repositories Conference in Southampton / UK verlautbaren, dass sie sich bei ihrer Metadatenstrategie stärker an den Standards und Methoden des W3C orientieren wollen [22].

Und nicht zuletzt sei auf die unzähligen Startups¹⁵ verwiesen, die innovative Methoden aus der Forschung in markttaugliche Produkte und Dienstleistungen umsetzen und dabei nicht unbeträchtliche Finanzierungserfolge durch Risikokapitalgeber verbuchen können oder durch große Player übernommen und vermarktet werden. So bietet Reuters nach einer Übernahme des Technologieanbieters ClearForest seit Anfang des Jahres 2008 den kostenlosen Web-Dienst OpenCalais¹⁶ zur Nutzung an, der unstrukturierten Text automatisch auf Basis des Semantic Web-Standards RDF¹⁷ annotiert. Eine Lektüre der Allgemeinen Geschäftsbedingungen [19] lässt klar erkennen, dass Reuters den Metadatenmarkt als neues Geschäftsfeld ins Auge gefasst hat.

In Anlehnung an den US-Theoretiker Yochai Benkler [3] lässt sich das Social Semantic Web als vielschichtiges klassen- und kulturübergreifendes Phänomen verstehen, das im Sinne der sozialen Produktion von Informationsgütern die organisationale/unternehmerische Ebene ebenso adressiert, wie den individuellen Nutzer.

Wo ist das „Soziale“ im Social Semantic Web?

Web 2.0 und Semantic Web stellen zwei Paradigmen zur Datenbewirtschaftung dar. Diese unterscheiden sich in der Art der Inhaltserschließung (z. B. manuell vs. maschinell) und der Art der Wissensorganisation (z. B. expertengetrieben vs. selbstorganisiert) [20]. Abbildung 1 veranschaulicht prototypisch das Zusammenspiel der unterschiedlichen Paradigmen bei der Herausbildung eines Social Semantic Web.

¹⁵ Exemplarisch zu nennen sind wie Radar Networks, Semantic Arts, Talis, Top-Quadrant, Franz, Ontotext, OpenLink, Powerset u. v. a. m.

¹⁶ Siehe www.opencalais.com

¹⁷ RDF (Resource Description Framework) ist ein zentraler Datenbeschreibungsstandard für das Semantic Web und wurde im Jahr 2004 als offizielle W3C Recommendation verabschiedet [W3C 2004].

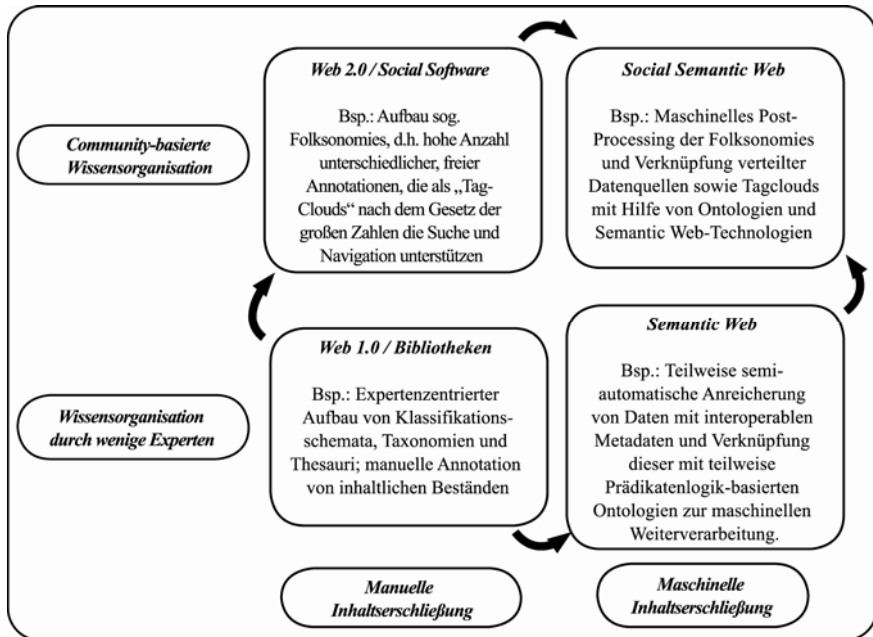


Abb. 1. Entwicklungspfade zum Social Semantic Web entlang der Dimensionen Wissensorganisation und Inhaltserschließung [20]

Fügt man die beiden Paradigmen zusammen, so ergeben sich zwei komplementäre Entwicklungspfade, die bei der Herausbildung des Social Semantic Web eine wichtige Rolle spielen [25]:

Einerseits „**Semantically Enabled Social Software**“, also die funktionelle Erweiterung von Social Software zur strukturierten Anreicherung von Web 2.0 Content mit maschinenverarbeitbaren Metadaten; andererseits als „**Socially Enabled Semantic Web**“, die kollaborative Bereitstellung von großen strukturierten Datenbeständen zur Ermöglichung von Mashups und Rich Content Applikationen.

Obwohl diese beiden Szenarien analytisch voneinander getrennt werden können, ist in der Praxis eine Konvergenz beobachtbar, die in Teilbereichen zu einer nahtlosen Verschmelzung von Social Software und semantischen Technologien führt. Im Folgenden werden die beiden Szenarien skizziert.

Semantically Enabled Social Software

Betrachtet man das wohl prominenteste Beispiel eines aktiven und wohl erfolgreichen Social Software Projektes, nämlich Wikipedia, so stellt man schnell fest, dass der überwiegende Content-Anteil unstrukturiert vorliegt. Im Laufe der Zeit haben sich jedoch zahlreiche strukturgebende Merkmale herausgebildet, die von der Community mittlerweile als mehr oder weniger fix betrachtet werden. So existieren zu jeder Wiki-Seite, in der eine Stadt beschrieben wird, strukturierte Angaben darüber, in welcher Region oder in welchem Land die Stadt liegt, oder Angaben über entsprechende Geo-koordinaten. Wie in Abb. 2 dargestellt, werden diese Metadaten übersichtlich und strukturiert zu einer Factbox zusammengefasst und sind damit für die maschinelle Verarbeitung bereits sehr gut geeignet.

Ein Ansatz, der explizit das Ziel verfolgt, Social Software mit der Vision des Semantic Web zu verbinden, ist z. B. bei Semantic-Wiki-Systemen zu finden. Hier werden erweiterte Möglichkeiten angeboten, Inhalte eines Wikis auf Basis von Ontologien zu annotieren. Muss in gegenwärtigen Systemen die inhaltliche Erschließung noch weitestgehend manuell erfolgen, werden in naher Zukunft semi-automatische Verfahren der Annotati-on – z. B. mittels „Tag Recommender“ – sowie der Datenextraktion – z. B. mittels Wrappern oder Screen Scrapern – die Nutzer bei der inhaltlichen Erschließung unterstützen und dabei generierte Metadaten in einer maschinenverarbeitbaren Form zur Verfügung stellen.



Location of Vienna in Austria
Coordinates: 48°12'N 16°21'E

State	Austria
Government	
- Mayor	Michael Häupl (SPÖ)
Area	
- City	414.90 km² (160.2 sq mi)
- Land	395.51 km² (152.7 sq mi)
- Water	19.39 km² (7.5 sq mi)
Population (2007)	
- City	1,668,737
- Density	4,011/km² (10,388.4/sq mi)
- Metro	2,081,714
Time zone	CET (UTC+1)
- Summer (DST)	CEST (UTC+2)
Website:	www.wien.at

Abb. 2. Beispiel einer Factbox zu Wien: Strukturierte Metadaten in Wikipedia

Socially Enabled Semantic Web

Während also bei „Semantically Enabled Social Software“ darauf fokussiert wird, wie Social Software um semantische Funktionalitäten angereichert werden kann, um letztlich die Generierung von strukturierten Daten zu unterstützen, liegt der Fokus des „Socially Enabled Semantic Web“ auf der Frage, wie Semantic Web-taugliche Daten der Allgemeinheit zur Verfügung gestellt werden können, um Rich Content Anwendungen am Web zu ermöglichen. Aus dieser Perspektive geht es also nicht um die technologische Realisierbarkeit, sondern um den sozio-kulturellen Prozess zur Generierung einer kritischen Masse an maschinenverarbeitbaren Metadaten.

Erste Ansätze wurden vom FOAF-Community-Projekt¹⁸ unternommen, das zum Ziel hat, Personen dazu zu animieren, persönliche Informationen über Interessen oder soziale Beziehungen in einem einheitlichen maschinenverarbeitbaren Format zu veröffentlichen. Auf Basis einer hinreichenden Anzahl von FOAF-Daten können Anwendungen wie z. B. Expertensuchmaschinen entwickelt werden.

Als neueres Beispiel eines Socially Enabled Semantic Web machen sich *DBpedia* und das damit verbundene Community-Projekt *Linking Open Data* [14] den Umstand zunutze, dass im Web Daten in strukturierter Weise gesammelt und auch frei zur Verfügung gestellt werden. Das DBpedia-Dataset beinhaltet (Stand: Mai 2008) 218 Millionen Triples, die aus Wikipedia extrahiert wurden, und bildet gemeinsam mit anderen Quellen wie Geonames (60 Millionen Triples), Musicbrainz (50 Millionen Triples), dem US Census (700 Millionen Triples) und einer Unmenge an kleinteiligen Daten-Repositories die derzeit wohl umfangreichste maschinenverarbeitbare Datenbasis im Internet.

Diese Quellen bilden nicht nur eine wichtige Basis für das in Entstehung befindliche Semantic Web, sondern sind auch der Ausdruck einer auf Kollaboration und Selbstorganisation aufbauenden Informationskultur, die als Alternative zu einer Top-Down, von Experten geleiteten Wissensorganisation im Web verstanden werden muss. Denn in einem Socially Enabled Semantic Web werden Ordnungssysteme (Ontologien) und darauf basierende Anwendungen nicht von oben vorgegeben, sondern von den Nutzern bottom-up auf Grundlage der zur Verfügung stehenden Datenbasis und (syntaktischen) Standards bedarfsgetrieben entwickelt.¹⁹ Erst dadurch ist

¹⁸ FOAF ist die Abkürzung für Friend-Of-A-Friend und bezeichnet einen Beschreibungssstandard für soziale Netzwerke.

¹⁹ Hier wäre spannend zu untersuchen, ob sich der Bazaar-Governance Ansatz, wie er aus der Open Source Forschung hervorgegangen ist [5], auch auf die kollaborative Generierung und Adaption von Ontologien wie etwa FOAF, SIOC, SKOS etc. anwendbar ist.

gewährleistet, dass Daten beliebig und kollaborativ miteinander kombiniert und wieder zur Verfügung gestellt werden können.

Eine kommunikationswissenschaftliche Perspektive

Zieht man den Erkenntnisstand der Kommunikationswissenschaft zur Internetökonomie heran [35; 40], so lässt sich die Entwicklung des Social Semantic Web mittels zweier reziproker Wirkungsprinzipien beschreiben: (1) die Erhöhung von (sozialer) Konnektivität und (2) die Erhöhung von Kontextualisierung. Abbildung 3 veranschaulicht das Zusammenspiel dieser beiden Dimensionen.

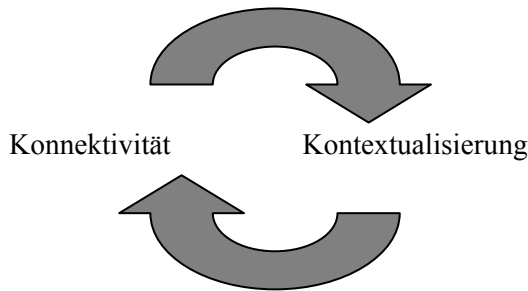


Abb. 3. Reziproke Verstärkung von Konnektivität und Kontextualisierung durch das Social Semantic Web

Unter Konnektivität sollen Aspekte der Vernetzung von Personen, Inhalten, Diensten und Daten, sowie dabei entstehende Kosten, wie Suchkosten, Verhandlungskosten, Maßnahmen zur Vertrauensbildung, Rückbindungsmaßnahmen etc. zusammengefasst werden.

Unter Kontextualisierung fallen hingegen Aspekte des Beschreibens, Ordners, Filterns und Bewirtschaftens von Daten und Information, und damit Kosten, die im Zuge der Anreicherung mit Metadaten, des Datenmanagements, der Herstellung von Interoperabilität etc., kurz: der Konnektivitätssteigerung entstehen.

Beide Dimensionen beeinflussen und regulieren sich gegenseitig. So bedingt eine gesteigerte Konnektivität im Zuge höherer Datenverfügbarkeit und Vernetzung im Internet einen höheren Grad an Kontextualisierung, wie sich an der Nutzung von Internet-Suchmaschinen, aber auch an den vielfältigen neuen Visualisierungs- und Navigationsmethoden wie Tag Clouds, Clustering-Verfahren, Recommender-Systemen, facettiertem Browsing etc. sehen lässt.

Wendet man die Prinzipien Konnektivität und Kontextualisierung auf das Web 2.0 und das Semantic Web an, so fällt auf, dass beide Web-Entwicklungen diese Prinzipien beinhalten, jedoch mit diametral entgegengesetzter Gewichtung. Denn während der Fokus des Web 2.0 auf der Herstellung von Konnektivität und dem „Front-End“ von Applikationen liegt, verfolgt das Semantic Web das Ziel der weitläufigen Kontextualisierung von Daten und adressiert primär das „Back End“. Mit anderen Worten: Beide Entwicklungen stehen komplementär zueinander, sowohl aus technologischer als auch aus organisationaler Sicht. Und sieht man von der Diskussion über die technologische Machbarkeit eines Semantic Web im Sinne von Tim Berners-Lee ab, so scheint offensichtlich, dass die Kontextualisierung ein inkrementelles Prinzip der weiteren Webentwicklung ist, wenn auch bei weitem anspruchsvoller zu realisieren und aus Gründen der Komplexitätsreduktion dem technisch unversierten Nutzer weitgehend verborgen.²⁰

Rekapitulieren wir: Vom Web 1.0 zum Web 2.0

Um gleich mit einem weit verbreiteten Missverständnis aufzuräumen: Das Web 1.0, also das „Web der Dokumente“, wird nicht vom Web 2.0, dem „Web der Services“, ersetzt oder verdrängt, sondern angereichert und teilweise überlagert.

Man kann sich das so vorstellen: Den Kern des „alten“ Web machen vorwiegend (X)HTML-Dokumente aus, die aufeinander via Hyperlink verweisen. Einige wenige Service-Provider verursachen im Web 1.0 den Großteil der Seitenzugriffe. Es werden Seiten, die in Google höher gerankt sind, leichter gefunden, da sie einen höheren PageRank haben²¹. Jede Website hat eine Homepage und viele Unterseiten, die zumeist baumartig organisiert sind, teilweise als „statische Dokumente“ am Webserver liegen, oder als dynamische Seiten aus einer Datenbank heraus generiert werden. Zusammengefasst ausgedrückt: Das Web 1.0 ist das „Web der Dokumente“, Dokumente, die von einigen wenigen für viele passive Benutzer publiziert werden.

²⁰ Dies soll jedoch nicht bedeuten, dass der interessierte Nutzer kein Recht auf die Einsicht und Transparenzierung von Kontextualisierungsmethoden hat und dieses auch einfordern muss. Leider ist die Informationsgebarung von Unternehmen und Behörden bei der Offenlegung ihrer Kontextualisierungsmethoden und -strategien bisher mehr als unbefriedigend.

²¹ Wobei der Begriff PageRank eigentlich vom Namen des Google-Mitbegründers Larry Page abgeleitet ist und die Relevanz einer Webseite im WWW berechnet.

Der entscheidende Schritt in Richtung Web 2.0 ist im Wesentlichen auf zwei Entwicklungen zurückzuführen:

User generieren Content

Dieses Grundprinzip im Web 2.0 hätte eigentlich schon in der Anfangsphase des WWW selbstverständlich sein sollen, ist doch gerade diese Möglichkeit der interaktiven Mediennutzung von Anfang an ein spannendes, neues Element im Internet gewesen. Die Nutzungsforschung [29] aber hat uns gelehrt, dass zunächst nur eine geringe Anzahl von Usern auch aktiv Inhalte erstellt und im Web publiziert hat (z. B. in Form einer eigenen Homepage, seltener noch in Form eines eigenen Web-Shops). Die Gründe dafür sind wohl mannigfaltig, u. a. die zu geringen Bandbreiten Anfang des Milleniums, komplizierte User-Interfaces, der zu geringe Kenntnisstand des durchschnittlichen Users, die fehlenden Vermarktungsmodelle und Marktplätze und damit verbundene extrinsische Motivationsfaktoren, oder einfach nur der Umstand, dass es nicht unbedingt als „cool“ galt, im Web besonders aktiv zu sein.

Die Popularität des „Social Web“-Gedankens stieg mit den wachsenden Partizipationsmöglichkeiten, sodass „jeder“ seinen eigenen Shop betreiben kann (Amazon, Ebay), dass jeder wertvolle Inhalte (auch zum Gemeinwohl anderer) publizieren kann (Wikipedia, Blogs), dass jeder Profil zeigen und sich damit für andere interessant machen kann (Facebook, MySpace, Xing) oder dass auch jeder sein eigenes Video drehen oder Photos schießen und publizieren kann und damit zumindest für „drei Minuten berühmt sein kann“ (Flickr, YouTube). Zusätzlich wurden RSS-Feeds gang und gäbe, um (aktuelle) Inhalte schneller im Web verbreiten zu können, bzw. wurde das einfache Einbinden von Content aus anderen Portalen, z. B. Videos aus YouTube in das eigene Weblog, eine einfache Fingerübung, ja sogar das Erstellen eigener Mashups mit Werkzeugen von Yahoo, Google oder IBM. Parallel dazu stieg das kommerzielle Interesse an Communities und den Möglichkeiten ihrer systematischen Bewirtschaftung.

Neben klassischen Medienhäusern waren es also plötzlich auch die Millionen User, die durch die Nutzung kostenloser Tools dafür sorgten, dass sich Content und Nachrichten über das Web verbreiten konnten. *Sie* waren es, die über die Wichtigkeit und Qualität eines Inhaltes, eines „Memes“ oder einer Nachricht bestimmten, und nicht mehr nur einige wenige Medienhäuser. Im Dezember 2006 kürte das renommierte Nachrichtenmagazin „Time“ daher „YOU“ als Person des Jahres 2006²² und die „Revolution“

²² <http://www.time.com/time/magazine/article/0,9171,1569514,00.html>, zuletzt aufgerufen am 9.10.2007

war vollzogen. Die kollektiven Ehren wurden jedem Web 2.0 User zuteil und das „Social Web“ war in aller Munde. Parallel dazu begann die Debatte darüber, ob dieses Mal wieder das Platzen einer Dot-Com-Blase, wie schon im Jahre 2001, bevorstand, oder ob sich das Internet nun zu einem tragfähigen Modell aus ökonomischer Sicht entwickeln werde.

Web Services zu Mashups verknüpfen

Internet-Provider stellen nicht mehr einfach nur in Browsern darstellbare Dokumente zur Verfügung, sondern auch Services, die via wohl definierter Programmierschnittstellen – kurz APIs genannt – abgefragt und miteinander kombiniert werden können. Ein Beispiel: Ein Web Service (oder Webdienst) namens „Geonames“ liefert u. a. wertvolle Länderinformationen. Nach Aufruf z. B. der Adresse <http://ws.geonames.org/countryInfo?country=AT> werden die wichtigsten Stammdaten des Landes mit dem Ländercode „AT“ (der als Wert für den Parameter „country“ übergeben wurde), also Österreich, als XML-Dokument zurückgegeben. Die Informationen können auch via Browser angezeigt werden:

```
<geonames>
  <country>
    <countryCode>AT</countryCode>
    <countryName>Austria</countryName>
    <isoNumeric>040</isoNumeric>
    <isoAlpha3>AUT</isoAlpha3>
    <fipsCode>AU</fipsCode>
    <continent>EU</continent>
    <capital>Wien</capital>
    <areaInSqKm>83858.0</areaInSqKm>
    <population>8184691</population>
    <currencyCode>EUR</currencyCode>
    <languages>de-AT,tr,sr,hr</languages>
    <geonameId>2782113</geonameId>
    <bBoxWest>9.533569</bBoxWest>
    <bBoxNorth>49.018883</bBoxNorth>
    <bBoxEast>17.166386</bBoxEast>
    <bBoxSouth>46.407494</bBoxSouth>
  </country>
</geonames>
```

Abb. 4. Strukturierte Länderinformationen aus der Geonames-Datenbank

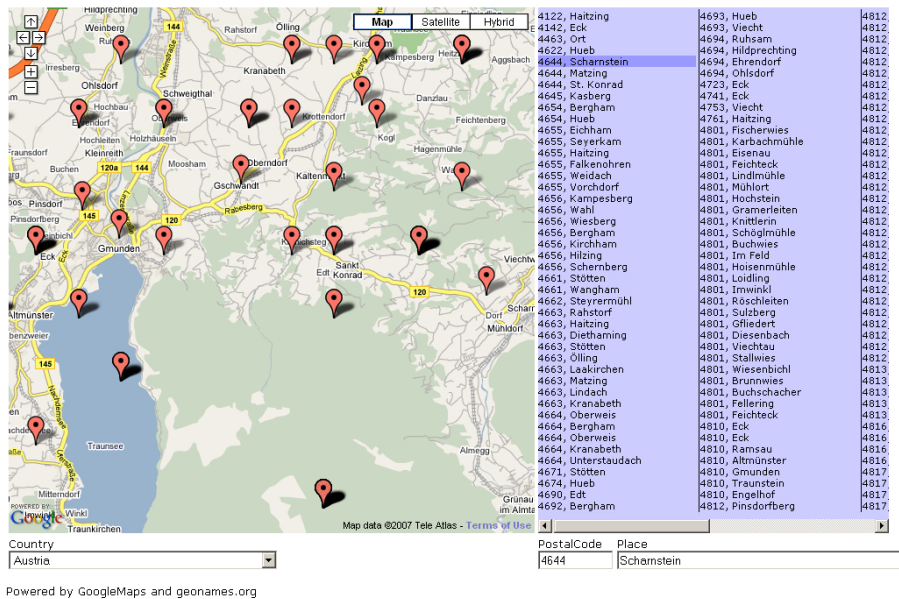


Abb. 5. Screenshot von „Postal Code Mapper“ für Österreich als Beispiel eines Mashups

Ein weiteres Service von Geonames erlaubt so genanntes „Reverse Geocoding“, d. h. nach Angabe von Längen- und Breitengrad wird jenes Land ermittelt, in dem dieser Punkt liegt: <http://ws.geonames.org/countrycode?lat=47.03&lng=10.2> gibt also den String „AT“ zurück.

Die laut ProgrammableWeb²³, einem Verzeichnisdienst verfügbarer APIs, populärsten Web Services werden von GoogleMaps angeboten. Verbindet man nun zwei Webdienste miteinander zu einer neuen Web-Applikation (oder rekursiv wieder zu einem weiteren, neuen Web Service), so spricht man von einem „Mashup“. Ein Beispiel dafür findet sich in Abb. 5: Der „Postal Code Mapper“²⁴ verbindet Geonames und GoogleMaps zu einer Anwendung, die es Nutzern erlaubt, für einen Kartenausschnitt alle Orte und ihre Postleitzahl anzuzeigen und darüber zu navigieren. Die „Webseite“ vereint also, für den User nicht sichtbar, zwei oder mehrere Web-Services (und die darunter liegenden Datenbanken) zu einer neuen Anwendung.

ProgrammableWeb zählt derzeit (Stand: Mai 2008) 745 APIs, u. a. in den Kategorien „Advertising“, „Blogging“, „Email“, „Dating“, „Finance“, „Mapping“, „Search“, „Tagging“ und insgesamt rund 3.000 Mashups, die

²³ <http://www.programmableweb.com/> (zuletzt aufgerufen am 09.10.2007)

²⁴ <http://www.xs4all.nl/~mnankman/Web20Stuff/PCMapper.html>, aufgerufen am 09.10.2007

darauf aufsetzen. Ist das der Nukleus für gänzlich neue Geschäftsmodelle im Web?

Wie das Semantic Web vom Web 2.0 begonnen hat zu lernen

Während der Begriff „Semantic Web“ und damit verbundene Konzepte spätestens seit dem viel zitierten Artikel „The Semantic Web“ von Tim Berners-Lee, James Hendler und Ora Lassila, veröffentlicht im Jahre 2001 im „Scientific American“ [4], vor allem in akademischen Communities herumgeisterte und bis heute weltweit ein kontinuierliches aber konstantes Echo hervorrufen konnte, erlebte das Web 2.0 Konzept (und vor allem der Begriff selbst) seit dem Jahre 2005, als im Umfeld des US-amerikanischen O'Reilly-Verlags der Begriff erstmals geprägt wurde, einen dramatischen Zulauf und wachsendes öffentliches Interesse. Ein Blick auf den Web-Dienst „Google Trends“ (Abb. 6) veranschaulicht diesen Trend.

Warum wurde das Semantic Web vom Web 2.0 „überholt“? Oder war damit das Semantic Web am Ende gar ein „überholtes Konzept“? Eines ist klar: Schreiben die Medien über ein Thema, so wird danach gesucht, doch dieser Umstand sagt kaum etwas darüber aus, ob sich dahinter ein tragfähiges, nachhaltiges Konzept verbirgt. Dennoch war schnell klar: Was der Semantic Web Community in all den Jahren nicht gelang, nämlich ein „neues“, besseres Internet zu ermöglichen, war innerhalb von zwei Jahren im Zuge der „Web 2.0 Revolution“ Realität geworden. Ein regelrechter Boom an neuen Webdiensten und -anwendungen setzte ein, Risikokapitalgeber witterten das große Geschäft, Medien berichteten fast täglich über neue, tolle Angebote im Web 2.0 und teure Kongresse zum Thema waren schnell ausgebucht. Der Begriff selbst wurde mit zahlreichen Bedeutungsebenen vorwiegend positiv „aufgeladen“: „basisdemokratisch“, „revolutionär“, „User-freundlich“, „neue Geschäftsmodelle“, „praktisch“, „einfach

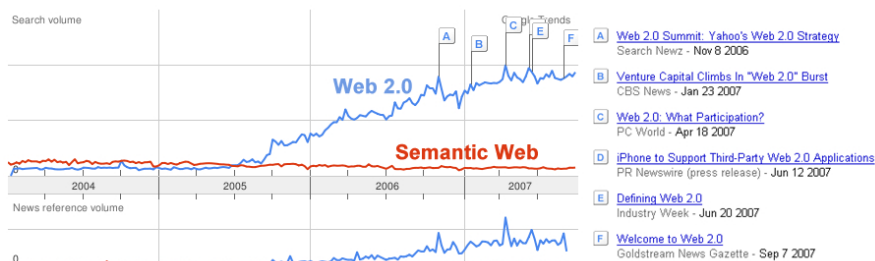


Abb. 6. Suchanfragen bei Google bzw. Nachrichtenaufkommen im Zeitverlauf zu den Themen „Web 2.0“ bzw. „Semantic Web“

bedienbar“ oder „nicht technologiebezogen“ hörte man da und man konnte damit das große Publikum ansprechen, plötzlich war auf vielen Softwareprodukten die Etikette „Web 2.0“ zu sehen.²⁵

Parallel zu dieser Entwicklung wurden – vor allem von den Proponenten des Semantic Web – zunächst die Unterschiede zwischen diesen beiden Ansätzen eines „Internet der nächsten Generation“ betont, und für kurze Zeit wurde dabei vor allem die nicht-wissenschaftliche Fundierung und das Unexakte am Web 2.0 Ansatz stark kritisiert. 2005 wurde mit dem Aufsatz von Peter Mika [17] zwar erstmals der Versuch unternommen, eine Brücke zwischen Web 2.0-Folksonomien und Semantic Web-Ontologien zu schlagen, dennoch ist der Begriff „Web 2.0“ auf den Webseiten der ISWC²⁶ im selben Jahr (noch) nicht existent. Jedoch im Jahr 2006 wird dem Thema auf der ISWC ein Keynote-Vortrag und sogar ein eigenes Web 2.0-Panel gewidmet. Auf der WWW2007 schließlich wurde die Beziehung zwischen Web 2.0 und dem Semantic Web systematisch reflektiert [1] und ein Plädoyer dafür abgegeben, beide Ansätze miteinander zu kombinieren, um damit verbesserte Anwendungen wie z. B. Blog-Systeme auf Basis von Semantic Web-Technologien zu realisieren. Diese Entwicklung ist symptomatisch für die Einsicht, dass sowohl Web 2.0-Ansätze als auch die Technologien des Semantic Web vor allem eines unterstützen: „Intelligente“, nutzerzentrierte, Web-basierte Anwendungen im Internet und Intranet.

Zusammenfassend lässt sich feststellen, dass das Semantic Web also keine „Alternative“ zum Web 2.0 ist, sondern zunächst eine Reihe interessanter Möglichkeiten zu Zwecken der strukturierten Erschließung von „Web 2.0 Content“ bietet (wie z. B. im Linking Open Data Project²⁷), aber auch zur kollaborativen Anreicherung von Content mit maschinenlesbaren Metadaten (wie z. B. durch Semantic Wikis). Gemeinsam sind diese beiden Strömungen bzw. technologischen Entwicklungen Ausdruck eines Trends in Richtung „Social Semantic Web“, dem 2007 bereits auf zwei internationalen Konferenzen²⁸ das Hauptaugenmerk galt.

²⁵ Nicht nur im IT-Business war das Attribut „2.0“ plötzlich sehr en vogue: „Enterprise 2.0“, „Office 2.0“, „Real Estate 2.0“, „Erde 2.0“, „Stasi 2.0“ oder gar „Family 2.0“ waren in den Medien zu finden. Doch auch der 2.0-Hype hat spätestens seit dem desaströsen Berliner Ableger von Tim O'Reillys Business-Konferenz Web 2.0 Expo im November 2007 und darauf folgenden Kritiken an Glanz verloren [28].

²⁶ International Semantic Web Conference, findet jährlich statt und gilt als der „Weltgipfel“ der akademischen Semantic Web Community.

²⁷ <http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>, aufgerufen am 10.10.2007

²⁸ I-SEMANTICS in Graz (<http://triple-i.tugraz.at/>) bzw. CSSW in Leipzig (http://sabreconference.wifa.uni-leipzig.de/frontend/index.php?folder_id=43)

Conclusio: Social Semantic Web

„Social Semantic Web“ ist eine Paraphrase, die eine Transformations- und eine Konvergenzbewegung gleichzeitig umschreibt, keinen Zustand. Der Begriff an sich hat keine Chance, über Jahre hinweg einen Trend zu ebnet oder gar ein langfristiges Phänomen zu beschreiben. „Social Semantic Web“ wird in wenigen Jahren als Begriff verschwunden und – so die These der Autoren – im Web als Selbstverständlichkeit aufgegangen sein. Auch wird damit keine langfristige Vision umschrieben, da bereits sämtliche technologischen Bausteine zur Realisierung vorhanden sind, die in Kürze im „Backend“ einer Web-Anwendung fast schon selbstverständlich laufen könnten.

Nachdem das Web zu „kollektivem Bewusstsein“ gekommen war und in Form eines Social Webs um Unmengen neuer Ressourcen (User, Gruppen), Relationen (Ratings, Tags) und Inhalte (Blogs, Wikis) angereichert worden war, wurde auch die Frage immer virulenter, wie sich denn das Social Web nun organisieren könnte, um auch plattform- und applikationsübergreifend den User mit Informationen versorgen zu können. Der Wettlauf um die besten Methoden, Technologien und Applikationen zum (1) Information Retrieval, zur (2) Informationsvernetzung und (3) Informationsaggregation hatte begonnen. Einen wesentlichen Mehrwert bieten dabei Semantic Web Technologien:

1. **Informationszugänge:** Sie basieren nicht nur auf untypisierten Such- bzw. Stichwörtern, sondern vermehrt auf „Facetten“: Personen, Orte oder Organisationen gehören zu den wichtigsten „Dingen“, nach denen gesucht wird, also wollen die zugrunde liegenden Stellen eines Textes bzw. Web-Ressourcen auch als solche erkannt werden. Semantische Technologien und Metadaten-Schemata des Semantic Web tragen dazu bei, dass diese Zusatzinformationen mit geringem Mehraufwand größtenteils automatisiert generiert werden können.
2. **Informationsvernetzung:** Die automatisierte Vernetzung von Informationen im Netz basiert im Regelfall auf dem Prinzip der semantischen Ähnlichkeit. Nachrichten, Blogs oder User mit ähnlicher Bedeutung, ähnlichen Profilen oder Inhalten wollen als „nahe liegende“ weitere Stationen im Zuge eines Browsings erfahren werden. Die Methoden zur Erkennung von Ähnlichkeiten werden dementsprechend ausgefeilter und basieren nicht mehr nur auf simpler Begriffsebene, sondern auf Wissensmodellen und Ontologien.
3. **Informationsaggregation:** Aus Mashups werden Semantic Mashups, aus Wikis werden Semantic Wikis, aus Blogs Semantic Blogs, d. h. Web Services oder auch Desktop-Anwendungen bauen auf einer „Semantic Web Middleware“ auf. RDF wird zum universellen Data Access Layer bzw. zur ultimativen Mashup-Sprache, auf deren Basis „Model Driven Architecture“ zur Realität wird.

Die wachsende Adaption semantischer Technologien zu Zwecken der strukturierten Erschließung von „Web 2.0 Content“, aber auch der Einsatz von Social Software zur kollaborativen Anreicherung von Web Content mit maschinenlesbaren Metadaten sind Ausdruck des Trends in Richtung Social Semantic Web, was von einigen Beobachtern auch plakativ als Web 3.0 bezeichnet wird.

An der technologischen Machbarkeit des Social Semantic Web wird an vielen Ecken und Enden gearbeitet. Jedoch die zunehmende Verfügbarkeit von Daten und die Möglichkeit ihrer Vernetzung verändern nachhaltig die wirtschaftlichen, sozialen und kulturellen Bedingungen des Internets und bedürfen einer eingehenden Betrachtung, die über die technologische Machbarkeit hinausgeht. Hier sind vor allem die Sozial- und Geisteswissenschaften gefordert ihren Beitrag zu leisten.

Dieser Forderung kann in diesem Band nur ansatzweise entsprochen werden, zumal nicht der Platz vorhanden ist, um all die auftretenden technischen, wirtschaftlichen, sozialen und politischen Fragen zufriedenstellend zu beantworten. Dennoch sei festgehalten, dass die Transformation des Internet zu einem allumfassenden, universellen Kommunikations- und Informationsmedium kontinuierlich voranschreitet. Das Social Semantic Web sollte in diesem Prozess zwar als temporäre, aber nichts desto weniger als disruptive Entwicklung im Auge behalten werden.

Literatur

1. Ankolekar, Anupriya; Krötzsch, Markus; Tran, Thanh; Vrandecic, Denny (2007). The Two Cultures. Mashing up Web 2.0 and the Semantic Web. In: <http://www2007.org/papers/paper777.pdf>, aufgerufen am 10.04.2008
2. Auer, Sören; Bizer, Christian; Müller, Claudia; Zhdanova, Anna V. (2007). The Social Semantic Web 2007. Proceedings of the 1st Conference of Social Semantic Web. Bonn: Gesellschaft für Informatik
3. Benkler, Yochai (2006). The Wealth of Networks. How Social Production Transforms Markets and Freedom. New Haven: Yale University Press
4. Berners-Lee, Tim; Hendler, James; Lasilla, Ora (2001). The Semantic Web. In: http://www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21, aufgerufen am 09.10.2007
5. Demil, Benoit; Lecocq, Xavier (2006). Neither Market, nor Hierarchy, nor Network: The Emergence of Bazaar Governance. In: Organization Studies, 27/2006, S. 1447 – 1466
6. Die Presse (2008). Yahoo kündigt semantische Suche an. 17.03.2008. In: <http://diepresse.com/home/techscience/internet/yahoo/370355/index.do?parentid=183130&act=0&isanonym=0>, aufgerufen am 10.04.2008