Data and Service Integration

Sonia Bergamaschi – Università di Modena e Reggio Emilia
Francesco Guerra – Università di Modena e Reggio Emilia
Andrea Maurino – Università di Milano – Bicocca
Carlo Batini – Università di Milano - Bicocca

Service integration

Outline

- What is service integration
- Web Service Basic
  - Description (SQL DDL)
  - Discovery (QUERY)
  - Composition (JOIN)

- The state of the art (with respect to the problem of data & service integration)
- Future Trends (with respect to the problem of data & service integration)

Service integration

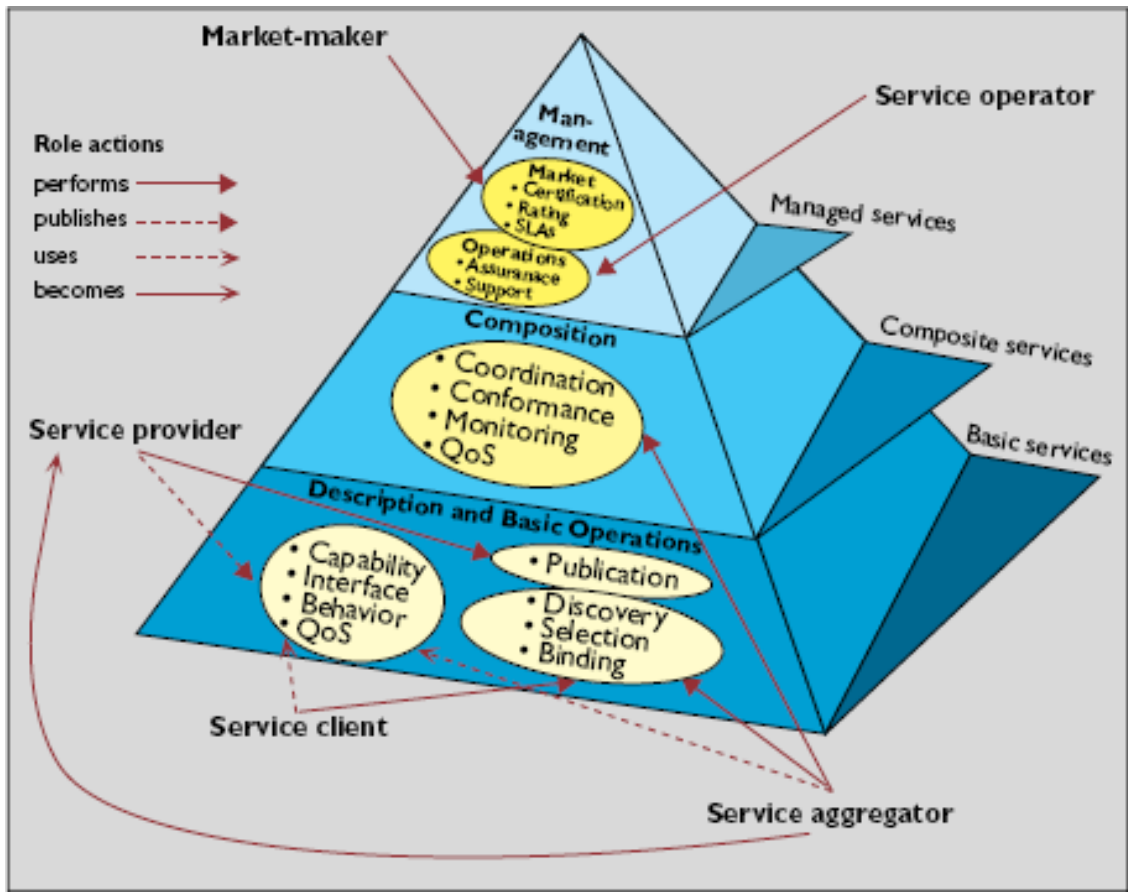
- Any process by which services are appropriately integrated or combined at either the level of direct contact with the individual client or between providers serving these individuals.
- Integrated services can be provided by an individual, a company that assumes responsibility for providing integrated services to the client.
- Service integration can be realized by means of Service Oriented Computing
- Services are self-describing, open components that support rapid, low-cost composition of distributed applications. [Papazoglou and Georgakopulos 2003]
- The application of SOC on the Web is manifested by Web services.

Web Services

- A Web service is a software system designed to support interoperable machine-to-machine interaction over a network. It has an interface described in a machine-processable format (specifically WSDL). Other systems interact with the Web service in a manner prescribed by its description using SOAP-messages, typically conveyed using HTTP with an XML serialization in conjunction with other Web-related standards

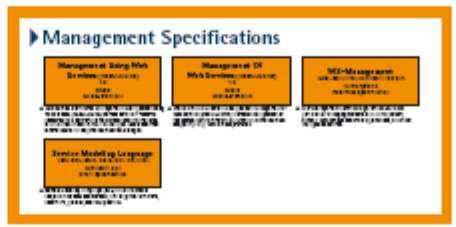
SOA Pyramid

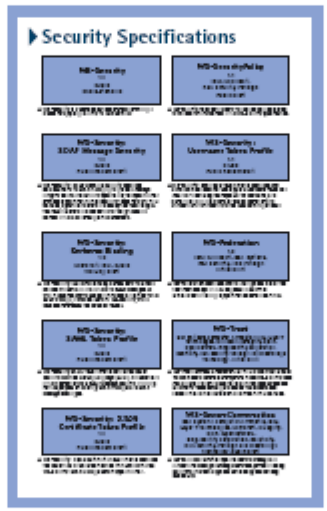
- Basic Service
  - Description
  - Pubblication
  - Discovery
  - Selection
- Composite service
  - Orchestration
- Managed Services
  - SLA
  - Configuration

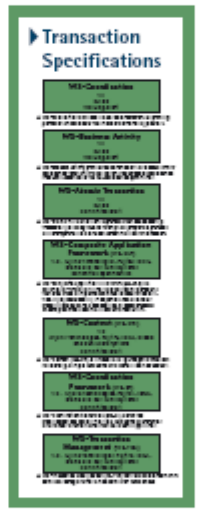


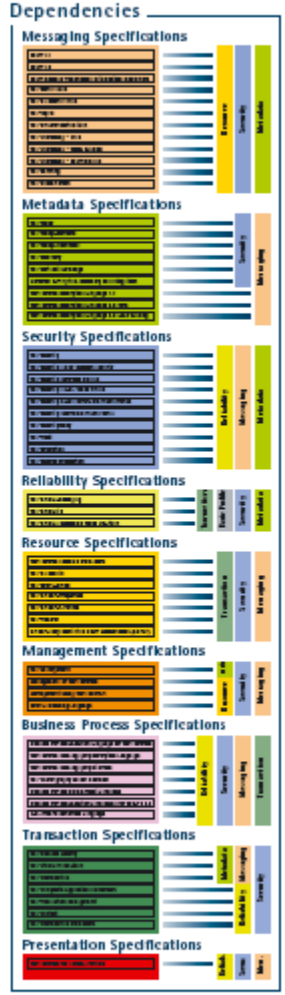
- Papazoglou and Gerogakopoluos 2003

Web Services Standards Overview









innoQ Deutschland GmbH
Halskestraße 17
D-40880 Ratingen
Phone +49 2102 77 162-100
info@innoq.com · www.innoq.com

innoQ Schweiz GmbH
Gewerbestrasse 11
CH-6330 Cham
Phone +41 41 740 01 11

Version 3.0 · February 2007

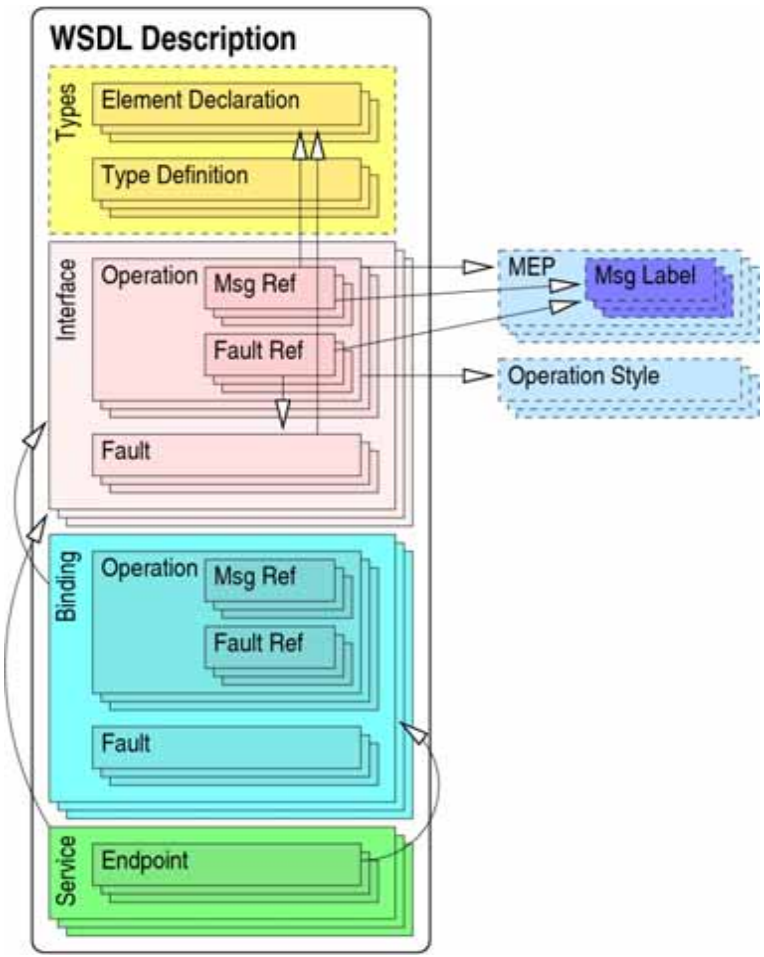
Service Description 1/3

- Web Services Description Language Version 2.0 (WSDL 2.0) provides a model and an XML format for describing Web services. [WSDL20 2007]
- WSDL 2.0 introduce the distinction between the description of the abstract functionality offered by a service (what) and concrete details of a service description (how and where that functionality is offered).
- At an abstract level, WSDL 2.0 describes a Web service in terms of the messages it sends and receives;
  - messages are described independent of a specific wire format using a type system, typically XML Schema.

Service Description 2/3

- Abstract Service
  - An *operation* associates a **message exchange pattern** with one or more messages.
  - A *message exchange pattern* identifies the sequence and cardinality of messages sent and/or received as well as who they are logically sent to and/or received from.
  - An *interface* groups together operations without any commitment to transport or wire format.
- At a concrete level,
  - a *binding* specifies transport and wire format details for one or more interfaces.
  - An *endpoint* associates a network address with a binding.
  - A *service* groups together endpoints that implement a common interface.

Service Description 3/3



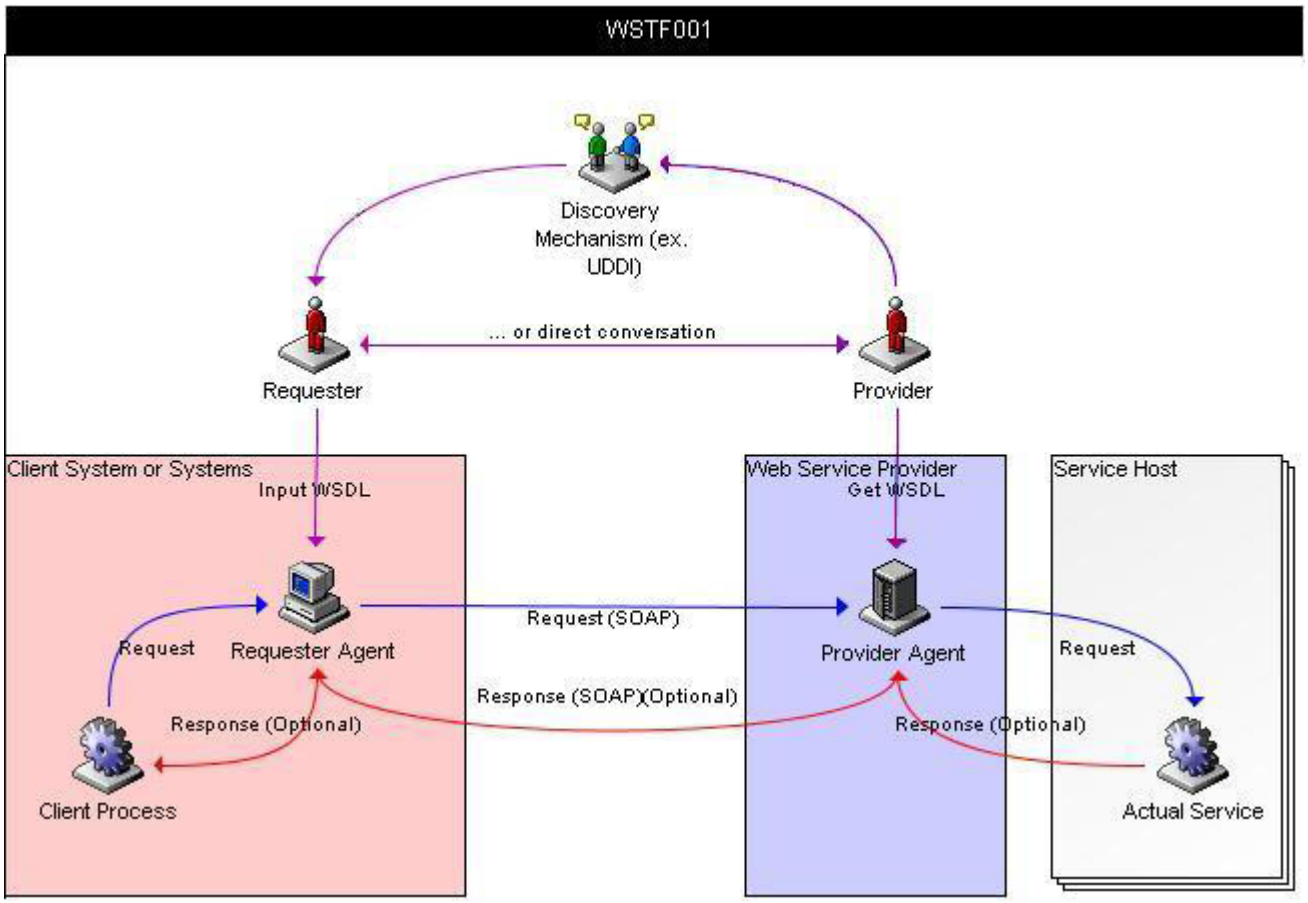
Service Publication 1/2

- Universal Description Discovery and Integration registry [UDDI20 2002]
- UDDI is a web-based registry that exposes information about Web Services. These registries are run by multiple Operator Sites, and can be used by anyone who wants to make information available, as well as anyone that wants to find that information.
- The benefit of having access to this information is to provide a mechanism that allows others to discover what technical programming interfaces are provided for interacting with a business for such purposes as electronic commerce, etc.

Service Publication 2/2

- The information that a business can register includes several kinds of simple data
  - “who, what, where and how”.
- WHO (AKA White Pages): Simple information about a business – such as name, business identifiers and contact information
- WHAT (AKA Yellow Pages): classification information that includes industry codes and product classifications, as well as descriptive information about the services that the business makes available.
- WHERE & HOW (Green Pages): provisioning of information about the URL through which each type of service is accessed and registering references (tModels) to information about interfaces and other properties of a given service.

Service Discovery



Service Composition 1/2

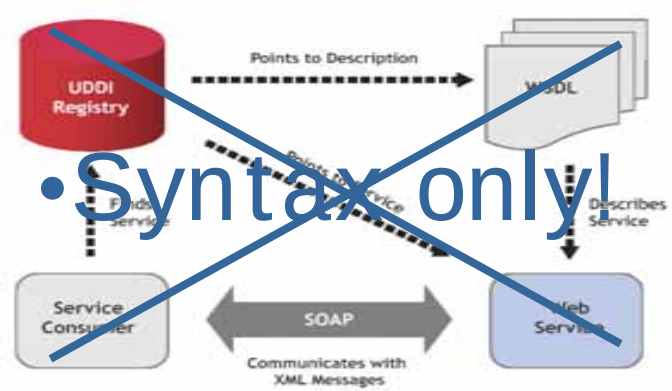
- Models for business interactions typically assume sequences of peer-to-peer message exchanges, both request-response and one-way, within stateful, long-running interactions involving two or more parties.
- A formal description of the message exchange protocols used by business processes in their interactions is needed.
- The Business Process Execution Language for Web Services (BPEL4WS) is a language to specify business processes and business interaction protocols.
  - By defining a model and a grammar for describing the behavior of a business process based on interactions between the process and its partners.
- WS-BPEL also introduces systematic mechanisms for dealing with business exceptions and processing faults.

Service Composition 2/2

- The business interaction protocols are called abstract processes. They are used to specify public and visible message exchange between different parties involved in a business protocol and they do not reveal the internal behavior or the implementation of the involved parties.
- The executable processes on the other hand are like workflow descriptions represented using basic and structured activities specifying a pattern of execution of web services.
- The process model defined by WS-BPEL is based on the WSDL based service description model.
- The services (described as partners in BPEL spec) that the process invoke/reply using basic activity are represented using their WSDL description.
- An executable process itself can be a Web service by itself and the interface of that process can be represented using WSDL.

SOA meets the Semantic Web

- In 2001 the WWW founder proposed the idea of “The semantic Web” [BHL01]
- Where semantics of information on the web is defined, making it possible for the web to understand and satisfy the requests of people and machines to use the web content
- From the merriage between Web service and the semantic web in 2001 semantic web services are born
- Two waves can be identified within the semantic web services (SWS)



The first wave

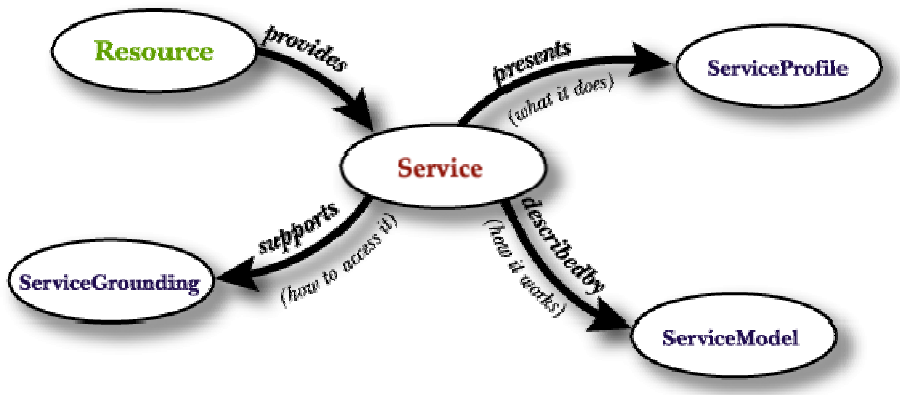
- DAML-S Darpa Agent Model Language for Services [AB+01] then evolved in OWL-S Web Ontology Language for services [MP+01]
  - USA based
- WSMO Web service modeling Ontology [RK+05] [FK+08]
  - EU based
- Other approach
- Meteor-s [PO+04]
- All approaches want to add a new semantic layer on the top of the Web Service stack (WSDL/SOAP/UDDI/WS-BPEL) and provide more flexible (and less syntax only) tools for discovery and composition

OWL-S

High-level description of a service and its provider

- description of service (human readable)
- specification of functionalities service provides
- functional attributes (requirements and capabilities)

- Profile used for
  - populating service registries
    - Also QoS
  - automated service discovery
  - matchmaking

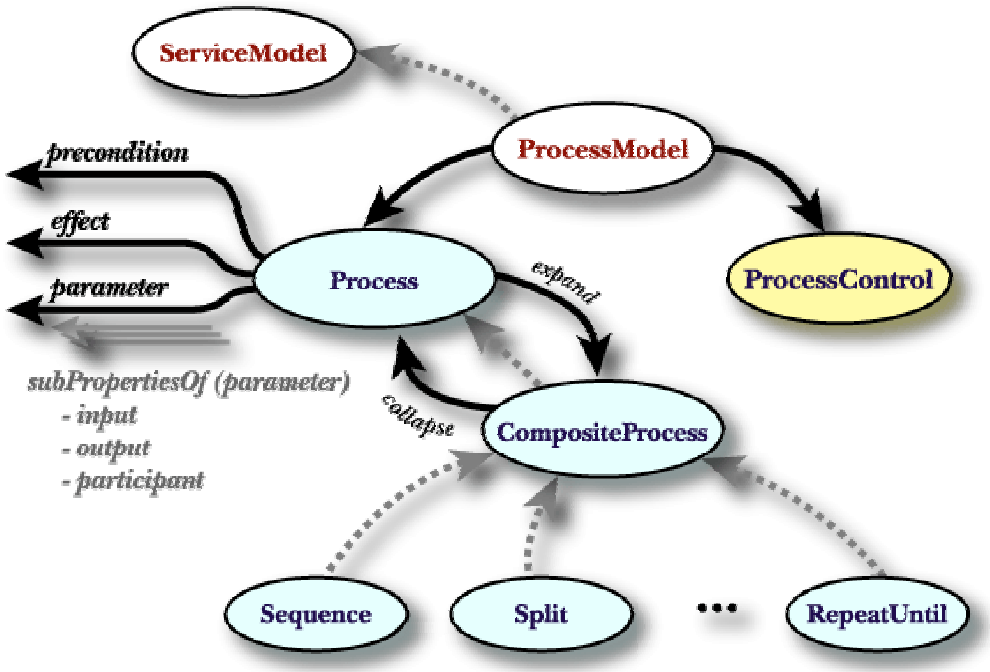


OWL-S Service Profile

- **Preconditions**
  - Set of conditions that should hold prior to service invocation
- **Inputs**
  - Set of necessary inputs that the requester should provide to invoke the service
- **Outputs**
  - Results that the requester should expect after interaction with the service provider is completed
- **Effects**
  - Set of statements that should hold true if the service is invoked successfully.
- **Service type**
  - What kind of service is provided (eg selling vs distribution)
- **Product**
  - Product associated with the service (eg travel vs books vs auto parts)

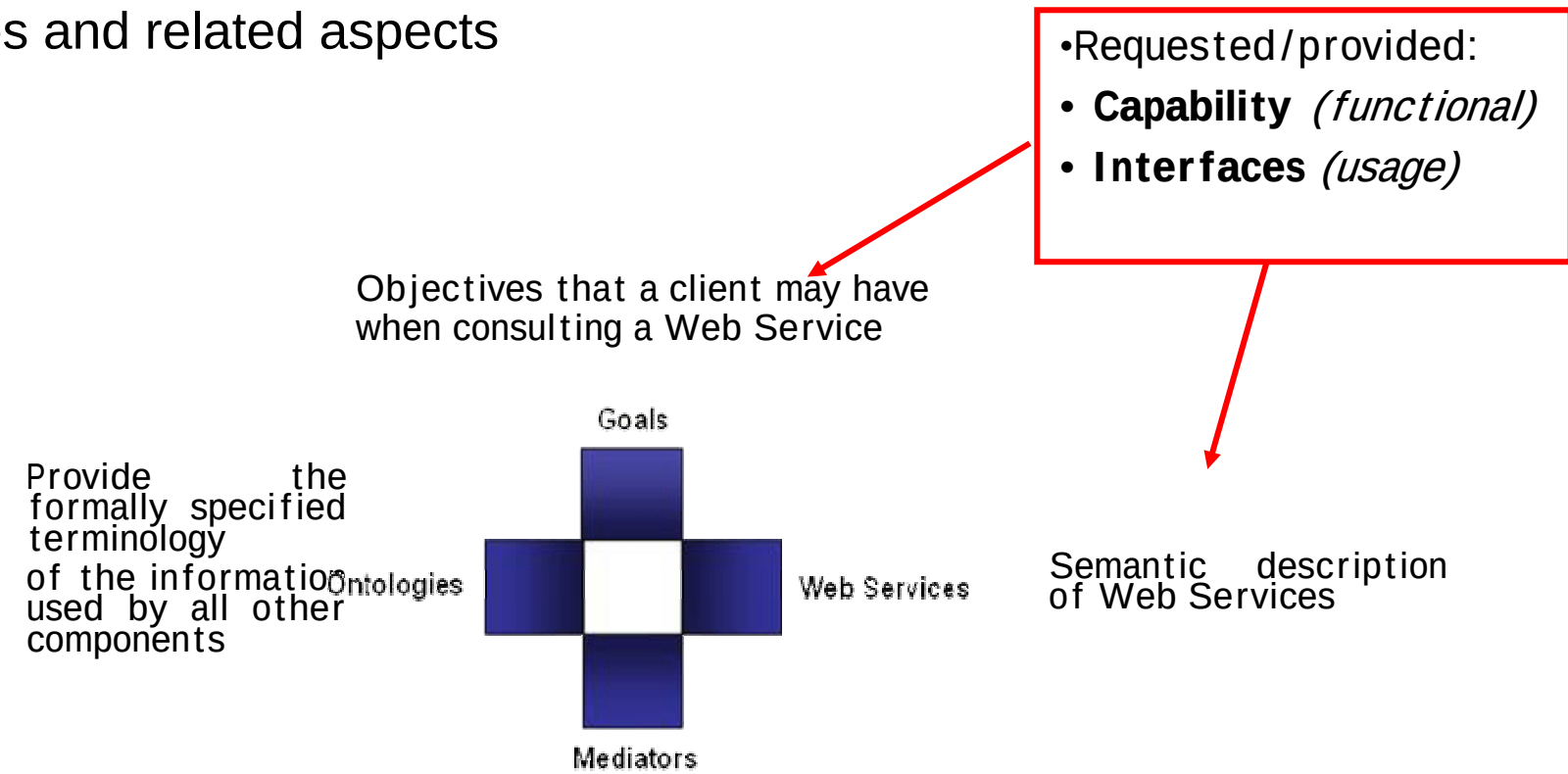
OWL-S Process

- A Process represents a transformation (function).
  - ***Inputs***: the inputs that the process requires
  - ***Preconditions***: the conditions that are required for the process to run correctly
  - ***Outputs***: the information that results from (and is returned from) the execution of the process
  - ***Results***: a process may have different outcomes depending on some condition
    - ***Condition***: under what condition the result occurs
    - ***Constraints on Outputs***
    - ***Effects***: real world changes resulting from the execution of the process



WSMO

- WSMO is an ontology and conceptual framework to describe Web services and related aspects



Web services

- **Pre-conditions**
  What a web service expects in order to be able to provide its service. They define conditions over the input.
- **Assumptions**
  Conditions on the state of the world that has to hold before the Web Service can be executed and work correctly, but not necessarily checked/checkable.
- **Post-conditions**
  describes the result of the Web Service in relation to the input, and conditions on it.
- **Effects**
  Conditions on the state of the world that hold after execution

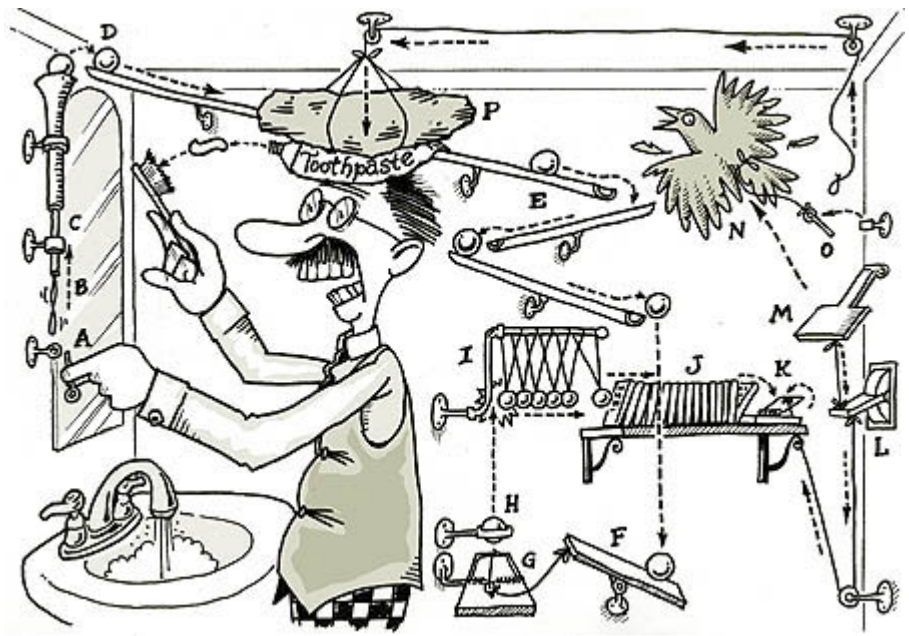
OWL-S, WSMO and Web Service

	OWL-S	**WSMO**	***Web Services Infrastructure***
Discovery ***What it does***	Profile	Web Services (capability)	*UDDI API*
Choreography ***How is done***	Process Model	Orchestration + choreography	*WS-BPEL*
Invocation ***How to invoke***	Grounding+ WSDL/SOAP	Grounding	*WSDL/SOAP*

The new wave

- A lot of papers published, a lot of founded project developed.
- Some conclusions:
- All semantic approaches are too heavy for real usage.
  - It is hard to “ontologize” the world
- Reasoning tools cannot operate in real time with big ontologies
- The answer is “kiss”

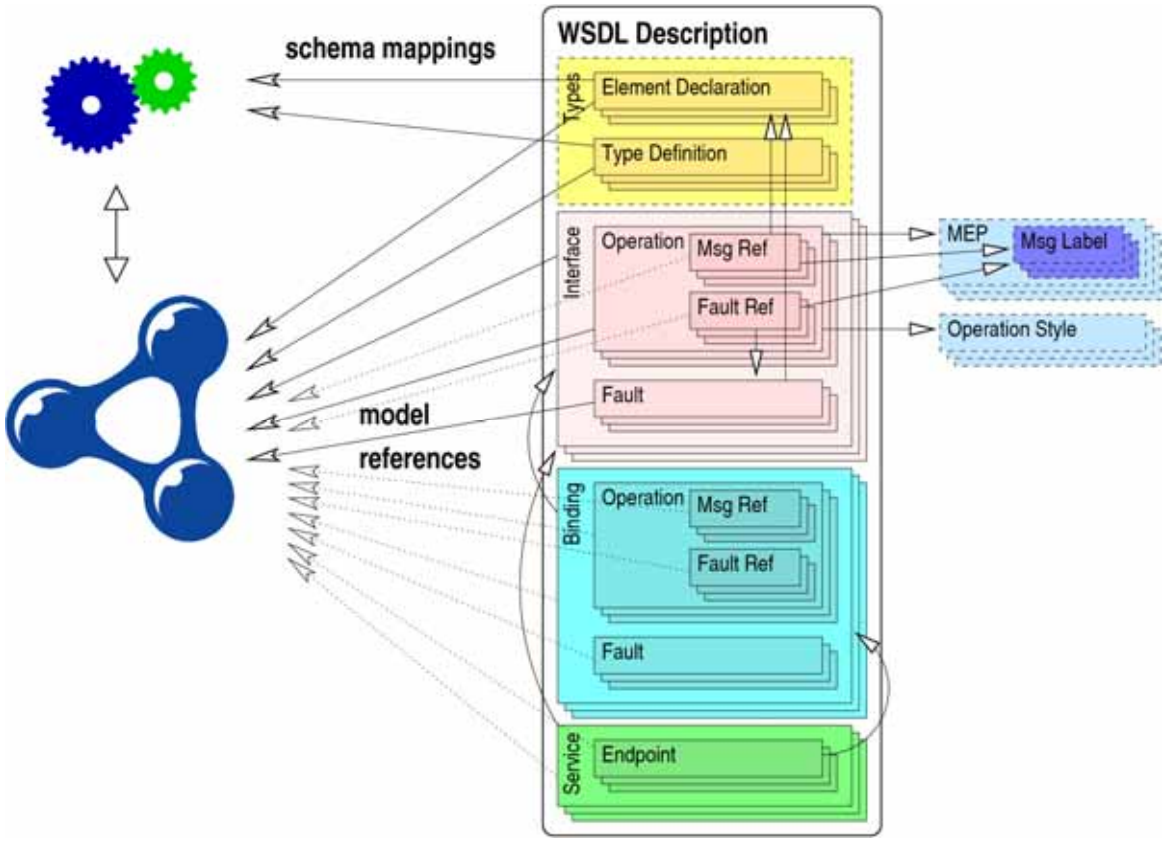
“Keep It Simple Stupid”



The second generation of SWS

- SA-WSDL [SAWSDL07] Semantic annotation for WSDL
  - SAWSDL defines how to add semantic annotations to various parts of a WSDL document such as input and output message structures, interfaces and operations.
  - The semantic annotations reference a concept in an ontology or a mapping document.
- Four service semantic can be added
- Functional
  - What the service does
- Information model
  - For handling data
- Behavioral
  - How the client talks to the service
- Nonfunctional
  - Policies, QoS, price, location etc.

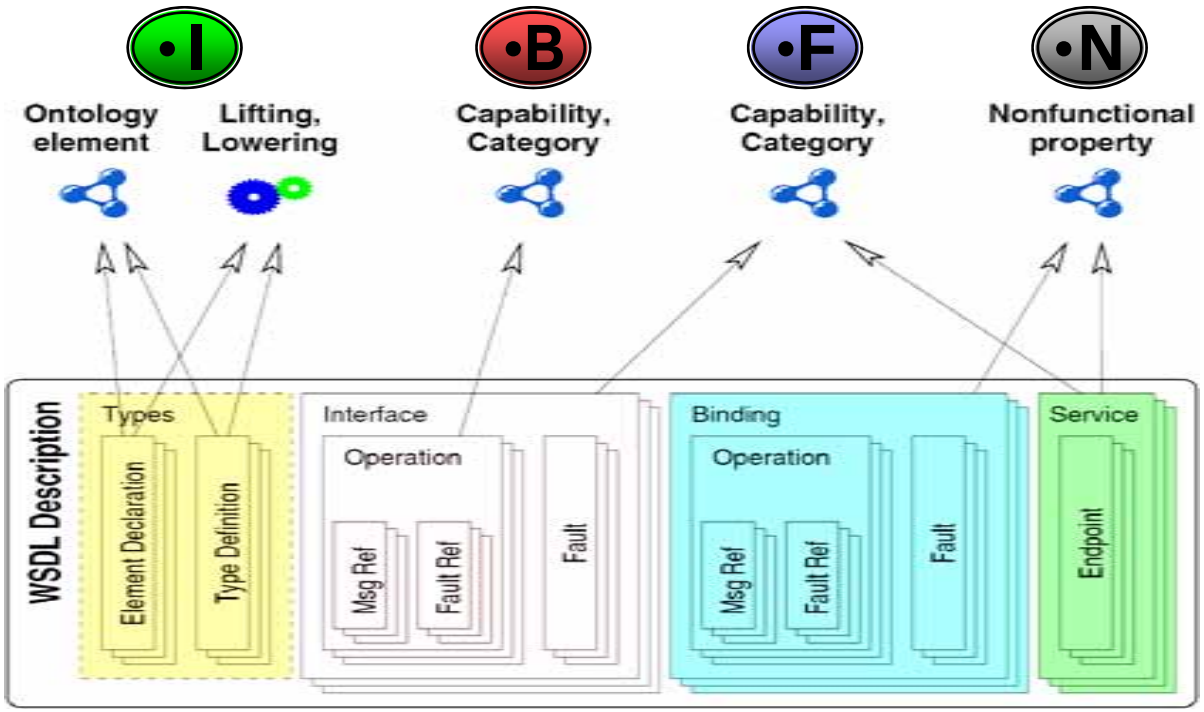
SAWSDL



•slides from Jacek Kopecký

WSMO-Lite

- WSDL-based service model
- SAWSDL-based annotations

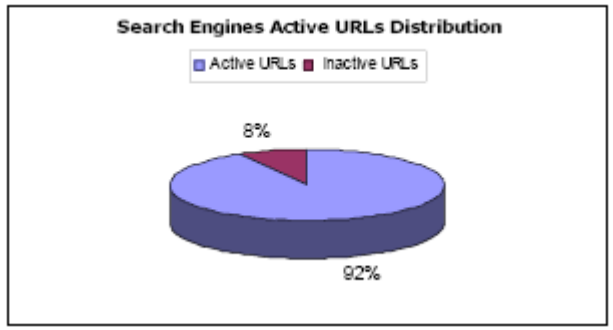


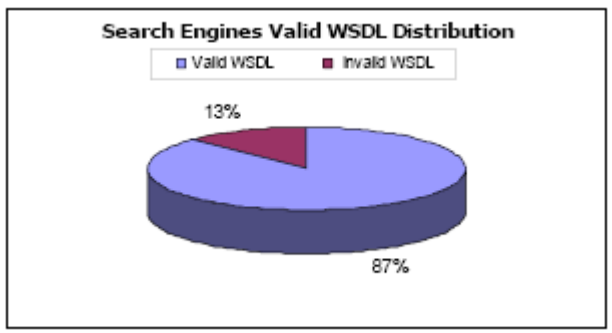
New trends in service description

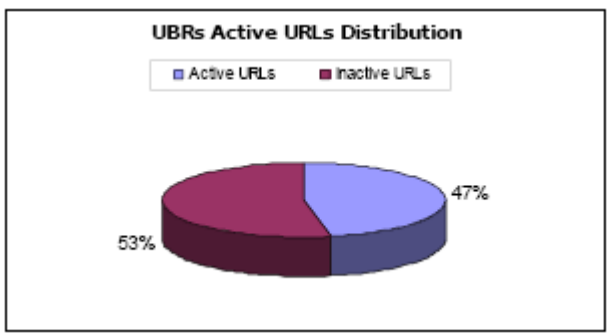
- Semantic annotation
  - Move first generation semantic approach towards the new wave
  - *OWL-S [PW+07]*
  - WSMO [VK+08]
  - SAWSDL [KV+07]
- NFP Description
  - Policy center metamodel [CD+08]
- QoS Description
  - See selelection and compisition
- QoS evaluation
  - There is the need to measure online QoS
    - [VA09][LC+09]
    - [ZL+07]
  - Composition [MJ+08]

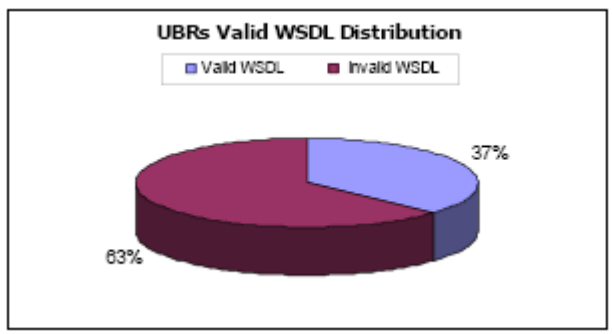
New trends in service pubblication

- Centralized registry
  - By example URBE [PP09]
  - Semantic oriented [SP07]
- Distributed registry
  - See later
- Other kinds of repository/registry
  - WWW
    - [AM08] about 5000 WS









New trends in service discovery

- The basic Web service model is an hybrid P2p where UDDI is the registry.
  - P2P oriented discovery [AB07] [ZL+09]
  - Mobile ad Hoc Service [AK+06][IS09]
  - Grid [LY+08]
  - Cloud
- Semantic Web service discovery [PT09]
  - Matric space based [DG+09]
  - Graph techniques.[BK07]
  - P2P [SS08]
  - User Feedback [AK+09]
- Compostion oriented
  - Semantic based [BC+09]

New trends in service discovery

- Secure Web Service Discovery
  - Collaborative environment[SB07]
- Matchmaking
  - SAWSDL[KK+09]:
  - OWL-S [KB+09]
  - IRS-III [DC+08]

New trends in service selection

- NFP based selection
  - WSMO [TR+07]
  - Discovery and selection engine[CC+08]
  - PCM based [CD+07]
- QoS based selection
  - Trust and Reputation [DH+08]
  - Federated UDDI [ZCL+08]
  - Constraint Programming [MW+08a]
- Link analysis [MC+08]

New trends in service composition

- WS-BPEL main limitations
  - Does not address conformance and QoS
  - Deals with connectivity only, not with correctness
  - Turing-complete language, more about implementation than specification
  - Cannot verify properties of a composition result
- Other problems
  - Automatic composition
  - Dynamic camposition

New trends in service composition

- QoS oriented composition
  - Declarative [BR+07]
  - Genetic algorithm[YR08]
  - Mixed integer Programming [AR08]
  - Planning [RL+08]
  - Semantic based [Lécué09]
- Semantic oriented composition
  - Context-based [MG+07]
  - Functional similarity [SL+09]
- Adaptive composition
  - Optimization [AP07]
  - Aspect oriented [NP+07]
  - Reflective [ZS+08]

New trends in service composition

- Other approaches
  - CSP [ST07]
  - Control Flow [BK+09]
  - Choreography [MK+09]
  - WSMO Choreography [RK+08]

Mashup

- Automatic composition [PM+08]
- Language [MW+07]
- Framework
- [AR+09]
- [WL+09]

Service and data integration

Data as a Service (DaaS)

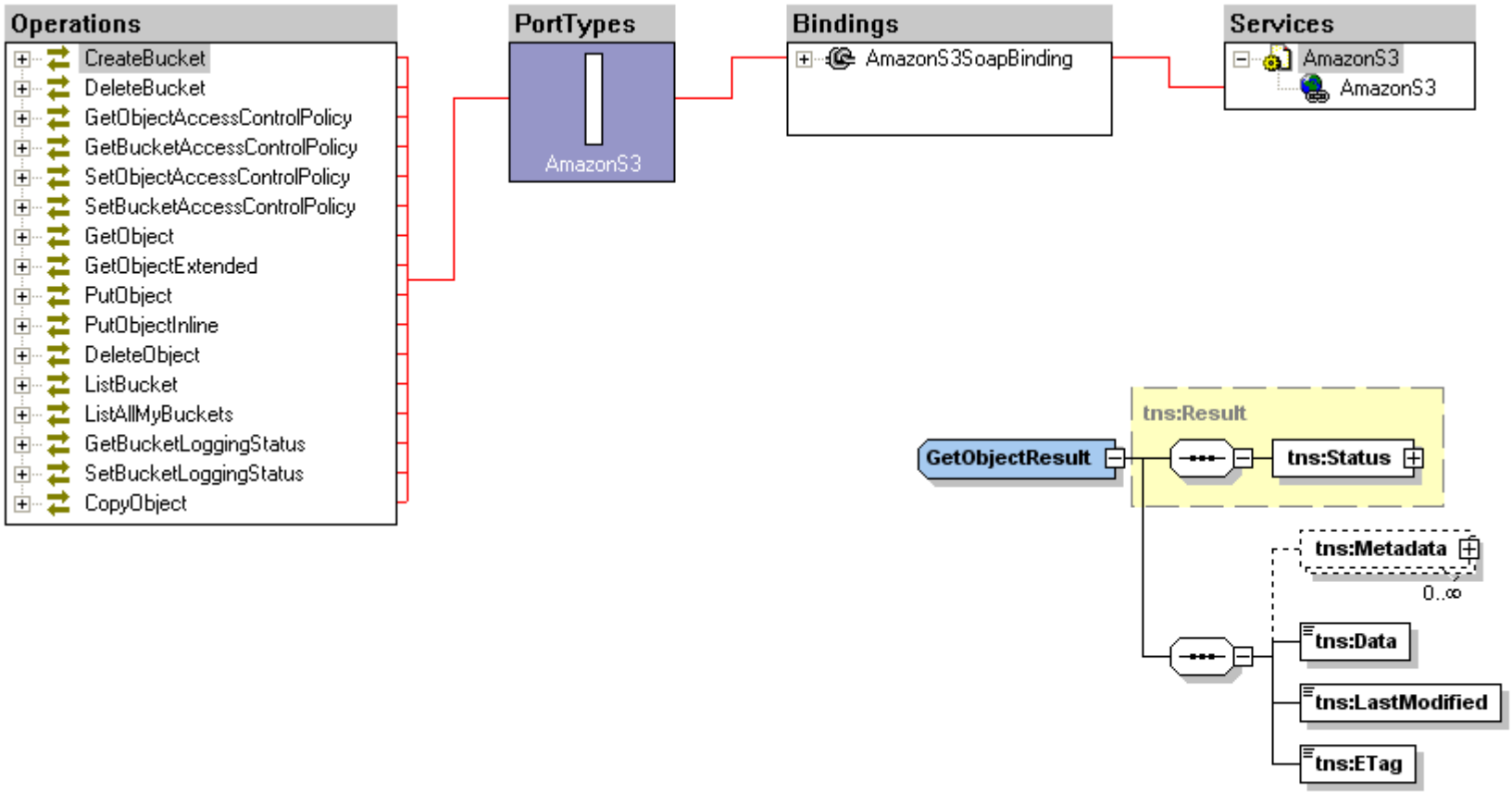
IDEA!!! ANY DATA SOURCE CAN BE EXPOSED AS A (WEB) SERVICE

Example of DaaS

- Amazon Simple Storage Service
- Amazon S3 provides a web services interface that can be used to store and retrieve
  - any amount of data
  - at any time,
  - from anywhere on the web.
- It gives any developer access to the same highly scalable, reliable, fast, inexpensive data storage infrastructure that Amazon uses to run its own global network of web sites.

- Pair Name Value storage system (object, metadata, data)
- Storage as a Service

S3 wsdl


Operations
CreateBucket
DeleteBucket
GetObjectAccessControlPolicy
GetBucketAccessControlPolicy
SetObjectAccessControlPolicy
SetBucketAccessControlPolicy
GetObject
GetObjectExtended
PutObject
PutObjectInline
DeleteObject
ListBucket
ListAllMyBuckets
GetBucketLoggingStatus
SetBucketLoggingStatus
CopyObject
PortTypes
AmazonS3
Bindings
AmazonS3SoapBinding
Services
AmazonS3
AmazonS3
tns:Result
GetObjectResult
tns:Status
tns:Metadata
0..∞
tns:Data
tns:LastModified
tns:ETag

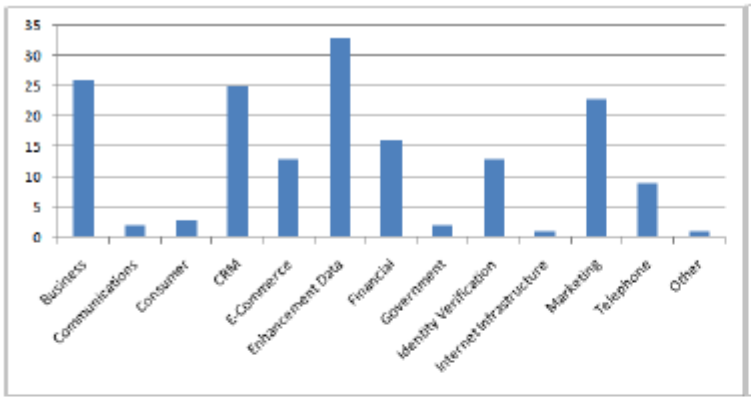
Example of DaaS

- SQL Azure
- SQL Azure will deliver a rich set of integrated services that enable you to perform relational queries, search, reporting, analytics, integration and synchronize data with mobile users, remote offices and business partners.
- It provides **Web services** that enable relational queries, search, and data synchronization with mobile users, remote offices and business partners. It can store and retrieve structured, semi-structured, and unstructu

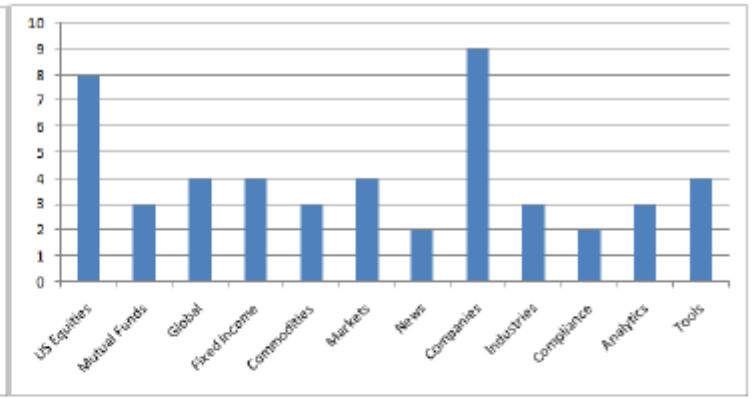
Example of DaaS

- According to [TD09] DaaSs can be categorized into
- Retrieve- only (read-only)
  - DaaS which only provides data based on existing data sources,
    - StrikeIron Address Validation and Xignite-Realtime
- CRUD DaaS
  - DaaS where the consumer can create, retrieve, update and delete data.
- The latter category can be roughly divided into:
  - infrastructure-based DaaS (Storage as a Service)
    - the consumer defines their own data schema.
  - application-specific DaaS
    - offer common data structure for the consumer to store and manage their common data.
      - Examples are the Amazon Flexible Payment Service7 and the XWebCheckOut8

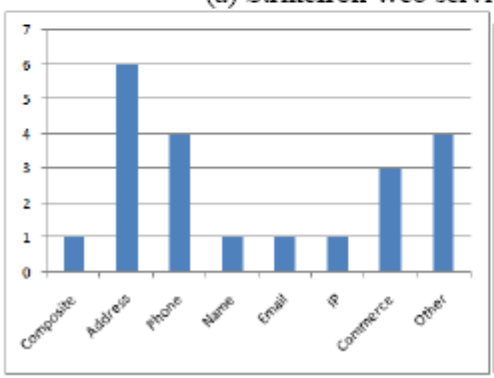
Existing DaaS [TD09]



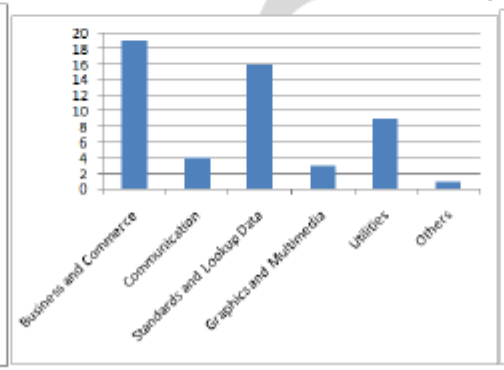
(a) StrikeIron Web services



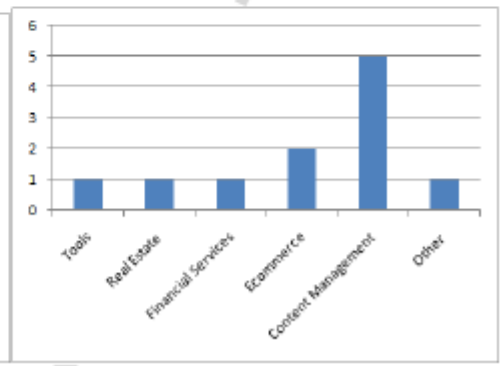
(b)Xignite Web services



(c) ServiceObjects Web services



(d)WebserviceX Web Services



(e) XWebservice Web services

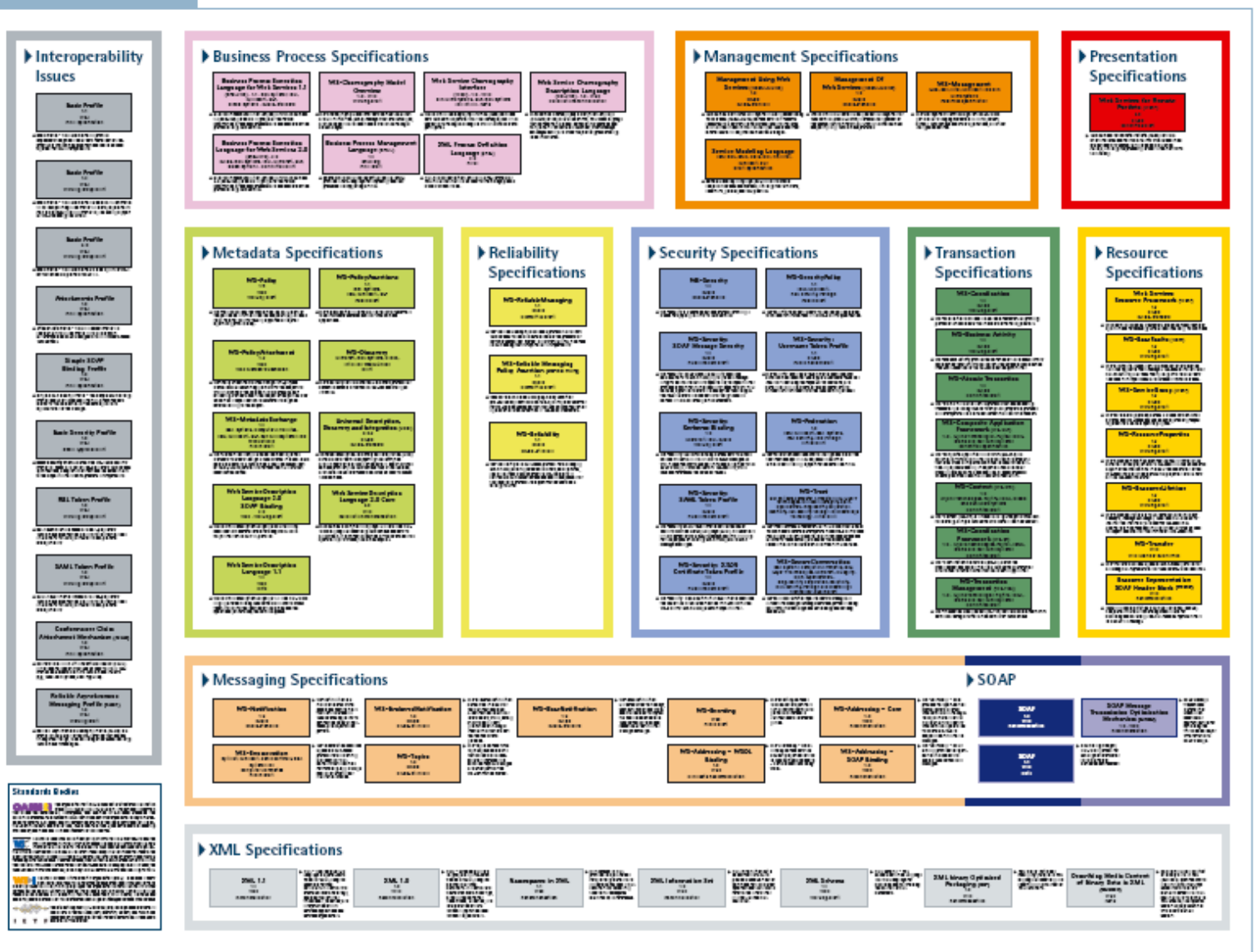
CRUD?

- What are the typical database operation?
- Create (e.g. Create Table ….)
- Query (e.g Select * from)
- Update (Update salary from emp….)
- Drop (Drop table)

- ….
- Create
- Read
- Update
- Delete

- CRUD

Web Services Standards Overview



innoQ Deutschland GmbH
Halskestraße 17
D-40880 Ratingen
Phone: +49 2102 77 162-100
info@innoq.com · www.innoq.com

innoQ Schweiz GmbH
Gewerbestrasse 11
CH-6330 Cham
Phone: +41 41 740 01 11

REST

- Representational State Transfer is a pattern of resource operations [Fielding2000]
- The traditional SOAP-based approach to Web Services uses full-blown remote objects with remote method invocation and encapsulated functionality,
- REST deals only with data structures and the transfer of their state.
- REST is basically a client server stateless application based over HTTP

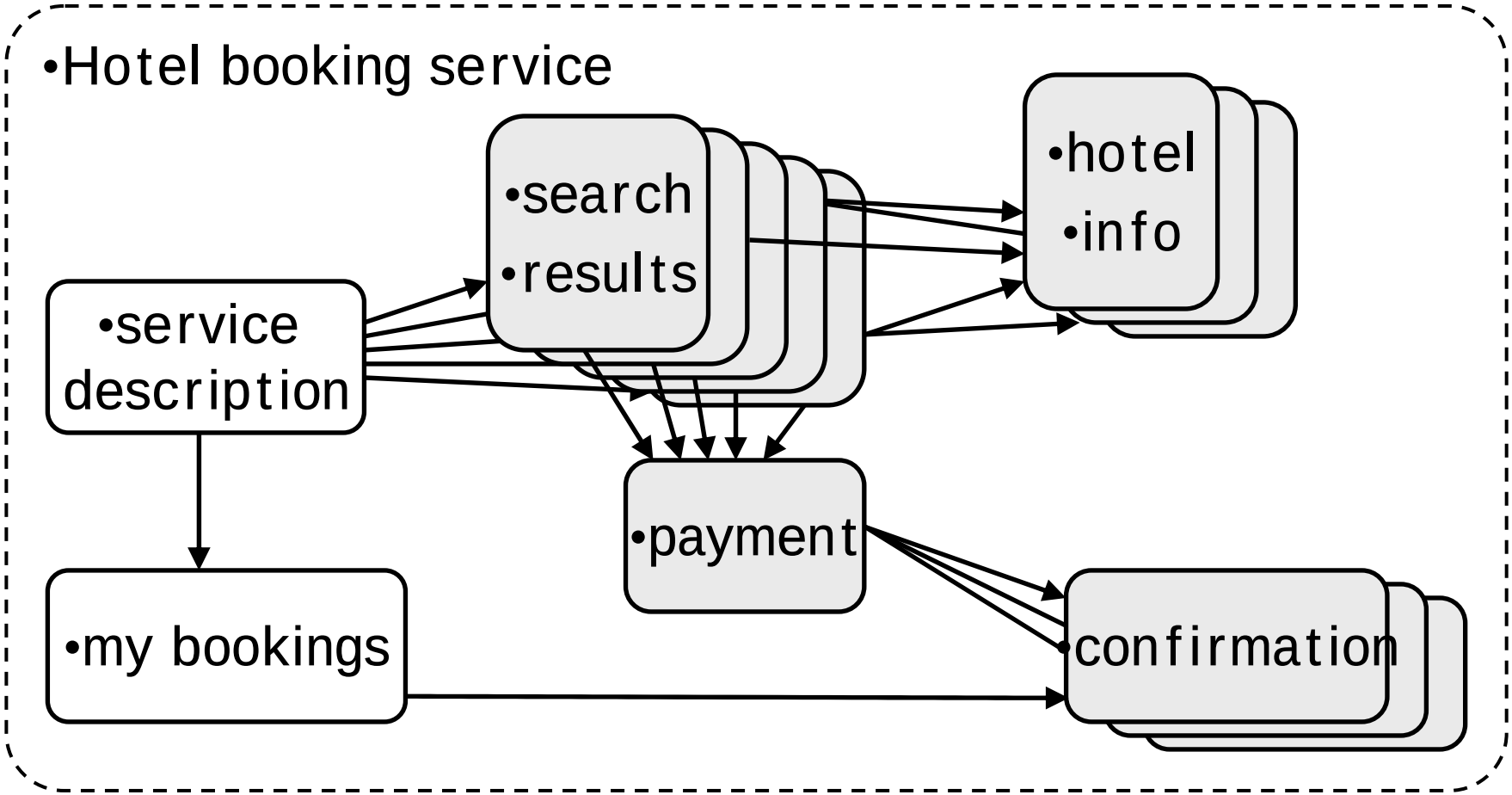
REST

- A RESTful Web service[RR07] is:
  - A set of Web resources
  - Interlinked
  - **Data-centric, not functionality-centric**
  - Machine-oriented

REST

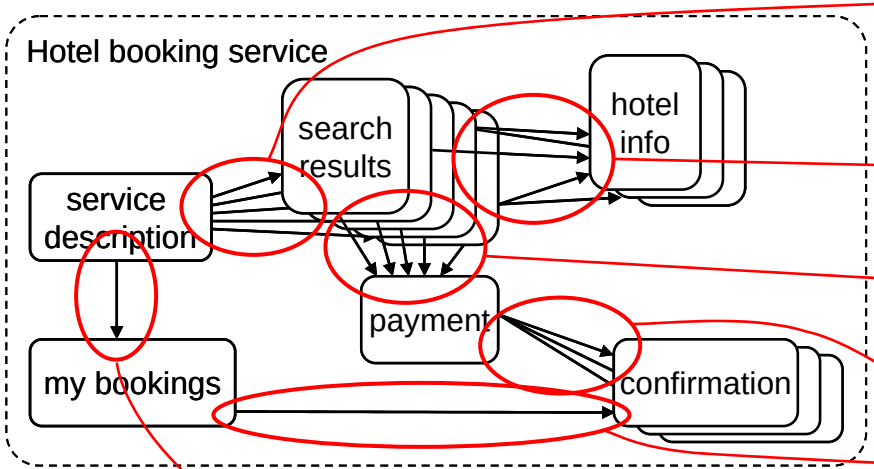
- The uniform interface that any REST interface must provide is considered fundamental to the design of any REST service.
- Individual resources are identified in requests, for example using URIs in web-based REST systems.
  - The resources are conceptually separate from the representations that are returned to the client.
- the server does not send its database,
  - HTML, XML or JSON that represents some database records encoded in UTF-8,

Example of RESTfull services



•Example from Dieter Fensel

REST example



search(date, city)

→ list of hotels & rates

getHotelDetails(hotel)

reserve(rate, creditCard)

→ confirmationID

getConfirmationDetails(confID)

listMyBookings()

→ list of confirmationIDs

Rest meets crud

- At the core of REST based design is a set of state transfer operations universal to any data storage and retrieval system.
- These operations, as commonly interpreted on the web, are referred to by CRUD
- There is an informal mapping of CRUD operations onto the commands provided by the HTTP protocol: POST, GET, PUT, and DELETE, respectively.
- These commands identify the particular CRUD operation being requested of the resource identified by the URL endpoint.

CRUD MEETS REST

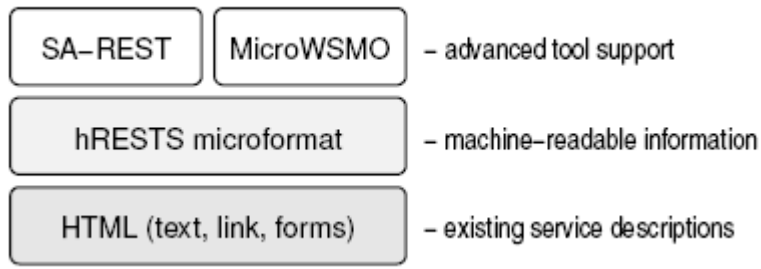
RESTful Web Service HTTP methods				
Resource	GET	PUT	POST	DELETE
Collection URI such ashttp://example.com/resources/	**List** the members of the collection complete with their member URIs for further navigation. For example list all the cars for sale.	Meaning defined as 'replace the entire collection with another collection'.	**Create** a new entry in the collection where the ID is assigned automatically by the collection. The ID created is usually included as part of the data returned by this operation.	Meaning defined as 'delete the entire collection'.
Member URI such ashttp://example.com/resources/7HOU57Y	**Read** a representation the addressed member of the collection expressed in an appropriate MIME type	**Update** the addressed member of the collection or **create** it with the specified ID.	It would imply treating the addressed member as a collection in its own right and creating a new subordinate of it.	**Delete** the addressed member of the collection.

Rest description: WADL

- While SOAP based Web services have a WSDL document that defines their operations,
  - no standard equivalent for REST services.
- Web Application Description Language (WADL) [Hadley09] from Sun Microsystems, as a specification for describing REST services.
- A WADL document is designed to be a simple alternative to WSDL for use with XML/HTTP Web applications.
- It provides a description of Web applications in a simpler format than WSDL while also defining how to generate the URI for each operation and defining the format of the input and output parameters.
- Is there the need of a rest description languages?

hREST

- "There's usually an HTML page"
  - There's no WSDL for Web apps
  - APIs described mostly in **text**
- Identifying machine-readable parts
  - Service, its operations
  - Resource address, HTTP method
  - Input/output data format
- **hRESTS** microformat [SK+08][KG+08]
  - SA-rest [SG+07]



hREST

- hRESTs provides constructs to markup operations and data elements in an API description.
- RESTful services are often described as Web APIs using HTML.
- The lack of a model like WSDL makes it difficult to use conventional service discovery approaches.

```
<p>Description of the
   ACME Hotels service:</p>
 <p>
   The operation <code>getHotelDetails</code> is
   invoked using the method <span>GET</span>
   at <code>http://example.com/h/{id}</code>,
   with the ID of the particular hotel replacing
   the parameter <code>id</code>.
   It returns the hotel details in an
   <code>ex:hotelInformation</code> document.
 </p>
```

hREST

```
<div class="service" id="svc">
  <p>Description of the
     <span class="label">ACME Hotels</span> service:</p>
  <div class="operation" id="op1">
    <p>
     The operation <code class="label">getHotelDetails</code> is
     invoked using the method <span class="method">GET</span>
     at <code class="address">http://example.com/h/{id}</code>,
     with <span class="input">the ID of the particular hotel replacing
        the parameter <code>id</code>.</span>
     It returns <span class="output">the hotel details in an
        <code>ex:hotelInformation</code> document.</span>
  </p></div>
</div>
```

hREST

- HTML for RESTful Service Description
- Introduces the service model structure
  - `service` (+ `label`)
  - `operations` (+ `address`, `method`)
  - `input`, `output`
- Can also be in RDFa
- Basis for extensions:
  - MicroWSMO adds semantic annotations

MicroWSMO

- Extends hRESTS
  - `model` for model references
  - `lifting, lowering`
- Applies WSMO-Lite semantics

MicroWSMO

```
<div class="service" id="svc">
  <p><span class="label">ACME Hotels</span> is a
    <a rel="model" href="…/ecommerce/hotelReservation">
      hotel reservation</a> service.</p> …
  <div class="operation" id="op1"><p> …
    <span class="input">A particular hotel ID replaces the param
      <a rel="model"  href="…/onto.owl#Hotel">
        <code>id</code></a>
      (<a rel="lowering" href="…/hotelID.xslt">lowering</a>).
    </span>. …
  </p></div>
</div>
```

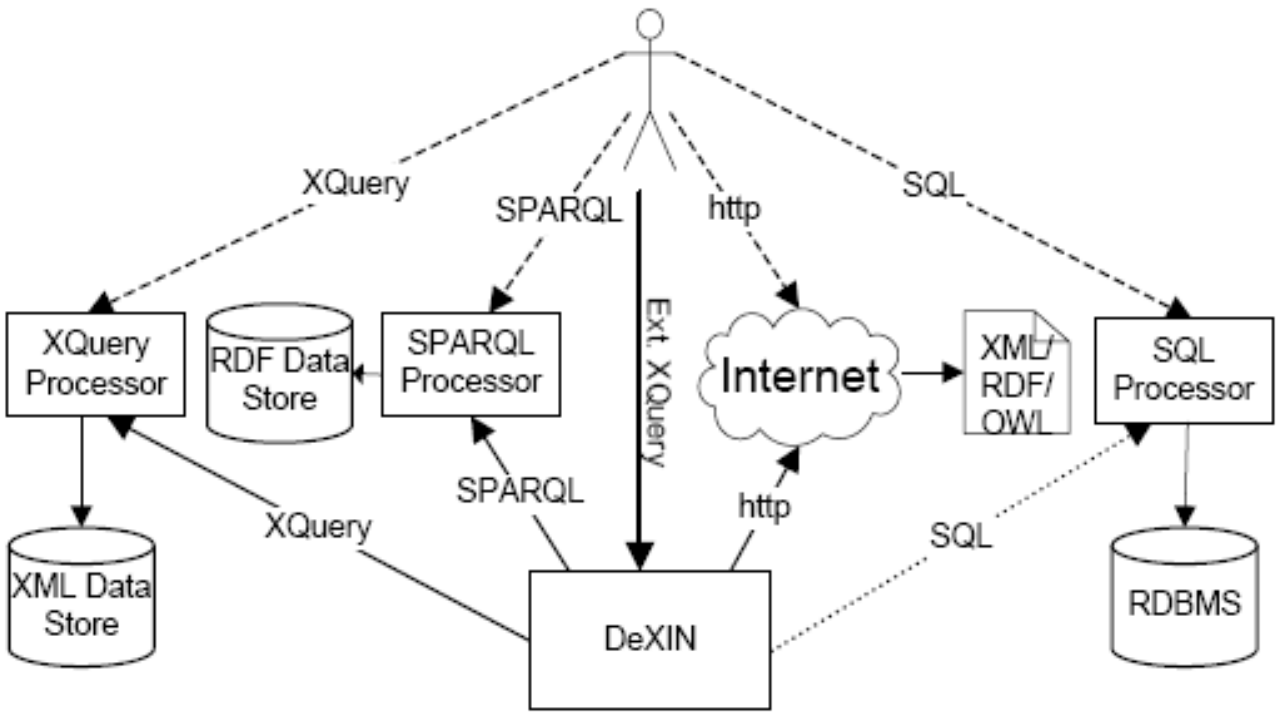
Semantics Implied in Web

- Hypermedia → behavioral semantics
  - Links become available through interaction
- Uniform interface → functional semantics
  - GET, PUT, DELETE have known effects
  - GET is *safe,* PUT and DELETE idempotent
- Self-description → information model
  - Operation output data can specify what it is

DeXIN

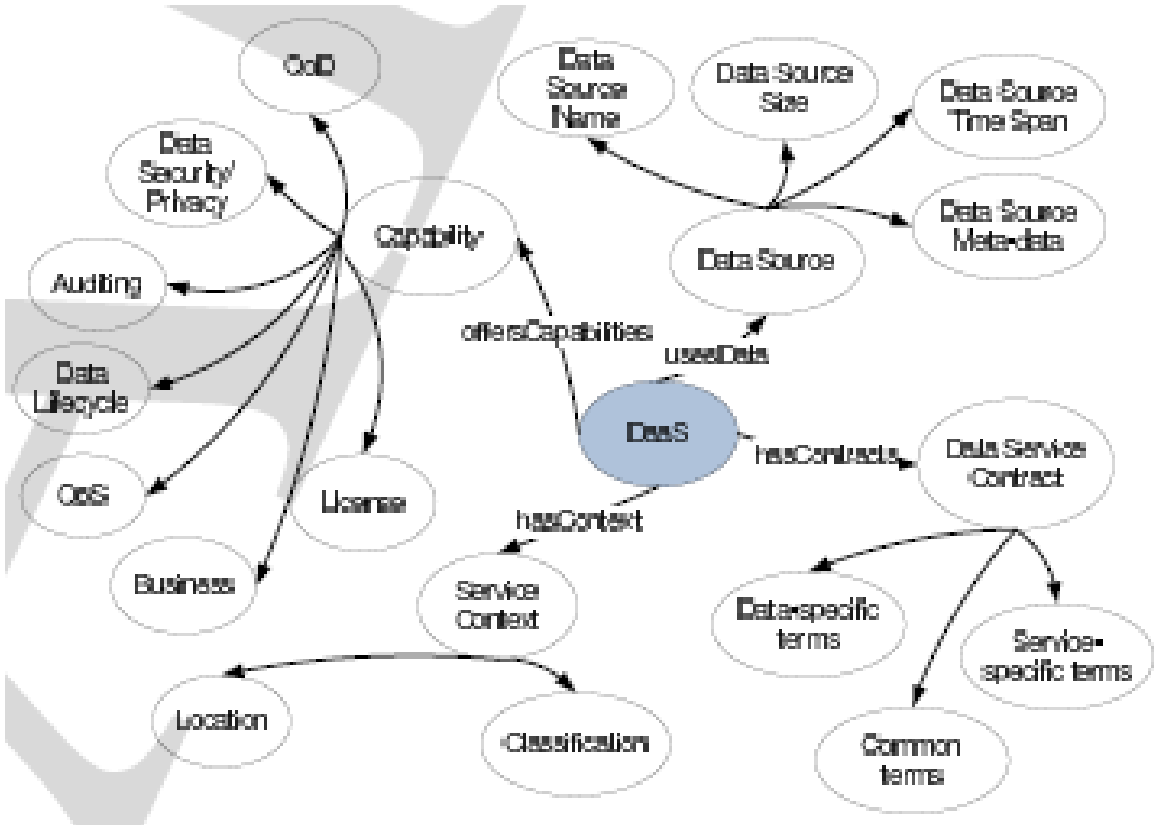
- DeXIN (**Distributed extended XQuery for data INtegration)** [PT+09]integrates multiple, heterogeneous, highly distributed and rapidly changing web data sources in different formats, e.g. XML, RDF and relational data.
- DeXIN is a RESTful data integration web service which integrates heterogeneous distributed data sources, including data services (DaaS – data as a service).
- At the heart of DeXIN is an
  - XQuery extension that allows users/applications to execute a single query against distributed, heterogeneous web data sources or data services.

DeXIN


XQuery
SPARQL
http
SQL
XQuery Processor
RDF Data Store
SPARQL Processor
Ext. XQuery
Internet
XML/ RDF/ OWL
SQL Processor
SPARQL
http
XQuery
SQL
XML Data Store
DeXIN
RDBMS

DaaS Description

- In [DT09] the problem of QoS-QoD (Data quality) is proposed



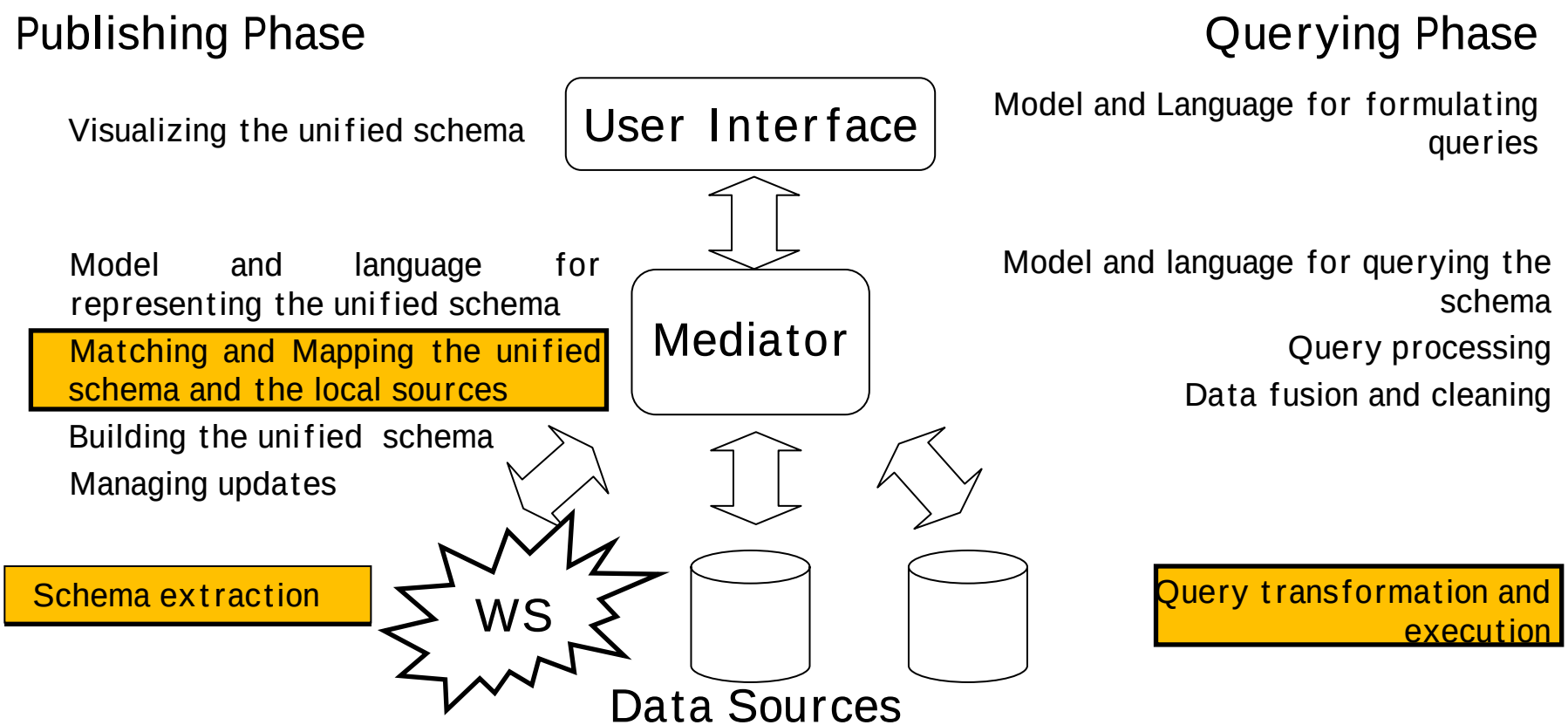
Service as Data:

Idea! Services return data after a request...

Databases return data after a query....

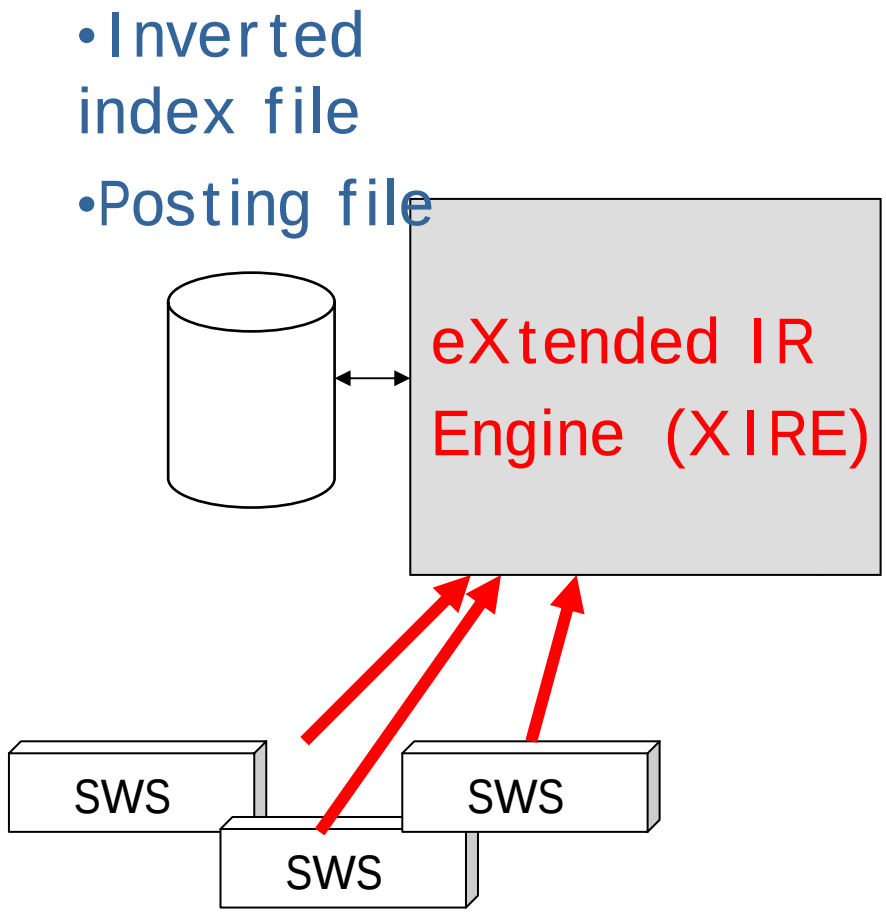
Service as Data

- If we consider a Web service as a data source we can integrate it with other data soruces



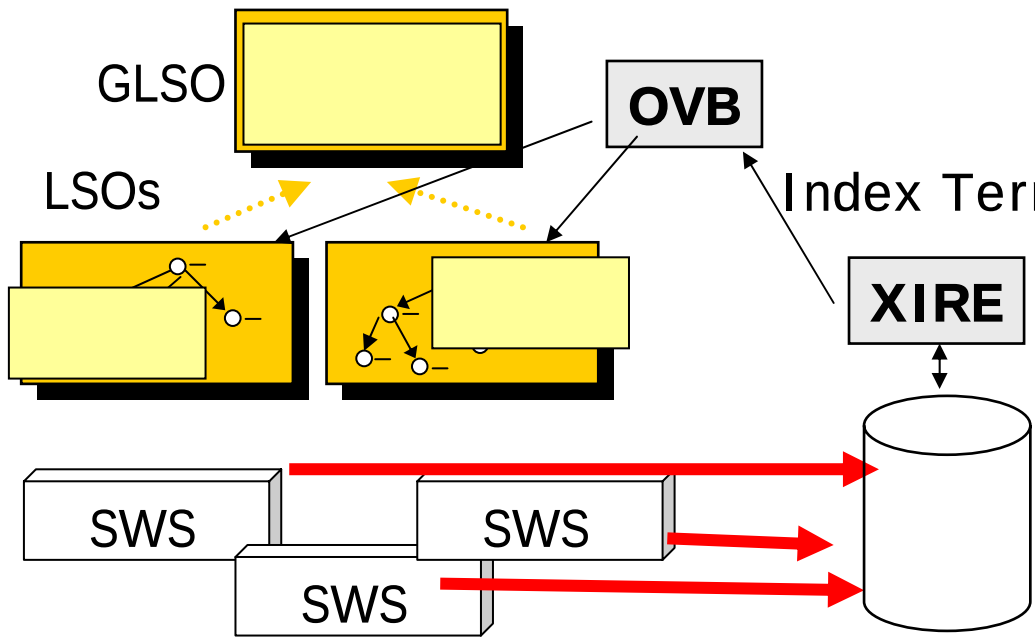
Schema Extraction

- Phase 1
- Each semantic web services description is indexed into the search engine
- The inverted index file is build by considering only a subset of the whole description
  - input/output/pre/postcondition OWL-s
  - Capabilities in WSMO
- The result is a inverted index file and a posting file
  - Set of terms
  - Weighted terms with respect to specific document section



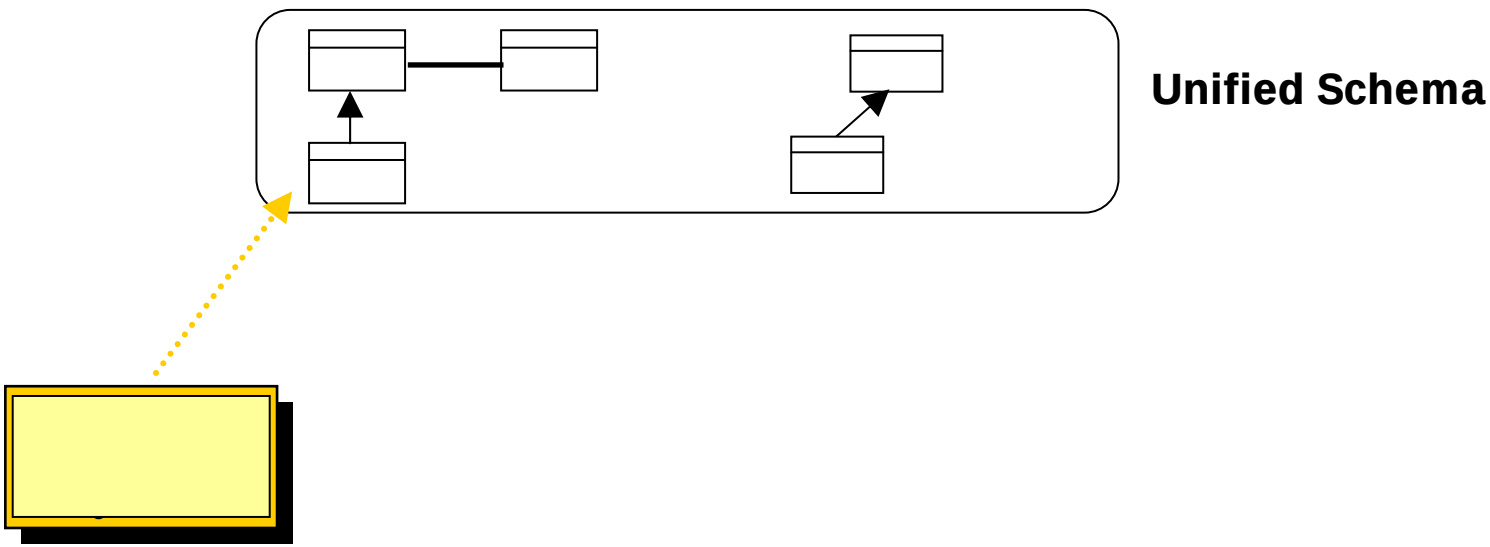
Schema Extraction

- Starting from terms of the inverted index file a Light Service Ontology (LSO) is built by the Ontology View Builder
  - By means of SWS imported ontologies
  - Starting terms are marked as “Index Term”
- The LSO is the integrated with the Global LSO (GLSO)
  - Integration based on automatic techniques [syntactic properties only (e.g. synonymous from WordNet)]

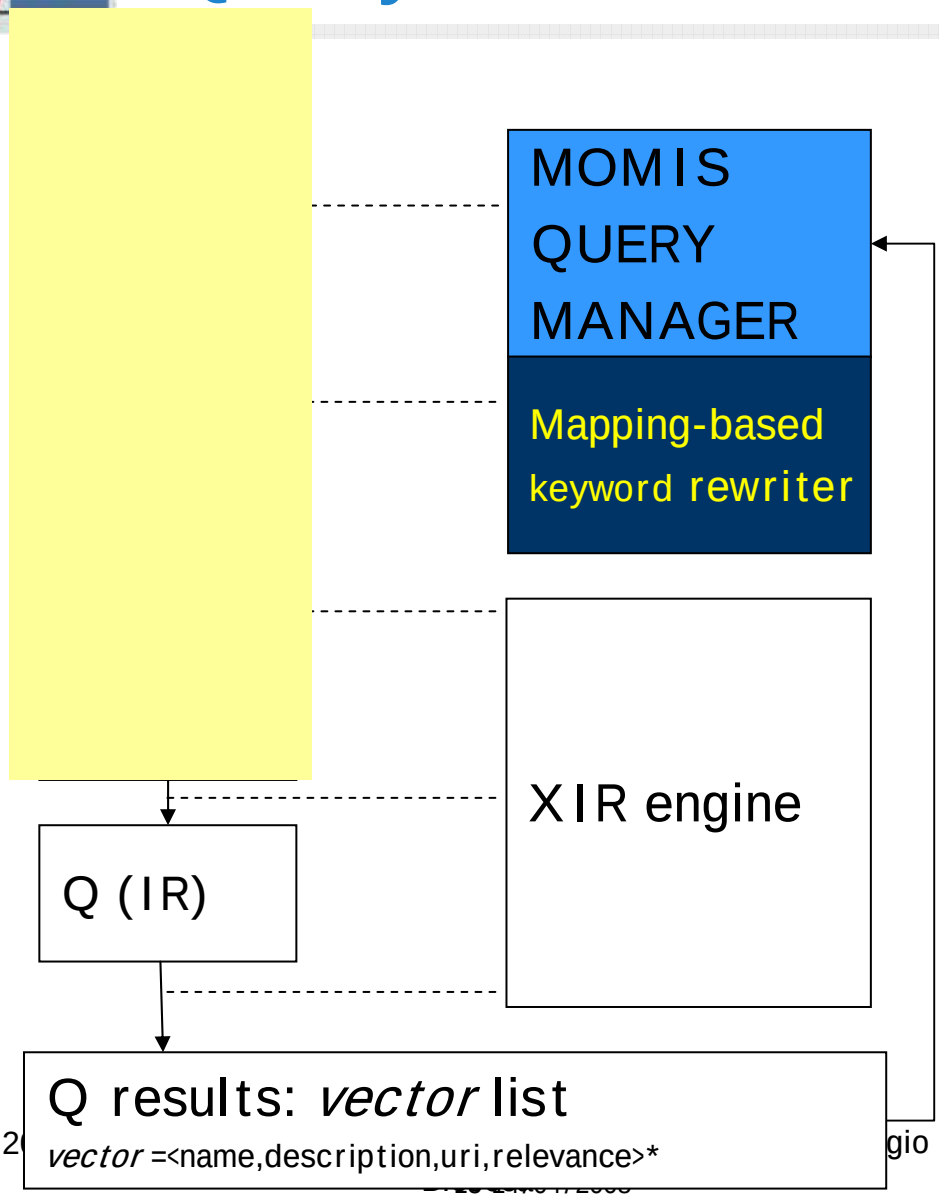


Matching and Mapping

- The GLSO is an ontololgy that can be integrated in any wrapper meditor architecture able to manage ontologies (e.g. MOMIS)



Query execution



An user submits an ODLI3 query

Keyword extraction

Terms extraction phase

Mapping SPDO-GLSO

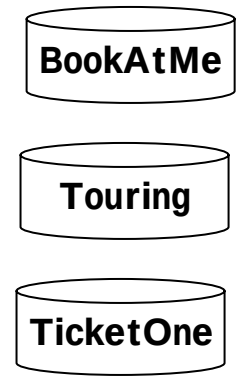
If the search term no match with any concept of GLSO it is discarded in the following

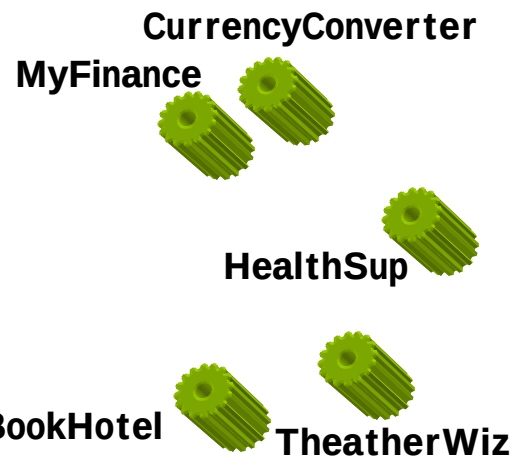
LEGENDA:
Q (O)= ODLi3 query
Q (K)= set of keywords
Q (WK)= set of weighted keywords
Q (IR)= Query input of IR engine

Service and Data integration: An example

Problem Context

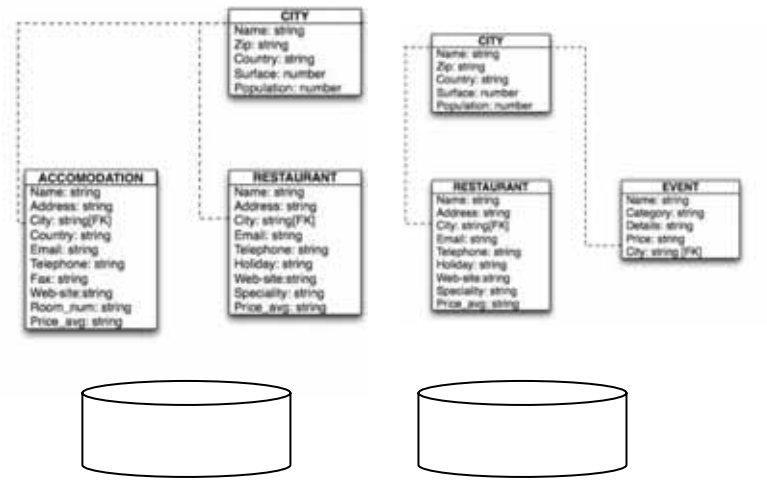
- A semantic peer environment, e.g. virtual touristic district
  - Different data sources
    - **BookAtMe**
      - Internation hotels
    - **Touring**
      - Italian hotels, restaurants and cities
    - **TicketOne**
      - Italian cultural events.
  - Different services (inside and outside the peer)
    - different domains such as **Economy**, **Communication**, **Education**, **Food**, **Medical**, **Travel** and **Weapon**

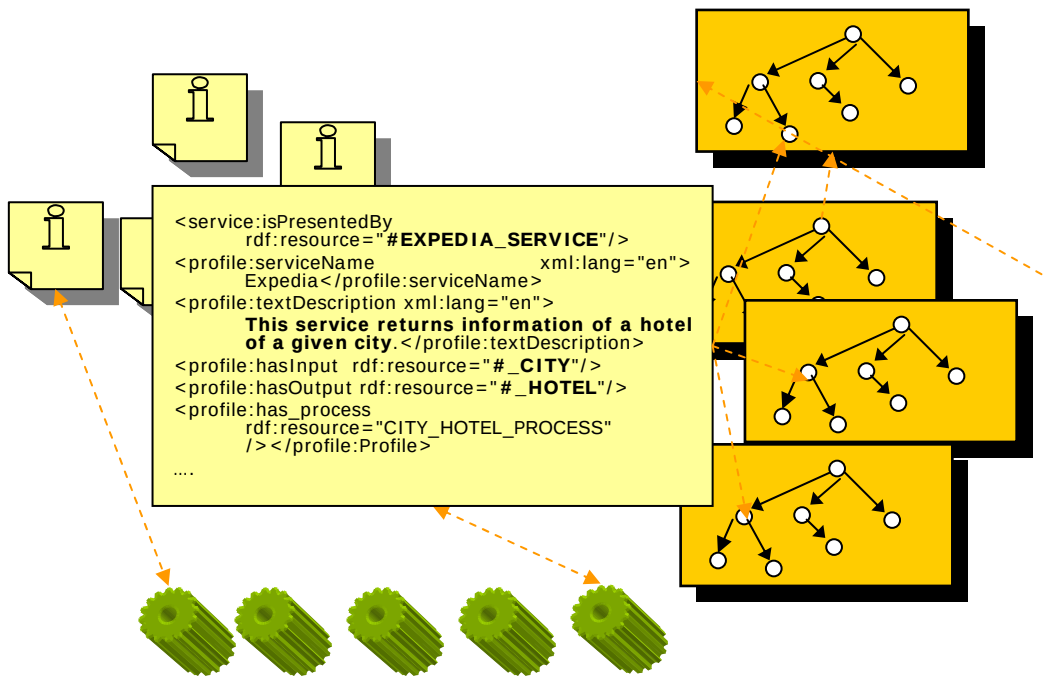




Motivation (2)

```
Select Name, Country
from Accommodation
Where City='Modena'
```





Data and services

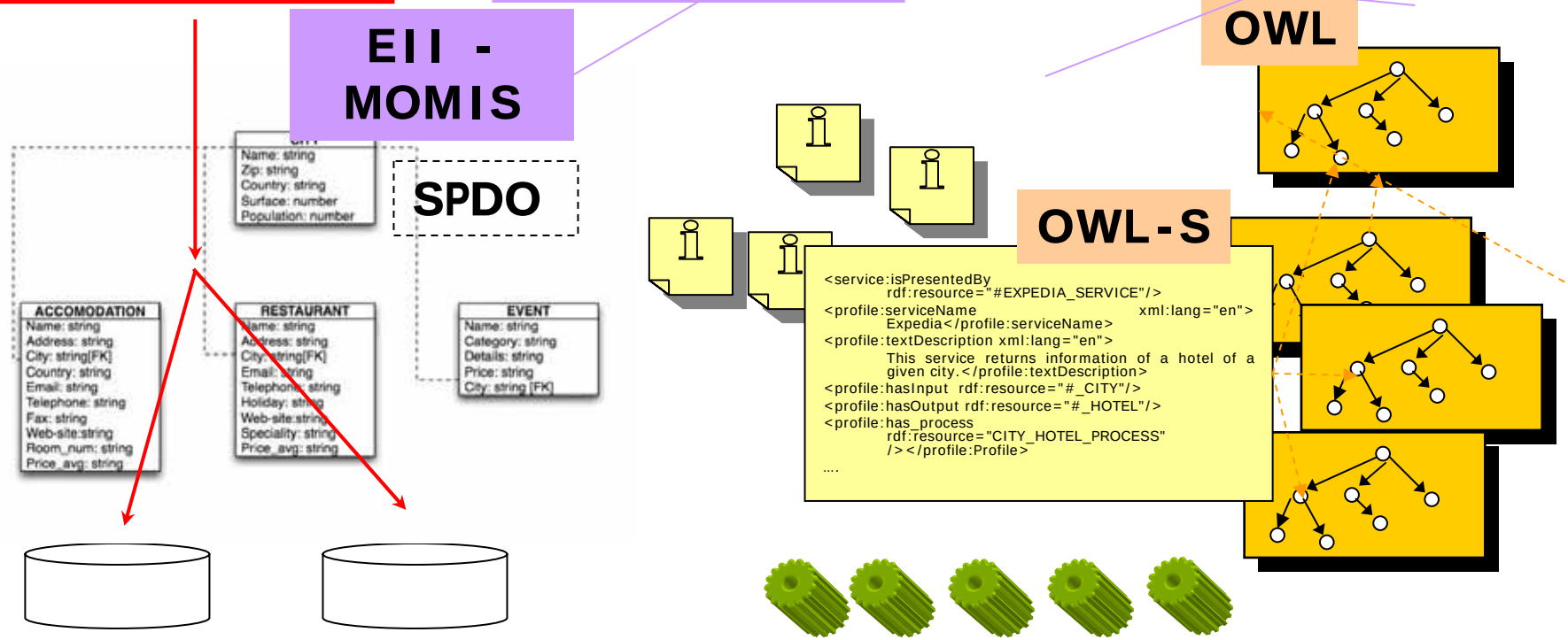
- are described with **different models**
- are retrieved/discovered with **different techniques**
  - different (not always overlapping) communities.
- provide a **complementary vision about the available resources**

Assumptions

```
Select Name, Country
from Accommodation
Where City='Modena'
```

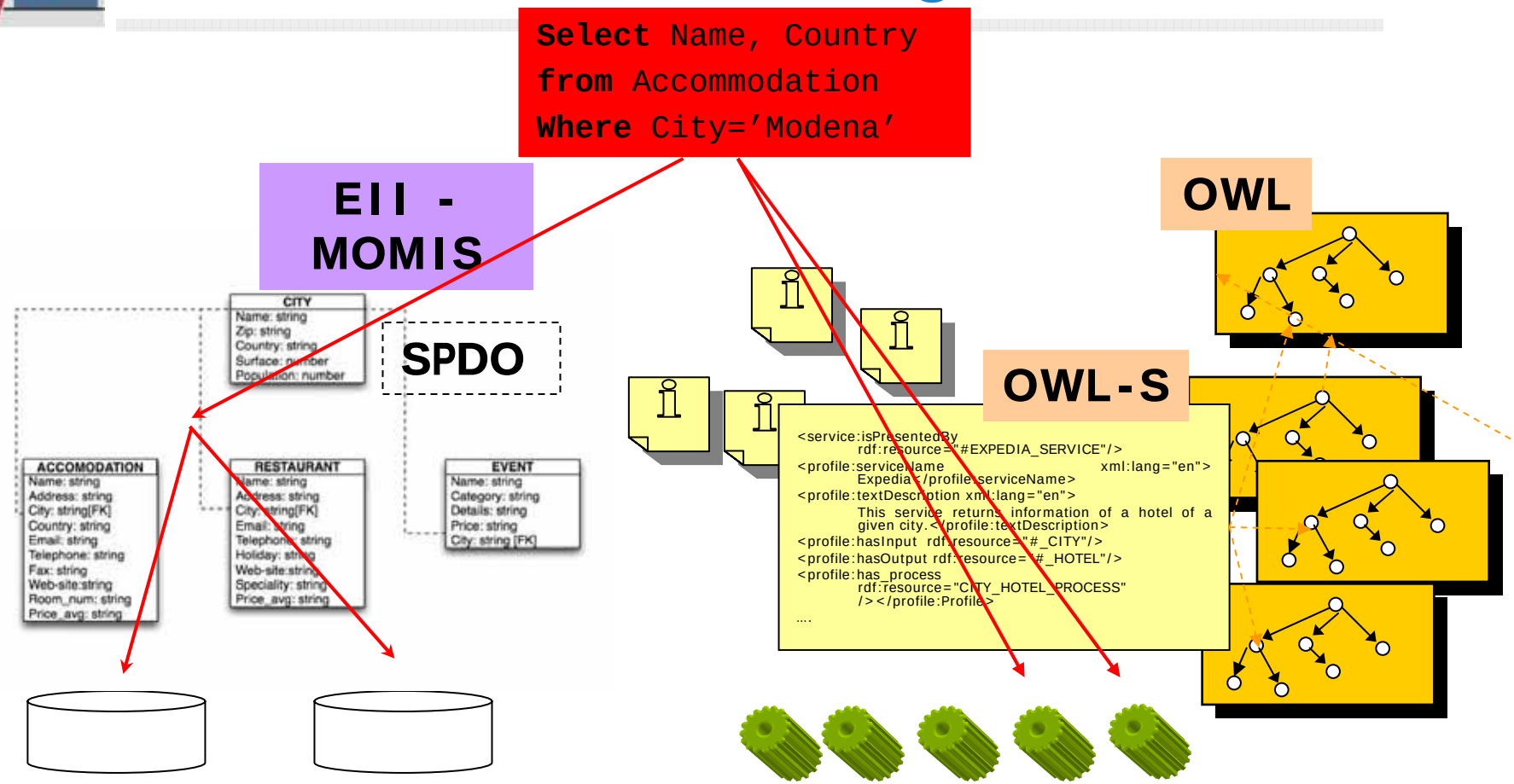
mediator-based data integration system

OWLS-TC 2.0. Benchmark descriptions and ontologies



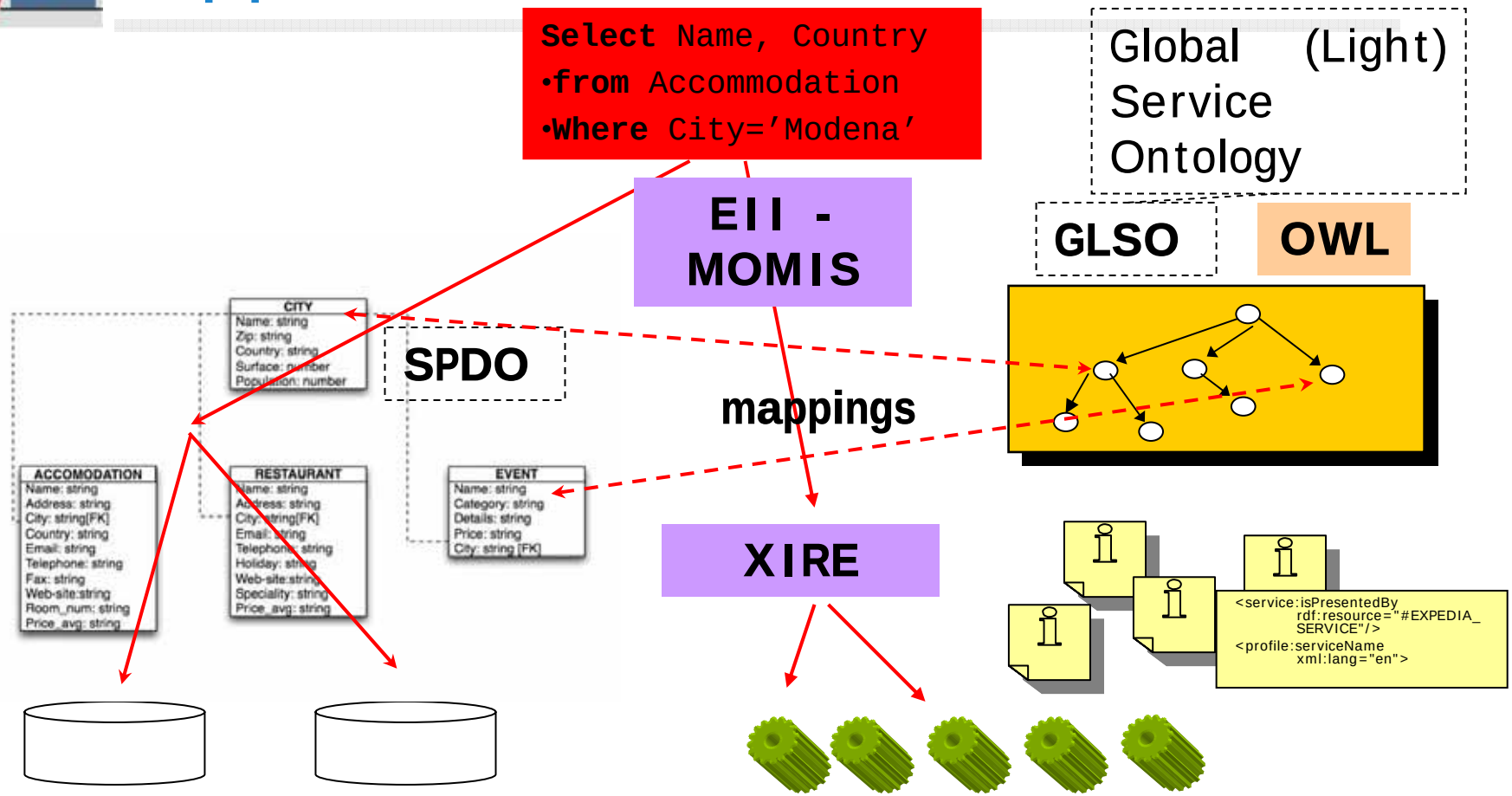
MOMIS provides a global virtual view of data - the **Semantic Peer Data Ontology** (**SPDO**).

Services as Data Target Goal



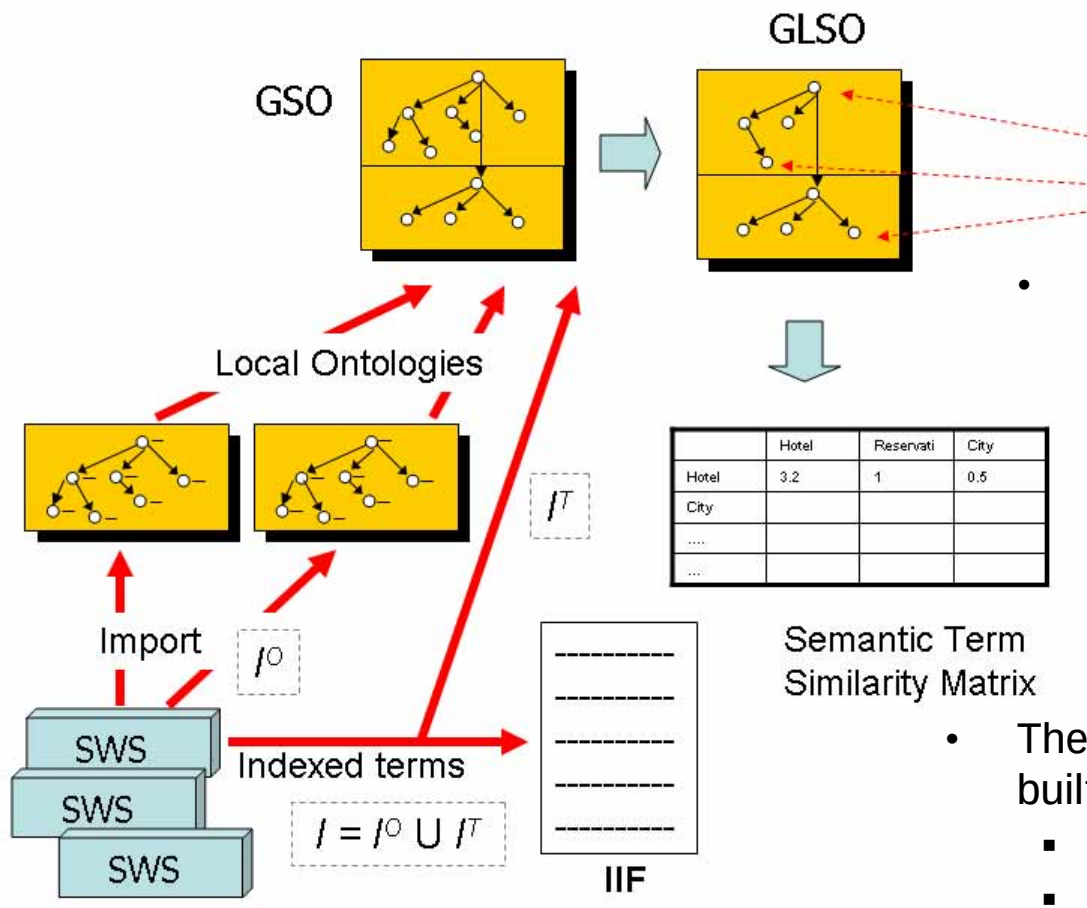
- **Semantic-based data and service integration (services as data):**
  - retrieve, among the many **services** available, the ones that are **related to the query**, according to the semantics of the terms involved in the query.

Approach Overview



- **Design Time**: Service indexing, SPDO, GLSO and mapping construction
- **Query Time**: Query rewriting (from SPDO-based/SQL-like to GLSO keywords and) and IR engine-based service retrieval

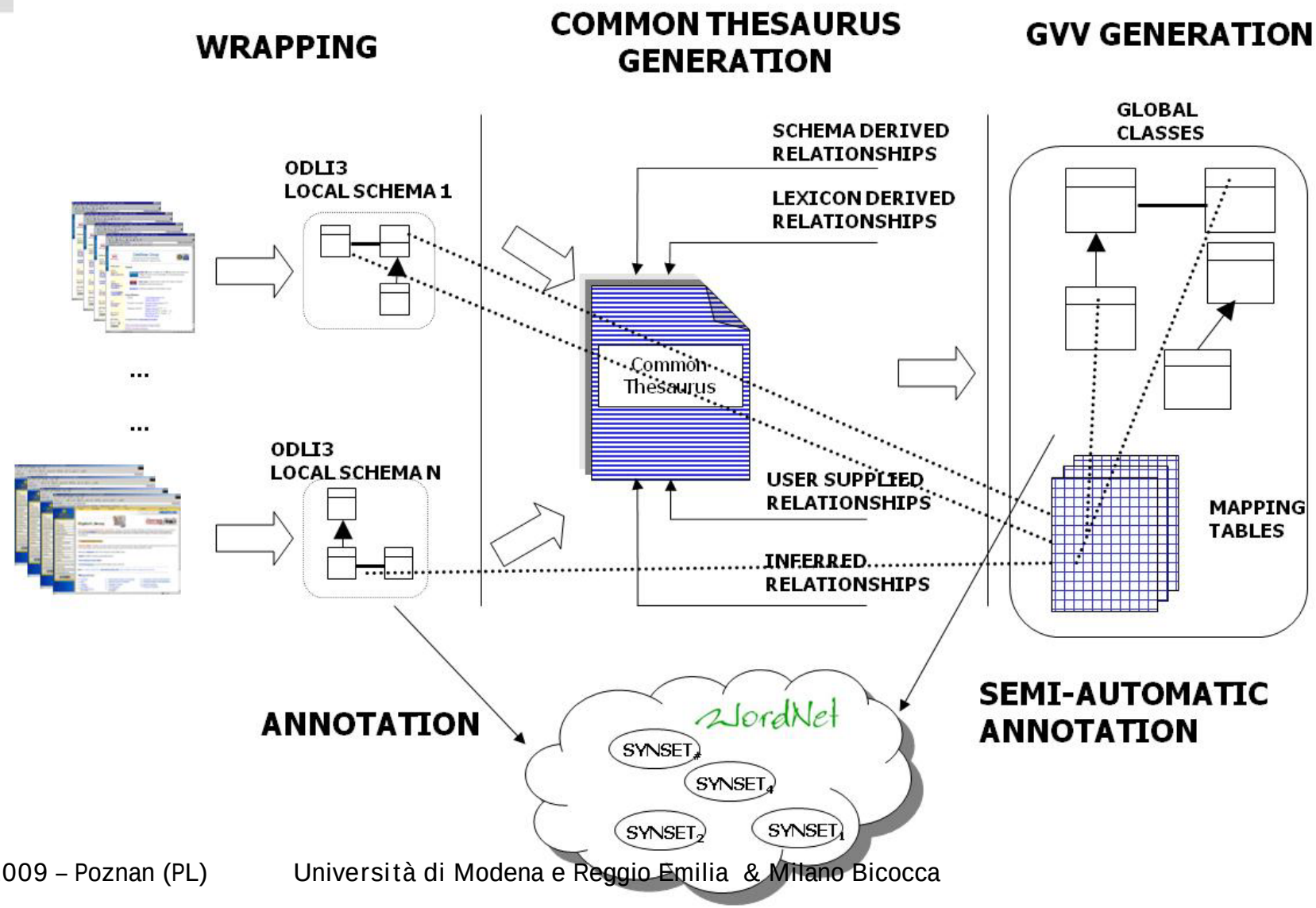
Building the Global Data and Service View



- The SPDO is built by exploiting the MOMIS integration system

- The Global Light Service Ontology is built by means of the following steps:
  - Service indexing
  - Global Service Ontology (GSO) construction
  - Global Light Service Ontology (GLSO) construction and Semantic Similarity Matrix (SSM) definition

MOMIS

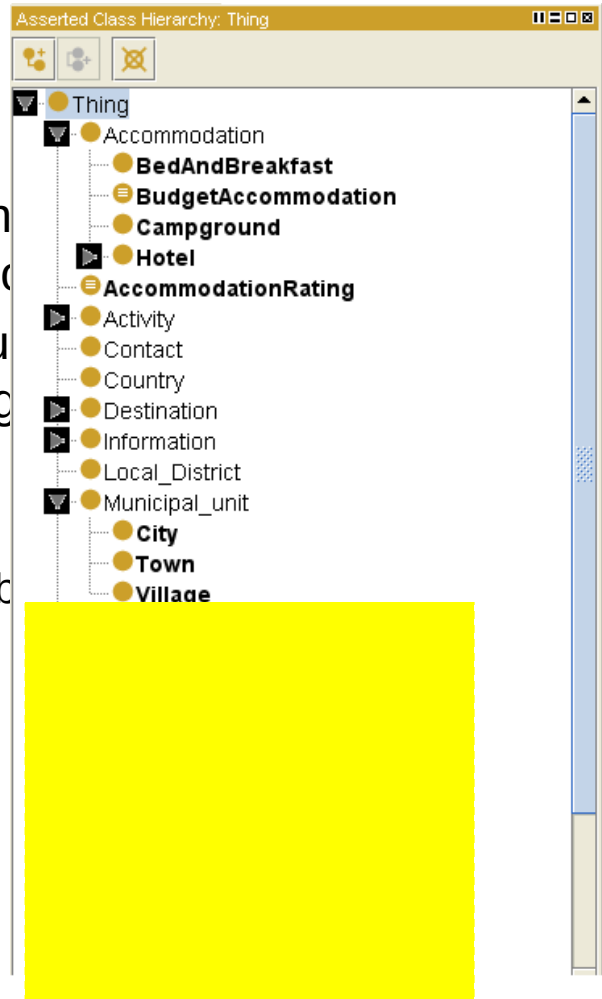

WRAPPING
COMMON THESAURUS GENERATION
GVV GENERATION
ODLI3 LOCAL SCHEMA 1
ODLI3 LOCAL SCHEMA N
SCHEMA DERIVED RELATIONSHIPS
LEXICON DERIVED RELATIONSHIPS
Common Thesaurus
USER SUPPLIED RELATIONSHIPS
INFERRED RELATIONSHIPS
GLOBAL CLASSES
MAPPING TABLES
ANNOTATION
WordNet
SYNSET#
SYNSET4
SYNSET2
SYNSET1
SEMI-AUTOMATIC ANNOTATION

Service Indexing

- Full text indexing from six specific sections of the service descriptions:
  - service name,
  - Service description,
  - input,
  - output,
  - pre-condition
  - post-condition
- A set of index terms I that will be part of the dictionary is extracted.
  - I^O= the set of index terms consisting of **ontology terms**
  - I^T = the set index terms extracted from **textual descriptions**
- The indexing structure is based on a “structured document” approach, where inverted file structure consists of:
  - a **dictionary file** based on I,
  - a **posting file**, with a list of references to the services’ sections where the considered term occurs

GSO construction

- The GSO is built by
  - **Loosely merging** each service ontology wh
    in some service description (plus transitive cl
    - Service Ontologies are merged withou
      similar concepts across the different integ
      - Not optimal choice but:
        - Consistency is preserved
        - The GSO building process need to b
  - Introducing a class ***Terms*** **in the GSO**
    extracted from text



GLSO construction and Semantic Similarity Matrix

- GSO is reduced to the GLSO (Global Light Service Ontology)
  - The GSO may result extremely large in size and only a subset of the terms of the ontologies are relevant to the SWS descriptions.
  - Ontology modularity techniques from **[Jimenez-Ruiz et Al. 2008]**
- The Semantic Similarity Matrix (SSM), which is exploited later on for query expansion at query time, is computed.
  - The SSM is defined by analyzing the GLSO structure, according to some semantic similarity measures developed in literature [Bernstein et al. 2005]

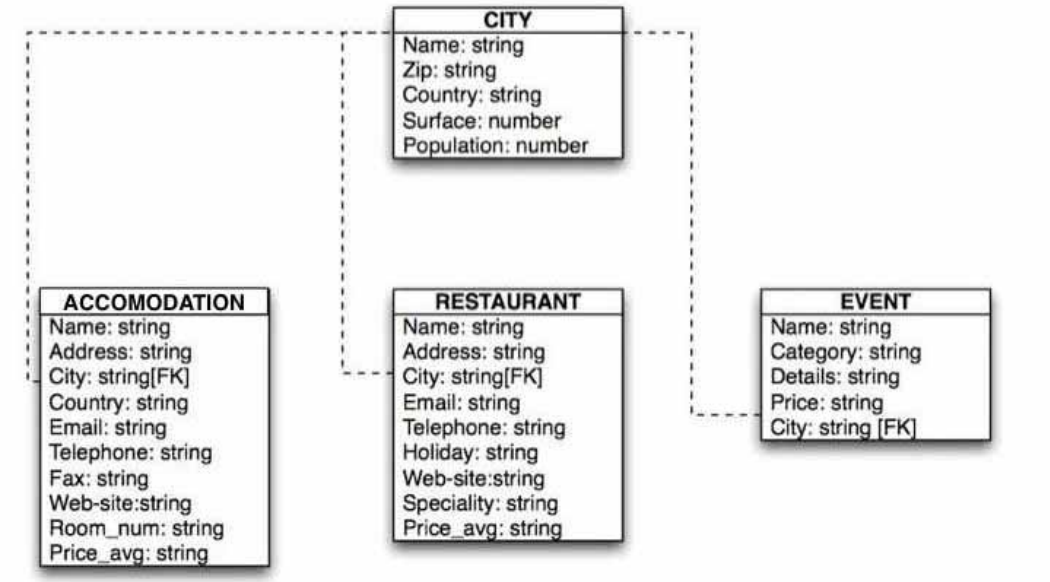
•" http ://127.0.0.1/ ontology / travel . owl # NationalPark ": ! hash
•" http :// www .owl - ontologies . com/ City . owl # Country ": ! double 0.5
•" http :// www .owl - ontologies . com/ City . owl # Village ": ! double 0.5
•" http ://127.0.0.1/ ontology / travel . owl # Accommodation ": ! double
•0.6666666666666666
•" http ://127.0.0.1/ ontology / travel . owl # BedAndBreakfast ": ! double
•0.6666666665631864

	Hotel	Reservati	City
Hotel	3.2	1	0.5
City			
....			
...			

Mapping of Data and Service Ontologies

- The modified clustering algorithm of MOMIS takes as input the SPDO and the GLSO with their associated metadata and generates **a set of clusters** of classes belonging to the SPDO and the GLSO
- Mappings are automatically generated exploiting the clustering result
  - Clusters containing classes belonging to the SPDO and the GLSO are used for the SPDO to GLSO mappings
    - each SPDO class in the cluster is mapped to each GLSO class in the cluster
  - A cluster containing only SPDO classes or only GLSO classes are not exploited for the mapping generation

Mapping Examples



SPDO fragment

```
Hotel
Hotel.Denomination
Hotel.Location
Hotel.Country
```

GLSO fragment

The following mappings are generated with the application of our technology:

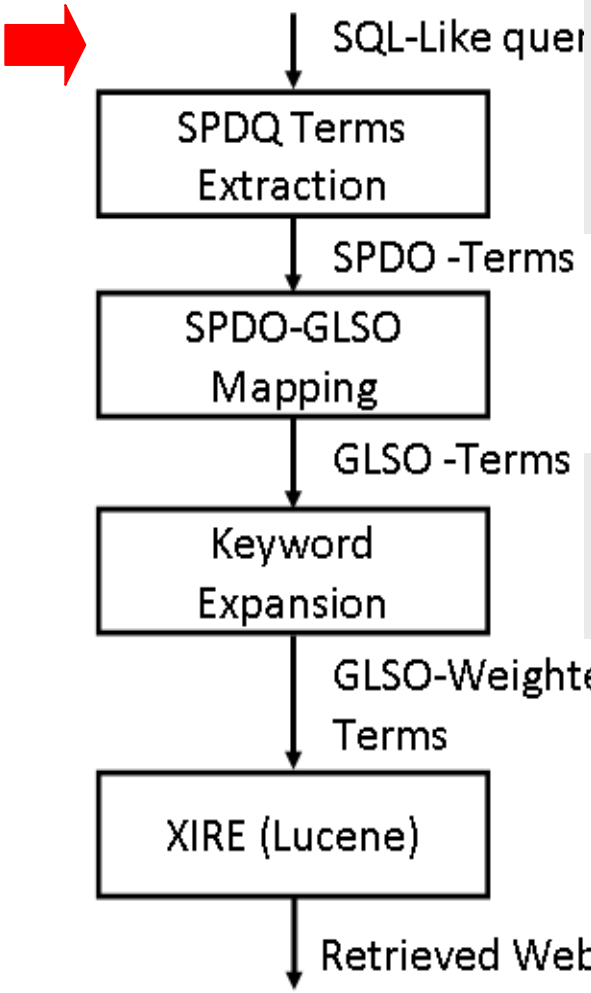
```
Accommodation --> Hotel
Accommodation.Name --> Hotel.Denomination
Accommodation.City --> Hotel.Location
Accommodation.Country --> Hotel.Country
```

Data and Service Retrieval

```
select <select_attribute_list>
from <from_class_list>
where <condition>
```

- The answer to this query is a **data set** from the data sources together with a **set of services** related to the concepts appearing in the query and then to the retrieved data.
- The query processing is divided into two simultaneously executed steps:
  - data set from the data sources is obtained with a **query processing on an integrated view**
    - The results are obtained by exploiting the **MOMIS Query Manager** which rewrites the global query as an equivalent set of queries expressed on the local schemata (local queries), by means of an unfolding process
  - a set of services related to the query is obtained by a **service retrieval process** based on
    - **query rewriting and expansion**.
    - the **XIRE** (eXtended Information Retrieval Engine) component, which is a service search engine based on the vector space.

Services Retrieval: Query Rewriting



Extraction of:

- all the classes in the "FROM" clause,
- all the attributes and the values in the "SELECT" and "WHERE" clauses
- all their ranges defined in the ontology

K^{SPDO}#1 {*Acc.Name, Acc.City, Acc.Country, Accommodation, Modena, City*}

Expansion by:

- similar terms in the Semantic Similarity Matrix up to a given threshold

q#1 <*(Hotel,1), (denomination,1), (location,1), (country,1), (City,1),(Municipal_Unit,05) (Capital_City,0,3), (LuxuryHotel,03),...* >

FUTURE DIRECTION

Future Directions

- Data and service integration is a recent new research area
- We expect more researches in both Data as a Service and Service as a data approaches
  - Who will win?
  - Probably both (it depends on domain)
- In particular
- DaaS
  - Statefull transaction
  - QoS/QoD
- Service as Data
  - Description model
  - Query model
- In both cases the use of semantic technology is mandatory

Future Directions

- The New frontier
- **Sensor as data/ Sensor as a Service [li08]**
- **Toward an integration of data/service/sensor**
  - **Unique description model**
  - **Unique query/composition language**
- **Is it possibile?**

REFERENCES

References

- Papazoglou and Georgakopulos 2003. M.P. Papazoglou and D. Georgakopoulos, Service-Oriented Computing, Communications Of The Acm, October 2003/Vol. 46, No. 10, 25-28
- WSDL20 2007. Chinnici, Moreau, Ryman, Weerawarana (eds). Web Services Description Language (WSDL) Version 2.0 Part 1: Core Language. W3C http://www.w3.org/TR/wsdl20/ (last visited 8/1/2009)
- UDDI20 2002. Bellwood (Eds). UDDI Version 2.04 API Specification. http://uddi.org/pubs/ProgrammersAPI-V2.04-Published-20020719.htm (last visited 8/1/2009)
- WS-BPEL20 2007. Jordan, Edvemon (EDS). Web Services Business Process Execution Language Version 2.0 http://docs.oasis-open.org/wsbpel/2.0/OS/wsbpel-v2.0-OS.html (last visited 8/1/2009)

- [MSH01]. McIlraith, S.A.; Son, T.C.; Honglei Zeng Semantic Web services IEEE Intelligent Systems 2001 16,(2), 46- 53
- [HM00]. J. Hendler and D. McGuinness, “The DARPA Agent Markup Language,” *IEEE Intelligent Systems, vol. 15, no. 6, Nov./Dec. 2000,* pp. 72–73.
- [AB+01] A. Ankolekar, M. Burstein, J. Hobbs, O. Lassila, D. Martin, S. McIlraith, S. Narayanan, M. Paolucci, T. Payne, K. Sycara, H. Zeng), "DAML-S: Semantic Markup for Web Services", in *Proceedings of the International Semantic Web Working Symposium (SWWS)*, July 30-August 1, 2001.
- [MP+04]. David Martin, Massimo Paolucci, Sheila McIlraith, Mark Burstein, Drew McDermott, Deborah McGuinness, Bijan Parsia, Terry Payne, Marta Sabou, Monika Solanki, Naveen Srinivasan, Katia Sycara, "Bringing Semantics to Web Services: The OWL-S Approach",*Proceedings of the First International Workshop on Semantic Web Services and Web Process Composition (SWSWPC 2004)*, July 6-9, 2004, San Diego, California, USA.

- [RK+05]. Dumitru Roman, Uwe Keller, Holger Lausen, Jos de Bruijn, Rubén Lara, Michael Stollberg, Axel Polleres, Cristina Feier, Christoph Bussler, and Dieter Fensel: Web Service Modeling Ontology, *Applied Ontology*, 1(1): 77 - 106, 2005.
- [PO+04] Abhijit Patil, Swapna Oundhakar, Amit Sheth, Kunal Verma, METEOR-S Web service Annotation Framework, Proceeding of the World Wide Web Conference, July 2004
- [SAWSDL07]. Farell, Lausen (eds). Semantic Annotations for WSDL and XML Schema. W3C Recommendation 28 August 2007. http://www.w3.org/TR/sawsdl/ (last visited 8/1/2009)
- *[PW+07]M. Paolucci, M. Wagner, and D. Martin* Grounding OWL-S in SAWSDL. ICSOC 2007

- [Vk+08] T. Vitvar , J. Kopecky , J. Viskova , D. Fensel: *WSMO-Lite Annotations for Web Services ,* The 5th European Semantic Web Conference 2008 , Springer, June, 2008, Tenerife, Spain.
- [VK+07] T. Vitvar , J. Kopecky , M. Zaremba , D. Fensel: *WSMO-Lite: Lightweight Descriptions of Services on the Web ,* The 5th IEEE European Conference on Web Services (ECOWS) , IEEE CS Press, November, 2007, Halle, Germany.
- [KV+07] J. Kopecky , T. Vitvar , C. Bournez , J. Farrell: *SAWSDL: Semantic Annotations for WSDL and XML Schema,* IEEE Internet Computing, IEEE CS Press, 2007, 11 (6).
- [CD+08] M. Comerio, F. De Paoli, A. Maurino, M. Palmonari, A Meta-model for Non-functional Property Descriptions of Web Services, International Conference on Web Services 2008

- [VA09] Le-Hung Vu and Karl Aberer. Towards Probabilistic Estimation of Quality of Online Services. IEEE International Conference on Web Services 2009
- [LC+09] Pei Li, Marco Comerio, Andrea Maurino, Flavio De Paoli An Approach to Non-functional Property Evaluation of Web Services. IEEE International Conference on Web Services 2009
- [PP09] Plebani, Pernici. URBE: Web Service Retrieval Based on Similarity Evaluation. IEEE Transaction on Knowledge and Data Engineering, jannuary 2009
- [SP07] Marta Sabou, Jeff Pan: Towards semantically enhanced Web service repositories. J. Web Sem. 5(2): 142-150 (2007)

- [AM08] Eyhab Al-Masri and Qusay H. Mahmoud. Investigating Web Services on the World Wide Web. In WWW conference 2008
- [PT09] Thomi Pilioura, Aphrodite Tsalgatidou: Unified publication and discovery of semantic Web services. TWEB 3(3): (2009)
- [PT06] Michael Pantazoglou, Aphrodite Tsalgatidou, George Athanasopoulos: Discovering Web Services and JXTA Peer-to-Peer Services in a Unified Manner. ICSOC 2006: 104-115
- [AK+06] Andronikos Nedos, Kulpreet Singh, Siobhán Clarke: Mobile Ad Hoc Services: Semantic Service Discovery in Mobile Ad Hoc Networks. ICSOC 2006: 90-103

- [IS09] Islam, N. Shaikh, Z.A. A Novel Approach to Service Discovery in Mobile Adhoc Network. IEEE Networking and Communications Conference, 2008. INCC 2008.
- [AB07] Ayorak, E. Bener, A.B. Super Peer Web Service Discovery Architecture ICDE 2007
- [LY08] Li M. Yu B. Rana O., Wang Z. Grid service discovery with rough sets. IEEE Transactions on Knowledge and Data Engineering. 20(6) 851-862 2008
- [DG09] *Stefan Dietze, Alessio Gugliotta, John Domingue.* Exploiting Metrics for Similarity-based Semantic Web Service Discovery. ICWS 2009
- [BK07] Umesh Bellur, Roshan Kulkarni Improved Matchmaking Algorithm for Semantic Web Services Based on Bipartite Graph Matching. ICWS2007

- [SS08] Dimitrios Skoutas, Dimitris Sacharidis, Verena Kantere, Timos K. Sellis: Efficient Semantic Web Service Discovery in Centralized and P2P Environments. International Semantic Web Conference 2008: 583-598
- [ZL+09] *Yiming Zhang, Ling Liu, Dongsheng Li, Feng Liu, Xicheng Lu* DHT-Based Range Query Processing for Web Service Discovery. ICWS2009
- [SB07] Mohamed Shehab, Kamal Bhattacharya, Arif Ghafoor Web services discovery in secure collaboration environments, 1, (5),2007
- [BC+09] Antonio Brogi, Sara Corfini, Razvan Popescu. Semantics-based composition-oriented discovery of Web services TOIT 8, (4), 2008

- [KK+09] Matthias Klusch, Patrick Kapahnke, Ingo Zinnikus. SAWSDL-MX2: A Machine-learning Approach for Integrating Semantic Web Service Matchmaking Variants. ICWS 2009
- [KB+09] Matthias Klusch, Benedikt Fries and Katia Sycara. OWLS-MX: A hybrid Semantic Web service matchmaker for OWL-S services. Journal of Web Semantics: Science, Services and Agents on the World Wide Web, 7, 2, 2009, Pages 121-133
- [DC+08]. John Domingue, Liliana Cabral, Stefania Galizia, Vlad Tanasescu, Alessio Gugliotta, Barry Norton, Carlos Pedrinaci. IRS-III: A broker-based approach to semantic Web services. Journal of Web Semantics: Science, Services and Agents on the World Wide Web,, 6,2, 2008, *109-132*
- [TR+07] I. Toma, D. Roman, D. Fensel, B. Sapkota, and J.M. Gomez. A Multi-criteria Service Ranking Approach Based on Non-Functional Properties Rules Evaluation. ICSOC 2007

- [DH+08] Dong, H., Hussain, F. K., and Chang, E. 2008. Quality of service (QoS) based service retrieval engine. In *Proceedings of the 6th international Conference on Advances in Mobile Computing and Multimedia* (Linz, Austria, November 24 - 26, 2008). G. Kotsis, D. Taniar, E. Pardede, and I. Khalil, Eds. MoMM '08. ACM, New York, NY, 405-408
- [CC+08] Alessio Carenini, Dario Cerizza, Marco Comerio, Emanuele Della Valle, Flavio De Paoli, Andrea Maurino, Matteo Palmonari, and Andrea Turati. GLUE2: A Web Service Discovery Engine with Non-Functional Properties. ECOWS 2008
- [CD+07] M Comerio, F De Paoli, A Maurino, M Palmonari. Nfp-aware semantic web services selection. EDOC 2007

- [ZCL08] C. Zhou, L. Chia, B. Lee. QoS-Aware Web Services Discovery with Federated Support for UDDI. In Quantitative Quality of Services for grid computing (Wang, Chen, Jie eds) 2008 IGI Global
- [AR+08] Mohammad Alrifai, Thomas Risse, Peter Dolog, Wolfgang Nejdl A Scalable Approach for QoS-Based Web Service Selection. In ICSOC Workshops pp. 190-199, 2008.
- [MW+08] Qian Ma, Hao Wang, Ying Li, Guotong Xie, Feng Liu, "A Semantic QoS-Aware Discovery Framework for Web Services," icws, pp.129-136, 2008 IEEE International Conference on Web Services, 2008
- [BR+07] *F. Baligand, N. Rivierre, and T. Ledoux.* A Declarative Approach for QoS-Aware Web Service Compositions. ICSOC 2007
- [MC+08] L. Mei, W. Chan, T. Tse. An Adaptive Service Selection Approach to Service Composition. ICWS 2008

- [FK+08] Dieter Fensel, Mick Kerrigan, Michal Zaremba (eds.) Implementing semantic web services: the SESA framework. Springer 2008
- [YR08] Xinfeng Ye, Rami Mounla. A Hybrid Approach to QoS-aware Service Composition ICWS 2008
- [AR08] M. Alrifai,T. Risse. Combining global optimization with local selection for efficient QoS-aware service composition. WWW conference 2008
- [MG+07] Michael Mrissa, Chirine Ghedira, Djamal Benslimane, Zakaria Maamar, Florian Rosenberg, Schahram Dustdar. A context-based mediation approach to compose semantic Web services TOIT, 8,1, 2007
- [RL+08] Kaijun Ren Xiao Liu Jinjun Chen Nong Xiao Junqiang Song Weimin Zhang. A QSQL-based Efficient Planning Algorithm for Fully-automated Service Composition in Dynamic Service Environments. SCC 2008

- [AK+09] Anna Averbakh, Daniel Krause,Dimitrios Skoutas Exploiting User Feedback to Improve Semantic Web Service Discovery. ICWS 2009
- [Lècuè09] Freddy Lécué, Optimizing QoS-Aware Semantic Web Service Composition. ICWS 2009
- [AP07] Ardagna, D.; Pernici, B.; Adaptive Service Composition in Flexible Processes IEEE Transactions on Software Engineering, Volume 33, Issue 6, June 2007 Page(s):369 - 384
- [NP+07] N.C. Narendra, K. Ponnalagu, J. Krishnamurthy, and R. Ramkumar Run-Time Adaptation of Non-functional Properties of Composite Web Services Using Aspect-Oriented Programming. ICSOC 2007
- [ZS+08] Yanlong Zhai Hongyi Su Shouyi Zhan. A Reflective Framework to Improve the Adaptability of BPEL-based Web Service Composition. SCC 08

- [ST07] Markus Stumptner and Rajesh Thiagarajan Service Composition With Consistency-based Matchmaking: A CSP-based Approach ECOWS 2007
- [SL+09] Dong-Hoon Shina, , Kyong-Ho Leea, and Tatsuya Sudab. Automated generation of composite web services based on functional semantics. Journal on Web Semantics: Science, Services and Agents on the World Wide Web in press
- [BK+09] Piergiorgio Bertoli, Raman Kazhamiakin, Massimo Paolucci, Marco Pistore, Heorhi Raik, Matthias Wagner. Control Flow Requirements for Automated Service Composition. ICWS 2009
- [MK+09] Saayan Mitra, Ratnesh Kumar, Samik Basu A Framework for Optimal Decentralized Service-Choreography. ICWS 09
- [RK+08] Dumitru Roman and Michael Kifer: Semantic Web Service Choreography: Contracting and Enactment. ICSW 2008

- [Hadley09] M. Hadley, Web Application Description Language (WADL), 2009, http://wadl.dev.java.net/wadl20090202.pdf.
- [SK+08] Amit P. Sheth, Karthik Gomadam, and Ajith Ranabahu. Semantics enhanced Services: METEOR-S, SAWSDL and SA-REST. Data engineering 31(3), 8-13, 2008
- [SG+07] A. P. Sheth, K. Gomadam, and J. Lathem. Sa-rest: Semantically interoperable and easier-to-use services and mashups. *IEEE Internet Computing, 11(6):91–94, 2007.*
- [KG+08]Jacek Kopecký, Karthik Gomadam, Tomas Vitvar, "hRESTS: An HTML Microformat for Describing RESTful Web Services," wi-iat, vol. 1, pp.619-625, 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2008

- [IP+09] Muhammad Intizar Ali, Reinhard Pichler, Hong-Linh Truong, and Schahram Dustdar. DeXIN: An Extensible Framework for Distributed XQuery over Heterogeneous Data Sources. Proceedings of 11th International Conference on Enterprise Information Systems, ICEIS 2009, Milan, Italy, May 6-10, 2009.
- [Fielding2000] Fielding, Roy Thomas, Architectural Styles and the Design of Network-based Software Architectures, Doctoral dissertation, University of California, Irvine, 2000
- [RR07] Leonard Richardson, Sam Ruby, RESTful Web Services, O'Reilly, May 2007
- [ZL+07] *L. Zeng, H. Lei, and H. Changl,* Monitoring the QoS for Web Services, ICSOC 2007

- [MJ+08] Mukherjee, D., Jalote, P., and Gowri Nanda, M. 2008. Determining QoS of WS-BPEL Compositions. In *Proceedings of the 6th international Conference on Service-Oriented Computing* (Sydney, Australia, December 01 - 05, 2008).
- [PM+08] Podorozhny, Rodion; Carlson, Michael; Ngu, Anne; Zeng, Liang-Zhao. Automatic service mash up for composite applications icsoc 2008
- [MW+07] A Domain-Specific Language for Web APIs and Services Mashups *E.M. Maximilien, H. Wilkinson,N. Desai,and S. Tai ICSOC 2007*
- [AR+09] Osama Al-Haj Hassan, Lakshmish Ramaswamy and John A. Miller "MACE: A Dynamic Caching Framework for Mashups", IEEE 7th International Conference on Web Services (ICWS-2009), July 2009

- [WL+09] Eric Wohlstadter, Peng Li, Brett Cannon, "Web Service Mashup Middleware with Partitioning of XML Pipelines," icws, pp.91-98, 2009 IEEE International Conference on Web Services, 2009
- [LI08] L. li A Framework for Semantic Sensor Network Services. ICSOC 2008
- [BHL01] Berners-Lee, Tim; James Hendler and Ora Lassila (May 17, 2001). "The Semantic Web". *Scientific American Magazine*.