Semantic web services in corporate memories

Moussa Lo, Fabien Gandon

1

Combing two stacks of recommendations

2

Semantic web

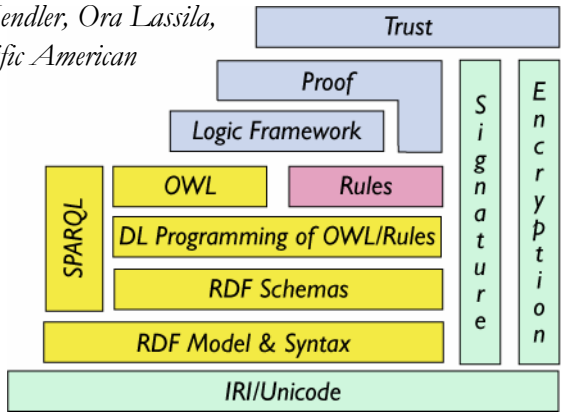
web ++

web + metadata for applications

The semantic cake

"The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation."

— *Tim Berners-Lee, James Hendler, Ora Lassila, The Semantic Web, Scientific American May 2001*

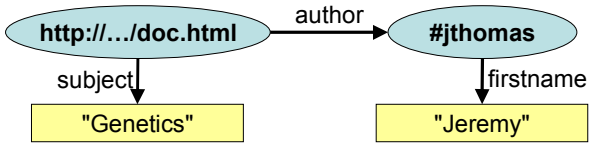


Resource Description Framework

- Represent assertions about (Web) ressources
  e.g.: doc.html has for author Jeremy and is about Genetics
- Model of triples `(subject, property, value)`

```
(http://…/doc.html , author , #jthomas)
(#jthomas , firstname , "Jeremy")
(http://…/doc.html , subject , "Genetics")
```

- XML syntax to exchange these on the Web
- Crawl triple stores and build graphs:



SparQL Query Language

- Query triple stores, SQL for the semantic Web
- Example : retreive long papers on genetics and sort them by alphabetic order of their title.

```
SELECT ?articles ?title
 WHERE {
  ?article rdf:type eg:Article .
  ?article eg:subject eg:Genetics .
  ?article eg:title ?title .
  ?article eg:nbpage ?nbpage .
  FILTER ( ?nbpage > 5 )
 }
 ORDER BY ?title
```

RDF Schema and OWL

- Publish and define the vocabulary used in the triples
  e.g. author, first name, article, etc.
- Describe hierarchies of concepts and relations
  e.g: article *is a kind of* document
  firstname *is a kind of* designation
  firstname *applies to* persons

- Give formal characterization of types and relations
  e.g. hasSpouse *is a symmetric relation*
  hasChild *is the inverse of* hasParent
  a Mother *is a* woman *with at least one* child

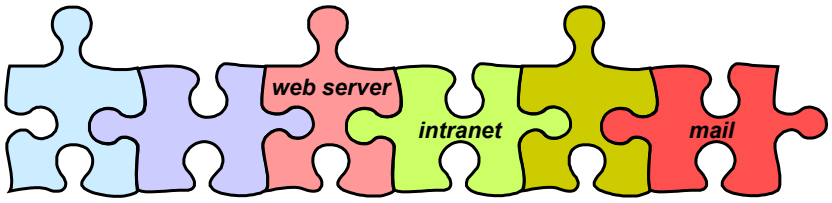
- OWL cake with three flavours: Lite / DL / Full

On intranets and extranets too,
a little semantics goes a long way.

corporate semantic web

- **Past work…**
  - Objectives: Offer methodological and software support (i.e. models, methods and tools) for construction, management and diffusion of corporate memories.
  - Corporate memory : Explicit and persistent materialization of crucial knowledge and information of an organization to ease access, sharing and reuse by the members of the organization in individual and collective tasks.
  - Corporate memories as corporate semantic webs

Corporate web & intranet



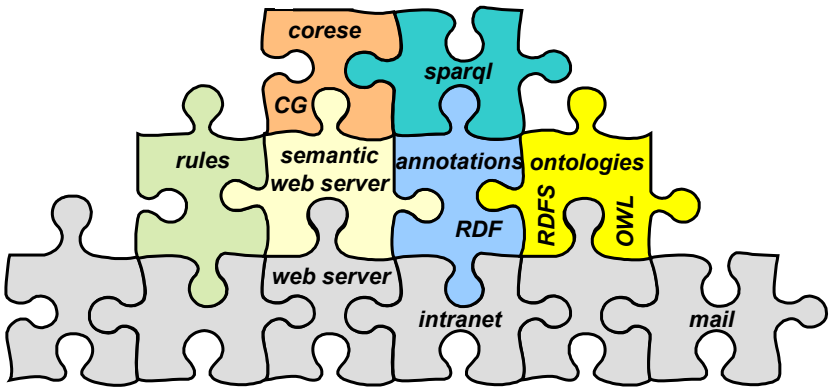
13

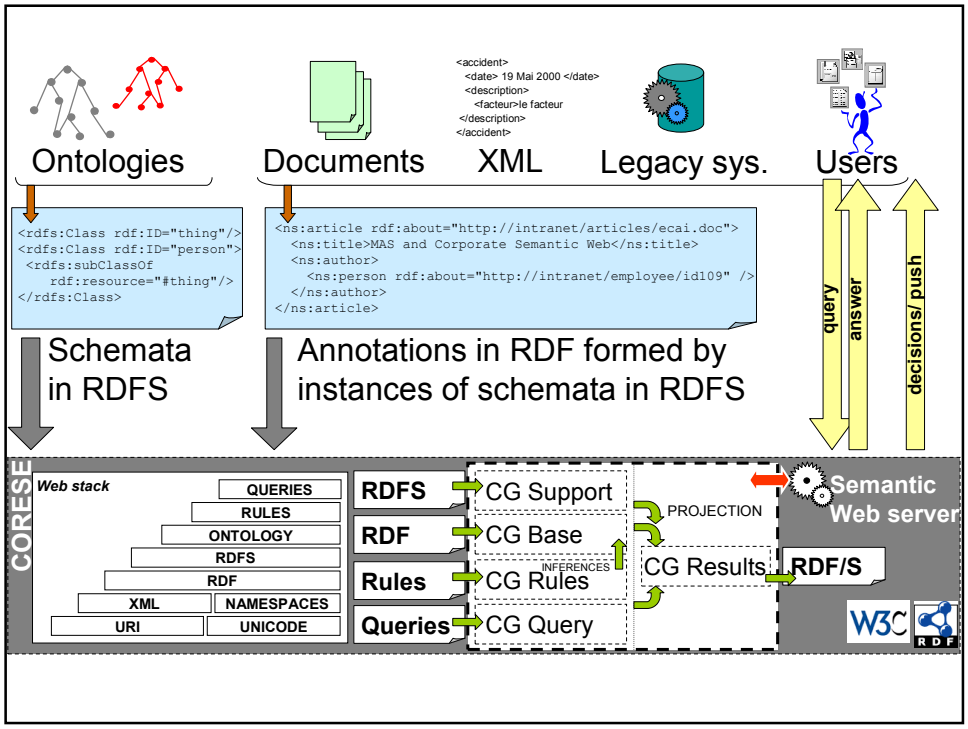
Corporate semantic Web

- Resources: persons, documents (XML, HTML...), services, software, hardware, etc.
- Ontologies: describing the conceptual vocabulary shared by the organization communities
- Semantic annotations: on these resources (e.g. persons' skills, document contents, characteristics of services/software/hardware), using the vocabulary defined in the ontologies
- Diffusion on the intranet / corporate web.

14

Corporate semantic web





SeWeSe

- To reduce the amount of time spent to develop new semantic web applications.
- JSP and Servlet technologies
- Provides a set of filters, servlets, JSP tags and libraries as well as some templates to build new applications

```
<ul>
  <stl:for-each-result
    query="SELECT ?name WHERE { ?x humans:name ?name }">
    <li>${name}</li>
  </stl:for-each-result>
</ul>
```

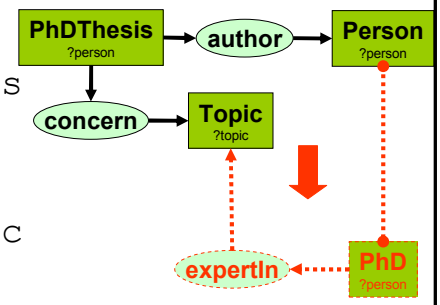
17

Production rules

Classify a resource

If a person wrote a Ph.D. thesis on a suject the s/he is a doctor and an expert on that subject.

```
?person author ?doc
?doc rdf:type PhDThesis
?doc concern ?topic
→
?person expertIn ?topic
?person rdf:type PhD
```



18

Web services & Enterprise application

- Transversal use of enterprise modeling
  - End of 90's: enterprise **modeling for KM**
  - In the past 4 years: technology and **application integration** can benefit from these models too
- Evolution of KM scenarios
  - Until end of 90's focus on: knowledge capture, storage, access and diffusion
  - More and more often: computation, decision, routing, transformation,… **knowledge workflow**
- Unified and integrated access to **knowledge sources and corporate applications**

19

Corporate information systems evolution

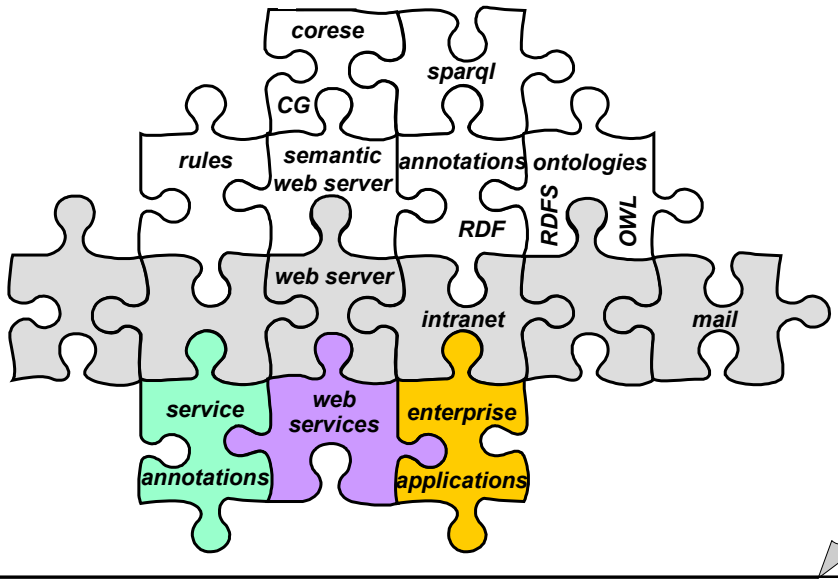
- Corporate memories on intranets providing:
  - information **capture** services;
  - information **storage** services;
  - information **computation** and **inference** services;
  - information **flows management** services;
  - information **mediation** services;
  - information **presentation** services.
- Resources may be **internal or external** (external standard library & online services)
- Interoperate smoothly and **integrate workflows** at the business layer.

20

Integrating corporate application

Integrating corporate services

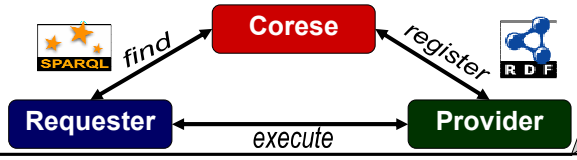
Corporate web services



23

Corporate semantic web services

- Consider services just like other (web) resources and annotate them with the vocabulary defined in semantic web schemas (e.g. OWLS)
  - Types of services (directory, buying service, delivery, etc.)
  - Types of inputs (people names, ISBN, dates & places, etc.)
  - Types of outputs (phone, confirmation, etc.)
  - Quality of service, duration, cost, etc.
- Rely on semantic search engines to discover services and match a request; enhance UDDI with inferences
  e.g. a phonebook *is a kind of* directory



24

Service description

```
<service:Service rdf:ID="PosteService_Secretaire">
   <service:presents
       rdf:resource="#Profile_Poste_Service_Secretaire"/>
   <service:describedBy rdf:resource="#PosteSecretaire"/>
   <service:supports rdf:resource="#PosteGrounding_Secretaire"/>
</service:Service>

<profile:Profile rdf:ID="Profile_Poste_Service_Secretaire">
    <service:presentedBy rdf:resource="#PosteService_Secretaire"
    <profile:has_process rdf:resource="#PosteSecretaire"/>
    <profile:serviceName>PosteSecretaire</profile:serviceName>
    <profile:hasInput rdf:resource="#PosteSecr_input"/>
    <profile:hasOutput rdf:resource="#PosteSecr_output"/>
</profile:Profile>
```

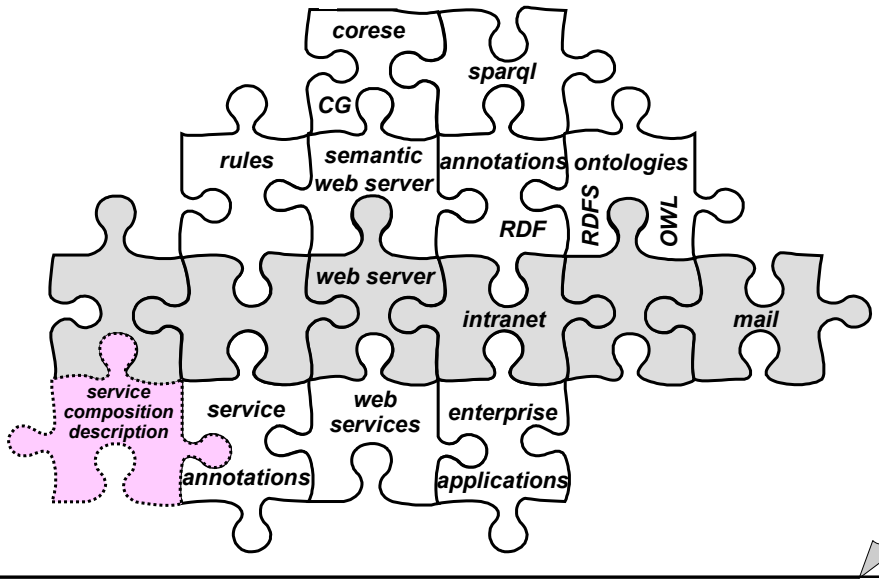
25

Service description

```
<process:AtomicProcess rdf:ID="PosteSecretaire">
 <process:hasInput>
   <process:Input rdf:ID="PosteSecr_input">
    <process:parameterType>&xsd;#string</process:parameterType>
    <process:semanticType rdf:resource="&doc;#EmployeeName"/>
   </process:Input>
 </process:hasInput>
 <process:hasOutput>
   <process:Output rdf:ID="PosteSecr_output">
     <process:parameterType>&xsd;#string</process:parameterType>
     <process:semanticType rdf:resource="&doc;#AssistantPhone"/>
    </process:Output>
  </process:hasOutput>
</process:AtomicProcess>
```

26

Corporate semantic web applications



27

Composing services

- Link output(s) of a service with input(s) of another; compose processes to create applications
- Different types of composition:
  - Manual composition e.g. an IT manager describes a useful composition and makes it public
    Macro-recording interface
  - Semi-automatic composition: assist selection
  - Simple composition: find sequences of services
  - Fully automatic composition (???)
    - Capture and decompose end-users' needs?
    - Even small examples seem to require a lot of domain knowledge
    - Controlled workflow description

28

Discover paths between resources

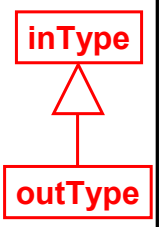
Find a link between two persons
(with maximal length of 4 relations)

```
?x rdf:type ex:Person
?y rdf:type ex:Person
?x ex:relation[4] ?y
```

29

Definining composable services

```
<cos:rule>
    <cos:if>
   ?s1 rdf:type proc:Process
   ?s2 rdf:type proc:Process
   ?s1 proc:hasInput ?input
   ?s2 proc:hasOutput ?output
   ?s1 != ?s2
   ?input proc:semanticType ?inType
   ?output proc:semanticType ?outType
   ?outType rdfs:subPropertyOf ?inType
    </cos:if>
    <cos:then>
        ?s2 proc:composable ?s1
    </cos:then>
</cos:rule>
```

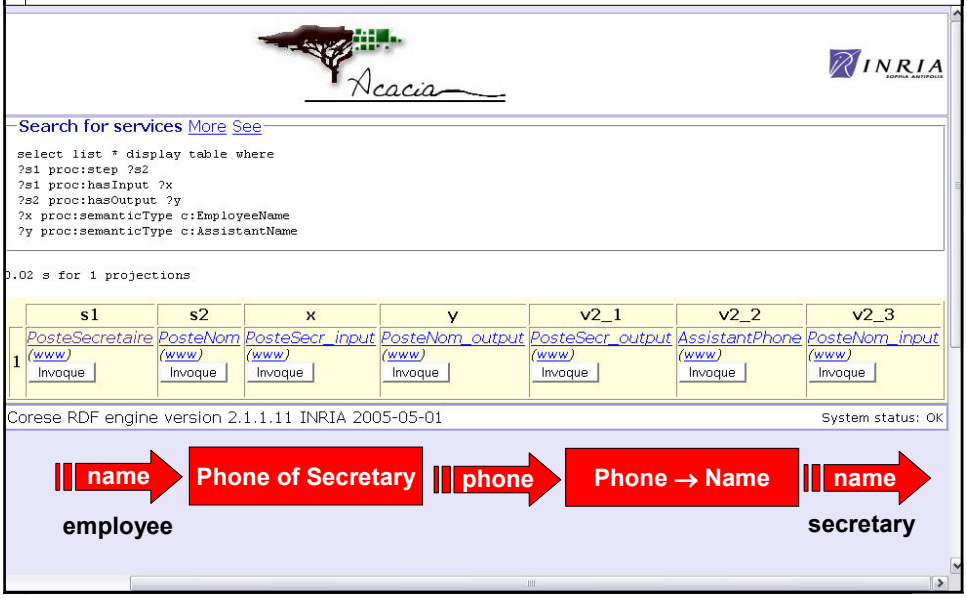


30

Finding composable services

```
?s1 all::proc:composable[N] ?s2
?s1 proc:hasInput ?param1
?s2 proc:hasOutput ?param2
?param1 proc:semanticType c:BookName
?param2 proc:semanticType c:BookBuyNotification
```

Automatic sequences



A memory with no intelligence
is doomed to decay

An intelligence with no memory
is doomed to go around and around

Composing services and memory

- **Services for memory & memory for services**
  - Intelligence: composing services and knowledge
  - Mapping input types to queries
    - associate to service inputs a predicate to identify candidate values
    - formally define these predicates using rules
    - invocation form (pre)populated potential inputs.
  - Corese as a semantic web service itself
    - to use the result of a query over the corporate memory as a service input;
    - to use a service output to add knowledge to the memory

36

Input description & extension

```
<process:AtomicProcess rdf:ID="PosteSecretaire">
  <service:describes rdf:resource="#PosteService_Secretaire"/>
  <process:hasInput>
    <process:Input rdf:ID="PosteSecr_input">
      <process:parameterType>&xsd;#string</process:parameterType>
      <process:semanticType rdf:resource="&doc;#EmployeeName"/>
    </process:Input>
  </process:hasInput>

  <process:hasOutput>
    <process:Output rdf:ID="PosteSecr_output">
      <process:parameterType>&xsd;#string</process:parameterType>
      <process:semanticType rdf:resource="&doc;#AssistantPhone"/>
    </process:Output>
  </process:hasOutput>
</process:AtomicProcess>
```

37

Rule defining the predicate

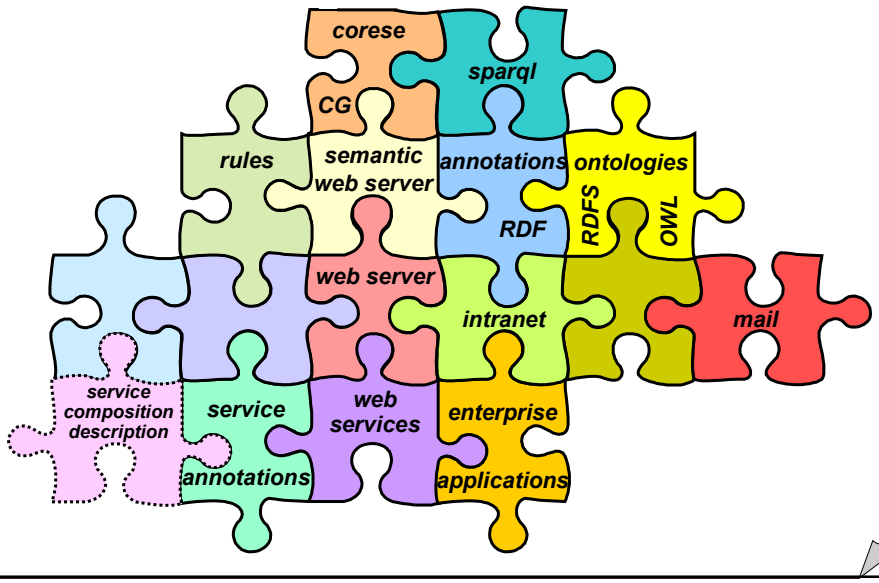
```
<cos:rule>
    <cos:if>
     ?x rdf:type c:Employee
     ?x c:Name ?n
    </cos:if>
    <cos:then>
     ?x c:EmployeeName ?n
    </cos:then>
</cos:rule>

(…)

<c:Employee rdf:ID='ML'>
    <c:Name>Moussa Lo</c:Name>
</c:Employee>
```

38

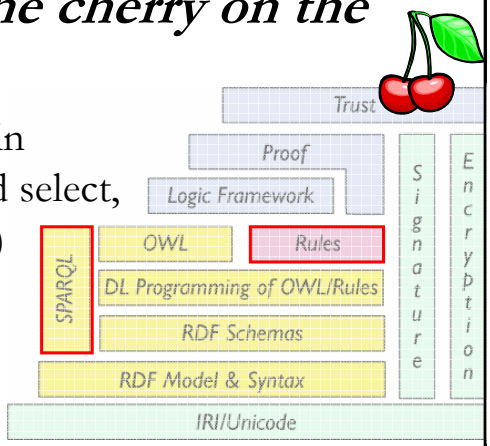
Corporate semantic web puzzle



39

Perspectives or "*the cherry on the cake*"

- SPARQL and data flows in choreographies (input and select, output and assert/update)
- Rules in workflows and semantic integration
- Visual frameworks to manage services, composition, choreography, orchestration, etc.
- Dynamic interfaces generation adaptation.
- Current application scenarios…

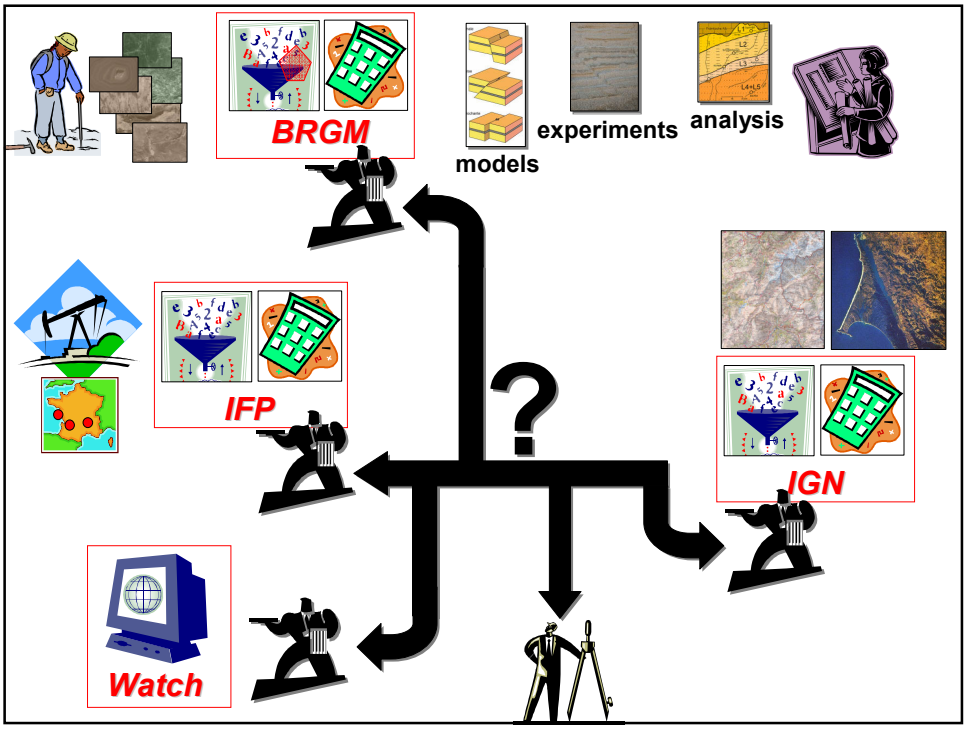


The semantic web cake © W3C

40

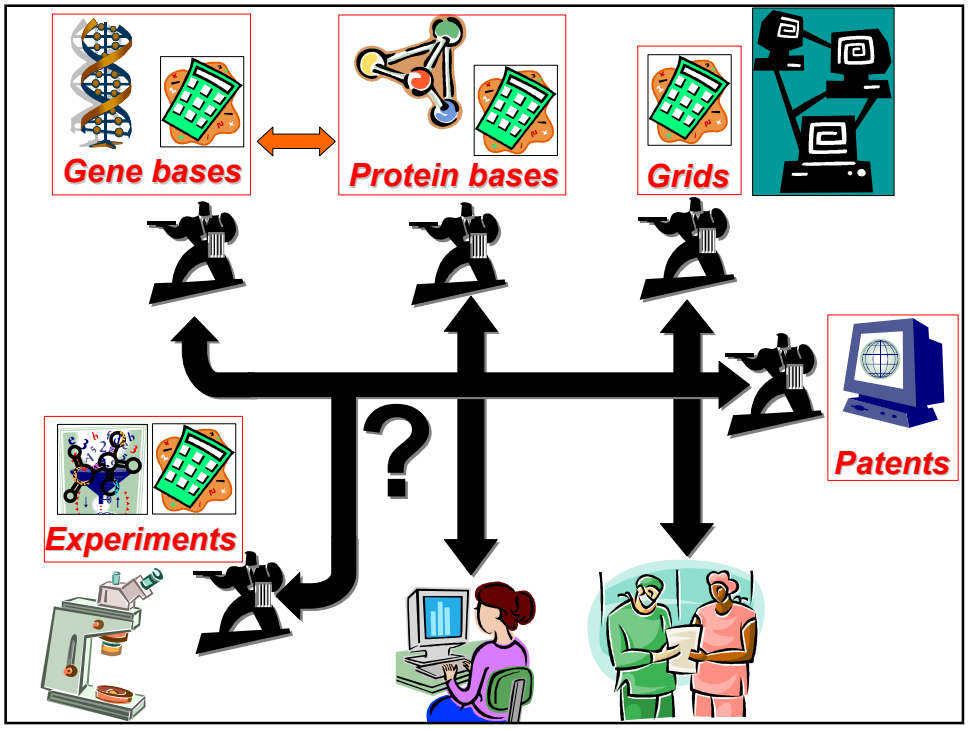
Scenario #1: the eWok hubs

- Members: IFP, INRIA, ENSMA, EADS, BRGM, ENSMP, CRCFAO
- Cooperation between different organisations (companies, institutes, etc.) participating in an engineering workflow: projects to capture et and store CO_2 reusing oil drillings.
  - Integrate information bases and domain/expert tools from different partners
  - Requirements: "*a set of communicating portals providing web applications (for humans), web services (for machines) and information resources (for humans and machines).*"



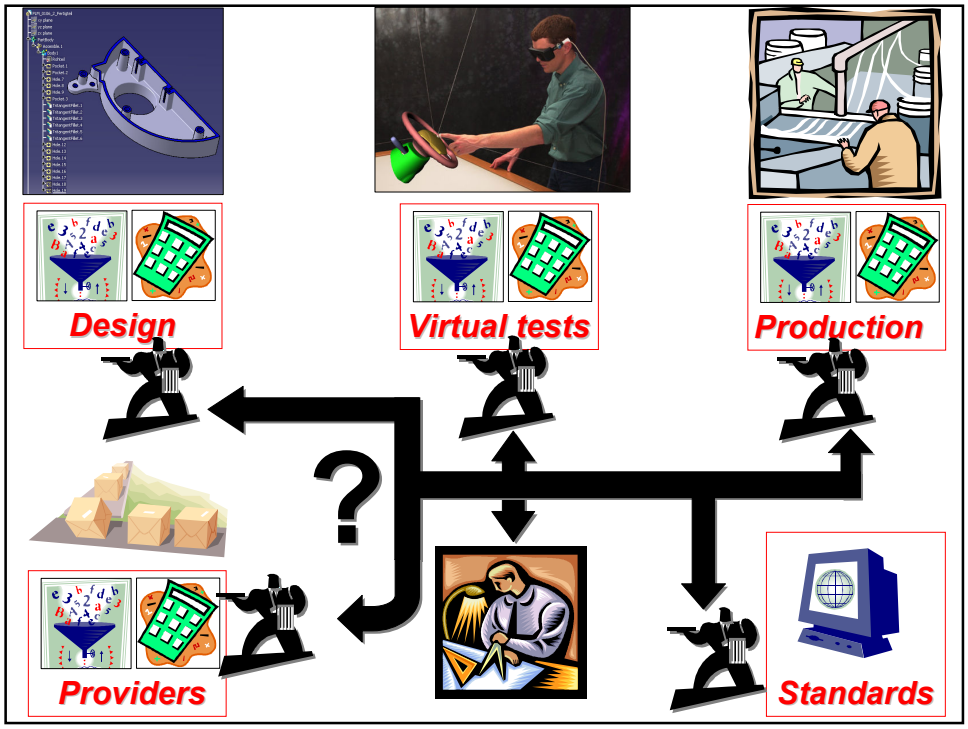
Scenario #2: SeaLife world-wide lab

- Members: TU Dresden, Hariot-Watt University, City University, University of Manchester, Scionics, INRIA
- Computational and data infrastructure to facilitate researches in Life sciences
  - Cooperation of geographically distributed organisations
  - Seamless integration of their computational and data resources
- Requirements: *"Browsers that can automatically identify entities such as protein and gene names, molecular processes, diseases, types of tissue, etc. and the relationships between them, in any Web document, collect these entities and then apply further analyses to them using applicable Web and Grid services."*



Scenario #3: SevenPro and eDesign

- Members: Semantic Systems, INRIA, Fraunhofer, Czech Technical University, Living Solids, Italdesign-Giuigiaro, Fundiciones del Estanda
- Allow the integration of tools used in industrial design
  - Mining of engineering knowledge in multimedia repositories
  - 3D interaction with that knowledge
- Requirements: "*Engineering environments integrating CAD tools, document repositories, ERP, virtual reality rendering, corporate Databases, etc. to improve the process of product engineering and development in manufacturing and engineering companies.*"



- < Questions />