

Web 3.0: das Ende von Google

Eduard Schibrowski

Universität Konstanz, Information Engineering, Database and Information Systems Group,
Webservices and Semantic Web
eduard.schibrowski@uni-konstanz.de
<http://www.inf.uni-konstanz.de/dbis/teaching/ws0708/web/>

Abstract. Der Begriff "Web 3.0" soll eine Verknüpfung zwischen (Web 2.0) und (Semantik Web) darstellen. Die neue Technologie soll in Zukunft die klassische Suchtechnologie von Google ersetzen und dadurch eine neue Zeit eröffnen, die Zeit in der die natürliche Sprache dem Computer immer näher gebracht wird und in der die ganzen Webinhalte in einem logischen Zusammenhang zueinander stehen. Bevor man das Konzept, die Vorteile und Nachteile und die Technologie hinter dem Web 3.0 einführt, ist eine Erklärung die wichtigsten Eigenschaften von (Web 1.0) und (Web 2.0) notwendig. Die Darstellung des Tools ("GroupMe") dient der Verdeutlichung der eingeführten Technologien. Da die neue Entwicklung, aus der heutigen Sicht betrachtet, mehr eine Vision der Zukunft als eine Tatsache darstellt, werden im zweiten Teil des Papers eine Reihe von Meinungen von Spezialisten im Bereich präsentiert. Viele offene Fragen im Zusammenhang mit den neuen Entwicklungen und mit dem Schicksal der aktuellen Suchtechnologien prägen heutige Diskussionen in Fachforen. Einige davon, zusammen mit den häufigsten Antworten schließen das Paper ab.

1 Motivation

Über die Jahre hat das Internet eine immer größere Bedeutung erreicht. Es gibt immer mehr Firmen die eine Internetpräsenz haben und die Anzahl der Benutzer kannte in den letzten zehn Jahren ein exponentielles Wachstum. Dadurch ist die Menge der Daten und Inhalte im Internet kaum noch überschaubar. Darüber hinaus ist der Computer, zumindest was das Web anbelangt, rein zum darstellen der Informationen da, die Funktion der Verarbeitung der Information ist dem Computer im Web entnommen worden. Aus diesen Gründen ist die Idee von Semantik Web entstanden, um Inhalten Bedeutung und Bezug zu verleihen.

Relevant in diesem Zusammenhang ist die Vision von Tim Berners-Lee, in der, er die Zukunft als eine totale Vernetzung sieht. Es geht dabei nicht nur um die Vernetzung der einzelnen Computer und Netzwerke, sondern die Daten und die Informationen stehen in einem logischen Zusammenhang miteinander in Verbindung.

"The vision I have for the Web is about anything being potentially connected with anything" ¹

2 Auf dem Weg zu Web 3.0

Bevor man sich an die Ideen und Konzepte von Semantik web widmet, ist der lange weg dahin von großer Bedeutung. In dem Bild (Fig.1) sehen wir was alles in Web 1.0 vorhanden war, was sich mit der Zeit als Web 2.0 etabliert hat und welche Neuigkeiten Web 2.0 mit sich gebracht hat.

2.1 Web 1.0

Eine grosse Anzahl von Diensten und Technologien waren in Web 1.0 bereits vorhanden. Erwähnenswert sind die statischen Webseiten und Downloaddienste, einfache Suche und sogar die ersten Content Management Systeme. Um sich einen Überblick über den Umfang der Möglichkeiten von Web 1.0 zu verschaffen, ist eine Aufzählung der bekanntesten Anbieter der Zeit sehr hilfreich.

¹ Berners-Lee, Tim; Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web; HarperBusiness; 2000

Web 1.0	Web 2.0	Neu
DoubleClick	Google AdSense	personalisiert
Ofoto	Flickr	tagging, community
Akamai	BitTorrent	P2P
mp3.com	Napster	P2P
Britannica Online	Wikipedia	community, free content
personal websites	blogging	dialog
Evite	upcoming.org and EVDB	
domain name speculation	search engine optimization	
page views	cost per click	pay for participation
screen scraping	web services	interoperability
publishing	participation	
CMS	wikis	flexibility, freedom
directories (taxonomy)	tagging ("folksonomy")	community, freedom
stickiness	syndication	open content

Fig. 1. Neuigkeiten von Web 2.0 gegenüber von Web 1.0

- DoubleClick, das Internet Marketing Dienste anbietet
- Ofoto, bietet Möglichkeiten zum Sharen von Photos
- Akamai Technologies, Inc. ist einer der weltweit größten Anbieter für die Auslieferung und Beschleunigung von Online-Inhalten, zum Beispiel die Lastverteilung von WWW-Inhalten, auch als "Content Distribution Network" bekannt.
- EVDB steht für "Events and Venues Database"

Schon mit Web 1.0 waren Begriffe wie "screen scraping", "domain name speculation", "stickiness" zum täglichen Gebrauch der Fachleute bekannt.

2.2 Web 2.0

Mit Web 2.0 werden die Inhalte personalisiert, Media Daten mit Tags versehen für ein schnelleres Finden im Netz, entwickeln sich P2P Netzwerke, es wird immer mehr auf ein soziales Netz aufgebaut in der die Benutzer interagieren, freie Inhalte, Flexibilität und Freiheit treten in den Vordergrund. Trends wie Blogs, Wikis, Tagging setzen sich immer mehr durch. Aber am aller Wichtigsten in der Entwicklung von Web 2.0 ist die Verzweigung in einer technischen und einer sozialen Entwicklung. Dank der neuen AJAX "Asynchronous JavaScript and XML" Technologie werden die Benutzerschnittstellen flüssiger. AJAX bezeichnet ein Konzept der asynchronen Datenübertragung zwischen einem Server und dem Browser, das es ermöglicht, innerhalb einer HTML-Seite eine HTTP-Anfrage durchzuführen, ohne die Seite komplett neu laden zu müssen (wie es im Web 1.0 der Fall war). Das eigentliche Novum besteht in der Tatsache, dass nur gewisse Teile einer HTML-Seite oder auch reine Nutzdaten sukzessiv bei Bedarf nachgeladen werden, womit AJAX eine Schlüsseltechnik zur Realisierung des Web 2.0 darstellt. Darüberhinaus gewinnt die Verlagerung von Desktopelementen im Web an Bedeutung. Mit GMail, Yahoo Widgets wird möglich dass man Kalender, Notizen, Aufgaben und sogar Dokumente in verschiedenen Formaten nicht mehr auf dem eigenen Computer erstellt und verwaltet werden, sondern im Web. Auf der anderen Seite bedeutet die soziale Entwicklung eine zunehmende Interaktion zwischen den Benutzern im Internet, indem ein Kollektiv von Nutzern zum Mehrwert beitragen (Wiki, Tagging). Der Mehrwert wird auch durch die Rekombination einzelner Elemente von verschiedenen Webseiten geschaffen. Freie Inhalte setzten sich durch, das Web wird eine globale Plattform in der sich die meisten Projekte in Beta-Stadium befinden.



Fig. 2. Ueberblick auf die Komponenten und Eigenschaften von Web 2.0

3 Web 3.0

Wie am Anfang des Papers erwähnt wurde, ist Web 3.0 als eine Ergänzung von Web 2.0 mit semantischen Elementen zu betrachten. Im Rahmen von Semantic Web werden Webseiten mit maschinenlesbaren Annotationen angereichert. Die Idee ist, dass durch derartige Erweiterungen, Abfragen in der natürlichen Sprache erlaubt werden. Zum Beispiel, anstatt dass man nach "Paris Europa" sucht, kann man eine Formulierung der Form "Paris liegt in Europa" verwenden. Das Semantic Web bietet Lösungsansätze für die Probleme, welche sich aus der Fülle der im Web präsenten Informationen ergeben, beispielsweise das Auffinden relevanter Informationen, die Integration von Informationen aus verschiedenen Quellen und das Problem des impliziten Wissens. Es werden Standards zu Repräsentation von Informationen eingeführt, die deren Austausch ermöglichen sollen. Methoden zu automatischen Schlussfolgerungen sollen implizites Wissen erschließen.

3.1 Technologien

Das Semantic Web basiert auf Technologien, die allgemein als semantische Technologien bekannt sind. Zum Ersten ist es nötig, einheitliche, offene Standards für die Beschreibung von Informationen zu vereinbaren, die es letztlich ermöglichen sollen, Informationen zwischen verschiedenen Anwendungen und Plattformen auszutauschen und zueinander in Beziehung zu setzen. Eine grundlegende Anforderung an diese Standards, neben deren formal klarer Definition, besteht in der Flexibilität und Erweiterbarkeit. Genau diesem Ziel der Schaffung solcher Standards für das Semantic Web hat sich das World Wide Web Consortium (W3C) ²verschrieben und bereits grundlegende Standards XML, RDF und OWL, Spezifikationssprachen ausführlich definiert. XML ist dabei nicht im engeren Sinne dem Semantic Web zuzuordnen, sondern eher dem klassischen Web. RTF's und OWL hingegen sind sogenannte Ontologiesprachen, die speziell für die Verwendung im Semantic Web entwickelt worden. XML ist zwar eine wichtige Basistechnologie für die Erstellung strukturierter Dokumente, insbesondere durch die einheitlich, standardisierte Art und Weise Informationen darzustellen. Zu dem bietet XML als Metasprache die Möglichkeit eigene Markupsprachen zu definieren. Allerdings sind XML Tags im Grunde nicht viel besser als die natürliche Sprache, zumindest aus Sicht des Semantic Web. Es fällt die Möglichkeit, die Bedeutung von Annotationen auf eine Art zu kodieren, die maschinenseitige Verarbeitung ermöglicht, bis zur automatischen Herleitung von nicht explizit gegebenem Wissen. Daher dient XML aus Sicht des Semantic Web primär als Grundlage, um die weiteren Sprachen RDF und OWL zu definieren.

² <http://www.w3.org/>

RDF Das Resource Description Framework (RDF) ist eine formale Sprache zur Bereitstellung von Metadaten im World Wide Web. RDF umfasst ein sogenanntes Tripel aus Subjekt, Prädikat und Objekt, das es ermöglicht, neben Attribut-Wert-Paaren auch Beziehungen zwischen Ressourcen darzustellen. RDF-Schema (RDF-S) ist ein XML-Standard, mit dem für eine Klasse von ähnlichen Ressourcen wie z. B. Webseiten, festgelegt werden kann, wie die Zuordnung von Eigenschaften erfolgen soll und welche grundlegenden Beziehungen zu anderen Eigenschaften bestehen.³

OWL Die Web Ontology Language (kurz OWL) ist eine Spezifikation des W3C, um Ontologien anhand einer formalen Beschreibungssprache zu erstellen, publizieren und verteilen zu können. Es geht darum, Terme einer Domäne und deren Beziehungen formal so zu beschreiben, dass auch Software (z.B. Agenten) die Bedeutung verarbeiten ("verstehen") kann (können). OWL ist somit ein wesentlicher Bestandteil der Semantic-Web-Initiative von Tim Berners-Lee. OWL basiert technisch auf der RDF-Syntax und historisch auf DAML+OIL, und geht dabei über die Ausdrucksmächtigkeit von RDF-Schema weit hinaus. Zusätzlich zu RDF und RDF-Schema werden weitere Sprachkonstrukte eingeführt, die es erlauben, Ausdrücke ähnlich der Prädikatenlogik zu formulieren.⁴

Ontologien haben in der Regel zwei Bestandteile:

Taxonomien Taxonomien definieren Objekte (z. B. Adresse) sowie Beziehungen zwischen Objekten. Ein Beispiel dazu: Das Objekt Postleitzahl gehört zum Objekt Stadt. So lassen sich sehr viele Beziehungen darstellen und sogar ineinander verschachteln: Postleitzahl gehört zu Stadt. Städte haben eigene Homepages. Den entsprechenden Link zu der Homepage können wir also entsprechend festhalten und bei Bedarf auswerten. Wir haben also eine wichtige Information erhalten, obwohl die ursprüngliche Datenbank das Objekt Postleitzahl nicht mit der Homepage der Stadt verknüpft hatte.

Inferenz-Regeln Richtig spannend wird die Sache aber erst durch sogenannte Inferenz-Regeln. Wieder ein Beispiel: Wenn eine "Postleitzahl" mit einem bestimmten Land assoziiert ist, und eine Adresse diese Postleitzahl benutzt, dann muss die Adresse auch mit dem Land assoziiert sein. Mit dieser Information könnte ein Programm eine Adresse automatisch so formatieren, wie es das entsprechende Land verlangt.

3.2 Vorteile und Nachteile von Web 3.0

Vorteile Die bisher vorgestellten Konzepte zum Semantic Web haben gemeinsam, dass sie Wissen über inhaltliche Zusammenhänge in Form von Schlagwörtern repräsentieren, die durch bestimmte Formalismen verknüpft sind. Das Semantic Web tritt aber nicht zuletzt an, um eine weitere Ausbaustufe des Internets zu werden. Das Internet setzt sich allerdings nicht aus Schlagwortnetzen zusammen, sondern aus einer überwältigenden Menge von Webseiten, Texten, Bildern, Videos, Dokumenten - kurz Ressourcen, die jeweils über eine eindeutige Adresse, die URL auffindbar sind. Eine bessere Ordnung in dieses Ressourcenuniversum zu bringen ist eine zentrale Hoffnung, die mit Semantic-Web-Technologien verbunden ist. Eine der Ziele von Semantic Web ist die Wiederverwendung von Wissen über sprachliche Barrieren hinweg, die wiederum eine aggregierte Suche über mehrere Seiten erlauben sollte. Durch den Einsatz von Techniken der künstlichen Intelligenz soll die Qualität der Suche verbessert werden, indem die Fehler und Widersprüche identifiziert und eliminiert werden.

Weitere Vorteile des semantischen Webs sind:⁵

* digitale Signaturen: Die Herkunft einer Information lässt sich eindeutig nachweisen. Entsprechende Programme könnten also die Glaubwürdigkeit einer Information abschätzen. So kann das Vertrauen der Nutzer in die gewonnenen Informationen gestärkt werden. Ebenfalls ist denkbar, dass, gestützt auf diese

³ <http://www.e-teaching.org/technik/vernetzung/semanticweb/>

⁴ www.wikipedia.org

⁵ http://www.net-wiki.de/index.php?title=Semantic_Web

glaubwürdigen Daten, die Zusammenarbeit im Unternehmen und von Unternehmen durch das Semantic Web ausgeweitet werden kann.

* Verkettung von Diensten: Einzelne Dienste (z.B. bestimmte Suchmaschinen oder spezialisierte Datenbanken) könnten ihre Serviceleistungen über entsprechende semantische Strukturen anpreisen, so dass ein Computer sie auslesen und ihren Nutzen abschätzen kann. Anschließend könnte er für eine beliebige Anfrage eine passende Kette zusammenstellen, in der er eine grosse Anzahl von Informationen durch verschiedene Dienste schickt und dem Nutzer anschliessend nicht eine unüberschaubare Vielzahl von Links gibt, sondern nur die Informationen nennt, die wirklich alle Filter durchlaufen haben. Die gezielte Auffindung von Informationen wird also stark verbessert.

* Automation: Wenn alle Geräte über entsprechende Datenbanken verkettet sind, lassen sie sich einfach automatisieren. Hierzu ein Beispiel: Ein Nutzer hört laute Musik, als das Telefon klingelt. Praktisch wäre es, wenn sich die Musik automatisch leise drehen würde, sobald er den Hörer abhebt, und wieder lauter wird, wenn das Gespräch beendet ist. Mit semantischen Strukturen ist so etwas nicht nur möglich, sondern auch einfach: Ein Nutzer muss jetzt nämlich nicht mehr seinen Computer so einstellen, dass er leiser wird, wenn das Telefon klingelt; und dann seinen Fernseher so einstellen, dass er leiser wird, wenn das Telefon klingelt; und dann seinen DVD-Spieler so einstellen, dass... sondern alle Geräte, die über die Information Lautstärke kann geregelt werden verfügen, werden automatisch leiser, wenn das Telefon klingelt.

Nachteile Obwohl das Semantic Web bereits seit vielen Jahren propagiert wird, ist bisher nur in einzelnen Vorzeige-Projekten eine Realisierung zu erkennen. Die vorgestellten Semantic-Web-Standards sind zwar technisch gesehen Standards, aber noch weit davon entfernt, durch verbreiteten Gebrauch faktisch akzeptierte Standards darzustellen. In der Praxis ergeben sich erhebliche Probleme. Sie reichen von der Frage, wie detailliert man semantische Relationen modellieren soll, will oder kann, bis hin zur politischen Überlegung, ob es Autoritäten geben soll, die zentrale Bedeutungen festlegen.

Zunächst bedeutet die zusätzliche Anreicherung von Daten mit maschinenlesbaren Informationen auch zusätzliche Arbeit. Redakteure müssen neben der Erstellung des Artikels auch noch für die korrekte Verschlagwortung sorgen.

Die Konzepte des Semantic Web sind hochkomplexe Formalismen. Das bleiben sie selbst dann, wenn es gelänge, eine intuitive Anwendung zu entwickeln. In erster Linie bleibt der Aufbau semantischer Netze damit Experten vorbehalten. Stärker noch als Enzyklopädien definieren semantische Netze, was die Welt ist, was wahr ist und was falsch. Damit kommt einer sehr kleinen Gruppe von Menschen eine ausgesprochen grosse Definitionsmacht zu.

Eine alternative Herangehensweise stellen Social-Tagging-Systeme dar, die genutzt werden können, um aus einer grossen Menge von individuell vergebenen Schlagwörtern dezentrale Ontologien (Folksonomies) aufzubauen. Daraus ergibt sich allerdings das umgekehrte Problem: Was einen Tag ausmacht wird über den Mittelwert der Masse aller, die Objekte mit diesem Tag versehen haben, definiert. Doch gibt es reichlich Beispiele für populäre Irrtümer.

Die Popularität von Tagging in diversen Web 2.0 Anwendungen macht dennoch Hoffnung, dass zwei Grundannahmen des Semantic Web richtig sind: Zum Einen geben sich viele Nutzer Mühe, Daten mit semantisch korrekten Metadaten anzureichern. Zum Anderen scheinen sich Tags als sinnvolle und verständliche Suchmöglichkeit neben der Volltextsuche etabliert zu haben.

4 GroupMe

In der Darstellung der Probleme von Semantic Web und deren Lösungsansätze wurden die Tagging Systeme erwähnt. An dieser Stelle wird eine Anwendung präsentiert, die eine neue Art vom sozialen

Tagging System darstellt. Die Anwendung heisst GroupMe und verbindet neben Web 2.0 Ansätze von Semantic Web Technologien.

Die neuartige Web 2.0"-Anwendung GroupMe!" fasst multimediale Webinhalte in Themengruppen zusammen. Durch seine innovative Benutzeroberfläche und den Einsatz von Semantic-Web-Technologien bietet GroupMe!" neue Möglichkeiten, Informationen zu erschliessen. Die neue Gruppenstruktur und semantische Annotationen verbessern die zielgerichtete Suche nach Inhalten.

4.1 Funktionen von GroupMe

Die Anwendung basiert auf einer einfachen Idee, die den Benutzer erlaubt Web Ressourcen zu gruppieren. Also eine der wichtigsten Funktionen ist die Erstellung von Gruppen von Multimedia Web Ressourcen mit Drag and Drop. Den Benutzern ist eine gewisse Freiheit überlassen die Gruppen nach eigenen Wünschen zu formatieren und zu strukturieren, basierend auf dem Konzept von Mind Maps. Der Sinn der Sache ist, dass sehr viele relevante Ressourcen in Gruppen zusammengefasst werden. Die Gruppen stellen verschiedene Themenbereiche dar. Innerhalb der Anwendung ist den Benutzern die Suche von Ressourcen von anderen Benutzern selbstverständlich gestattet. Eine eigene API steht den Benutzern zur Verfügung und bietet die Möglichkeit die Anwendung mit weiteren Funktionen zu erweitern.

4.2 Benutzerschnittstelle von GroupMe

Ein Überblick über die Möglichkeiten und Funktionen von GroupMe wird im Bild (Fig.3) dargestellt.⁶

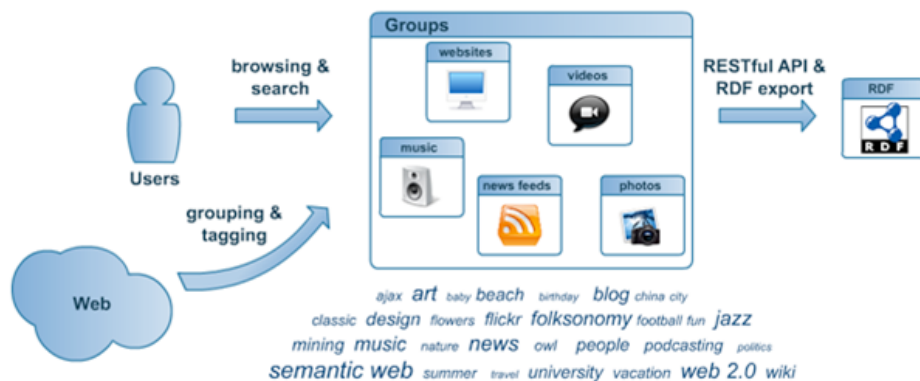


Fig. 3. Ueberblick auf die Eigenschaften von GroupMe

Eine GroupMe!"-Gruppe ist eine thematische Sammlung von Links, die auf relevante Webseiten, Bilder und Videos verweisen.

Ein wichtiger Unterschied zu gewöhnlichen Linksammlungen ist die graphische Darstellung der Links und Gruppen. Links werden dem Medientyp entsprechend visualisiert: Bilder und Videos können direkt in der Gruppe betrachtet werden, bei News-Feeds, wie Informationen aus Blogs, werden aktuelle Schlagzeilen angezeigt, und Musikdateien können direkt abgespielt werden. Durch diese Form der Visualisierung bekommt der Betrachter einer Gruppe einen sofortigen Überblick über das Thema und seinen Inhalt. Mit dem GroupMe!-System können Web-Inhalte somit effizienter erfasst und verwaltet werden als mit traditionellen Wissensmanagement-Systemen.

⁶ <http://groupme.org/GroupMe/>

5 Zukunftsvision und Debatte

Eine Reihe von Wissenschaftlern, die sich mit dem Thema auseinandergesetzt haben, haben die Welt von Semantic Web und des Web 3.0 nach eigenen Vorstellungen in verschiedenen Publikationen präsentiert.

John Markoff

John Markoff, dessen Artikel⁷ am 12. November 2006 in der New York Times über Web 3.0 veröffentlicht wurde, beschrieb Web 3.0 als Web 2.0 erweitert durch die künstliche Intelligenz. Für Markoff ist das Web 3.0 eine Wissensdatenbank des Alltagswissens, die weiterentwickelt wird, um Anwendungen der künstlichen Intelligenz zu ermöglichen. Dabei werden alle Inhalte als logische Aussagen formuliert, die auf der Prädikatenlogik aufbaut.

Nova Spivack und Tobias Kollmann

Spivack und Kollmann zählen eine Reihe von Begriffen, die im engen Zusammenhang von Web 3.0 stehen:⁸ Semantik Web, Mikroformate, Suche nach natürlichen Sprachen, Daten gewinnend, lernfähige Maschinen, neue Technologien für Benutzerfreundlichkeit. Wichtig dabei ist die Tatsache, dass die Kundenwünsche im Mittelpunkt stehen. Man kann in deren Vorstellung die stark geprägte wirtschaftliche Seite von Web 3.0 erkennen. Die Autoren geben als Beispiel die Webseite Askerus, die Urlaubsdienste anbietet. Auf der Webseite der Firma wird dem Benutzer anhand einer Reihe von Schritten nach seinen Wünschen gefragt und dementsprechend werden ihm Urlaubsangebote geliefert. Das ganze hat auch einige Nachteile, wie die Pflichtregistrierung auf der Webseite und meistens nach einer mühsamen und aufwendigen Befragung bekommt der Benutzer doch nicht die erwünschten Informationen.

Karl Kamino

Eine Welt in der man die Algorithmen durch Redakteure ersetzt, wird von Karl Kamino präsentiert. In dieser Welt werden die Anfragen individuell bearbeitet. Hierfür bringt er das Beispiel von der Firma Mahalo⁹. Die Mitarbeiter der Firma haben verschiedene Themenbereiche zusammengefasst, sehr gut strukturiert, so dass die Benutzer können innerhalb dieser Bereiche eine Suche sehr effizient durchführen. Der Nachteil ist allerdings, dass viele Themenbereiche noch nicht erfasst wurden und deswegen wird man bei einer tieferen Suche zu Google weitergeleitet.

Robert Basic

Eine interessante Vorstellung stammt von Robert Basic, der sich das zukünftige Web als ein virtuelles Web vorstellt. Alle Seiten sind 3D und simulieren unsere reale Welt, also alles wird wie im Second Life. Alle sollen sich in virtuellen Spielen aufhalten und dort miteinander kommunizieren.

Sramana Mitra

Sramana Mitra¹⁰ versucht das Konzept Web 3.0 in einem mathematischen Konzept darzustellen.

$$\text{Web 3.0} = (4C + P + VS)$$

□ 3C = Content, Commerce and Community

⁷ <http://www.nytimes.com/2006/11/12/business/12web.html?ex=1320987600&en=254d697964cedc62&ei=5088>

⁸ Spivack Kollmann, Article: What is the Semantic Web, Actually?, New York Times, November 12, 2006

⁹ www.mahalo.com

¹⁰ <http://www.sramanamitra.com/>

☐ 4te C = Context

☐ P = Personalisierung

☐ VS = Vertikale Suche

Web 3.0 erhält also gegenüber Web 2.0 noch den Kontext dazu.

Andere Meinungen

In den heftigen Debatten bezogen auf Semantic Web kann man von pesimistischen bis optimistischen Meinungen. Also viele sehen das Konzept von Web 3.0 als einen Marketing Trick zum Hochpuschen von Semantic Web, eine graue Welt in der die Benutzer kommerziell ausgenutzt werden. Auf der anderen Seite sehen viele das Web 3.0 als eine neue Generation von Business Anwendungen. Über alle diese Meinungen hinweg macht sich die Vorstellung bemerkbar, dass Web 3.0 die Internetwelt komplett neu organisiert und das die aktuellen Suchmaschinen und die klassischen Technologien überflüssig werden.

6 Häufig gestellte Fragen

Viele Fragen rund um Semantic Web drehen sich um die begrenzte Möglichkeit an menschlichen Ressourcen von Google, die Wissensbereiche zusammenzufassen und der Menschheit frei zur Verfügung zu stellen.

”Google and other companies do not have the resources in man power (i.e. the thousands of volunteers Wikipedia has) who would help create those ontologies for the large set of knowledge domains that Wikipedia covers. Wikipedia does, and is positioned to do that better and more effectively than anyone else. Its hard to see how Google would be able create the ontologies for all domains of human knowledge (which are continuously growing in size and number) given how much work that would require. Wikipedia can cover more ground faster with their massive, dedicated force of knowledgeable volunteers.”¹¹

6.1 Qualität der Information

Der Aufbau und die Stellung der Onthologien werden den einzelnen Benutzern überlassen und basiert auf der Idee von free community. Natürlich stellt sich die Frage, ob alle die von den Freiwilligen erstellten Informationen erstens korrekt und zweitens relevant sind. Die Frage lautet also: Wie unterscheide ich zwischen guter und schlechter Informationen? Die allgemein akzeptierte Antwort ist, dass es keine Notwendigkeit besteht die Information zu selektieren, falls die Information aus einer vertrauenswürdigen Onthologie stammt. Von Wikipedia 3.0 abgeleitete Information ist als vertrauenswürdig anzusehen.

6.2 Relevanz der Information

In der Zusammenfassung von Wissensbereichen werden Informationen unterschiedlich gewichtet und dementsprechend werden die Abfragen unterschiedlich beantwortet. Die Frage die jetzt auftritt ist: Wer oder was entscheidet welche Informationen für mich als Endbenutzer relevant sind? Auf der Suche nach einer Antwort für diese Frage fand ich im Mittelpunkt einen interessanten Ansatz. Für die Antwort auf die Benutzerfragen werden Annahmen über den Sinn der Abfrage getroffen. Man kann sich den Ansatz wie eine Art Data Mining Klassifikationsverfahren vorstellen, basierend auf Wahrscheinlichkeiten wie der bekannte Algorithmus ”Naive Bayes Classifier”. Der Algorithmus würde in dem Fall verschiedenen Gewichtungen einen Begriff zuweisen je nach dem mit welchen anderen Begriffen er in der Abfrage in Verbindung steht.

¹¹ <http://evolvingtrends.wordpress.com/2006/06/26/wikipedia-30-the-end-of-google/>

6.3 Web 3.0 als ultimativer Standard

Im wirklichen Sinne des Wortes bringt Web 3.0 viele neue Standardisierungen mit. Werden diese Standardisierungen aber den Benutzern gedrängt? oder wird der Benutzer verpflichtet sich an verschiedene Standards zu halten? Viele behaupten, dass die neuen Standards ein Grund sind warum Semantic Web und die neuen Technologien sich schleppend imponieren. Die Tatsache ist aber, dass für die Benutzer überhaupt keine Aneignung der Standards notwendig sind. Nur diejenigen, die sich an die Erstellung der Ontologien beteiligen, müssen sich an die Standards halten. Zu den Standards gehören nicht nur Regeln für den Aufbau der Ontologien, sondern auch die zugehörigen Erstellungstools. Es gibt bereits auf dem Markt Anwendungen für die Erstellung von Ontologien, die diese Standards beinhalten (Protege).

7 Mikroformate - eine Alternative?

Mikroformate sind einfache semantische Erweiterungen für HTML, die verteilten Systemen ermöglicht strukturiert auf die vorhandenen Daten zuzugreifen. Anders gesagt sind die Mikroformate simple Standardisierungen wie bestimmte Anwendungsfälle in HTML ausgezeichnet werden sollen um diese maschinenlesbar zu machen.

Das folgende Beispiel zeigt die Verwendung von Mikroformaten:

```
<div class="vcard">  
  <a class="url fn" href="http://blog.no-panic.at/">Florian Beer</a>  
</div>
```

Fig. 4. Mikroformate

In der reinen Darstellung auf der Webseite, würde sich durch diese Notierung als vCard nichts ändern, jedoch kann man sehr einfach einen Parser schreiben der die Konventionen versteht und daraus ableiten kann, dass der Autor der Webseite (`http://blog.no-panic.at` den `FullName(fn Florian Beer` hat. Die Entwicklung dieser Formate wurde als offener Prozess gestaltet. Es wurde auf der Projektseite `http://microformats.org` ein Wiki, eine Sammlung von Webseiten, die von jedem Benutzer angesehen und editiert werden kann.

8 Schlusswort

Über viele Jahre hinweg wurde das Internet als Technologie erlebt, die es erlaubt, Daten, Informationen oder Medien zu publizieren und zu verteilen. Die Rollenverteilung der beteiligten Personen war zweiteilig: Zum einen gab es aktive Ersteller von Webinhalten, die - teils kommerziell, teils privat - Informationen einstellten und publizierten. Zum anderen gab es passive Konsumenten, die sich lediglich die bereitgestellten Inhalte ansahen und auch gar keine andere Option hatten, als die Informationen zu empfangen und zu konsumieren.

Mit der Geburt von Web 2.0 wurde die Internetwelt revolutioniert. Der Communitygedanke steht ganz klar im Vordergrund und bildet nicht zuletzt die Basis vieler neuer Onlinegeschäftsideen, an die vor einigen Jahren noch nicht zu denken war. Aber das Web 2.0 bietet noch mehr: Mit Techniken wie Weblogs können Unternehmen jeder Art in einen Dialog mit ihren Kunden treten und ihren Internetauftritt bereichern - auch wenn die eigene Kasse nur wenig Geld für Marketing und Webangebot bereithält.

Das Web und die damit in Verbindung stehende Geschäftswelt hat mittlerweile eine neue Stufe der Entwicklung erreicht. Eine neue Revolution in der Informationswissenschaft ist zu erwarten. Die Frage ist ob das Semantic Web, das von vielen als "Totgeburt" bezeichnet wird, erstens überleben wird, und zweitens, ob es sich als ein reifes Konzept und stabile Technologie etablieren wird. Tatsache ist, dass die Mehrheit der Webdesigner sich nie mit XML als Sprache des semantischen Webs anfreunden konnte, ganz zu schweigen von RDF, dessen Sinn und Zweck vielen Entwicklern ein Rätsel blieb.

Im Rahmen eines Forschungsprojekts beim Deutschen Forschungszentrum für künstliche Intelligenz wird versucht das Semantic Web über die Hintertür auf die Rechner der Anwender und der Entwickler zu bekommen. Leo Sauermann, Semantic Web Forscher erklärt in einem Interview für das "XML und Web Services Magazin" die Grundidee seiner Forschung:

"Die Grundidee ist es, Menschen zu ermöglichen ihre Dateien nach ihren eigenen Vorstellungen zu organisieren, und da diese Vorstellungen in den Köpfen der Nutzer meist auf der Ebene von Projekten, Terminen, Personen, Themen oder dem Zeitlichen geordnet sind, sollten die Personal Computer genauso funktionieren. Im Semantic Desktop geht es nun darum, die Technologien des Semantic Web zu verwenden, um diese semantischen Strukturen in den Köpfen der Benutzer auch am Desktop zu etablieren. Ziel ist ein semantisches Netz, das alle Dateien, Kontakte, Termine und Projekte miteinander verbindet. Das gilt für persönliche Daten genauso wie Informationen aus Groupware und anderen kollaborativen Systemen. Der Social Semantic Desktop soll die semantische Zusammenarbeit zwischen Personen in diesen Umgebungen ermöglichen." ¹²

¹² <http://xml-magazin.de/itr/news/psecom,id,36225,nodeid,68.html>

References

1. www.mahalo.com
2. <http://www.sramanamitra.com/>
3. <http://www.w3.org/>
4. <http://protege.stanford.edu/>
5. Semantic Web; Pascal Hitzler, Markus Krötzsch u.a. Springer Verlag; 2007
6. Entrepreneurs See a Web Guided by Common Sense, John Markoff, The New York Times, November 12, 2006
7. Diplomarbeit: Das semantische Web und wie die Microformats die Entwicklung vorantreiben werden, Florian Beer, Eisenstadt, 2007
8. <http://groupme.org/GroupMe/>
9. <http://microformats.org>
10. Semantic web services; Studer Rudi, Stephan Grimm; Springer Verlag Berlin; 2007
11. Semantic Web (Concepts, Technologies and Applications); Karin Breitman, Marco Antonio Casanova; Springer Verlag London; 2007
12. Berners-Lee, Tim; Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web; HarperBusiness; 2000
13. <http://evolvingtrends.wordpress.com/2006/06/26/wikipedia-30-the-end-of-google/>

Table of Contents

Web 3.0: das Ende von Google	1
<i>Eduard Schibrowski</i>	
1 Motivation	1
2 Auf dem Weg zu Web 3.0	1
2.1 Web 1.0	1
2.2 Web 2.0	2
3 Web 3.0	3
3.1 Technologien	3
RDF	4
OWL	4
3.2 Vorteile und Nachteile von Web 3.0	4
Vorteile	4
Nachteile	5
4 GroupMe	5
4.1 Funktionen von GroupMe	6
4.2 Benutzerschnittstelle von GroupMe	6
5 Zukunftsvision und Debatte	7
6 Häufig gestellte Fragen	8
6.1 Qualität der Information	8
6.2 Relevanz der Information	8
6.3 Web 3.0 als ultimativer Standard	9
7 Mikroformate - eine Alternative?	9
8 Schlusswort	11