Semantic Integration in real life

Jürgen Angele

ESWC 2005
**Founded:** 1999 (Spin Off Univ. Karlsruhe)

**Team:** 40 Employees

**Context:** “Semantic Europe” (~ 200 R&D)
- AIFB Karlsruhe
- FZI, Karlsruhe
- DERI Galway, Irland
- DERI Innsbruck, Austria

**Products:** - OntoStudio, OntoBroker,...

**Technology:**
- Technology Leader (Gartner Group, Forrester Research)
- Vision: SemanticWeb
Our customers and partners
Introduction
Kinds of Integration Problems

*It is generally estimated that for each $1 spent for an application, companies spend on average $5 to $9 for the integration.*

What is the **problem** of information integration?

- **structural heterogeneity** – different application systems store their data in different structures
- **semantic heterogeneity** – intended meaning of information items is different in the various application systems
- **inconsistency and redundancy problems** – data in different application systems might be partially inconsistent or redundant
What are Ontologies?

„People can’t share knowledge if they do not speak a common language.“ [Davenport & Prusak, 98]

• Ontologies **standardize** and **formalize** the meaning of words through concepts

• Ontologies enable a better **communication** between
  – **Humans and/or**
  – **Machines**

• Ontologies **integrate** different conceptualisations
XSLT Transformation

```xml
<article>
  <articleid>a-5634</articleid>
  <category>printer</category>
  <name>hp81</name>
  <price currency='USD'>500</price>
  <producer>hp</producer>
  <resolution>1960 dpi</resolution>
  <type>laser</type>
</article>

<xsl:template match="article">
  <Artikel ArtNr='a-5634'>
    <Typ>Laserdrucker</Typ>
    <Name>hp81</Name>
    <Preis Waehrung='E'>625</Preis>
    <Hersteller>hp</Hersteller>
    <Aufloesung>1960 dpi</Aufloesung>
    ....
  </Artikel>
</xsl:template>

<xsl:template match="article">
  <Artikel ArtNr='{./articleid}'>
    <Name><xsl:value-of select="name"/></Name>
    <Preis Waehrung='E'><xsl:value-of select="price*1.25"/></Preis>
    <Hersteller><xsl:value-of select="producer"/></Hersteller>
    <Aufloesung><xsl:value-of select="resolution"/></Aufloesung>
    ....
  </Artikel>
</xsl:template>
```

Ontology

- Representation Language: F-Logic, WRL
- standards: RDF, OWL
if X is a Drucker and hasTyp Laser, then X is a laserprinter
Motivation Semantic Information Integration
End Users dealing with Multiple Systems

- Customer Service Delays
- Rising Costs

- Inaccurate Information
- Processing Delays

- Incomplete View of Business
- Reporting Delays

Multiple Interfaces

Different Formats

Different Meanings
EII Value Propositions

- Single View
- Business Agility
- Increased Productivity
Single View

- Aggregating data from multiple systems
- Presenting relevant information in the user’s terminology
- Giving different perspectives into the same information

“Can I get a single view of ... ?”

...Customer?
...Citizen?
...Patient?
...Policy Holder?
...Revenue?
...Supply Chain?

...ANY Entity!
...ANY Process!
Business Agility

“What are $$ by Region?”

<table>
<thead>
<tr>
<th>Reg1</th>
<th>$$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terr1</td>
<td>$</td>
</tr>
<tr>
<td>Terr2</td>
<td>$</td>
</tr>
<tr>
<td>Terr3</td>
<td>$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reg2</th>
<th>$$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terr4</td>
<td>$</td>
</tr>
<tr>
<td>Terr5</td>
<td>$</td>
</tr>
<tr>
<td>Terr6</td>
<td>$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reg3</th>
<th>$$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terr7</td>
<td>$</td>
</tr>
<tr>
<td>Terr8</td>
<td>$</td>
</tr>
<tr>
<td>Terr9</td>
<td>$</td>
</tr>
</tbody>
</table>
Business Agility

- Minimizing impact of change
- Ease of maintenance
- Rapid implementation of new strategies

Restructure!
- Remove 1 Region
- Split Territories

“What are $$ by Region?”

<table>
<thead>
<tr>
<th>Region</th>
<th>Territories</th>
</tr>
</thead>
<tbody>
<tr>
<td>RegA</td>
<td></td>
</tr>
<tr>
<td>Terr1</td>
<td>$</td>
</tr>
<tr>
<td>Terr2</td>
<td>$</td>
</tr>
<tr>
<td>Terr3</td>
<td>$</td>
</tr>
<tr>
<td>Terr4</td>
<td>$</td>
</tr>
<tr>
<td>RegB</td>
<td></td>
</tr>
<tr>
<td>Terr5</td>
<td>$</td>
</tr>
<tr>
<td>Terr6</td>
<td>$</td>
</tr>
<tr>
<td>Terr7</td>
<td>$</td>
</tr>
<tr>
<td>Terr8</td>
<td>$</td>
</tr>
<tr>
<td>Terr9</td>
<td>$</td>
</tr>
</tbody>
</table>
“Should this insurance policy be canceled?”

(it is a lousy payer, but his cousin is CEO of an important customer)
Increased Productivity 2

- Capturing business rules directly in the Information Model
- Determining the optimal system access
- Bringing every user to the same level of effectiveness and productivity

“Authorize this charge?”

(Need to extend limit, check credit)
EII Value Propositions

- **Single View**
  - Aggregating data from multiple systems
  - Presenting relevant information in the user’s terminology
  - Giving different perspectives into the same information

- **Business Agility**
  - Minimizing impact of change
  - Ease of maintenance
  - Rapid implementation of new strategies

- **Increased Productivity**
  - Capturing business rules directly in the Information Model
  - Determining the optimal system access
  - Bringing every user to the same level of effectiveness and productivity
Concept
Relational Databases

- Mapping to Databases
- Mapping to Ontologies
Mapping in OntoStudio
Mapping in OntoStudio
Integration of several databases

which cars have a multitronic and an TDI V6 engine?

Car 54
A: Address[zip->>Z] and temperature(Z,T) and T > 25 -> warm(A).
Unstructured Information

example

Who has java programming skills and knows customer Bigdeal AG?
Java skills?

employees with java skills:
S. Maier maier@firma.de
G. Nial nial@firma.de
S. Uper uper@firma.de

S. Maier: „… In Java Version 1.4 this function has been implemented

employee-DB

H. Müller müller@firma.de
S. Maier maier@firma.de
F. Schmidt schmidt@firma.de
employees with java skills:
S. Maier maier@firma.de
G. Nial nial@firma.de
S. Uper uper@firma.de

answer:
S. Uper uper@firma.de

project reports
Titel: Effizienz mit Onto
Team: H. Fleissig  
S. Uper
Kunde: Bigdeal AG
......
Semantic Information Integration
Applications
Audi: Semantic Testcar Configuration

**Background**
- Complex dependencies decrease the speed of development
- Knowledge is distributed over different departments

**Goal**
- Design of a Semantic Guide for
  - capturing the dependencies
  - Configuration of components
- Integration into existing order system
- Engineers can concentrate on creative efforts
Ontologies represent the meaning of information

Represent the **meaning of information**

- Concepts and hierarchies (Car, has_Part, Engine, Body, ...)
- Synonyms (Engine, Motor)
- Attributes and relations (Part_ID, designed_for_power, controls)
- Other

“An ontology is a hierarchically structured set of terms for describing a domain that can be used as a skeletal foundation for a knowledge base.” Swartout, Patil, Knight and Russ.
Ontologies represent the logic of information

Represent the logic of information
- Rules to define constraints (Chassis has to be designed for the power of the engine)
- Rules for defining any functional, logical, geometrical, chronological dependencies (has_Power influences gearbox and tires)
- Rules for information integration (value “Engine has_power” is stored in “PDM p, Table t1”; value “designed_for_power” is stored in “CAT c, Table t2”)
- Rules to define different contexts

“Ontologies are the backbone of semantic technologies. They enable companies to integrate information, make them tangible and re-usable.” Prof. Dr. Rudi Studer.
Rule 2: The maximum power of the motor must not exceed the one of the brakes: \( P_{\text{motor}} < | P_{\text{brakes}} | \)

\[
\text{FORALL } X,Y,Z_1,Z_2,Z_3 \\
\text{message(“The motor’s maximum power exceeds the one of the brakes.”)} \\
\langle- X: \text{motor}[\text{maximum_power}]->Z_1 \rangle \text{ AND} \\
\quad Y: \text{brake}[\text{maximum_power}]->Z_2 \rangle \text{ AND} \\
\quad \text{abs}(Z_1,Z_3) \text{ AND lessorequal}(Z_2,Z_3).
\]
Problems with IT support of R&D processes

R&D processes

Application → System → Application

Experience

Semantic Middleware

Service → DB
Service → CAX
Service → PDM

What if:
- You could integrate existing data
- You could add the experience of engineers to the applications
Accelerate R&D and Customer Service

Background

• Critical success factors for development of complex products
  • Time-To-Market
  • Service Quality

Problem

• Complex dependencies and variants
• Heterogeneous Sources
• Difficulty to transfer engineers’ know-how

Solution

• Semantic Customer Service Support

Why did you change the supplier?

Better product available: 65%

Unsatisfied with service: 20%

Source: Wards Automotive Yearbook, McKinsey analysis

Chrysler/DaimlerChrysler since 1998 (merger), Ford, General Motors.
Customer Service Support

Background

- 65% of all customer in the manufacturing industry change their suppliers because there are not satisfied with the service

- Service engineers spend a lot of time with known problems

Goal

- Capturing and usage of engineers and experts know-how
- Decision support for choosing the right solution
- Increase customer satisfaction

Implementation

- Semantic Customer Service Support
## Problem: Endstufe defekt

**Auftrag:** DC 3809/77

**Gerät Id:** KR C2 DC3490

### Kategorien

**Bauteil > Bauteilgruppe 1 > Endstufe**

**Defekt: physischer Defekt > Baugruppen1-Defekt**

### Lösung

<table>
<thead>
<tr>
<th>#</th>
<th>PROBLEM</th>
<th>Lösung</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Kabel defekt oder fehlerhaft montiert</td>
<td>Durch Kabeldefekt oder fehlerhafte Montage kann Überstrom/Unterspannung ausgelöst werden. Dadurch wird Defekt der Endstufe ausgelöst. Kabelkontrolle, gegebenenfalls Tausch</td>
</tr>
</tbody>
</table>

### DOKUMENTE

- TSN Technische Service Nachrichten vom 12.02.2003
- Dokumentation KR C2 Service Elektrik Rachortel vom 17.04.2003
- TSN Technische Service Nachrichten vom 12.05.2002

**AUTOR**

- [ ]

**NUTZEN**

- [ ]

**PROZESS**

- [ ] Übernahme in Auftrag
- [ ] Übernahme in Auftrag mit Änderung
- [ ] Neue Problemstellung erbracht
- [ ] Lösung nicht hilfreich

**Dokumente:**

- Technische Service Nachrichten vom 23.09.2003

**Letzte Aktualisierung:** 28.10.2003 16:38
EII Product
Software AG’s
Enterprise Information Integrator 2.1
EII v2.1 Architecture

1. **User Application**
   - Web Service
   - Semantic Server
   - Metadata Manager

2. **Integrator Studio**
   - Import Metadata
   - Physical Model

3. **Build Query**
   - Business Model

4. **Deploy Web Service**
   - WSDL

5. **Map to Business Model**

Start → Metadata Manager → XML → Physical Model → ... → End

© 2003 ontoprise GmbH

WWW.ONTOPRISE.DE
Provider of technology and applications enabling semantic solutions
EII Differentiator

<table>
<thead>
<tr>
<th>Virtual Federated DB</th>
<th>Model-Driven Virtual Federated DB</th>
<th>Model-Driven Semantic Integration</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Static Coded Queries</td>
<td>• Generated Queries</td>
<td>• Dynamic Queries</td>
</tr>
<tr>
<td>• Very Brittle</td>
<td>• No Optimization</td>
<td>• Optimized Access</td>
</tr>
<tr>
<td>• Single Use Case</td>
<td>• No Intelligence</td>
<td>• Intelligent Access</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Multiple Use Cases</td>
</tr>
</tbody>
</table>
EII v2.1 Active Projects

- Florida Community College of Jacksonville
  - Registration and Course Scheduling in Student Portal
- SIRVA
  - Financial Reporting, Single Customer View
- CompuCredit
  - Debt Collections, Single Customer View
- South Carolina Retirement Systems
  - Account Statement Generation, Single Customer View
- Alcatel
  - *Intelligent Storage System* – Mobile Services Single View
- CBIG (Internal Software AG)
  - Single Customer View
- Bundeswertpapierverwaltung
  - Part of BPM, Single Customer View
EII v2.1 Other Opportunities

- Pharmaceutical
  - Single View of Brand Marketing Strategies
  - Financial Reporting of Product
- Large Grocer
  - Business Analytics
- 2 Very Large Insurance Companies
  - Single Policyholder Views
- Automobile Manufacturer
  - Semantic Matching
- US State Agency
  - Unemployment Compensation Fraud
- A National Bureau of Statistics
  - Single Citizen Views
- A National Security Agency
  - Metadata Repository of Services and an Enterprise Data Model
- Independent Software Vendor
  - Information Model and Business Rules Engine
- Global System Integrator
  - Multiple Projects
Future
The Future: Evolution of Integration

- **Data Transport**: ... to transport data
- **Data & Information Integration**: ... to have a common understanding of the data of an enterprise
- **Application Integration**: ... to integrate application logic and data between two or more applications
- **Process Integration**: ... to automate business operations, tasks and transactions of business processes
- **Service-Driven Integration**: ... to allow for flexible and dynamic integration of suppliers, customers and partners

**Added value by integration**

**Business Complexity**

© 2003 ontoprise GmbH
The Vision of a (Semantic) Web of Services

Dynamic

Web Services
UDDI, WSDL, SOAP

Semantics Services

Static

WWW
URI, HTML, HTTP

Semantic Web
RDF, RDF(S), OWL
Thank you!

Prof. Dr. Jürgen Angele
angele@ontoprise.de
+49 (0)721 509 809 0