

# SEMANTIC WEB TECHNOLOGIES I

Lehrveranstaltung im WS09/10

Dr. Andreas Harth  
Dr. Sebastian Rudolph  
M.Sc. Anees ul Mehdi

# AGENDA



- Vorstellung der Dozenten
- Organisatorisches zur Vorlesung
- Was ist das "Semantic Web"?
- "Semantic Karlsruhe"

# AGENDA

- **Vorstellung der Dozenten**
- Organisatorisches zur Vorlesung
- Was ist das "Semantic Web"?
- "Semantic Karlsruhe"

# VORSTELLUNG DER DOZENTEN



## **Dr. Andreas Harth**

- Diplom Informatik an der FH Würzburg
- Promotion (Informatik) am Digital Enterprise Research Institute, National University Ireland, Galway



## **Dr. Sebastian Rudolph**

- Staatsexamen (Lehramt) an der TU Dresden
- Promotion (Mathematik) am Institut für Algebra an der TU Dresden



## **M.Sc. Anees ul Mehdi**

- Master am International Center for Computational Logic an der TU Dresden

# AGENDA

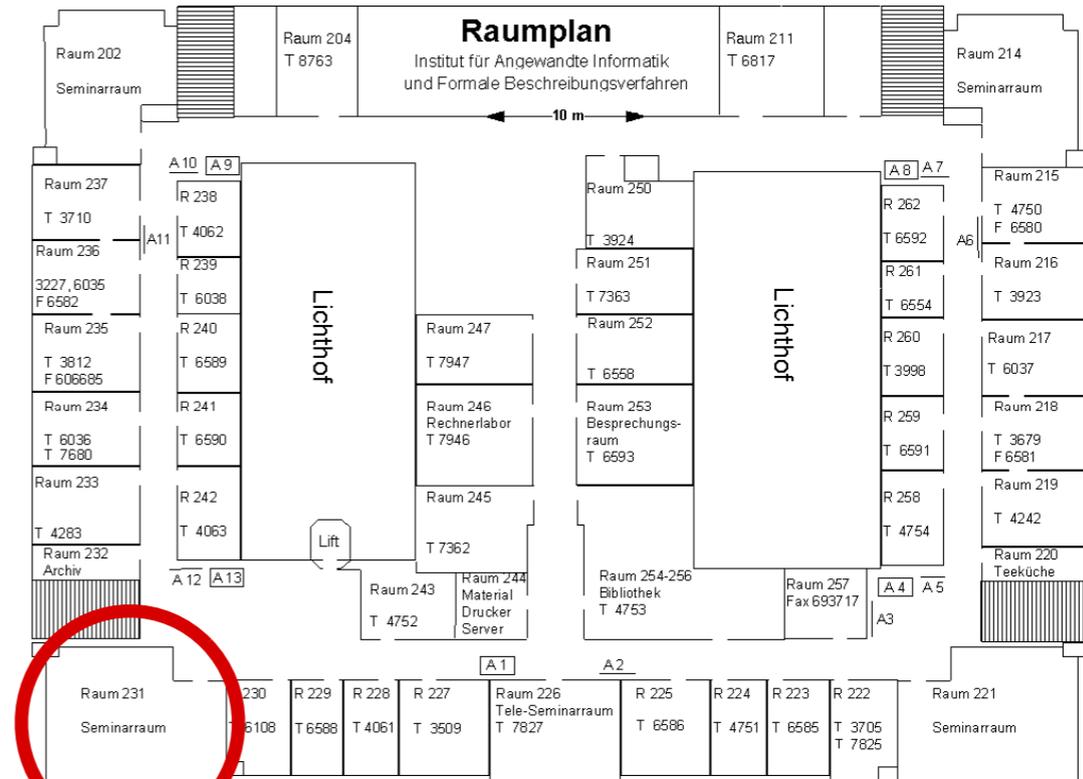
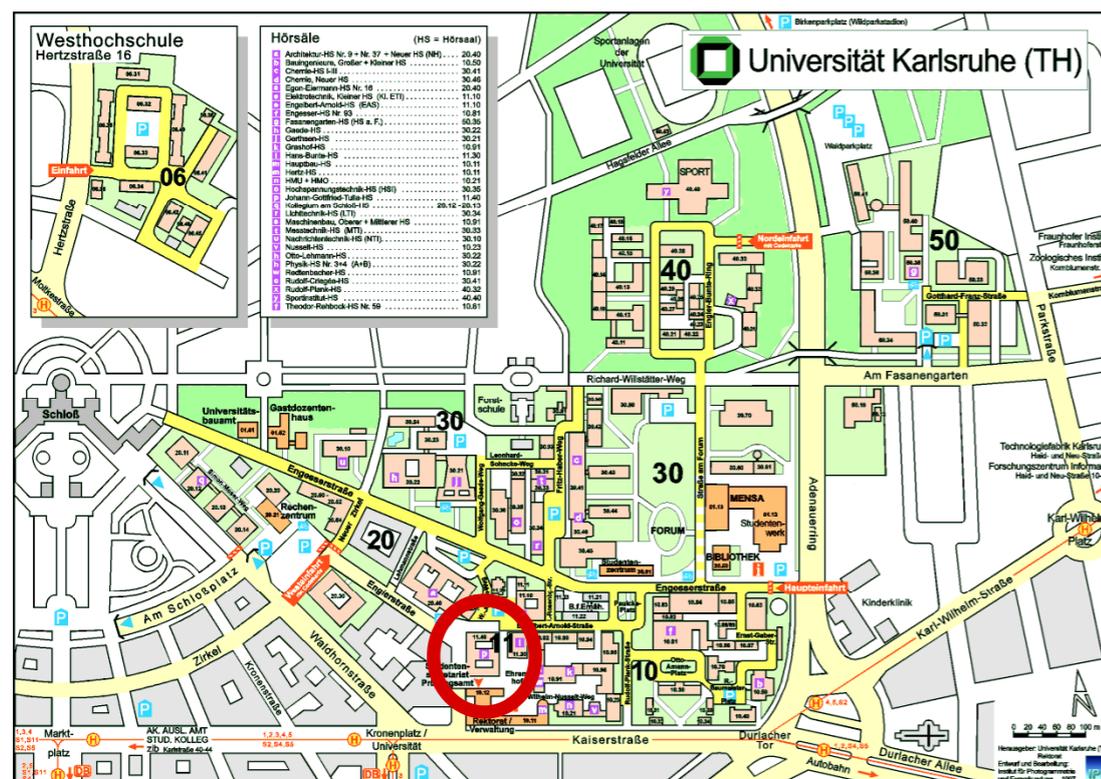
- Vorstellung der Dozenten
- **Organisatorisches zur Vorlesung**
- Was ist das "Semantic Web"?
- "Semantic Karlsruhe"

# ORGANISATORISCHES: ZEIT UND ORT



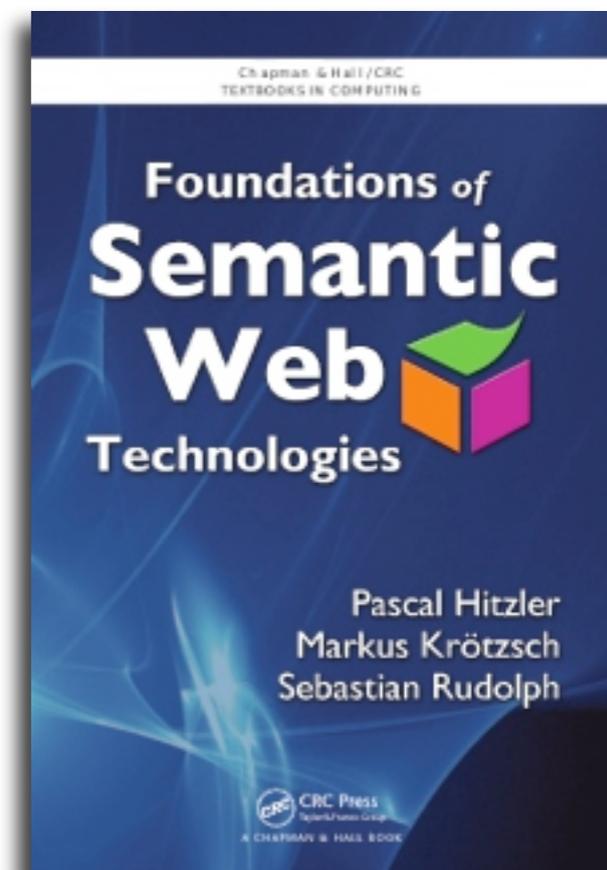
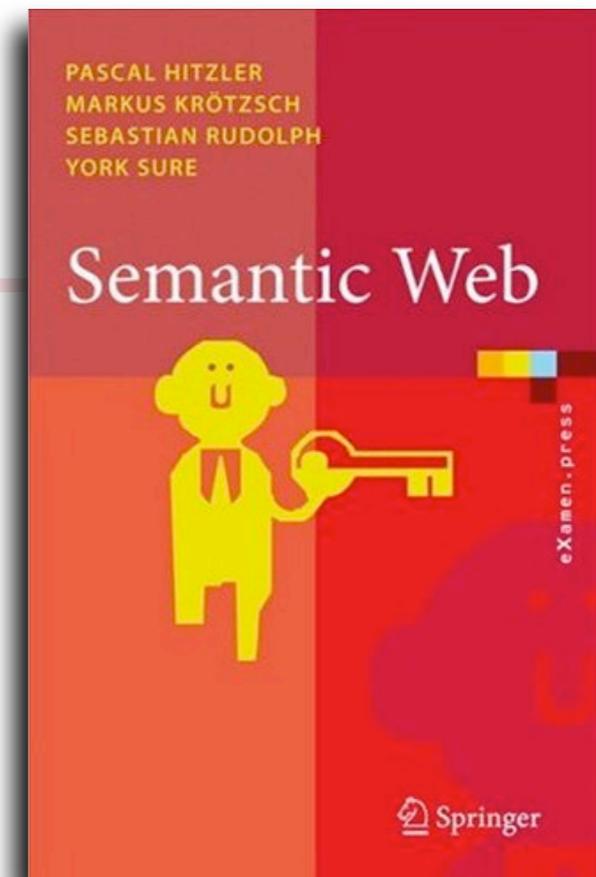
- Vorlesung: mittwochs, 11:30 - 13:00
- Übung: montags, 9:45 - 11:15
- Ort: Raum 231, Geb. 11.40
- Webseite mit aktuellen Infos:

[http://www.semantic-web-grundlagen.de/wiki/SWebT1\\_WS10/11](http://www.semantic-web-grundlagen.de/wiki/SWebT1_WS10/11)



# LITERATUR

- Hitzler, Krötzsch, Rudolph, Sure  
"Semantic Web. Grundlagen."  
Springer-Verlag
- Hitzler, Krötzsch, Rudolph  
"Foundations of Semantic Web  
Technologies"  
CRC Press



# AGENDA



- Vorstellung der Dozenten
- Organisatorisches zur Vorlesung
- **Was ist das "Semantic Web"?**
- "Semantic Karlsruhe"

# DAS WEB



Das Web flankiert den Übergang von der Industrie- zur Informationsgesellschaft und bietet die Infrastruktur für eine neue Qualität des Umgangs mit Information hinsichtlich Beschaffung wie auch Bereitstellung.



- hohe Verfügbarkeit
- hohe Aktualität
- geringe Kosten

## Kommerzialisierung in allen Größenordnungen



**ebay** Einloggen oder Neu anmelden

Kategorien ▾ Motors Express Shops

zurück Kategorie: Computer > Apple > MacBook / MacBook Pro > MacBook Pro 15"

### Apple Macbook Pro 15" 2,33 GHZ!!!! glossy

Bieter oder Verkäufer dieses Artikels? [Einloggen](#) zur Statusabfrage

 [Größeres Bild](#)

Aktuelles Gebot: **EUR 1.450,00**

Ihr Maximalgebot: EUR  [Bieten >](#)  
(Geben Sie mindestens EUR 1.460,00 ein)

Angebotsende: **54 Minuten 18 Sekunden**  
(23.10.07 17:48:17 MESZ)

Versandkosten: **EUR 12,00**  
Versicherter Versand  
Service nach: [Deutschland](#)

Versand nach: Weltweit  
Artikelstandort: Hamburg, Deutschland  
Übersicht: [36 Gebot\(e\)](#)  
Höchstbietender: [m\\*\\*\\*](#) (23 ⭐) 

Weitere Möglichkeiten: [Diesen Artikel beobachten](#)

Lassen Sie sich benachrichtigen per [Instant Messenger](#)  
[An einen Freund senden](#)

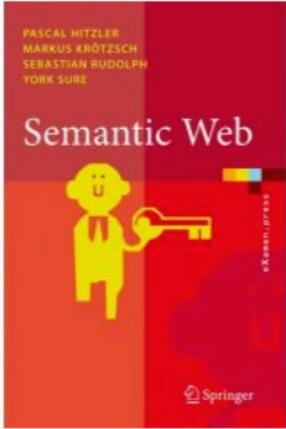
Angebots- und Zahlungsdetails: [Anzeigen](#)

amazon.de  WUNSCHZETTEL | [MEIN KONTO](#) | [HILFE](#) | [IMPRESSUM](#)

HOME MEIN SHOP **BÜCHER** ENGLISH BOOKS ELEKTRONIK & FOTO MUSIK DVD KAUFEN & LEIHEN SOFTWARE GAMES KÜCHE, HAUS & GÄRTEN SPIELWAREN & KINDERWELT SPORT & FREIZEIT UHREN BABY **NEU** SCHUHE & HANDTASCHEN

ERWEITERTE SUCHE | STÖßERN | BESTSELLER | NEUHEITEN | Hörbücher | Taschenbücher | Fachbücher | Preis-Hits | Bücher verkaufen

Suche



**Semantic Web. Grundlagen (eXamen.press) (Taschenbuch)**  
von [Pascal Hitzler](#) (Autor), [Markus Krötzsch](#) (Autor), [Sebastian Rudolph](#) (Autor), [York Sure](#) (Autor)

**Preis: EUR 24,95** Kostenlose Lieferung. [Siehe Details.](#)

**Verfügbarkeit:** Dieser Artikel ist noch nicht erschienen. Reservieren Sie sich Ihr Exemplar jetzt und Sie erhalten es pünktlich zum Erscheinungstermin. Verkauf und Versand durch **Amazon.de**. Geschenkverpackung verfügbar. Zustellung durch **DHL**.

**Noch 4 Tage** bis zum Erscheinungstermin von [Harry Potter Band 7](#). Sichern Sie sich jetzt [Ihr Exemplar mit Liefergarantie -- sonst geschenkt!](#)

**Preis: EUR 24,95**  
Vorbestellbar  
Verkauf und Versand durch **Amazon.de**

Menge:

[Jetzt vorbestellen](#)

oder

[Loggen Sie sich ein](#), um 1-Click® einzuschalten.

[Auf meinen Wunschzettel](#)

[Auf die Hochzeitsliste](#)

[Einem Freund weitersagen](#)

[Größeres Bild](#)

[Verleger: So können Kunden in diesem Buch suchen.](#)

[Bewerten](#)

[Frage](#)

[Zu meinen bevorzugten Verkäufern hinzufügen](#)

[Andere Artikel des Verkäufers](#)

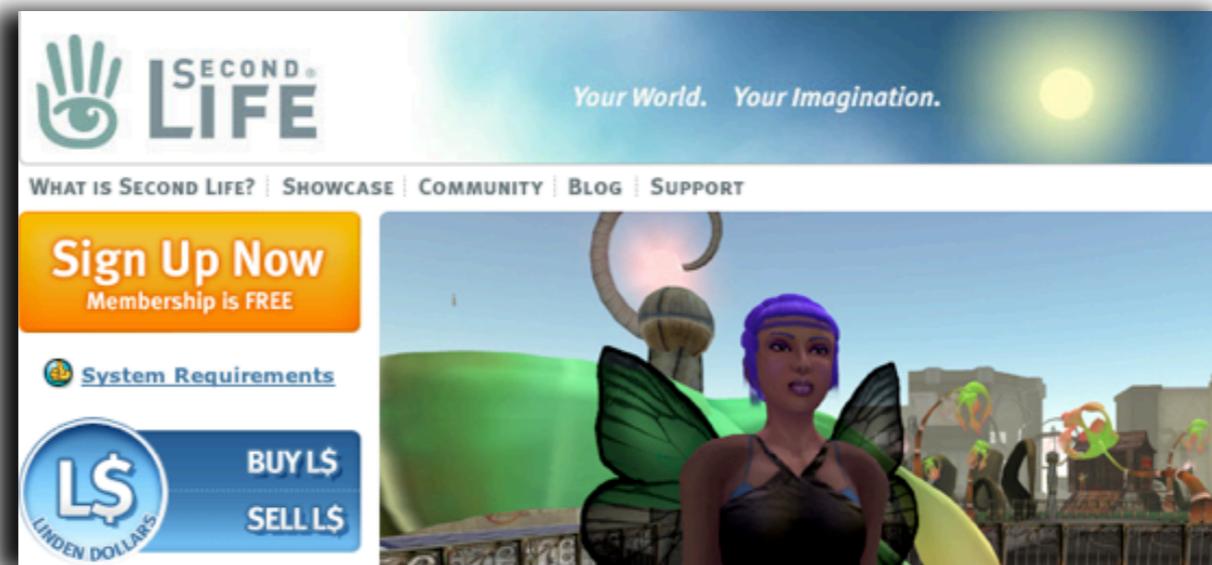
### Sicher kaufen

1. **Sehen Sie sich das Bewertungsprofil des Verkäufers an**  
Bewertungspunkte: 29 | 100% Positiv  
[Bewertungskommentare lesen](#)
2. **Informieren Sie sich über den Käuferschutz**  
Lesen Sie unsere [Tipps zum sicheren Kauf](#)

# DAS WEB

weitere Lebensbereiche werden "webisiert":

- Behörden, Verwaltung (eGovernment)
- Ausbildung (eLearning, eEducation)
- Sozialkontakte (Social-Networking-Plattformen, Partnerbörsen)
- Alltag?



# WARUM SEMANTIC WEB?



## Exkurs: Syntax vs. Semantik

- **Syntax**  
(von grch. συνταξις – *Zusammenstellung, Satzbau*) steht für die (normative) Struktur von Daten, d.h. sie charakterisiert, was "wohlgeformte" Daten sind.
- **Semantik**  
(grch. σημαντικός – *zum Zeichen gehörend*) steht für die Bedeutung von Daten, d.h. sie charakterisiert beispielsweise, welche inhaltliche Schlussfolgerungen sich ziehen lassen.

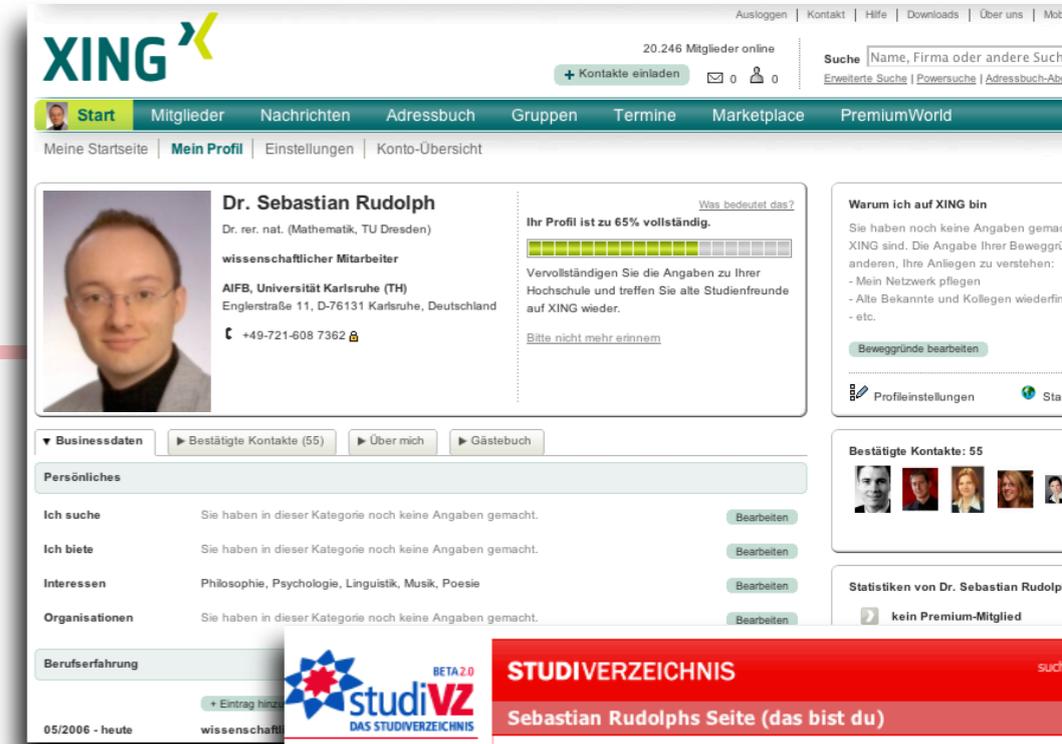
$4+)= ($   
syntaktisch falsch  
--

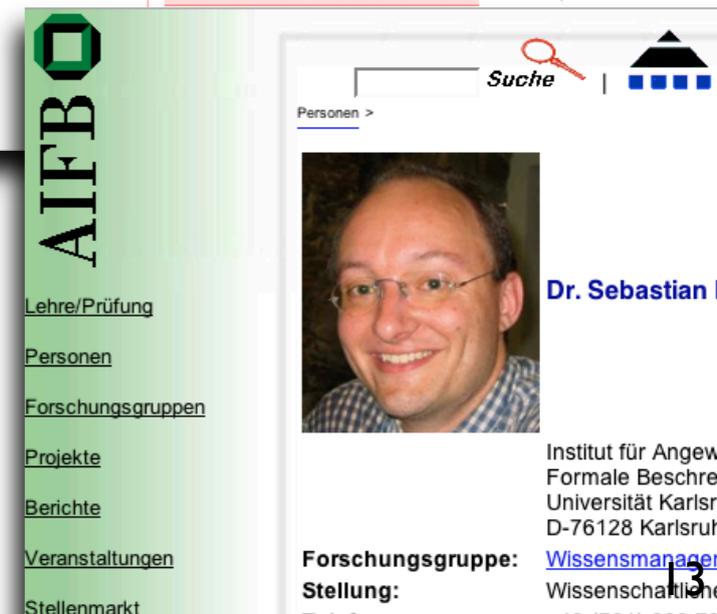
$3+4=12$   
syntaktisch richtig  
semantisch falsch

$3+4=7$   
syntaktisch richtig  
semantisch richtig

# PROBLEME DES WEB

- Fülle an Informationen
- ausgerichtet auf Menschen als Endnutzer
  - Erfassen der Bedeutung einer Webseite
  - Unabhängig von konkreter Repräsentation
  - Bilden von Zusammenhängen





# PROBLEME DES WEB



- **Lokalisierung** von Information problematisch
- heutige Suchmaschinen gut, aber stichwortbasiert
- wünschenswert:  
inhaltliche,  
*semantische Suche*



# PROBLEME DES WEB

- **Heterogenität** der vorhandenen Information auf verschiedensten Ebenen:
  - Zeichenkodierung (z.B. ASCII vs. Unicode)
  - verwendete natürliche Sprachen
  - Anordnung von Information auf Webseiten
- *Informationsintegration*

## Semantic Web Technologies I & II: Intelligente Systeme im WWW

Winter 2007/08

**Dozenten:** PD Dr. Pascal Hitzler,  
Dr. Sebastian Rudolph

**Betreuer:** M.Sc. Markus Krötzsch

**Umfang:**

2+1 SWS (Vorlesung+Übung),  
4.5 Leistungspunkte

**Zeit & Ort:**

Vorlesung: wöchentlich Mittwoch 11:30 bis 13:00



로그인 /

문서 토론 편집 0

## 시맨틱 웹

위키백과 — 우리 모두의 백과사전

**시맨틱 웹**(Semantic Web)은 인터넷과 같은 분산환경에서 리소스 각종 화일, 서비스 등)에 대한 사이의 관계-의미 정보(Semantic Web)가 처리할 수 있는

**위키백과**  
우리 모두의 백과사전 둘러보기

- 대문
- 사용자 모임
- 요즘 화제

# PROBLEME DES WEB

- **implizites Wissen**, d.h. Informationen, sind nicht explizit spezifiziert, folgen aber aus der Kombination gegebener Daten
- formallogische Methoden erforderlichlich
- *automatisches Schlussfolgern*



# PROBLEME DES WEB

## Lösungsansätze:

- I. Ad hoc: Verwendung von KI-Methoden zur Auswertung bestehender unstrukturierter Informationen im Web
- II. A priori: Strukturierung der Web-Informationen zur Erleichterung der automatisierten Auswertung:  
→ **Semantic Web**

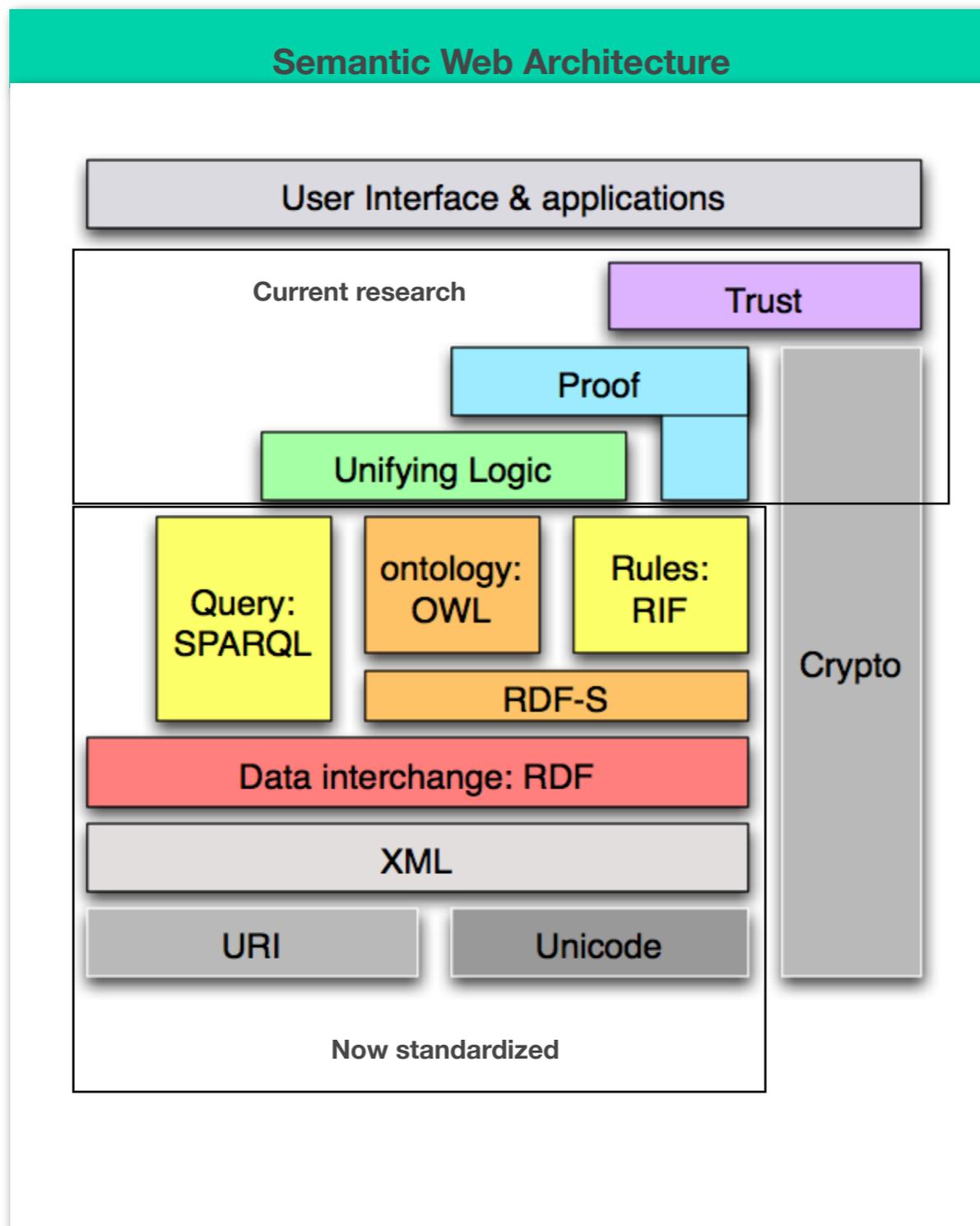
# SEMANTIC WEB - ZUTATEN



Zwei essentielle Voraussetzungen zur Realisierung:

1. offene Standards zur Beschreibung von Informationen
  - klar definiert
  - flexibel
  - erweiterbar
2. Methoden zur Gewinnung von Informationen aus derlei Beschreibungen

# SEMANTIC WEB - STANDARDS



# AGENDA



- Vorstellung der Dozenten
- Organisatorisches zur Vorlesung
- Was ist das "Semantic Web"?
- **"Semantic Karlsruhe"**

# „Semantic Karlsruhe“: Semantic Technologies in Karlsruhe

Applied Research & Technology Transfer

Innovative Application Scenarios

Basic Research  
Applied Research

Applied Research

Product Development & Consulting



**Semantic Web Infrastructure**  
**Ontology Management**  
**Social Software & Web 2.0**  
**Cloud Computing**  
**Complex Event Processing**  
**Data, Text & Web Mining**

**Knowledge Management**  
**eBusiness**  
**eScience**  
**Service Engineering & Management**

# Who are we? ...

## Semantic Technologies Research Group

AIFB

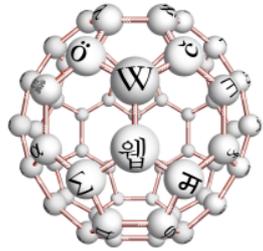
FZI



+ 3 external Ph.D. students at SAP Research

+ ca. 40 people at Ontoprise

# Selected Projects, Partners, and Customers



Semantic MediaWiki



Bundesministerium für Wirtschaft und Technologie



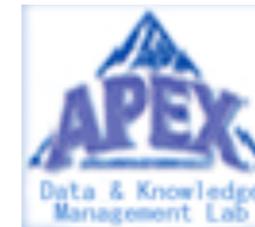
KompetenzNetzwerk Wissensmanagement



Deutsche Forschungsgemeinschaft DFG



Bundesministerium für Bildung und Forschung



# Interesse an Lehrveranstaltungen?

## ■ Wintersemester

- Angewandte Informatik I
- Wissensmanagement
- Knowledge Discovery
- Semantic Web Technologies I

## ■ Sommersemester

- Grundlagen der Informatik I
- Semantic Web Technologies II

## ■ Seminare/Praktika, z.B.



- Linked Open Data basierte Web 3.0 Anwendungen
- Ambient Assisted Living
- Service Science, Management und Engineering
- Collective Intelligence
- AI for Decision Making and Game Playing Computers
- Graphbasierte Wissensrepräsentation
- Developing Business Models for the Semantic Web
- Knowledge Discovery and Text Mining

## ■ Lehrangebot gilt für alle Studienrichtungen: Informationswirtschaft, Wirtschaftsingenieurwesen, technische VWL, Wirtschaftsmathematik

## ■ Wahlmöglichkeiten entsprechend Prüfungsordnung (Diplom bzw. Bachelor/Master)

# Interesse an Mitarbeit?

- **AIFB-WBS [1]** und **FZI-IPE [2]**: Mitarbeit an praxisrelevanten internationalen Forschungsprojekten im Rahmen von
  - HiWi-Tätigkeiten
  - Seminaren / Praktika
  - Studien- / Diplomarbeiten
- ggf. in Zusammenarbeit mit ontoprise GmbH
- und ausgewählten Partnern, z.B.
  - Google
  - Accenture
  - British Telecom
  - Wikimedia
  - Telefonica
  - SAP

**Semantic Web Infrastructure**  
**Ontology Management**  
**Social Software & Web 2.0**  
**Cloud Computing**  
**Complex Event Processing**  
**Data, Text & Web Mining**  
**Knowledge Management**  
**eBusiness, eScience**  
**Service Engineering**  
**& Management**

[1] <http://www.aifb.kit.edu/web/Wissensmanagement>

[2] <http://www.fzi.de/ipe/>

# XML UND URIs

Dr. Sebastian Rudolph

Einleitung und XML

Einführung in RDF

RDF Schema

Logik - Grundlagen

Semantik von RDF(S)

OWL - Syntax und Intuition

OWL - Semantik und Reasoning

OWL 2

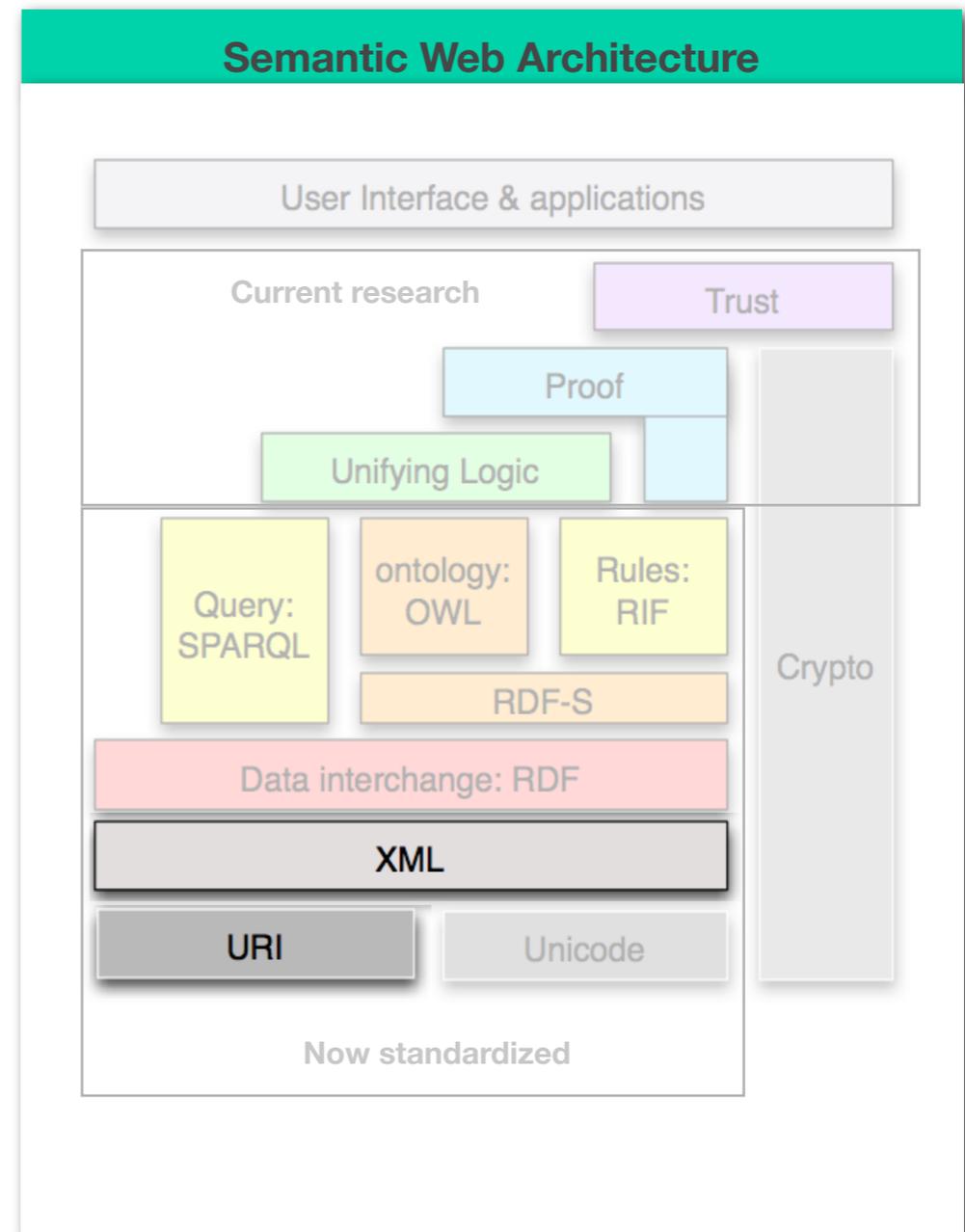
SPARQL - Syntax und Intuition

Konjunktive Anfragen / Einführung Regelsprachen

Regeln für OWL

Ontology Engineering

Semantic Web - Anwendungen



# AGENDA



- XML - Motivation/Idee
- XML - Syntax
- URIs
- Namensräume

# AGENDA



- XML - Motivation/Idee
- XML - Syntax
- URIs
- Namensräume

# ANNOTATION MIT MARKUPSPRACHEN



- Grundidee des Markup: versehen von (unstrukturiertem) Text mit zusätzlicher Information (bzw. Struktur)
- synonym: *auszeichnen*, auch: *annotieren* von Text
- Text = Daten  
Zusatzinformation = *Metadaten*

# ANNOTATION MIT MARKUPSPRACHEN

- häufige Markup-Strategie: Einschließen des zu annotierenden Textes in sogenannte *tags* (engl.: Etikett, Schild):  
  

<Tag-Bezeichner>...	Text	...</Tag-Bezeichner>
<i>öffnendes Tag</i>		<i>schließendes Tag</i>
- Zusatzinformation wird von verarbeitenden Programmen gelesen und interpretiert

# ANNOTATION MIT MARKUPSPRACHEN



- prominentestes Beispiel: HTML  
tags kodieren Darstellungsinformationen:  
<i>Dieses Buch</i> hat den Titel <b>Semantic Web Grundlagen</b>.
- Darstellung im Browser:  
*Dieses Buch* hat den Titel **Semantic Web Grundlagen**.
- Strategie auch geeignet zur inhaltlichen Annotation, z.B.:  
<Buch>Dieses Buch</Buch> hat den Titel <Titel>Semantic Web Grundlagen</Titel>.

# ANNOTATION MIT MARKUPSPRACHEN

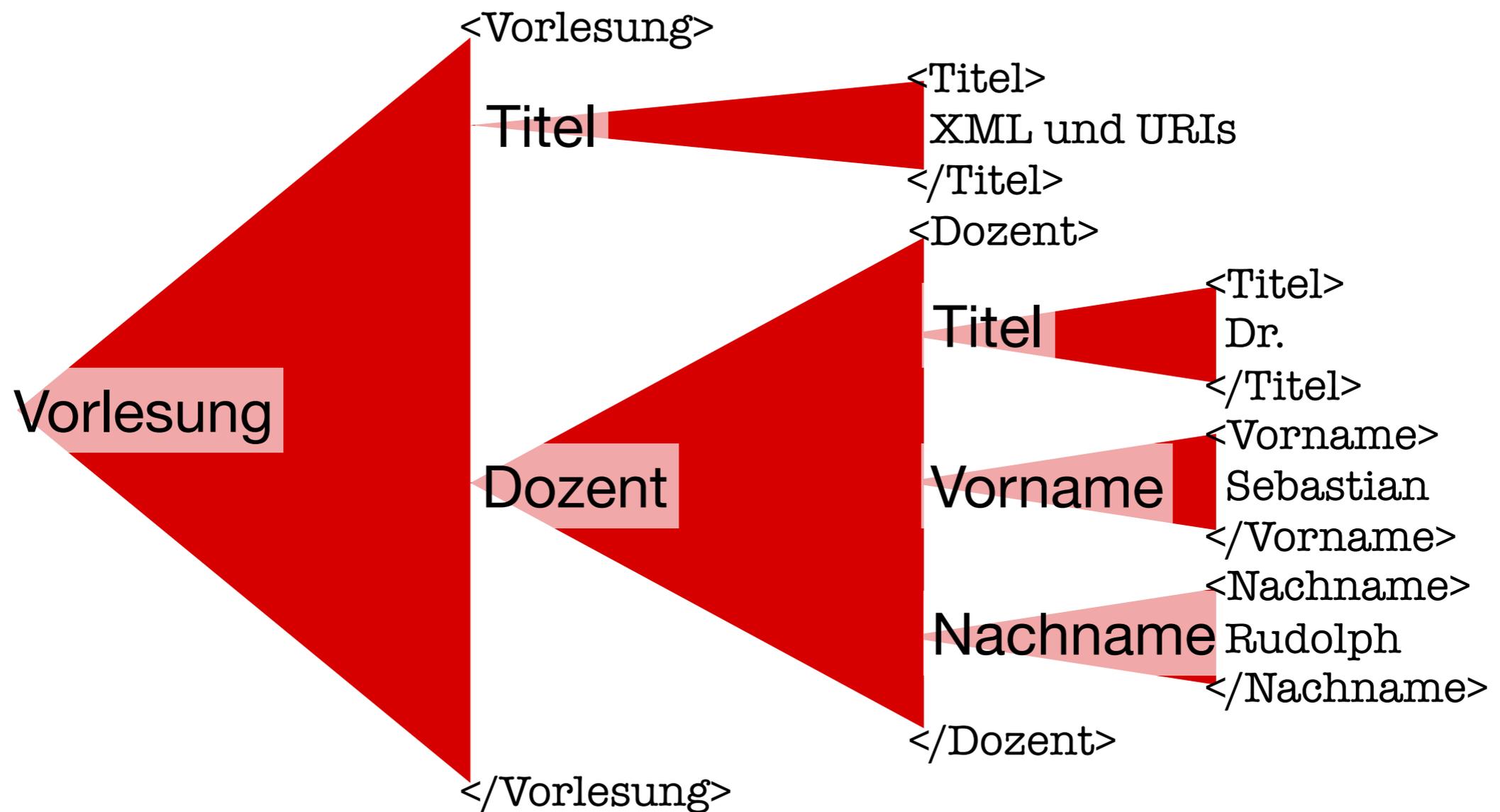
- Verschachtelung von Tags erlaubt:

```

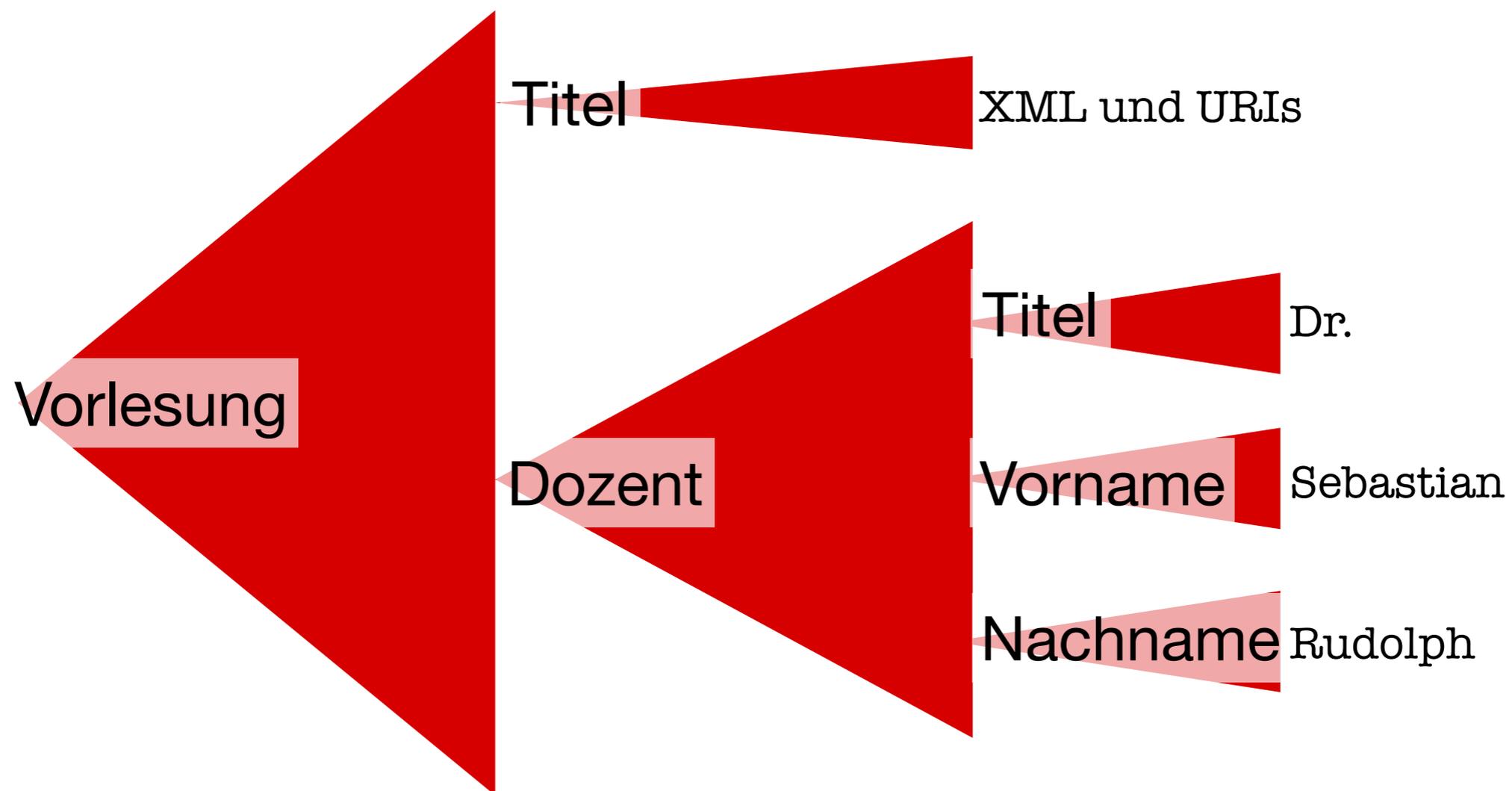
<Vorlesung>
  <Titel>
    XML und URIs
  </Titel>
  <Dozent>
    <Titel>
      Dr.
    </Titel>
    <Vorname>
      Sebastian
    </Vorname>
    <Nachname>
      Rudolph
    </Nachname>
  </Dozent>
</Vorlesung>
  <Titel>
    XML und URIs
  </Titel>
  <Dozent>
    <Titel>
      Dr.
    </Titel>
    <Vorname>
      Sebastian
    </Vorname>
    <Nachname>
      Rudolph
    </Nachname>
  </Dozent>

```

# ANNOTATION MIT MARKUPSPRACHEN



## Baumstruktur



# AGENDA



- XML - Motivation/Idee
- **XML - Syntax**
- URIs
- Namensräume

# XML



- eXtensible Markup Language
- Ursprung: strukturierter Text ( $\text{HTML4.0} \in \text{XML} \subset \text{SGML}$ )
- Web-Standard (W3C) zum Datenaustausch:
  - Ein- und Ausgabedaten von Anwendungen können mittels XML beschrieben werden
  - Industrie muss sich nur noch auf standardisierte Beschreibung (= Vokabular) einigen
- Komplementärsprache zu HTML:
  - HTML beschreibt die Präsentation
  - XML beschreibt den Inhalt
- Datenbank-Sichtweise: XML als Datenmodell für semistrukturierte Daten

# XML-SYNTAX (I) PRÄAMBEL



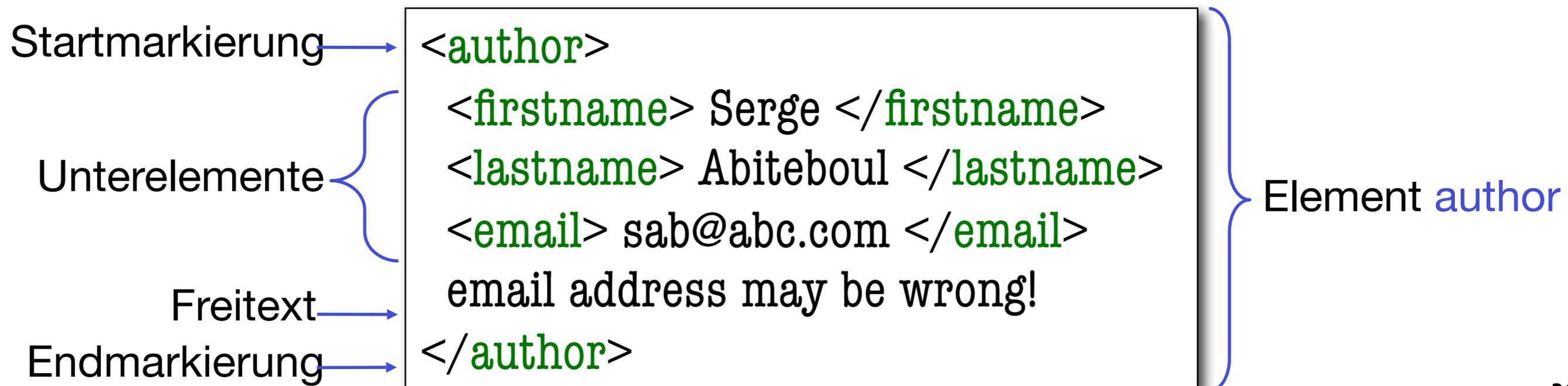
- XML-Dokument ist Textdokument
- beginnt mit Deklaration, die Versionsnummer des verwendeten Standards und optional die Zeichenkodierung enthält, z.B.:

```
<?xml version="1.0" encoding="utf-8"?>
```

# XML-SYNTAX (2) – XML-ELEMENT



- XML-Element (engl. element):
  - Beschreibung eines Objekts, die durch passende Markierungen (tags) geklammert ist
  - Inhalt eines Elements: Text und/oder weitere Elemente (beliebige Schachtelung möglich)
  - Leere Elemente: `<year></year>` kurz: `<year/>`
  - "äußerstes" Element heißt Wurzelement (nur eines pro XML-Dokument)



# XML-SYNTAX (3) – XML-ATTRIBUT

AIFB 

- XML-Attribut (engl. attribute):
  - Name-Zeichenkettenwert-Paar in Start- oder selbstschließendem Tag
  - Assoziiert mit einem Element
  - Alternative Möglichkeit, Daten zu beschreiben

Attribut **email**



```
<author email="sab@abc.com">  
  <firstname> Serge </firstname>  
  <lastname> Abiteboul </lastname>  
</author>
```

Weitere denkbare Beschreibung derselben Daten:

```
<author firstname="Serge" lastname="Abiteboul" email="sab@abc.com"/>
```

# AGENDA



- XML - Motivation/Idee
- XML - Syntax
- URIs
- Namensräume

# URIs - IDEE



- URI = Uniform Resource Identifier
- dienen zur weltweit eindeutigen Bezeichnung von Ressourcen
- Ressource kann jedes Objekt sein, was (im Kontext der gegebenen Anwendung) eine klare Identität besitzt (z.B. Bücher, Orte, Menschen, Verlage, Beziehungen zwischen diesen Dingen, abstrakte Konzepte usw.)
- in bestimmten Domänen ähnliches bereits realisiert: ISBN für Bücher

# URIs - SYNTAX



- Erweiterung des URL-Konzeptes; nicht jede URI bezeichnet aber ein Webdokument (umgekehrt wird als URI für Webdokumente häufig deren URL verwendet)
- Beginnt mit dem sogenannten URI-Schema das durch ":" vom nachfolgenden Teil getrennt ist (z.B.: http, ftp, mailto)
- häufig hierarchisch aufgebaut

# SELBSTDEFINIERTER URIs



- nötig, wenn für eine Ressource (noch) keine URI existiert (bzw. bekannt ist)
- Strategie zur Vermeidung von (ungewollten) Überschneidungen:  
Nutzung von `http`-URIs einer eigenen Webseite
- ermöglicht auch Ablegen einer Dokumentation zur URI an dieser Stelle

# BESCHREIBENDES VS. BESCHRIEBENES



- Trennung von URI für Ressource und deren Dokumentation durch URI-Referenzen (durch "#" angehängte Fragmente) oder content negotiation
- z.B.: als URI für Shakespeares "Othello"  
<http://de.wikipedia.org/wiki/Othello>  
nicht geeignet, besser  
<http://de.wikipedia.org/wiki/Othello#URI>

# AGENDA



- XML - Motivation/Idee
- XML - Syntax
- URIs
- Namensräume

# XML-NAMENSRAÜME: MOTIVATION



- XML-Dokumente besitzen Element- und Attributnamen (“Markup Vocabulary”) mit allgemeiner Gültigkeit
- Eine XML-Anwendung basiert auf allgemeiner Interpretation dieser Namen
- Ein XML-Dokument soll Markup-Vokabular aus mehreren ‘Dictionaries’ enthalten können. (Erinnerung: XML-Dokument muss keine DTD haben.)
- Namespaces zur Vermeidung von Namenskonflikten.

# XML-NAMENSÄÄUME



- XML Namespaces sind ähnhlich zu Modul-Konzepten in Programmiersprachen
- Disambiguierung von Tag-Namen durch Verwendung unterschiedlicher “Präfixe”
- Ein Präfix wird vom lokalen Namen separiert durch ein “:”, so entstehen prefix:name Tags

# NAMENSRAUM-BINDUNGEN



- Präfixe werden belegt mit Namespace URIs, indem ein Attribut `xmlns:prefix` bei dem relevanten Element oder einem seiner Vorgängerelemente eingefügt wird: `prefix:name1, ..., prefix:namen`
- Der Wert des `xmlns:prefix`-Attributes ist eine URI, welche (für XML Schemata) auf eine Beschreibung auf eine Beschreibung der Namespace Syntax verweisen kann aber nicht muss
- Ein Element kann Bindings nutzen für mehrere (unterschiedliche) Namespaces durch Verwendung separater Attribute `xmlns:prefix1, ..., xmlns:prefixm`

# BEISPIEL: OHNE NAMENSÄÄUME



```
<Vorlesung>  
  <Titel>  
    XML und URIs  
  </Titel>  
  <Dozent>  
    <Titel>  
      Dr.  
    </Titel>  
    <Vorname>  
      Sebastian  
    </Vorname>  
    <Nachname>  
      Rudolph  
    </Nachname>  
  </Dozent>  
</Vorlesung>
```

Titel ist  
mehrdeutig  
verwendeter  
Tagname

# ZWEI VERSCHIEDENE NAMENSRÄUME



```
<Vorlesung xmlns:lv="http://www.semantic-web-Grundlagen/Lehrveranstaltungen"
           xmlns:person:="http://www.semantic-web-Grundlagen/Person" >
  <lv:Titel>
    XML und URIs
  </lv:Titel>
  <lv:Dozent>
    <person:Titel>
      Dr.
    </person:Titel>
    <person:Vorname>
      Sebastian
    </person:Vorname>
    <person:Nachname>
      Rudolph
    </person:Nachname>
  </lv:Dozent>
</lv:Vorlesung>
```

Titel wurde  
disambiguiert  
durch  
Verwendung der  
Prefixe lv und  
person