

ONTOLOGIEN: STAND UND ENTWICKLUNG DER SEMANTIK FÜR DAS WORLD WIDE WEB

von Katrin Weller

Einleitung: Die Vision vom Semantic Web

Die Idee zu einem semantischen Web wurde maßgeblich geprägt (wenn auch nicht initiiert) durch eine Veröffentlichung von Tim Berners-Lee, James Hendler und Ora Lassila im Jahre 2001. Darin skizzieren die Autoren ihre Version von einem erweiterten und verbesserten World Wide Web: Daten sollen so aufbereitet werden, dass nicht nur Menschen diese lesen können, sondern dass auch Computer in die Lage versetzt werden, diese zu verarbeiten und sinnvoll zu kombinieren. Sie beschreiben ein Szenario, in dem „Web agents“ dem Nutzer bei der Durchführung komplexer Suchanfragen helfen, wie beispielsweise „finde einen Arzt, der eine bestimmte Behandlung anbietet, dessen Praxis in der Nähe meiner Wohnung liegt und dessen Öffnungszeiten mit meinem Terminkalender zusammenpassen“. Die große Herausforderung liegt hierbei darin, dass Informationen, die über mehrere Webseiten verteilt sind, gesammelt und zu einer sinnvollen Antwort kombiniert werden müssen. Man spricht dabei vom Problem der Informationsintegration (Information Integration). Diese Vision der weltweiten Datenintegration in einem Semantic Web wurde seither vielfach diskutiert, erweitert und modifiziert, an der technischen Realisation arbeitet eine Vielzahl verschiedener Forschungseinrichtungen.

Einigkeit besteht dahingehend, dass eine solche Idee nur mit der Hilfe neuer bedeutungstragender Metadaten verwirklicht werden kann. Benötigt werden also neue Ansätze zur Indexierung von Webinhalten, die eine Suche über Wortbedeutungen und nicht über bloße Zeichenketten ermöglichen können. So soll z.B. erkannt werden, dass es sich bei „Heinrich Heine“ um den Namen einer Person handelt und bei „Düsseldorf“ um den Namen einer Stadt. Darüber hinaus sollen auch Verbindungen zwischen einzelnen Informationseinheiten festgehalten werden, beispielsweise dass Heinrich Heine in Düsseldorf wohnte. Wenn solche semantischen Relationen konsequent eingesetzt werden, können sie in vielen Fällen ausgenutzt werden, um neue Schlussfolgerungen zu ziehen.

Nehmen wir beispielsweise den Fall, dass eine Person A einen Bruder (Person B) hat, der wiederum einen Sohn hat (Person C). Mit dem richtigen Hintergrundwissen lässt sich daraus ableiten, dass Person B folglich der Onkel von Person C ist (Beispiel modifiziert nach Daconta et al., 2003).

Derartige Informationen über die Bedeutung einzelner Worte oder anderer Informationseinheiten und deren Beziehungen untereinander sollen ergänzend an bestehende Webseiten angefügt werden und so eine zusätzliche „semantische“ Ebene zum traditionellen Internet bilden, wie Berners-Lee, Hendler und Lassila (2001) es beschreiben:

- „The Semantic Web is not a separate Web but an extension of the current one, in
- which information is given well-defined meaning, better enabling computers and
- people to work in cooperation.“

Den Entwicklern geht es dabei nicht darum, einem Computer tatsächlich beizubringen, Informationen zu verstehen (wie es in der Künstlichen-Intelligenz-Forschung oftmals angestrebt wird), sondern vielmehr darum, Internetdienste für praktische Anwendungsszenarien zu entwickeln (Hitzler et al., 2008). Computer sollen also nicht lernen, „normale“ Webseiten zu lesen, sondern die Webseiten sollen mit zusätzlichen Informationen versehen werden, die für den Computer interpretierbar sind.

Indexierung im Sinne des Semantischen Webs basiert also auf komplexen Metadaten oder sogenannten „Semantic Markups“ (Pan, 2007). Das sind maschinenlesbare Index-Terme, welche einer Webseite (bzw. einzelnen Informationseinheiten innerhalb einer Webseite) zugewiesen werden. Aus technischer Sicht erfordert diese Idee neue Entwicklungen im Bereich maschinenlesbarer Auszeichnungssprachen, die solch eine feingliedrige Indexierung ermöglichen.

Über die technische Ebene hinaus benötigt das Semantische Web noch etwas anderes: hochwertige Wissensrepräsentationsmodelle, welche das Wissen eines oder mehrerer Interessensgebiete in einer allgemein akzeptierten Form verzeichnen. Man braucht ein gemeinsames kontrolliertes Vokabular für die Indexierung im Sinne eines Semantic Web, das wesentliche Begriffe und deren Beziehungen zueinander umfasst, eine Ontologie.

Für die Realisierung der Semantic Web Vision werden also sowohl standardisierte und kompatible Technologien zur Erstellung und Anwendung von semantischen Metadaten benötigt als auch Ontologien als komplexe und ausdrucksstarke Wissensrepräsentationssysteme.

Des Weiteren reicht es nicht, diese Techniken und Wissensmodelle bereitzustellen, sie müssen darüber hinaus auch noch auf die bestehenden Internetseiten angewandt werden. Das alles macht die Idee vom Semantic Web zu einer enormen Herausforderung für alle Beteiligten – und wirft vielfach Zweifel daran auf, ob ein solches semantisches Web jemals umfassend realisiert werden kann. Eine beachtliche Forschergemeinschaft hat diese Herausforderung angenommen und arbeitet hoch engagiert an den Grundlagen des Semantic Web. Die Semantic Web Science Association¹ (SWSA) hat sich dabei als eine treibende Kraft etabliert und auch das World Wide Web Consortium² (W3C) unterstützt die Entwicklung von Standards. Legg (2007) nennt das Semantic Web auch „the most ambitious project the W3C has scaffolded so far“.

In Bezug auf die technische Infrastruktur sind dabei bereits einige wesentliche und grundlegende Innovationen gelungen. Shadbolt, Berners-Lee und Hall (2006) nennen vor allem die folgenden Aspekte als bemerkenswerte Meilensteine der Semantic-Web-Forschung:

- Ontologiesprachen zur standardisierten Wissensrepräsentation (wobei die Web Ontology Language OWL³ mit ihren drei Dialekten derzeit die wohl bekannteste Ontologiesprache ist),
- Regelmechanismen und zugehörige Algorithmen für das automatische Schließen zur Ableitung impliziter Informationen (*Reasoning/ Inference Technologies*),
- Speichertechnologien für große Ontologien (insbesondere sogenannte *Triple Stores*),
- Abfragesprachen, um Informationen in Ontologien zu suchen, und
- Methoden, um strukturierte Informationen aus Webseiten zu extrahieren (*Information Extraction*).

Hinzu kommen noch zahlreiche Programme, die bei der Erstellung und Verwaltung von Ontologien helfen, sogenannte Ontologie-Editoren (das wohl bekannteste Beispiel hierfür ist Protégé⁴).

Ontologien als neue Methode der Wissensrepräsentation

Diese technologischen Entwicklungen bilden also den Grundstein für semantische Indexierung im WWW, enthalten jedoch per se noch keinerlei semantische Informationen. Eine Ontologiesprache allein macht noch keine Ontologie. Erst wenn mit einer solchen Sprache Wissensbausteine umgesetzt werden, kann eine semantische Repräsentation entstehen, die wiederum zur Indexierung des Webs herangezogen werden kann. Andererseits ist die Formalisierung in einer Ontologiesprache notwendig, damit ein Wissensmodell auch tatsächlich im Sinne des Semantic Web eingesetzt wird – um beispielsweise automatische Schlussfolgerungen und Konsistenzprüfungen zu ermöglichen (Gómez-Pérez et al., 2004). Ontologien sollen also die technischen und semantischen Anforderungen an eine zukunftsfähige Wissensrepräsentationsmethode für das semantische Web erfüllen.

¹ The Semantic Web Science Association Web (SWSA): <http://www.iswsa.org/>.

² World Wide Web Consortium (W3C): <http://www.w3.org>. Die Seiten des W3C zum Thema Semantic Web sind unter der URL <http://www.w3.org/2001/sw/> zu finden.

³ Web Ontology Language (OWL): <http://www.w3.org/TR/owl-guide/>.

⁴ Protégé: <http://protege.stanford.edu/>.

Die Bezeichnung Ontologie leitet sich ab aus dem griechischen *Ontologia*, was in etwa soviel bedeutet wie „Lehre vom Sein“. Es bezeichnete ursprünglich eine philosophische Disziplin, die sich mit den Grundstrukturen und dem Wesen allen Seins befasst. Im Rahmen der Computerwissenschaft wurde der Term Ontologie aufgegriffen und mit einer eigenen Bedeutung versehen. Allgemein gefasst bezeichnet Ontologie hier eine formale Konzeptualisierung eines Wissensbereichs („a formal conceptualization of a knowledge domain“, Gruber, 1993). Zahlreiche weitere Begriffserklärungen wurden und werden in der Literatur zitiert, eine endgültige und eindeutige Definition, die von allen beteiligten Forschungsgruppen geteilt wird, gibt es dabei allerdings bis heute nicht.

Die Formalisierung von Themenbereichen in Ontologien erfolgt typischerweise in Form von *Klassen* (bzw. *Konzepten*) und *Instanzen* (bzw. *Individuals*) – die Terminologie von Ontologiebestandteilen und deren genaue Ausprägungen können jedoch von Ontologiesprache zu Ontologiesprache variieren. Klassen repräsentieren dabei Allgemeinbegriffe, also abstrakte Konzepte eines Interessensgebiets, und sind hierarchisch über IS-A Beziehungen strukturiert (im Sinne der Hyponymie); Instanzen repräsentieren Individualbegriffe, also konkrete Vertreter der einzelnen Klassen, die in der jeweils untersten Hierarchieebene angesetzt werden.

Beispielsweise könnte es eine Klasse „Personen“ mit einer Unterklasse „Schriftsteller“ geben, welche wiederum mit einer Instanz „Heinrich Heine“ verbunden ist. Konzepte bzw. deren Instanzen können weiter in ihrer Bedeutung spezifiziert werden, oft geschieht dies durch sogenannte *Properties*, welche die Konzepteigenschaften über semantische Relationen darstellen. Zum Einsatz von *Properties* gibt es verschiedene Möglichkeiten: entweder erfasst die *Property* eine Beziehung zwischen zwei Konzepten/Instanzen (z.B. „ist_verwandt_mit“ oder „ist_einTeil_von“) oder die *Property* weist einem Konzept/ einer Instanz eine Eigenschaft mit einem gewissen Wert zu (z.B. „hat_Geburtsjahr“ mit einer Jahreszahl als Wert, ohne dass alle denkbaren Jahreszahlen als Instanzen in die Ontologie aufgenommen werden müssen). Auf diese Weise können Fakten aus dem betrachteten Themengebiet in der Ontologie gespeichert werden. Die Ontologie ist nicht mehr nur Indexierungsvokabular, sondern wird selbst zu einer Wissensbasis. Abbildung 1 zeigt, wie diese Ontologiebausteine mit Hilfe des Ontologie-Editors Protégé erstellt werden. Anschließend werden die Daten z.B. als OWL-Datei ausgegeben und weiterverarbeitet.

Schließlich besteht in formalen Ontologien noch die Möglichkeit, Klassen mit Hilfe von *Restrictions* (oder *Axiomen*) formal fest zu definieren. Diese fixieren Aussagen, die generell für die betrachtete Domäne gültig sind. Gehen wir davon aus, dass in einer Ontologie bereits eine Klasse „Schriftsteller“ mit verschiedenen Schriftstellern als Instanzen sowie eine *Property* „hat_Nationalität“ mit den verschiedenen Nationalitäten, die den Schriftstellern zugewiesen werden, verfügbar sind. Eine neue Klasse „deutsche Schriftsteller“ könnte dann formal so definiert werden, dass es sich dabei um die Untermenge der Klasse „Schriftsteller“ handelt, für die die *Property* „hat_Nationalität“ auf jeden Fall den Wert „deutsch“ annimmt. Ein Reasoner kann dann eine Instanz der Klasse „Schriftsteller“, die diese Bedingung erfüllt, automatisch der Klasse „deutsche Schriftsteller“ zuweisen. Je nach Ontologiesprache stehen unterschiedlich ausdrucksstarke *Restrictions* für die formale Definition von Klassen zur Verfügung.

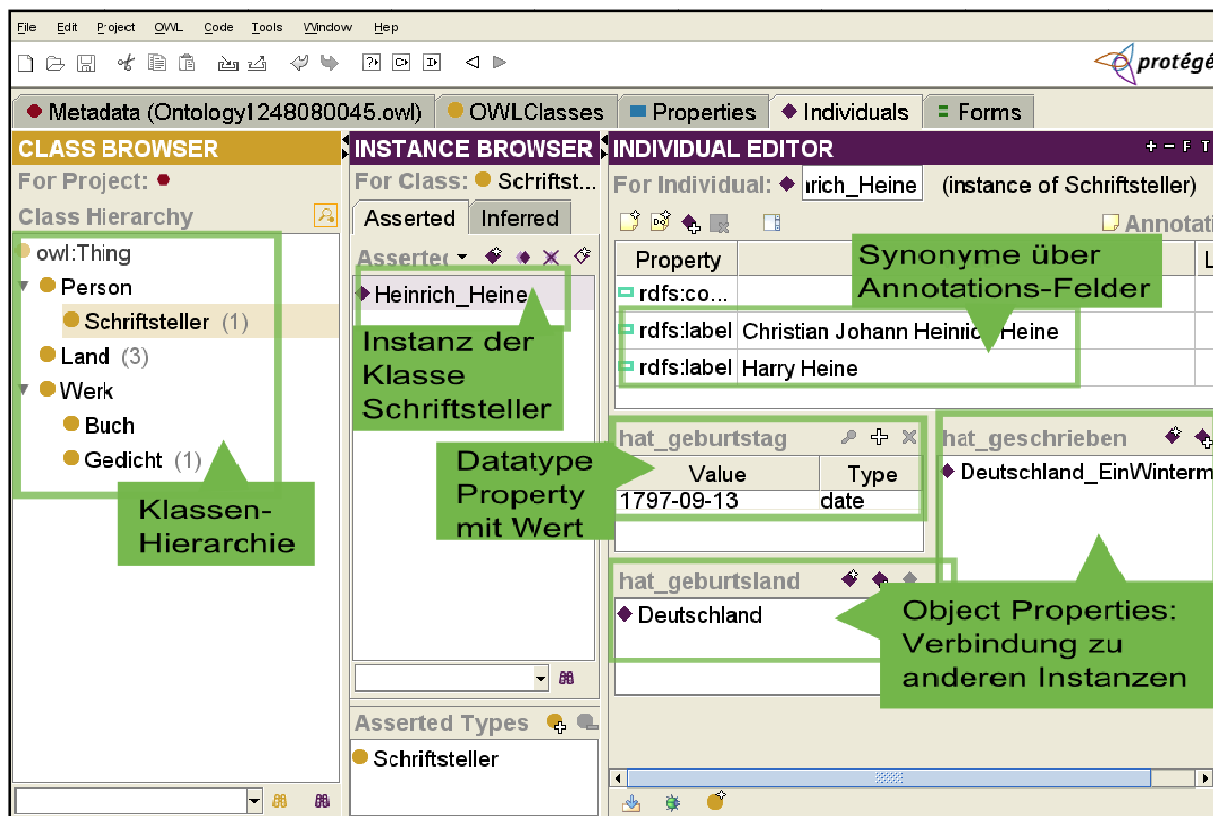


Abbildung 2: Ein Beispiel für typische Ontologie-Elemente in der Benutzeroberfläche des Ontologie-Editors

Eine Ontologie, die all diese Möglichkeiten ausschöpft, unterscheidet sich deutlich von klassischen Methoden der Wissensrepräsentation wie Thesauri und Klassifikationen. Mit Klassenhierarchien und Properties können Ontologien nicht nur das Standard-Repertoire semantischer Relationen abdecken (Hyponymie, Meronymie, Synonymie und Assoziationsrelation) (Peters & Weller, 2008a), sie ermöglichen auch den Einsatz weiterer spezifizierter Verbindungen zwischen den Konzepten eines Interessensgebietes und bilden so die Grundlage für verfeinerte semantische Repräsentationen. Der Einsatz von Datatype Properties eröffnet weitere Möglichkeiten für die präzise Definition einzelner Konzepte und die Einbindung von Fakteninformationen über Instanzen. Die formale Unterscheidung zwischen Klassen und Instanzen ist so bislang in keinem klassischen Wissensrepräsentations-Typus realisiert. Darüber hinaus wird eine Ontologie durch den Einsatz einer standardisierten Ontologiesprache maschinenlesbar und interpretierbar.

Mit der Ausdrucksmächtigkeit von Ontologien ist umgekehrt jedoch auch ein Problem verbunden: Diese detaillierte Darstellung lässt sich in der Praxis oft nur für eng begrenzte Themenbereiche umsetzen. Nur wenige Projekte wie die Cyc⁵ Ontologie versuchen sich daran, detailliert universelles Weltwissen abzubilden. Andere Ontologieprojekte befassen sich eher mit der Erstellung von Wissensmodellen für spezialisierte, eng abgegrenzte Themengebiete, wie etwa verschiedenen Ontologien im Bereich der Life Sciences (Bodenreider & Stevens, 2006). Dies liegt zum einen an dem hohen Aufwand, den die Erstellung eines feingliedrigen Wissensrepräsentationssystems mit sich bringt. Zum anderen liegt es bis zu einem gewissen Grad auch noch an Schwierigkeiten in der Handhabung der großen Datenmengen, die hinter einem Wissensmodell stecken, das in einer formalen Ontologiesprache verfasst ist. Die praktische Umsetzung von Ontologien unter Ausnutzung des Modellierungsspektrums von Properties und Restrictions erfordert Fachkenntnisse und einige Übung.

Mit der Vision eines neuen, komplexeren und hoch-ausdrucksstarken Begriffssystems können aktuelle Ontologieprojekte so auch oftmals noch nicht mithalten. Viele Ontologien, die derzeit für den Ein-

⁵ Cyc: <http://www.cyc.com>.

satz im Semantic Web entwickelt und diskutiert werden, übertreffen kaum die Ausdrucksfähigkeit traditioneller Thesauri. Oft zeichnen sie sich vor allem dadurch aus, dass sie in einer Ontologiesprache wie OWL abgespeichert wurden – ohne jedoch die formalen Möglichkeiten auszuschöpfen. Die aktuelle Lage von Ontologien als Mittel zur semantischen Anreicherung des WWW sieht also noch recht verhalten aus.

Die Umsetzung der Semantic Web Idee

Trotz zahlreicher ambitionierter Forschungsprojekte ist das Semantic Web insgesamt noch deutlich davon entfernt, weltweit Realität zu werden und das „herkömmliche“ Internet abzulösen. Doch woran liegt das? Verschiedene Erklärungsversuche wurden und werden diskutiert – ebenso wie einzelne Lösungsansätze.

Der Mangel an verfügbaren, ausdrucksstarken Ontologien hängt wechselseitig zusammen mit dem Fehlen eines publikumswirksamen Semantic-Web-Dienstes: Solange eine überzeugende Muster-Anwendung fehlt, sind wenige bereit, Zeit und Aufwand in die Erstellung von Ontologien zu investieren. Umgekehrt können semantische Dienste ohne zu Grunde liegende Ontologien nicht flächendeckend in Kraft treten. So bleiben Anwendungen wie ontologie-basierte Sucherweiterungen oder automatische Faktenextraktion aus Texten bislang begrenzt auf spezialisierte Einzelanwendungen für eng begrenzte Fach-Communities. Gesucht wird jedoch die sogenannte „Killer App“ des Semantic Web, ein Dienst, der gleichzeitig überzeugend effektiv und einfach nutzbar ist und so große Teile der Internetnutzerschaft einbezieht und ein Leitbild für nachfolgende Dienste werden kann. Bislang scheiterte die Praxis an einem solchen Angebot.

Der wohl am leichtesten nachvollziehbare Grund für den ausbleibenden Durchbruch des Semantic Web ist die enorme Größe des World Wide Web. Die Anreicherung des gesamten WWW mit semantischen Metadaten ist eine schier endlose Aufgabe. Das bedeutet einerseits, dass ein riesiger (und stetig wachsender) Indexierungsaufwand zu bewältigen ist, und andererseits, dass eine Vielzahl von Themenbereichen durch Ontologien erfasst und abgedeckt werden muss. Ein umfassendes Semantic Web wird daher in direkter Zukunft sicherlich noch keine Realität.

Von semantischen Technologien werden wir aber dennoch im Internet profitieren können. Der Forschungsschwerpunkt hat sich von einem einzigen, allumfassenden semantischen WWW weg verlagert, hin zu einzelnen semantischen Diensten (*Semantic Applications*) für den normalen Internetnutzer, also nicht nur für spezielle Expertengruppen. Immer mehr Internetdienste nutzen technische Grundlagen des Semantic Web innovativ aus (Carstens, 2008), auch für Einsatzbereiche in geschlossenen Systemen wie Intranets werden Anwendungen erarbeitet. Einige der jüngsten Entwicklungen konnten beispielsweise als Demos auf den letzten International Semantic Web Conferences begutachtet werden (siehe z.B. Bizer & Joshi, 2008), darunter neue semantische Recommender-Systeme, Systeme für semantische Bildersuche oder die Verknüpfung typischer Computerprogramme des täglichen Lebens zum „semantischen Desktop“. Die Entstehung einer Semantic Web Killer Application ist damit möglicherweise schon deutlich näher gerückt.

Mit der Fokussierung auf ein dezentralisiertes Semantic Web bestehend aus verschiedenen Einzeldiensten, die sich der gleichen semantischen Technologien bedienen, gehen einige weitere Herausforderungen einher. Wesentlich wird sein, dass diese Einzeldienste letztlich sinnvoll miteinander interagieren können. Die Verbindung von Datenbeständen wird derzeit zu einem der Schwerpunktthemen der Semantic-Web-Forschung, damit einzelne isolierte „Data Silos“ verhindert werden (Bizer et al., 2007). Aber auch die Möglichkeiten zur Verbindung verschiedener Ontologien müssen noch stärker in den Blickpunkt rücken, als es bislang der Fall ist. Hier werden derzeit zwar zahlreiche Techniken und Algorithmen für das Verknüpfen (*Mapping*) und Verschmelzen (*Merging*) von Ontologien entwickelt, es fehlen jedoch noch Ideen, wie dies im Rahmen der tatsächlichen Nutzung des Internets praktisch umgesetzt werden kann. Welche Konsequenzen ergeben sich, wenn zahlreiche kleine Einzelontologien miteinander interagieren müssen?

Insgesamt wurde bisher bei der Semantic-Web-Forschung der Schwerpunkt vor allem auf technologische Entwicklungen gesetzt. So fand auch eine Standardisierung bislang auf technischer Ebene statt, insbesondere bezogen auf Ontologiesprachen als standardisierte Austauschformate. Auf semantischer Ebene fehlen jegliche Standardisierungen bislang: Darüber, wie bestimmte Bedeutungen einheitlich repräsentiert werden sollen, wird nur am Rande der Semantic Web Community diskutiert. Das heißt vor allem, dass beim Aufbau einer Ontologie keinerlei Vorgaben verfügbar sind, wie bestimmte Zusammenhänge abzubilden sind. Klassendefinitionen werden von jedem Ontologie-Entwickler selbst gewählt und in der Regel höchstens in kleinen Teams diskutiert. Auch kann derzeit jeder, der eine Ontologie aufbaut, beliebige Properties zur Verbindung von Ontologie-Klassen erstellen und einsetzen. Gemeinsame, gar weltweit gültige Vereinbarungen z.B. bezüglich der Formulierung von Teil-Ganzes-Beziehungen zwischen Konzepten fehlen.

Mit der Wende in Richtung einzelner Semantic Applications scheint auch der Ansatz einer gemeinsamen Wissensrepräsentation zunächst aus dem Blick verloren. Dabei gilt die Idee einer Konsens-Abbildung im Sinne der „Shared Conceptualization“ eigentlich als eine der Grund-Definitionen für Ontologien (Borst, 1997). In einem Semantic Web, das aus zahlreichen Einzelontologien in verschiedenen Anwendungen besteht, ist das sinnvolle Verknüpfen von Ontologien anhand gleicher bedeutungstragender Elemente entscheidend.

Das Fehlen von Standardisierungen setzt sich auf der Ebene der ontologie-basierten Indexierung fort. Strategien für das tatsächliche Indexieren von Web-Inhalten mit Ontologien sind generell noch nicht ausreichend entwickelt, so fehlen auch Richtlinien für ein einheitliches Vorgehen. Experimentiert wird derzeit vor allem mit verschiedenen automatisierten und semi-automatischen Verfahren. Zur Qualitätssicherung muss hier jedoch auch der Nutzer stärker mit in den Prozess einbezogen werden. Ansätze aus dem Bereich des Nutzer-fokussierten Web 2.0 könnten hier wertvolle Impulse liefern.

Folksonomy trifft Ontologie: Ein Ausblick auf das Social Semantic Web

Als ein Lösungsansatz für die Probleme des Semantic Web zeichnet sich derzeit die Verbindung von semantischen Technologien mit Web 2.0 (bzw. Social Web) Strategien ab. Man spricht hier von einem Social Semantic Web als neue Zielrichtung. Web 2.0 bezeichnet dabei eine in jüngster Vergangenheit stark zunehmende Tendenz im Internet, Dienste bereitzustellen, die maßgeblich von der aktiven Beteiligung ihrer Nutzer leben. Internetnutzer kommentieren beispielsweise das Tagesgeschehen, schreiben Nutzerberichte oder Lexikonartikel, teilen private Fotos und Videos mit anderen, legen Steckbriefe über sich an und vernetzen sich mit Freunden und Bekannten – und die Internetnutzer *indexieren* selbst ihre auf den verschiedenen Plattformen eingestellten Inhalte mit Hilfe freier Keywords, sogenannten Tags. Mit dem *Social Tagging*, dem unkontrollierten Verschlagworten durch Nutzer-Communities ist somit ein neues Indexierungsmodell geboren: die *Folksonomie* (Peters, 2009).

Durch das völlige Fehlen einer Vokabularkontrolle sowie jeglicher Regeln kann die Folksonomy als Gegenpol zu klassischen Methoden der Wissensrepräsentation aufgefasst werden. Und auch im Vergleich mit Ontologien als Semantic-Web-Indexierung erscheint sie zunächst als völlig konträre Herangehensweise. Letztlich wird jedoch auch das Semantic Web nicht ohne die Einbeziehung von Nutzer-Beiträgen zustande kommen können – und muss sich daher auch mit solchen Alternativlösungen zunehmend auseinandersetzen. Sowohl bei der Erstellung umfassender Ontologien als auch bei ihrer Anwendung auf Dokumente ist die Mithilfe einer interessierten Nutzergemeinschaft von großer Bedeutung. Die Semantic-Web-Forschergemeinschaft befasst sich bereits seit einiger Zeit mit Werkzeugen zum gemeinschaftlichen Ontologieaufbau. Derartige Ansätze werden nun massiv ausgebaut und in verschiedener Hinsicht durch Kombinationen von Semantic und Social Web erweitert werden (Ankolekar et al., 2007; Blumauer & Pellegrini, 2008; Greaves, 2007; Hotho & Hoser, 2007), darunter beispielsweise semantische Wikis (Buffa et al., 2008; Lange et al., 2008) oder Software für semantisches Blogging (Bojars et al., 2008). Die Forschungsfragen sind dabei sinngemäß vor allem ‚Was kann das Semantic Web vom Erfolg des Web 2.0 lernen?‘ und ‚Wie können semantische Technologien den Wert von Web 2.0 Angeboten noch steigern?‘ (Mika & Greaves, 2008).

Für das Semantic Web-Ziel der Annotation von Webseiten mit semantischen und maschinenlesbaren Metadaten kann das Web 2.0 durch die Einbeziehung der Nutzer dabei vor allem zwei Aspekte beisteuern: Arbeitskraft und das gesammelte Wissen der Community. Die Einbeziehung von großen Internet-Nutzergruppen ist sowohl für den Aufbau von Ontologien als auch in der Indexierung denkbar. Auch die Pflege von Ontologien (z.B. das langfristige Einpflegen neuer Instanzen) kann hierdurch auf einer umfassenden Ebene realisiert werden. Folksonomy-Nutzer haben den Sinn von inhaltsbeschreibenden Metadaten für die Organisation ihrer eigenen Dokumente (darunter auch viele nicht-textuelle Dokumente wie Fotos und Videos) erkannt und wenden Tags zum persönlichen Informationsmanagement an. Wenn es gelingt, den Nutzern auch die Vorteile kontrollierter Vokabulare und ontologie-basierter Schlussfolgerungsmechanismen spielerisch näher zu bringen, ist ein großer Schritt in Richtung des Semantic Web getan. Dies ist natürlich auch nicht ohne Aufwand möglich, insbesondere für die Gestaltung ansprechender Nutzer-Schnittstellen liegen hier einige Herausforderungen. Erste Ansätze finden sich vor allem im Bereich des *Tag Gardening* (Peters & Weller, 2008b).

Durch die Einbeziehung einer Community kann auch die Vision einer gemeinsam akzeptierten Bedeutungsebene wieder aufgenommen werden. Durch die direkte (über Diskussionen) und indirekte (z.B. über die Auswertung nutzergenerierter Tagging-Daten) Nutzer-Interaktion kann eine neue Form von „shared conceptualization“ definiert werden.

Die praktische Umsetzung ist auch hier natürlich noch mehr Vision als Realität, rückt aber zunehmend in das Bewusstsein der Forschung.

Fazit

Das Semantic Web ist die Vision eines erweiterten World Wide Web, in dem inhaltsbeschreibende, computerlesbare Metadaten eine neue Bedeutungsebene erschließen. In technischer Hinsicht ist das Semantic Web mit verschiedenen Ontologiesprachen, Ontologieeditoren und Reasonern bereits gut ausgestattet. Dennoch fehlt der entscheidende Punkt: ausgereifte, bedeutungstragende Ontologien sowie Strategien für ihre flächendeckende Anwendung im WWW.

Ontologien als Methoden der Wissensrepräsentation bieten zwar eine reiche semantische Grundstruktur, doch ihre weitreichende Anwendung auf das WWW wird derzeit noch durch zahlreiche Schwierigkeiten behindert. Dazu zählen insbesondere der enorme Aufwand, der für die Erstellung komplexer Wissensrepräsentationsmodelle aufgebracht werden muss und die Schwierigkeiten bei der Einigung auf eine gemeinsame Darstellungsform für eine große Nutzer-Community. Auf viel weniger kontrolliertem Wege hat das Web 2.0 mit seinen Folksonomy-basierten Social-Software-Diensten bereits für eine neue Bedeutungsebene im WWW gesorgt. Social Tagging ist der erste Schritt in Richtung einer neuen Web-Navigation basierend auf Wortbedeutungen. Die automatische Integration verschiedener Informationseinheiten z.B. für semantische Suchsysteme kann damit jedoch nicht realisiert werden. Langfristig muss sich das Soziale mit dem Semantischen Web in verschiedener Hinsicht verbinden, so dass nutzergenerierte Basis-Metadaten durch Ontologien in maschinenlesbare Semantik umgewandelt werden können.

Literatur

- Ankolekar, A., Krötzsch, M., Tran, T., & Vrandečić, D. (2007). The Two Cultures: Mashing Up Web 2.0 and the Semantic Web. In 16th International World Wide Web Conference (WWW 2007). Banff, Alberta, Canada (pp. 825–834). Red Hook, NY: Curran.
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, 284(5), 34–43.
- Bizer, C., Heath, T., Ayers, D., & Raimond, Y. (2007) Interlinking Open Data on the Web. In Proceedings of the Demonstrations Track 4th European Semantic Web Conference (ESWC2007). Innsbruck, Austria. Retrieved November 10, 2008, from <http://www.eswc2007.org/demonstrations.cfm>.

- Bizer, C., & Joshi, A. (Eds.) (2008). ISWC2008 Posters and Demonstrations: Proceedings of the Poster and Demonstration Session at the 7th International Semantic Web Conference (ISWC2008). Karlsruhe, Germany. CEUR Workshop Proceedings, 401.
- Blumauer, A., & Pellegrini, T. (Eds.) (2008). *Social Semantic Web: Web 2.0 – Was nun?* Berlin: Springer.
- Borst, W. (1997). *Construction of Engineering Ontologies for Knowledge Sharing and Reuse*. PhD Thesis, University of Twente, Netherlands. Retrieved November 18, 2008, from University of Twente, Netherlands: <http://doc.utwente.nl/17864/>.
- Bodenreider, O., & Stevens, R. (2006). Bio-Ontologies: Current Trends and Future Directions. *Briefings in Bioinformatics*, 7(3), 256–274.
- Bojars, U., Breslin, J. G., & Finn, A. Decker S. (2008). Using the Semantic Web for Linking and Reusing Data Across Web 2.0 Communities. *Journal of Web Semantics*, 6(1), 21–28.
- Buffa, M., Gandon, F. L., Sander, P., Faron, C., & Ereto, G. (2008). SweetWiki. *Journal of Web Semantics*, 6(1), 84–97.
- Carstens, C. (2008). Semantic Web Applications In- and Outside the Semantic Web. In S. Auer, S. Schaffert, & T. Pellegrini (Eds.). *Proceedings of I-SEMANTICS'08. International Conference on Semantic Systems*. Graz, Austria (pp. 25–33). Graz: J.UCS.
- Daconta, M. C., Obrst, L. J., & Smith, K. T. (2003). *The Semantic Web. A Guide to the Future of XML, Web Services and Knowledge Management*. Indianapolis, Indiana: Wiley.
- Gómez-Pérez, A., Fernández-López, M., & Corcho, O. (2004). *Ontological Engineering: Advanced Information and Knowledge Processing (3rd Print)*. London: Springer.
- Greaves, M. (2007). Semantic Web 2.0. *IEEE Intelligent Systems*, 22(2), 94–96.
- Gruber, T. (1993). A Translation Approach to Portable Ontology Specification. *Knowledge Acquisition*, 2(5), 199–220.
- Hitzler, P., Krötzsch, M., Rudolph, S., & Sure, Y. (2008). *Semantic Web*. Berlin, Heidelberg: Springer.
- Hotho, A., & Hoser, B. (Eds.) (2007). *Bridging the Gap between Semantic Web and Web 2.0: Workshop located at the European Semantic Web Conference (ESWC 2007)*. Innsbruck, Austria. Retrieved November 13, 2008, from <http://www.kde.cs.uni-kassel.de/ws/eswc2007/program.html>.
- Lange, C., Schaffert, S., Skal-Molli, H., & Völkel, M. (Eds.) (2008). *The Wiki Way of Semantics: Proceedings of the 3rd Semantic Wiki Workshop (SemWiki 2008) at the 5th European Semantic Web Conference (ESWC 2008), Tenerife, Spain*. CEUR Workshop Proceedings, Vol. 360.
- Legg, C. (2007). Ontologies on the Semantic Web. *Annual Review of Information Science and Technology*, 41, 407–451.
- Pan, J. Z. (2007). OWL for the Novice: A Logical Perspective. In C. J. O. Baker & K.-H. Cheung (Eds.), *Semantic Web: Revolutionizing Knowledge Discovery in the Life Sciences* (pp. 159–182). Boston, MA: Springer.
- Peters, I. (2009). *Folksonomies: Indexing and Retrieval in Web 2.0. (Knowledge and Information. Studies in Information Science; Vol. 1)*. München: Saur.
- Peters, I., & Weller, K. (2008a). Paradigmatic and Syntagmatic Relations in Knowledge Organization Systems. *Information - Wissenschaft & Praxis*, 59(2), 100–107.
- Peters, I., & Weller, K. (2008b). Tag Gardening for Folksonomy Enrichment and Maintenance. *Webology*, 5(3).
- Shadbolt, N., Berners-Lee, T., & Hall, W. (2006). The Semantic Web Revisited. *IEEE Intelligent Systems*, 21(3), 96–101.