

Projektgruppe



Mirko Rose

Reale Anwendungsgebiete des Semantic-Web

4. Juni 2010

- Motivation
- Technologien des Semantic-Web
- Anwendungen des Semantic-Web
- Probleme des Semantic-Web
- Interessante Einsatzmöglichkeiten für unsere PG
- Zusammenfassung und Ausblick

- Netz aus Vielzahl verschiedener **Dokumenten**
 - Wo sind die relevanten Informationen?
 - In Mrd. Websites
- ➔ Aber als unstrukturierter Text!

First-person shooter computer games

Ungefähr 2.310.000 Ergebnisse (0,32 Sekunden)

[Top PC Games - Best PC Video Games - Best PC Games - Top PC Video](#)

[...](#) - [[Diese Seite übersetzen](#)]

F.E.A.R. 3 is a heart-pounding first-person shooter ... shooter, and a faithful follow-up to one of the best first-person shooters... - vor 5 Stunden
www.gamespot.com/gamespot
gefunden - Im Cache - Ähr

Googles Antwort

[TenSpot: Top Ten First-Person Shooters - Features at GameSpot](#) - [[Diese S](#)

Before we dive into the list, we should first define a **first-person shooter**. FPS games

- Netz aus Vielzahl verschiedener **Daten**
- Wo sind die relevanten Informationen?
→ In der Antwort auf die gestellte Frage!

title	title
kkrieger	Call of Duty 2
A Spy in H.A.R.M.'s Way	Call of Duty 2: Big Red One
A/X-101	Call of Duty 4: Modern Warfare
ARMA II	Call of Duty: Modern Warfare 2
Absolute Zero	Call of Duty: Modern Warfare
Action Half-life	Call of Duty: Modern Warfare
Action Quake II	Call of Duty: Modern Warfare
Alien Breed 3D	Carnivores
Alien Breed 3D II: The Killing Grounds	Carnivores

DBPedia's Antwort

Das Web wie es heute ist

Mediziner



Wirkstoff X hilft gegen
Krankheit Z

Publiziert im WWW

Botaniker



Wirkstoff X ist in
Pflanze Y

Publiziert im WWW

Nicht verknüpft!
Wissen da, aber
nicht verfügbar

Das Ideal des Semantic Web

Mediziner



Wirkstoff X hilft gegen
Krankheit Z

Publiziert im SW

Botaniker



Wirkstoff X ist in
Pflanze Y

Publiziert im SW

Automatisch verknüpft:
*Pflanze Y hilft gegen
Krankheit Z*

Wie können Computer verstehen?

- Menschen erkennen Semantik in Dokumenten
 - Aber manuelle Suche zu aufwendig
- Bedeutungen schwer verständlich für Computer
 - Synonyme (Aufzug, Lift, Fahrstuhl)
 - Homonyme (Beinhaltung)
 - abstrakte Begriffe (Freiheit)



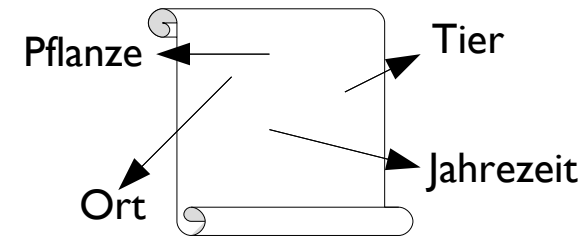
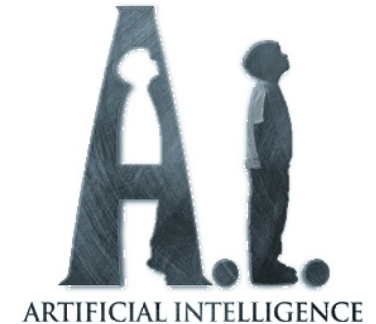
PHOTO-ILLUSTRATION BY RANDY POLLACK



Wie können Computer verstehen?



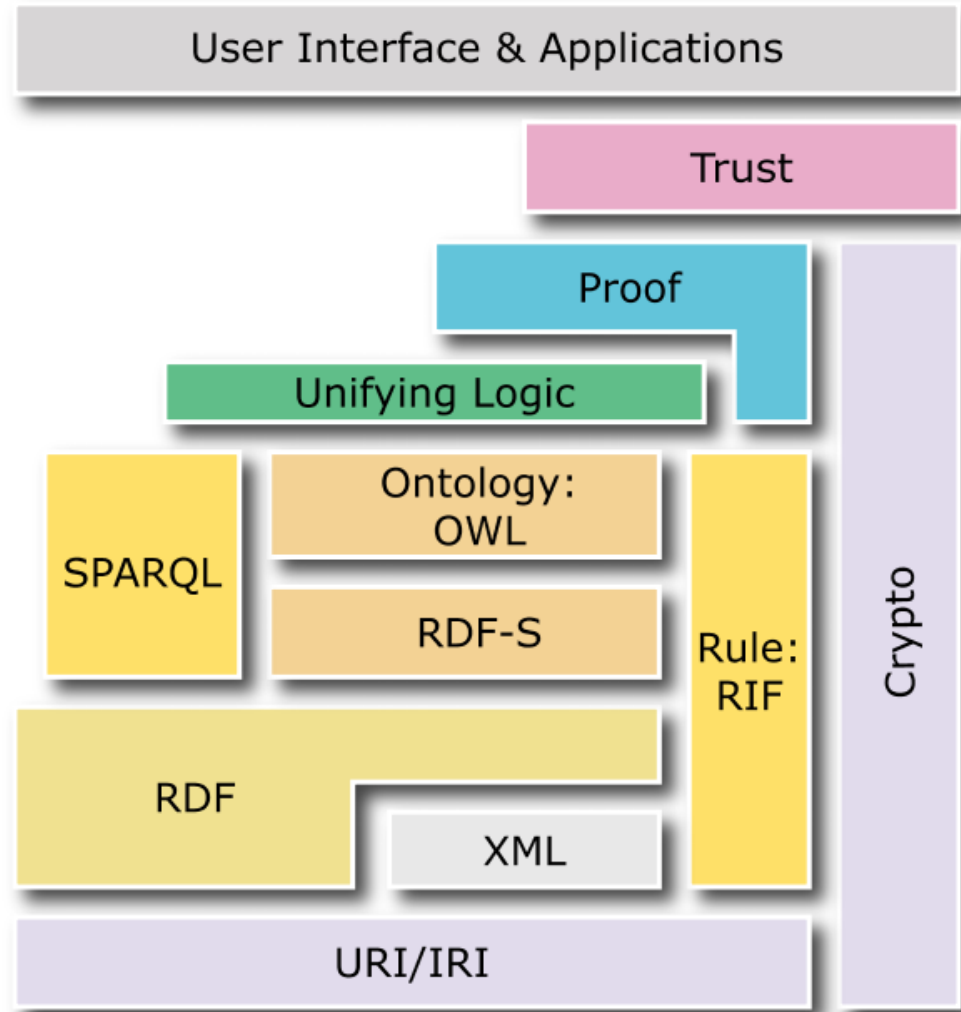
- Künstliche Intelligenz
 - *Information Driven* Themen unserer PG
- Zusätzliche Daten in Dokumente bringen
 - Informationen annotieren und verknüpfen



➔ **Semantic Web als metadatenbasiertes Rahmenwerk für maschinelle Schlussfolgerungen im WWW**

Technologien des Semantic-Web

Semantic-Web-Architektur
Empfehlung des W3C

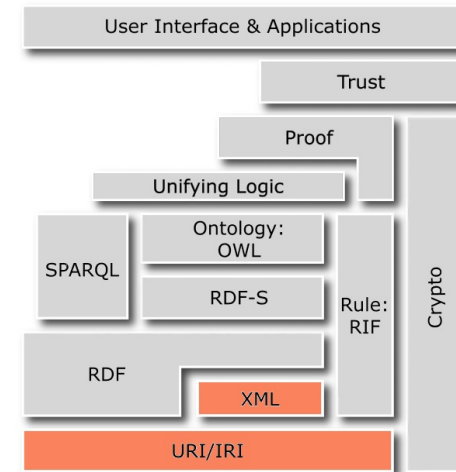


Wie werden die Daten gehalten?

- URI / IRI (ASCII vs. Unicode)
 - **Eindeutige Identifizierung** eines Objekts im Web
 - *http://name.de/hansmeier*
 - *mailto:Adam@eden.org*

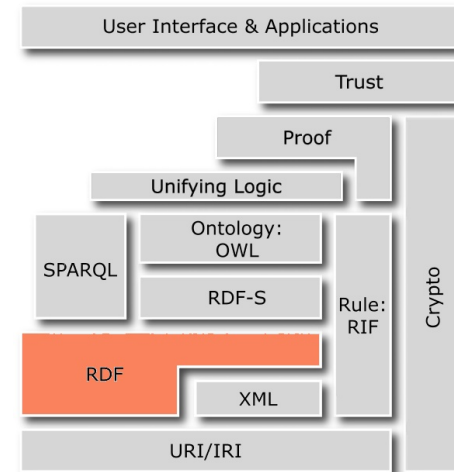
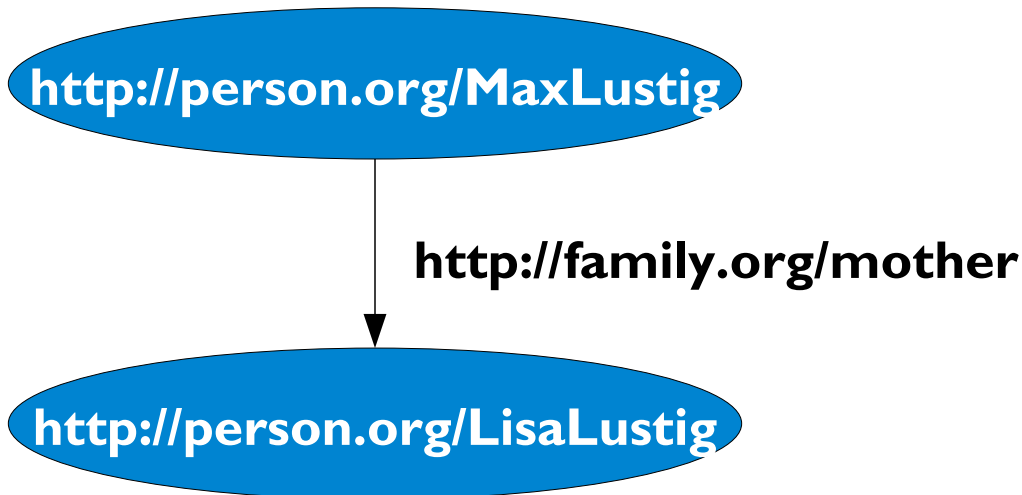
- XML

- Strukturierung der Daten



Wie werden die Daten gehalten?

- RDF erzeugt Beziehung zwischen Daten
 - Wissen in Form von Tripeln
 - Subjekt (URI)
 - Prädikat (URI)
 - Objekt (URI oder Identifier)

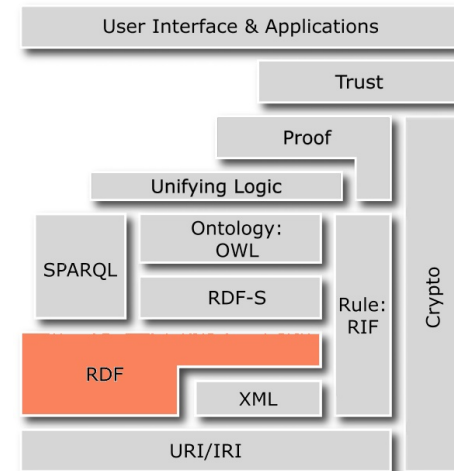


Wie werden die Daten gehalten?

- Webadressen als URIs (<http://www.acdc.com>)
 - Menschenlesbare Informationen beim Aufruf der URI
- RDF im XML Format

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

```
<rdf:Description rdf:about="http://person.org/maxlustig">
  <family:mother>
    <rdf:Description rdf:about="http://person.org/lisalustig"/>
  </family:mother>
</rdf:Description>
</rdf:RDF>
```



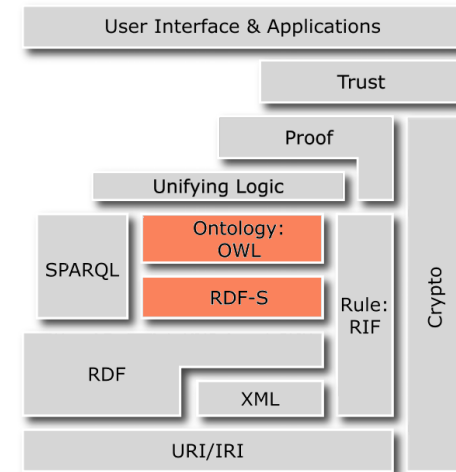
Wissen in Form von Ontologien

- Problem: RDF schränkt nicht ein!

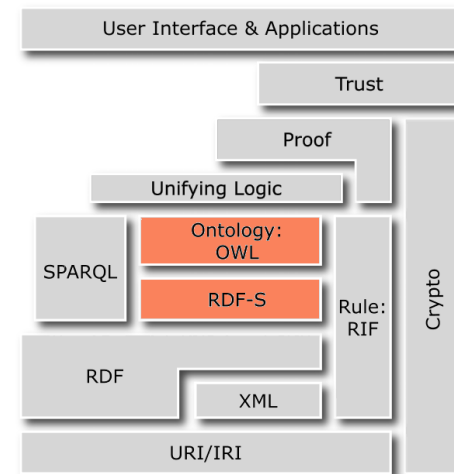


- Lösung: RDF-S (neu: Klassen, Vererbung, Gültigkeitsbereiche)

```
<rdf:Property rdf:ID="http://family.org/mother">  
  <rdfs:domain rdf:resource="#person"/>  
  <rdfs:range rdf:resource="#person"/>  
</rdf:Property>
```

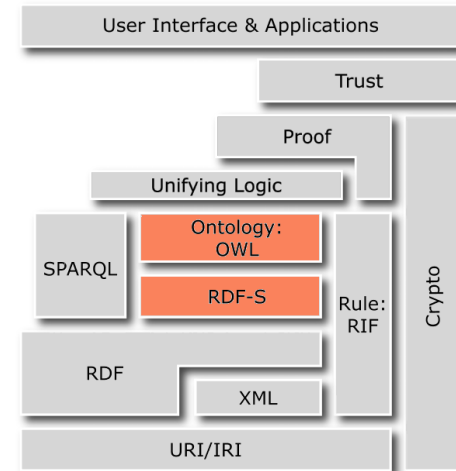
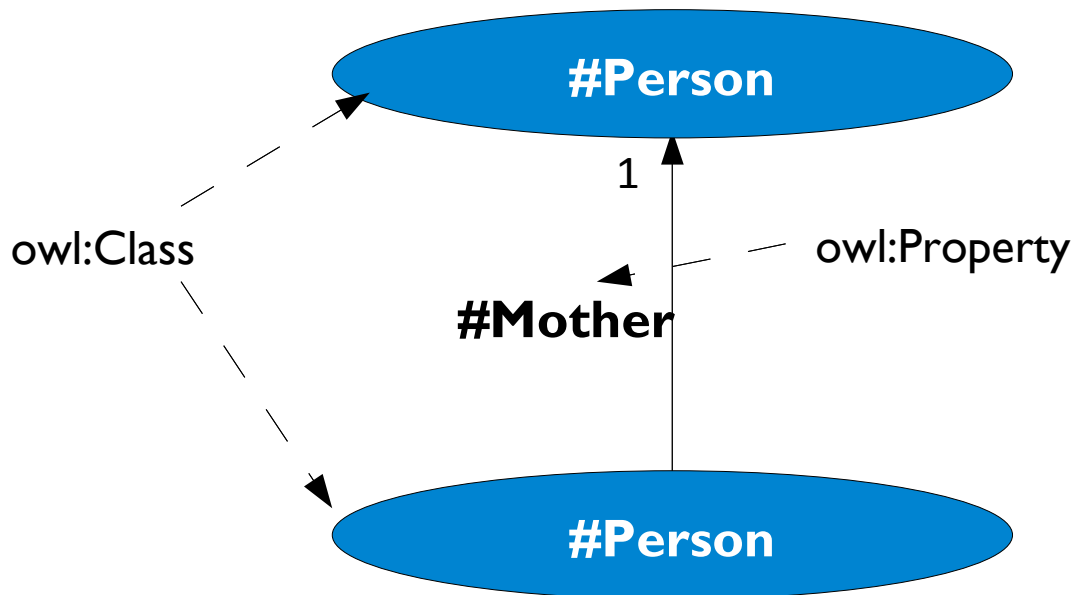
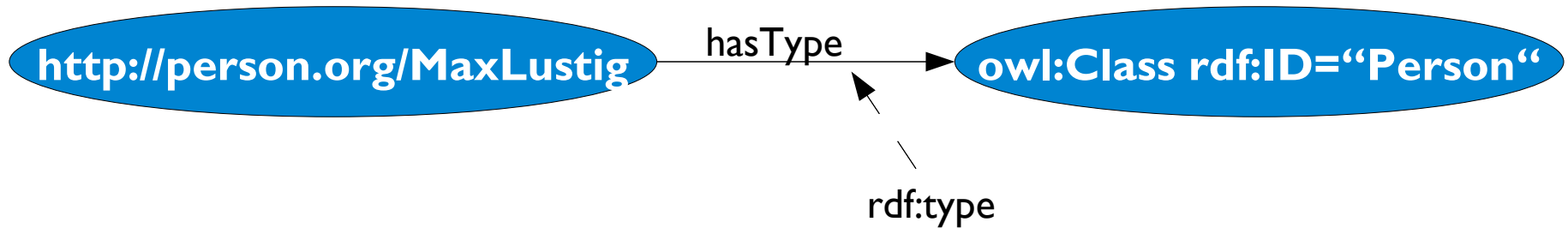


- **Ontologie** = Durch Zusammenfassen von Klassen mit festgelegten Eigenschaften ergibt sich das Modell eines Wissensgebiets
- RDF-S nicht viel mehr als Taxonomie (= Klassenhierarchie)
- **OWL (Web Ontology Language)**
 - **Erweitert die Möglichkeiten**
 - Datentypen
 - Mengenoperationen
 - Stärkere Einschränkungen (Kardinalitäten)
 - Transitivität, Symmetrie uvm.



Wissen in Form von Ontologien

- Kleiner Ontologieausschnitt

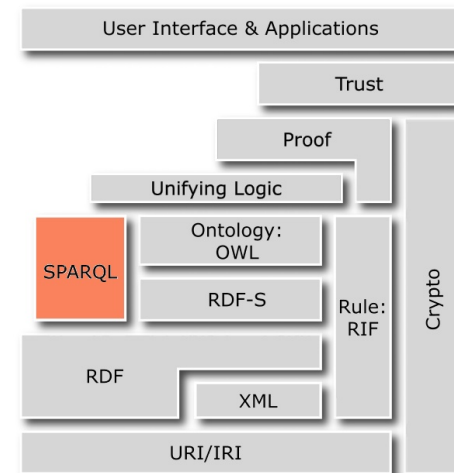


- **SPARQL Protocol and RDF Query Language**

- Anfragesprache für RDF

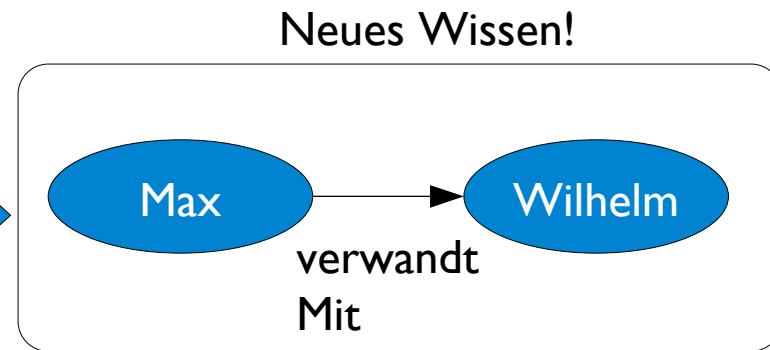
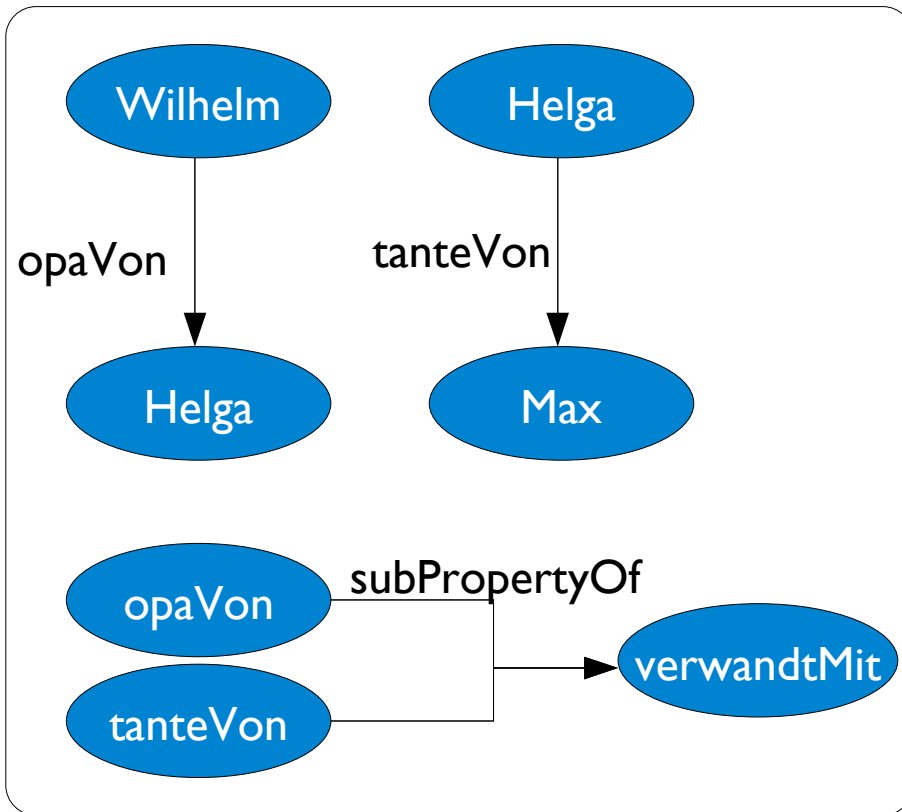
- *Alle Namen und Emails der Manager der Firma in Afrika*

```
PREFIX abc: <http://aCompany.org/companyOntology#>
SELECT ?name ?email
WHERE {
  ?x abc:employee ?name.
  ?y abc:email ?email.
  ?x abc:hasEmail ?y.
  ?x abc:isManagerOf ?z.
  ?z abc:isInContinent abc:africa.
}
```

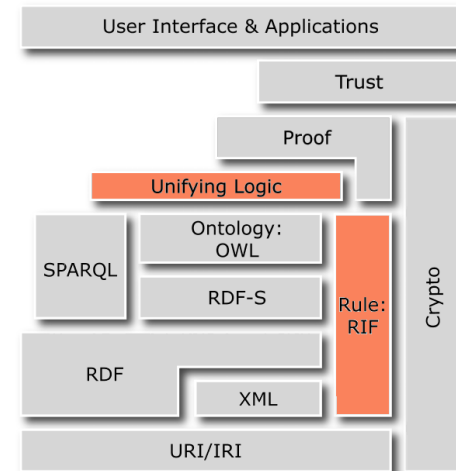


Auf neues Wissen stoßen

- Inferenz-Regeln erschließen neues Wissen in Ontologie
 - Transitivität, Symetrie etc. ausnutzen

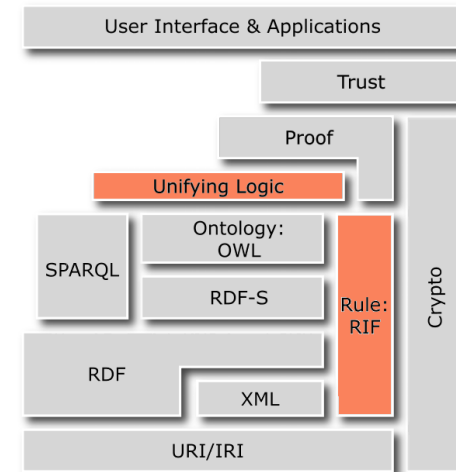


Wissensbasis



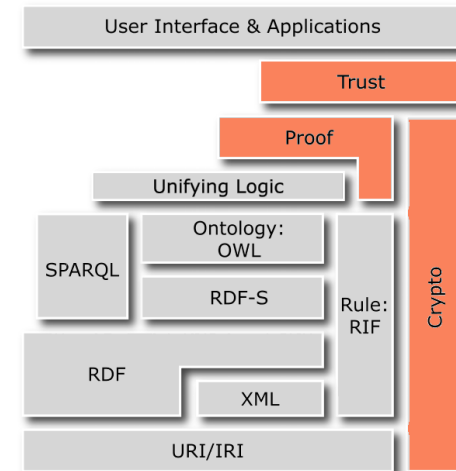
Auf neues Wissen stoßen

- Eigene Regeln erschließen neues Wissen
 - Wenn $x:\text{Haustiere} > 3$ Dann
neue Verbindung $x \rightarrow \text{Tierfreund}$
- Ontologie Mapping durch Regeln
- RIF (Rule Interchange Format)
 - Standard Format für Regeln
 - Zum Austausch von Regeln verschiedener Regelsysteme

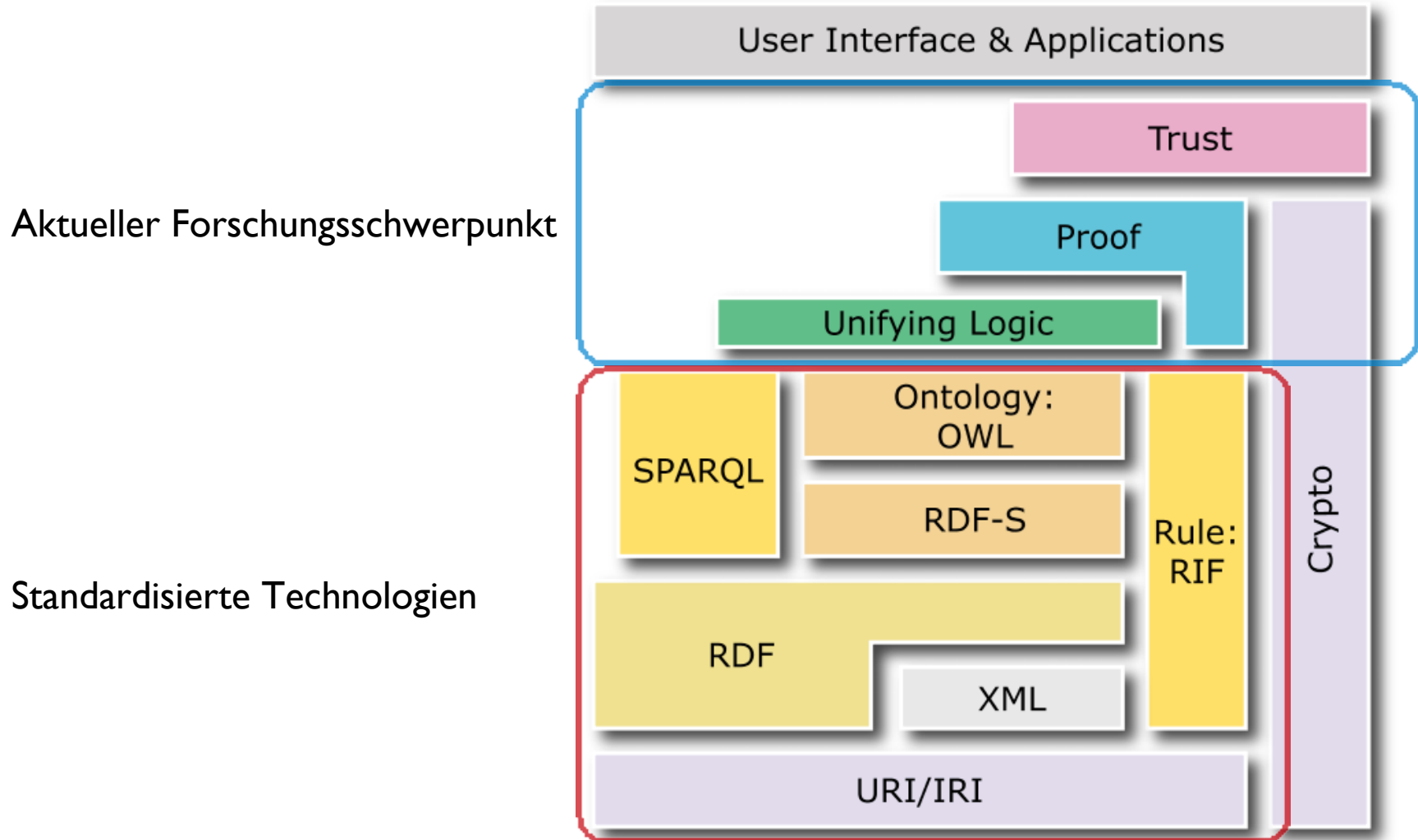


Wissen prüfen und ihm vertrauen

- Proof Ebene
 - Korrektheit der RDF-Aussagen prüfen
- Trust Ebene
 - Wer darf RDF-Aussagen treffen?
 - Wem kann man vertrauen?
- Crypto



Wdh.: Technologien des Semantic-Web



Anwendungen im Semantic-Web



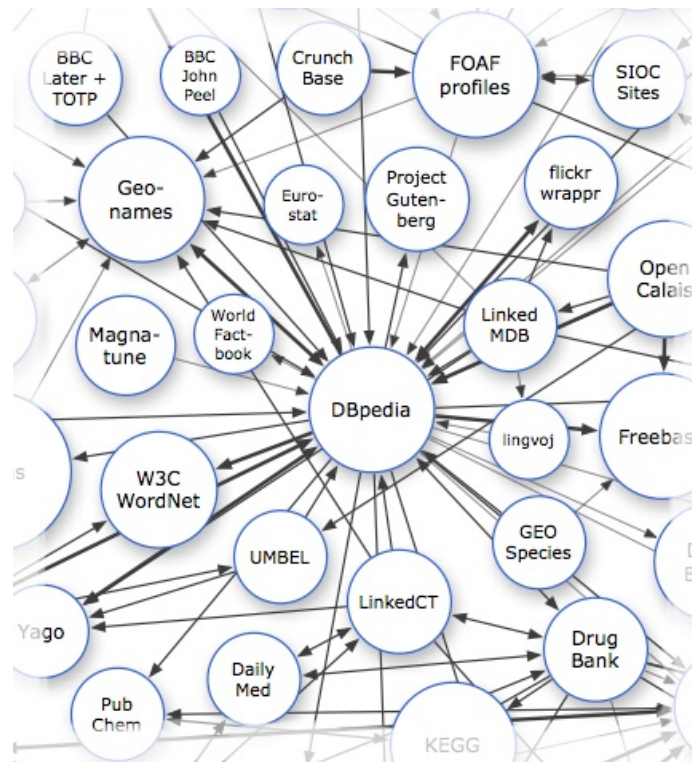
- Allgemeine Anwendungsgebiete
- DBpedia und Linked Open Data
- OnToKnowledge Tool
- Andere Anwendungen, Tools und Ontologien

- Die Realität ist, es gibt nicht **ein** Semantic Web
 - Viele verschiedene kleine, mittlere und große Semantic Webs
 - Viel im Knowledge-Management
 - Telefon-Call-Center
 - Tourist-Informationen
 - Personalmanagement
 - Aber auch E-Commerce / E-Business
 - Semantische Produktinformationen
 - Produktinformationsaustausch zwischen Unternehmen
- Ziel: Verknüpfung der einzelnen „Inseln“ des Semantic-Web

- Bereitet Informationen aus Wikipedia semantisch auf
 - Anfangs per Hand
 - Automatische Umwandlung strukturierter Informationen (bsp. Info-Kasten)
- Semantische Technologien
 - RDF (über 1 Mrd. verschiedener Tripel)
 - SPARQL
 - OWL (260 Klassen und 1.200 verschiedene Eigenschaften)
- Alle Datensätze und die Ontologie sind downloadbar

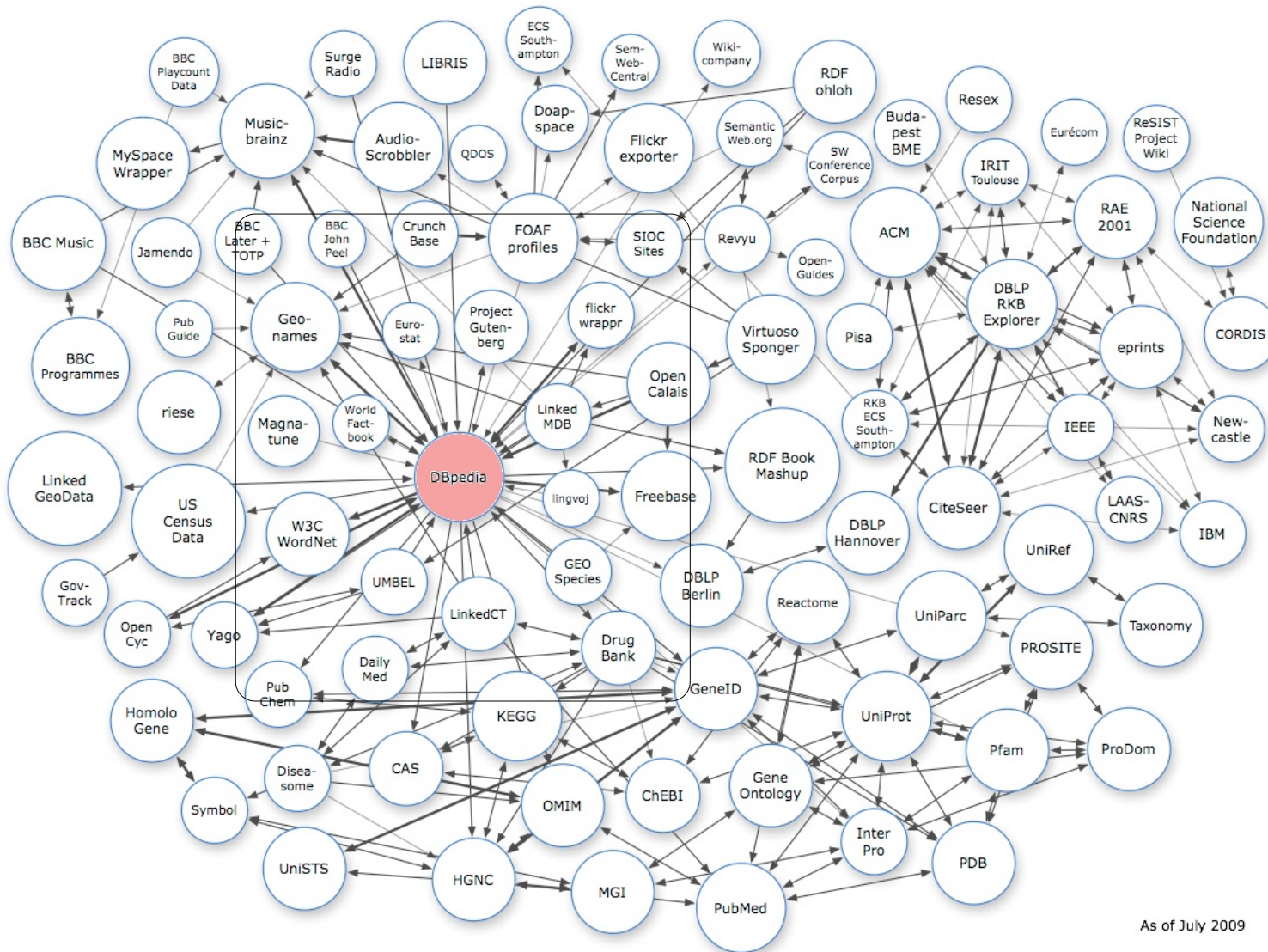
DBpedia und Linked Open Data

- Verknüpfung der „Inseln“ hat begonnen



Linked Open Data
Zusammenschluss

DBpedia und Linked Open Data



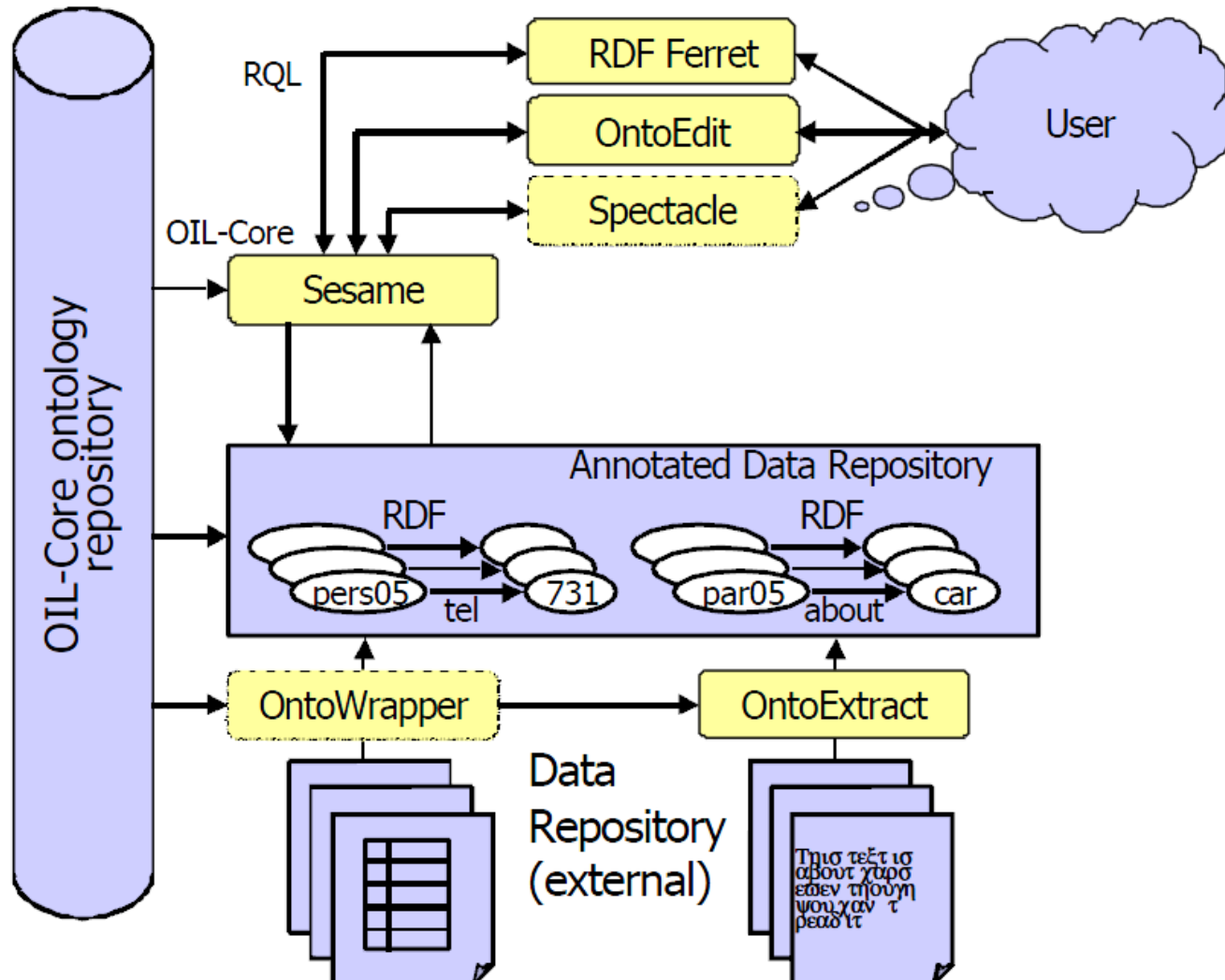
As of July 2009

- *All soccer players, who played as goalkeeper for a club that has a stadium with more than 40.000 seats and who are born in a country with more than 10 million inhabitants*

Google

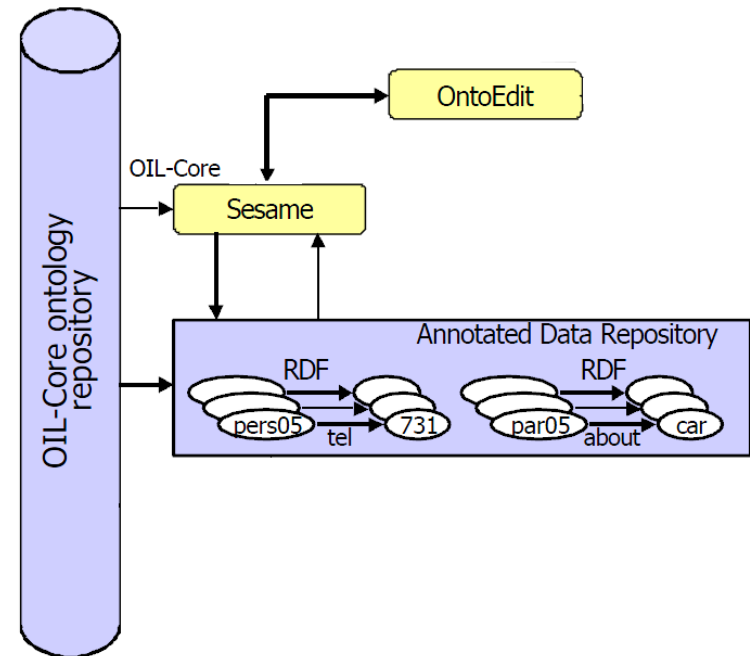
DBpedia Suche

- Entwickelt an der Universität Karlsruhe und anderen europäischen Universitäten und Einrichtungen
 - Von 1999-2002
- Knowledge-Management-Tool
 - Benutzt Technologien des Semantic-Web
- 3 Kernpunkte
 - Erwerb von neuen Informationen
 - Halten und Pflege von vorhandenen Informationen
 - Zugriff auf vorhandene Informationen

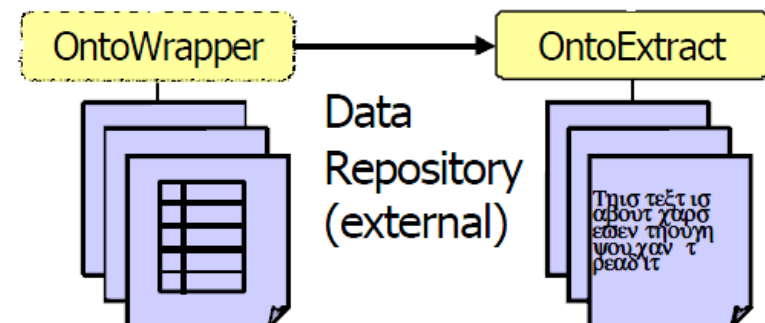


OnToKnowledge Tool – Halten / Pflege

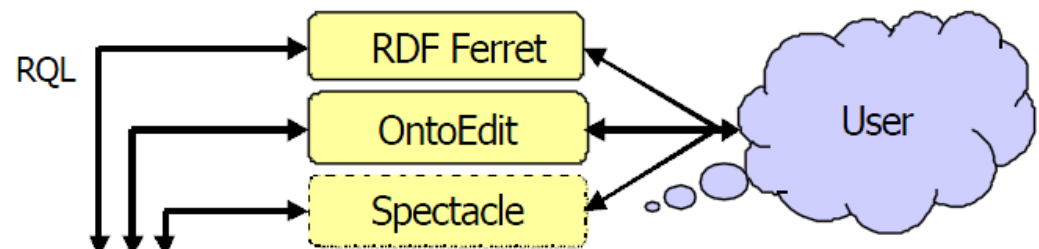
- **Ontology Inference Layer**
 - Vorläufer von OWL
- **Sesame**
 - RDF-Triple-Store Datenbank
 - Auch SPARQL als Query Language
 - Wird bis heute weiterentwickelt
- **OntoEdit**
 - Ontologie Editor



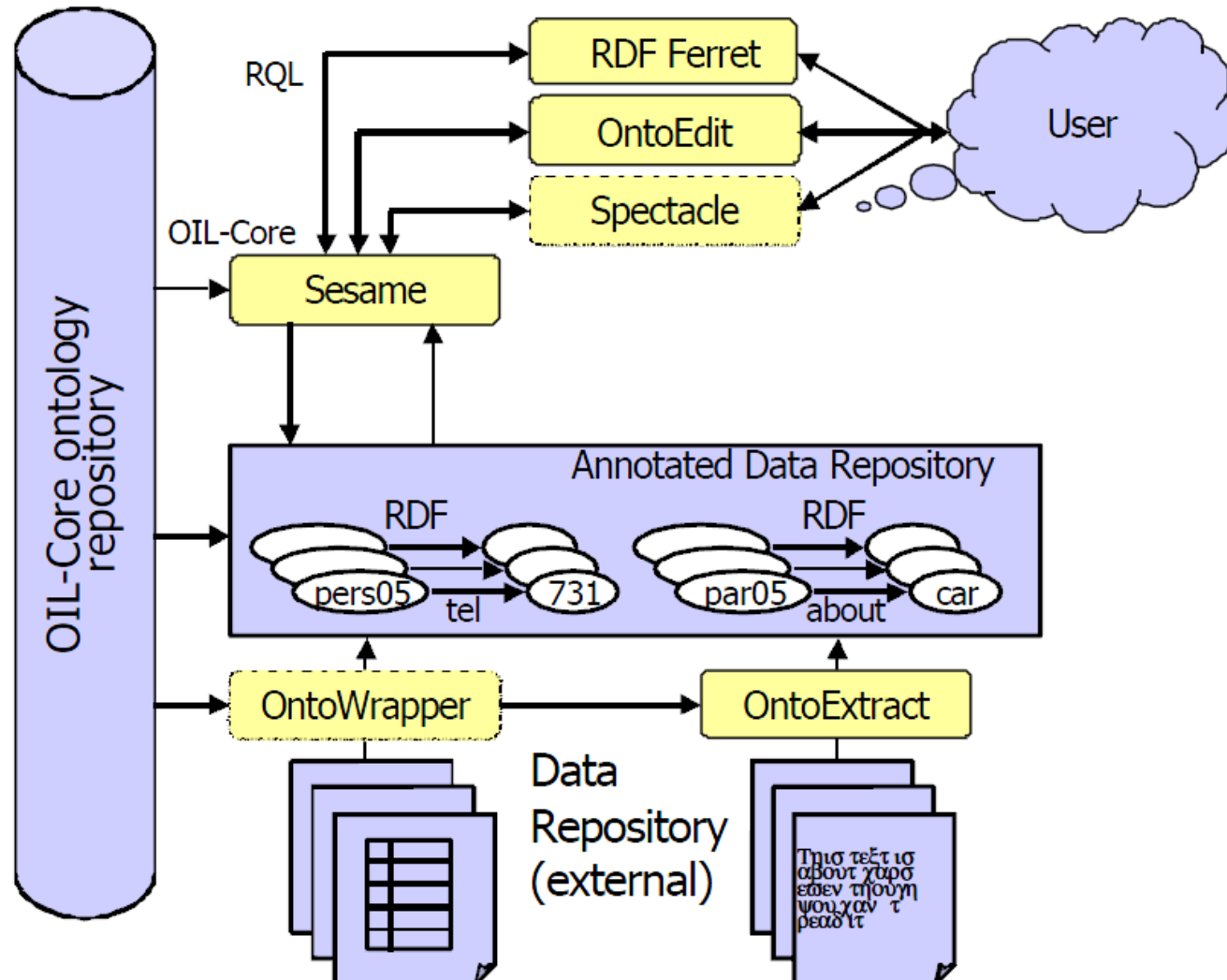
- OntoWrapper
 - Extrahiert semantische Informationen aus **strukturierten** Daten (mithilfe vorgegebener Transformationsregeln)
- OntoExtract
 - Extrahiert semantische Information aus **unstrukturierten** Daten



- RDF Ferret
 - Suchen in den Daten (sowohl RDF basiert als auch Volltext)
 - RQL eine der Abfragesprachen von Sesame
- Spectacle
 - Ontologiebasierte Visualisierung der Daten



Nochmal: OnToKnowledge Tool



- Große, weiterentwickelte Ontologien
 - Linked-Open-Data (DBpedia, FriendOfAFriend, SchemaWeb, theGeneOntology uvm.)
- Semantic Web Rule Language
 - Regelsprache für den Logic-Layer
 - $hasParent(?x1,?x2) \wedge hasBrother(?x2,?x3) \Rightarrow hasUncle(?x1,?x3)$
 - Übersetzt Regeln in spezielle XML-Syntax

Andere Anwendungen, Tools und Ontologien



- Protégé
 - Ontologieeditor und Knowledge-Base-Framework
 - Java-Anbindung
 - Über 150.000 registrierte Nutzer

- Jena
 - Java-Framework zum Erstellen von Semantic Web Applikationen
 - RDF, OWL, SPARQL

Andere Anwendungen, Tools und Ontologien



- Mondeca ITM
 - kommerzielles Tool
 - Gleiche Art wie OnToKnowledge

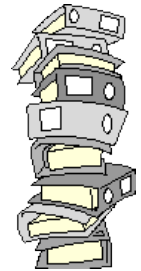
- Semantic MediaWiki
 - Erweitert Wiki-Software um das Einfügen von semantischen Informationen innerhalb Verweisen
 - *"London" is the capital city of `[[capital of::England]]` and of the `[[capital of::UnitedKingdom]]`*

Probleme des Semantic-Web



- Probleme mit formaler Wissensrepräsentation im Allgemeinen
- Weitere Probleme

- Zusätzlicher Overhead (Zeit, Kosten)
- *Knowledge-Engineers* werden gebraucht (Kosten, Weiterbildung)
- Verständigung bei Zusammenführung von Ontologien
 - Formalisierung von abstrakten Begriffen wie *Freiheit*



Weitere Probleme und offene Fragen

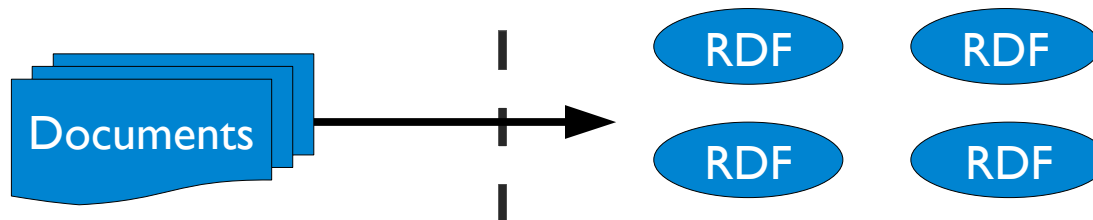
- Befugnisse und Vertrauen
- Meta-Crap wie im normalen WWW
- Datenschutz und Privatsphäre
- Verändertes Suchverhalten
 - Kein „nicht gesucht, aber dennoch interessant“ mehr
- Staatliche Zensur und Überwachung



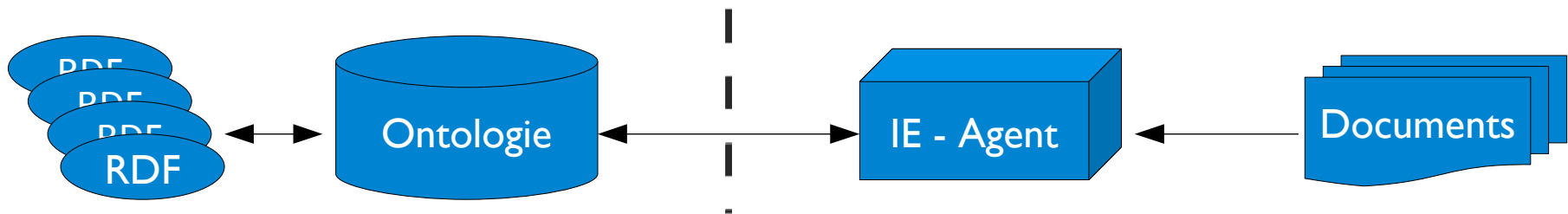
EINE ZENSUR FINDET
NICHT STATT

Interessante Einsatzmöglichkeiten für unsere PG

- Datenhaltung in RDF



- Unterstützung des automatischen Wissenserwerbs (bsp. IE) durch Ontologien und vorhandenes Wissen



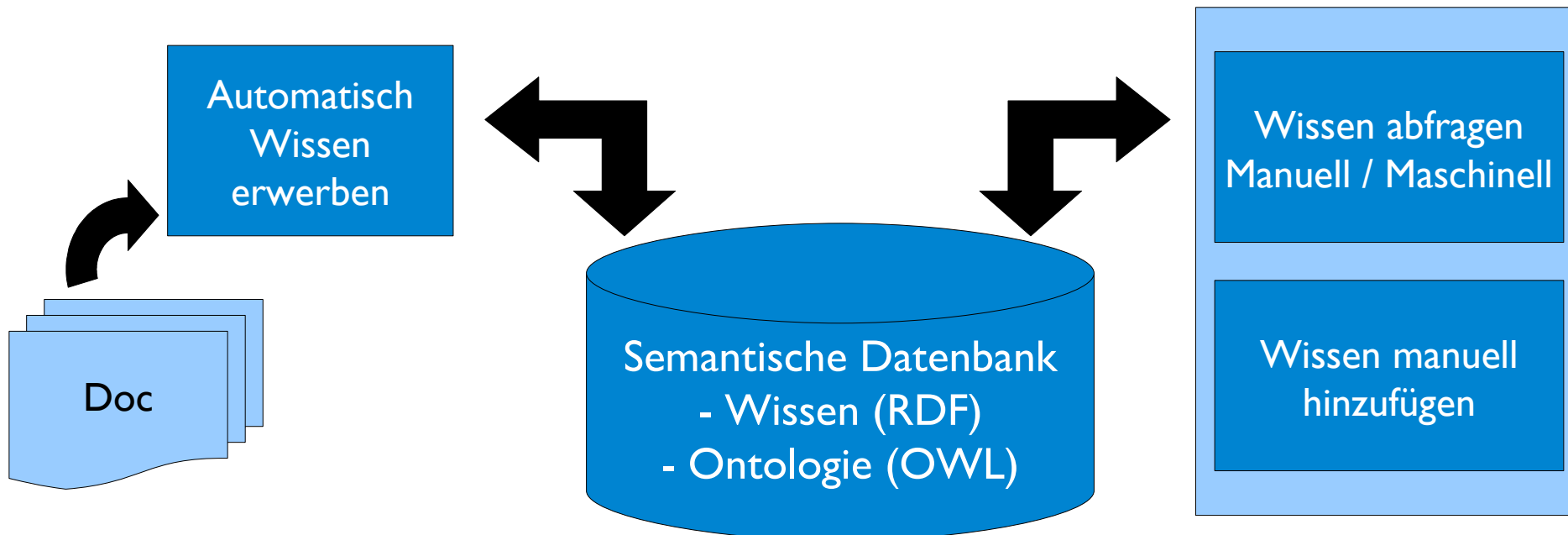
Interessante Einsatzmöglichkeiten für unsere PG


- Semantic-Web als Backend für die Haltung und Abfrage von Informationen

Informationen-Erwerben

Informationen-Halten

Auf Informationen-Zugreifen



- Semantic Web hat Stärken im Bereich der Ontologien
 - Linked Open Data
 - Maschinelle Wissensextraktion notwendig
 - Mensch kann nicht alle Dokumente semantisch nachbearbeiten
 - Muss außerdem unterstützt werden beim semantischen Annotieren von neuen Dokumenten
 - Gute Tools notwendig
 - Semantische Web Services als zukünftiger Schwerpunkt
-  Vereinfachung der Tools und Konzentration auf normalen User um Siegeszug wie WWW anzutreten

Vielen Dank für die Aufmerksamkeit!

Ende