Chapter 6
Applications

Grigoris Antoniou
Frank van Harmelen

Lecture Outline

1. Horizontal Information Products at Elsevier
2. Data Integration at Audi
3. Skill Finding at Swiss Life
4. Think Tank Portal at EnerSearch
5. E-Learning
6. Web Services
7. Other Scenarios

Elsevier – The Setting

- Elsevier is a leading scientific publisher.
- Its products are organized mainly along traditional lines:
  - Subscriptions to journals
- Online availability of these journals has until now not really changed the organisation of the productline
- Customers of Elsevier can take subscriptions to online content

Elsevier – The Problem

- Traditional journals are vertical products
- Division into separate sciences covered by distinct journals is no longer satisfactory
- Customers of Elsevier are interested in covering certain topic areas that spread across the traditional disciplines/journals
- The demand is rather for horizontal products

Elsevier – The Problem (2)

- Currently, it is difficult for large publishers to offer such horizontal products
  - Barriers of physical and syntactic heterogeneity can be solved (with XML)
  - The semantic problem remains unsolved
- We need a way to search the journals on a coherent set of concepts against which all of these journals are indexed

Elsevier – The Contribution of Semantic Web Technology

- Ontologies and thesauri (very lightweight ontologies) have proved to be a key technology for effective information access
  - They help to overcome some of the problems of free-text search
  - They relate and group relevant terms in a specific domain
  - They provide a controlled vocabulary for indexing information

Elsevier – The Contribution of Semantic Web Technology (2)

- A number of thesauri have been developed in different domains of expertise
  - Medical information: MeSH and Elsevier's life science thesaurus EMTREE
- RDF is used as an interoperability format between heterogeneous data sources
- EMTREE is itself represented in RDF

Elsevier – The Contribution of Semantic Web Technology (3)

- Each of the separate data sources is mapped onto this unifying ontology
  - The ontology is then used as the single point of entry for all of these data sources

Lecture Outline

1. Horizontal Information Products at Elsevier
2. Data Integration at Audi
3. Skill Finding at Swiss Life
4. Think Tank Portal at EnerSearch
5. E-Learning
6. Web Services
7. Other Scenarios

Audi – The Problem

- Data integration is also a huge problem internal to companies
  - It is the highest cost factor in the information technology budget of large companies
  - Audi operates thousands of databases
- Traditional middleware improves and simplifies the integration process
  - But it misses the sharing of information based on the semantics of the data

Audi – The Contribution of Semantic Web Technology

- Ontologies can rationalize disparate data sources into one body of information
- Without disturbing existing applications, by:
  - creating ontologies for data and content sources
  - adding generic domain information
- The ontology is mapped to the data sources giving applications direct access to the data through the ontology

Audi – Camera Example

```
<SLR rdf:ID="Olympus-OM-10">
    <viewFinder>twin mirror</viewFinder>
    <optics>
        <Lens>
            <focal-length>75-300mm zoom</focal-length>
            <f-stop>4.0-4.5</f-stop>
        </Lens>
    </optics>
    <shutter-speed>1/2000 sec. to 10 sec.</shutter-speed>
</SLR>
```

Audi – Camera Example (2)

```
<Camera rdf:ID="Olympus-OM-10">
    <viewFinder>twin mirror</viewFinder>
    <optics>
        <Lens>
            <size>300mm zoom</size>
            <aperture>4.5</aperture>
        </Lens>
    </optics>
    <shutter-speed>1/2000 sec. to 10 sec.</shutter-speed>
</Camera>
```

Audi – Camera Example (3)

- Human readers can see that these two different formats talk about the same object
  - We know that SLR is a kind of camera, and that **fstop** is a synonym for **aperture**
- Ad hoc integration of these data sources by translator is possible
- Would only solve this specific integration problem
- We would have to do the same again when we encountered the next data format for cameras

Audi – Camera Ontology in OWL

```
<owl:Class rdf:ID="SLR">
		<rdfs:subClassOf rdf:resource="#Camera"/>
</owl:Class>
<owl:DatatypeProperty rdf:ID="f-stop">
		<rdfs:domain rdf:resource="#Lens"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="aperture">
		<owl:equivalentProperty rdf:resource="#f-stop"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="focal-length">
		<rdfs:domain rdf:resource="#Lens"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="size">
		<owl:equivalentProperty rdf:resource="#focal-length"/>
</owl:DatatypeProperty>
```

Audi – Using the Ontology

- Suppose that an application A
  - is using the second encoding
  - is receiving data from an application B using the first encoding
- Suppose it encounters **SLR**
  - Ontology returns “**SLR** is a type of **Camera**”
  - A relation between something it doesn’t know (**SLR**) to something it does know (**Camera**)

Audi – Using the Ontology (2)

- Suppose A encounters **f-stop**
  - The Ontology returns: "**f-stop** is synonymous with **aperture**"
- Bridges the terminology gap between something A doesn't know to something A does know
- Syntactic divergence is no longer a hindrance

Lecture Outline

1. Horizontal Information Products at Elsevier
2. Data Integration at Audi
3. Skill Finding at Swiss Life
4. Think Tank Portal at EnerSearch
5. E-Learning
6. Web Services
7. Other Scenarios

Swiss Life – The Setting

- Swiss Life is one of Europe's leading life insurers
  - 11,000 employees, $14 billion of written premiums
  - Active in about 50 different countries
- The most important resources of any company for solving knowledge intensive tasks are:
  - The tacit knowledge, personal competencies, and skills of its employees

Swiss Life – The Problem

- One of the major building blocks of enterprise knowledge management is:
  - An electronically accessible repository of people's capabilities, experiences, and key knowledge areas
- A skills repository can be used to:
  - enable a search for people with specific skills
  - expose skill gaps and competency levels
  - direct training as part of career planning
  - document the company's intellectual capital

Swiss Life – The Problem (2)

- Problems
  - How to list the large number of different skills?
  - How to organise them so that they can be retrieved across geographical and cultural boundaries?
  - How to ensure that the repository is updated frequently?

Swiss Life – The Contribution of Semantic Web Technology

- Hand-built ontology to cover skills in three organizational units
  - Information Technology, Private Insurance and Human Resources
- Individual employees within Swiss Life were asked to create “home pages” based on form filling driven by the skills-ontology
- The corresponding collection could be queried using a form-based interface that generated RQL queries

Swiss Life – Skills Ontology

```
<owl:Class rdf:ID="Skills">
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:onProperty rdf:resource="#HasSkillsLevel"/>
            <owl:cardinality
    rdf:datatype="&xsd;nonNegativeInteger">
            1</owl:cardinality>
        </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>
<owl:ObjectProperty rdf:ID="HasSkills">
    <rdfs:domain rdf:resource="#Employee"/>
    <rdfs:range rdf:resource="#Skills"/>
</owl:ObjectProperty>
```

Swiss Life – Skills Ontology (2)

```
<owl:ObjectProperty rdf:ID="WorksInProject">
    <rdfs:domain rdf:resource="#Employee"/>
    <rdfs:range rdf:resource="#Project"/>
    <owl:inverseOf rdf:resource="#ProjectMembers"/>
</owl:ObjectProperty>
<owl:Class rdf:ID="Publishing">
    <rdfs:subClassOf rdf:resource="#Skills"/>
</owl:Class>
<owl:Class rdf:ID="DocumentProcessing">
    <rdfs:subClassOf rdf:resource="#Skills"/>
</owl:Class>
```

Swiss Life – Skills Ontology (3)

```
<owl:ObjectProperty rdf:ID="ManagementLevel">
    <rdfs:domain rdf:resource="#Employee"/>
    <rdfs:range>
        <owl:oneOf rdf:parseType="Collection">
            <owl:Thing rdf:about="#member"/>
            <owl:Thing rdf:about="#HeadOfGroup"/>
            <owl:Thing rdf:about="#HeadOfDept"/>
            <owl:Thing rdf:about="#CEO"/>
        </owl:oneOf>
    </rdfs:range>
</owl:ObjectProperty>
```

Lecture Outline

1. Horizontal Information Products at Elsevier
2. Data Integration at Audi
3. Skill Finding at Swiss Life
4. Think Tank Portal at EnerSearch
5. E-Learning
6. Web Services
7. Other Scenarios

EnerSearch – The Setting

- An industrial research consortium focused on information technology in energy
- EnerSearch has a structure very different from a traditional research company
  - Research projects are carried out by a varied and changing group of researchers spread over different countries
  - Many of them are not employees of EnerSearch

EnerSearch – The Setting (2)

- EnerSearch is organized as a virtual organization
- Owned by a number of firms in the industry sector that have an express interest in the research being carried out
- Because of this wide geographical spread, EnerSearch also has the character of a virtual organisation from a knowledge distribution point of view

EnerSearch – The Problem

- Dissemination of knowledge key function
- The information structure of the web site leaves much to be desired
- It does not satisfy the needs of info seekers, e.g.
  - Does load management lead to cost-saving?
  - If so, what are the required upfront investments?
  - Can powerline communication be technically competitive to ADSL or cable modems?

EnerSearch – The Contribution of Semantic Web Technology

- It is possible to form a clear picture of what kind of topics and questions would be relevant for these target groups
- It is possible to define a domain ontology that is sufficiently stable and of good quality
  - This lightweight ontology consisted only of a taxonomical hierarchy
  - Needed only RDF Schema expressivity

EnerSearch – Lunchtime Ontology

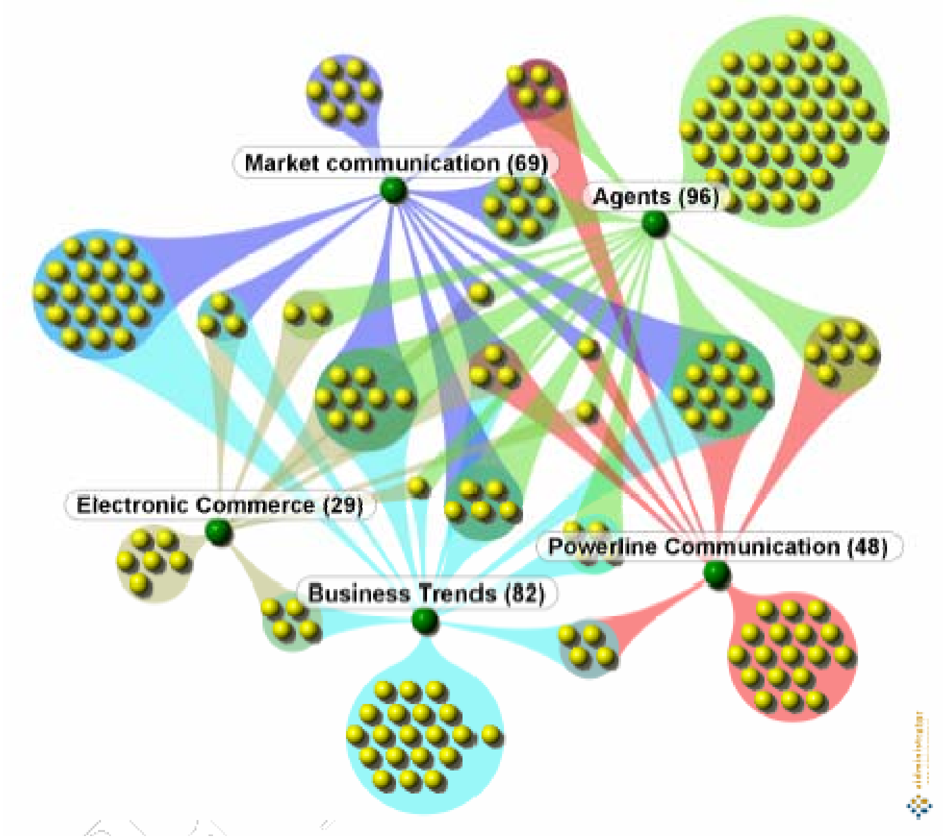
...

- **IT**
  - **Hardware**
  - **Software**
  - **Applications**
  - **Communication**
    - **Powerline**
    - **Agent**
  - **Electronic Commerce**
    - **Agents**
      - **Multi-agent systems**
      - **Intelligent agents**
    - **Market/auction**
      - **Resource allocation**
      - **Algorithms**

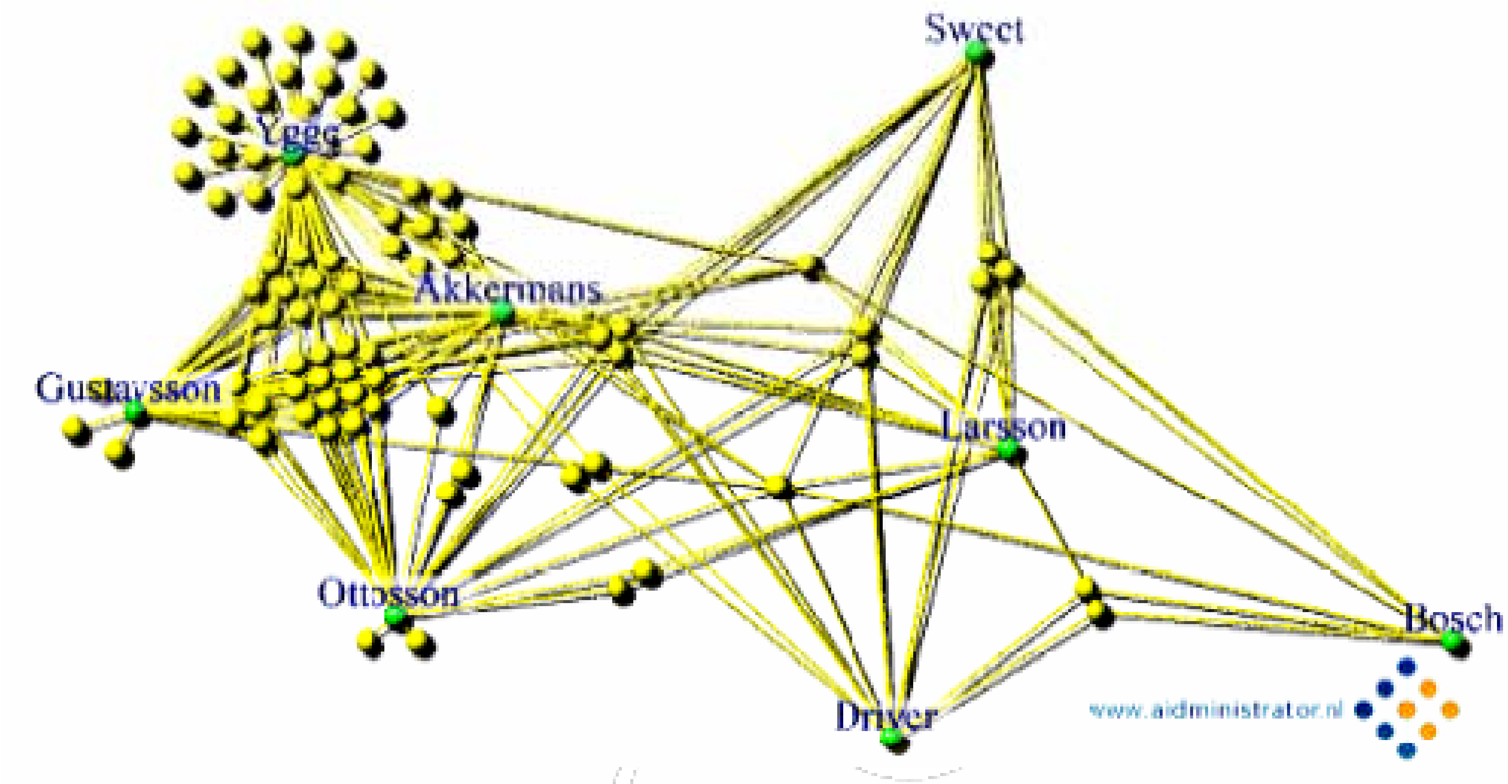
EnerSearch – Use of Ontology

- Used in a number of different ways to drive navigation tools on the EnerSearch web site
  - Semantic map of the EnerSearch web site
  - Semantic distance between EnerSearch authors in terms of their fields of research and publication

Semantic Map of Part of the EnerSearch Web Site



Semantic Distance between EnerSearch Authors



EnerSearch – QuizRDF

- QuizRDF aims to combine
  - an entirely ontology based display
  - a traditional keyword based search without any semantic grounding
- The user can type in general keywords
- It also displays those concepts in the hierarchy which describe these papers
- All these disclosure mechanisms (textual and graphic, searching or browsing) based on a single underlying lightweight ontology

Lecture Outline

1. Horizontal Information Products at Elsevier
2. Data Integration at Audi
3. Skill Finding at Swiss Life
4. Think Tank Portal at EnerSearch
5. E-Learning
6. Web Services
7. Other Scenarios

E-Learning – The Setting

- Traditionally learning has been characterized by the following properties:
  - Educator-driven
  - Linear access
  - Time- and locality-dependent
  - Learning has not been personalized but rather aimed at mass participation

E-Learning – The Setting (2)

- The changes are already visible in **higher education**
  - Virtual universities
  - Flexibility and new educational means
  - Students can increasingly make choices about pace of learning, content, evaluation methods

E-Learning – The Setting (3)

- Even greater promise: **life long learning activities**
  - Improvement of the skills of its employees ic critical to companies
  - Organizations require learning processes that are just-in-time, tailored to their specific needs
  - These requirements are not compatible with traditional learning, but e-learning shows great promise for addressing these concerns

E-Learning – The Problem

- E-learning is not driven by the instructor
- Learners can:
  - Access material in an order that is not predefined
  - Compose individual courses by selecting educational material
- Learning material must be equipped with additional information (metadata) to support effective indexing and retrieval

E-Learning – The Problem (2)

- Standards (IEEE LOM) have emerged
  - E.g. educational and pedagogical properties, access rights and conditions of use, and relations to other educational resources
- Standards suffer from lack of semantics
  - This is common to all solutions based solely on metadata (XML-like approaches)
  - Combining of materials by different authors may be difficult
  - Retrieval may not be optimally supported
  - Retrieval and organization of learning resources must be made manually
  - Could be done by a personalized automated agent instead!

E-Learning – The Contribution of Semantic Web Technology

- Establish a promising approach for satisfying the e-learning requirements
  - E.g. ontology and machine-processable metadata
- Learner-centric
  - Learning materials, possibly by different authors, can be linked to commonly agreed ontologies
  - Personalized courses can be designed through semantic querying
  - Learning materials can be retrieved in the context of actual problems, as decided by the learner

E-Learning – The Contribution of Semantic Web Technology (2)

- Flexible access
  - Knowledge can be accessed in any order the learner wishes
  - Appropriate semantic annotation will still define prerequisites
  - Nonlinear access will be supported
- Integration
  - A uniform platform for the business processes of organizations
  - Learning activities can be integrated in these processes

Ontologies for E-Learning

- Some mechanism for establishing a shared understanding is needed: ontologies
- In e-learning we distinguish between three types of knowledge (ontologies):
  - Content
  - Pedagogy
  - Structure

Content Ontologies

- Basic concepts of the domain in which learning takes place
- Include the relations between concepts, and basic properties
  - E.g., the study of **Classical Athens** is part of the *history of* **Ancient Greece**, which in turn is part of **Ancient History**
  - The ontology should include the relation "*is part of*" and the fact that it is transitive (e.g., expressed in OWL)
- COs use relations to capture synonyms, abbreviations, etc.

Pedagogy Ontologies

- Pedagogical issues can be addressed in a pedagogy ontology (PO)
- E.g. material can be classified as lecture, tutorial, example, walk-through, exercise, solution, etc.

Structure Ontologies

- Define the logical structure of the learning materials
- Typical knowledge of this kind includes hierarchical and navigational relations like **previous**, **next**, **hasPart**, **isPartOf**, **requires**, and **isBasedOn**
- Relationships between these relations can also be defined
  - E.g., **hasPart** and **isPartOf** are inverse relations
- Inferences drawn from learning ontologies cannot be very deep

Lecture Outline

1. Horizontal Information Products at Elsevier
2. Data Integration at Audi
3. Skill Finding at Swiss Life
4. Think Tank Portal at EnerSearch
5. E-Learning
6. Web Services
7. Other Scenarios

Web Services

- Web sites that do not merely provide static information, but involve interaction with users and often allow users to effect some action
- **Simple Web services** involve a single Web-accessible program, sensor, device
- **Complex Web services** are composed of simpler services
  - Often they require ongoing interaction with the user
  - The user can make choices or provide information conditionally

A Complex Web Service

- User interaction with an online music store involves
  - searching for CDs and titles by various criteria
  - reading reviews and listening to samples
  - adding CDs to a shopping cart
  - providing credit card details, shipping details, and delivery address

Web Services – Contribution of Semantic Web Technology

- Use machine-interpretable descriptions of services to automate:
  - discovery, invocation, composition and monitoring of Web services
- Web sites should be able to employ a set of basic classes and properties by declaring and describing services: ontology of services

DAML-S and OWL-S

- DAML-S is an initiative that is developing an ontology language for Web services
- It makes use of DAML+OIL
- It can be viewed as a layer on top of DAML+OIL
- OWL-S is more recent version on top of OWL

Three Basic Kinds of Knowledge Associated with a Service

- Service profile
  - Description of the offerings and requirements of a service
  - Important for service discovery
- Service model
  - Description of how a service works
- Service grounding
  - communication protocol and port numbers to be used in contacting the service

Service Profiles

- Describe services offered by a Web site
- A service profile in DAML-S provides the following information:
  - A human-readable description of the service and its provider
  - A specification of the functionalities provided by the service
  - Additional information, such as expected response time and geographic constraints
- Encoded in the modeling primitives of DAML-S:
- E.g. classes and properties defined in DAML+OIL

Service Profiles (2)

```
<rdfs:Class rdf:ID="OfferedService">
   <rdfs:label>OfferedService</rdfs:label>
   <rdfs:subClassOf rdf:resource=
   "http://www.daml.org/services/daml-s/
   2001/10/Service.daml#"/>
</rdfs:Class>
```

Service Profiles (3)

- Properties defined on this class:
  - **intendedPurpose** (range = string)
  - **serviceName** (range = string)
  - **providedBy** (range = is a new class, **Service-Provider**, which has various properties)

Functional Description of Web Services

- **input** describes the parameters necessary for providing the service
  - E.g., a sports news service might require the following input:
  - date, sports category, customer credit card details.
- **output** specifies the outputs of the service
  - In the sports news example, the output would be the news articles in the specified category at the given date

Functional Description of Web Services (2)

- **precondition** specifies the conditions that need to hold for the service to be provided effectively
  - The distinction between inputs and preconditions can be illustrated in our running example:
  - The credit card details are an input, and preconditions are that the credit card is valid and not overcharged
- **effect** specifies the effects of the service
  - In our example, an effect might be that the credit card is charged $1 per news article

Service Models

- Based on the key concept of a process, which describes a service in terms of
  - inputs, outputs, preconditions, effects, and
  - its composition of component subprocesses
- **Atomic processes** can be directly invoked by passing them appropriate messages; they execute in one step
- **Simple processes** are elements of abstraction; they have single-step executions but are not invocable
- **Composite processes** consist of other, simpler processes

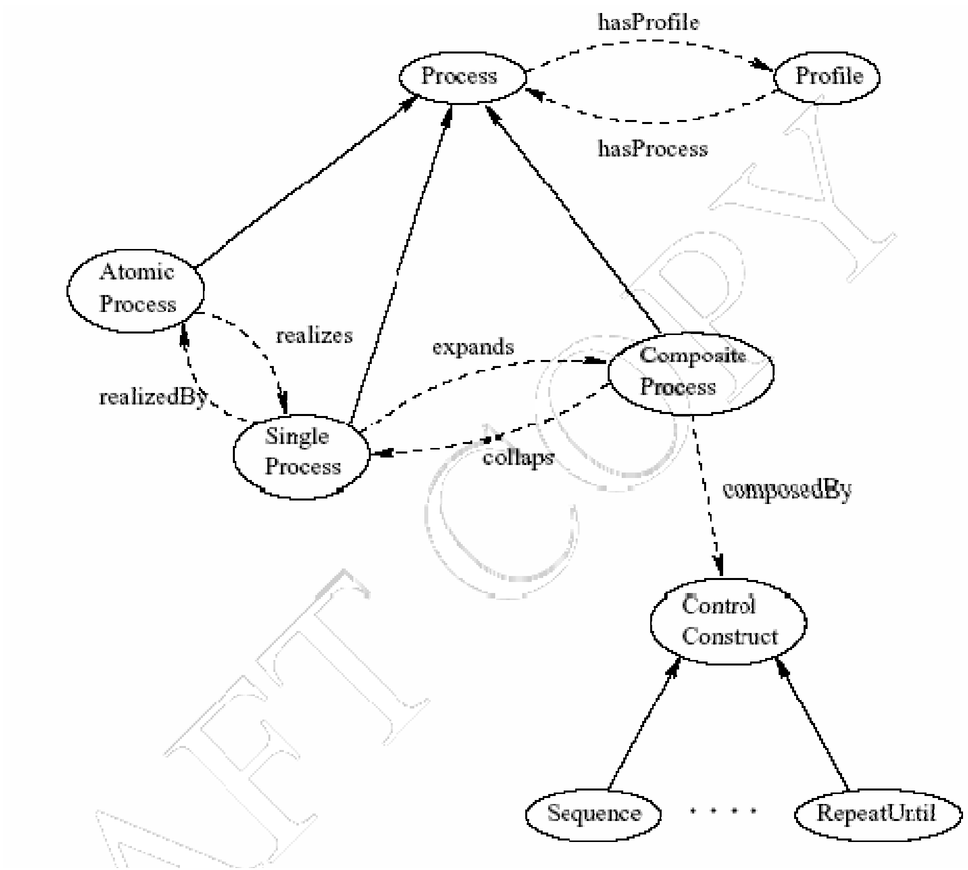
Composition of Processes

- A composite process is composed of a number of control constructs:

```
<rdf:Property rdf:ID="composedBy">
        <rdfs:domain rdf:resource="#CompositeProcess"/>
        <rdfs:range rdf:resource="#ControlConstruct"/>
</rdf:Property>
```

- Control constructs offered by DAML-S include:
  – sequence, choice, if-then-else and repeat-until

Top Level of the Process Ontology



Lecture Outline

1. Horizontal Information Products at Elsevier
2. Data Integration at Audi
3. Skill Finding at Swiss Life
4. Think Tank Portal at EnerSearch
5. E-Learning
6. Web Services
7. Other Scenarios

Multimedia Collection Indexing at Scotland Yard

- Theft of art and antique objects
- International databases of stolen art objects exist
  - It is difficult to locate specific objects in these databases
  - Different parties are likely to offer different descriptions
  - Human experts are needed to match objects to database entries

Multimedia Collection Indexing at Scotland Yard – The Solution

- Develop controlled vocabularies such as the Art and Architecture Thesaurus (AAT) from the Getty Trust, or Iconclass thesaurus
- Extend them into full-blown ontologies
- Develop automatic classifiers using ontological background knowledge
- Deal with the ontology-mapping problem

Online Procurement at Daimler-Chrysler – The Problem

- Static, long-term agreements with a fixed set of suppliers can be replaced by dynamic, short-term agreements in a competitive open marketplace
- Whenever a supplier is offering a better deal, Daimler-Chrysler wants to be able to switch
- Major drivers behind B2B e-commerce

Online Procurement at Daimler-Chrysler – The Solution

- Rosetta Net is an organization dedicated to such standardization efforts
- XML-based, no semantics
- Use RDF Schema and OWL instead
  - Product descriptions would “carry their semantics on their sleeve”
  - Much more liberal online B2B procurement processes would exist than currently possible

Device Interoperability at Nokia

- Explosive proliferation of digital devices:
  - PDAs, mobiles, digital cameras, laptops, wireless access in public places, GPS-enabled cars
- Interoperability among these devices?
- The pervasiveness and the wireless nature of these devices require network architectures to support automatic, ad hoc configuration
- A key technology of true ad hoc networks is *service discovery*

Device Interoperability at Nokia (2)

- Current service discovery and capability description require a priori identification of what to communicate or discuss
- A more attractive approach would be "*serendipitous interoperability*"
  - Interoperability under "unchoreographed" conditions
  - Devices necessarily designed to work together

Device Interoperability at Nokia (3)

- These devices should be able to:
  – Discover each others' functionality
  – Take advantage of it
- Devices must be able to "understand" other devices and reason about their functionality
- Ontologies are required to make such "unchoreographed" understanding of functionalities possible