



Prof. Dr. Dr. h.c. mult. Wolfgang Wahlster
CEO des Deutschen Forschungszentrums für Künstliche Intelligenz (DFKI GmbH)



Web 3.0 = Semantisches Web + Web 2.0

Durch die Konvergenz der aktuellen Informatik-Trends „Semantisches Web“ und „Web 2.0“ ergibt sich die Chance, als Web 3.0 eine innovative Wissensinfrastruktur zu schaffen, die völlig neuartige Formen des Wissenszugriffs, des Wissensmanagements und der webbasierten Dienste, Anwendungen und Geschäftsmodelle ermöglicht.

Die Konvergenz von Semantischem Web und Web 2.0

Das klassische Web 1.0 hat den weltweiten Zugang zu digital gespeicherter Information drastisch verbessert. Aber dort sind die Inhalte nur maschinenlesbar, ohne maschinell verstehbar zu sein. Da Wissen meist in natürlicher Sprache präsentiert wird, sind die gefundenen Dokumente bislang nur für den Menschen voll verständlich. Das semantische Web basiert auf der inhaltlichen Beschreibung digitaler Dokumente mit standardisierten Vokabularen, die eine maschinell verstehbare Semantik haben. Damit wird der Übergang von einem syntaktischen „Netz aus Verweisstrukturen“ zu einem semantischen „Netz aus Inhaltsstrukturen“ vollzogen.

Den Kern der semantischen Webtechnologie bilden Markierungssprachen, die eine formale Syntax und Semantik haben und in Form einer Ontologie eine standardisierte Begrifflichkeit zur Beschreibung digitaler Inhalte bereitstellen. Semantische Markierungssprachen wie OWL erlauben die weltweite Verbreitung und gemeinsame Nutzung von Ontologien. Bei der Suche nach Informationen können Schlussfolgerungen den Suchprozess beschleunigen oder gesuchte Information ableiten. Methoden des maschinellen Lernens, des Data Mining und des Text, Speech und Video Mining reduzieren beim Aufbau und bei der Wartung von On-

Pressekonferenz am 4. Oktober 2006 um 12:30 Uhr
„Mensch - Semantik - Business: Wohin steuert das Web?“
TU Dresden, Rektorat, Festsaal, Mommsenstraße 11, Dresden



tologien den manuellen Aufwand. Software-Werkzeuge unterstützen das Erstellen, Editieren, Visualisieren, Auswerten, Verschmelzen und Validieren von Ontologien.

Das Web 2.0 ist noch ein primär syntaktisches Web, das aber durch die private Annotation von multimedialen Dokumenten und deren Bereitstellen für die Öffentlichkeit eine große soziale Dynamik entfaltet. Dadurch, dass die klassische Unterscheidung zwischen Autoren, Editoren und Konsumenten im Web 2.0 verschwindet und jeder Webnutzer abwechselnd in allen diesen drei Rollen auftreten kann, entsteht als soziales Massenphänomen mit rasanter Geschwindigkeit eine dynamische Wissensinfrastruktur, die stark von Privatnutzern in der Freizeit aufgebaut wird. Allerdings fehlt den ad-hoc Markierungen im Web 2.0 die Systematik und Präzision der Ontologien im semantischen Web, auf der letztlich auch ein maschinelles Verstehen basieren muss. Der Aufbau des semantischen Webs verlief bisher aufgrund der hohen Investitionskosten zur Erstellung semantisch markierter digitaler Inhalte schleppend. Aktuell sind weniger als 1% der weltweit im Internet verfügbaren Inhalte im Sinne des semantischen Webs aufbereitet. Wenn es gelingt, die Ontologien des semantischen Webs in die Massenbewegung des Web 2.0 einzubringen, so dass die privaten Endnutzer in einfacher und effizienter Weise bei der Erstellung von Wikis, Blogs und Mashups auf wohlfundierte Ontologiewerkzeuge zurückgreifen können, dann kann in der nächsten Dekade ein Web 3.0 entstehen, das die Vorteile des semantischen Web und des Web 2.0 verknüpft und deren Nachteile überwindet. Der top-down Ansatz zur Organisation von Wissen im semantischen Web wird durch den bottom-up Ansatz der Web 2.0 Bewegung komplementär ergänzt.

Von SMARTWEB zu QUAERO

Das vom BMBF geförderte und vom DFKI geleitete SmartWeb-Konsortium demonstriert den intuitiven Zugang zum Semantischen Web durch multimodalen Dialog über ein mobiles Endgerät (MDA oder Smartphone). Der Benutzer kann mit **SmartWeb** seine Fragen frei in Umgangssprache formuliert in das mobile Gerät einsprechen und mit Zeigegesten auf dem Bildschirm unterstützen. Beispielsweise wird die Frage „Wie wird das Wetter morgen?“ ortssensitiv durch Zugriff auf einen der Wetterprognosedienste im Internet beantwortet. Die Faktenfrage „Wie hoch ist der Mount Fuji?“ wird nicht durch eine Vielzahl von Dokumenten beantwortet, in denen der Benutzer die Höhe selbst heraussuchen muss, sondern prägnant mit der gewünschten Angabe „3776 m“. SmartWeb bietet nicht



nur dem mobilen Nutzer jederzeit und überall einen multimodalen Webzugriff, sondern stellt selbst Auto- und Motorradfahrern komponierte Webdienste und Medien aller Art bereit. Die Weiterentwicklung der situationsadaptiven Dialogkomponente zum offenen und intuitiven Zugang zur digitalen Wissensinfrastruktur ist auch eine Basistechnologie für das geplante QUAERO-Vorhaben.

Das deutsch-französische Leuchtturmprojekt **QUAERO** wurde als strategisches Technologieprojekt konzipiert, mit dem die in Deutschland und Frankreich bestehenden Standortstärken im Bereich innovativer Technologien beim Informations- und Wissenszugang für das Internet der Zukunft gebündelt und ausgebaut werden sollen. QUAERO ist ein Teil der Hightech-Strategie der Bundesregierung, deren Umsetzung in sechs jährlichen Treffen der Forschungsunion Wirtschaft-Wissenschaft begleitet und weiterentwickelt wird. Mit QUAERO soll die Wettbewerbsfähigkeit Deutschlands und Europas beim Zugang und der Nutzung von digital verfügbarem Wissen als wichtigste Ressource des 21. Jahrhunderts und mit Blick auf den globalen Inhalte-Wettbewerb verbessert werden.

Das lateinische Wort QUAERO bedeutet nicht nur "ich suche", sondern kann auch mit "ich frage" oder "ich forsche" übersetzt werden. Dies ist insofern eine treffende Bezeichnung für das Leuchtturmprojekt als es um die Erforschung und Entwicklung von innovativen multimodalen Internet-Diensten geht, die statt Schlüsselwort-basierten Suchmaschinen dialogische Frage-Antwort-Systeme bereitstellen.

Aus Sicht des deutschen Industriekonsortiums ist QUAERO keine Allianz gegen Google oder andere etablierte Suchmaschinen für das bestehende Web 1.0, sondern es geht um die netzbasierte Wissensinfrastruktur der Zukunft, in der sich neue Märkte für digitale Güter und Dienste eröffnen.

In einem breit angelegten Prozess der spezifischen Ausgestaltung von QUAERO erhielten die führenden Know-how-Träger aus Wirtschaft und Wissenschaft in Deutschland die Chance, sich als Partner in das QUAERO-Konsortium einzubringen. Mit Abgabe der Anträge zum vom BMWi vorgegebenen Zeitpunkt 8. August 2006 ist dieser Prozess abgeschlossen worden. Ein schlagkräftiges deutsches Konsortium unter Federführung der Bertelsmann-Tochter Empolis entstand, das neben deutschen Großunternehmen wie Siemens und SAP, mittelständische Unternehmen, Verbände und öffentliche Einrichtungen umfasst. Als Forschungspartner haben sich neben dem DFKI zahlreiche Institute der Fraunhofer-Gesellschaft und verschiedene Universitätsinstitute an dem Antrag betei-



ligt, der am 18./19. September einer Evaluation durch externe Fachgutachter unterzogen wurde.

Der Mehrwert des Web 3.0

Die Diversität und die Dynamik von Wissen sowie die Heterogenität in der Wissensrepräsentation machen eine einzige weltweit standardisierte Universal-Ontologie als Grundlage für das Web 3.0 unmöglich. Vielmehr werden Peer-to-Peer-Mechanismen (P2P) wie heute schon auf der Netzwerkebene auch die Wissensebene im Web 3.0 wesentlich bestimmen. Die semantische Wissensintegration, das Übersetzung und der inhaltliche Abgleich zwischen verschiedenen Branchenontologien sowie eine inhaltlich-gesteuerte P2P-Suche basieren auf Webontologie-Sprachen wie OWL, für die eine fundierte modelltheoretische Semantik und effiziente Inferenzsysteme existieren.

Softwarewerkzeuge für die Erstellung, die Wartung, die Wiederverwendung, den Austausch und die Adaption von Ontologien stellen wichtige Basiskomponenten für das Web 3.0 dar. Branchen-Ontologien werden künftig die wesentlichen Begriffe, Beziehungen und Prozesse für eine bestimmte Branche definieren, so dass eine Interoperabilität von Webdiensten erreicht wird. Metadaten werden die kontextabhängige Transformation digitaler Inhalte für unterschiedliche Ausgabegeräte und Nutzergruppen unterstützen.

Der Gefahr der Informationsüberflutung kann durch das Web 3.0 begegnet werden. Information wird in Zukunft dem Nutzer bedarfs- und situationsgerecht sowie personalisiert zur Verfügung gestellt. Durch die Verwendung von Ontologien können Hyperlinks inhaltlich klassifiziert werden, so dass ein semantisch gesteuertes Surfen und ein Retrieval mit inhaltlichen Ähnlichkeitsmetriken möglich wird. Durch Zugriff auf semantische Webseiten können intelligente Suchagenten im Web 3.0 mit hoher Präzision gezielt Informationen im Internet auffinden, filtern und kombinieren, um den Benutzer zeitaufwändige Rechercharbeit abzunehmen.

<http://www.dfki.de/~wahlster/>