

## Semantic Web

- Wieso Semantic Web?
- Überblick
- Wesentliche Technologien

## Ausgangslage: Wissen schwierig im Web zu ermitteln

- Aktuelles World Wide Web
- Auffinden von Wissen schwierig
- Vorhandene Technologien zur Suche: Suchmaschinen, im Kern: Information Retrieval
- Dienste: Google, Yahoo, AltaVista, MetaCrawler etc.
- Probleme:
  - Recall limitiert: Alles relevante gefunden?
  - Precision limitiert: Wieviele gefundene Dokumente sind relevant?
  - Relevanz fragwürdig: Nur auf Wort-Vektoren basierend

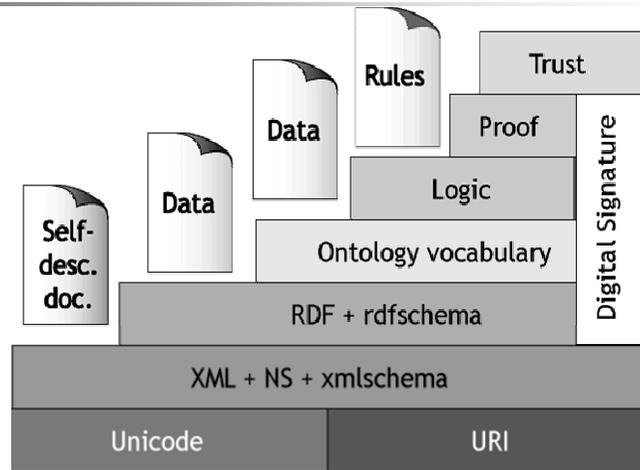
## Zielsetzung

- Bedeutung von Informationen maschinenverarbeitbar markieren
- Damit Unterstützung von
  - Entdeckung gehaltvollerer Informationen
  - Datenintegration
  - Bessere Navigation
  - Automatisierung von Aufgaben

## Aktueller Ansatz: Semantic Web

- Idee: Definition von standardisierten Technologien zur maschinenverarbeitbaren Beschreibung von Wissen in einem großen Netz
- Prinzipien
  - Jede Information kann mit URIs identifiziert werden
  - Ressourcen und Beziehungen sind semantisch typisiert
  - Partielle Informationen sind akzeptabel
  - Absolute Wahrheit wird nicht benötigt
  - Evolution als Entwicklungsprinzip
  - Minimalistisches Design der Technologien

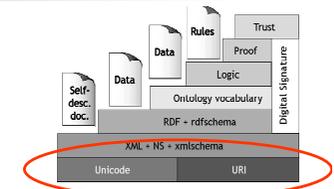
## Schichten des Semantic Web



Aus: Tim Berners-Lee: <http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html>

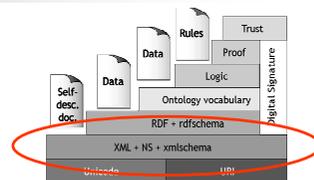
## Bestandteile des Semantic Web

- Unicode
  - Universelles Zeichenrepertoire
  - Umfasst alle Zeichen der Welt
  - Ist fortlaufend in Erweiterung
  - Als ISO 10646 normiert
- URI
  - Universeller Bezeichner
  - Klare Syntax
  - Bei Interpretation im Netz: URL



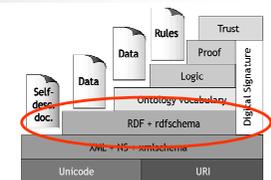
## Bestandteile des Semantic Web

- XML
  - Sprache zur Definition von Auszeichnungssprachen
- Namespaces
  - Konzept zur Verwendung mehrerer XML Sprachen in einem Dokument ohne Namenskonflikte
- XML Schema
  - Modellierungssprache, die XML um Datentypen, Strukturierungs- und Wiederverwendungsmöglichkeiten erweitert



## Bestandteile des Semantic Web

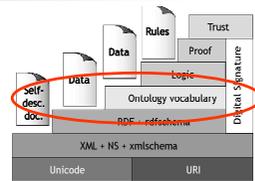
- RDF
  - Modell und Syntax zur Annotation von Metadaten
- RDF Schema
  - Sprache zur Definition von Metadatenformaten und einfachen Ontologien



## Bestandteile des Semantic Web

### Ontologien

- Vokabulare
- Begriffsbeziehungen (Unterkategorie, Untereigenschaft, Wertebereiche, ..., selbstdefinierte)
- Sprachen für Web-Ontologien: DAML+OIL, OWL



## Bestandteile des Semantic Web

### Logik

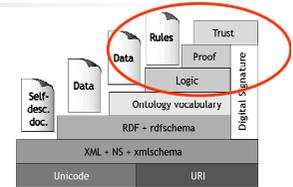
- Semantik auf logischer Basis
- Ableitungsregeln

### Proof

- Konsistenz
- Ableitung (Inferenz)

### Trust

- Am Beginn der Entwicklung



## Maschinelle Auswertung

- Metadaten: menschliche Nutzer und intelligente Anwendungen
- Ziele
  - Sammlung von dezentralen Webinhalten
  - Informationsverarbeitung
  - Ergebnisaustausch mit anderen Programmen  
=> gemeinsame Ontologien für Semantikbeschreibungen
- Framework
  - Konsistenzüberprüfung (Seiteninhalt, Metadaten, Ontologie(n))
  - Klassifikation (Informationsreduktion)
  - Anfragen (Inferenz-Service: logisches Schließen)

## Mögliche Anwendungen

### Semantische Suche:

- Anhand von Ontologien mittels Inferenz sinnvolle(re) Dokumente und Informationen finden

### B2C-Commerce: (auch B2B)

- Kunde: Produkt mit bestimmten Anforderungen (Gestaltung, Fähigkeiten, Leistung, Preis, ...) gesucht  
Was passt am besten?  
=> U.U. interaktive Anpassung der Spezifikation

- ...

## Zielvision

Tim Berners-Lee:

- Automatische Koordination von Terminen mit Suche von geeigneten Dienstleistern:
  - Geschwisterpaar will Vater zum Arzt und zur Behandlung begleiten
  - Wo sind geeignete Ärzte, KG-Praxen? (in der Nähe)
  - Wann sind Termine passend frei? (Präferenzen, Wegezeit)
  - Müssen andere Termine dafür umgelegt werden? (Priorität)

## Zustand heute

- Das Semantic Web existiert nicht (Aussage vom 18.2.2003)
- Es ist allgemein akzeptiert, dass es existieren sollte
- Wege zu einem Semantic Web
  - Entdeckung der "Killer Applikation"
  - Bottom-Up Entwicklung getrieben von Anwendern
  - Von Firmennetzwerken bzw. speziellen Anwendungsdomänen zu allgemeiner Verfügbarkeit
- Viele Methoden / Technologien / Werkzeuge existieren schon (KI)  
=> für das Semantic Web verfügbar machen / integrieren

## Herausforderungen

- Wissen aufbereiten
  - Entwicklungsumgebung
  - Wissensmanagement
  - Organisatorische Einbindung
  - Retrospektive Erschließung von Wissen
- Wissen nutzen
  - Inferenzen
  - Skalierbare Tools
  - Endnutzer