Toward a Unified View of Data and Services

Carlo Batini
Matteo Palmonari
Andrea Maurino
University of Milan-Bicocca
Italy

Sonia Bergamaschi
Francesco Guerra
Domenico Beneventano
Antonio Sala
DBGroup, University of Modena and Reggio Emilia
Italy

Emanuele Della Valle
Dario Cerizza
Andrea Turati
CEFRIEL
Italy

Fausto Rabitti
Claudio Gennaro
ISTI – CNR Pisa
Italy

DB Group @ unimo

Scenario and goals

- Scenario: networks of (semantic) peers sharing knowledge and inter-operating through service-based interactions
  - NEP4B project: Networked Peers for Business (3-year basic research project)
  - This presentation is focused on the Semantic Peer, in reality the goal of the NeP4B project is to create a network of Semantic Peers
- Each semantic peer
  - provides a unified access to different **data** sources referring to the same domain.
  - provides a number of related **services**
- Two complementary aspects
  - DATA management: data integration
  - SERVICE management: Web Services (WS), Semantic WS Framework, ...
- Data and Services are usually represented with different models and queried by different tools.

Our approach aims at providing users data and services of a domain.

Where we start from...

- Data management:
  - A data-integration system (MOMIS): unified view of data
    - *ODLi3 l*anguage
- Service management:
  - Semantic WS approach: WSMO framework
    - semantic service descriptions
    - discovery engine [WSMO compliant]
    - *WSML-Flight* language
- Integration?
  - TASK#1: to find services related to queries on data
    - E.g. When a user wants to have a funny night attending an event in his/her town, first he/she searches for the events occuring in that night and, then, after selecting one of them, he/she may invoke a service for buying the ticket.

Data vs. Services: differences

- Data:
  - Query answering through schemata/ontologies
  - Data behind the concepts via mappings (e.g. all the instances of “Hotel” in some databases)
  - Multimedia data sources
- Semantic Web Services:
  - Discovery (complex mechanism)
  - Ontologies are exploited to describe services
- Different conceptual languages with different expressivity
  - *ODLi3* (OWL) vs. *WSML-Flight*

Service Descriptions vs. Service Ontologies

```
...
webservice _
   "http://.../hotel_Miramare_reservation"
...
   capability BookRooms
      precondition
         // checkIn and checkOut dates must
         // be valid dates and time
         // checkIn must be before checkOut
         definedBy
         ...
      assumption
         // credit card should be valid
         definedBy
         ...
      postcondition
         definedBy
               BookingRoomsAxiom
      effect
         // the total price will be deducted
         // from the credit card
         definedBy
         ...
```

```
room[
      hasType=>roomTipology,
      hasMaxNumberOfPersons=>integer,
      hasPricePerNight=>price,
      hasFacility=>>roomFacility
].

hotel[
      hasCategory=>hotelCategory,
      locatedIn=>location,
      hasFacilities=>>hotelFacility,
      checkIn=>timeInterval,
      checkOut=>timeInterval,
      hasRooms=>>room,
      acceptPaymentMethods=>>paymentMethod
].
```

```
bookingRooms[
   rooms->> {
         r1:room[hasType->singleRoom,
               hasMaxNumberOfPersons->1,
               hasPricePerNight->65.00,
               hasFacility->>{television,
                     airconditioning}],
         r2:room[hasType->twinRoom,
               hasMaxNumberOfPersons->2,
               hasPricePerNight->110.00,
               hasFacility->>{television,
                     airconditioning}],
   },
   checkIn->_#:timeInterval[
         start->_#:time[
               hourOfDay->11,
               minuteOfHour->30,
               secondOfMinute->0],
         end->_#:time[
               hourOfDay->20,
               minuteOfHour->0,
               secondOfMinute->0],
   checkOut->_#:timeInterval[
         start->_#:time[
               hourOfDay->8,
               minuteOfHour->0,
               secondOfMinute->0],
         end->_#:time[
               hourOfDay->10,
               minuteOfHour->0,
               secondOfMinute->0],
         ].
```

The Peer Virtual View (PVV)

A global **Peer Virtual View (PVV)** providing the connections between the two worlds is built with:

- a **Semantic Peer Data Ontology (SPDO)** of the data
  - i.e. a common representation of all the data sources belonging to the peer, expressed with the ODLI3 language.
- a set of **Light Service Ontologies (LSOs)**
  - i.e. ODLi3 ontologies whose elements have a number of relevant services associated, e.g. the service “miramare.booking” is associated to the concept “Hotel” with a relevance degree of 0.9 .
- a set of **SPDO-LSOs mappings**
  - which connects data representations to service descriptions.
    - To be used by the query transformation engine to translate a query for data into a query for retrieving services.

Building the Data Ontology: MOMIS

MOMIS* (Mediator envirOnment for Multiple Information Sources) is a framework to perform information extraction and integration of heterogeneous, structured and semistructured, data sources

Semantic Integration of Information

- A common data model ODLI3 (derived from ODL-ODMG and I3) & mapped into OLCD description logics

Tool-supported techniques to construct the Global Virtual View (GVV)

- Local sources wrapping
- Local Schema Annotation w.r.t. a common lexical ontology (WordNet)
- Semi-automatic discovery of relationships between local schemata
- Clustering techniques to build the GVV & mappings between the GVV and local schemata (Mapping Table)
- automatic GVV Annotation w.r.t. a common lexical ontology & OWL exportation

Global Query Management

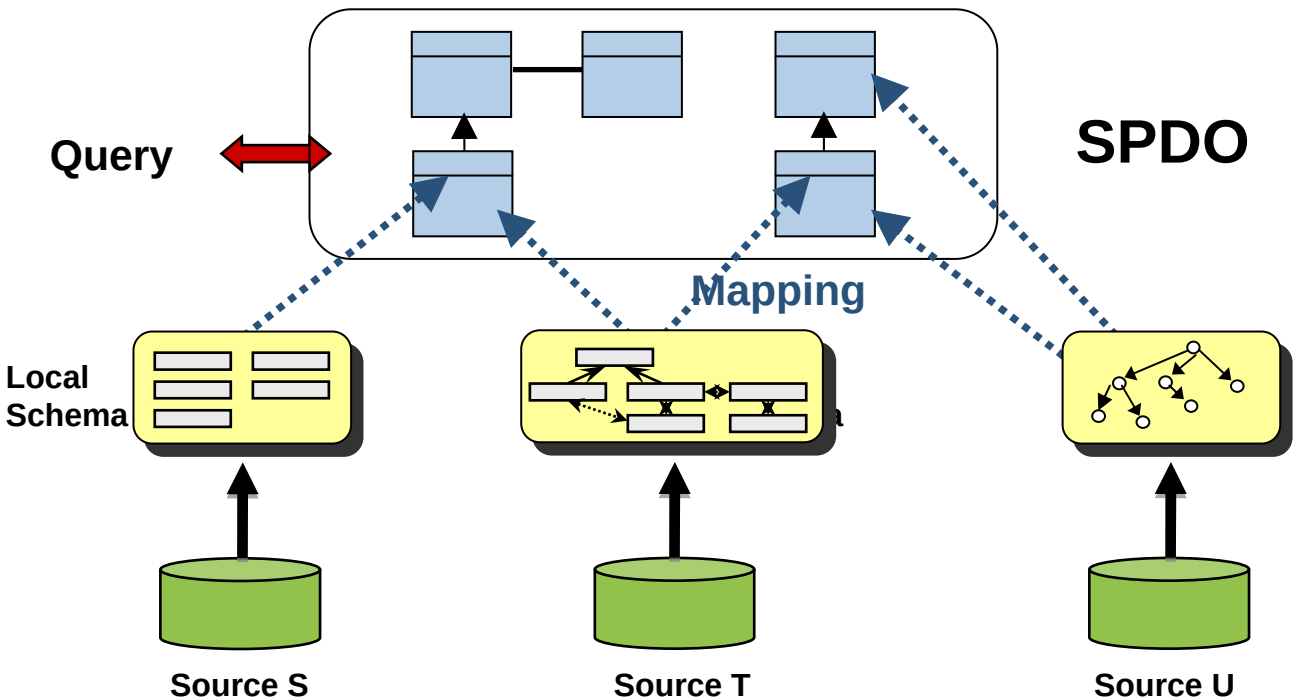
- Including services and multimedia data sources

* D. Beneventano, S. Bergamaschi, F. Guerra, M. Vincini: "Synthesizing an Integrated Ontology ", IEEE Internet Computing Magazine, September-October 2003,42-51.

S. Bergamaschi, S. Castano, M. Vincini "Semantic Integration of Semistructured and Structured Data Sources", SIGMOD Record Special Issue on Semantic Interoperability in Global Information, Vol. 28, No. 1, March 1999.

Querying Data Ontology (SPDO) with MOMIS

To answer a query expressed on the SPDO (global query) we exploit the MOMIS **Query Manager** which rewrites the global query as an equivalent set of queries expressed on the local schemata (local queries); this query translation is carried out by considering the mappings between the SPDO and the local schemata.



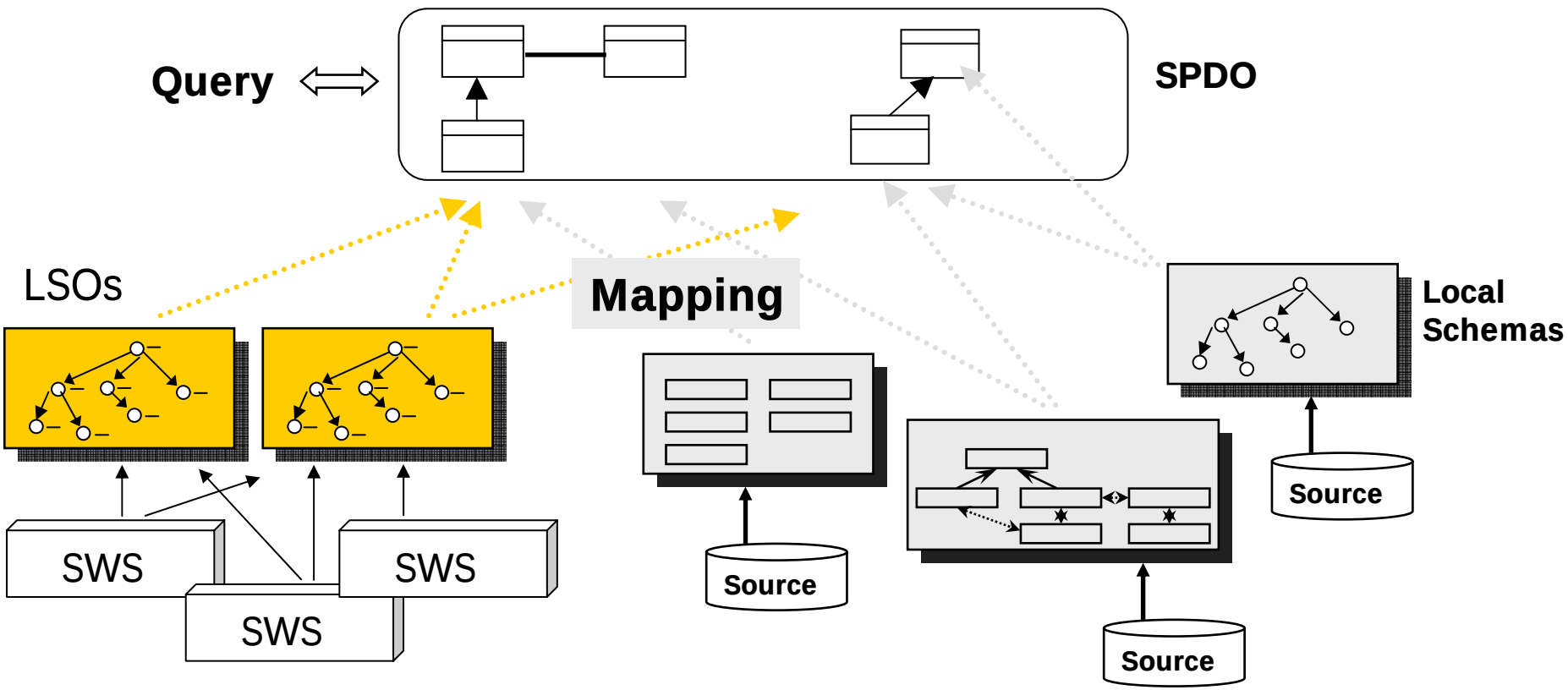
Building of the Light Service Ontologies (LSO) from SWS descriptions

- *Service ontologies (WSML-Flight)*:
  - the data ontologies exploited within SWSs (e.g. describing concepts such as Hotel, Room, TimeInterval)
  - not the ontological descriptions of SWSs (e.g. the description of a “room reservation service”)

1. SWS tagging
   - Relevant concepts associated to SWS descriptions
2. Service ontologies translation
   - Core aspects of service ontologies transformed into light-weight ODLi3 ontologies
3. LSOs creation
   - Elements of ODLi3 ontologies are mapped into services relevant to them

The Query Transformation Module

Queries for retrieving data are solved by the Query Transformation Module, providing also a number of services relevant w.r.t. the queries



Query processing w.r.t. services

- The list of web services related to the query is computed in two steps:
  1. Exploiting the mapping between SPDO and LSOs classes, a query **Q is rewritten** w.r.t. the LSOs.
     - classes and attributes of Q are substituted by the corresponding classes and attributes of the LSO, thus obtaining a query Q' on the LSOs.
  2. Exploiting the *rsm* (relevant service mapping) function the **services related** to the atoms (classes and attributes) of the query Q are obtained as result.
     - relevance degrees w.r.t. the query atoms can be aggregated and exploited for **ranking**

Matching between two Local Multimedia Sources

Global Class

```
Interface BuildingCatalog() {
attribute string        product;
attribute double        price;
attribute integer       delivery_time;
attribute double        weigh;
attribute double        size;
attribute Text          description;
attribute Image         image;
};
```

Join attribute

LMS1

```
Interface BuildingProduct() {
attribute string        product_name;
attribute double        price;
attribute integer       delivery_time;
attribute Text          description;
attribute Image         image;
};
```

mappings

LMS2

```
Interface Products() {
attribute string        product;
attribute double        unitary_price;
attribute double        weigh;
attribute double        size;
attribute Text          characteristics;
attribute Image         Photo;
};
```

Complex query elaboration

```
select from BuildingCatalog
where price < 100 and
image ~ "3654-photo.jpg" and
description ~ "glue for tiles"
```

Query unfolding

```
select from BuildingProduct
where price < 100 and
image ~ "3654-photo.jpg" and
description ~ "glue for tiles"
```

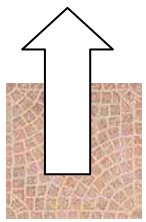
```
select from Products
where unitary_price < 100 and
photo ~ "3654-photo.jpg" and
characteristics ~ "glue for tiles"
```

LMS1

LMS2

Complex query elaboration

```
select from BuildingCatalog
where price < 100 and
image ~ "3654-photo.jpg" and
description ~ "glue for tiles"
```

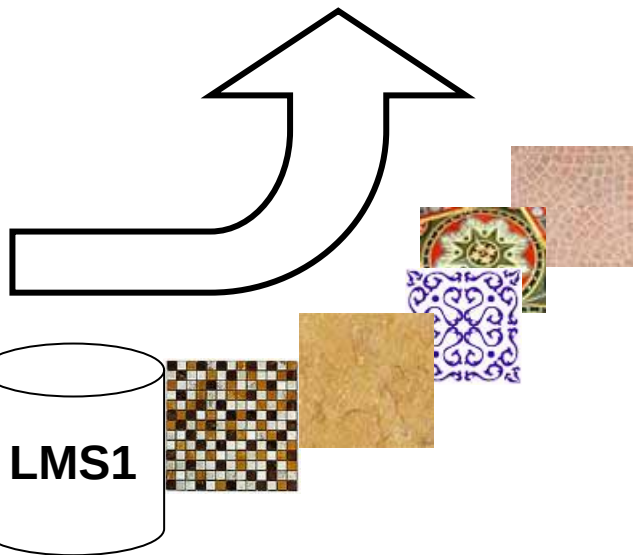


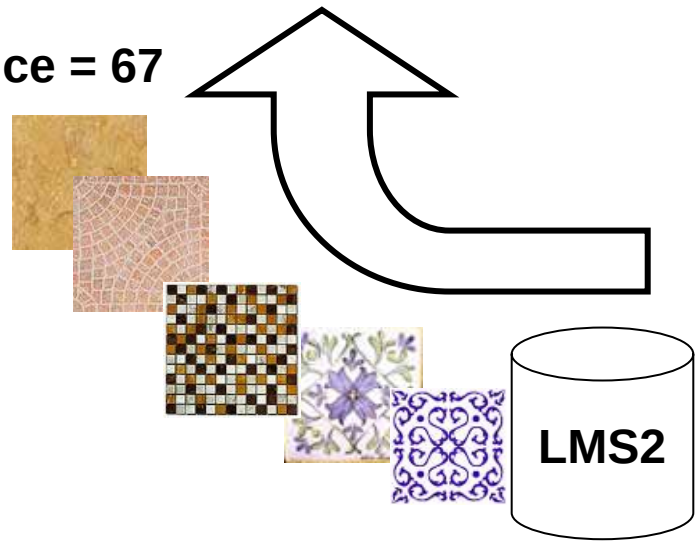
price = 55

price = 85

price = 67

Full Outer Join





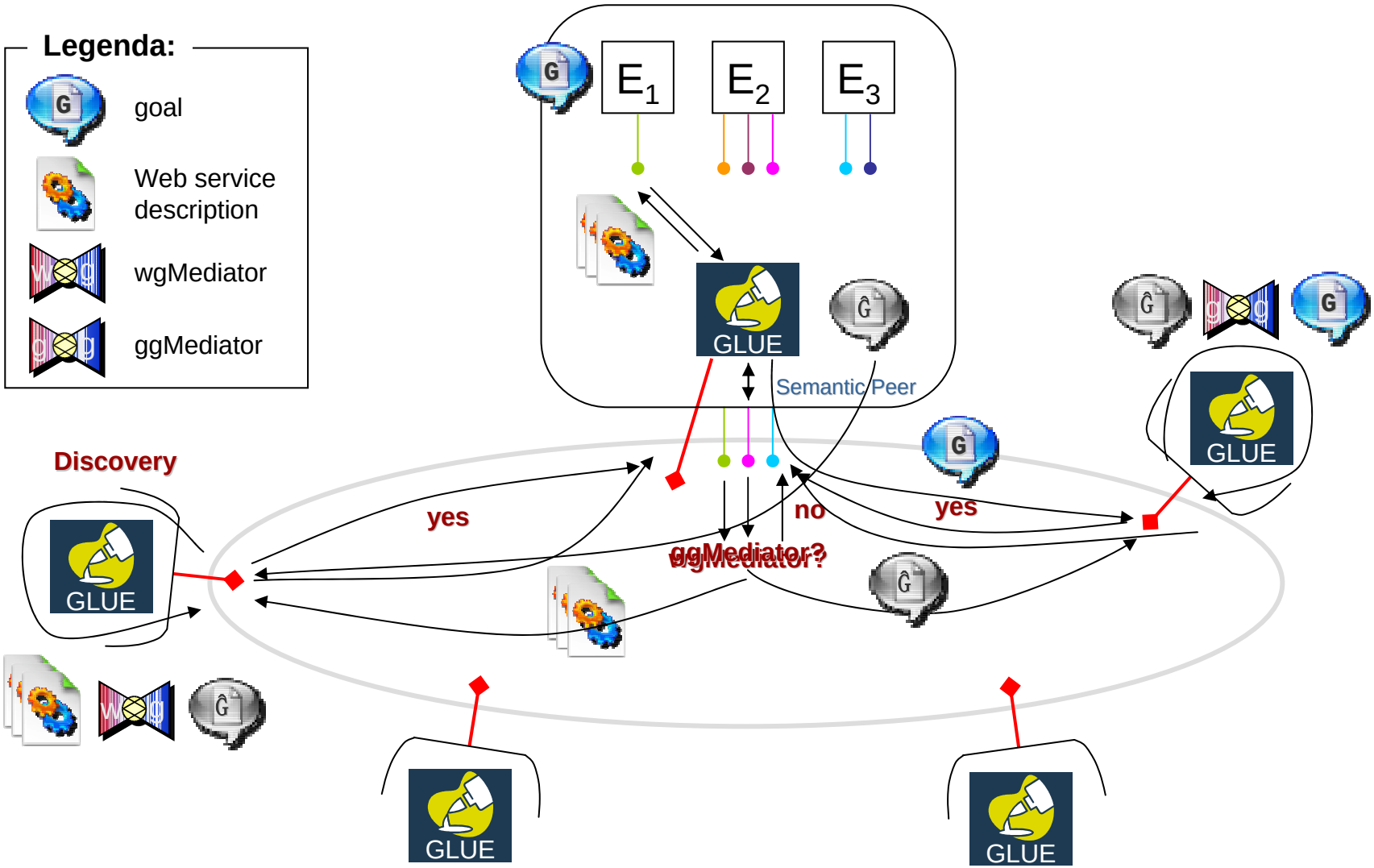
Conclusion and to-do list

- We defined an approach to provide a unified view of data and services for a semantic peer within a network
- To do list:
  - Extension of the MOMIS Query Manager to retrieve services and multimedia data in a semantic peer
  - Querying and providing services in a p2p network

References

- Project Web Site: http://www.dbgroup.unimo.it/nep4b
- M. Palmonari, F. Guerra, A. Turati, A. Maurino, D. Beneventano, E. Della Valle, A. Sala, D. Cerizza - **Toward a unified View of Data and Services - Position paper International Workshop on Semantic Data and Service Integration 2007 (SDSI07)**
- S. Bergamaschi, L. Po, A. Sala, and S. Sorrentino, "**Data source annotation in data integration systems", Fifth International Workshop on Databases, Information Systems and Peer-to-Peer Computing (DBISP2P), VLDB 2007 33st International Conference on Very Large Data Bases. University of Vienna, Austria, September 24, 2007**
- S. Bergamaschi, F. Guerra, M. Orsini, C. Sartori, "**Extracting Relevant Attribute Values for Improved Search", IEEE Internet Computing, vol. 11, no. 5, pp. 26-35, Sept/Oct, 2007 (special issue on Semantic-Web-Based Knowledge Management)**
- F. Mandreoli, R. Martoglia, S. Sassatelli, W. Penzo and S. Lodi, "**Semantic Peer, Here are the Neighbors You Want!", In Proceedings of the 11th International Conference on Extending Database Technology (EDBT 2008), March 2008, Nantes, France.**
- S. Bergamaschi, S. Castano, D. Beneventano, M. Vincini: "**Semantic Integration of Heterogeneous Information Sources", Special Issue on Intelligent Information Integration, Data & Knowledge Engineering, Vol. 36, Num. 1, Pages 215-249, Elsevier Science B.V. 2001**

SWS inter-peer discovery



Thank you very much for your attention…