

# SEMANTISCHES WEB

Wolfgang Wahlster  
DFKI

**Das World Wide Web (WWW) hat den weltweiten Zugang zu digital gespeicherter Information drastisch verbessert. Aber**



**bislang waren die Inhalte im WWW nur maschinenlesbar, ohne maschinell verstehbar zu sein. Da Information im WWW 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 Vokabularien, die eine maschinell verstehbare Semantik haben. Damit wird der Übergang von einem „Netz aus Verweisstrukturen“ zu einem „Netz aus Inhaltsstrukturen“ vollzogen.**

Drei Ebenen von Markierungen

Inhalt	OWL/ DAML/OIL
Struktur	XML
Form	HTML
WWW-Dokument	

## Die Technologie

Den Kern der Technologie 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 im WWW. Die semantische Markierung durch Metadaten stellt neben der Markierung für das Layout (HTML) und die Struktur (XML) eines Dokumentens die dritte Ebene einer Dokumentbeschreibung dar. Bei der Suche nach Informationen können Schlussfolgerungen den Suchprozess

beschleunigen oder gesuchte Information ableiten. Methoden des maschinellen Lernens, des Data Mining und des Text Mining reduzieren beim Aufbau und bei der Wartung von Ontologien den manuellen Aufwand. Software-Werkzeuge unterstützen das Erstellen, Editieren, Visualisieren, Auswerten, Verschmelzen und Validieren von Ontologien.



Semantisches Surfen  
Bild: DFKI

## Problem und Lösung

Der Gefahr der Informationsüberflutung kann durch das semantische Web 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, sodass ein semantisch gesteuertes Surfen und ein Retrieval mit inhaltlichen Ähnlichkeitsmetriken möglich wird. Durch Zugriff auf semantische Webseiten können intelligente Suchagenten mit hoher Präzision gezielt Informationen im Internet auffinden, filtern und kombinieren, um den Benutzer zeitaufwändige Rechercharbeit abzunehmen.

## Die Relevanz

Informationstechnische Dienste werden durch eine ontologische Beschreibung ihrer Voraussetzungen, Funktionalitäten und Schnittstellen zu semantischen Webdiensten, die durch Agentensysteme automatisch auffindbar, ausführbar und konfigurierbar werden. Über eine gemeinsame Semantik, die mithilfe einer Ontologiesprache eingeführt wird, können heterogen gespeicherte Daten in einem gemeinsamen Datenmodell zusammengeführt, inhaltlich selektiert und für unterschiedliche Präsentationsformen aufbereitet werden.

## Die Prognose

Das semantische Web wird in der nächsten Dekade schrittweise das bestehende WWW ablösen und für den Benutzer stark vereinfachte Dienste im Bereich der Informationssuche, des E-Business, des E-Government, des E-Recruiting und des Wissensmanagements anbieten. Branchen-Ontologien werden künftig die wesentlichen Begriffe und Beziehungen für eine bestimmte Branche definieren, sodass eine Interoperabilität von Webdiensten erreicht wird. Metadaten werden die Transformation digitaler Inhalte für unterschiedliche Ausgabegeräte und Nutzergruppen unterstützen.

