

Ulrich Ultes-Nitsche

Web 3.0 – wohin geht es mit dem World Wide Web?

Grundlagen zum Social Semantic Web

In diesem Übersichtsartikel stelle ich in groben Zügen das Web 3.0 und die um dieses Konzept existierenden Technologien vor. Er soll als Einleitung zu den in dieser Ausgabe vorliegenden Fachbeiträgen dienen. Vor allem Neulinge, die sich mit diesem Thema zum ersten Mal beschäftigen, sollen von dieser Einführung profitieren können. Natürlich hoffe ich, dass auch Web-3.0-Kenner ihre Vorstellungen und Ideen hier wiederfinden. Grob gesprochen beinhaltet das Web 3.0 die Möglichkeit, dass Webseiten die Bedeutung ihrer Inhalte in einer eindeutigen, maschineninterpretierbaren Form mitteilen können, woraus sich vielfältige Möglichkeiten für das automatische Auffinden von Inhalten und für die am Web angebotenen Dienste ergeben.

Inhaltsübersicht

- 1 Vom Web 2.0 zum Web 3.0
 - 1.1 Ein Beispiel
 - 1.2 Die Grundidee des Web 3.0
- 2 Die Struktur von Ressourcen im Web 3.0
 - 2.1 Ontologien
 - 2.2 Eine existierende Ontologiesprache: OWL
 - 2.3 Ressourcenbeschreibungen
 - 2.4 Ressourcenstruktur
- 3 Soziale Webanwendungen
 - 3.1 Wikis
 - 3.2 Collaborative Tagging
- 4 Probleme
 - 4.1 Sicherheit
 - 4.2 Vertrauen
- 5 Kritische Würdigung und persönliche Einschätzung
- 6 Literatur

1 Vom Web 2.0 zum Web 3.0

Hinter dem Schlagwort Web 3.0 versteckt sich vor allem die Idee eines mit Bedeutung (Seman-

tik) angereicherten World Wide Web, das sogenannte Semantic Web. Das Web 2.0 lieferte uns bereits neue Möglichkeiten der sozialen Interaktion, sein Nachfolger, das Web 3.0, ermöglicht uns jetzt noch zusätzlich, nach dedizierten Inhalten maschinenunterstützt zu suchen, Inhalte in Relation zueinander zu setzen und Ähnliches. Hierzu müssen Webinhalte natürlich in geeigneter, d.h. maschineninterpretierbarer Form angeboten werden.

Webinhalte, die in maschinenlesbarer Form ihre Bedeutung »anbieten«, erlauben erst, die Idee von Webdiensten zu realisieren. Dies lässt sich vielleicht an einem Beispiel näher erläutern.

1.1 Ein Beispiel

Angenommen ich habe in einem Restaurant »Kaiserschmarren« gegessen. Ich weiß nicht genau, wie der zubereitet wird, möchte ihn aber gerne nachkochen. Im guten alten WWW hätte ich eine Suchmaschine bemüht, die mir zum Suchbegriff »Kaiserschmarren« Seiten von Restaurants geliefert hätte, die Kaiserschmarren auf der Speisekarte haben, vielleicht noch Firmen, die Kaiserschmarren tiefgekühlt anbieten, und ähnliche Seiten. Sicher wären auch Seiten mit Rezepten dabei gewesen, aber ich hätte in der Ergebnisliste der Suchmaschine selbst danach suchen müssen.

Im Web 2.0 wäre es vielleicht schon einfacher gewesen: Ich hätte ein soziales Forum aufgesucht, in dem sich Kaiserschmarrenliebhaber austauschen. Dort hätte ich mich über diverse Rezepte und Kniffe informieren können. Allerdings hätte ich auch hier erst einmal einen gewissen Suchaufwand betreiben müssen.

Die Vorgehensweise im Web 3.0 ist nun eine vollkommen andere: Der Web-3.0-Agent auf meinem Computer wird sich auf die Suche nach

Kaiserschmarrenrezepten machen. Infrage kommen nur Webseiten, die auch über sich sagen, dass sie Rezepte für Kaiserschmarren enthalten, und die meinem Agenten auch noch mitteilen, wie ihr Inhalt zu lesen ist, um daraus das Rezept zu extrahieren. Sollte mein Agent auch wissen, dass ich eine bestimmte Lebensmittelallergie habe, wird er nur Rezepte suchen, die bestimmte Inhaltsstoffe vermeiden. Kennt mein Web-3.0-Agent nicht alle Bestandteile des Rezepts, könnte er noch einen anderen Webdienst besuchen, der Informationen darüber liefert, welche Kochzutaten den Stoff enthalten, gegen den ich allergisch bin. Kennt der Web-3.0-Agent noch meine Vorlieben (z.B. wenig Rosinen, viel Zimt und einigermaßen viel Zucker), wird er ein geeignetes Rezept für mich finden. Weiß er darüber hinaus, welche Zutaten sich bereits in meiner Küche befinden, sucht er noch schnell nach einem Geschäft in meiner Nähe, das über das Web 3.0 mitteilt, dass es alle fehlenden Bestandteile des Rezeptes vorrätig hat, noch geöffnet ist und sogar innerhalb einer Stunde eine entsprechende Lieferung zu mir nach Hause machen kann. Er könnte dann sogar in meinem Namen die Bestellung aufgeben, wenn ich ihn entsprechend autorisiert habe.

1.2 Die Grundidee des Web 3.0

So weit die Vision des Web 3.0. Vieles existiert natürlich so noch nicht, aber viele Forscherteams arbeiten weltweit daran, dass diese Vision einmal Wirklichkeit werden könnte. Einer der bekanntesten Visionäre des Web 3.0 ist Sir Tim Berners-Lee, der »Vater« des WWW. Zusammen mit James Hendler und Ora Lassila hat er einen sehr empfehlenswerten Übersichtsartikel in *Scientific American* geschrieben [Berners-Lee et al. 2001]. Für ihn gilt die simple Gleichung in Abbildung 1.

Web 2.0 + Semantic Web = Web 3.0

Abb. 1: Gleichung für Web 3.0 nach Tim Berners-Lee

Statt mittels intelligenter Texterkennung zu versuchen, Webinhalte von ihrer Bedeutung zu verstehen, setzt Web 3.0 auf eindeutige Beschreibungssprachen der Semantik der Webinhalte. Damit ist die einzige Voraussetzung dafür, Web-3.0-Inhalte auf ihre Bedeutung hin zu analysieren, die Fähigkeit, diese formalen Beschreibungssprachen interpretieren zu können. Somit handelt es sich beim Web 3.0 um keine Anwendung von künstlicher Intelligenz – und es ist daher auch nicht den Grenzen künstlicher Intelligenz unterworfen –, sondern um eine Anwendung formaler Beschreibungssprachen, mit denen sogenannte Ontologien beschrieben werden.

Im Folgenden werde ich die meiner Ansicht nach wesentlichen Aspekte des Web 3.0 darstellen: Ontologien und RDF, soziale Webanwendungen sowie die Sicherheitsproblematik. Die Auswahl ist subjektiv getroffen mit der Zielsetzung, einen Überblick zu geben, um die nachfolgenden Beiträge dieser Ausgabe gut einordnen zu können.

Interessierten Leserinnen und Lesern ist der lesenswerte wissenschaftliche Übersichtsartikel [Horrocks & Patel-Schneider 2003b] zum *Semantic Web* empfohlen.

2 Die Struktur von Ressourcen im Web 3.0

2.1 Ontologien

Letztendlich ist das Ziel von Web 3.0, neben den bereits in Web 2.0 existierenden sozialen Aspekten, maschinenlesbare Wissensrepräsentationen im Web zur Verfügung zu stellen. Diese Wissensrepräsentationen heißen Ontologien [Uschold & Gruninger 1996].

Ontologien haben eine fest vorgegebene Struktur, und ein Web-3.0-Agent, der das Wissen aus der Repräsentation extrahieren möchte, muss die Kenntnis haben, wie die Elemente dieser Struktur semantisch zu interpretieren sind. Neben individuellen Elementen enthalten Ontologien typischerweise auch Eigenschaften. Diese drücken in logischer Form ein Wissen

über die Relation zwischen Elementen der Ontologie aus (dies sind typischerweise Einschränkungen über Wertekombinationen für die entsprechenden Elemente). Durch logische Ableitung lassen sich daraus aber auch andere Relationen entweder ableiten oder ausschließen. Ein gern verwendetes Beispiel zur Illustration dieses Sachverhalts ist das folgende: Angenommen wir haben eine Ontologie vorliegen, die Wissen über Personen darstellt. In dieser Ontologie gebe es verwandtschaftliche Relationen (z.B. über Eltern-Kind-Verhältnisse) sowie medizinisches Wissen (z.B. über Blutgruppen). Eine Eigenschaft wäre dann z.B. eine Aussage darüber, dass bestimmte Blutgruppenkombinationen eine Eltern-Kind-Beziehung zwischen Personen ausschließen (eine Person mit Blutgruppe o wird nie ein Elternteil mit Blutgruppe AB haben). Dies ist eine ausschließende Eigenschaft.

Eine oben angesprochene logische Ableitung wäre dann die Information, dass, wenn eine Eltern-Kind-Beziehung besteht und ein Elternteil Blutgruppe AB hat, das Kind nicht Blutgruppe o haben kann. Dies ist eine logische Ableitung, da die gegebene Eigenschaft zunächst etwas über mögliche Elternschaft in Abhängigkeit von Blutgruppen aussagt, wir aber, hieraus abgeleitet, eine Aussage über die mögliche Blutgruppe machen, abgeleitet aus einem Wissen über die Eltern-Kind-Beziehung.

2.2 Eine existierende Ontologiesprache: OWL

Die *Web Ontology Language* (OWL, nicht WOL) [W3C 2004] ist die wichtigste existierende Sprache zur Beschreibung von Ontologien. OWL wurde aus der Vorläufersprache DAML+OIL abgeleitet. Neben einer selbstverständlich eindeutigen Syntax hat OWL eine eindeutige Semantik; eine Eigenschaft, die ihre Vorgängersprachen nicht besaßen. OWLs eindeutige Semantik leitet sich aus der Äquivalenz von OWL zur Beschreibungslogik ab [Horrocks & Patel-Schneider 2003a].

OWL besteht aus:

- Klassen (benutzerdefinierte Klassen, die der Ontologie eine Struktur geben; Unterklassendefinitionen sind zulässig; z.B. eine Klasse Rezept mit der Unterklasse Kaiserschmarrenrezept),
- Instanzen (konkrete Ausprägungen von Klassen im Sinne der Objektorientierung; z.B. ein konkretes Kaiserschmarrenrezept),
- Eigenschaften (Relationen in Form logischer Prädikate zwischen Klassen, Instanzen und gegebenenfalls Datentypen; die Eigenschaft *Autor* könnte z.B. ein konkretes Rezept mit einer Person in Verbindung setzen) und
- Operationen (mengentheoretische Operation auf Klassen, die allgemein als Menge ihrer individuellen Ausprägungen (Instanzen) angesehen werden; z.B. könnte die Vereinigung von Kaiserschmarrenrezepten, Pfannkuchenrezepten usw. eine Rezeptinstanz Mehlspeisenrezepte liefern).

Ein Beispiel für eine konkrete OWL-Anwendung ist *GoodRelations*, eine Ontologie für kommerzielle Webangebote [Hepp 2008]. *GoodRelations* enthält ein reichhaltiges Vokabular, um alle notwendigen Elemente und Beziehungen in kommerziellen Webangeboten zu beschreiben. Beispielsweise gibt es ein Teilkokabular zur Preisbeschreibung, das neben dem eigentlichen Preis auch darüber Informationen liefert, welche Bezahlmöglichkeiten es gibt, wann der Preis gültig ist, ob aus der Zahlung Garantieansprüche erwachsen und Ähnliches.

2.3 Ressourcenbeschreibungen

Wie im vorherigen Abschnitt bereits angesprochen, tauchen in Ontologien Daten bestimmter Typen auf. Um eine geeignete Datendurchlässigkeit zu gewährleisten, braucht es im Web 3.0 Datenaustauschformate. Hierzu dient das RDF (*Resource Description Framework*; www.w3.org/RDF). Daten müssen eindeutig, d.h. maschineninterpretierbar, dargestellt werden. RDF gewährleistet dies. RDF ist eine Datenmodellie-

rungssprache ähnlich *Entity-Relationship-Diagrammen*. RDF ist XML-basiert (mit einer grafischen Interpretation). RDF Schema (RDFS) ist hierzu ein Mechanismus, die Syntax und einfache semantische Anforderungen an RDF-Dokumente auszudrücken, so wie XML Schemas für XML-Dokumente verwendet werden.

2.4 Ressourcenstruktur

Aus dem bisher Gesagten ergibt sich die in Abbildung 2 dargestellte Struktur für Web-3.0-Ressourcen.

Hierin ist SPARQL (SPARQL Protocol and RDF Query Language) die Schnittstelle zu den Anwendungen, die auf die Ressource zugreifen möchten [W3C 2008].

3 Soziale Webanwendungen

Wie in der Einleitung bereits erwähnt, lässt sich das Web 3.0 verkürzt als Web 2.0 + Semantic Web definieren. Was wir bisher vom Web 3.0 gesehen und besprochen haben, waren Aspekte des Semantic Web. Was noch fehlt, ist ein Hinweis auf ein paar der bereits im Web 2.0 existierenden sozialen Webanwendungen [Governor et al. 2009].

3.1 Wikis

Ein *Wiki* ist ein kollektiv editierbares Dokument, das häufig zusätzlich mit Foren und ähnlichen

Kommunikationsmitteln ausgestattet ist. In Wikis können Informationen von vielen Nutzerinnen und Nutzern gemeinsam erstellt und aktuell gehalten werden. Eines der prominentesten Beispiele für Wikis ist die Wiki-basierte Web-Enzyklopädie Wikipedia (www.wikipedia.org). Ein solches Wiki wird auch *syntaktisches Wiki* genannt.

Moderne Wikis erlauben es zusätzlich, ihren Inhalt in maschinenlesbarer Form zu kommunizieren. Diese Wikis heißen *semantische Wikis*. Hierbei muss von vornherein definiert sein, wie das im Wiki vorhandene Wissen nach außen repräsentiert wird: Zusätzlich zum Standardtext im Wiki muss es maschinenlesbare Elemente geben, die einer formalen Sprache folgen und damit eindeutig maschinell interpretierbar werden.

In Abbildung 3 sieht man die zu Anfang erwähnte Gleichung $\text{Web 3.0} = \text{Web 2.0} + \text{Semantic Web}$ besonders schön. Die Web-2.0-Anwendung Wiki wird um semantische Elemente erweitert und wird damit zu einer Web-3.0-Anwendung, die gegenüber ihrer Web-2.0-Schwester zusätzlich das Feature der Maschinenlesbarkeit der enthaltenen Informationen besitzt. Eine solche Erweiterung stellt natürlich auch Anforderungen an die Benutzer semantischer Wikis, die sich an die Regeln der formalen Beschreibungssprache halten müssen, um tatsächlich semantische Inhalte im Wiki zu erzeugen.

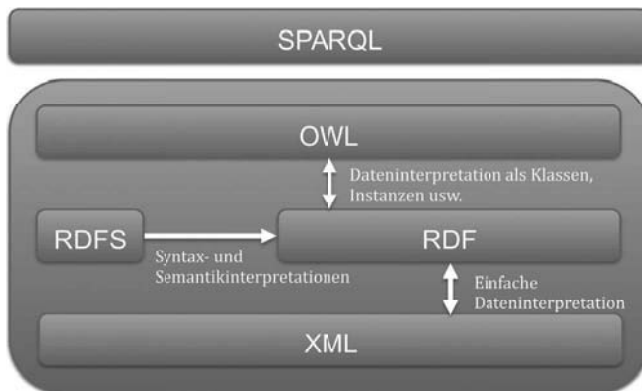


Abb. 2: Struktur für Web-3.0-Ressourcen



Abb. 3: Semantisches Wiki

3.2 Collaborative Tagging

Hinter dem Begriff *Collaborative Tagging* (auch *Social Tagging*, *Social Indexing* oder *Folksonomy* genannt) verbirgt sich die Idee, dass Nutzerinnen und Nutzer dieser Technologie Webinhalte mittels sogenannter *Tags* um Informationen in Form von Beschriftungen (Annotationen) ergänzen und dadurch Inhalte kategorisieren. Hierbei entstehen mehrfach annotierte Dokumente (z.B. durch die Nutzer Denise und Frank – vgl. Abb. 4).



Abb. 4: Collaborative Tagging

Collaborative Tagging ist auch ein interessantes sozialwissenschaftliches Forschungsthema, das sich mit kollaborativer Intelligenz beschäftigt sowie dem Entstehen von Wissensstrukturen aufgrund nicht moderierter Zusammenarbeit.

4 Probleme

4.1 Sicherheit

Das *Semantic Web* bietet unglücklicherweise einige Möglichkeiten zum Missbrauch. Da Ressourcen zumindest zum Teil von Maschinen ge-

lesen werden, die daraufhin weitere Entscheidungen automatisiert treffen können, ist es notwendig, dass Ressourcen auch korrekte Informationen liefern. Wenn einem Nutzer bzw. einer Nutzerin bereits vertrauenswürdige Ressourcen bekannt sind, muss unter anderem sichergestellt sein, dass ein Ressourcenzugriff auch tatsächlich bei der entsprechenden Ressource landet und nicht nur bei einer scheinbar identischen Kopie, die bewusst falsche Informationen liefert. Um dies zu gewährleisten, müssen kryptografische Verfahren eingesetzt werden, die die Authentizität der Ressourcen garantieren. Falls die Kommunikation mit der Ressource vertraulich stattfinden soll, braucht es ebenfalls Kryptografie.

Allerdings reicht Kryptografie alleine nicht aus. Sie schützt z.B. nicht gegen zwar bekannte Ressourcen, die sich jedoch bösartig verhalten. Web 3.0 ist daher eine Anwendung, die das Konzept des Vertrauens benötigt.

4.2 Vertrauen

Im Bereich der Informationssicherheit gibt es für Anwendungen wie das Web 3.0 seit einiger Zeit die Bestrebung, das soziale Konzept des Vertrauens auf die Kommunikation zwischen Maschinen zu übertragen. Bereits im Kontext von sicherer E-Mail-Kommunikation sind in den 1980er-Jahren Vertrauenskonzepte eingeführt worden (z.B. in PGP (Pretty Good Privacy) [Zimmermann 1995]). Zur Bildung von Vertrauen (im Kontext technischer Systeme) gibt es folgende Möglichkeiten:

Um Vertrauen aufzubauen, muss es eine Historie von Interaktionen der beteiligten Systeme geben, die von den Systemen gespeichert wird. Haben sich die Systeme über einen längeren Zeitraum korrekt verhalten, wird das Vertrauen in diese Systeme steigen. Außerdem werden Empfehlungsstrukturen verwendet, um Vertrauen aufzubauen (empfiehlt ein für mich vertrauenswürdigen System ein anderes System als vertrauenswertig, dann werde ich bereits zu Anfang ein höheres Vertrauen in dieses mir unbekannte System haben, als wenn ich eine erste Interaktion mit einem unbekanntem System durchführe). Letztendlich wird aus historischen Interaktionen und Empfehlungen ein numerischer Wert berechnet, der das Vertrauen relativ zu einer Vertrauensskala festlegt. Ob ein System mit einem anderen interagieren wird, hängt dann davon ab, ob der Vertrauenswert über einem vom System festgelegten Schwellwert liegt. Weiterhin helfen *Blacklists*, nicht vertrauenswürdige Systeme zu publizieren.

5 Kritische Würdigung und persönliche Einschätzung

In diesem Artikel habe ich kurz ein paar wesentliche Aspekte des Web 3.0 übersichtsartig vorgestellt: Die semantischen Eigenschaften des *Semantic Web* werden mit den sozialen Aspekten des Web 2.0 kombiniert. Damit ist das Web 3.0 eine echte Weiterentwicklung des Web 2.0.

Das Web 3.0 bietet sicher viele Möglichkeiten. Gleichwohl stellt es aufgrund seiner inhärenten Komplexität eine große Herausforderung dar. Sehr viele Teilsysteme müssen in konsistenter Weise miteinander interagieren, um die verschiedenen Aspekte des Web 3.0 zu realisieren. Die am häufigsten anzutreffende Kritik ist daher, dass das Web 3.0 zwar eine sehr schöne Wunschliste von zu realisierenden Komponenten darstellt, diese aber aufgrund der ungeheuren Komplexität des notwendigen Zusammenspiels nie zu den im Web 3.0 angestrebten Resultaten führen werden.

Dieser Kritik schließe ich mich zum Teil an, da ich auch nicht glaube, dass das Web 3.0, so wie stellenweise propagiert, in seinem vollen Umfang kommen wird (genau aus den genannten Komplexitätsüberlegungen). Dennoch ergeben sich aus der Forschung am Web 3.0 sehr viele interessante Ideen und Überlegungen, von denen wir zukünftig sicher profitieren werden (z.B. wie Computer grundsätzlich das, was sie können, in maschinenlesbarer Form vermitteln).

6 Literatur

- [Berners-Lee et al. 2001] *Berners-Lee, T.; Hendler, J.; Lassila, O.*: The Semantic Web. In: Scientific American Magazine, Mai 2001, www.scientificamerican.com/article.cfm?id=the-semantic-web; Zugriff am 22.11.2009.
- [Governor et al. 2009] *Governor, J.; Hinchcliffe, D.; Nickull, D.*: Web 2.0 Architectures. O'Reilly, 2009.
- [Hepp 2008] *Hepp, M.*: GoodRelations: An Ontology for Describing Products and Services Offers on the Web. In: Knowledge Engineering: Practice and Patterns, 16th International Conference EKAW 2008, Acitrezza, Italy, 2008, S. 332-347.
- [Horrocks & Patel-Schneider 2003a] *Horrocks, I.; Patel-Schneider, P. F.*: Reducing OWL Entailment to Description Logic Satisfiability. In: 2nd International Semantic Web Conference (ISWC2003), Florida, USA, 2003, S. 17-29, <http://ect.bell-labs.com/who/pfeps/publications/reducing-iswc.pdf>; Zugriff am 30.11.2009.
- [Horrocks & Patel-Schneider 2003b] *Horrocks, I.; Patel-Schneider, P. F.*: Three Theses of Representation in the Semantic Web. In: The 12th International World Wide Web Conference (WWW2003), Budapest, Hungary, May 2003, S. 39-47, <http://ect.bell-labs.com/who/pfeps/publications/representation-theses.pdf>; Zugriff am 01.12.2009.
- [Uschold & Gruninger 1996] *Uschold, M.; Gruninger, M.*: Ontologies: Principles, Methods and Applications. In: Knowledge Engineering Review, 11(2), 1996, S. 93-155.
- [W3C 2004] *World Wide Web Consortium*: OWL Web Ontology Language Reference. W3C Recommendation, 2004, www.w3.org/TR/owl-ref; Zugriff am 22.11.2009.

[W3C 2008] *World Wide Web Consortium: SPARQL Query Language for RDF*. W3C Recommendation, 2008, www.w3.org/TR/rdf-sparql-query; Zugriff am 01.12.2009.

[Zimmermann 1995] *Zimmermann, P.: PGP Source Code and Internals*. MIT Press, 1995. (Nicht mehr erhältlich; nützliche Informationen sind auf Phil Zimmermanns Homepage zu finden: www.philzimmermann.com/DE/background/index.html; Zugriff am 01.12.2009.)

Prof. Dr. Ulrich Ultes-Nitsche
Universität Fribourg
Departement für Informatik
Boulevard de Pérolles 90
CH-1700 Fribourg
uun@unifr.ch
www.unifr.ch



Herbert Kubicek, Steffen Brückner

**Businesspläne für
IT-basierte Geschäftsideen**

Betriebswirtschaftliche
Grundlagen anhand von Fallstudien

2010, 320 Seiten, Festeinband

€ 39,00 (D)

ISBN 978-3-89864-639-0



dpunkt.verlag

Ringstraße 19 B · D-69115 Heidelberg · fon: 0 62 21 / 14 83 40
fax: 0 62 21 / 14 83 99 · e-mail: bestellung@dpunkt.de · www.dpunkt.de