



Information Workbench for Interacting with Linked Data

Peter Haase

www.fluidops.com

fluid Operations GmbH



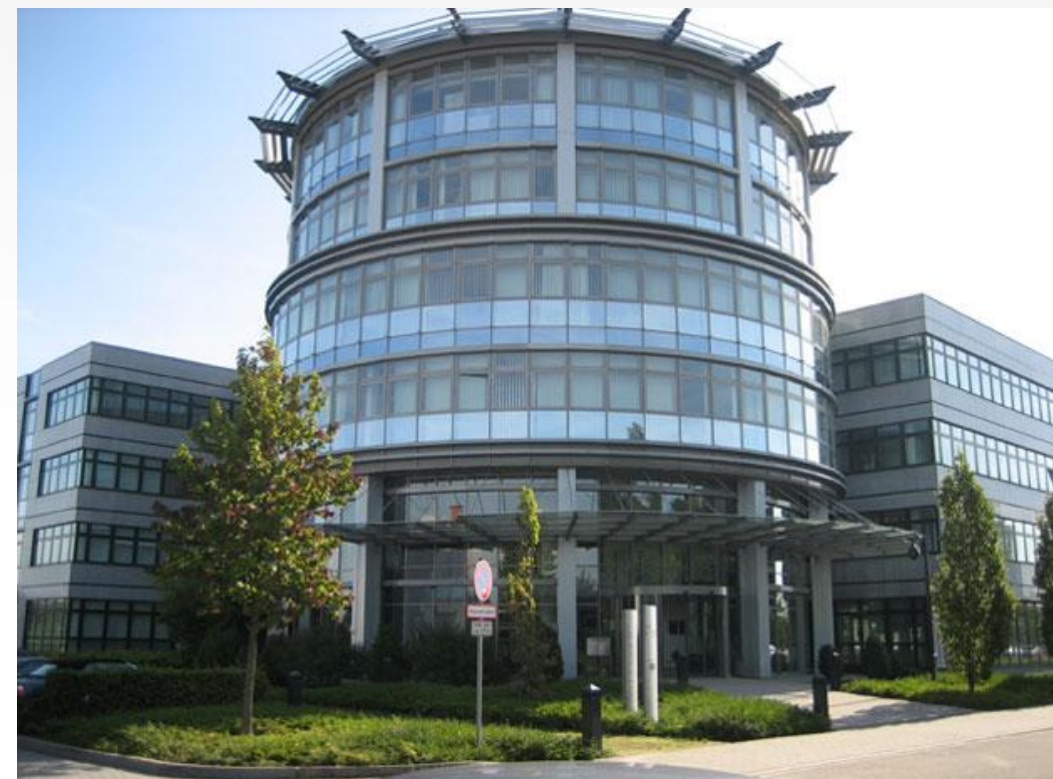
Providing management solutions for enterprise clouds

Founded Q1/2008

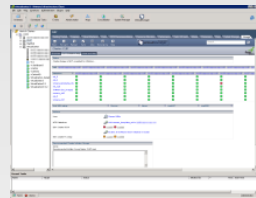
Headquarters in Walldorf / Germany, SAP Partner Port

Privately held, VC funded

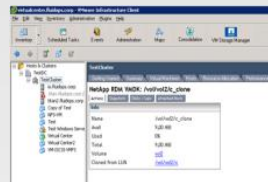
Currently 20 people, growing rapidly



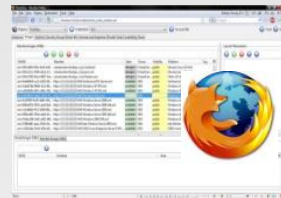
eCloud - Solution View



Web-UI



Native Plug-ins



Self Service Portal



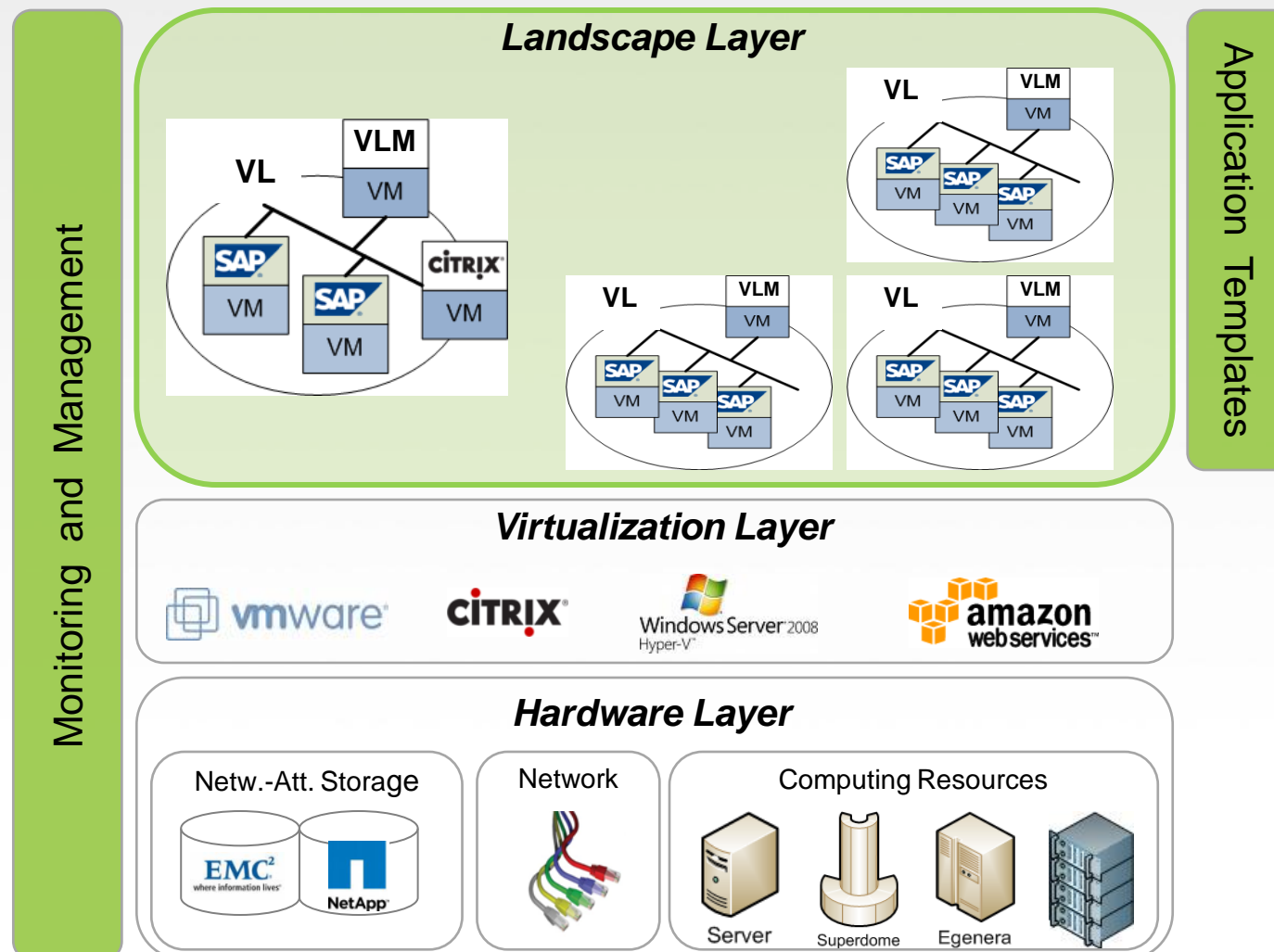
Dashboards

Business layer

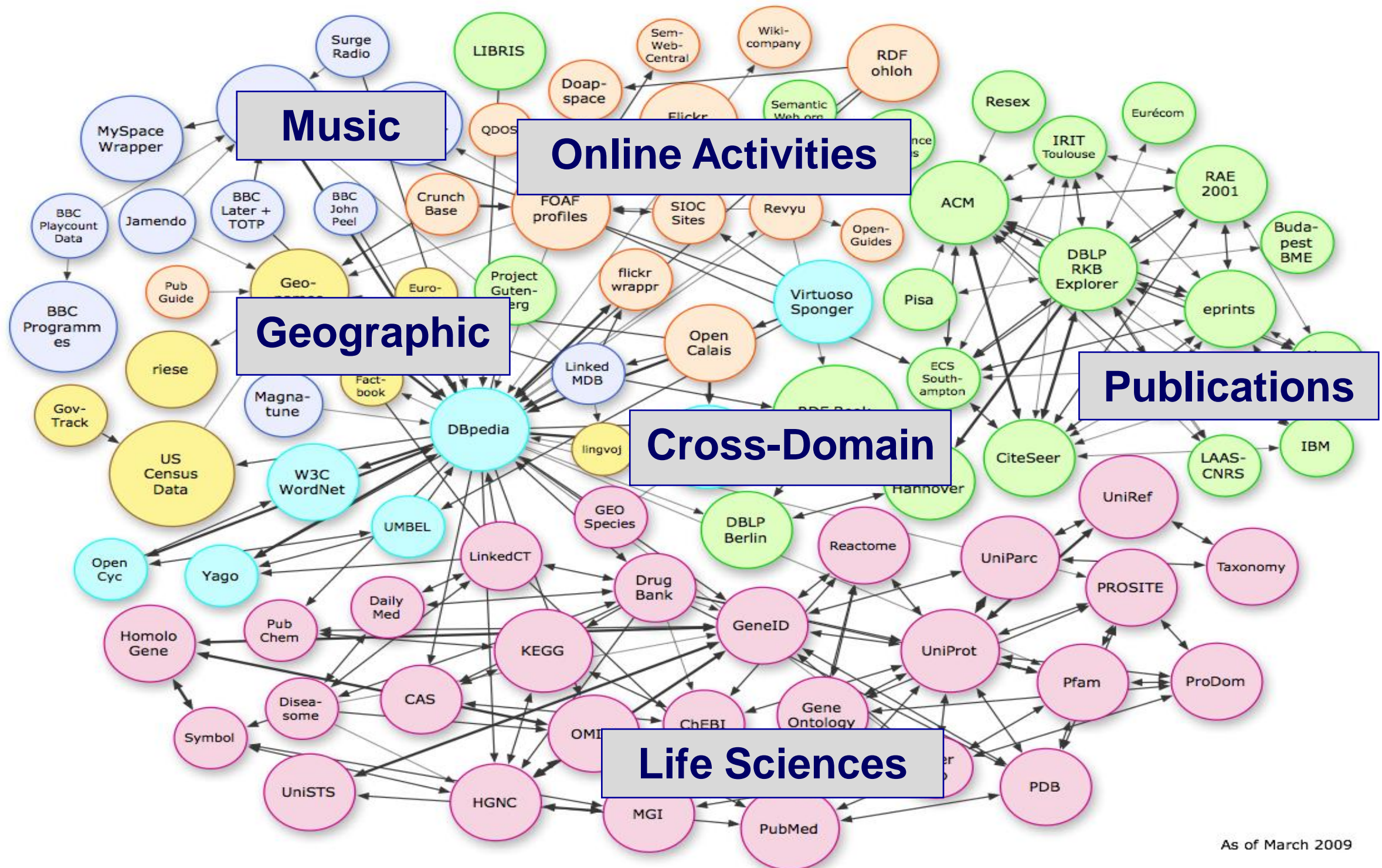
- Admin, business client & business sponsor cockpits
- Integrated monitoring and system management along the entire stack
- Landscape-level provisioning (LaaS)TM with unique application template support & automated cloning
- Unprecedented application development, testing and production productivity

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



Linked Data



Information Workbench



Addressing the entire **lifecycle of interacting** with the data

- Integration of data sources
- Content generation by the end user
- Provenance
- Search and Exploration
- Visualization
- Publishing

Integrated management of **heterogeneous data sources**

- Structured and unstructured
- Published and user-generated
- Static and dynamic
- Open domain and domain specific
- Public and enterprise

Information Workbench Architecture

Widget-based
User Interface

Semantic
Browser

Semantic
Wiki

Custom
Widgets

Data
Integration
& Indexing

Search and
Query
Processing

Presentation,
Interaction and
UI Customization

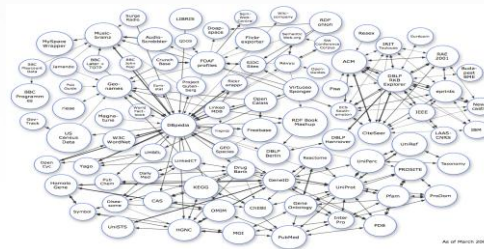
Semantic Data Store

Triple
Store

Unstructured
data

Keyword &
Structure
Index

Data
Provider



Semantic Search



Schema-agnostic: Expressive queries without knowledge of the schema

Hybrid Search: Structured queries combined with keywords across structured and unstructured data sources

Query interpretation: Translation of keywords into hybrid queries (conjunctive queries with structured and unstructured elements)

Faceted search: Iterative refinement process based on keywords and operations on facets

User Interfaces and Interaction Paradigms



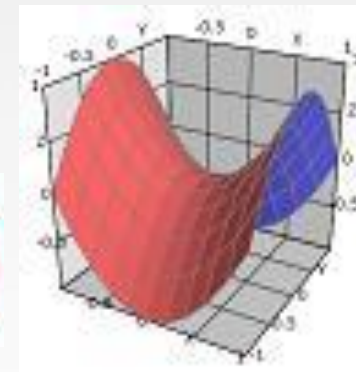
- How to realize a continuous, seamless and personal user experience across domains?
- Interfaces that operate over large amounts of data
- Meaningful aggregation of the data
- Data-driven composition of interface elements
- An interface that reacts to changes in the available data
- Schema-agnostic definition of the user interface

Widget-based user interface



Multiple paradigms for interaction: browsing, visualization,
editing, knowledge acquisition

Mashups with external sources



Automated selection of widgets based on available data

Customization and personalization

Demo: Data Sources in the Application



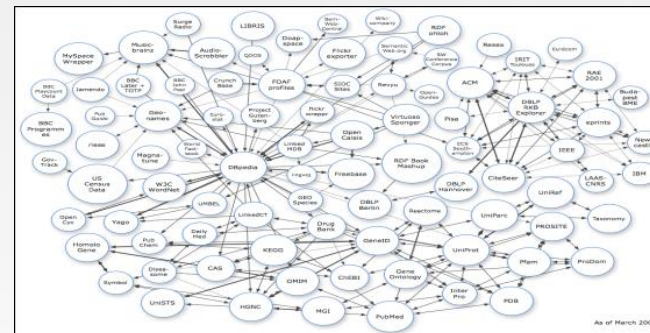
Entire English Wikipedia



WIKIPEDIA
The Free Encyclopedia

Data from Linked Open Data

- DBpedia
- YAGO
- ...



Data from Data.gov (US Government)

- E.g. live data about earthquakes

Many more



<http://iwb.fluidops.com/>



fluid

Operations

Operations

fluid

CONTACT US:

fluid Operations
Altrottstr. 31
Walldorf, Germany

Email: info@fluidOps.com
website: www.fluidOps.com
Tel.: +49 6227 3849-567

Semantic Web Appliances

Rapid Provisioning of Semantic Applications



Sample use case: I want an application to interact with biomedical datasets from LOD

- Application setup in the cloud
- Populated with relevant datasets

In <1 hour

- Extended with custom providers, widgets

In <1 day

1. Provisioning of infrastructure: storage, computational resources (IaaS)

- Handled by eCloudManager

2. Provisioning of information management platform (SaaS)

- Deployment of Information Workbench on arbitrary target infrastructure
 - E.g. Amazon EC2, VMware Image, ...

3. Provisioning of datasets (DaaS)

- Efficient and intuitive integration of different data sources
- Transparent federation

Data-as-a-Service



- Current way of publishing data:
 - Providing RDF dumps
 - Possibly SPARQL endpoint
- State-of-the-Art to integration/federation

Federation of SPARQL endpoints (DQP)	Centralized repositories (Warehousing)
Poor performance	Good performance
Limited set of services	Open for extensions
Read only	Write access possible (in theory)
Always up-to-date	Update problem
Selection of data sources possible	No selection of data sources

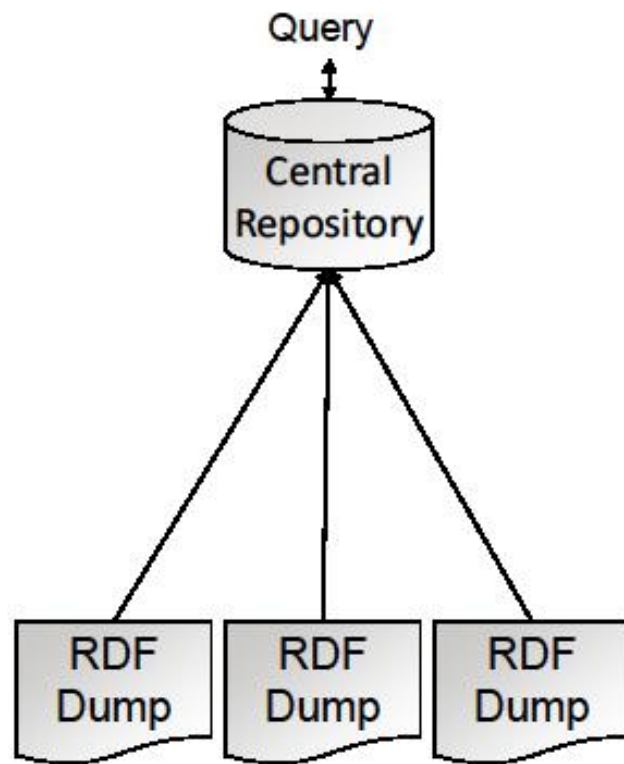
- Application requirements may differ
 - How long is the federation intended to exist (ad hoc queries vs. Long lived applications)
 - Are updates to the data required (Read only vs. Write access)
 - Processing capabilities (CPU, storage)
 - Characteristics of the queries: Types, frequency
 - How fix is the set of data sources

Data Components

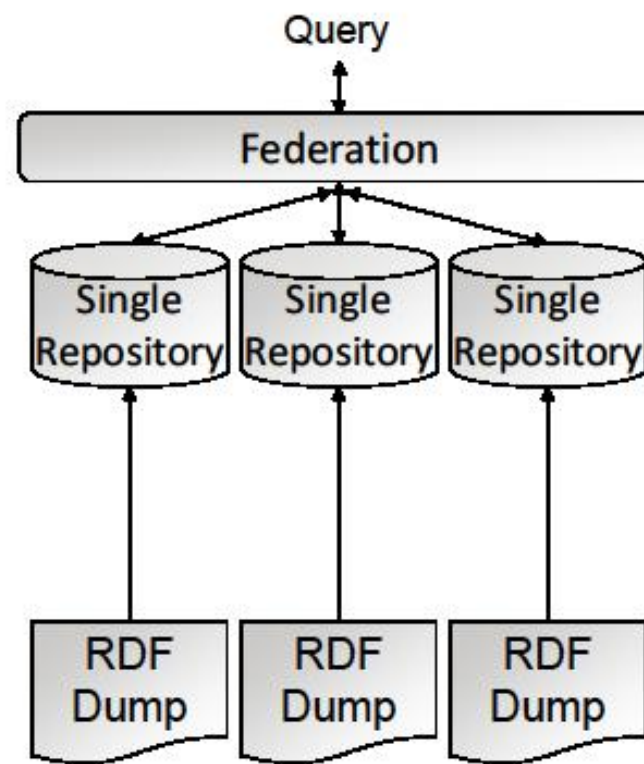


- Definition of Software Components:
 - "A software component is a unit of composition with **contractually specified interfaces** and explicit context dependencies only. A software component can be **deployed independently** and is **subject to composition** by third parties." (wikipedia.org)
- Applied to data components:
 - Interfaces: Components with precise interfaces and metadata
 - Deployment: Easy provisioning and integration in applications
 - Composition: Transparent access to atomic or composite units

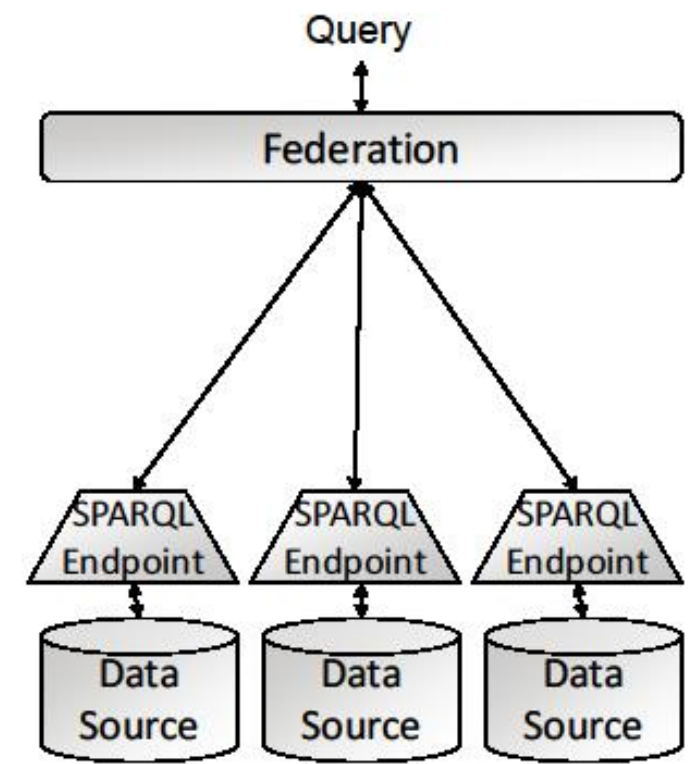
Alternatives for federation



a) Integration in a central repository



b) Federation over multiple single repositories



c) Federation over multiple SPARQL endpoints

Linked Data Federation

