

RDF und RDF Schema

Einführung in die Problematik
Von HTML über XML zu RDF

Kirsten Albrecht <3albrech@informatik.uni-hamburg.de>

Roland Illig <1illig@informatik.uni-hamburg.de>

Probleme des HTML-basierten Webs

- HTML ist eine Dokumentenbeschreibungssprache
- Enthält nur Daten
- Informationen werden erst beim menschlichen Benutzer erzeugt
- Computer können nur Daten verarbeiten
- Der Inhalt (Information) wird nicht erfasst
 - Probleme für Suchmaschinen, Einkaufsagenten, Datamining

Lösung durch XML?

- Allgemeine Beschreibungssprache
- Information und Darstellung getrennt
- Freie Definition der Begriffe und Zusammenhänge

Probleme von XML

- Keine eindeutige Schreibweise

```
<buch>  
  <titel>Studentenkochbuch</titel>  
  <autor>Heinz Otto</autor>  
  <preis>24 €</preis>  
</buch>
```

```
<buch titel="Studentenkochbuch"  
  preis="24 €" autor="Heinz Otto"/>
```

- Bedeutung der Daten nicht definiert
- Nur hierarchische Strukturen möglich

Lösung durch XML-Schema?

- Feste Definition der Syntax (Grammatik)
- Legt Elemente und Attribute fest

```
<xs:element name="buch">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="titel" type="xs:string" />
      <xs:element name="autor" type="xs:string" />
      <xs:element name="preis" type="xs:string" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

Probleme von XML-Schema

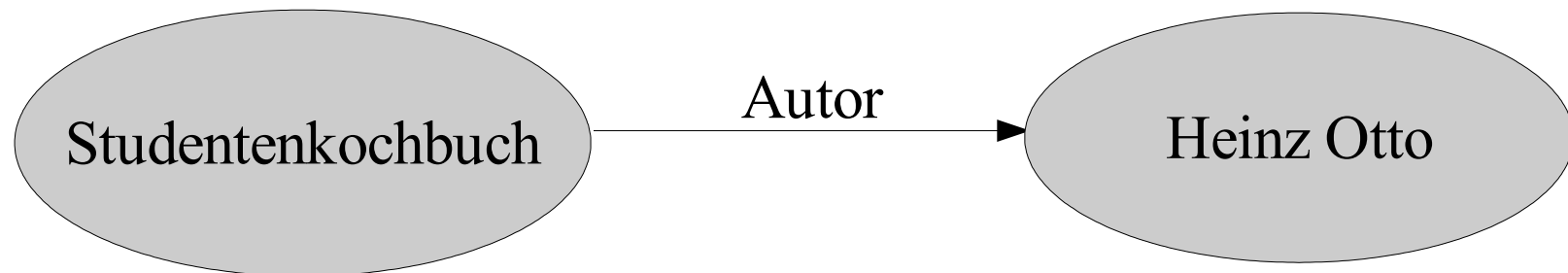
- Immer noch keine Definition der Semantik
- Beispiel: Was bedeutet der `<preis>` des Buches?
 - Wann war er gültig? (Währungsumrechnung)
 - Was genau enthält der Preis (Versand?)

Lösung durch RDF?

- RDF (Resource Description Framework) beschreibt Ressourcen (Dinge aller Art) und die Beziehungen zwischen ihnen
- Die Informationen werden in Dokumenten gespeichert

RDF: Statements

- RDF-Dokumente bestehen aus Statements
- Ein Statement besteht aus:
 - **Subjekt:** Die zu beschreibende Ressource
 - **Prädikat:** Die Eigenschaft der Ressource, die durch das Statement beschrieben werden soll
 - **Objekt:** Der Wert der Eigenschaft



Darstellung von RDF

- RDF ist zunächst ein abstraktes Modell
- Die Syntax bzw. Sprache ist nicht festgelegt
- Mögliche Sprachen sind:
 - N3 (wird hier nicht behandelt)
 - XML
 - Darstellung als Graphen

RDF und XML

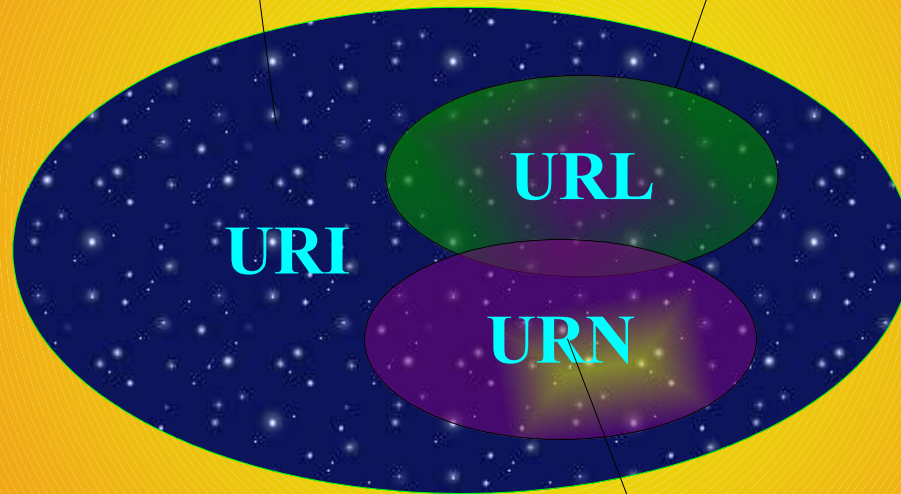
```
<rdf:Description rdf:about="StudentenKochbuch">  
  <author rdf:resource="HeinzOtto" />  
</rdf:Description>
```

- Bezeichner sollten eindeutig sein, wenn möglich sogar weltweit
- Lösungsansatz: URIs
 - weltweit eindeutig
 - lange Namen
 - können einen Hinweis geben, wie man an die Ressource gelangt

URI, URL und URN

Unique Resource Identifiers

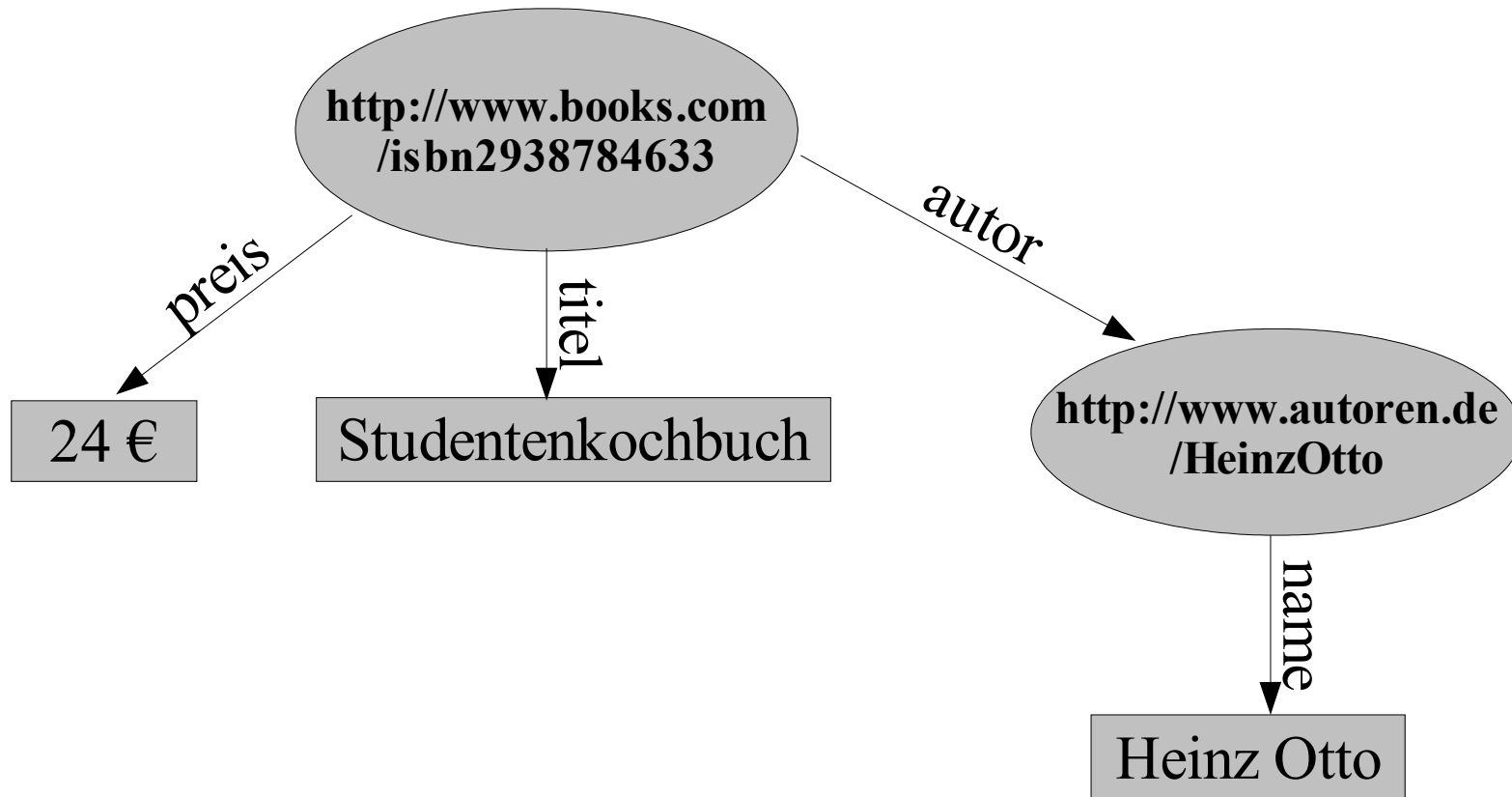
Uniform Resource Locators



Uniform Resource Names

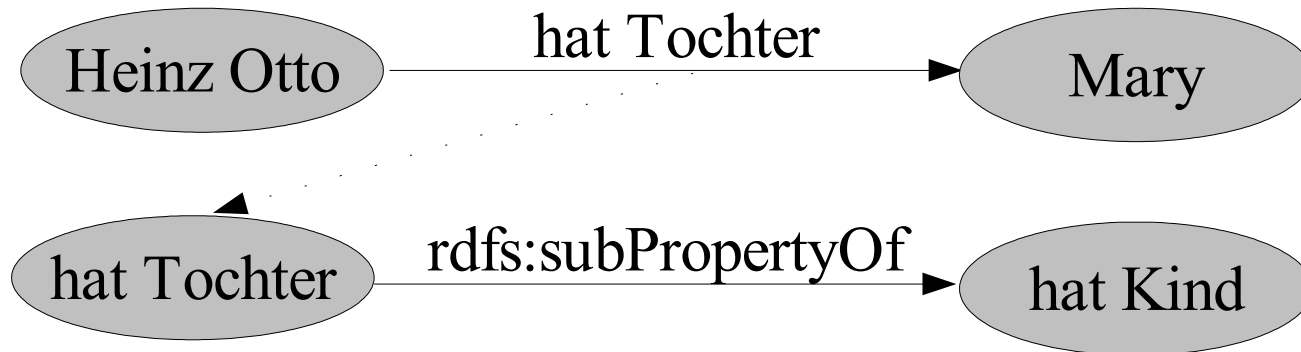
Komplexe Graphen

- Mehrere Statements kann man übersichtlich in Graphen darstellen.



Properties

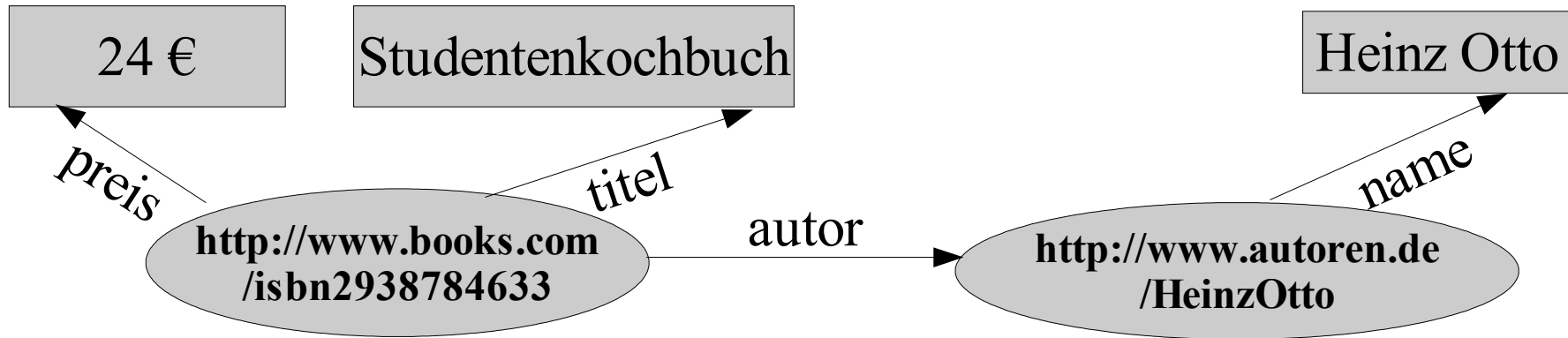
- RDF bietet die Möglichkeit Properties (Eigenschaften) näher zu beschreiben.
- dazu können die Properties durch URI identifiziert werden



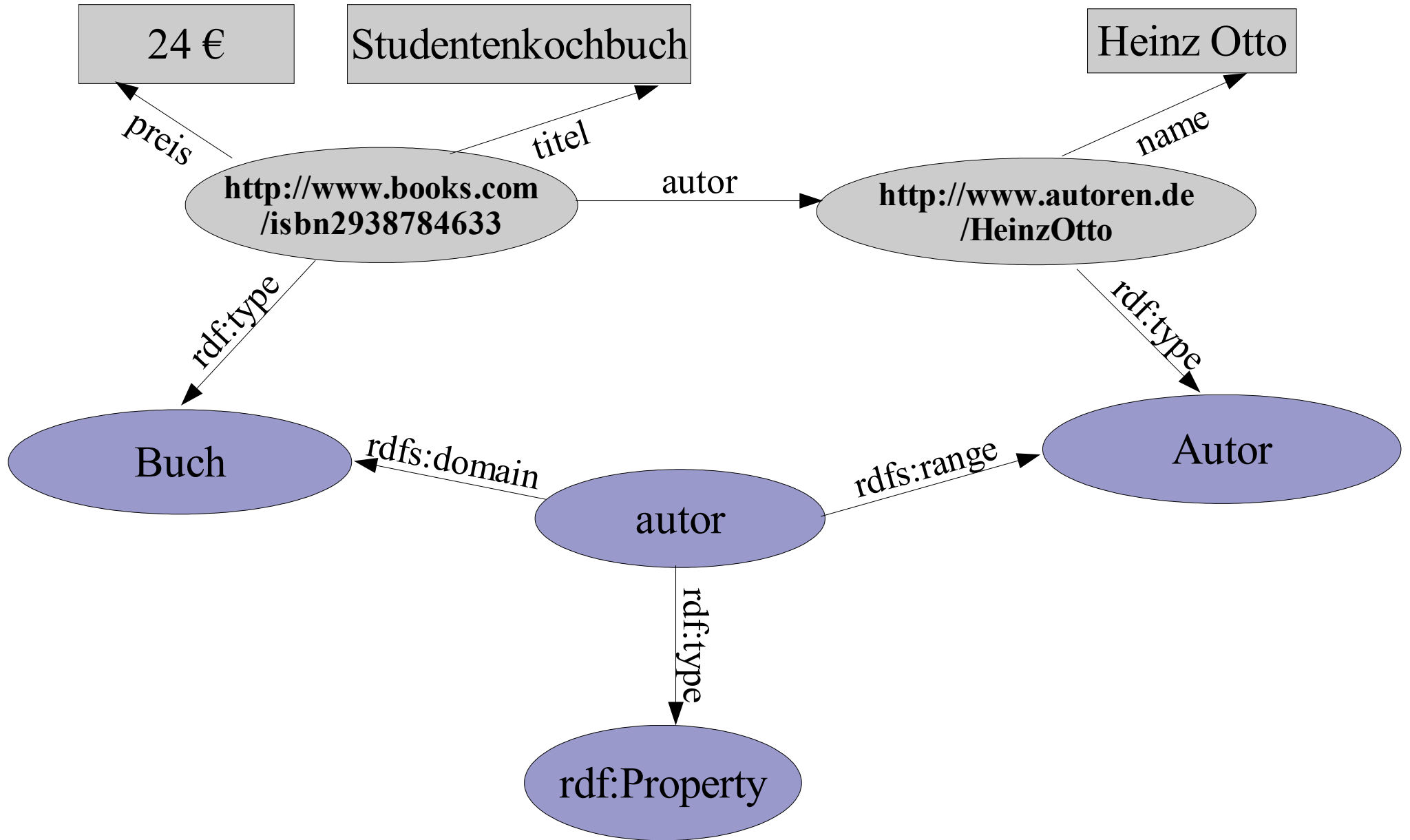
zusätzliche Konstrukte

- Ressourcen und Literale können in Containern zusammengefasst werden
 - Bag: ungeordnete Liste von Dingen
 - Sequence: geordnete Liste von Dingen
 - Alternative: eine Liste, aus der ein Ding ausgewählt wird

Probleme von RDF



RDF Schema



RDF Schema

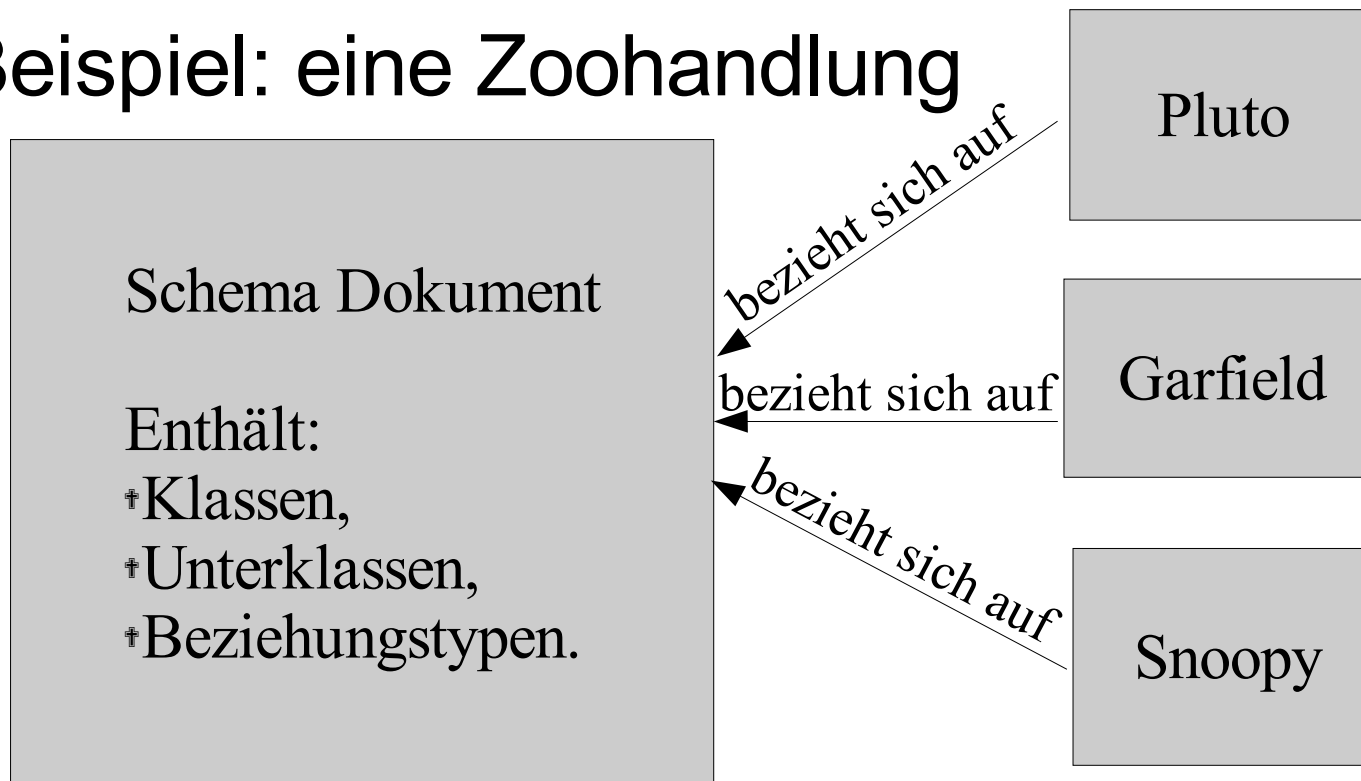
- RDF Schema stellt ein Standardvokabular zur Typisierung von RDF Ressourcen dar
- Es bietet Konzepte zur Beschreibung von Klassen von Ressourcen und Properties, sowie deren Zusammenhänge

RDFS Konstrukte

- `rdfs:class` Zur Definition von Klassen
- `rdfs:subClassOf` Zum Definieren von Klassenhierarchien
- `rdfs:subPropertyOf` Konstrukt zur Definition von Property-Hierarchien
- `rdfs:domain` zum Festlegen des Definitionsbereichs einer Property
- `rdfs:range` Zum Festlegen des Wertebereichs einer Property

Schemata

- Mit Hilfe des RDF Modells und der in RDF Schema definierten Konstrukte können nun Schemata gebildet werden
- Beispiel: eine Zoohandlung



Beispiel: Schemadokument

```
<rdf:RDF xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">

  <rdfs:Class rdf:ID="Haustier">
    <rdfs:comment>Klasse aller Haustiere</rdfs:comment>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Hund">
    <rdfs:comment>Der beste Freund des Menschen</rdfs:comment>
    <rdfs:subClassOf rdf:resource="#Haustier"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Futter" />

  <rdf:Property rdf:ID="frisstGerne">
    <rdfs:domain rdf:resource="#Haustier"/>
    <rdfs:range rdf:resource="#Futter"/>
  </rdf:Property >

</rdf:RDF>
```

Beispiel: Instanzdokument

```
<rdf:RDF xml:lang="de"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:zoo="http://zoohandlung.de/zoo.rdf">

  <zoo:Hund rdf:about="http://zoohandlung.de/pluto">
    <zoo:name>Pluto</zoo:name>
    <zoo:frisstGerne rdf:resource="http://zoohandlung.de/chappi"/>
  </zoo:Hund>

  <zoo:Futter rdf:about="http://zoohandlung.de/chappi">
    <zoo:name>Chappi</zoo:name>
    <zoo:preis>1,50 €</zoo:preis>
  </zoo:Futter>
</rdf:RDF>
```

Zusammenfassung

- Beschreibung von Metadaten mit deklarativer Semantik: sowohl über die verwendbaren Klassen, als auch über Properties können Statements formuliert werden
- Dies ermöglicht Computern die angebotenen Informationen eigenständig zu interpretieren
 - zum Beispiel kann eine Suchmaschine für Publikationen auch Bücher aufnehmen, da dies über die Klassenhierarchie definiert ist
- Für viele Anwendungsbereiche ist eine Modellierung mit RDFS möglich und ausreichend