

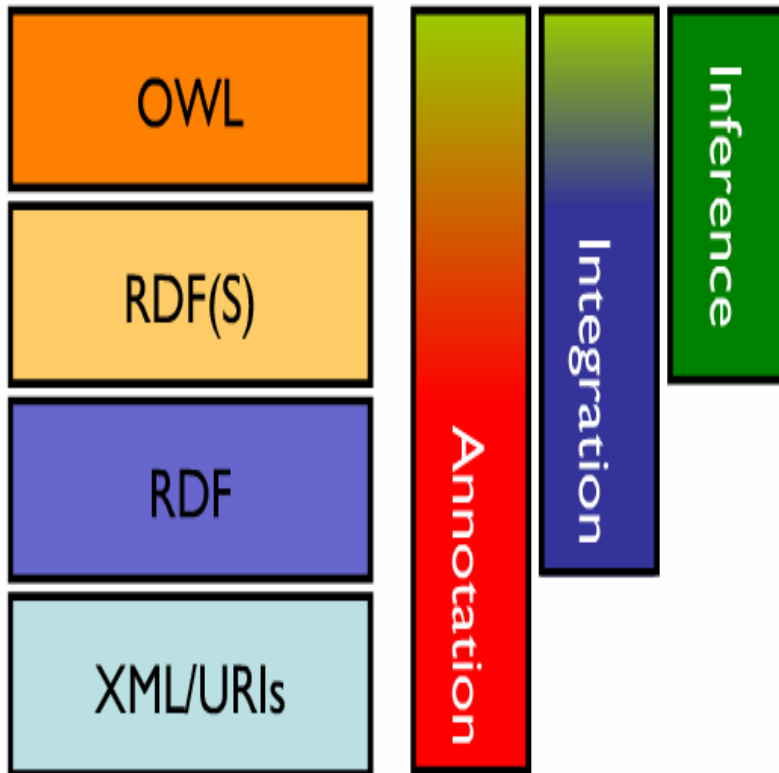
Intelligente Systeme im WWW: Semantic Web

RDF und RDF Schema

Dr. Raphael Volz
Institut AIFB, Universität Karlsruhe (TH)

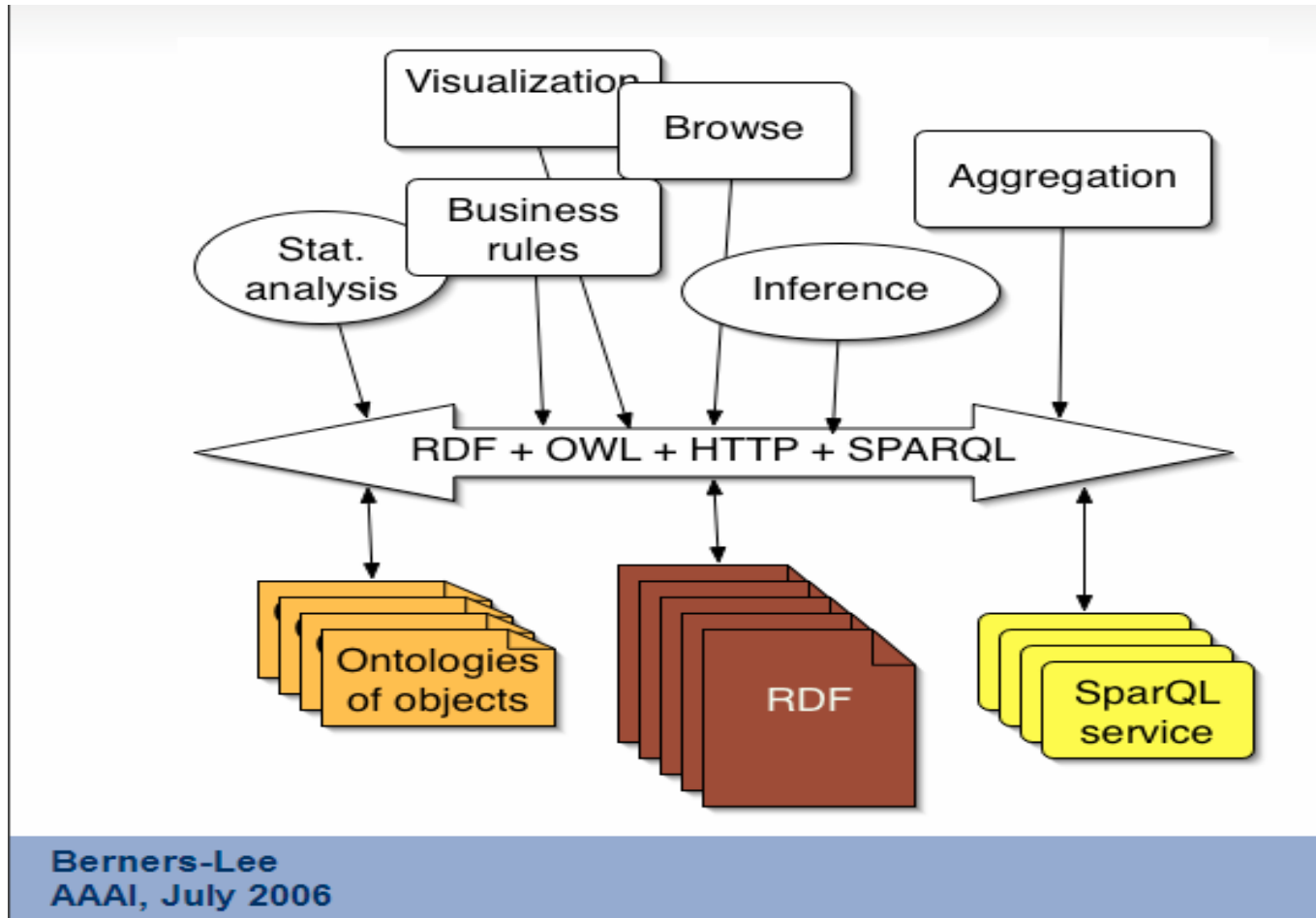
*(Most) Material courtesy
Prof. Dr. Heiner Stuckenschmidt
Uni Mannheim*

Semantic Web - High-Level Objectives



- **Annotation**
 - Associating metadata with resources
- **Integration**
 - Integration of information over several data sets (“Web of Data”)
- **Inference**
 - Reasoning over Data
 - Could be light-weight (RDF Schema)
 - Could be heavy-weight (OWL)

Further Uses of RDF...

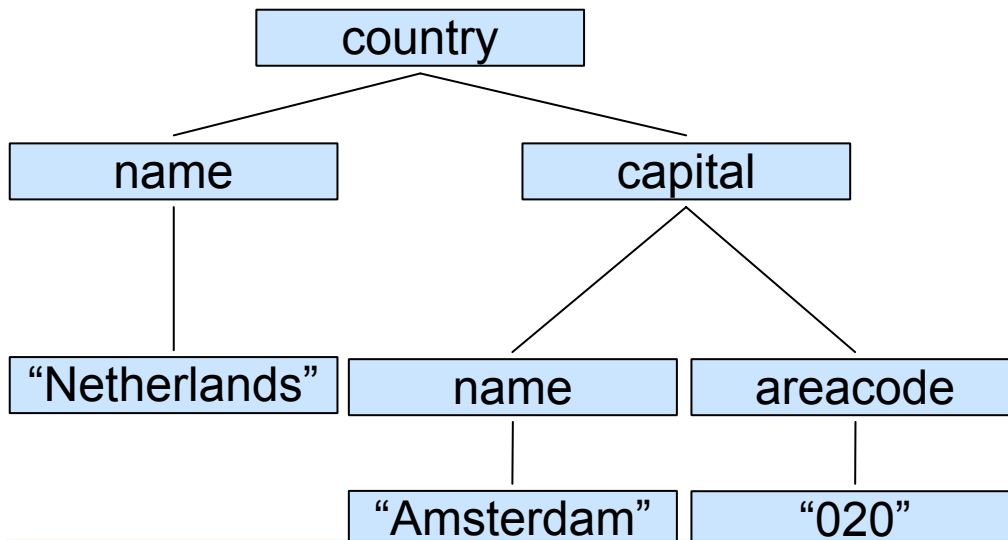


Inhalt

- Kurz: XML
- Was ist RDF
- RDF im Detail
 - Datenmodell
 - Syntax
- RDF Schema
 - Vocabulary
 - Schema Syntax und Inferenz
- Sesame: eine RDF Datenbank

Erinnerung: XML

```
<country name="Netherlands">  
  <capital name="Amsterdam">  
    <areacode>020</areacode>  
  </capital>  
</country>
```



- Syntax:
 - Elemente,
 - Attribute
 - Daten
- Datenmodell:
 - Baumstruktur

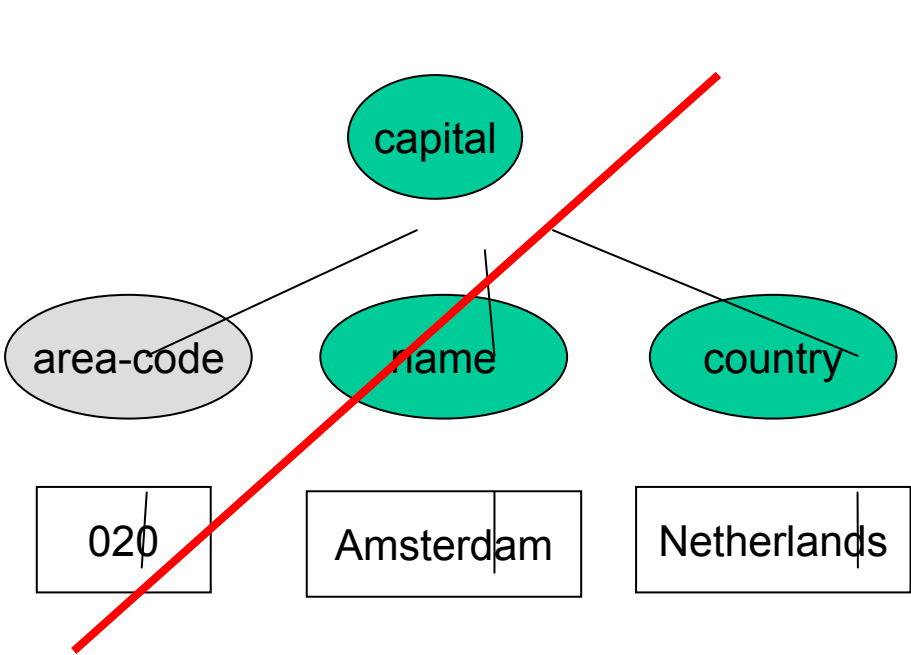
Probleme von XML

- Keine Festlegung auf eindeutige Struktur:
 - **ist country:**
 - Ein Object?
 - Eine Klasse?
 - Ein Attribut?
 - Eine Relation?
 - **Was bedeutet die Schachtelung ?**
- Keine Festlegung auf ein eindeutiges Vokabular
 - Ist **country** das gleiche wie **nation** ?

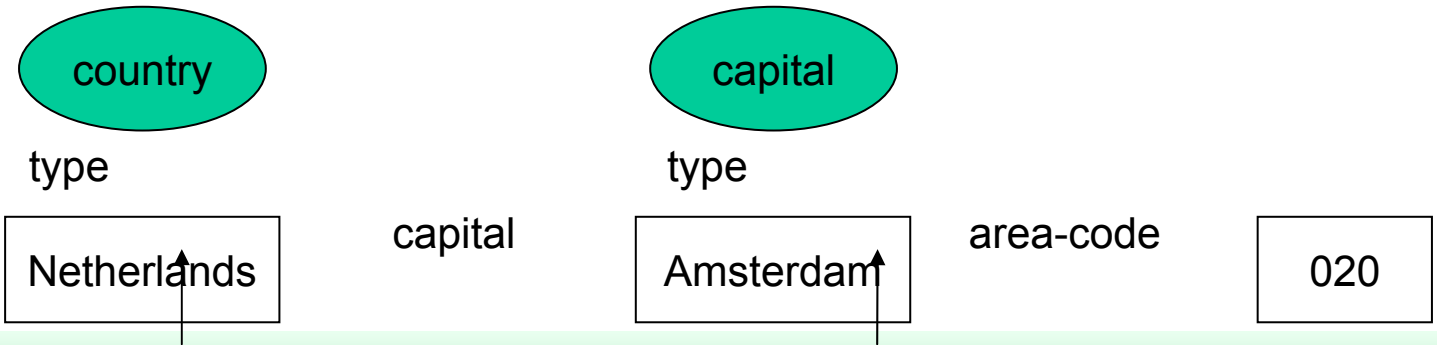
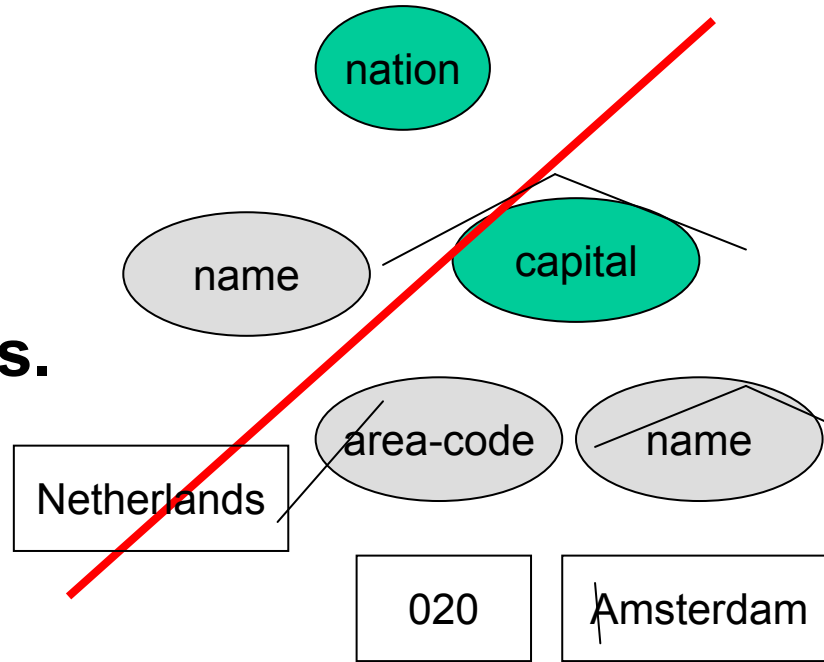
```
<capital areacode="020">  
  <name>Amsterdam</name>  
  <country>Netherlands</country>  
</capital>
```

```
<nation name="Netherlands">  
  <capital name="Amsterdam">  
    <areacode>020</areacode>  
  </capital>  
</nation>
```

Eindeutige Strukturen



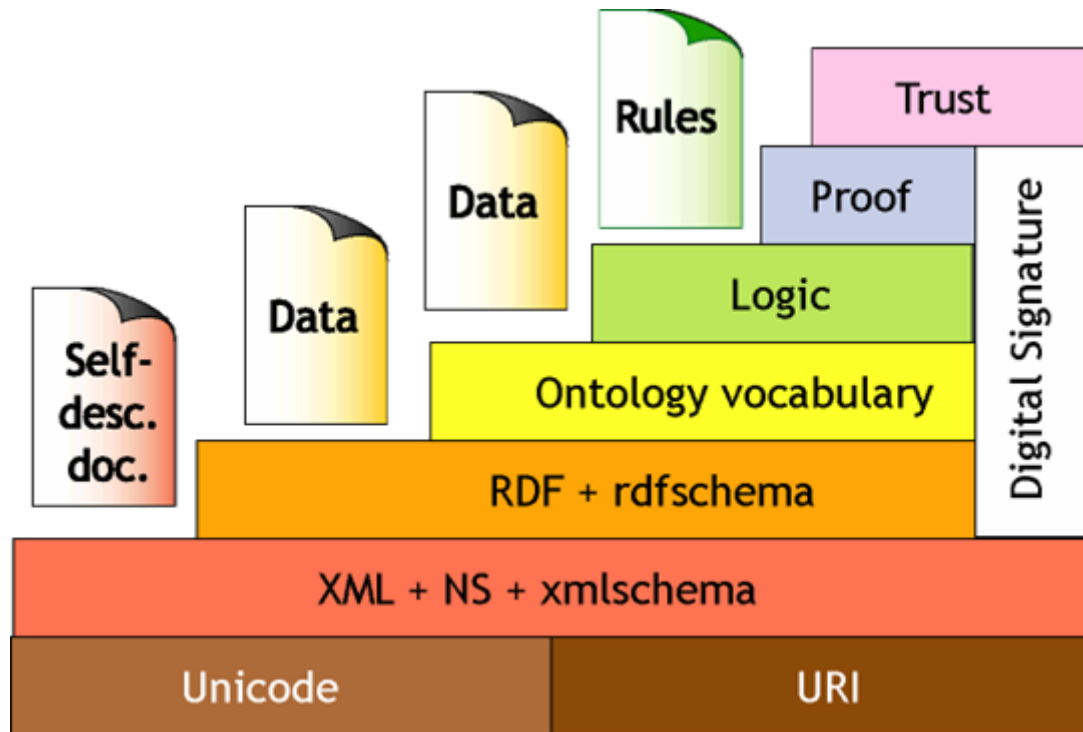
vs.



Was ist RDF?

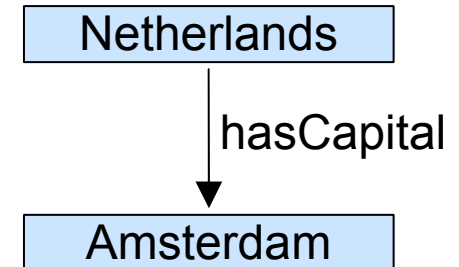
- RDF
 - “Resource Description Framework”
 - W3C Recommendation
(<http://www.w3.org/RDF>)
- RDF ist ein Datenmodell
 - Für **Metadaten** (daten über Daten)
 - RDF beschreibt die Bedeutung von Informationen in Maschinenlesbarer Form

Die Rolle von RDF im Semantic Web



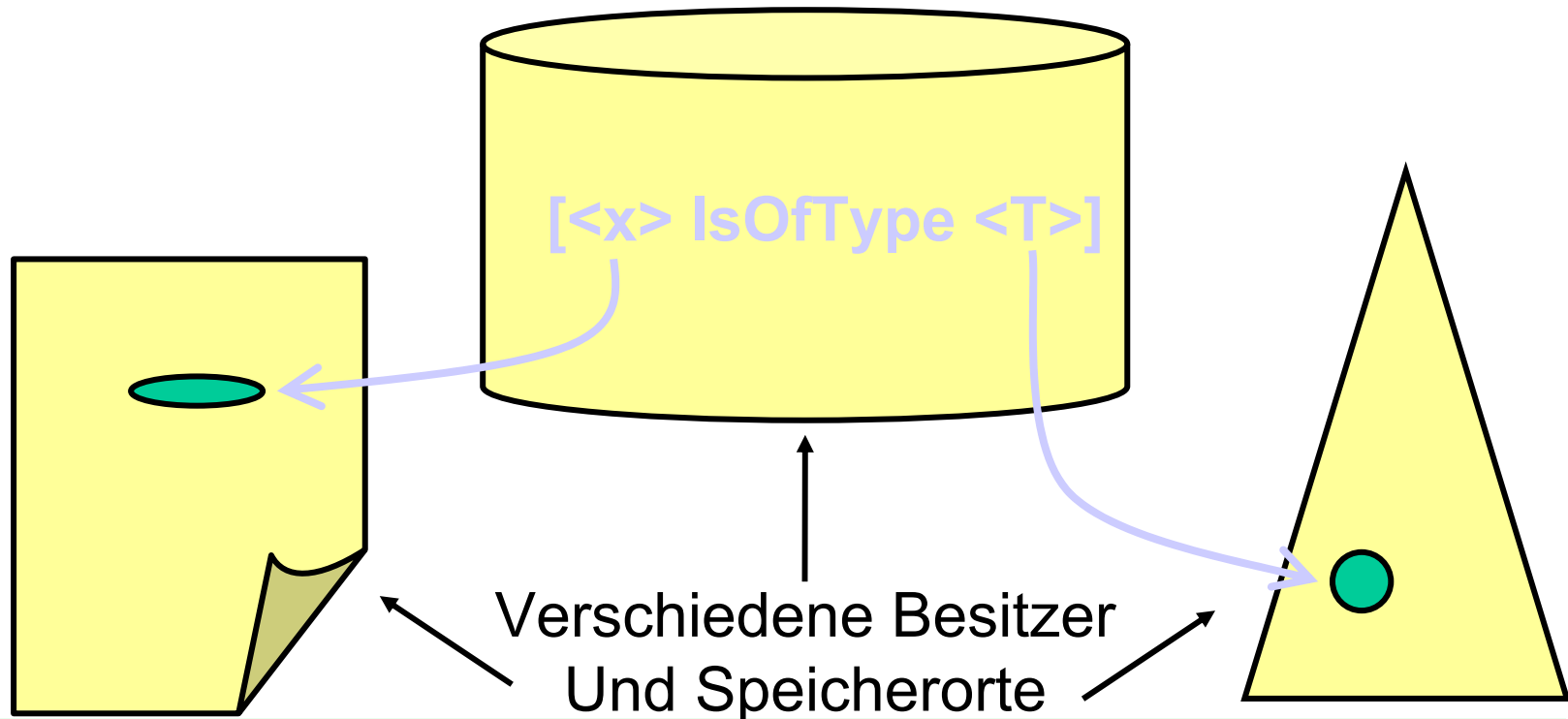
RDF im Detail: Das Datenmodell

- *statements* sind (subject, predicate, object) *triple*:
 - (Netherlands, hasCapital, Amsterdam)
- statements beschreiben *resources* (*Object*)
- *Eine resource kann alles sein, was eine URI besitzt*:
 - *Ein Dokument, ein Bild, ein Teil eines XML Dokuments*
 - <http://www.cs.vu.nl/index.html>
 - *ein Buch in einer Bibliothek*:
 - [isbn://5031-4444-3333](http://www.isbn.org/5031-4444-3333)



URI als Referenzen

- URIs können auf Informationen in unterschiedlichen Modellen verweisen:

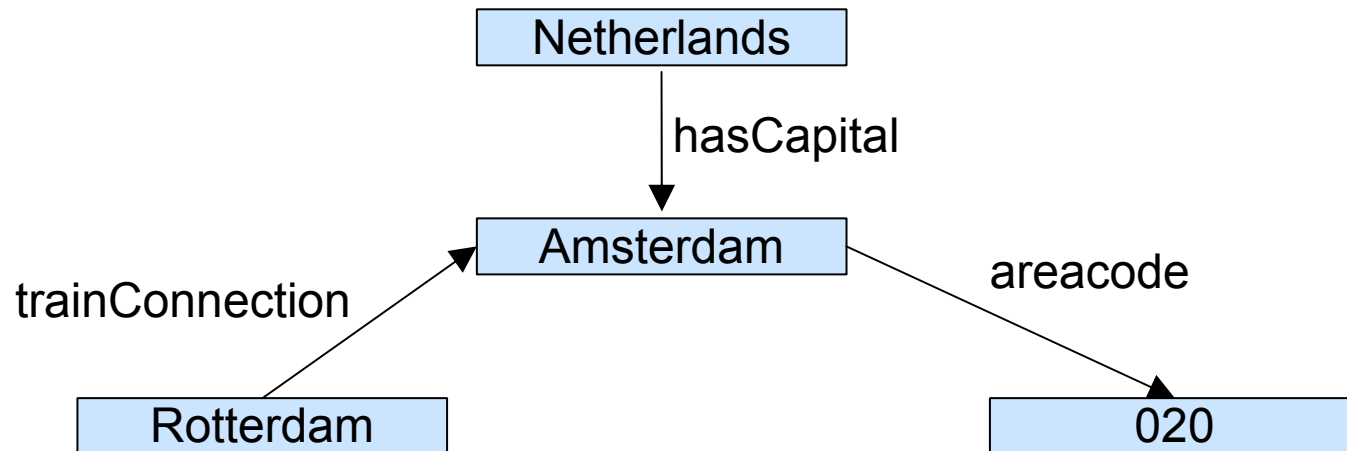


URIs

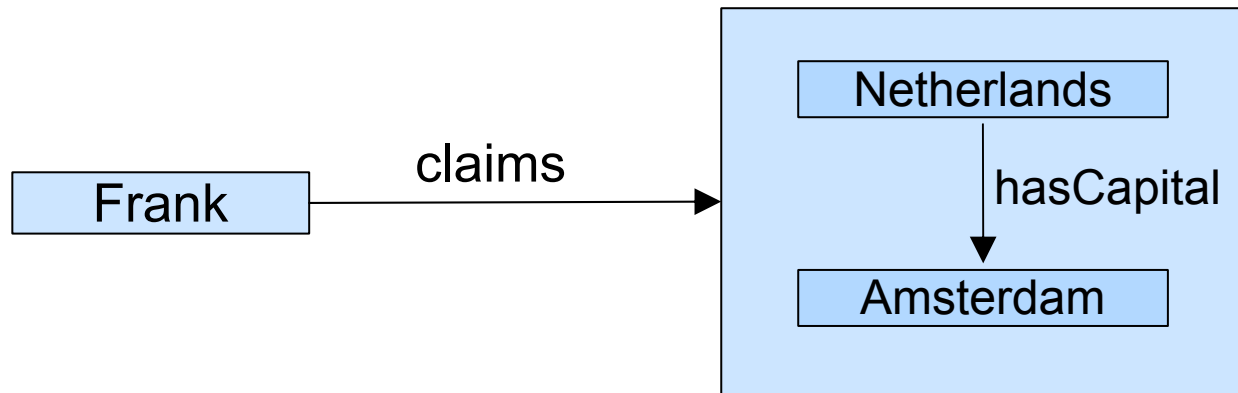
- URI = Uniform Resource Identifier
- “The generic set of all names/addresses that are short strings that refer to resources”
- URLs (Uniform Resource Locators) sind spezielle URIs, die Adressen von Webseiten angeben.
- In RDF werden häufig URLs verwendet – oft mit direkten Verweisen auf Teile des Dokuments:
 - <http://somedomain.com/some/path/to/file#fragmentID>

Verknüpfung von Statements

- Das subject eines Statements kann das object (oder auch predicate) eines anderen statements sein
- Mengen solcher Statements bilden semantische Netzwerke:



Reification: Ausagen über Aussagen



“Frank claims that the Netherlands
has a capital called Amsterdam”

RDF syntax: XML

- RDF hat eine spezielle XML-basierte Syntax
 - Jedes Description Element beschreibt eine resource
 - Jedes geschachtelte Element bezeichnet eine property
 - Das Attribut des property elements bezeichnet das object:

```
<Description about="http://www.countries.org/countries#Netherlands">  
  <hasCapital resource="http://www.cities.org/cities#Amsterdam"/>  
</Description>  
<Description about="http://www.cities.org/cities#Amsterdam">  
  <areacode>020</areacode>  
</Description>
```

RDF/XML syntax

- Wie in XML gibt es unterschiedliche Arten das gleiche aufzuschreiben, aber das Datenmodell bleibt das gleiche !

```
<Description about="http://www.countries.org/countries#Netherlands">
  <hasCapital resource="http://www.cities.org/cities#Amsterdam" />
</Description>
<Description about="http://www.cities.org/cities#Amsterdam">
  <areacode>020</areacode>
</Description>
```

```
<Description about="http://www.countries.org/countries#Netherlands">
  <hasCapital resource="http://www.cities.org/cities#Amsterdam" />
</Description>
<Description about="http://www.cities.org/cities#Amsterdam"
  areacode="020" />
```

```
<Description about="http://www.countries.org/countries#Netherlands">
  <hasCapital resource="http://www.cities.org/cities#Amsterdam">
    <areacode>020</areacode>
  </hasCapital>
</Description>
```

RDF/XML syntax: namespaces

- wie in XML werden namespaces eingesetzt, um tagnamen zu disambiguieren
- RDF eigene tags haben einen festgelegten Namespace, der Bezeichner ist standardmäßig 'rdf'

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:geo="http://www.geography.org/schema.rdf#"
         xmlns:words="http://www.dictionary.org/schema.rdf#">

  <rdf:Description rdf:about="#Netherlands">
    <geo:hasCapital rdf:resource="#Amsterdam"/>
    <words:hasCapital> N </words:hasCapital>
  </rdf:Description>
  <rdf:Description rdf:about="#Amsterdam">
    <geo:areacode>020</geo:areacode>
  </rdf:Description>
</rdf:RDF>
```

RDF(S) Syntax: Turtle

- Die XML Syntax ist nur EINE mögliche Form, RDF Modelle zu beschreiben.
- Hier eine andere:

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
@prefix geo: <http://www.geography.org/schema.rdf#>.
@prefix words: <http://www.dictionary.org/schema.rdf#>.
@prefix countries: <http://www.countries.org/countries#>.
@prefix cities. <http://www.cities.org/cities#>.

countries:Netherlands a geo:Country;
                        geo:hasCapital cities:Amsterdam;
                        words:hasCapital "N".
cities:Amsterdam geo:areacode "020".
```

Diskussion

- Vorteile von RDF :
 - Es wird explizit zwischen Objekten, Relationen und Konzepten unterschieden
 - Relationen zwischen Ressourcen werden explizit gemacht (auch über Modellgrenzen hinweg)
 - Das Datenmodelle ist invariant bzgl. unterschiedlicher Syntaktischer Varianten
- Offene Probleme
 - Es besteht keine Festlegung auf ein bestimmtes Vokabular:
 - **Sind country und nation nun das gleiche ?**
 - **Welche Eigenschaften kann country haben ?**

Metadaten

alleviates

<treatment>

<symptoms>

IS-A

<drug>

<drug administration>

林克星 根留台灣 可能增高

在愛慕者熱心奔走之下，華裔名指揮家林克星根留台灣的可行性又提升了幾分。兩廳院主任李炎、國家音樂廳樂團副團長黃奕明日前親赴林克星、石聖方高所拜會，並提出多場客席邀約。此外，台灣省立交響樂團團長陳澄達也早早「下訂」，邀請林克星赴台中購峰，從八月十三日起訓練省交，為期長達一個月。

在台灣諸多公家樂團中，陳澄達是以實際行動表達對林克星肯定的樂界人士之一，曾多次公開表示對林克星指揮才華的欽佩，而且幾乎每個樂季都邀請林克星客席演出。

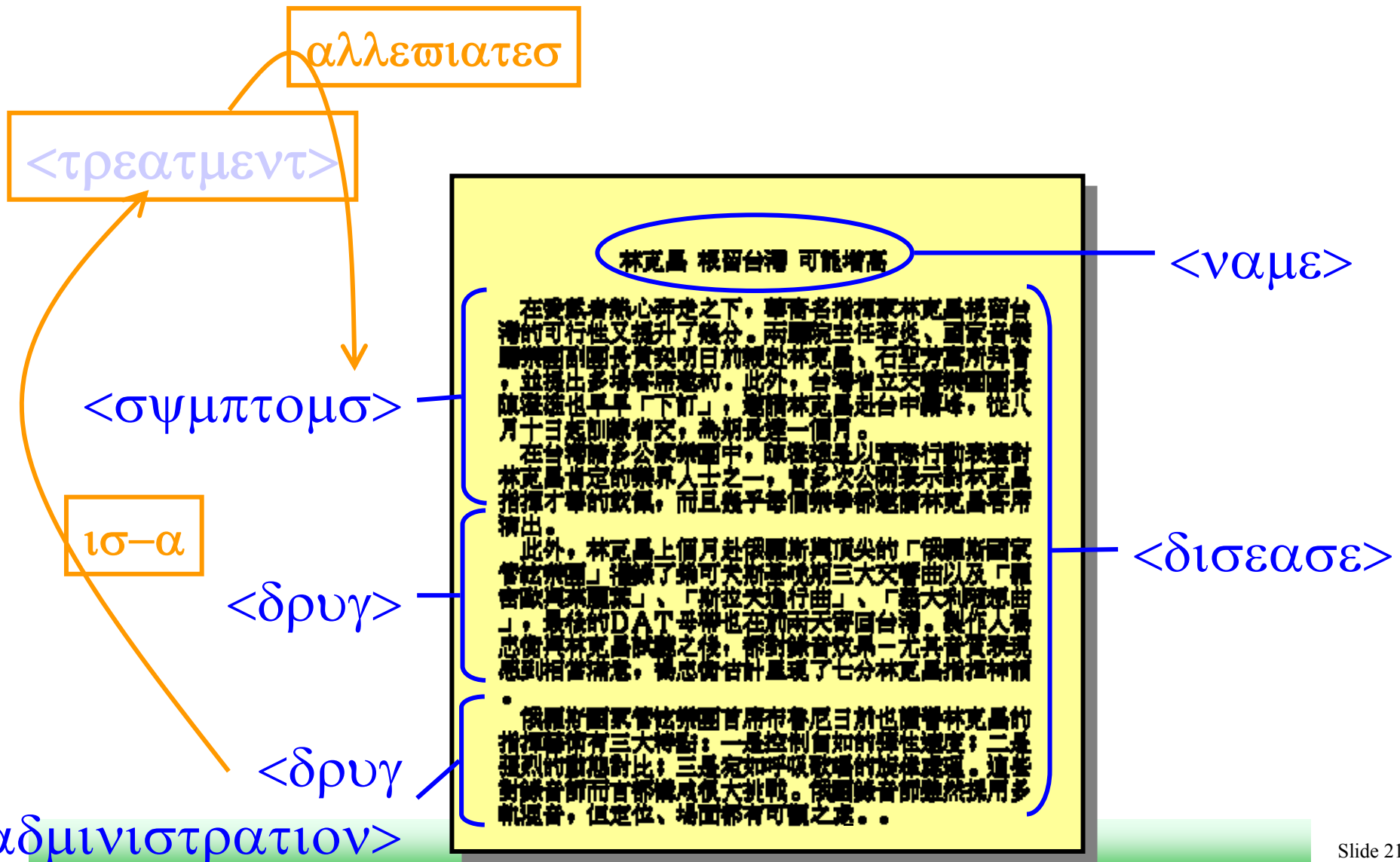
此外，林克星上個月赴俄羅斯與頂尖的「俄羅斯國家管絃樂團」灌錄了柴可夫斯基晚期三大交響曲以及「羅密歐與朱麗葉」、「斯拉夫進行曲」、「義大利隨想曲」，最後的DAT母帶也在前兩天寄回台灣。製作人楊忠衛與林克星試聽之後，都對錄音效果—尤其音質表現感到相當滿意，楊忠衛估計呈現了七分林克星指揮神韻。

- 俄羅斯國家管絃樂團首席布魯尼日前也讚譽林克星的指揮藝術有三大特點：一是控制自如的彈性速度；二是強烈的動態對比；三是宛如呼吸歌喉的旋律處理。這些對錄音師而言都構成很大挑戰。俄國錄音師雖然採用多軌混音，但定位、場面都有可觀之處。

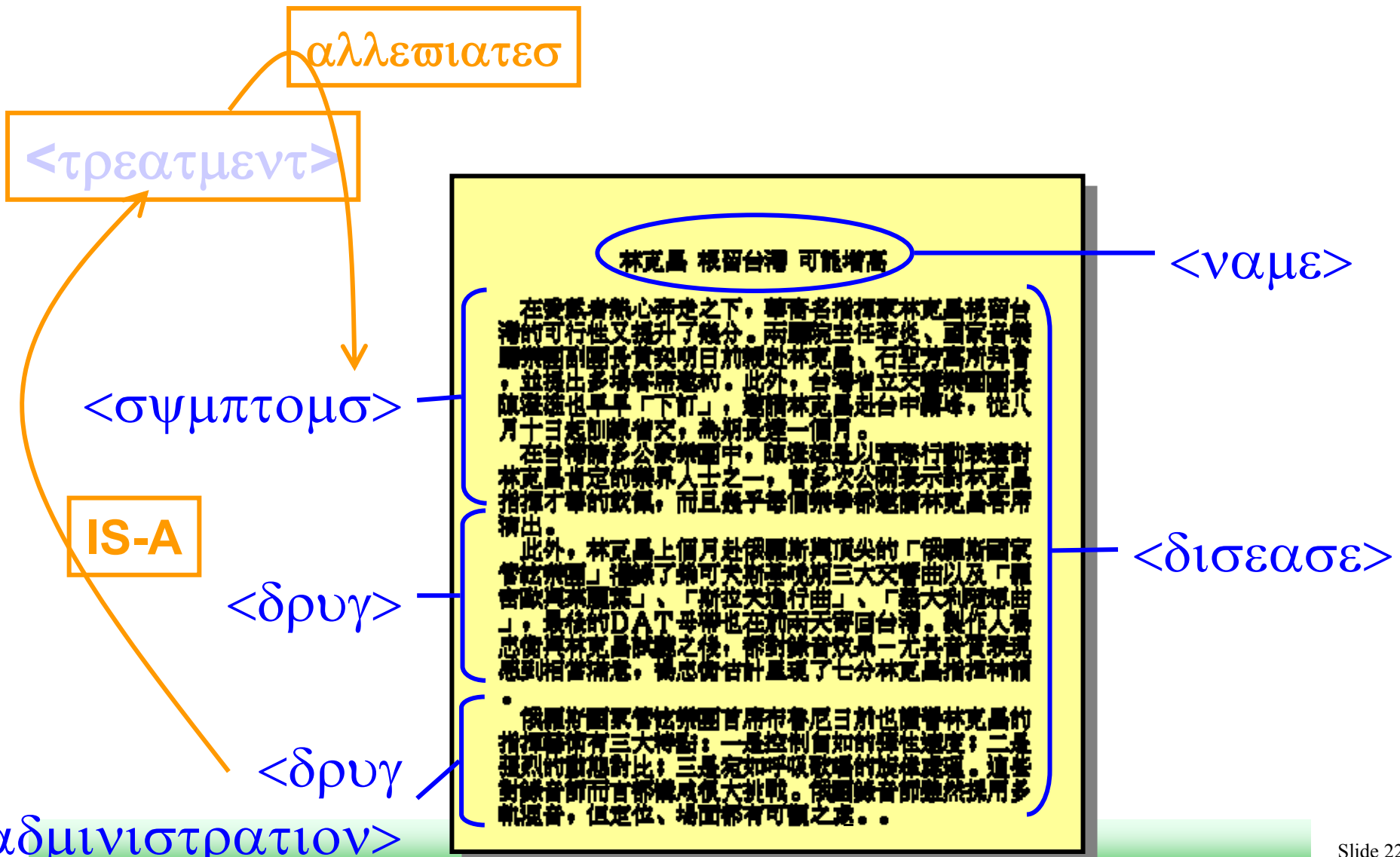
<name>

<disease>

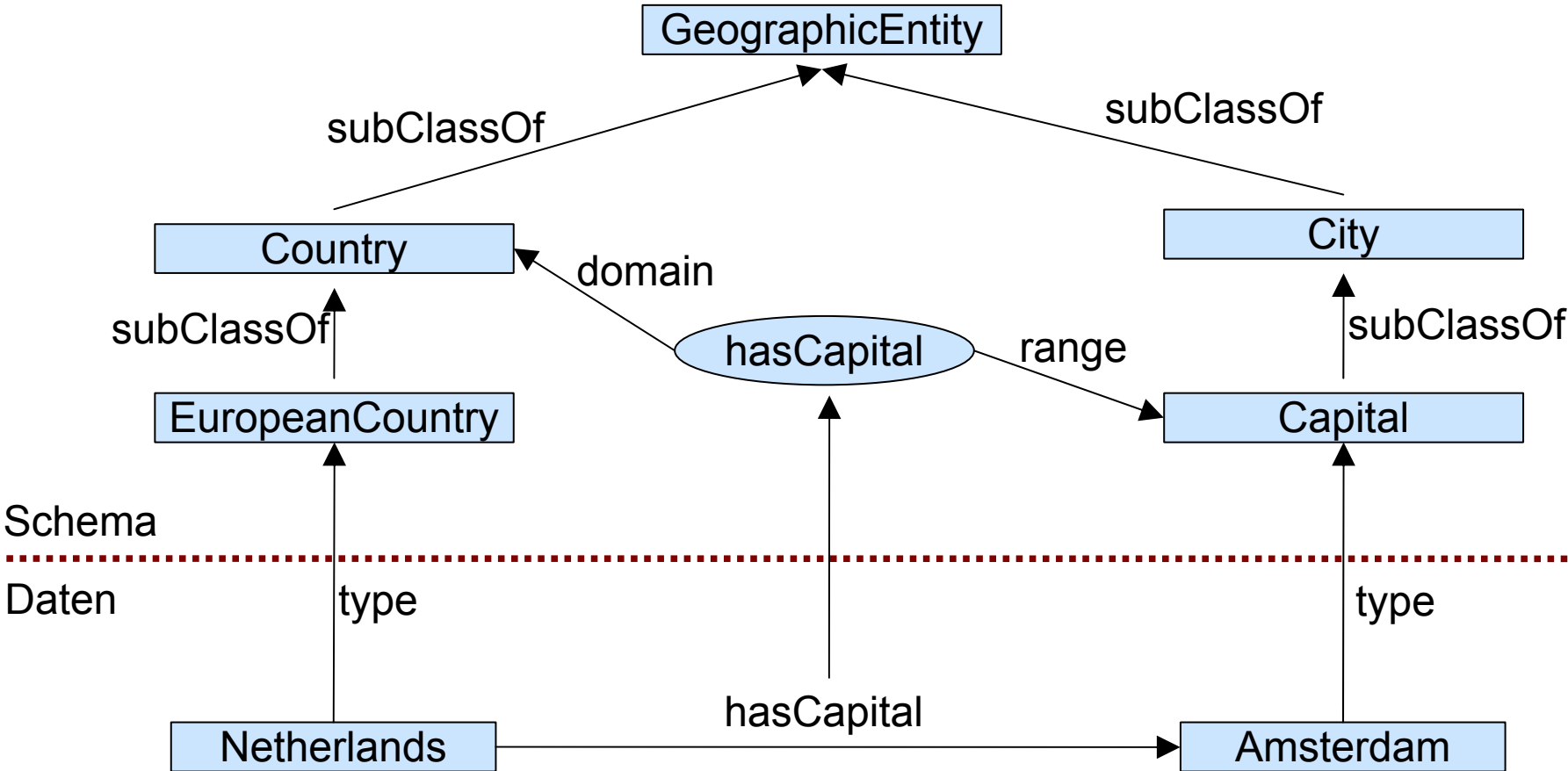
Metadaten in RDF



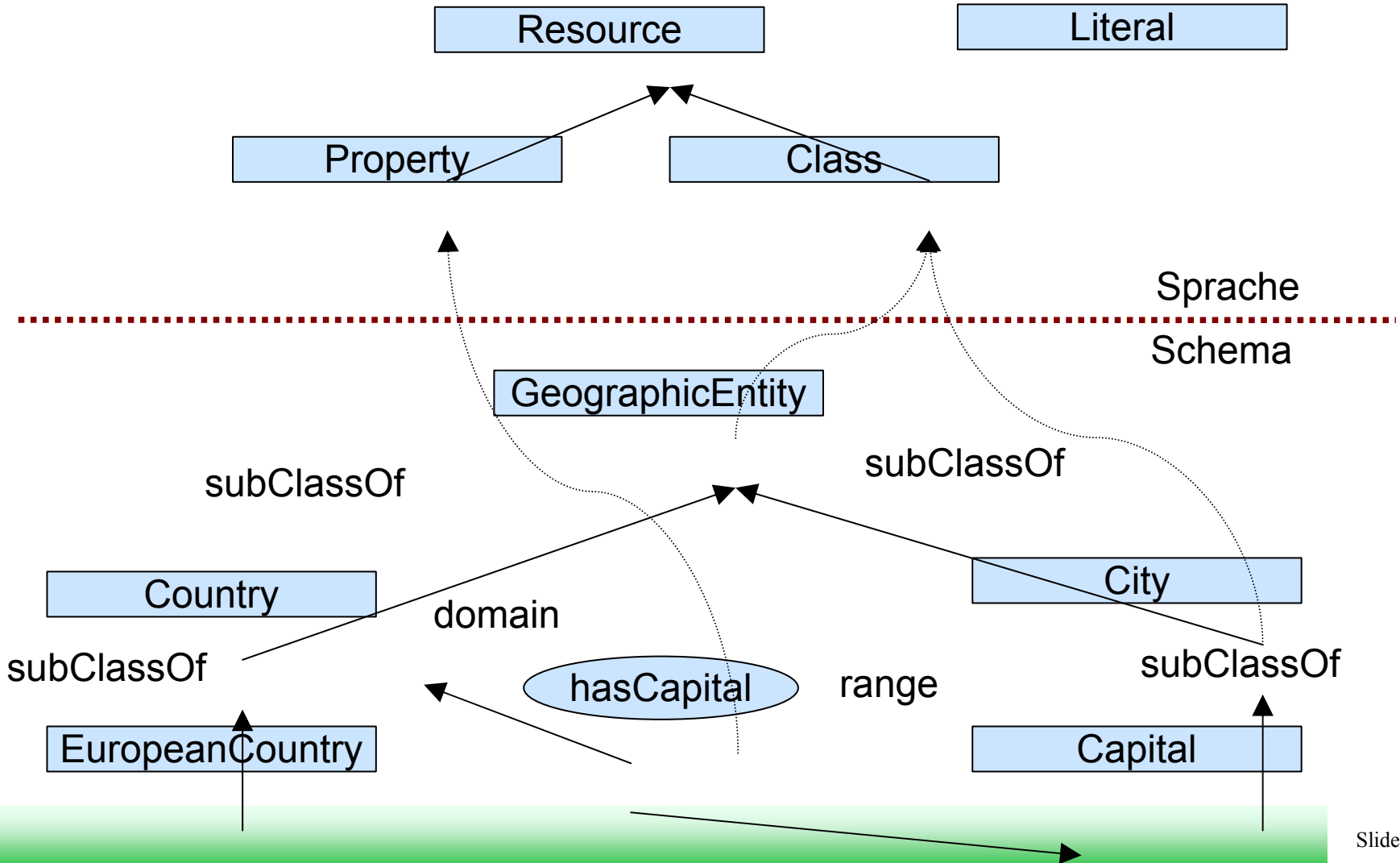
RDF Schema



RDF Schema



RDF Schema



Beobachtungen

- Relationen sind das wichtigste Instrument!
 - Relationen werden unabhängig von Klassen beschrieben
 - Unterschied zu gängigen Ansätzen im Software Engineering (z.B. UML) .
 - range/domain haben globale Bedeutung!
- Ein RDF Schema ist ein RDF Modell, in dem einige Relationen spezielle Bedeutung haben.
- Der Übergang zwischen den Metaebenen ist fließend
 - Ressourcen können gleichzeitig Instanzen und Klassen sein.

Die Semantik von RDF Schema

- Die Bedeutung spezieller Relationen wird im RDF Schema Standard definiert, z.B.
 - 2.3.2 rdfs:subClassOf

“This property specifies a subset/superset relation between classes. The rdfs:subClassOf property is transitive. If class A is a subclass of some broader class B, and B is a subclass of C, then A is also implicitly a subclass of C. Consequently, resources that are instances of class A will also be instances of class C, since A is a subset of both B and C. Only instances of rdfs:Class can have the rdfs:subClassOf property and the property value is always of rdf:type rdfs:Class. A class may be a subclass of more than one class.”

- Frage: gilt (A subclass A)?

Formale Semantik

- Welche statements lassen sich aus einem gegebenen Modell ableiten
- Menge von einfachen Ableitungsregeln
 - Einfach zu berechnen
 - Effizient in der Praxis
 - Keine kombinatorische Explosion
- Minimale Semantik ist bindend für alle Parteien.
- Darüber hinaus können anwendungsspezifisch weitere Regeln definiert werden

RDFS Semantik: Beispiel

- Netherlands **Type** EuropeanCountry
EuropeanCountry **subClassOf** Country
→ Netherlands **Type** Country
- aspirin alleviates headache
alleviates **range** symptom
→ headache **Type** symptom

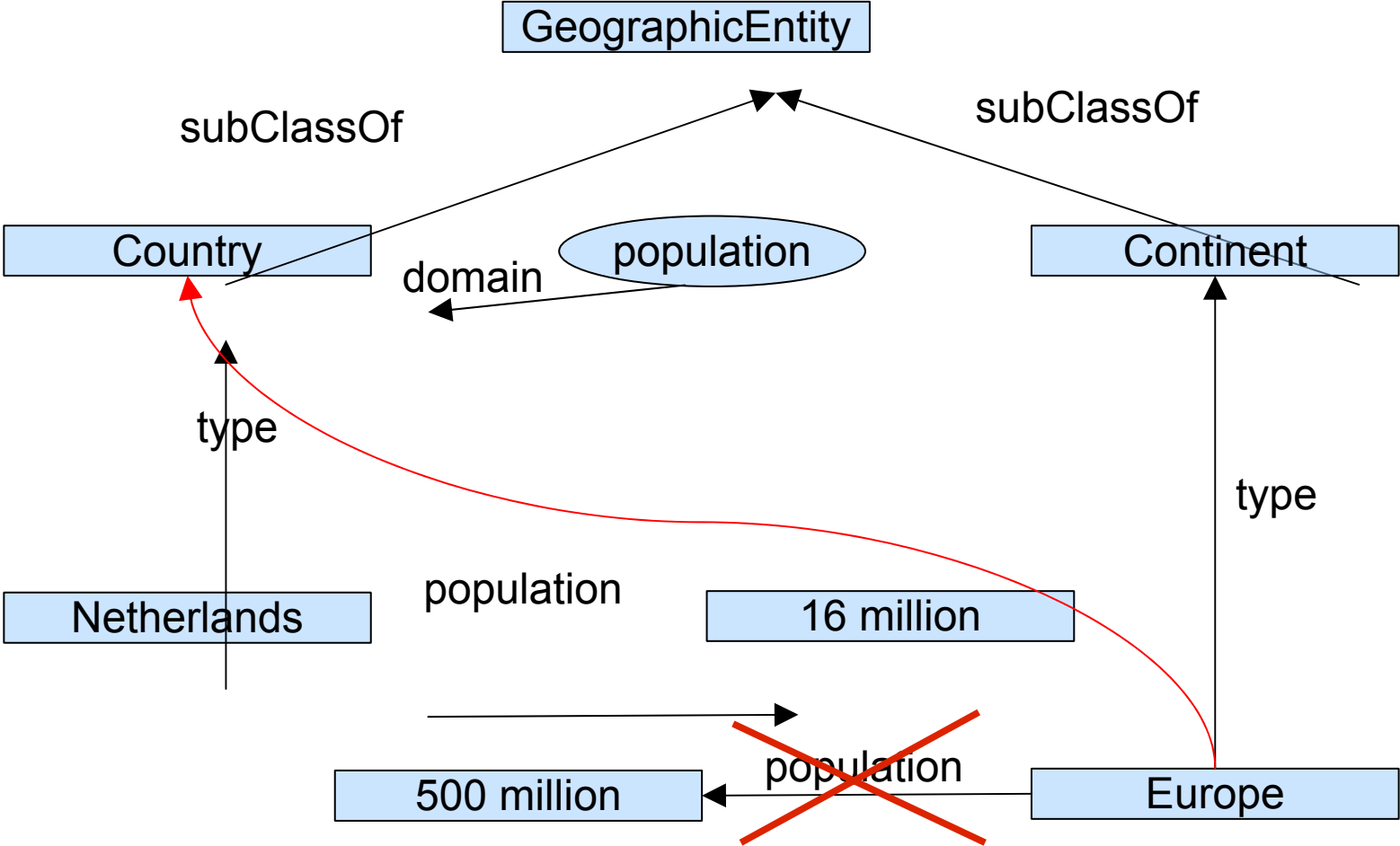
RDFS Semantik: Beispiel

- Νετηερλανδσ **Type** ΕυροπεανΧουντρψ
ΕυροπεανΧουντρψ **subClassOf** Χουντρψ
→ Νετηερλανδσ **Type** Χουντρψ
- ασπριν αλλεπιατεσ ηεαδαχηε
αλλεπιατεσ **range** σψμπτομ
→ ηεαδαχηε **Type** σψμπτομ

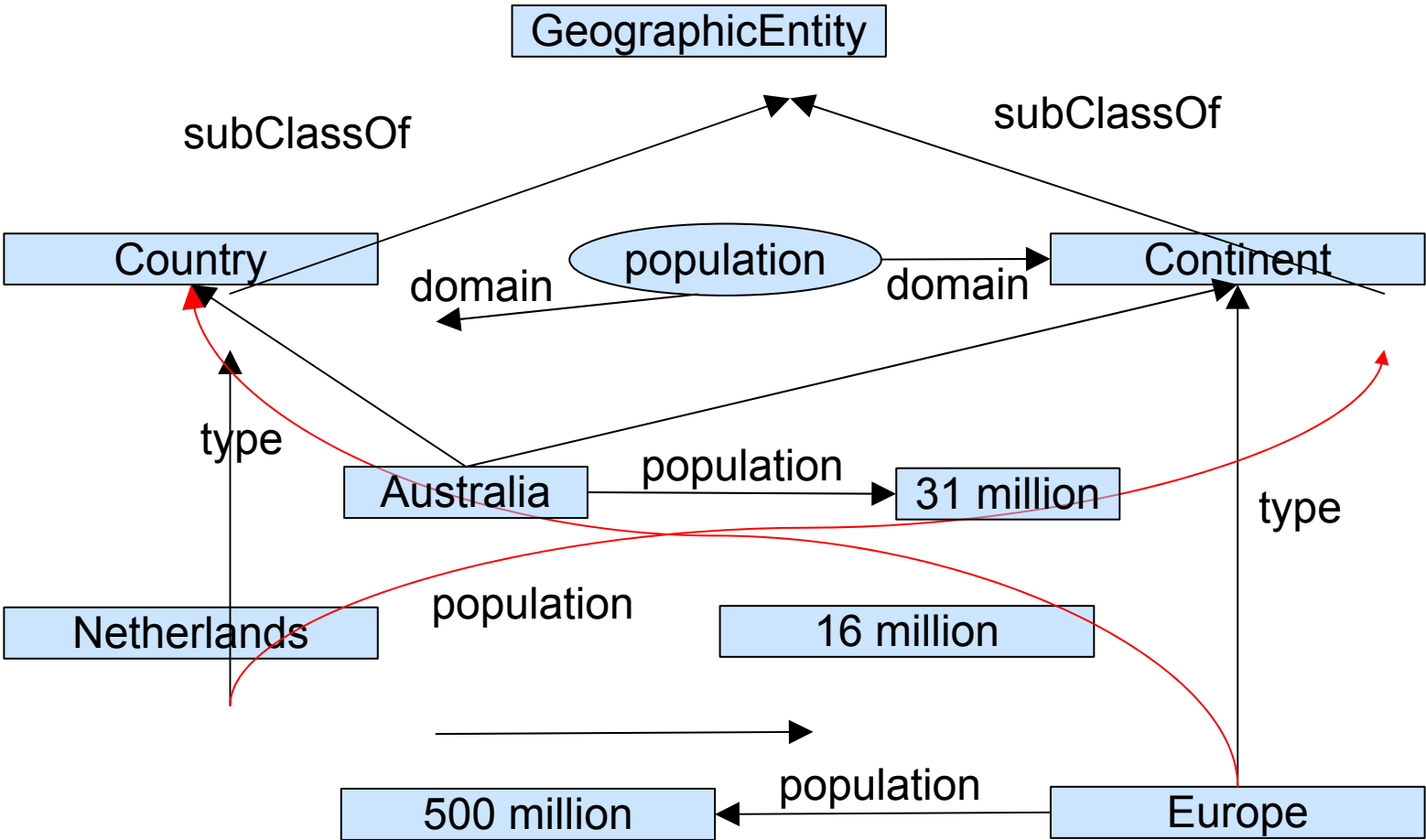
RDFS Semantik: Beispiel

- $X R Y + R \text{ domain } T \rightarrow X \text{ IsOfType } T$
- $X R Y + R \text{ range } T \rightarrow Y \text{ IsOfType } T$
- $T1 \text{ SubClassOf } T2 +$
 $T2 \text{ SubClassOf } T3 \rightarrow T1 \text{ SubClassOf } T3$
- $X \text{ Type } T1 +$
 $T1 \text{ SubClassOf } T2 \rightarrow X \text{ IsOfType } T1$
- $X \text{ Type Class} \rightarrow X \text{ SubClassOf } X$

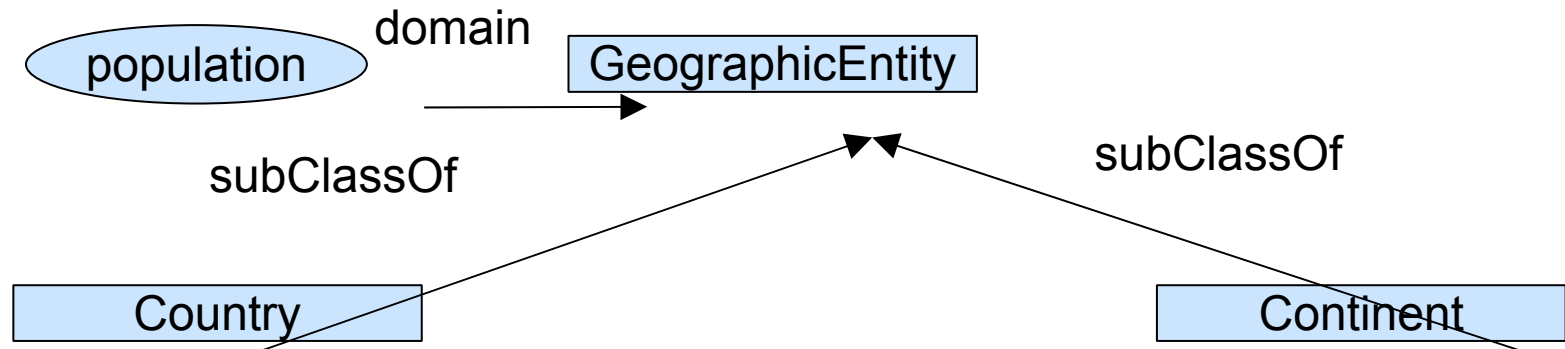
Globale Einschränkungen



Globale Einschränkungen



Globale Einschränkungen



- Solution
 - moving the domain restriction 'up in the hierarchy'
 - This solves the problem, but **over-generalization** is a danger:
 - **properties get very 'loose' restrictions**
 - **properties might be used on classes for which they are not meant (but it is allowed because the restriction could not be made more specific)**

RDF Schema Syntax

- Class definition

```
<rdf:Description rdf:about="#Country">  
  <rdf:type rdf:resource="#Class"/>  
  <rdfs:subClassOf rdf:resource="#GeographicEntity"/>  
</rdf:Description>
```

or shorter:

```
<rdfs:Class rdf:about="#Country">  
  <rdfs:subClassOf rdf:resource="#GeographicEntity"/>  
</rdfs:Class>
```

- Property definition

```
<rdfs:Property rdf:about="#hasCapital">  
  <rdfs:domain rdfs:resource="#Country"/>  
  <rdfs:range rdfs:resource="#Capital"/>  
</rdfs:Property>
```

Putting it all together

- Das schema file: <http://www.geography.org/schema.rdf>

```
<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:about="#Country">
    <rdfs:subClassOf rdfs:resource="#GeographicEntity"/>
  </rdfs:Class>

  <rdf:Property rdf:about="#hasCapital">
    <rdfs:domain rdfs:resource="#Country"/>
    <rdfs:range rdfs:resource="#Capital"/>
  </rdf:Property>
</rdf:RDF>
```

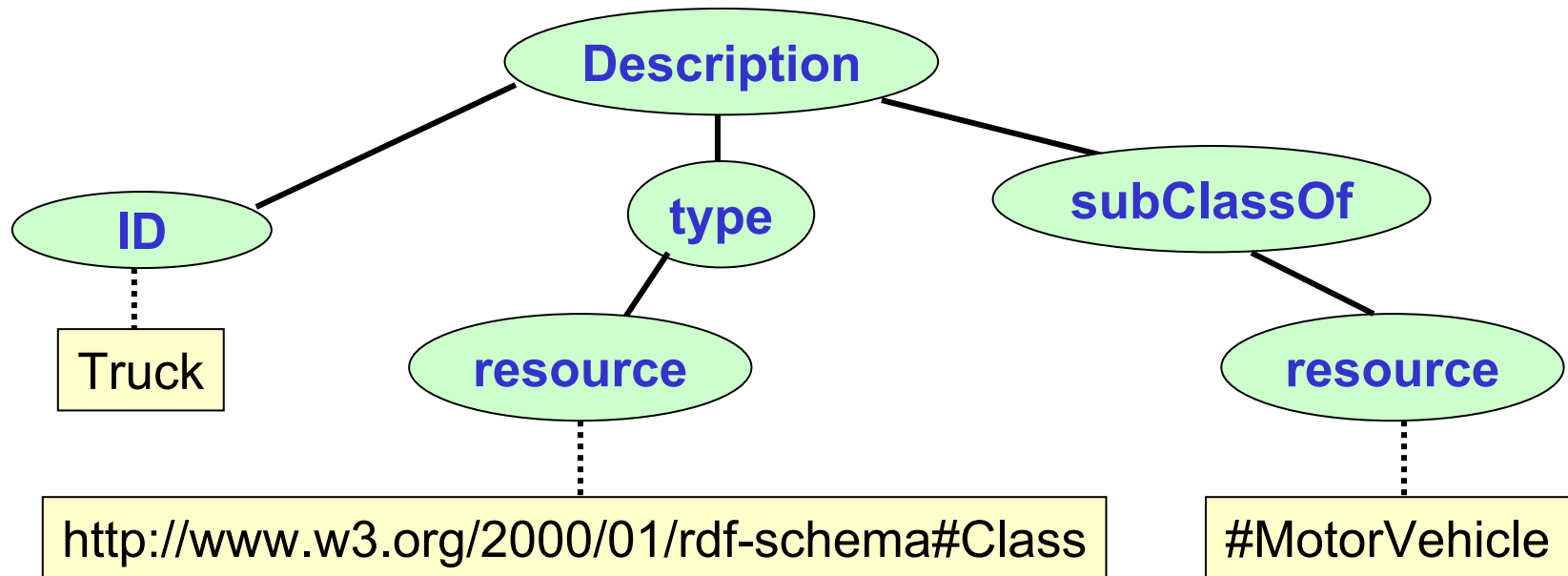
- Das Modell: kann irgendwo anders liegen

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:geo="http://www.geography.org/schema.rdf#">
  <geo:Country rdf:about="#Netherlands">
    <geo:hasCapital rdfs:resource="#Amsterdam"/>
  </geo:Country>
  <geo:Capital rdf:about="#Amsterdam"/>
</rdf:RDF>
```

Syntax und Interpretation

```
<rdf:Description ID="Truck">  
  <rdf:type resource="http://www.w3.org/2000/01/rdf-schema#Class"/>  
  <rdfs:subClassOf resource="#MotorVehicle"/>  
</rdf:Description>
```

- Interpretation als XML:



```

<rdf:Description ID="Truck">
  <rdf:type resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  <rdfs:subClassOf resource="#MotorVehicle"/>
</rdf:Description>

```

- Interpretation als RDF:
 - Anderes Datenmodell
 - **Description**, **ID** and **resource** haben eine festgelegte Bedeutung

subject**predicate****object**

1. 'Truck'

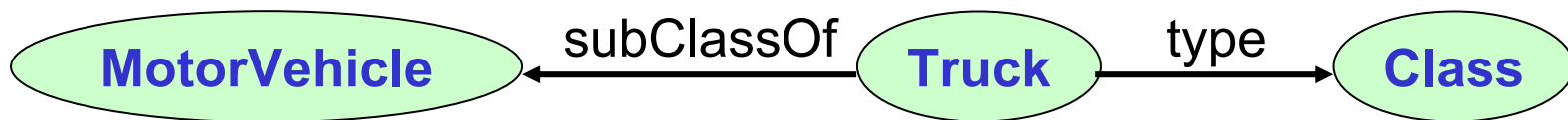
rdf:type

#Class

2. 'Truck'

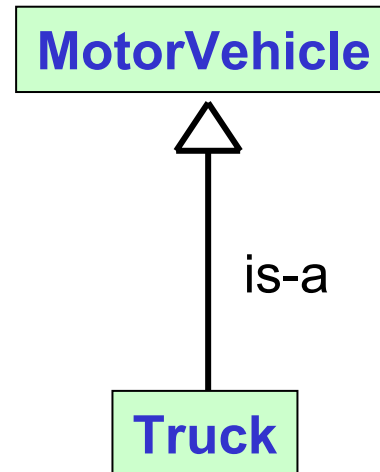
rdfs:subClassOf

#MotorVehicle



```
<rdf:Description ID="Truck">  
  <rdf:type resource="http://www.w3.org/2000/01/rdf-schema#Class"/>  
  <rdfs:subClassOf resource="#MotorVehicle"/>  
</rdf:Description>
```

- Interpretation als RDF Schema
 - Wieder anderes Datenmodell
 - `type` and `subClassOf` werden speziell interpretiert



Zusammenfassung

- RDF ist ein einfaches, Graph-basiertes Datenmodell für metadaten im Web
- RDF hat eine XML Syntax und kann so in Web Dokumente integriert werden
- Vorteile gegenüber XML
 - Datenmodell ist invariant gegenüber syntaktischen Variationen
 - Informationen aus unterschiedlichen Modellen können verbunden werden
 - Bestimmte Operationen, z.B. die Vereinigung zweier Modelle ist trivial
- RDF Schema erlaubt die Spezifikation von Hintergrundwissen
 - Vokabular: Class, SubClassOf, domain, range
 - Einfache Ableitungsregeln können implizites Wissen herleiten
 - RDF Schema ist RDF mit speziellen relationen