

**Diploma Thesis**



**Fraunhofer**  
Institut  
Informations- und  
Datenverarbeitung

---

# **Design and Implementation of an Automatic Semantic Annotation Service**

**Alina Kopp**

Fraunhoferstr. 1  
76131 Karlsruhe  
Alina.Kopp@iitb.fraunhofer.de

**Oberseminar**



27.02.2007  
Saarbrücken

# Risk and Crisis Management

---

## Issues

- **Common terminology**
- **Interoperability** of data, information and software
  - **Syntactic interoperability**
    - physical connectivity
    - data exchange
  - **Semantic interoperability**
    - common understanding of contents
- **Accessibility** and **availability** of data, information and software



# Improving Risk Management



Open Architecture  
and Spatial Data Infrastructure  
for Risk Management

- Platform independent
- Standards: ISO, OGC, W3C, OASIS



Information Society  
Technologies



Framework Programme 6  
of the European Commission

**Project start** September 2004  
**Duration** 36 months  
**Budget** 13.75 M€  
**Funding** 8.2 M€  
**Contract type** FP6 Integrated Project

## ORCHESTRA Consortium



OpenGIS®



Design and Implementation  
of an Automatic Semantic Annotation Service

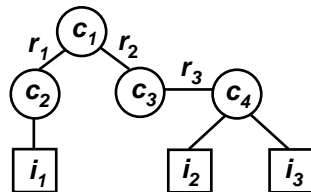
Fraunhofer



Institut  
Informations- und  
Datenverarbeitung

# Approach on Semantic Interoperability

## Semantic Web Approach

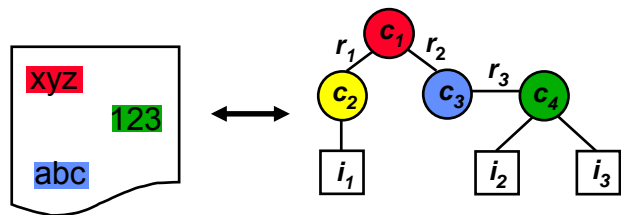


- Define ontologies for common domain and tasks of risk management

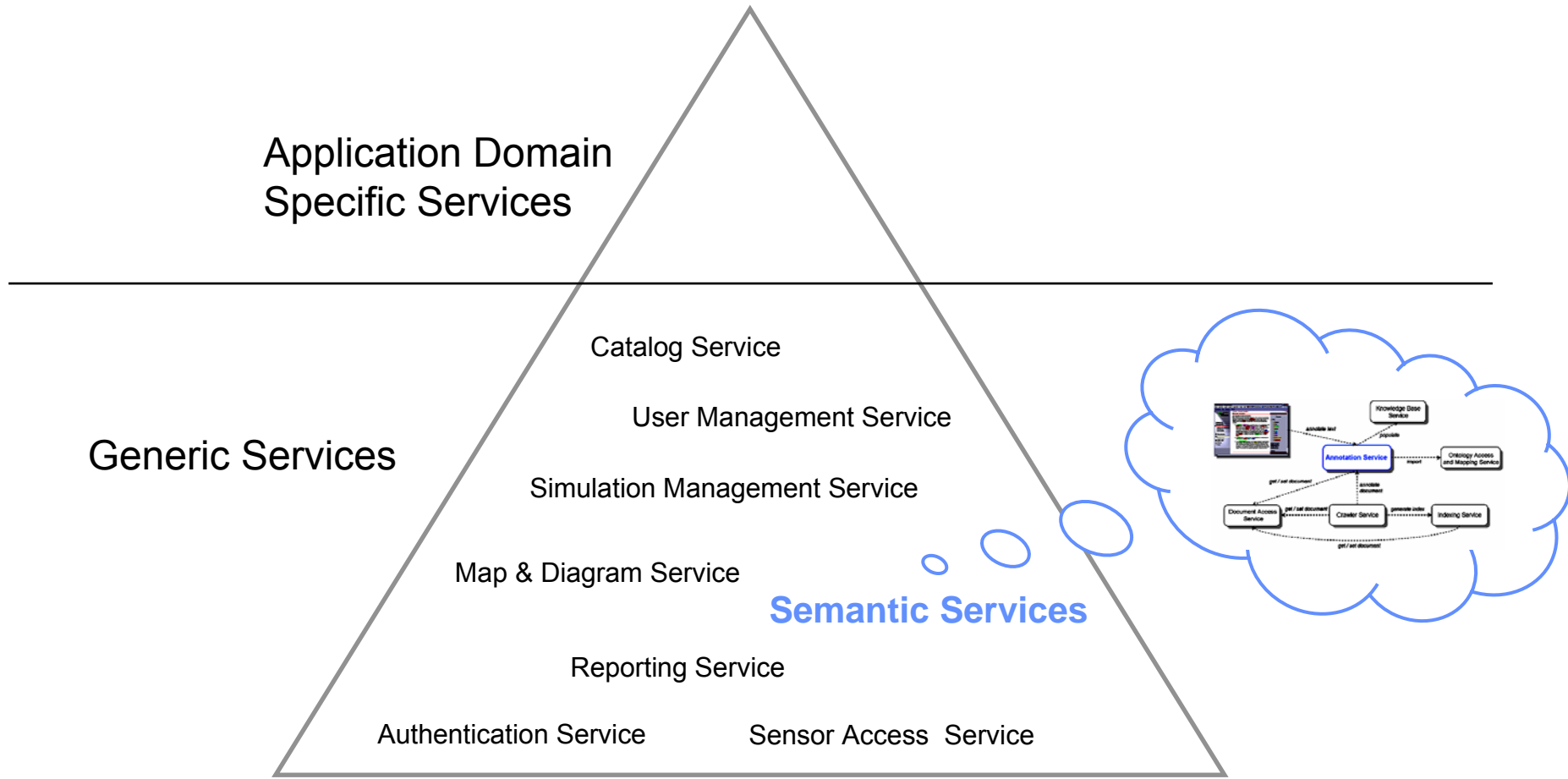
**Ontology** — formal representation of the knowledge associated with a particular subject area

- Define mapping of ontologies to formal descriptions of
  - documents
  - services
  - information models of applications

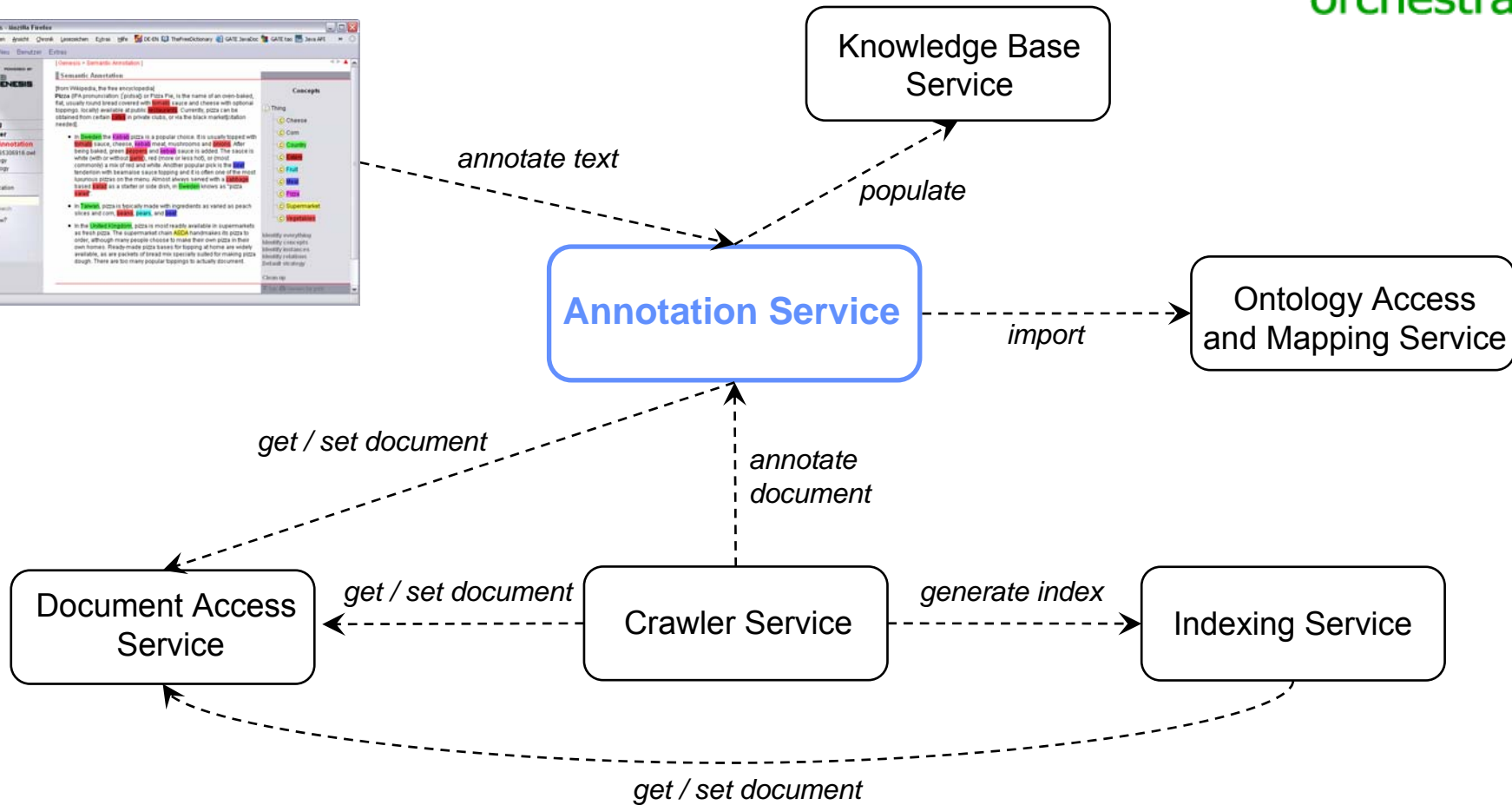
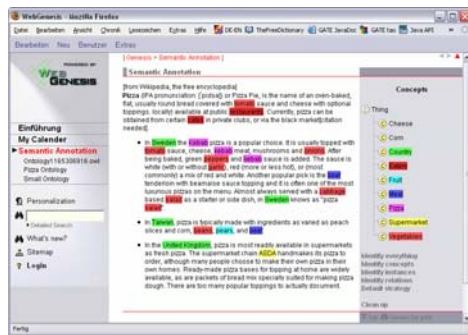
**Annotation** — mapping between a text passage and an ontology element (concept, element or relation)



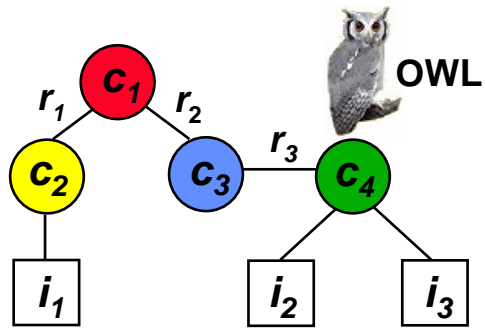
# Generic Service-Oriented Architecture



# Service-Oriented Architecture (Semantic Services)



# Annotation Service



xyz 123  
abc

- annotation strategy**
- identify concepts
  - identify instances
  - identify relations
  - ...

SOAP

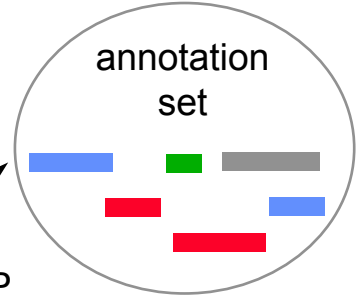
### Annotation Service

- + *annotateText*
- + *annotateDocument*
- + *annotateDataOntology*

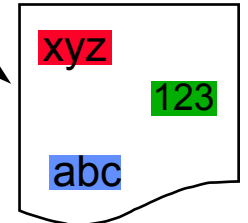
Information  
Extraction  
+  
NLP

Jena API  
by HP

Ontology  
Analysis



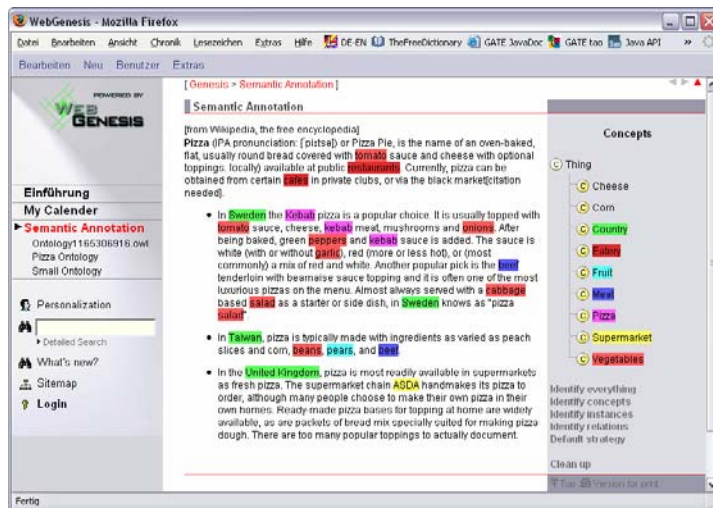
SOAP



# Annotation Service User Interface



## WebGenesis® — An Information, Knowledge and Community Management System



*annotateText (text, ontology, strategy)*



SOAP



*annotationSet*

### Annotation Service

- + *annotateText*
- + *annotateDocument*
- + *annotateDataOntology*



Fraunhofer

Institut  
Informations- und  
Datenverarbeitung

Design and Implementation  
of an Automatic Semantic Annotation Service



# Implementation

