

SEMANTIC WEB TECHNOLOGIES I

Lehrveranstaltung im WS09/10

M.Sc. Markus Krötzsch

Dr. Sebastian Rudolph

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



M.Sc. Markus Krötzsch

- Since Mar. 2005: PhD student and research assistant in Prof. Rudi Studer's group at Institute AIFB of Universität Karlsruhe (TH)
- Feb. 2005: Master of Science in Computational Logic, degree 1.0
- Jun.–Aug. 2004: visit at CWRU Cleveland
- 2003–2005: Study of Computational Logic at the ICCL of TU Dresden
- 2000–2003: Study of Computer Science at Technische Universität Dresden



Themen:

Wissensrepräsentation
Semantic Web
Semantische Wikis
Logik
Komplexitätstheorie

VORSTELLUNG DER DOZENTEN



Dr. Sebastian Rudolph

- 1995 – 2000 Studium Lehramt
Mathematik/Physik/Informatik,
TU Dresden
- 2000 – 2003 Stipendiat im Graduiertenkolleg 334,
TU Dresden
- 2003 – 2005 wissenschaftlicher Mitarbeiter an der
Professur für die Psychologie des
Lehrens und Lernens, TU Dresden
- 2006 Promotion in Mathematik, TU Dresden
- seit 2006 als Postdoc am AIFB



Themen:

formale Aspekte der
Wissensverarbeitung
Logik
Komplexitätstheorie
Formale Begriffsanalyse
NLP
E-Learning

AGENDA



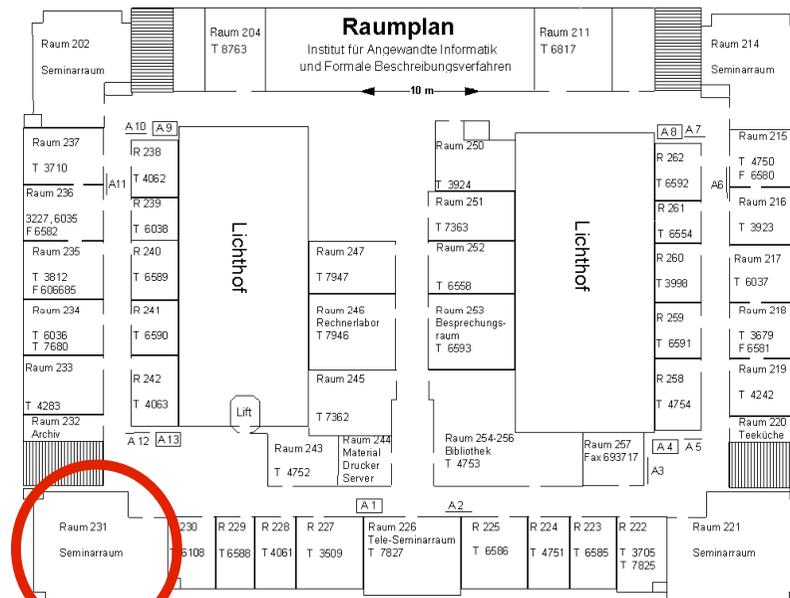
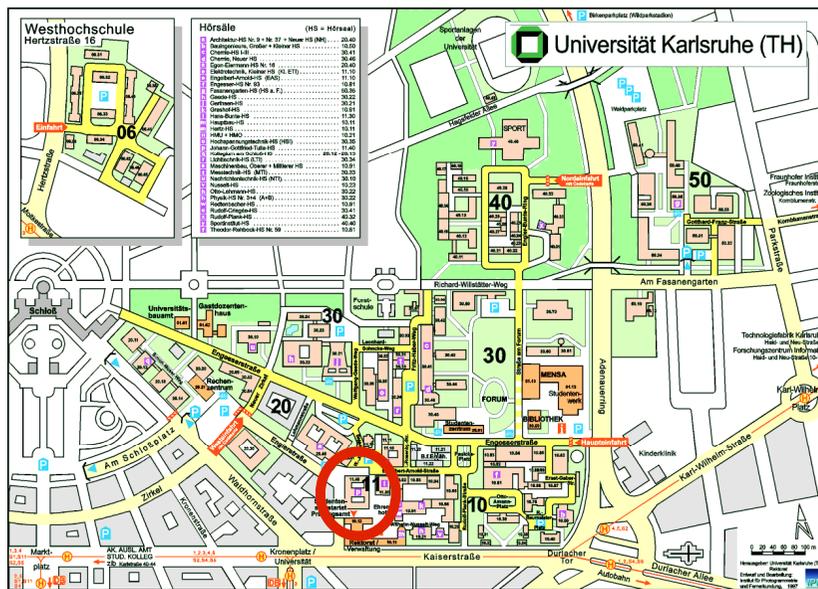
- Vorstellung der Dozenten
- **Organisatorisches zur Vorlesung**
- Was ist das "Semantic Web"?
- "Semantic Karlsruhe"

ORGANISATORISCHES: ZEIT UND ORT



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

http://www.semantic-web-grundlagen.de/wiki/SWebT1_WS09/10

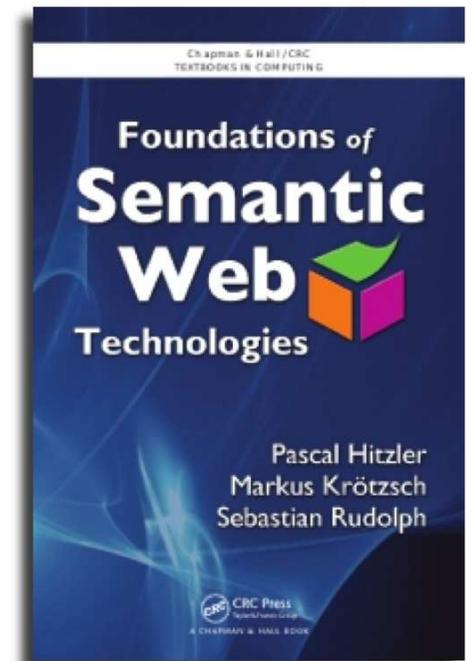
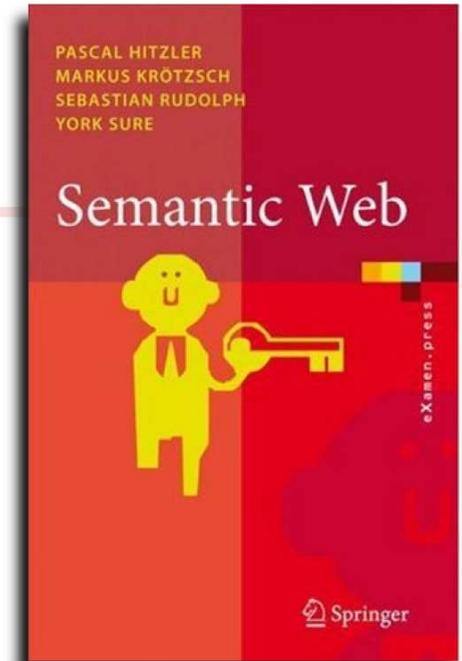


ORGANISATORISCHES: INHALT



Titel	Termin	Dozent
Einleitung und XML	21. Oktober 2009	Sebastian Rudolph
Einführung in RDF	28. Oktober 2009	Sebastian Rudolph
RDF Schema	4. November 2009	Sebastian Rudolph
Logik – Grundlagen	11. November 2009	Markus Krötzsch
Semantik von RDF(S)	18. November 2009	Markus Krötzsch
OWL – Syntax und Intuition	25. November 2009	Sebastian Rudolph
OWL – Semantik und Reasoning	2. Dezember 2009	Sebastian Rudolph
OWL 2	9. Dezember 2009	Markus Krötzsch
SPARQL – Syntax und Intuition	16. Dezember 2009	Markus Krötzsch
Semantik von SPARQL	13. Januar 2010	Markus Krötzsch
Konjunktive Anfragen/Einführung Regelsprachen	20. Januar 2010	Markus Krötzsch
Regeln für OWL	27. Januar 2010	Markus Krötzsch
Ontology Engineering	3. Februar 2010	Sebastian Rudolph
Semantic Web – Anwendungen	10. Februar 2010	Sebastian Rudolph

- Hitzler, Krötzsch, Rudolph, Sure
"Sematic 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 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

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

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

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 Gebote](#)

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

HOME MEIN SHOP **BÜCHER** ENGLISH BOOKS ELEKTRONIK & FOTO MUSIK DVD KAUFEN & LEIHEN SOFT-WARE GAMES KÜCHE, HAUS & GÄRTEN SPIELWAREN & KINDERWELT SPORT & FREIZEIT UHREN BABY 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 **DPD**.

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](#)

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

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.

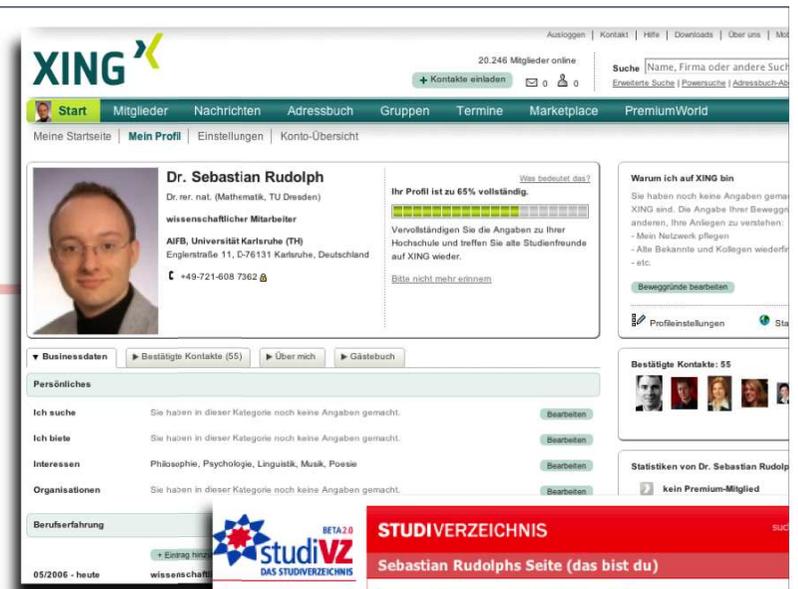
$4+)=($
syntaktisch falsch
--

$3+4=12$
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



- **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 - Intelligente Systeme im WWW

Winter 2007/08

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

Betreuer: M.Sc. Markus Krättsch

Umfang:

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

Zeit & Ort:

Vorlesung: wöchentlich Mittwoch 11:30 bis 13:00



PROBLEME DES WEB

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



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**

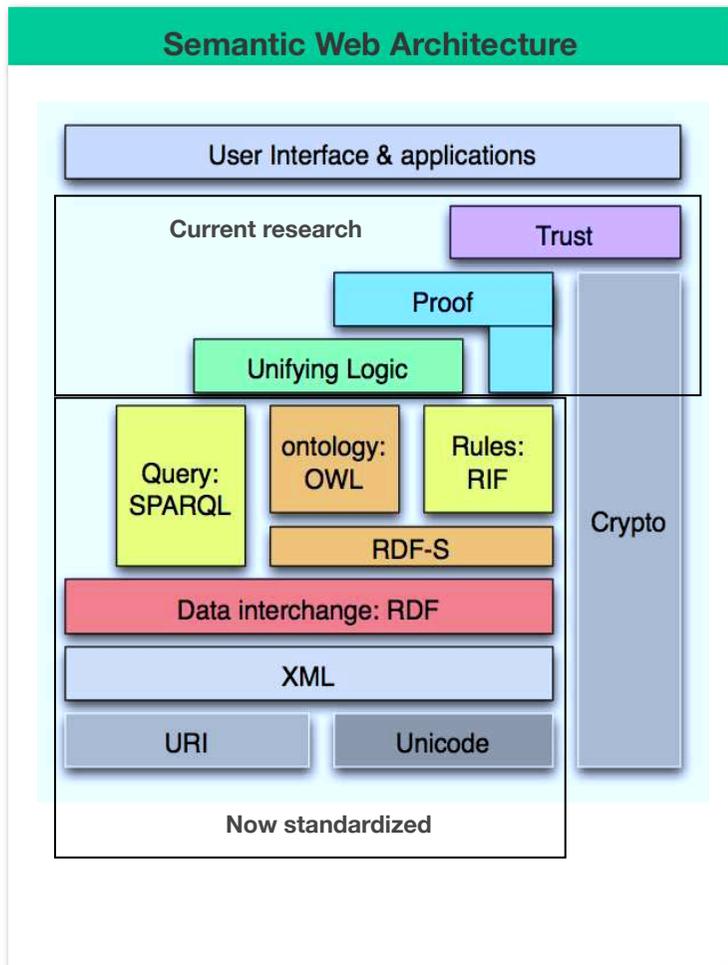
Zwei essentielle Voraussetzungen zur Realisierung:

1. offene Standards zur Beschreibung von Informationen

- klar definiert
- flexibel
- erweiterbar

2. Methoden zur Gewinnung von Informationen aus derlei Beschreibungen

SEMANTIC WEB - STANDARDS



Standardization Semantic Web

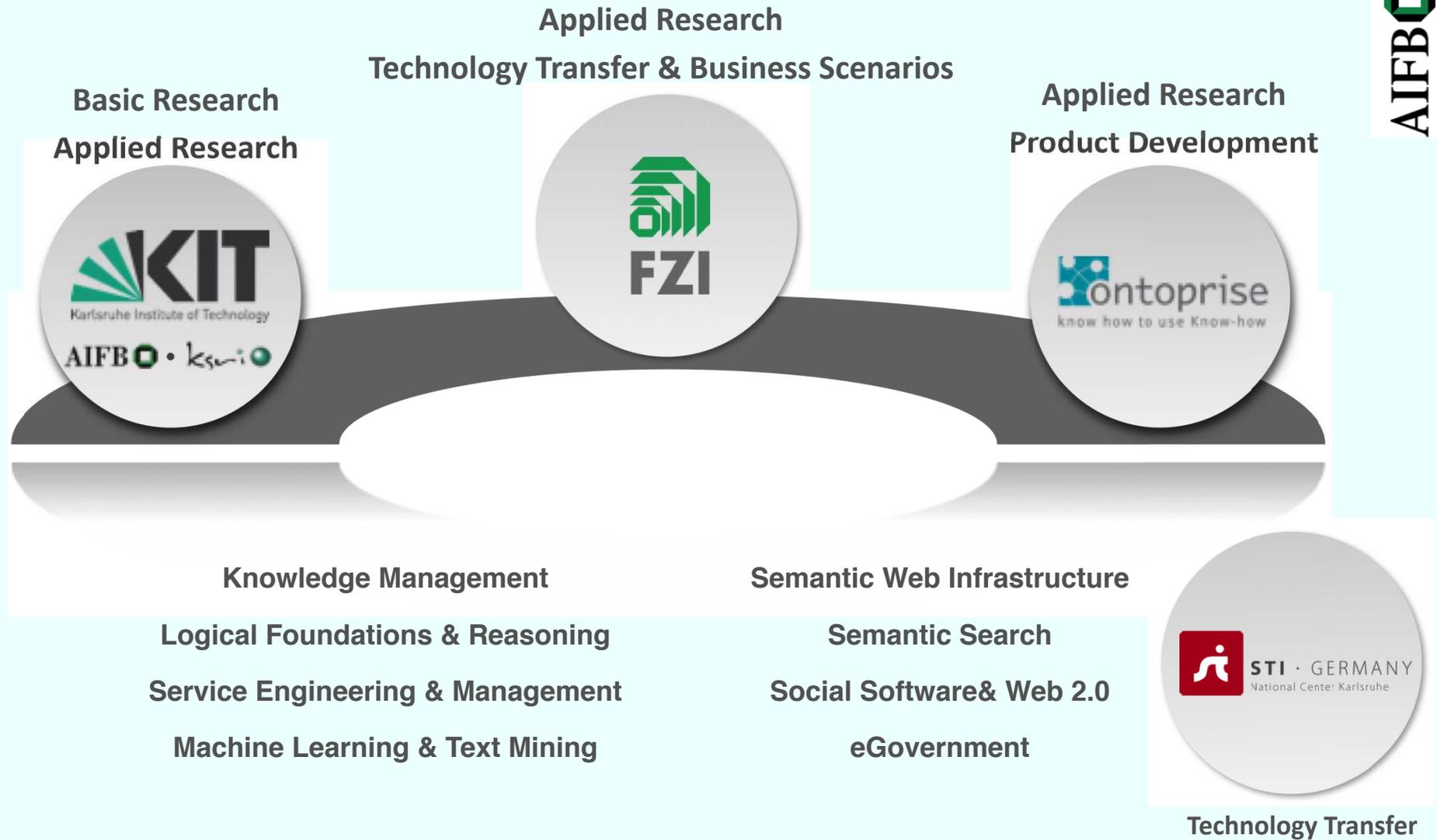
1994	<ul style="list-style-type: none"> • First public presentation of the Semantic Web idea
1998	<ul style="list-style-type: none"> • Start of standardization of data model (RDF) and a first ontology languages (RDFS) at W3C
2000	<ul style="list-style-type: none"> • Start of large research projects about ontologies in the US and Europe (DAML & Ontoknowledge)
2002	<ul style="list-style-type: none"> • Start of standardization of a new ontology language (OWL) based on research results
2004	<ul style="list-style-type: none"> • Finalization of the standard for data (RDF) and ontology (OWL)
2006	<ul style="list-style-type: none"> • Standardization of a query language (SPARQL, 6. April 2006) • Ongoing work on rule languages (SWRL, DL-safe rules, RIF) • Extension of OWL to OWL 1.1 / 2.0 • Ontology language of OMG based on UML (ODM)

AGENDA



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

Karlsruhe – Standort für semantische Technologien



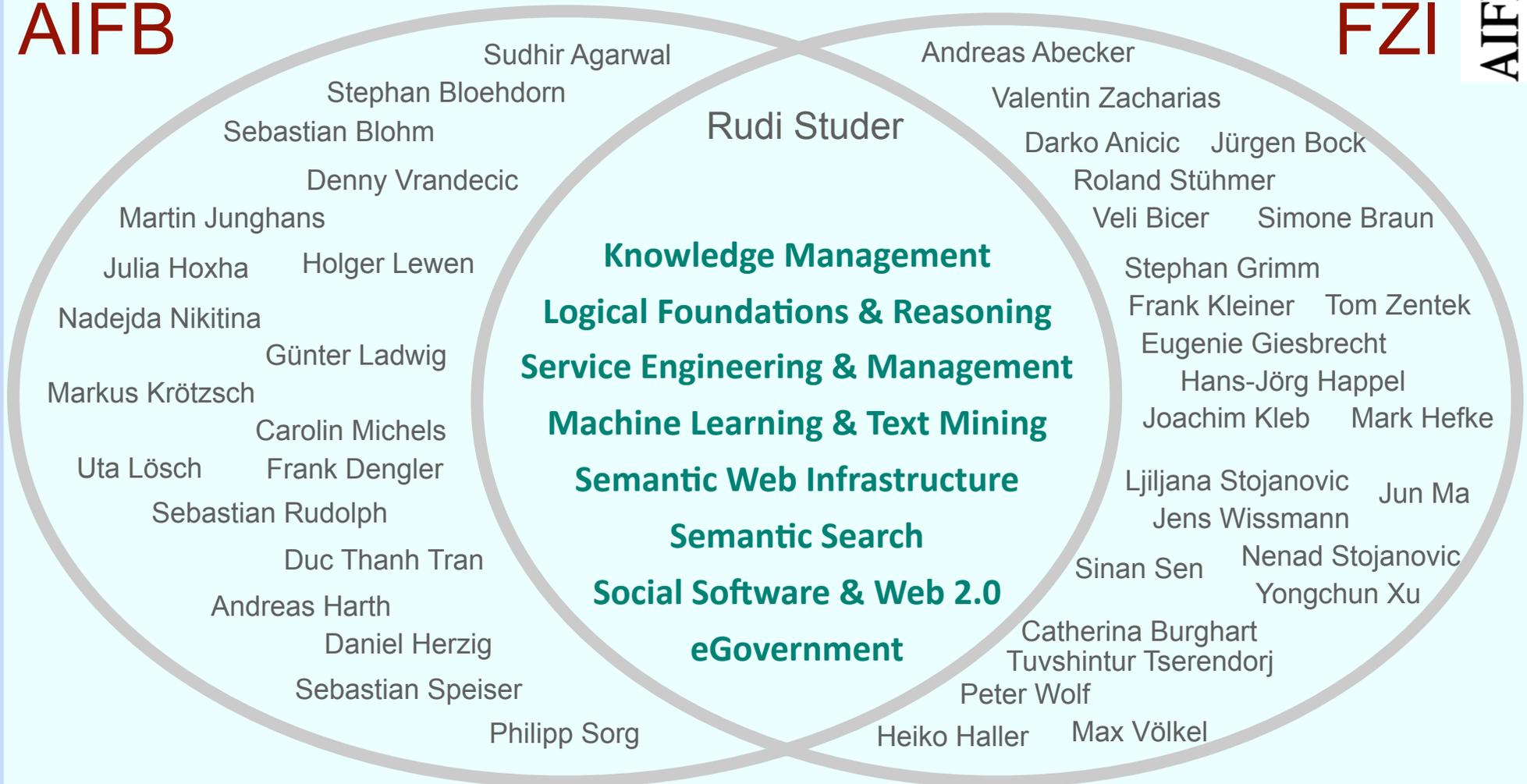
Who are we? ...

Semantic Technologies Research Group



AIFB

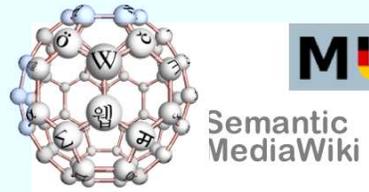
FZI



+ 3 external Ph.D. students at SAP Research

+ ca. 40 people at Ontoprise

Selected Projects, Partners and Customers



Interesse an Lehrveranstaltungen?

- **Wintersemester**
 - Angewandte Informatik I
 - Semantic Web Technologies I
 - Wissensmanagement
 - Knowledge Discovery
 - Seminar Wissensmanagement
 - Seminar/Praktikum Web Services
- **Sommersemester**
 - Semantic Web Technologies II
 - Service-oriented Computing II
 - Seminar/Praktikum Knowledge Discovery
 - Seminar Service Science, Management & Engineering
- Lehrangebot gilt für alle Studienrichtungen: Informationswirtschaft, Wirtschaftsingenieurwesen, technische VWL, Wirtschaftsmathematik
- Wahlmöglichkeiten entsprechend Prüfungsordnung (Diplom bzw. Bachelor/Master)

Interesse an Mitarbeit?

- **AIFB-WBS** und **FZI-IPE**: Mitarbeit an praxisrelevanten internationalen Forschungsprojekten im Rahmen von
 - HiWi-Tätigkeiten
 - Seminaren / Praktika
 - Studien- / Diplomarbeiten
 - ggf. in Zusammenarbeit mit ontoprise GmbH

Wissensmanagement	Semantic Web Infrastructure
Logische Grundlagen & Reasoning	Semantic Search
Service Engineering & Management	Social Software & Web 2.0
Machine Learning & Text Mining	eGovernment

<http://www.aifb.uni-karlsruhe.de/WBS>
<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

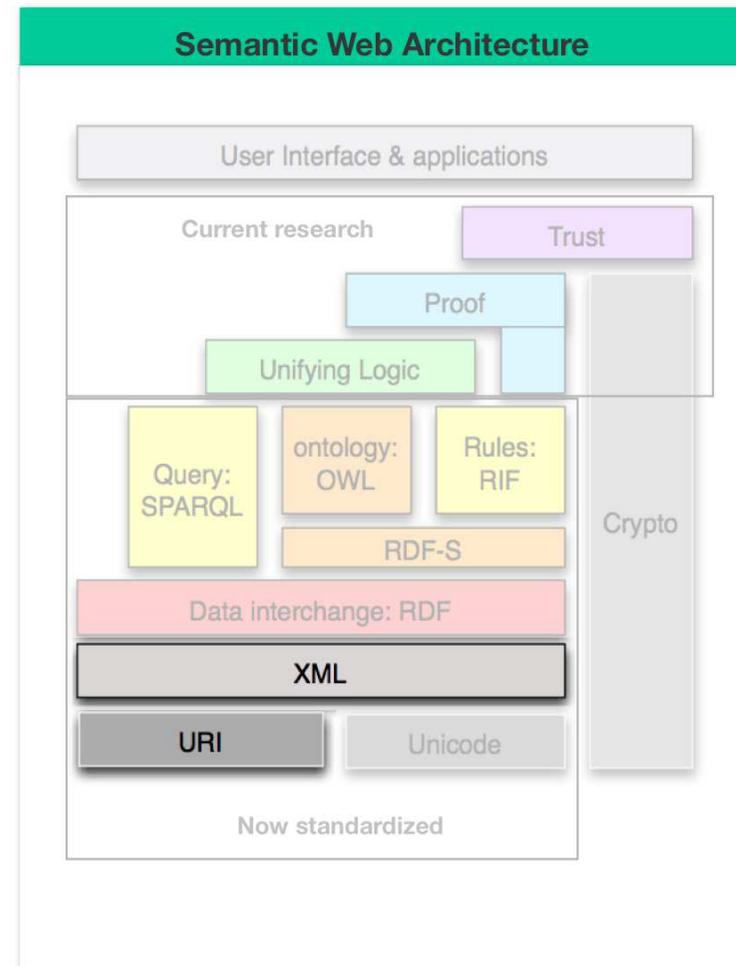
SPARQL - Syntax und Intuition

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>
öffnendes Tag *schließendes Tag*

- Zusatzinformation wird von verarbeitenden Programmen gelesen und interpretiert

ANNOTATION MIT MARKUPSPRACHEN



- prominentestes Beispiel: HTML
tags kodieren Darstellungsinformationen:
<i>Dieses Buch</i> hat den Titel Semantic Web Grundlagen.
- 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>.

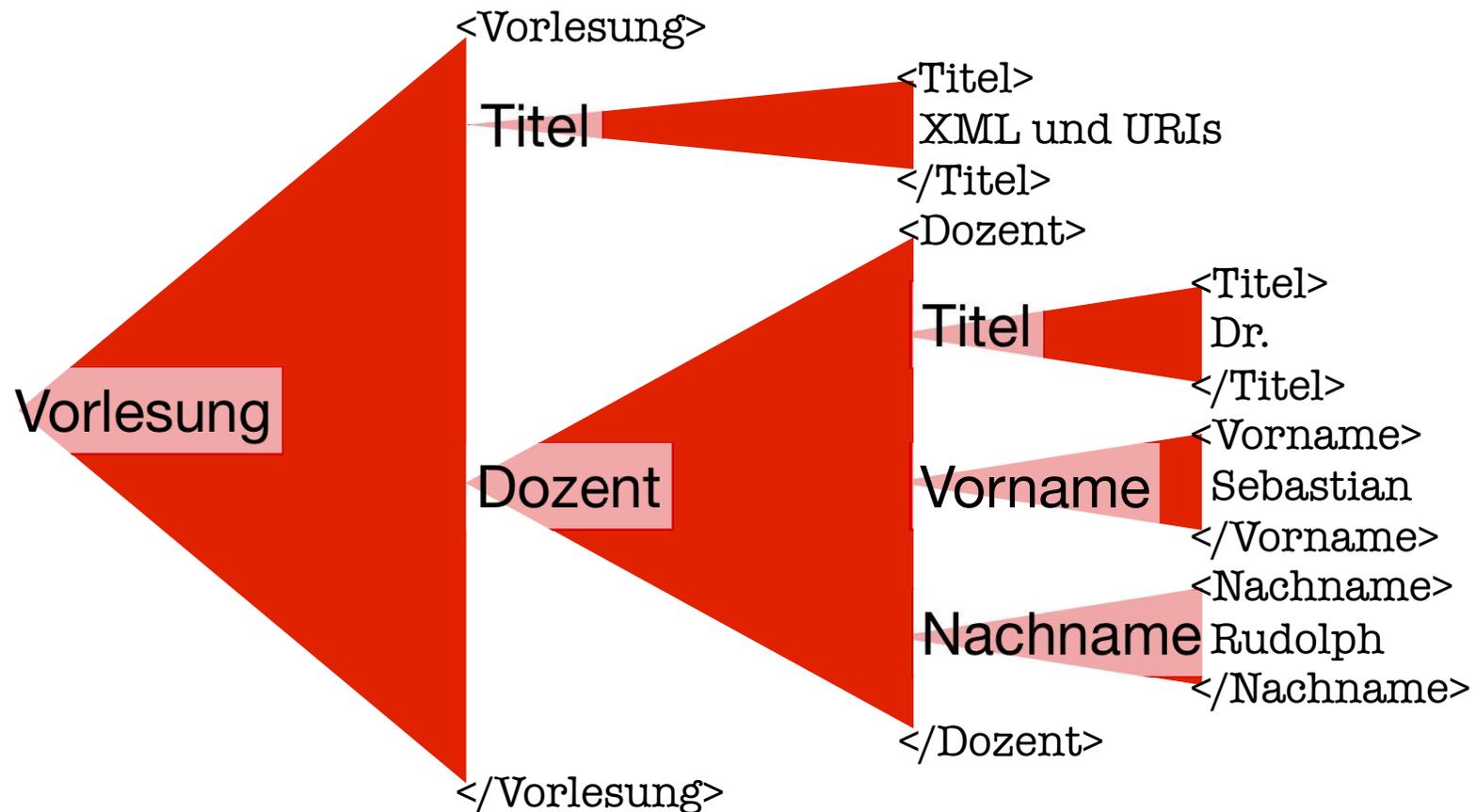
ANNOTATION MIT MARKUPSPRACHEN



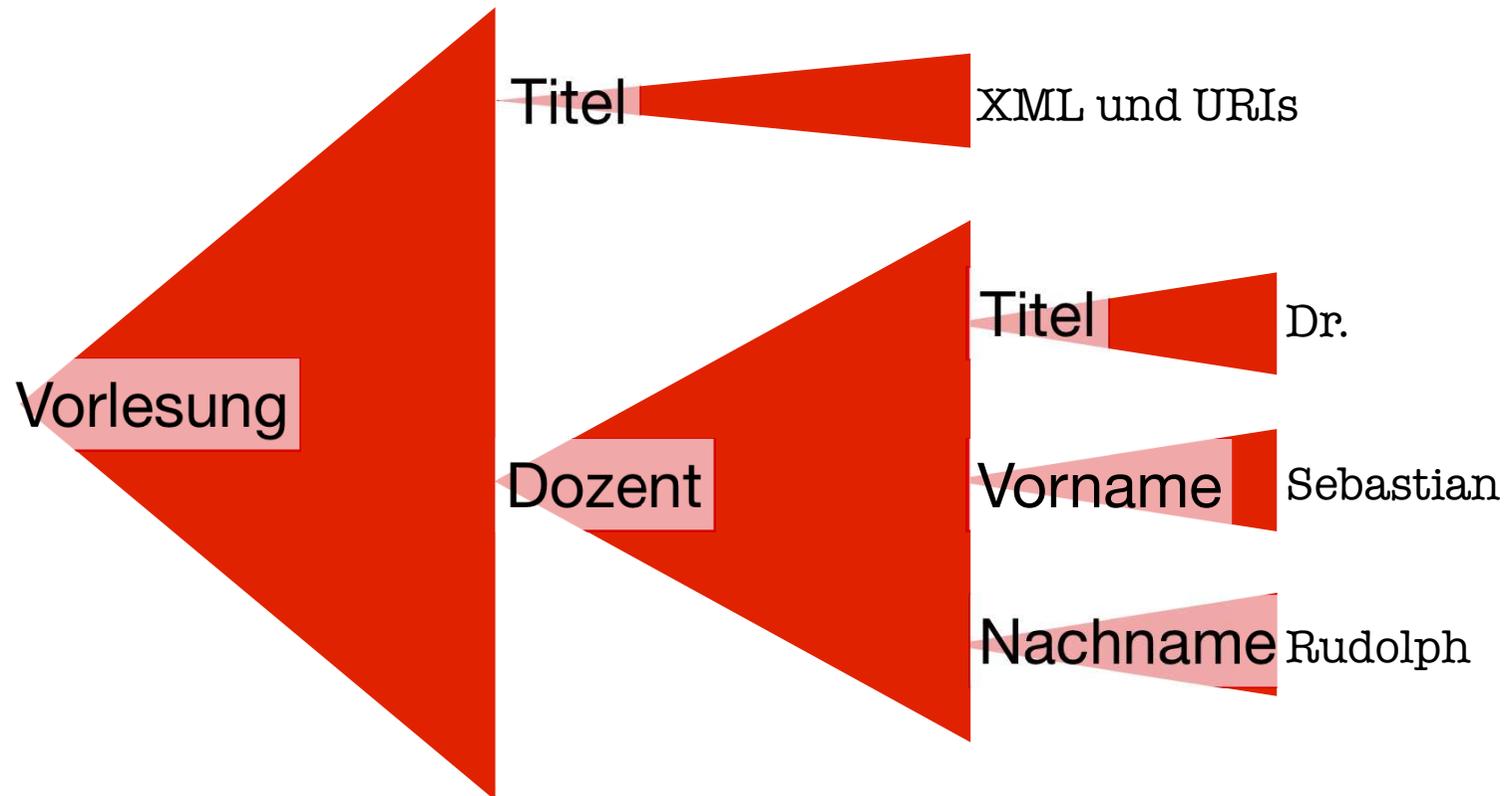
- Verschachtelung von Tags erlaubt:

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

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



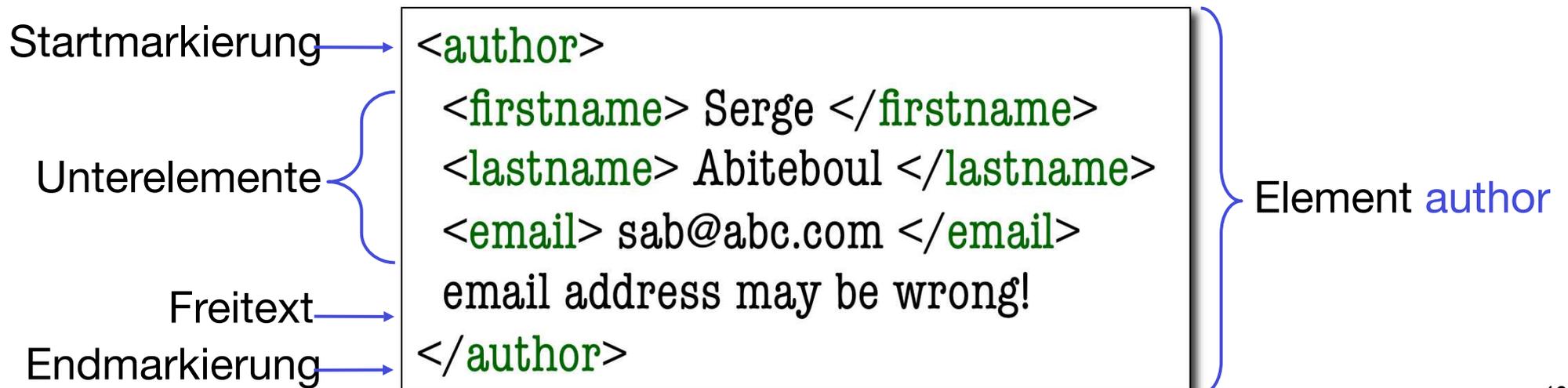
- XML-Dokument ist Textdokument
- beginnt mit Deklaration, die Versionsnummer des verwendeten Standards und optional die Zeichenkodierung enthält, z.B.:

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

XML-SYNTAX (2) – XML-ELEMENT

AIFB 

- 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)

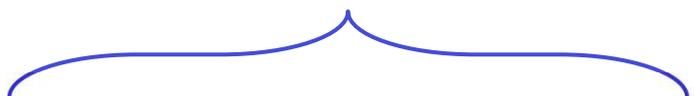


XML-SYNTAX (3) – XML-ATTRIBUT

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"/>
```

XML vs. HTML

AIFB

- HTML: festes Vokabular (Menge von tags) und Semantik (die Darstellung von Text)
- XML: freie Bezeichner zur Beschreibung von anwendungsspezifischer Syntax und Semantik
- $XML \subset SGML$

```

<h1> Bib </h1>
<p>
  <i> Foundations of Databases </i>
  Serge Abiteboul
  <br> Addison Wesley, 1997
<p>
  ...
  
```

HTML

```

<Bib id="01">
  <paper id="012">
    <title> Foundations of Databases </title>
    <author>
      <firstname> Serge </firstname>
      <lastname> Abiteboul </lastname>
    </author>
    <year> 1997 </year>
    <publisher> Addison Wesley </publisher>
  </paper>
  ...
</Bib>
  
```

XML

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

SELBSTDEFINIIERTE URIs

AIFB 

- 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-NAMENSÄÄUME: MOTIVATION

AIFB 

- 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-NAMENSRAÜME

AIFB 

- XML Namespaces sind ähnlich zu Modul-Konzepten in Programmiersprachen
- Disambiguierung von Tag-Namen durch Verwendung unterschiedlicher “Prefixe”
- Ein Prefix wird vom lokalen Namen separiert durch ein “:”, so entstehen prefix:name Tags
- Namespace-Bindungen werden von manchen Werkzeugen ignoriert, sog. “flache Namespaces”

NAMENSRAUM-BINDUNGEN

AIFB

- Prefixe 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 NAMENSRÄUME



```
<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 NAMENS RÄUME



```
<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