Semantic Web Technologies II

SS 2009

15.07.2009

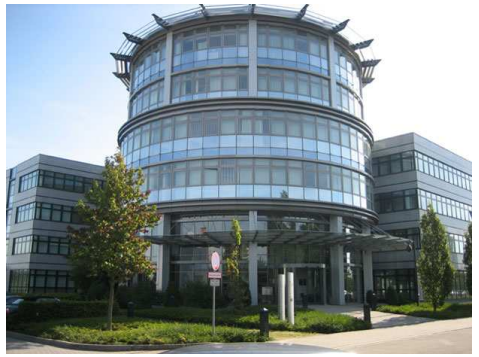
Semantic Cloud Management

Dr. Sudhir Agarwal
Dr. Stephan Grimm
Dr. Peter Haase
PD Dr. Pascal Hitzler
Denny Vrandecic

fluid Operations

- Mission Statement
  - solving today's complex **cloud computing** challenges with truly innovative technologies:
  - virtualization, cloud-based management, Web 2.0 and semantic technologies

- Key data
  - Growing startup with a spirited team
  - Offices in the SAP Partner Port Walldorf
  - Some of the largest IT companies as customers or development partners
  - http://www.fluidops.com/



Agenda

- Cloud Computing
  - Overview
  - Challenges related to semantics
- Semantic Technologies for Cloud Management
  - Information integration
  - Wiki-based documentation collaboration
  - Intelligent user interfaces
- Demo Information Workbench
- Open Research Topics

Cloud Computing

- Web based
- Virtualized
- On-demand
- Easy to use
- Scalable
- Pay as you go

Software as a Service

Platform as a Service

Infrastructure as a Service

Alexander Lenk, Thomas Sandholm, Markus Klems, Jens Nimis, Stefan Tai
What's Inside the Cloud? An Architectural Map of the Cloud Landscape
Workshop on Software Engineering Challenges in Cloud Computing @ ICSE 2009

fluid Operations @ SAP

- SAP Center of Excellence
  - ~ 800 Proof of Concept projects /year
  - Project = Management of an SAP landscape over the Internet
  - Project duration: between 2 weeks and 1 year
- Hardware

6 NetApp Filer

250 Blades

110 IWDF

2 HP EVA

IBM P6

HP Superdome

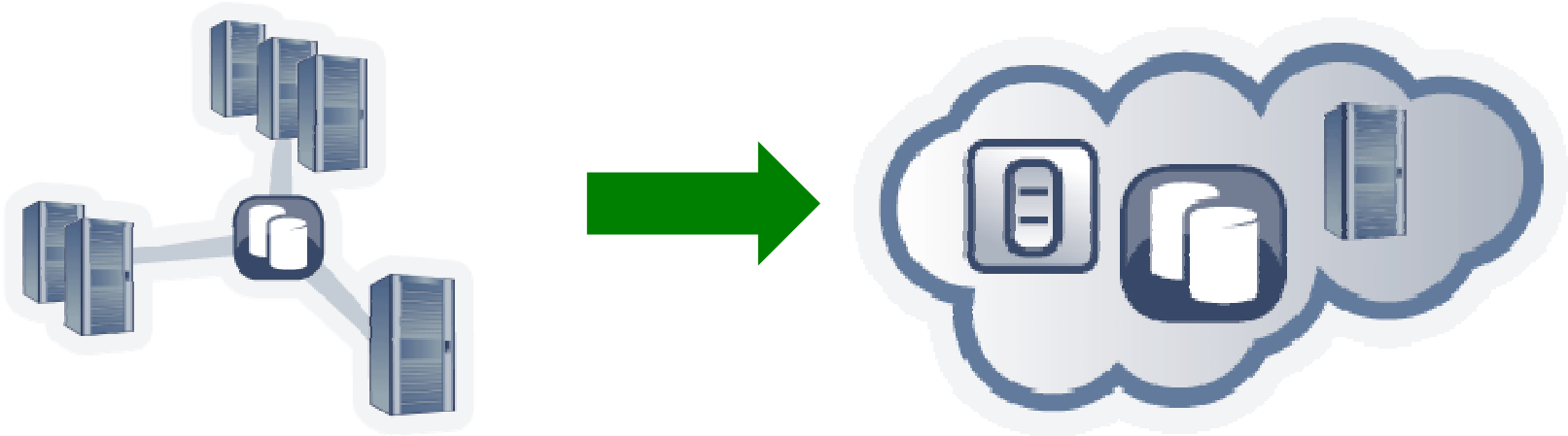
- *What if you could provision a complete SAP landscape in **10 minutes**?*

- *What if that landscape were already pre-seeded with all relevant **application content**?*

- Compare …
  - 8 weeks average for the same at SAP customer sites, 4 within SAP
  - 1 week to resurrect a (successful) test landscape
  - 15 min for an empty WinOS with Amazon EC2

Enterprise Clouds – *eCloud* Vision

- All resources of an ***adaptive, cloud-enabled*** IT environment can be ***set up, monitored, and maintained*** from a ***single, unified,*** and ***intuitive management console:***
  - Internal and external IT resources accessible across stack without vendor lock-in
  - High degree of automation and IT provisioning at click of button on the level of enterprise landscapes
  - Internal portal of private/public IT services with e.g. pay-as-you-go cost models



Manage IT like an eCloud

IT admins

Application customers

CXOs

eCloudManager Product Suite

fluidOps:
stack virtualization and ***semantic integration*** as foundational capabilities for efficient automation

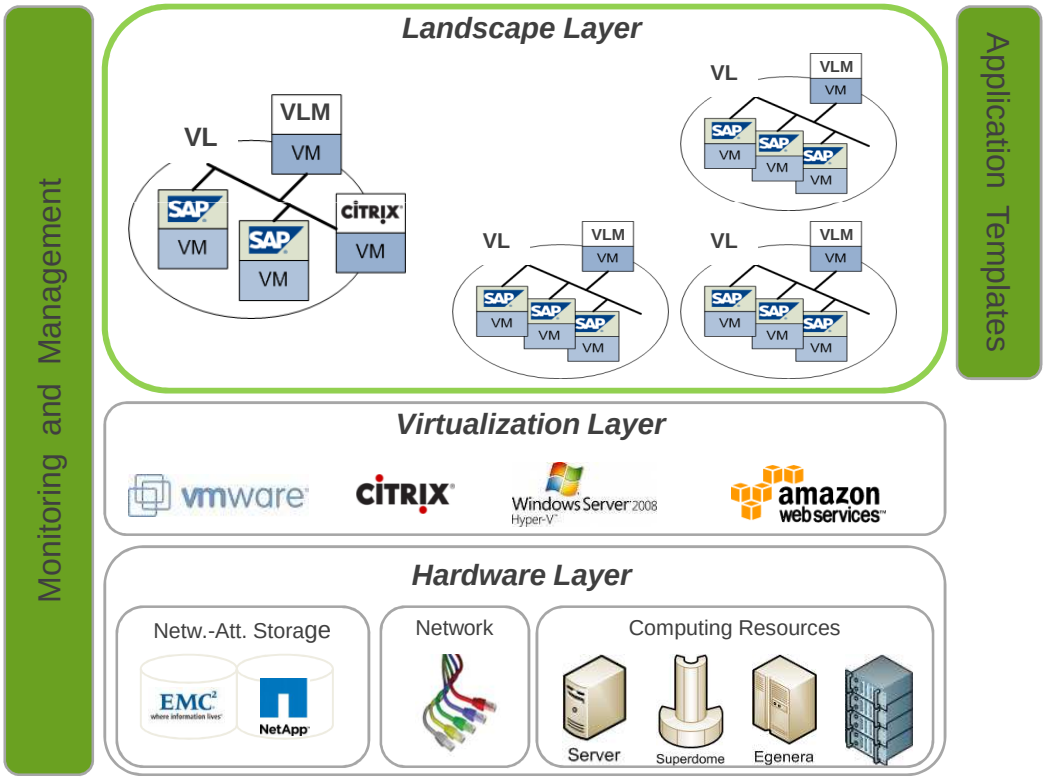
eCloud - Solution View


Web-UI


Native Plug-ins


Self Service Portal

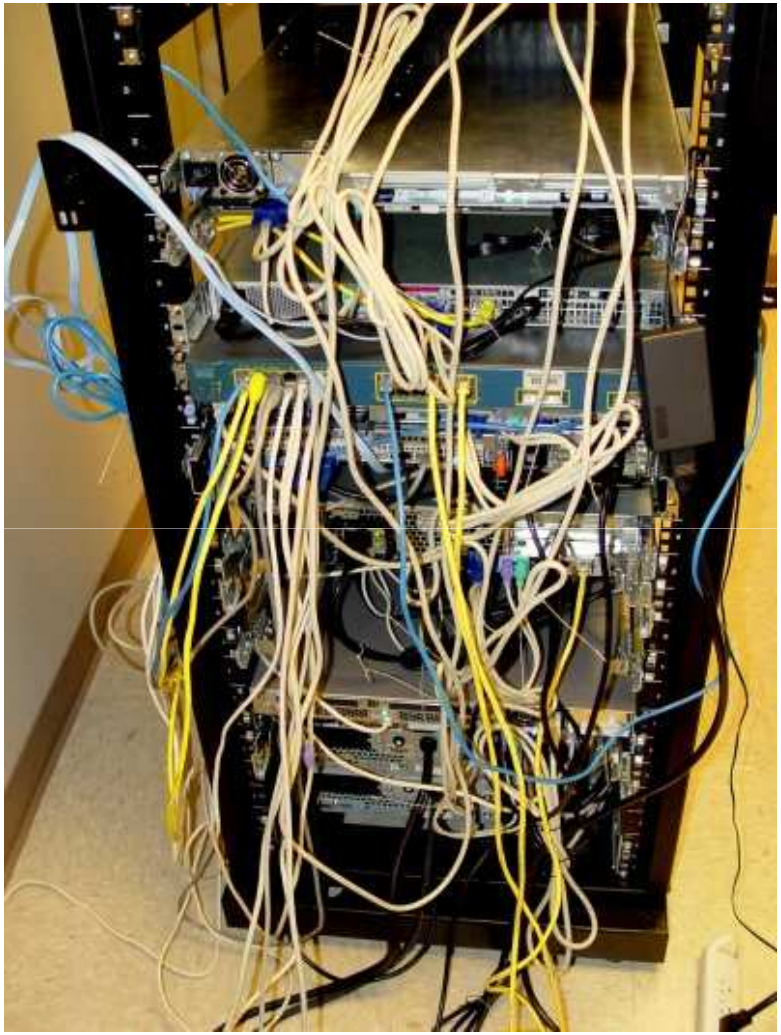

Dashboards



- Business layer
  - Admin, business client & business sponsor cockpits
  - Integrated monitoring and system management along the entire stack
  - Landscape-level provisioning with unique application template support & automated cloning
  - Unprecedented application development, testing and production productivity
- Infrastructure layer
  - Open platform for heterogeneous systems
  - Lower TCO via optimized HW utilization and maintenance automation
  - Flexibility of scaling up or down
  - Security and reliability

Challenge: Automation

- Automation
  - “as a Service” environments are always dynamic
- Virtualisation
  - Do we change the wires physically or virtually?

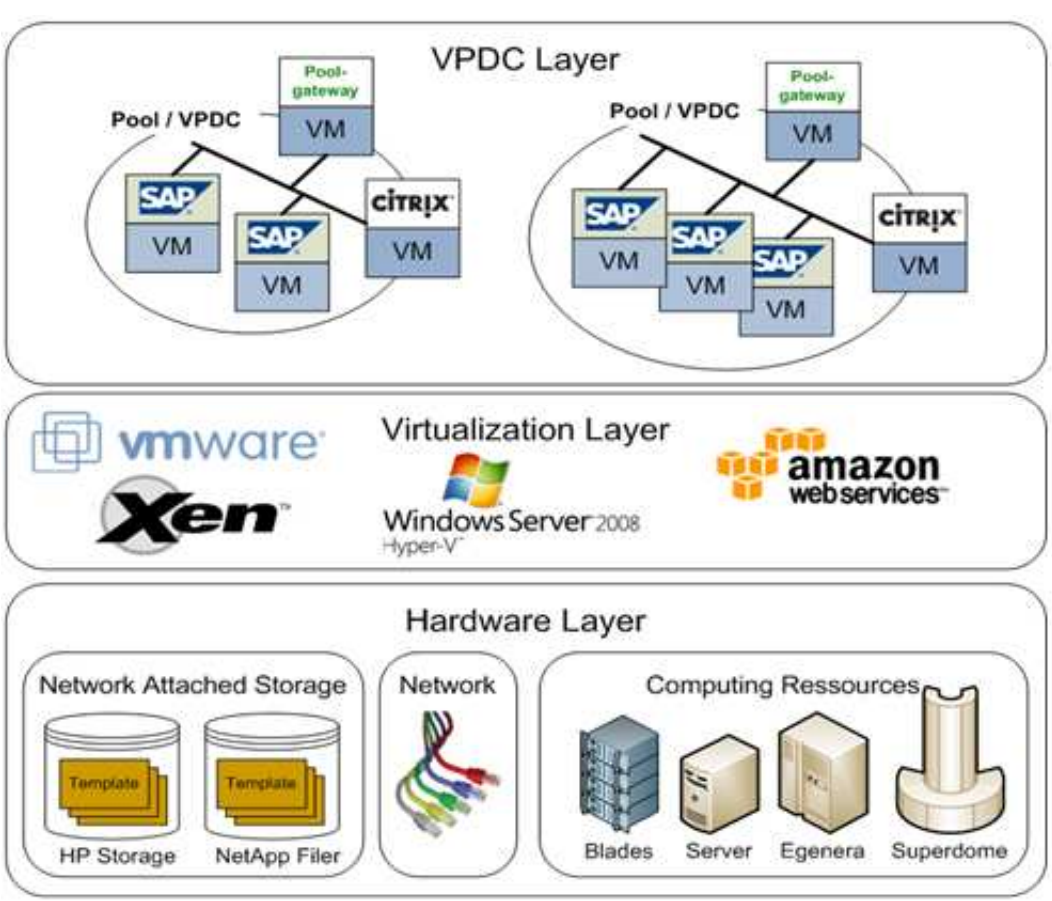


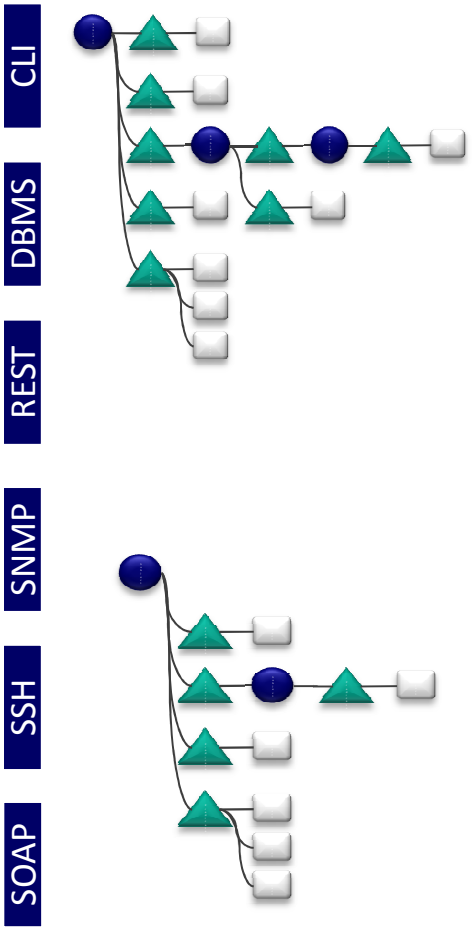
Prerequisite for Automation

- Orchestration of all layer
  - CPU, Storage, Network, Application
- Integration of interfaces
  - Data, Services

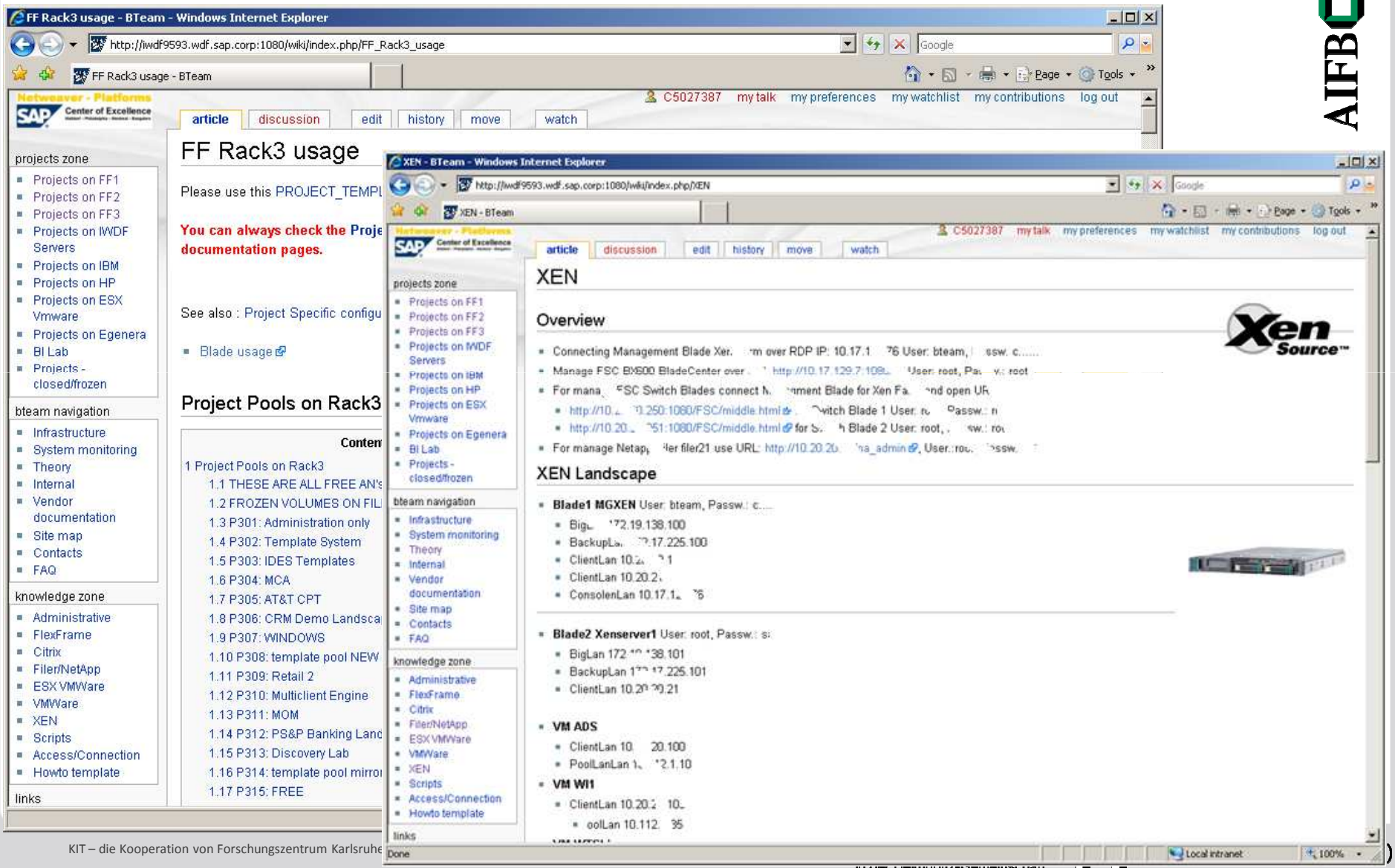
Cloud Infrastructure Management

Semantic Integration for Cloud Management





Documentation using Wikis



Documentation using Wikis

- Problems
  - Dynamic Data (e.g. IP adresses) need to be entered redundantly
  - Users need to use multiple, isolated applications
  - Lack of structure and semantics

Challenge: Intelligent User Interfaces

- Personalization and views on the data according to preferences and user roles
- Actionable information: Ability to interact with the data
  - E.g. Logging onto a server, starting a VM, cloning a system
  - Calling a person, notification via email, …

Agenda

- Cloud Computing
  - Overview
  - Challenges related to semantics
- Semantic Technologies for Cloud Management
  - Information integration
  - Wiki-based documentation collaboration
  - Intelligent user interfaces
- Demo Information Workbench
- Open Research Topics

Semantic Technologies for Cloud Management

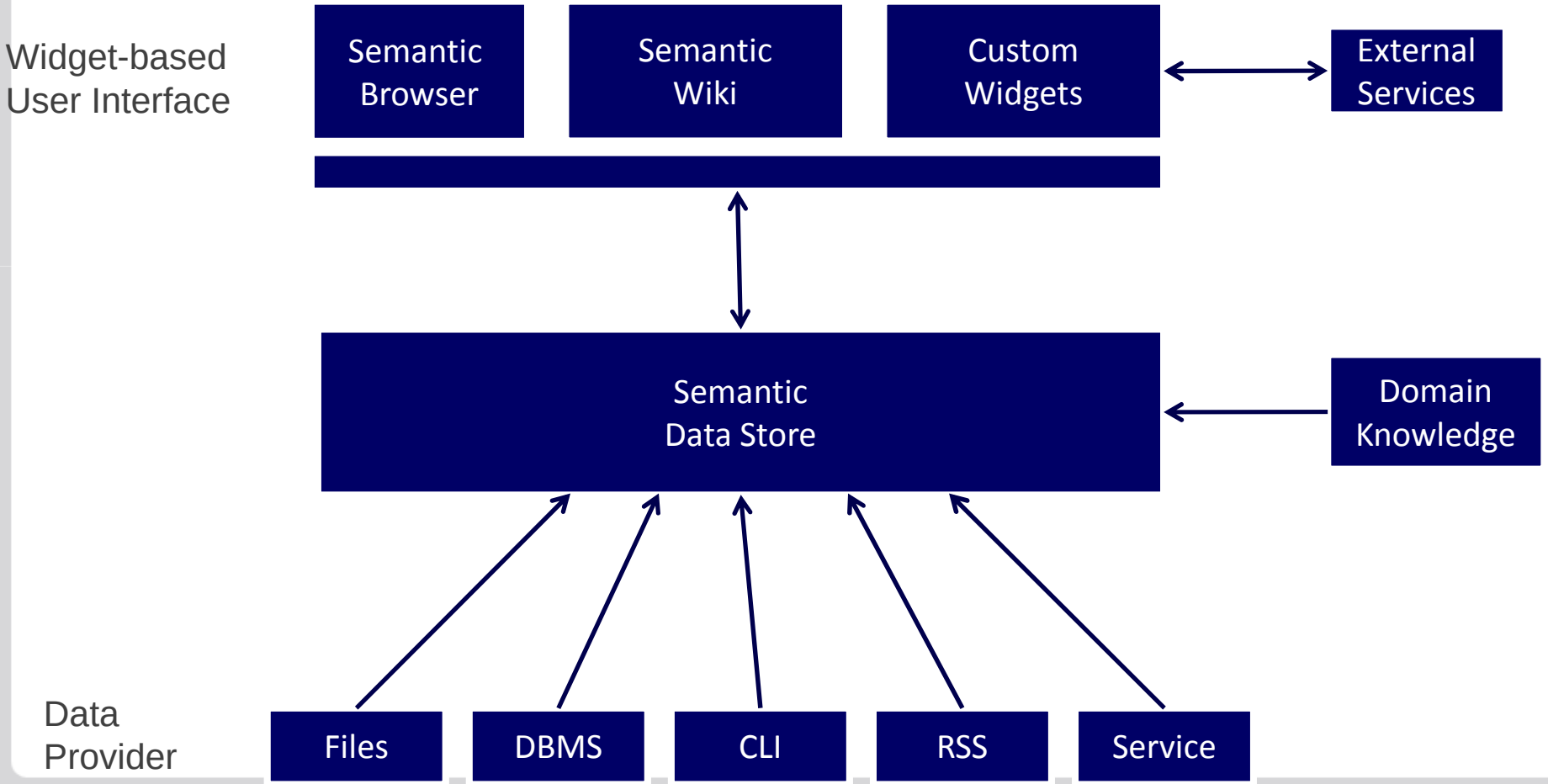
- Sample Application: Information Workbench
  - Collaborative data processing and annotation engine

- Uses of semantic technologies
  - Import and integrate live data from heterogeneous sources, semantic search across data sources
  - Annotation, documentation, collaboration
    - Semantic Wiki technology
  - Visualization driven by semantics
    - Customizable UI widgets
    - Web 3.0 mashup features

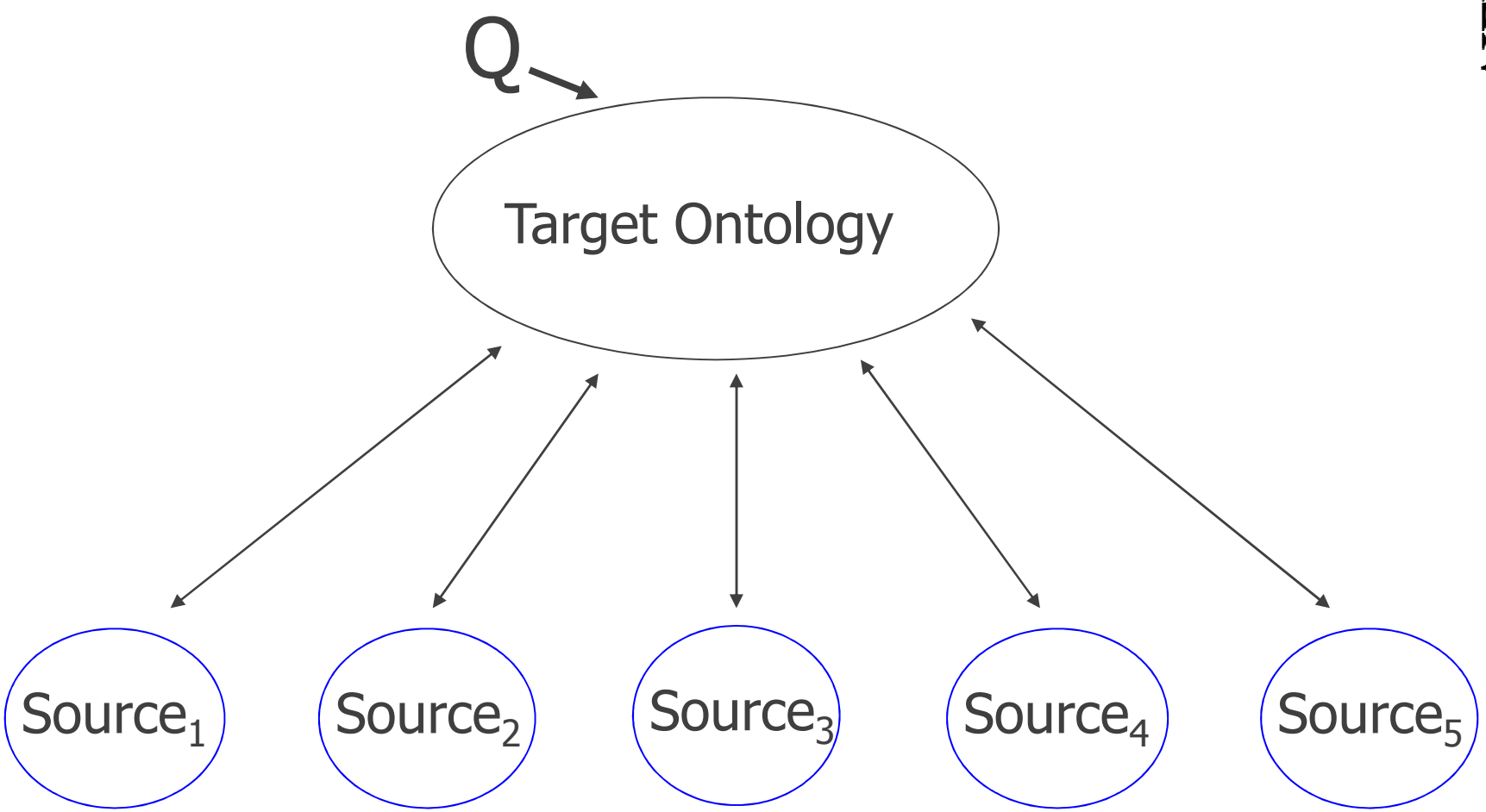
Example: Cloud Management at the SAP Center of Excellence

- Integration of data from diverse sources
  - User information
  - Infrastructure
  - SAP landscapes
  - Project databases
- Wiki-based collaborative annotations
  - Documentation tighty integrated with live data
  - No separate wiki needed
  - Avoiding redundancy of information
- Actionable information
  - E.g. logging onto a server, cloning an SAP system

Information Workbench - Architecture



Ontology-based Information Integration

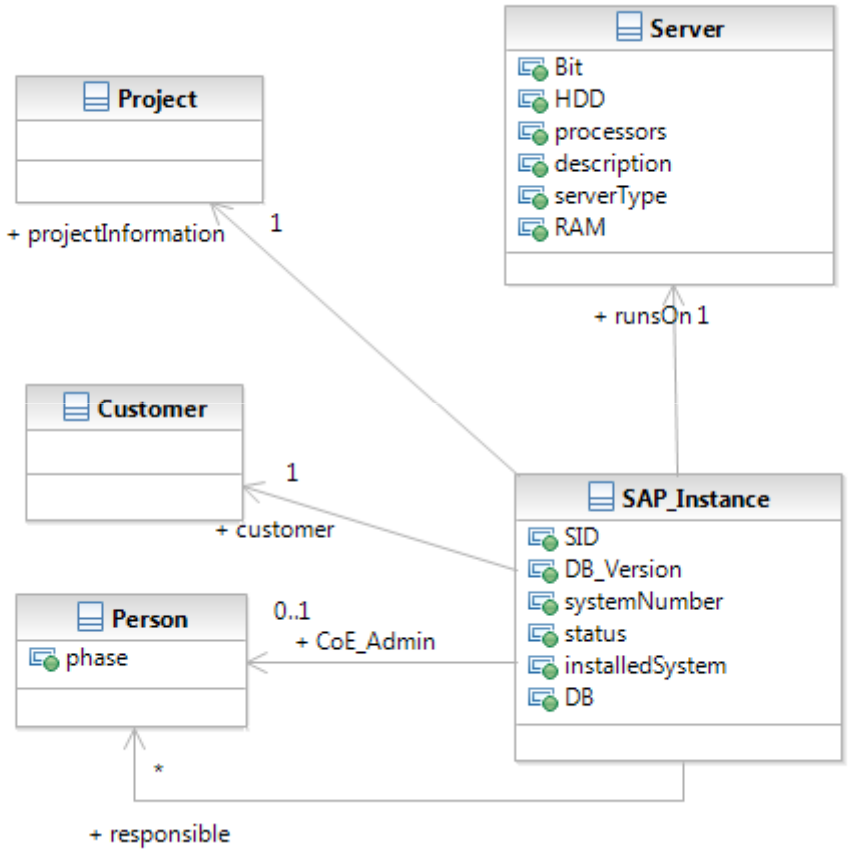


Modeling Data Sources using Ontologies

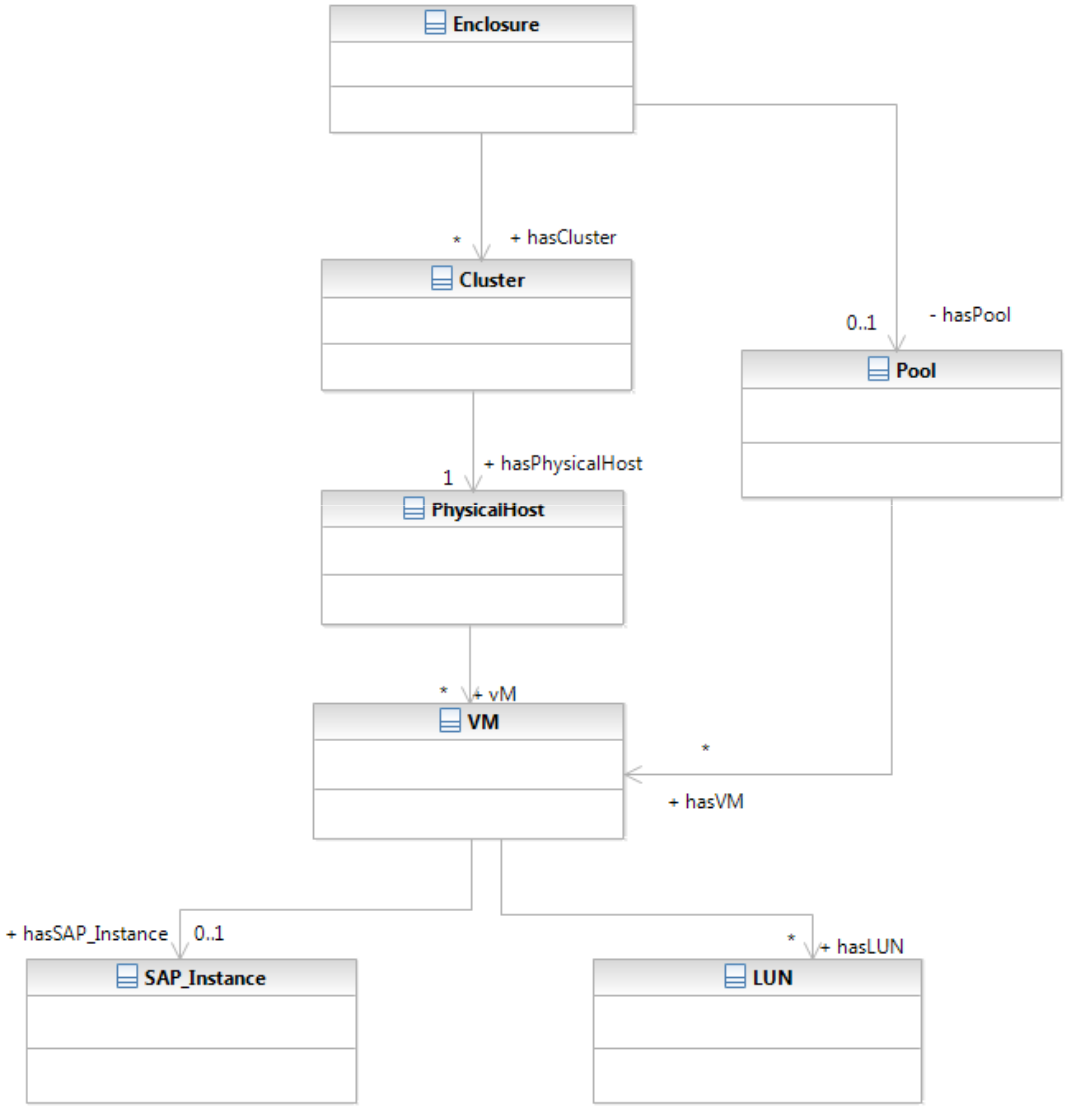
1. Reverse Engineering of data sources using domain knowledge
2. Definition of mapping of the source data model to the ontology
3. Population of the target ontology from the source instance data
   - One-time extraction for static data
   - Using caching and regular updates
   - Or runtime/online access for dynamic data sources

- Typical Problems:
  - Non-normalized sources
  - Imperfect data:
    - Syntactic heterogeneity (no keys)
    - Missing information
    - Obsolete information

Sample Ontology for a Data Source

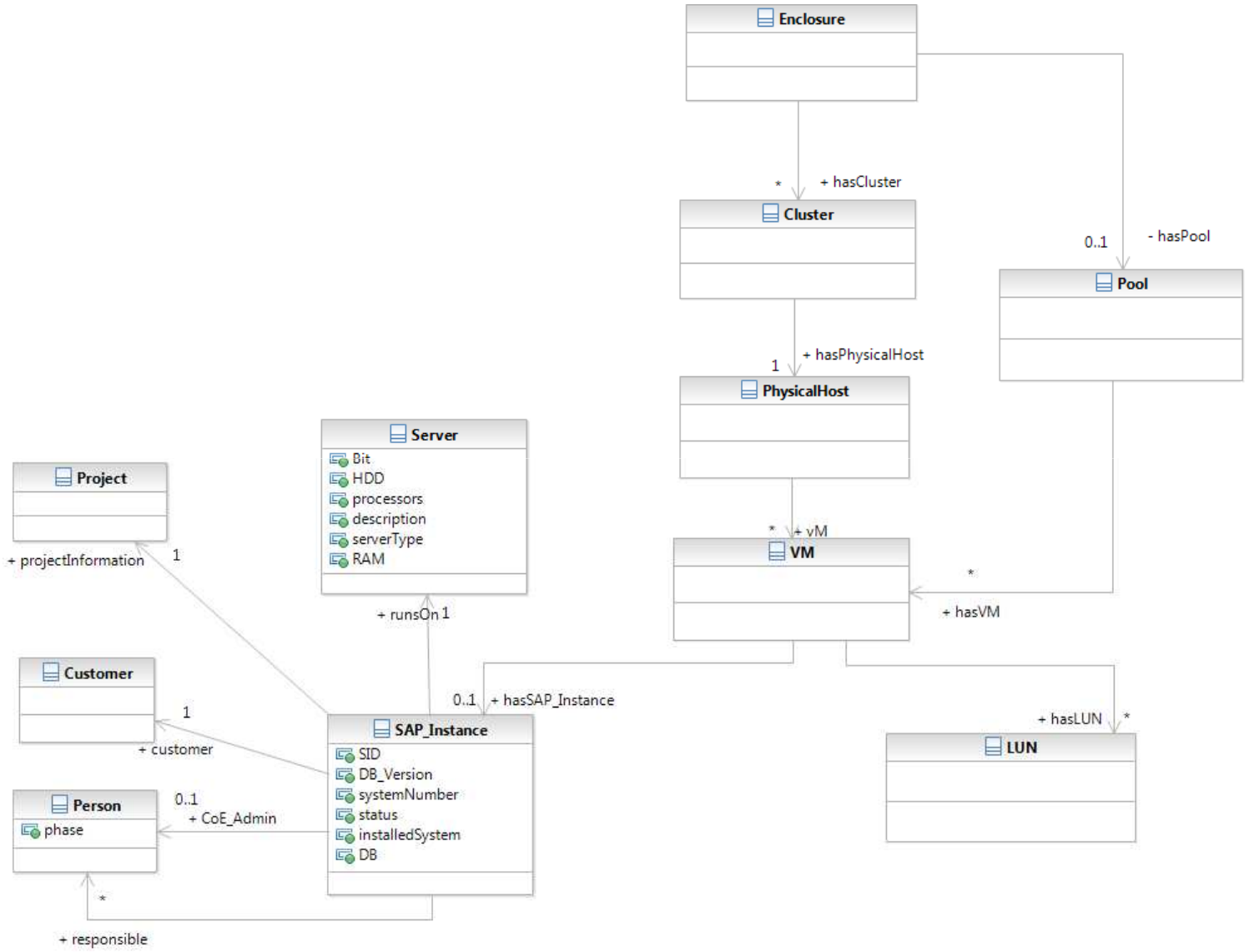


Cloud Management Ontology: Landscapes



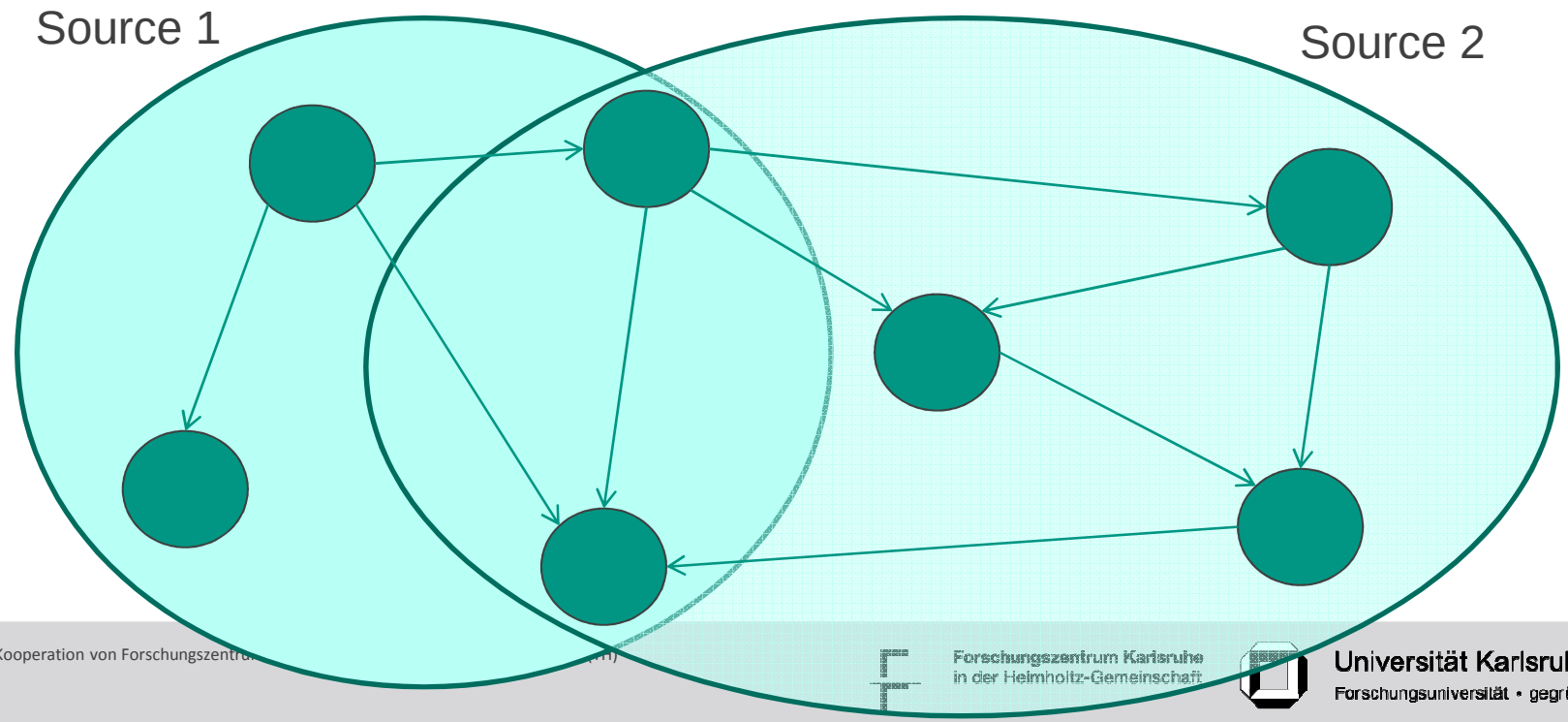
- Integration of data from different APIs (Providers)

Integrated Ontology (Fragment)



Managing Provenance of the Data

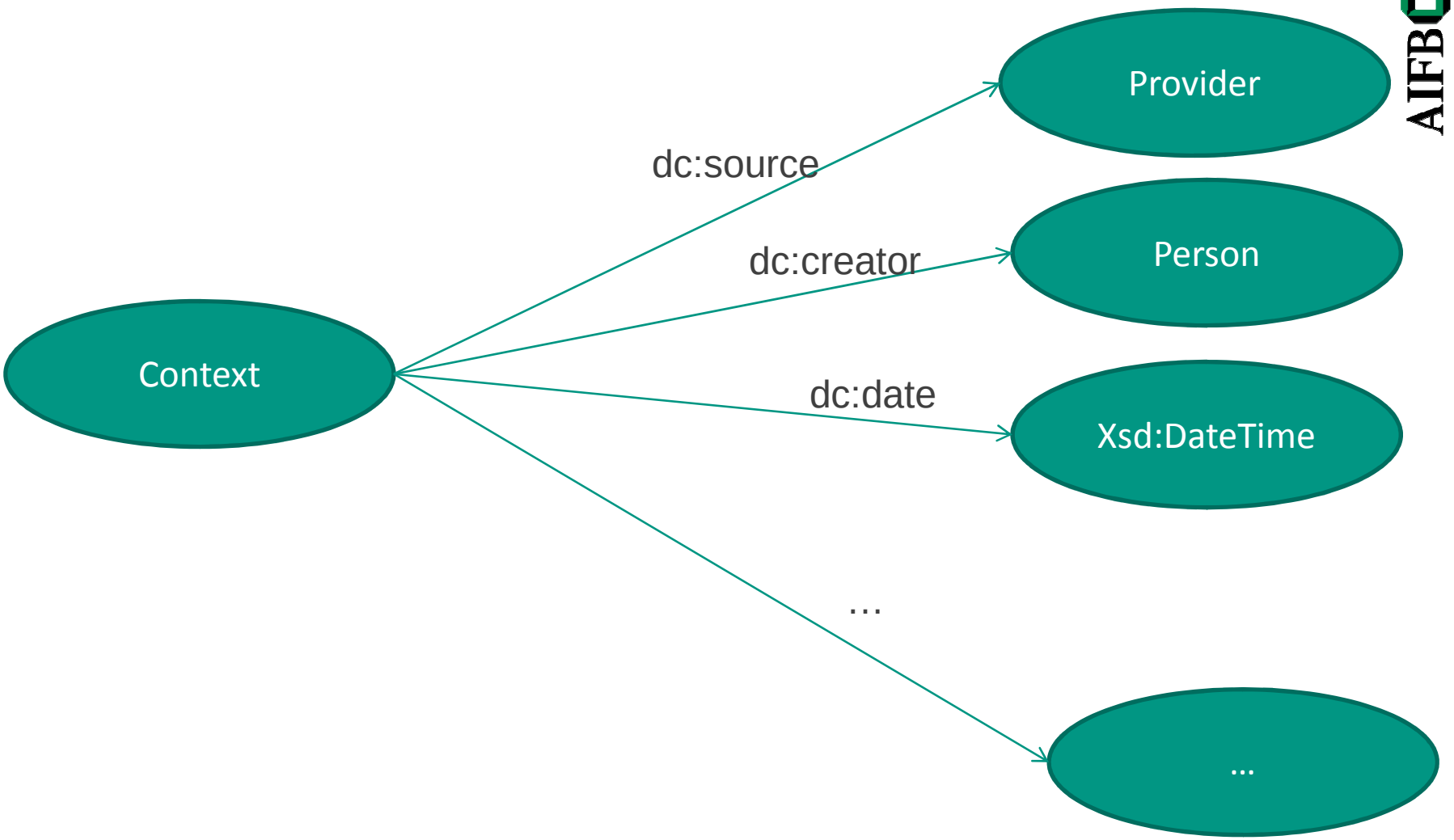
- Need to manage metadata about the origin of the data: where it comes from, when it was created, …
- Contexts can be used to group sets of statements that logically belong together, for example because they come from the same source



RDF Data and Contexts

- Quadruples instead of triples
- (subject, predicate, object, context)
- Context is just another resource about which additional statements can be made

Using Context to Express Provenance



Demo

Content: 'iwdf1081'
phaase Edit Widgets...

Search

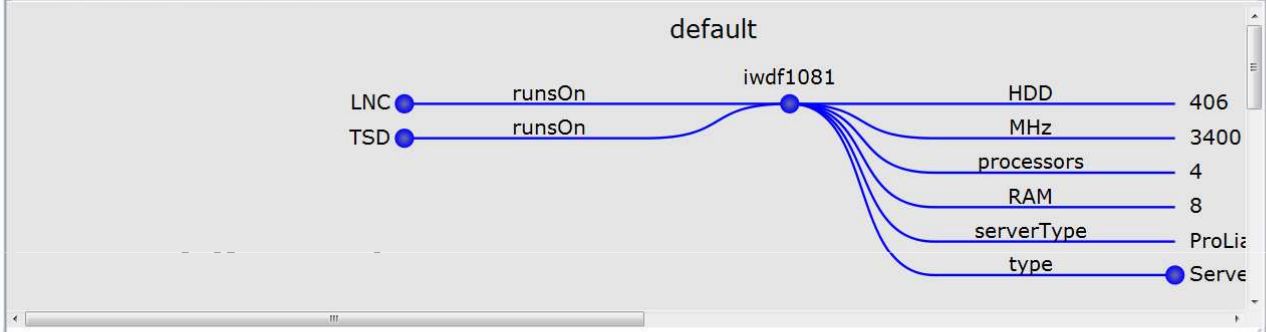
fluid Operation

Wiki for *iwdf1081*

View Blog Edit Revisions

(No text defined for this topic)

Graph for Subject *iwdf1081*



Add new Facts to *iwdf1081*

Define a new RDF statement
Predicate:
Object:
Add

Upload a document
Durchsuchen...
Select document to upload

Outgoing Statements from *iwdf1081*

Outgoing Statements from *iwdf1081*	
serverType	ProLiant DL380 G4
HDD	406
RAM	8
type	Server Server
processors	4
MHz	3400

Incoming Statements to *iwdf1081*

Incoming Statements to *iwdf1081*	
runsOn	LNC TSD

Facts from this Provider

Facts provided by this model

DataSources *iwdf1081*

systemsSAP.groovy
systemsHW.groovy

Semantic Wikis: From unstructured to structured content

- Wikis are well-established tool for annotation and documentation
- Do not enforce schema and structure up-front
- Semantic annotations allow seamless introduction of structure and semantics in an iterative manner

Intelligent User Interfaces: Widgets for Semantic Data

- Situation:
  - Either no schema available, or schema dynamically changing
- Idea:
  - Every widget is able to **display certain properties** of an entity
  - Widgets are **selected based on the properties** of an entity (types of triples)
  - Widgets can dynamically integrate **external data sources**
  - Widgets may allow for different types of **interaction**

Demo: Part 3

Karlsruhe Institute of Technology

Content: 'TATA'

haase Edit Widgets...

Search

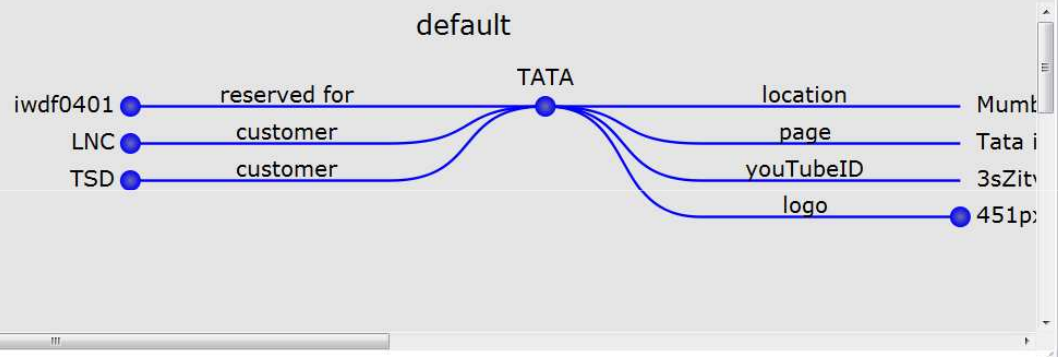
Wiki for *TATA*

View Blog Edit Revisions

Tata is located in Mumbai

Image *TATA*

Graph for Subject *TATA*



YouTube *TATA*

Outgoing Statements from *TATA*

Outgoing Statements from *TATA*	
youTubeID	3sZitve3SUw
logo	451px-tata_logo.png
location	Mumbai
page	Tata is located in [[location::Mumbai]]

GMap *TATA*

Add new Facts to *TATA*

Define a new RDF statement

Predicate:

Object:

Add

Upload a document

Durchsuchen...

Select document to upload

Summary

- Challenges in Cloud Management
  - Heterogeneous data sources covering technical and business aspects
  - Integration of structiured data and unstructuted annotations required
  - Loose and dynamic schemas

- Semantic Technologies for Cloud Management
  - Ontologies enable flexible information integration
  - Semantic wikis allow for (semi)-structured collaborative annotations
  - Widgets for semantic data allow for intelligent user interfaces

Program for students: fluid experience

- Internships / Working Students
  - Gain practical experience with cutting edge technologies
  - Professional working environment
  - Close interaction with the development and management staff
  - Attractive compensation
  - Flexible start dates and durations
  - Gourmet cafeteria, free soft drinks

- Thesis Students (Bachelor, Masters, Diplom)
  - Challenging research topics with practical relevance
  - Top class supervision and sponsorship of theses
  - Excellent chances for transfer of results into innovative products

fluid experience – Hot Topics

	Project topic
1	Scalable semantic data management in the cloud
2	Mashups and information integration in cloud-based environments
3	Managing internal IT like a cloud
4	Infrastructure-as-a-Service
5	Compute, storage and network virtualization
6	Load balancing and failover management in enterprise compute clouds
	… and many more!

Contact us: students@fluidops.com