



# Collibra

## Business Semantics-driven Integration of Service-oriented Applications

SOA Symposium, April 21 2010

Collibra nv/sa

<http://www.collibra.com>

Ransbeekstraat 230, 1120 Brussels, Belgium

Felix Van de Maele, CEO

[felix@collibra.com](mailto:felix@collibra.com)

+32 475 68 18 12

# Context

- ▶ **Felix Van de Maele**, co-founder and CEO of Collibra, commercializing + 100 person years of semantic technology research
- ▶ “Business Semantics-driven integration of Service-oriented Applications”
  - ▶ **What:** Aligning people and systems, from business definitions to service implementation
  - ▶ **How:** Open Standards (SBVR, XSD, UML, WSDL, OWL, ...)
  - ▶ **Why:** Data Governance
- ▶ Based on several **government cases:**
  - ▶ Flemish Government of Economy, Science and Innovation
  - ▶ Flemish Ministry of Education and Training
  - ▶ Fedict (Belgian Federal public service for ICT, responsible for e-Government and the Federal Service Bus)



An image can take the place of a word in a statement:



An object is not so attached to its name that one cannot find another for it which suits it better:



An object never performs the same function as its name or its image:



*Ceci*

*ripe.*

# What is ...

asked to >20 people in  
the Ministry of  
Education and Training

“Study Area”



sample from an inquiry of >20 people in the  
Ministry

# Well, it means ...

- One of the 16 categories mentioned in article 19 in which academic trainings are bundled (decree concerning universities)

*In public administrations the decrees dictate the semantics of business terms and how they relate to each other.*

*This business context is a crucial source to give meaning to metadata.*

- A group of structural components based on content-wise relations in technical and vocational education, as well as , in [...] (decree secondary education)



# Mind the Gap

Business

*Business communities, IT demand side*

**data utilisation:** how is my data utilised?

**data governance:** who is responsible for my data?

"Country Belgium  
a product last year"

"Person that buys our products  
with a certain frequency"

Manage the  
business context

"Customer"

**data quality:** how is my data evaluated?

**data provenance:** where does my data come from?

Operational/  
Technical



CRM



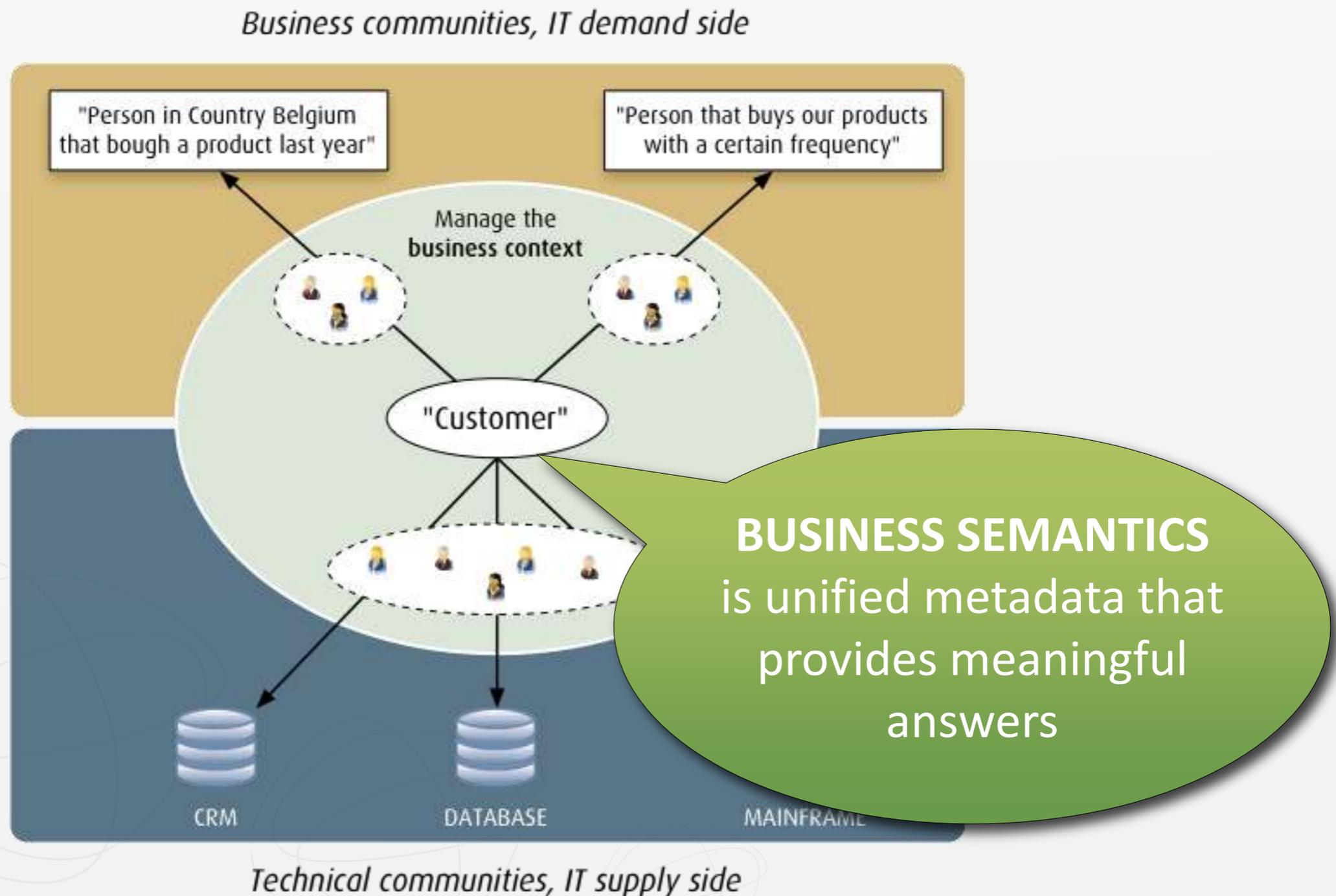
DATABASE

*Technical communities, IT supply side*

**data semantics:** what does my data mean?



# Bridging the Gap



## Political Sector Training and Education

Minister

Policy Council

*It is not about the metadata format.*

*The question is: How to deal with metadata in large organizations that need versatility in cross-referring business vocabularies ?*

Department of Education

Agency for Education Services

Agency for higher education

Agency for educational communication

Agency for quality insurance in education

Agency for educational infrastructure

# Political Sector Training and Education

Community that understands the same things and shares some terms

Minister

Policy Council

Community that understands the same things

Community that understands the same things

Department of Education

Agency for Education Services

Agency for higher education

Agency for educational communication

Agency for quality insurance in education

Agency for educational infrastructure



# Political Sector Training and Education

Community that understands the same things and shares some terms

Minister

Policy Council

Community that understands the same things

Community that understands the same things

Department of Education

Voc 1

Agency for

Voc 2

Agency for higher education

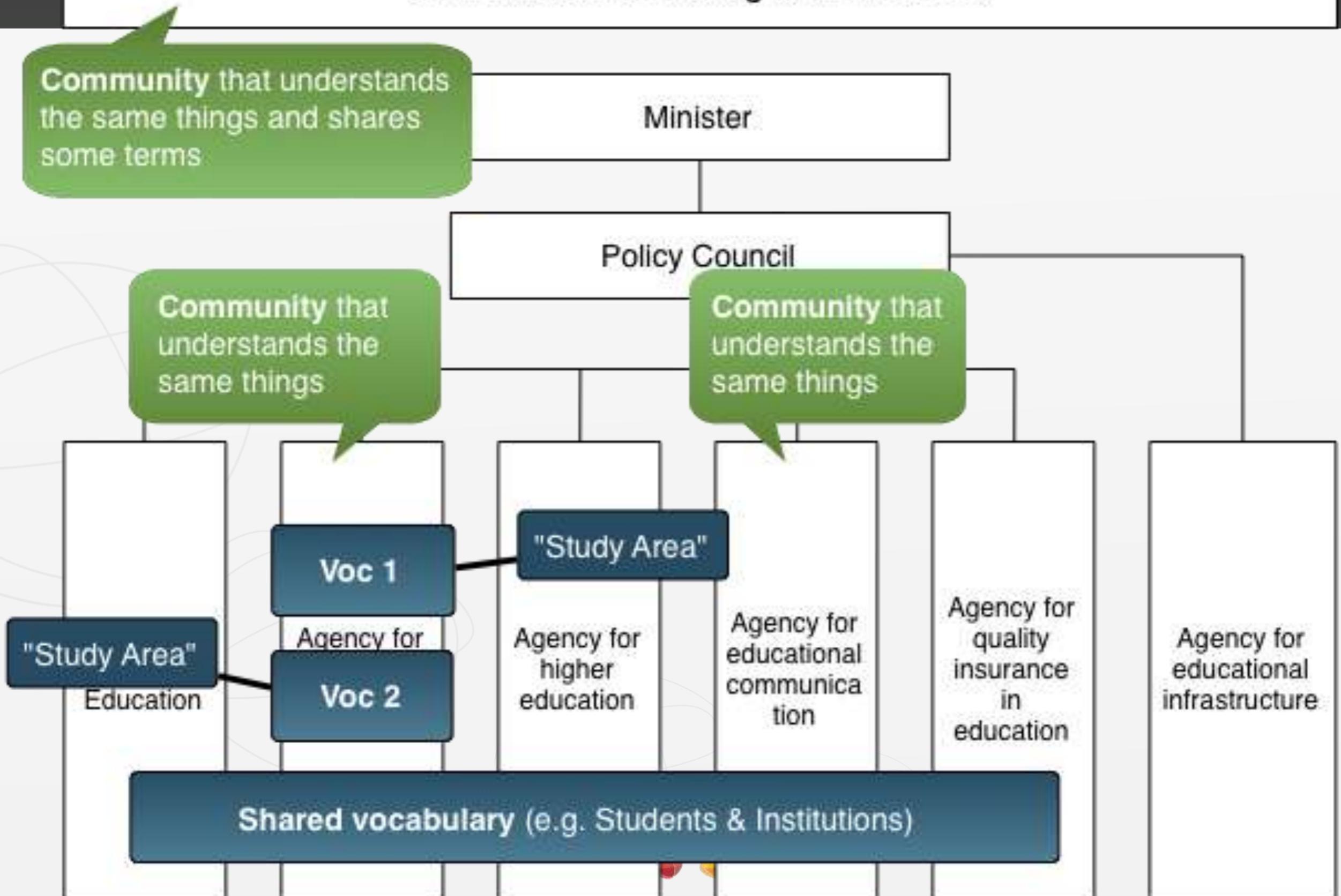
Agency for educational communication

Agency for quality insurance in education

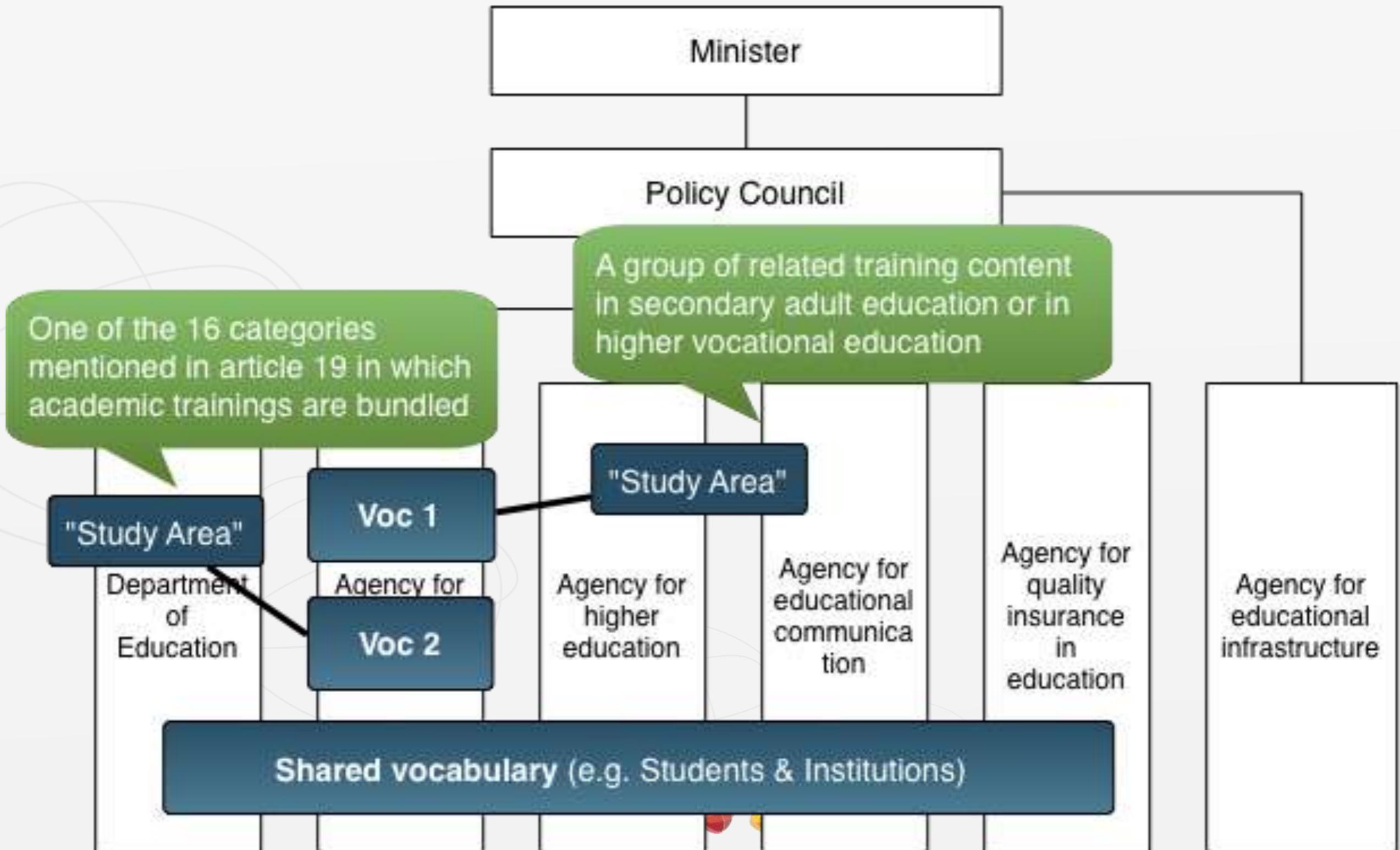
Agency for educational infrastructure

Shared vocabulary (e.g. Students & Institutions)

# Political Sector Training and Education



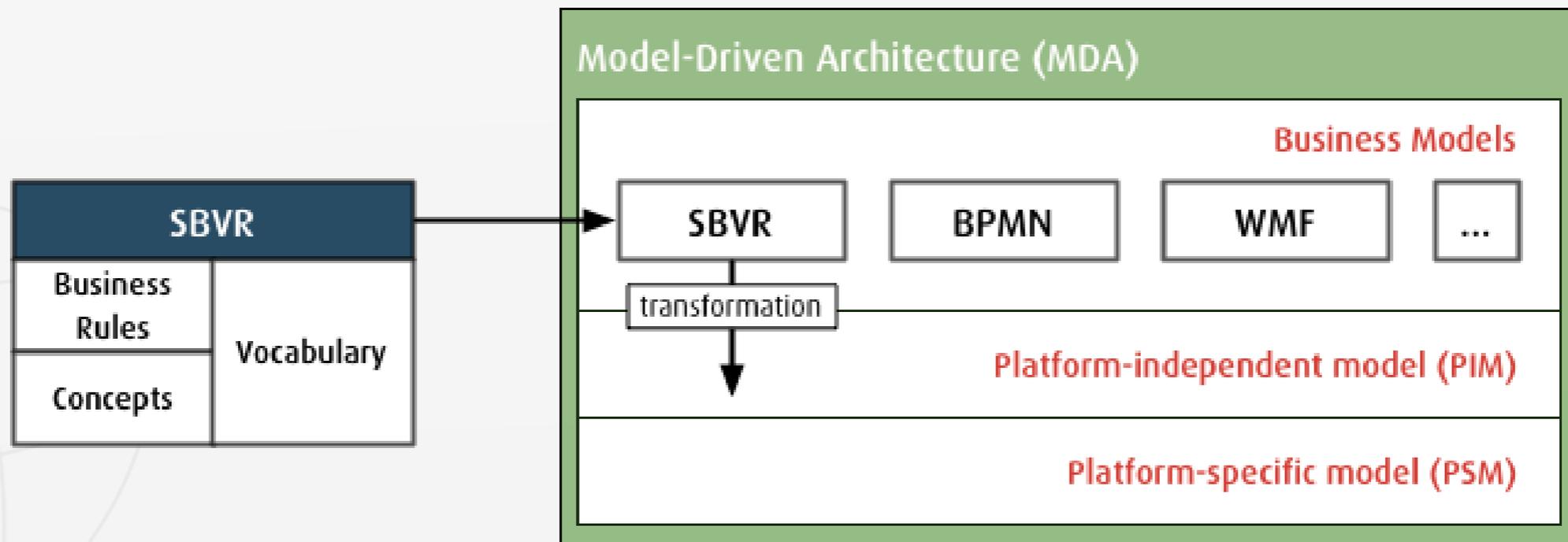
# Political Sector Training and Education





# SBVR Standard positioning

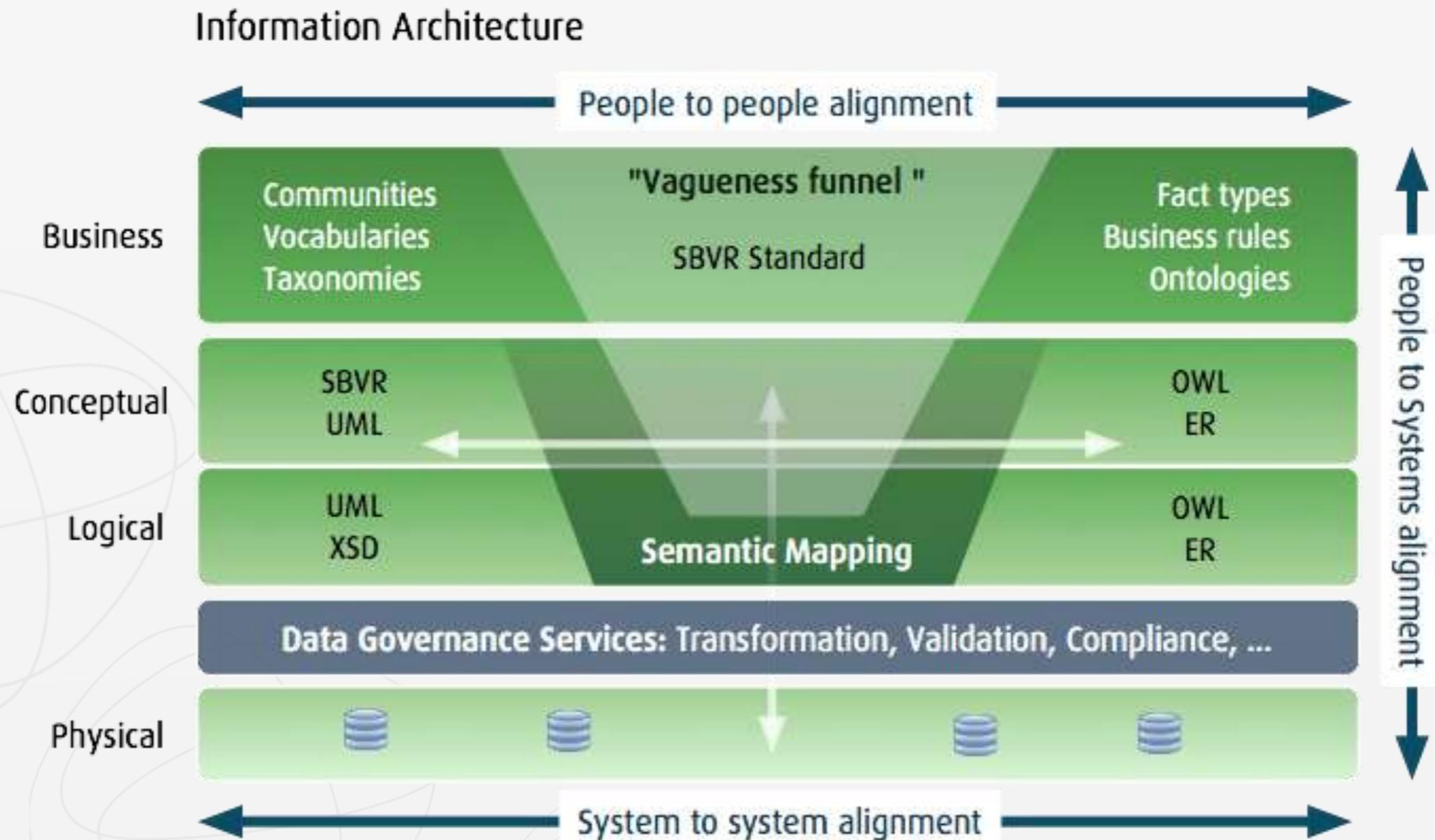
## OMG's Model-driven Architecture



- ▶ Terminological ontology
- ▶ Community-driven  
(*Semantic Community, Speech Community*)
- ▶ Rule-based approach
- ▶ Supports time-changeability
- ▶ Supports semantic stability
- ▶ Enables extensibility and reuse
- ▶ Business-friendly



# Information Architecture

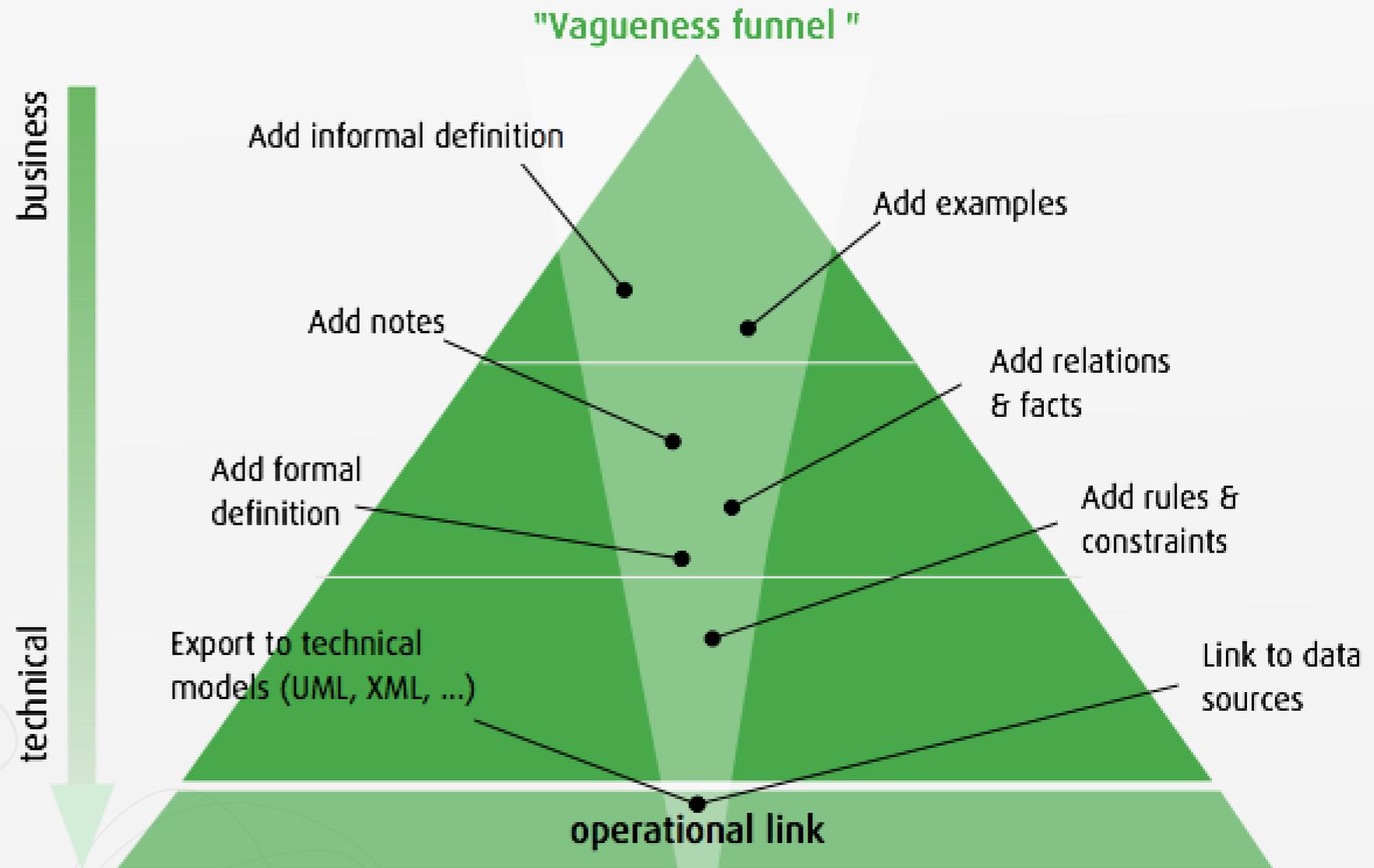


- ▶ *Combine enterprise 2.0 best-practices (community-driven, business-friendly) with semantic technology (integration & alignment)*



# Involve all stakeholders

- ▶ *Divide & Conquer*
- ▶ *Let everyone do what they know best*
- ▶ *Aligned decoupling*

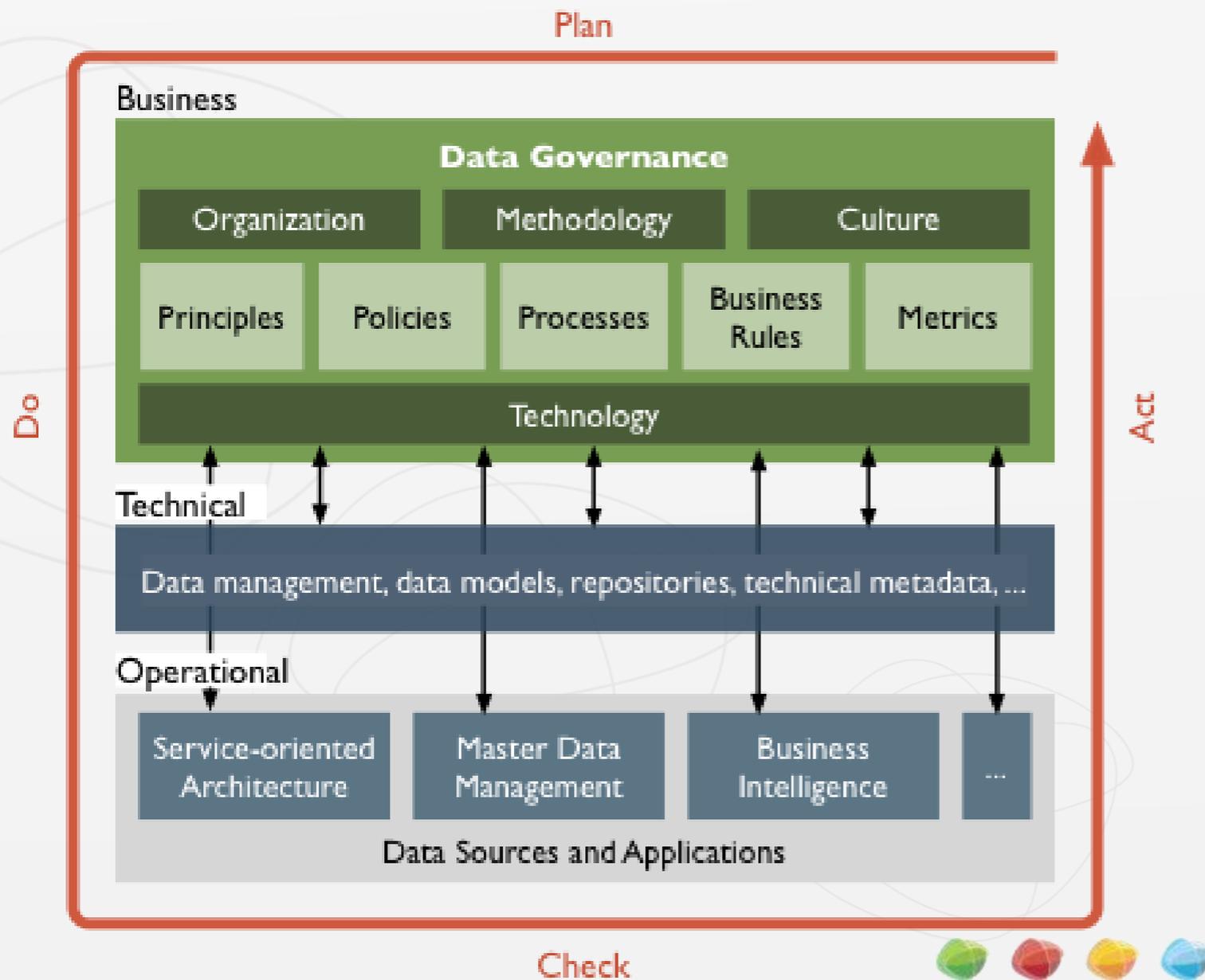


Basically, all we are using is RDF schema.



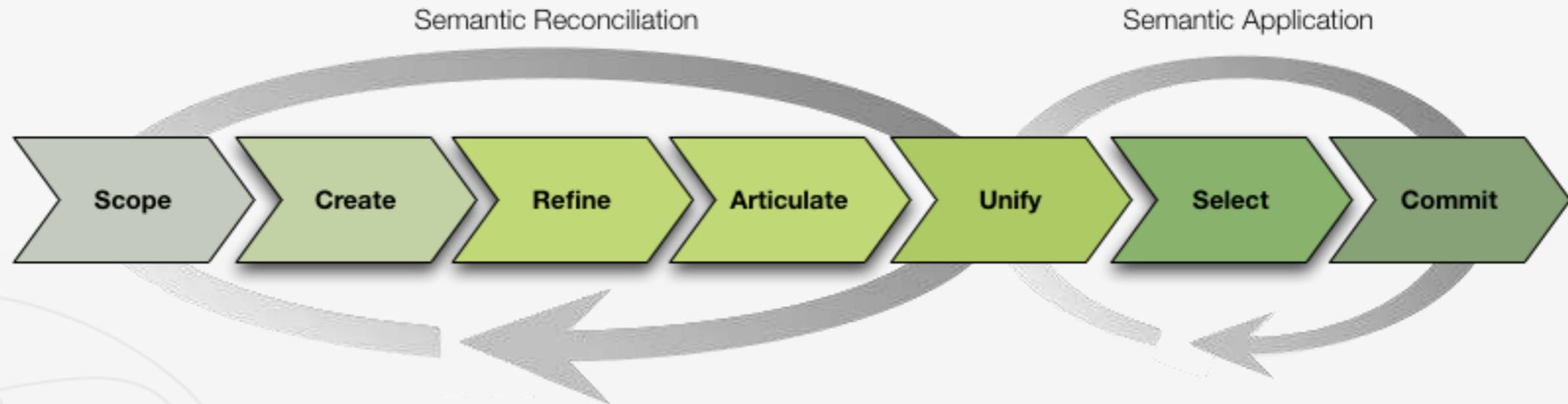
# Data Governance

**Data Governance** seeks to define and control the principles, policies, processes, rules and metrics needed to make sure enterprise data is in a consistent and useable form when it is consumed by people or systems.



*“If management is about running the business, governance is about seeing that it is run properly”*  
 (Robert I. Tricker)

# Agile Methodology



- ▶ **Two processes:**
  - ▶ Semantic Reconciliation
  - ▶ Semantic Application
- ▶ **Divide & Conquer:** *Involve all stakeholders at their own comfort level*
- ▶ **Decoupling:** *avoid a technical or business bottleneck*
- ▶ **Feedback loop:** *provide closed-loop feedback and transparency between business meaning and technical implementation*



# Fedict

- ▶ Federal Government for Information and Communication technology
- ▶ Manages the Federal Service Bus
  - ▶ Standardized exchange of data between applications
  - ▶ Access authentic data sources through web-services
  - ▶ To effectively exchange information between different government services
- ▶ Automatically generate service definitions and data models (XSD) from business definitions using specific service patterns (rules & structures)
- ▶ **Why:** Semantic Interoperability, business/IT alignment, documentation, reuse, knowledge sharing

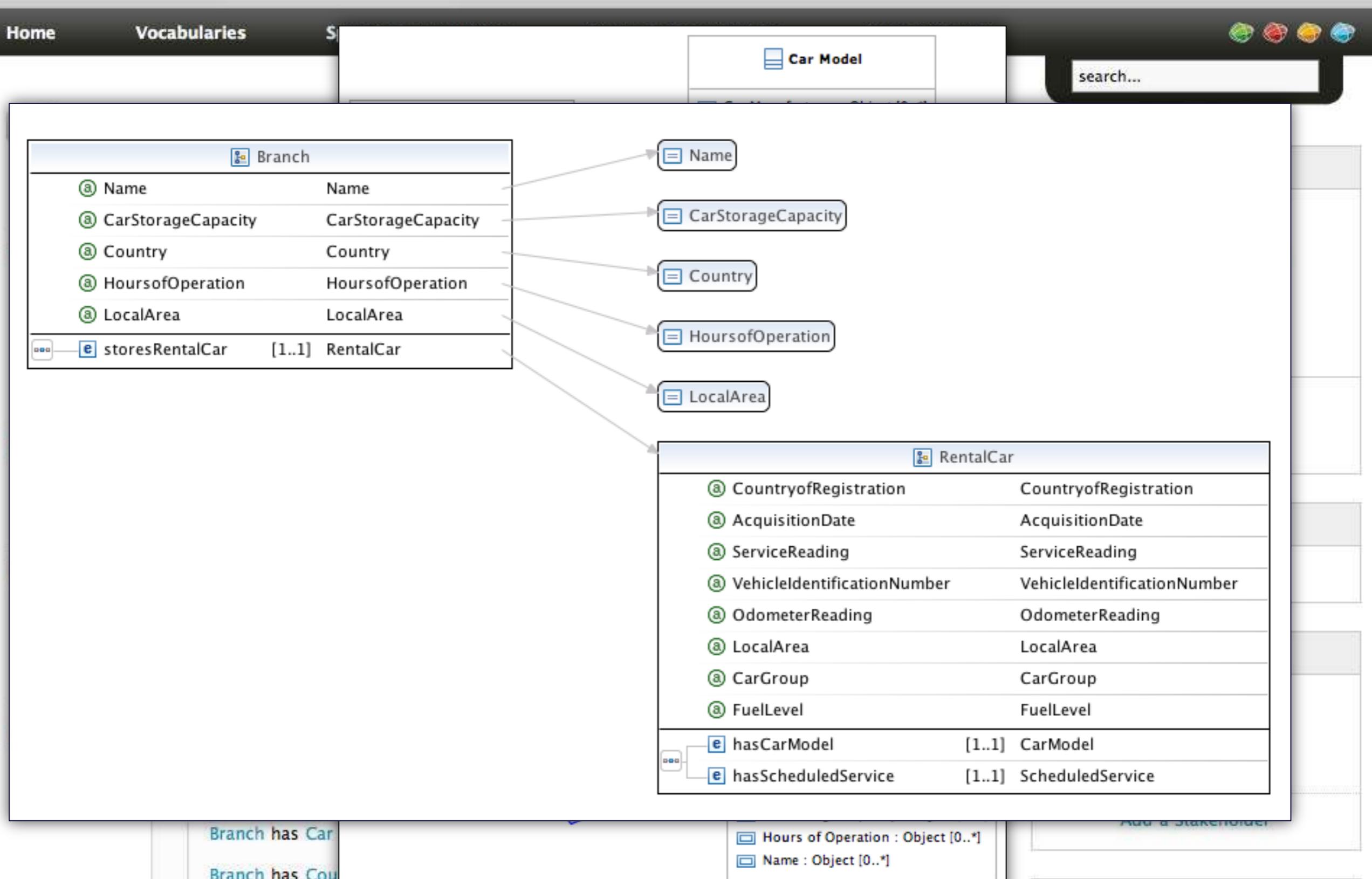


# Fl. Gov. Economy, Science & Innovation

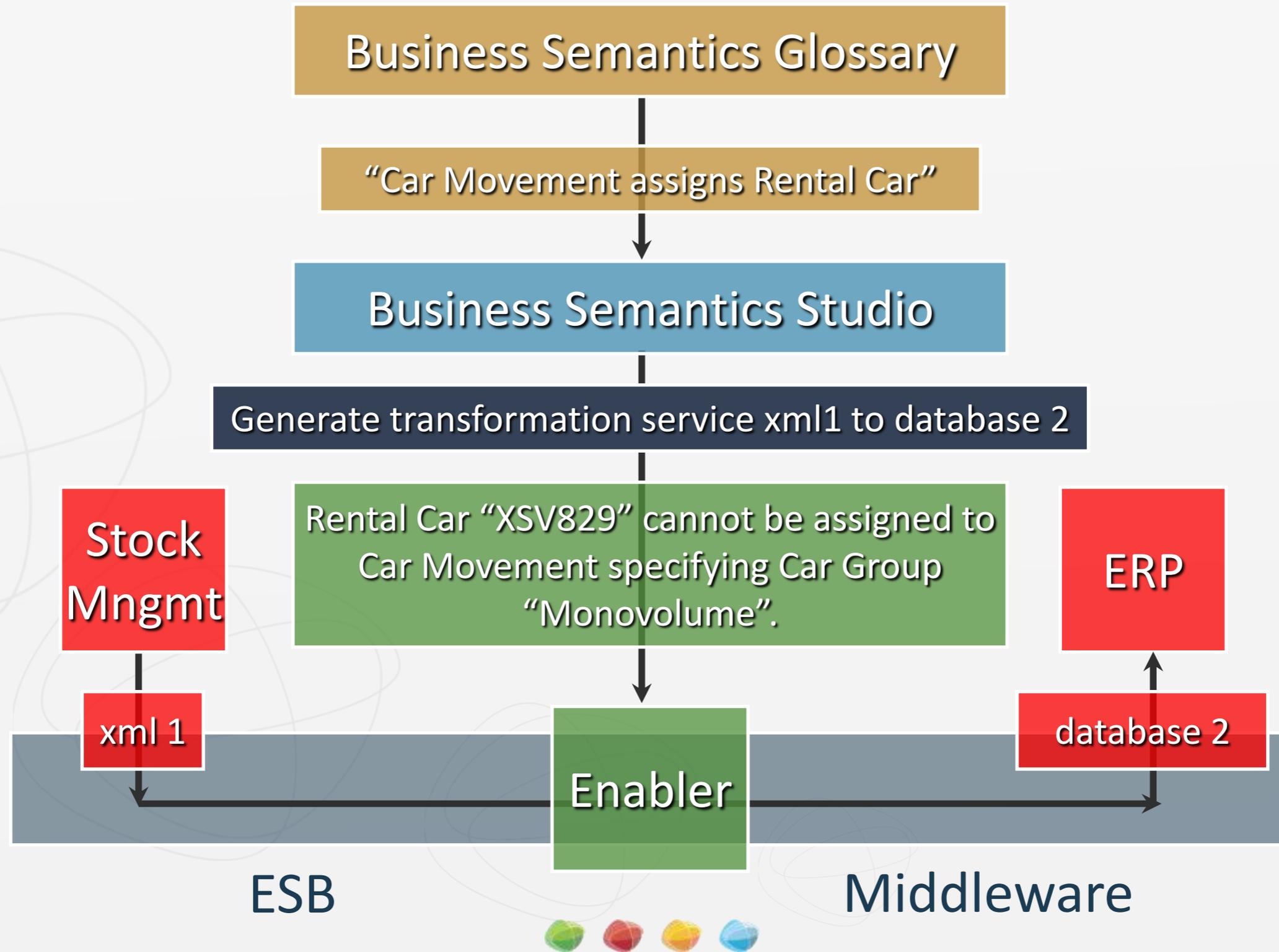
- ▶ FRIS program: Flanders Research Information Space
  - ▶ Aggregate all different data on research projects and grants from different public and private organizations
  - ▶ Validate this data against business rules (*e.g. No overlapping research grants*)
  - ▶ *Publish this aggregated data through a portal and web services*
- ▶ **Collibra:**
  - ▶ The agility to match the *business reality* with the *technical reality*
  - ▶ *Scalability & flexibility: from 6 sources now to +100 in the coming years*
  - ▶ *Ready for the future: Publish & consume linked data*



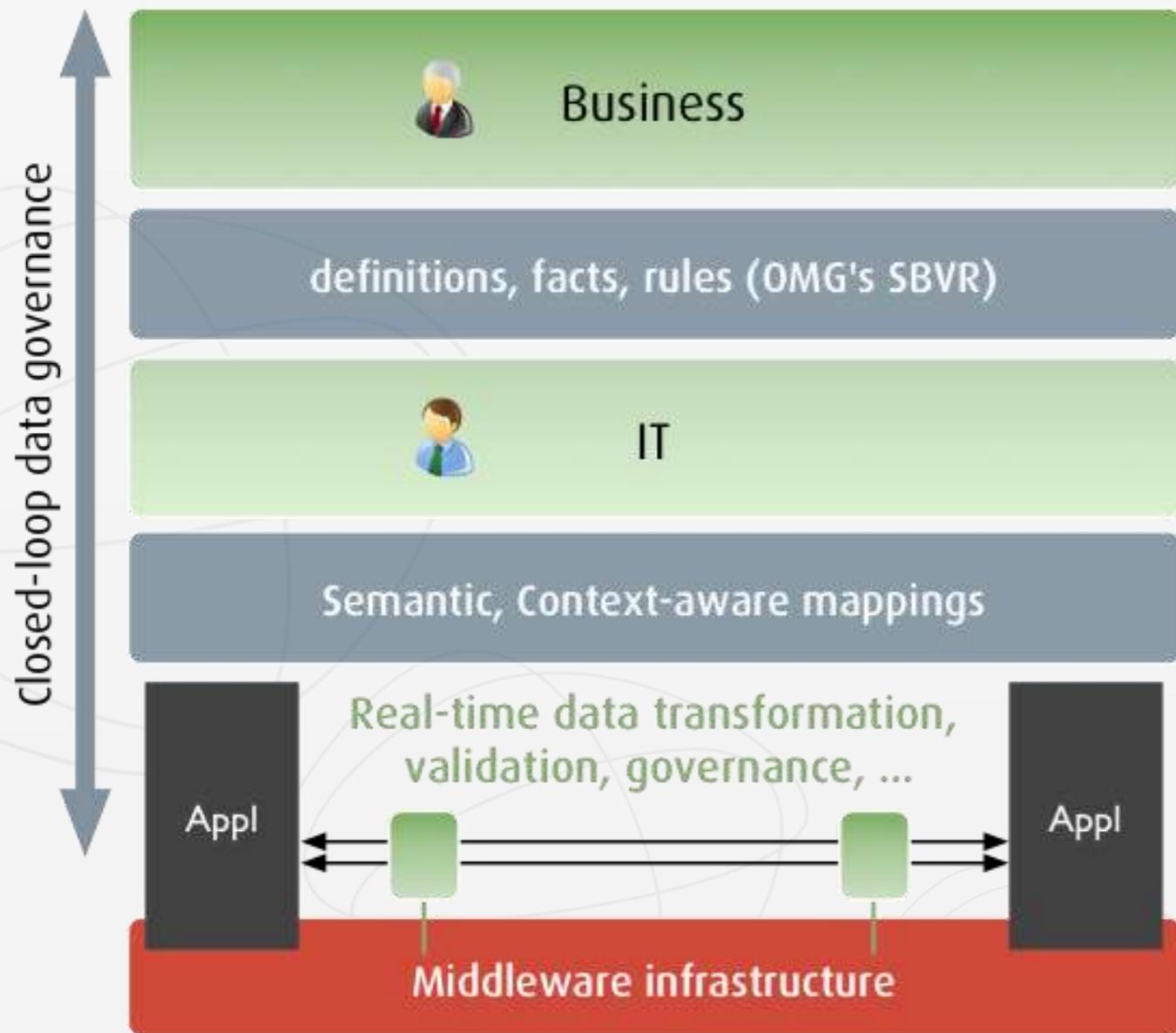
# Business – IT alignment for SOA



# Semantic Data Services



# Closed-loop alignment



- ▶ Involve business and technical stakeholders to capture and govern business definitions, - facts and - rules.
- ▶ Automatically generate technical data models (canonical models)
- ▶ Map existing data models onto the shared business definitions, facts and rules
- ▶ Full-cycle feedback loop between business definitions and technical implementation



# Conclusion

- ▶ Many SOA projects fail through ***lack of data understanding between stakeholders***
- ▶ ***SBVR*** enables domain experts to define and regulate the “business reality” using business definitions, -facts and -rules.
- ▶ These business semantics become the ***shared language*** for all human stakeholders as well as for service definitions.
- ▶ Automatic generation of technical data models and semantic mappings provide ***semantic interoperability***
- ▶ Data Services on top of your existing middleware provide ***closed loop transparency and governance.***

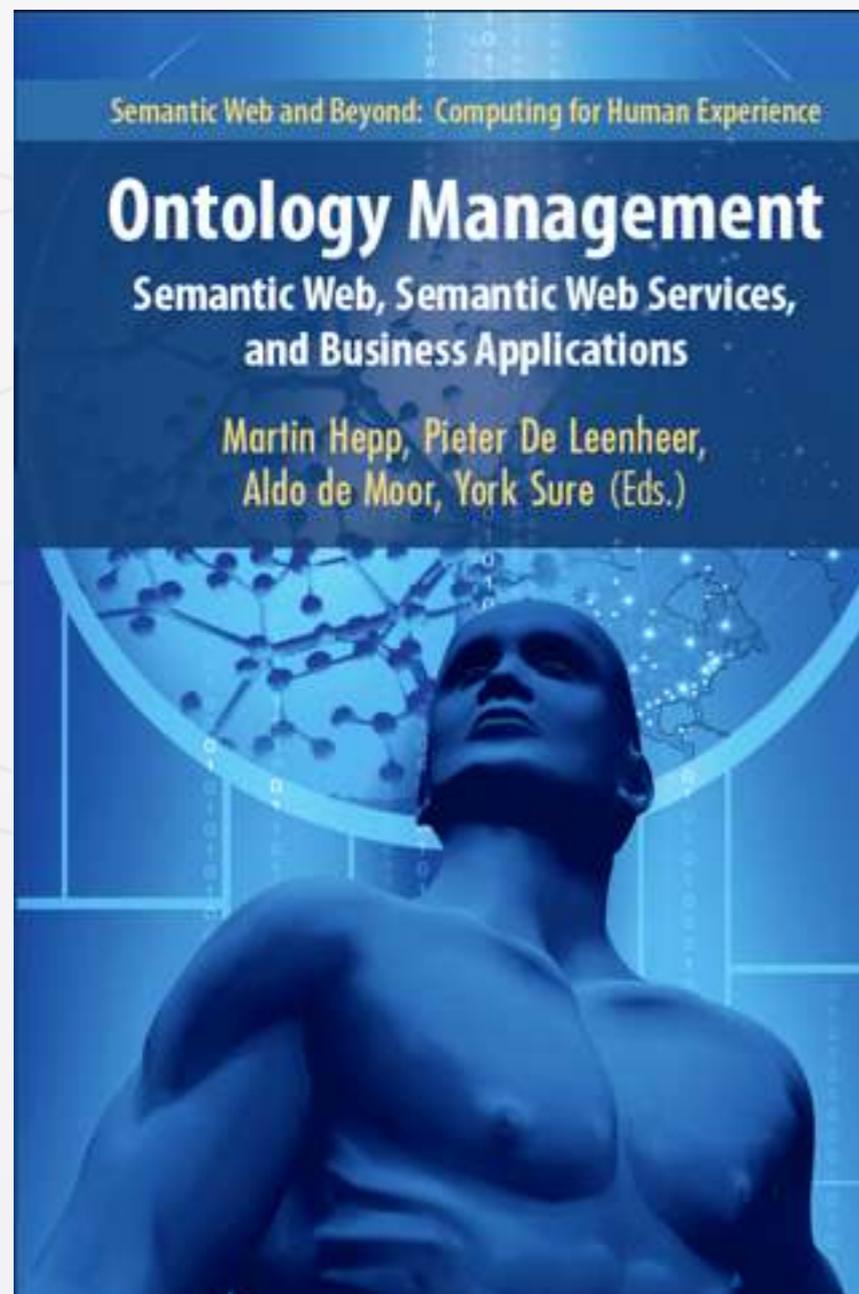


# Questions ?

Thank you

Felix Van de Maele

[felix@collibra.com](mailto:felix@collibra.com)



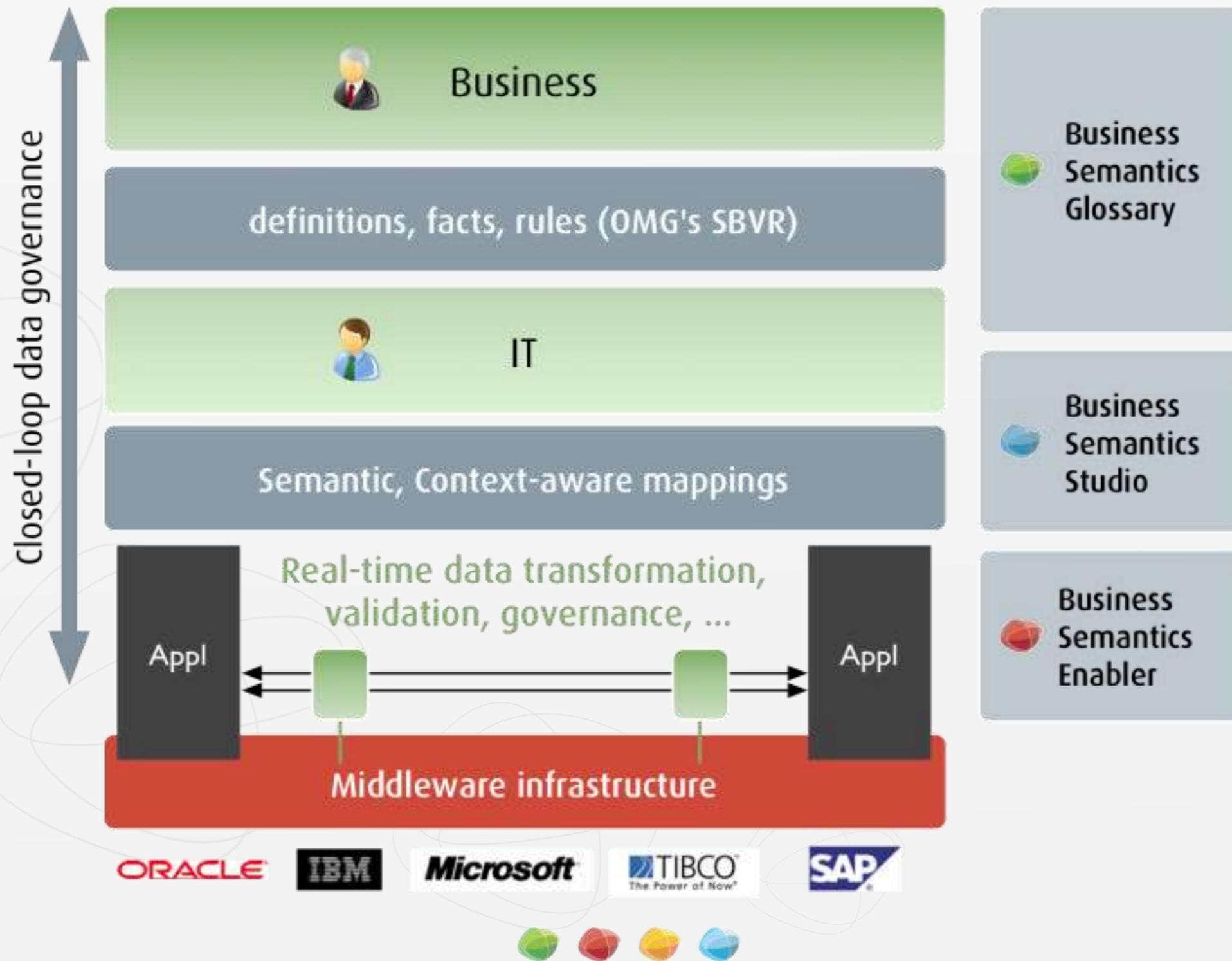
P. De Leenheer, S. Christiaens, and R. Meersman (2009) Business Semantics Management: a Case Study for Competency-centric HRM. In Journal of Computers in Industry: Special Issue about Semantic Web Computing in Industry. Elsevier.

P. De Leenheer and S. Christiaens (2008) Challenges and Opportunities for More Meaningful and Sustainable Internet Systems. In Proc. of Future of the Internet Symposium 2008 (Vienna, Austria), LNCS, SpringerM.

Hepp. (2007) Possible ontologies: How reality constrains the development of relevant ontologies. IEEE Internet Computing, 11(1):90–96



# Product suite



# Business Semantics Collibra

## Management

- ▶ *Business Semantics Glossary*  
*lets you collaboratively define and govern the meaning of your business assets in their business context and automatically generate new technical models (XML, UML, RDF/OWL, ...).*
- ▶ *Business Semantics Studio*  
*ties meaningful business context to your technical models and (legacy) data sources.*
- ▶ *Business Semantics Enabler*  
*automatically generates data transformation, validation and governance services on top of your existing middleware infrastructure*



# Business Semantics Glossary

Home
Vocabularies

**Car Movement**  
The EU-Rent Community >

- ⓐ Name
- ⓐ CarStorageCap
- ⓐ Country
- ⓐ HoursofOperat
- ⓐ LocalArea
- ⓐ storesRentalCa

```

classDiagram
    class ScheduledService {
        Service Depot : Object [0..*]
        Service Date : Object [0..*]
    }
    class CarModel {
        Car Manufacturer : Object [0..*]
        Engine Capacity : Object [0..*]
        Car Group : Object [0..*]
        Fuel Type : Object [0..*]
        Name : Object [0..*]
    }
    class RentalCar {
        Service Reading : Object [0..*]
        Acquisition Date : Object [0..*]
        Fuel Level : Object [0..*]
        Vehicle Identification Number : Object [0..*]
        Country of Registration : Object [0..*]
        Odometer Reading : Object [0..*]
        Car Group : Object [0..*]
    }
    class LocalArea {
        Operating Company : Object [0..*]
        Service Depot : Object [0..*]
    }
    class Branch {
        Country : Object [0..*]
    }

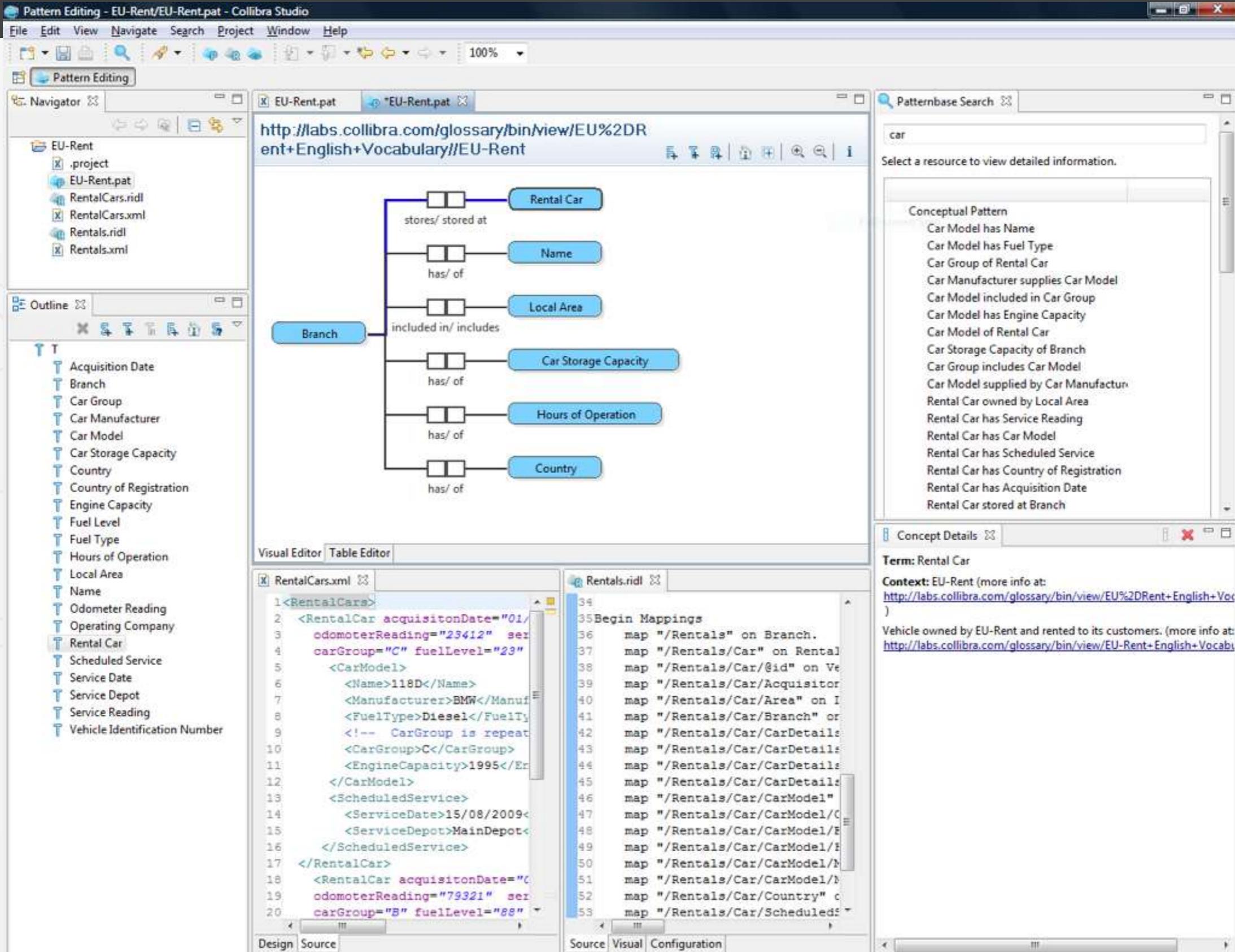
    ScheduledService "0..*" -- "0..*" RentalCar : has
    CarModel "0..*" -- "0..*" RentalCar : of
    RentalCar "0..*" -- "0..*" LocalArea : owns
    LocalArea "0..*" -- "0..*" Branch : includes
    RentalCar "0..*" -- "0..*" Branch : stored at
    
```

search...

Information

CountryofRegistration
AcquisitionDate
ServiceReading
VehicleIdentificationNumber
OdometerReading
LocalArea
CarGroup
FuelLevel
CarModel
ScheduledService

# Business Semantics Studio



The screenshot displays the Business Semantics Studio interface for a project named "EU-Rent". The main workspace shows a conceptual model with a central "Branch" entity connected to several other entities: "Rental Car", "Name", "Local Area", "Car Storage Capacity", "Hours of Operation", and "Country". The relationships are labeled as "stores/ stored at", "has/ of", and "included in/ includes".

On the left, the "Navigator" and "Outline" panels show the project structure and a list of entities including: Acquisition Date, Branch, Car Group, Car Manufacturer, Car Model, Car Storage Capacity, Country, Country of Registration, Engine Capacity, Fuel Level, Fuel Type, Hours of Operation, Local Area, Name, Odometer Reading, Operating Company, Rental Car, Scheduled Service, Service Date, Service Depot, Service Reading, and Vehicle Identification Number.

The bottom section shows two XML files: "RentalCars.xml" and "Rentals.ridl". The "RentalCars.xml" file contains XML data for two rental cars, including details like acquisition date, odometer reading, car group, fuel level, car model (BMW), fuel type (Diesel), engine capacity, and scheduled service information.

The "Rentals.ridl" file shows a list of mappings for the "Rental Car" term, such as:
 

```

35 Begin Mappings
36 map "/Rentals" on Branch.
37 map "/Rentals/Car" on Rental
38 map "/Rentals/Car/@id" on Ve
39 map "/Rentals/Car/Acquisitor
40 map "/Rentals/Car/Area" on I
41 map "/Rentals/Car/Branch" or
42 map "/Rentals/Car/CarDetails
43 map "/Rentals/Car/CarDetails
44 map "/Rentals/Car/CarDetails
45 map "/Rentals/Car/CarDetails
46 map "/Rentals/Car/CarModel"
47 map "/Rentals/Car/CarModel/C
48 map "/Rentals/Car/CarModel/E
49 map "/Rentals/Car/CarModel/E
50 map "/Rentals/Car/CarModel/M
51 map "/Rentals/Car/CarModel/M
52 map "/Rentals/Car/Country" c
53 map "/Rentals/Car/ScheduledS
    
```

On the right, the "Patternbase Search" panel shows a search for "car" and a list of conceptual patterns related to "Rental Car", such as "Car Model has Name", "Car Model has Fuel Type", and "Rental Car stored at Branch". The "Concept Details" panel provides information about the "Rental Car" term, including its context and a URL for more information.

# Business Semantics Enabler

