

# Seminar „DB-Aspekte des E-Commerce“

## Schwerpunkt: Anwendungen

Das Semantic Web – Ontologien, RDF, DAML+OIL

Boris Stumm



# Übersicht

- Einführung in das Semantic Web
- Grundlegende Begriffe wie URI, Metadaten, Schema, Ontologie
- Datenbeschreibungssprachen für das Semantic Web
  - RDF
  - RDF-Schema
  - DAML+OIL
- Beweise und Ziehen von Schlüssen
- Vertrauen und Konsistenz im Semantic Web
- Anwendungen
- Ausblick

# Einführung in das Semantic Web

- Das Internet heute
  - Benutzt von Menschen
  - keine Semantik und somit keine Nutzungsmöglichkeit für Maschinen
  
- Das Semantic Web von morgen
  - Viele Ontologien für verschiedene Anwendungsbereiche
  - Daten werden maschinenverständlich aufbereitet
  - Computer können das Internet „verstehen“
  - Weitgehende Automatisierung verschiedenster Tätigkeiten durch Einsatz von Agenten
  
- Der Weg zum Semantic Web
  - Verstehen von Informationen entweder durch das Verstehen natürlicher Sprache oder durch die Aufbereitung von Informationen in ein maschinenverständliches Format.

# URIs und Metadaten

- Unified Resource Identifier (URI)
  - Bezeichnet Ressourcen eindeutig
  - Jedermann kann URIs definieren → Mehrere URIs für dieselbe Ressource
  - Beispiele:  
`http://www.rhrk.uni-kl.de/~bstumm/`  
`mailto:stumm@informatik.uni-kl.de`
- Metadaten
  - Metadaten sind „Daten über Daten“
  - Anwendungen arbeiten oft auf Metadaten
  - Unterscheidung zwischen Metadaten und Daten ist nicht immer eindeutig

# Schemas und Ontologien

- Schemas zur Beschreibung von Ressourcen
  - Beispiel Dublin Core Element Set
- Ontologien zur Beschreibung der Zusammenhänge eines Anwendungsbereichs

*„Eine Ontologie stellt eine formale Beschreibung der Gegenstände und Beziehungen dar, die für eine Person oder Gruppe von Personen begriffsbildend sind“*
- Problem der Abbildung der Begriffe einer Ontologie auf eine andere

# Resource Description Framework (RDF)

- Modell zur Repräsentation von Metadaten
- W3C Recommendation
- Das RDF-Datenmodell hat drei wesentliche Objekttypen:
  - *Ressourcen*, repräsentiert durch ihre URI
  - *Eigenschaften* von Ressourcen (ebenfalls Ressourcen)
  - *Aussagen* der Form  
„Subjekt Prädikat Objekt“
- Vergleichbar mit OO-Technologien als auch mit ER-Diagrammen

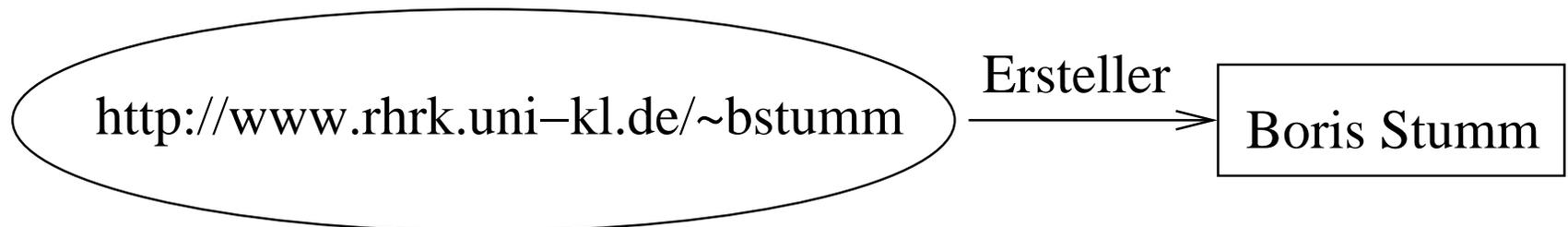
- Beispiel einer RDF-Aussage

Die Ressource „<http://www.rhrk.uni-kl.de/~bstumm/>“  
hat den Ersteller Boris Stumm

Subjekt: <http://www.rhrk.uni-kl.de/~bstumm/>

Prädikat: Ersteller

Objekt: „Boris Stumm“



## Definition von Ressourcen in anderen Ressourcen

- Oft werden mehrere Ressourcen in einem Dokument definiert
- URI des Dokuments:  
`http://www.rhrk.uni-kl.de/~bstumm/index.html`
- URIs der definierten Ressourcen:  
`http://www.rhrk.uni-kl.de/~bstumm/index.html#lebenslauf`
- Abkürzung von Ressourcenbezeichnern durch Ersetzung der Dokument-URI mit einem Kürzel:  
`bst:lebenslauf`
- Gebräuchliche Kürzel: `rdf:`, `rdfs:`, `daml:`, `dc:`

# RDF Serialisierung

- Notation 3 (N3)

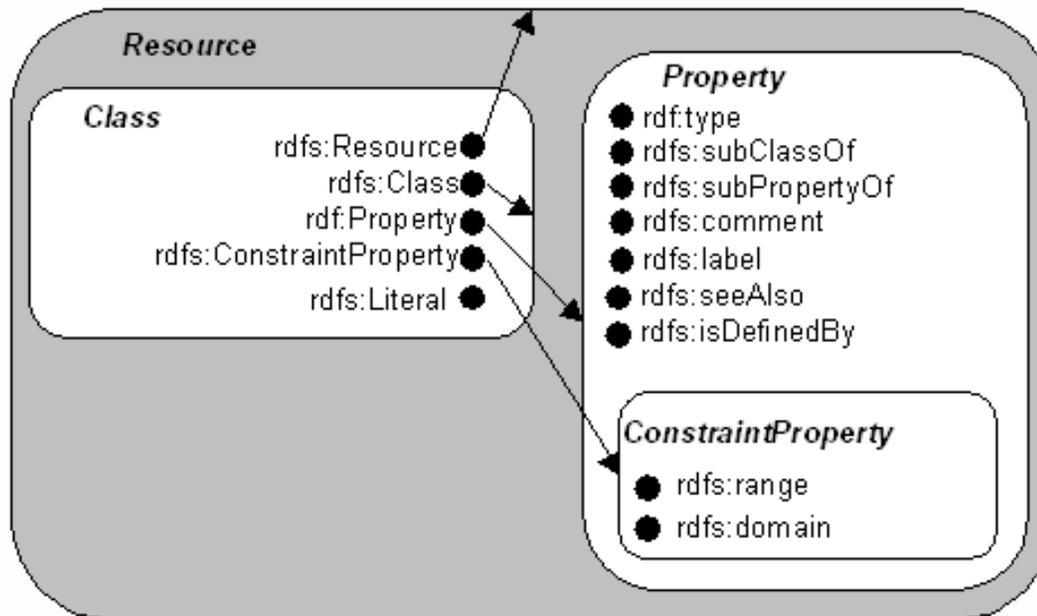
```
@prefix dc: <http://purl.org/dc/elements/1.1/> .  
<http://www.rhrk.uni-kl.de/~bstumm> dc:creator "Boris Stumm" .
```

- RDF/XML

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"  
  xmlns:dc="http://purl.org/dc/elements/1.1/">  
  <rdf:Description about="http://www.rhrk.uni-kl.de/~bstumm">  
    <dc:Creator>  
      Boris Stumm  
    </dc:Creator>  
  </rdf:Description>  
</rdf:RDF>
```

# RDF-Schema

- Erlaubt die Spezifikation von Schemas wie dem Dublin Core Element Set
- Klassen und Eigenschaften:



- Abbildungen zwischen Schemas nur durch `rdfs:subClassOf` möglich:  
Ersteller `rdfs:subClassOf dc:Creator` .

## DAML+OIL

- Darpa Agent Markup Language + Ontology Inference Layer
- Unterscheidung zwischen Objekten und Datentypen
- Klassendefinition durch
  - Vererbung
  - Mengenoperationen wie  
`daml:intersectionOf`, `daml:unionOf`, `daml:complementOf`
  - Aufzählung der Elemente
- Restriktionen auf Eigenschaften
  - Kardinalitätsrestriktionen
  - Typrestriktionen für Definitions- und Wertebereich
  - `daml:inverseOf`, `daml:TransitiveProperty`, ...
- Abbildungen zwischen Ontologien möglich mit `daml:equivalentTo`, `daml:sameClassAs`, `daml:inverseOf`, ...

- Beispiel einer Ontologie

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:daml="http://www.daml.org/2001/03/daml+oil#"
  xmlns="http://wwwdbis.informatik.uni-kl.de/.../ausarbeitung3.pdf#">

  <daml:Ontology rdf:about=""/>

  <daml:Class rdf:ID="Mensch">
    <daml:unionOf rdf:parseType="daml:collection">
      <daml:Class rdf:about="#Mann/>
      <daml:Class rdf:about="#Frau/>
    </daml:unionOf>
  </daml:Class>

  <daml:Class rdf:ID="Mann">
    <rdfs:subClassOf rdf:resource="#Mensch"/>
  </daml:Class>

  <daml:Class rdf:ID="Frau">
    <rdfs:subClassOf rdf:resource="#Mensch"/>
    <daml:disjointWith rdf:resource="#Mann"/>
  </daml:Class>
</rdf:RDF>
```

# Beweise und Ziehen von Schlüssen

- Schlußfolgern ist ein wichtiger Bestandteil des Semantic Web
- Beispiel 1:
  - Alle Studenten bekommen 50% Rabatt. Max ist ein Student. Also bekommt Max 50% Rabatt.
- Beispiel 2:
  - Schema A definiert `a:hatErsteller`, und Schema B definiert `b:istErstellerVon`.
  - Abbildung der Eigenschaften aufeinander:  
`a:hatErsteller daml:inverseOf b:istErstelltVon` .
- In den meisten Fällen reichen 2-3 Ableitungen
- Komplexe Schlußfolgerungen zu aufwendig und die Ergebnisse davon fragwürdig

# Open-World-Modell und Web of Trust

- Das Internet als offene Welt
  - dezentrales Internet
  - jeder kann beliebige (Fehl-)Informationen ins Internet stellen
  - Konsistenz von Informationen ist im allgemeinen nicht gewährleistet
- Das Web of Trust
  - verschiedene Stufen des Vertrauens
  - bei widersprechenden Aussagen kann die vertrauenswürdigste gewählt werden
  - Akzeptanz von Daten nur ab einem bestimmten Grad an Vertrauen

# Anwendungen

- Suchmaschinen
  - Angabe nicht nur von Suchwörtern, sondern auch deren Bedeutung
  - Bsp.: Suchwort „Koch“ soll ein Nachname sein und kein Beruf
  - Relevanz-Feedback
  
- Agenten
  - Reiseplaner
  - Termine suchen
  - billigste Anbieter finden
  
- Mozilla
  - SmartBrowsing und Aurora
  
- Open Directory Projekt
  - Datenbestand in RDF-Dateien frei verfügbar

# Ausblick

- Das Semantic Web steht noch am Anfang
- nicht alle Probleme werden gelöst
- ist aber unerlässlich für die Weiterentwicklung und effiziente Nutzung des Internet