

Andreas Blumauer (Geschäftsführer punkt. netServices)
Tassilo Pellegrini (Projektleiter Semantic Web School)

Semantisches Web – schon wieder eine Patentlösung für die Wissensgesellschaft?

Die Informationsüberflutung und der damit verbundene Zeit- und Kostenaufwand der Informationsbeschaffung nagen an der Produktivität von Unternehmen. Trotz – oder vielleicht gerade wegen – zunehmender Vernetzung und schier unerschöpflicher Wissensreservoirs fällt der Weg durch das Dickicht der modernen Informationsgesellschaft häufig schwer. Laut der Studie „Strategische Erfolgspotenziale wissenschaftlicher Portale“ (2004) von Mummert Consulting sind Probleme der Einschätzung von Informationsqualität und ihre unstrukturierte Darstellung besonders dringend. Menschengerechte Software, das heißt Anwendungen, die dem Wesen menschlichen Denkens Rechnung tragen und in der Lage sind, logische Schlussfolgerungen zu ziehen, versprechen Effizienzgewinne. Was auf den ersten Blick wie ein Exkurs in das Science Fiction Genre anmutet, ist durchaus mehr als eine unausgegrenzte Fantasterei. Das semantische Web ermöglicht neben erheblichen Verbesserungen der Usability, einen höheren Gebrauchswert verfügbarer Informationsbestände und effizientere Wissensströme. Mit Ontologie-Repräsentanzsprachen, wie dem Resource Description Framework (RDF), Topic Maps (XTM) und darauf basierenden Derivaten sind die Grundlagen für eine sanfte Revolution der Informationsbeschaffung und -verarbeitung gelegt.

Das semantische Web oder SemWeb, baut auf technische – zum Teil bereits bestehende – Bausteine und Standards. Es bietet einen umfassenden Methodenvorrat, der sich aus unterschiedlichen Teildisziplinen wie Computerwissenschaft, Psychologie oder Linguistik, rekrutiert. Diese Interdisziplinarität könnte gerade dort zum Vorteil gereichen, wo es darum geht, die Technologie an den Benutzer heranzuführen und auf eine große Bandbreite an Anwendungen zuzuschneiden.

Praktische Anwendungen

Die maßgeblichen Anwendungsfelder reichen dabei vom Wissensmanagement über den E-Commerce bis hin zur Enterprise Application Integration (EAI).

EAI ist eines der zentralen Anwendungsszenarien des semantischen Webs. Analysten der Gartner Group, IDC und MetaGroup registrierten in den vergangenen zwei Jahren ein kontinuierliches Wachstum der Ausgaben in diesem Segment. Dieses wird durch den komplementären Einsatz von Web Services als Integrationsschnittstelle vor allem zwischen .NET-Technologien und J2EE (Java 2 Enterprise Edition) noch zunehmen.

Die Einsatzmöglichkeiten sind aber bei weitem vielfältiger: Beispielsweise in Sachen Information Retrieval. Mithilfe von semantischen Technologien lassen sich Suchzeiten in Wissensdatenbanken und Dokumentenbeständen beträchtlich verkürzen und darüber hinaus Mehrwert aus der systematischen Vernetzung bestehender Informationsbestände generieren.

Verwendet man bei der Klassifikation der semantischen Anwendungen die bekannten Wissensbausteine des Wissenskreislaufmodells von Probst¹, so wird ersichtlich, dass Methoden des semantischen Webs die technische Basis zur Unterstützung nahezu des gesamten Wissensflusses einer Organisation sein können. Neben den bereits angesprochenen Prozessstufen der Nutzung von Wissen unterstützen semantische Technologien auch die Prozesse des Erwerbs, der Identifikation, der Entwicklung, der Verteilung und der Bewahrung von Wissen.

Dies spielt besonders im E-Learning eine wichtige Rolle: Durch Firmenglossare, Taxonomien und Ontologien vernetzte Wissensbestände und darauf aufbauende Darstellungsformen wird die Einarbeitungszeit in neue Themengebiete erheblich verkürzt und der Gebrauchswert betriebsinternen Wissens erhöht.

Ferner erhält das Verarbeiten von Information im Allgemeinen einen Schub, denn neue Informationen wie eingehende E-Mails oder Pressemeldungen lassen sich durch automatische Klassifikations-Software schneller erfassen und der Bearbeitung zuführen.

Und last but not least kann man die berühmt berüchtigte Suche nach der Nadel im Heuhaufen vermeiden, wenn es etwa darum geht, Wissensträger zu identifizieren. Gerade bei Mitarbeiterverzeichnissen können Kontext-Informationen einen maßgeblichen Beitrag zur Erhöhung der Interaktionsdichte leisten und so die Kommunikationseffizienz steigern.

Selbst bei der Bewertung von Wissen können semantische Technologien einen wichtigen Beitrag leisten, wenn es darum geht, Versionierungsprozesse nachzuzeichnen, Hot Papers zu identifizieren und die kontextbezogene Relevanz von Dokumenten zu analysieren.

Technisch betrachtet geht es dabei um die stärkere Einbindung der firmeninternen Informations- und Wissensprozesse in die bestehende IT. Ziel ist es, die Partizipation am Wissenskreislauf und die Bedienbarkeit der IT für den Anwender zu erleichtern.

Anders formuliert: Man will Software anbieten, die dem Menschen Arbeiten abnimmt und teilautomatisiert erledigt. Dies können neben den bereits dargelegten Anwendungen auch komplexe Anmelde- oder Ummelde-Vorgänge auf Ämtern oder Ausschreibungsverfahren sein.

¹ Vgl. Probst, G.; Raub, S.; Romhardt, K.: Wissen managen: Wie Unternehmen ihre wertvollste Ressource optimal nutzen. 3. Aufl. Frankfurt/Main: FAZ, 1999

Wie aber lässt sich dies technisch umsetzen? Ist die radikale Abkehr von bisher eingesetzten Systemen notwendig?

Die umfassende Revolution muss nicht stattfinden

Die Proponenten des semantischen Webs zielen keineswegs auf das Ablösen bestehender Systeme, sondern versuchen, eine kostenschonende Variante der Datenintegration mit Hilfe semantischer Technologien zu forcieren.

Einen wesentlichen Beitrag hierfür leisten Web Services, denen die im August 2004 veröffentlichte Studie „Kosten senken mit Web Services“ von SOREON Research eine Vielzahl von Kostensenkungspotentialen nachwies. Im Rahmen der Studie wurde aber auch sichtbar, dass eine Ersparnis nur möglich wurde, weil Web Services auf offene Standards bei der Interoperabilität und der Schnittstellendefinition setzen.

Aber mit den „gängigen“ Web Services heutiger Provenienz ist noch lange nicht das Ende der Fahnenstange erreicht. Deutlich erkennbar zeichnet sich eine neue Generation von Web Services als Ergänzung zum SemWeb ab. Diese semantischen Web Services gehen über die Möglichkeiten bisheriger Web Services auf der Grundlage von .NET und J2EE hinaus. Während letztere noch mit den Beschränkungen der bestehenden Webarchitektur und proprietären Standards kämpfen, sollen semantische Web Services für eine grundlegende Flexibilisierung sorgen.

Es handelt sich hierbei um autarke, sich selbst beschreibende, modulare Applikationen, die alles, was sich über das Web ausführen lässt, bewerkstelligen können – von einfachen Anforderungen bis hin zu komplexen Geschäftsprozessen.

Damit semantische Web Services wahr werden können, müssen die einzusetzenden Markup-Sprachen so aufgebaut sein, dass ein Rechner ihre Bedeutung erkennen kann.

Das Auffinden von geeigneten Web Services ist dafür als erstes zu gewährleisten. Via dem Standard UDDI (Universal Description, Discovery, and Integration Specification) lassen sich Web Services anbieten, identifizieren und nachfragen und so zu komplexeren endkundenspezifischen Services assemblieren. Dazu muss der Web Service seine Merkmale und Fähigkeiten so darlegen, dass der Agent seinen Zweck automatisch versteht.

Darüber hinaus ist es zwingend, dass der Agent in der Lage ist, mit dem Service zu interagieren, um ihn aufzurufen. Dies leisten semantische Dienste durch eine Liste, die spezifiziert, was der Agent tun muss, um sie auszuführen.

Der Agent muss ferner benötigte Services auswählen und verbinden können. Voraussetzung dafür ist, dass diese Services nahtlos zusammenarbeiten, denn nur so können kombinierte Ergebnisse eine gültige Lösung sein.

Und schließlich sollte der Agent die Merkmale der Dienste überprüfen und überwachen können, während sich diese im Einsatz befinden.

Das W3C (World Wide Web Consortium) steht mit der Weiterentwicklung geeigneter Markup-Sprachen kurz davor, diesen Anforderungen genügen zu können. In diesem Zusammenhang lässt sich die Web Markup Languages Pyramid² anführen, an deren Basis XML steht und die nach oben hin etwa durch RDF und die neue Web Ontology Language (OWL) sowie deren Derivat OWL-S (für Services) fortgeführt wird.

Gegenwärtig bleiben jedoch Bedenken in Hinblick auf Logik, Syntaxfehler und Vertrauenswürdigkeit der veröffentlichten Informationen bestehen. Allerdings versprechen Fortschritte im Bereich der Reasoning-Forschung und im breiten Einsatz von Digital Rights Management auch diese Probleme in den Griff zu bekommen.

Mythos Ontologie: Fundament des SemWebs

Neben Web Services bilden auch der Einsatz von Ontologien einen wesentlichen Baustein des semantischen Webs.

Ontologien beschreiben Wissensräume, Wissensobjekte und ihre Zusammenhänge. Ontologien helfen also Wissen zu strukturieren und aus unterschiedlichen Perspektiven abzubilden. Hierbei entstehen Metamodelle, die je nach hierarchischer Anordnung sehr abstrakt, wie im Falle der ISO-Terminologienormung, oder sehr praxisorientiert sein können, wenn es zum Beispiel darum geht eine Firmenstruktur als Metamodell abzubilden. Ontologien als Wissensmodelle sind der Grundstein für zahlreiche Softwarelösungen im semantischen Web, weil Ontologien die Verwendung von Begriffen regeln und so Information mit Bedeutung versehen.

Ein maßgeblicher Technologiebaustein ist dabei das RDF, das vom W3C entwickelt wurde. RDF ist eine XML-Anwendung zur Beschreibung einfacher semantischer Zusammenhänge und wird insbesondere zur Definition von Metadaten für Internet-Seiten verwendet. Kennzeichnend für RDF ist die Verbindung von Ressourcen auf der einen Seite mit Eigenschafts- sowie Wertepaaren zum Beschreiben der Ressourcen auf der anderen.

Entwickelte man das RDF vornehmlich dazu, den Informationsaustausch zwischen Maschinen zu automatisieren, existieren auch Standards zur Optimierung der Mensch-Maschine-Interaktion. Zu den wichtigsten zählen hierbei die sogenannten Topic Maps (XTM), die bereits in einer Reihe von

² Vgl. <http://www.w3.org/2000/Talks/0906-xmlweb-tbl/slide9-2.html>

Softwaretools zum Einsatz kommen. Topic Maps bauen allerdings auf andere Konventionen der Ontologiemodellierung.³

Ontologiebasierte Applikationen sind dort von Nutzen, wo es darum geht, einen Dialog zwischen dem User und der Maschine herzustellen und durch „Rückfragen“ der Maschine Suchanfragen immer feiner zu spezifizieren und zu kontextualisieren. Während das Gros der Suchmaschinen nicht imstande ist, die Suchanfrage des Anwenders im menschlichen Sinne „zu verstehen“, können ontologiebasierte Suchmaschine nach Dokumenten mit mehrdeutigen Begriffen wie „Kohl“, „Bank“ oder „Jaguar“ suchen und dem Kontext entsprechend die passenden Ergebnisse extrahieren. Semantische Suchmaschinen entsprechen so dem Wesen des menschlichen Denkens und erlauben im höheren Reifegrad sogar automatische, logische Schlussfolgerungen.

Dabei liegt die Wertschöpfung durch Ontologien weniger darin, dass Computer „denken“ lernen als darin, dass sie Informationen Bedeutung verleihen und die Qualität des Informationsflusses steigern.

Das semantische Web als Wirtschaftlichkeitsfaktor

Das semantische Web ist letztlich die logische Weiterentwicklung des Internets (und firmeninterner Intranets) zu einem intelligenten und menschengerechten Kommunikations- und Informationssystem.

Die Kosteneinsparungspotentiale liegen auf der Hand: Immer mehr Unternehmen haben erkannt, dass ihre Mitarbeiter zu viel Arbeitszeit damit verbringen, nach Dokumenten oder Ansprechpartnern zu suchen. Die Möglichkeiten, mit Hilfe semantischer Netze Information im Unternehmen kontextsensitiv und personalisiert aufzubereiten, münden in einer höheren Qualität der Informationsverarbeitung.

Das Zugpferd in Sachen wirtschaftlicher Nutzen findet sich aber sicherlich beim Thema Integration. Durch das Zusammenwachsen von semantischem Web und semantischen Web Services ergeben sich in diesem Zusammenhang immense Möglichkeiten.

Welche Nutzenpotentiale insgesamt im SemWeb liegen, kann man gegenwärtig nur schwer in Zahlen gießen. Die wirtschaftlichen Vorteile semantischer Systeme sind stets von der konkreten Anwendung abhängig und immer im Gesamtkontext des Zusammenspiels von Technologie, Mensch sowie Organisation zu betrachten.

³ Für einen Vergleich unterschiedlicher Softwaretools siehe Blumauer et al. (2003). Werkzeuge für Wissensmanagement. Plattform Wissenmanagement, abrufbar unter <http://www.pwm.at/41/artikel/14/14.html>

Fazit: Das semantische Web steht vor dem Markt-Rollout

Statt verfrüht ein Loblied auf das semantische Web anzustimmen, ist es besser, nun von Fall zu Fall den Nutzen zu untersuchen.

Das lehrt nicht zuletzt das Beispiel der künstlichen Intelligenz, die schon in den 50er Jahren zu Euphoriestürmen führte und große Erwartungen in die ersten Anwendungen weckte. Erfüllt haben sich diese nicht im gleichen Maße.

Das semantische Web ist jedoch weit mehr als „Artificial Intelligence“ und baut auf einer Entwicklung, die innerhalb weniger Jahre zu einem integralen Bestandteil der Unternehmensrealität geworden ist: das Internet.

Das semantische Web steht nun davor, seinen Kinderschuhen zu entwachsen: Tim Berners Lee, Director des W3C und Semantic-Web-Visionär, empfiehlt Unternehmen bereits bei der Anschaffung von Software darauf zu achten, dass eine Exportfunktion in RDF gegeben ist. Google & Co. experimentieren bereits jetzt mit den Einsatzmöglichkeiten semantischer Technologien⁴.

Unternehmen, die die Bedeutung der Ressource Wissen entsprechend ernst nehmen und bereits jetzt diese Entwicklungen antizipieren, haben in der Zukunft entscheidende Wettbewerbsvorteile, da sie auf den neuerlichen Akzelerationsschub des Internets besser vorbereitet sind, und nicht zuletzt aufgrund des höheren Problembewusstseins mit geringeren Migrationskosten konfrontiert sind.

Kontakt:

Andreas Blumauer (blumauer@punkt.at)
Tassilo Pellegrini (t.pellegrini@semantic-web.at)

Die Autoren Andreas Blumauer und Tassilo Pellegrini sind beide Projektmanager bei der Semantic Web School – Zentrum für Wissenstransfer mit Sitz in Wien/Österreich (SWS). Im Zentrum der Semantic Web School steht die Vermittlung von Praxiskompetenz bei der Integration von semantischen Technologien in Organisationen. Die SWS vernetzt Kompetenzträger und Interessierte im Bereich Semantic Web und hilft wissenschaftliche Erkenntnisse aus der Semantic Web Forschung für die unternehmerische Praxis nutzbringend einzusetzen. Die SWS ist ein Gemeinschaftsprojekt von punkt. netServices, Österreichische Nationalbibliothek, itl – Institut für technische Literatur und APA-IT. Die Durchführung eines modularen Lehrgangs erfolgt in Kooperation mit 13 weiteren Institutionen aus dem ganzen deutschsprachigen Raum. Weitere Informationen unter <http://www.semantic-web.at>.

⁴ Siehe dazu Google Suggest, Google Sets oder Personalized Web Search auf <http://labs.google.com/>