Proceedings of the 6th AGILE
April 24th-26th, 2003 – Lyon, France

LOCATION-BASED SPATIAL MODELLING USING ONTOLOGY

Shijun Yu[1], Marie-Aude Aufaure[2], Nadine Cullot[3], Stefano Spaccapietra[1]

[1]EPFL-I&C-LBD, 1015 Lausanne, Switzerland, {shijun.yu, stefano.spaccapietra}@epfl.ch
[2]LISI University Lyon 1, France, Marie-Aude.Aufaure@inria.fr
[3]University of Burgundy, Dijon, France, ncullot@u-bourgogne.fr

1. INTRODUCTION

The development of mobile communication technology has released users from the constraints of desktop-bound activities; however, the type and format of services offered are mostly monotonous and mundane, focusing on email and SMS exchanges. Emerging technologies, such as Bluetooth, UMTS [1], GPRS, W-LAN have made it possible to provide better mobile communication services using a variety of geographical scales and approaches. In particular, significant improvements in terms of data transfer rate, wireless connectivity, data-format transformation, display techniques and power supply, increasingly extend mobile communications from the pure text and audio transmission to mobile multimedia.

Beyond the technical breakthrough, new social and economical life patterns appear, as more users get equipped with portable and embeddable devices, e.g. GPS and other positional sensors, PDAs (portable digital assistants), smart watches [2], mobile phones [3], cameras and home appliances.

Currently, most services provide access via mobile phones to public data stores for public transport timetables, weather broadcast, correlative market tendency, book literary performance tickets, movies [3], etc. In the near future, the focus is expected to move towards more interactive services, i.e. full support of a range of personalized and location-based services. That will materialize the idea of a *mobile office* or sophisticated *way finding* services. An example scenario for the latter is described in sect. 2.

In particular, demand for location-based services has already appeared, and prompted research in different fields such as entertainment appliances, real-time environment monitoring (e.g. traffic control), information on the move (e.g. position-enabled tourism), or interactive mobile services (e.g. virtual bus). Relevant work includes, for instance, investigations into using a data warehouse as an information centre to provide location-based services [4]. The data warehouse integrates and distributes information on the basis of the users' particular situation, preferences and needs. The task involves dynamic information generation through the interaction with users, spatial-temporal information extraction from different dimensions and measures. In [5] Pfoser & al. propose an ad-hoc database model to locate correlative data stores and exchange similar information within a specific community. The model is mainly composed of Data Handlers, Data Stores and proxies, and uses ontologies to deal with the spatial relationship between the moving objects. The ODGIS (Ontology-Driven Geographical Information Systems) framework [6] developed a comprehensive usage of ontologies for gradual classification purposes, focusing on integrating different kinds of geographic information. Kuipers [7] proposed an intelligent cognition approach to space from an artificial intelligence viewpoint, mainly focusing on orientation (e.g. left, right and front, behind) issues, and speed and interaction with the small-scale environment.

This paper[1] proposes a mixed ontologies approach, aiming at a good compromise between two antagonist goals, knowledge sharing and autonomy. Section 2 presents an example scenario and gives its concrete description, uses examples to formalize the issue. Section 3 discusses the possible research challenges. Section 4 introduces our architecture and section 5 explains how to perform data mining and present data to users, and then proposes a specific hierarchical information organization. Section 6 presents future work and conclusions.

2. A SCENARIO

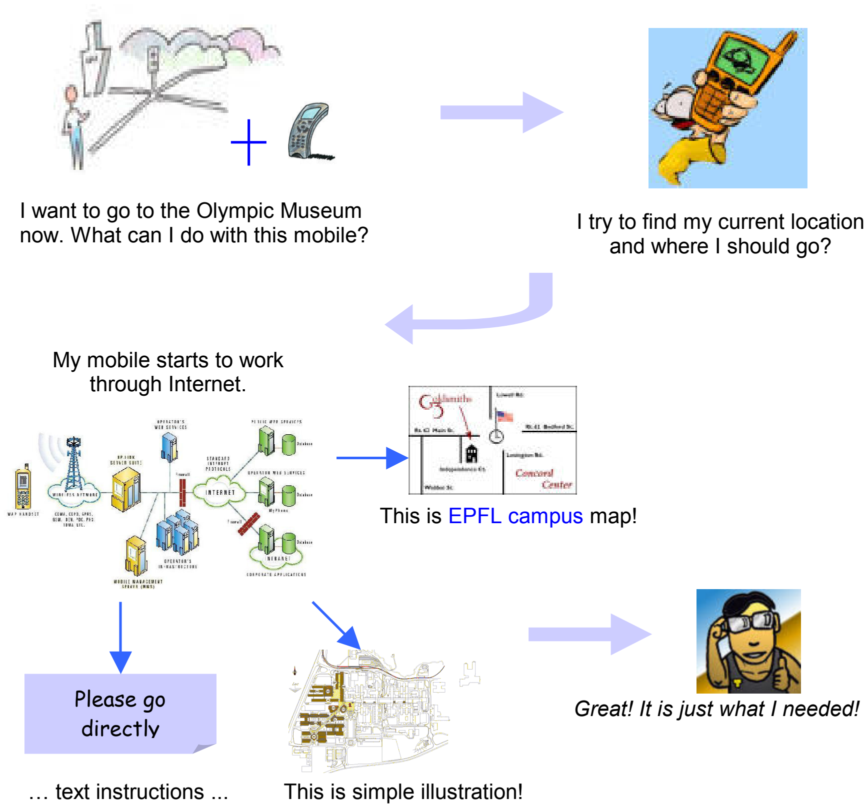

Fig. 1 A *wayfinding* scenario.

People are naturally nomadic. Thus, way finding is a frequent and fundamental activity in our daily life, and provides a significant framework for location-based services.

In the scenario shown in Fig. 1, the visitor uses her PDA to get information on her desired destination (the Olympic Museum) and on how to get there from EPFL, her current location. Information may be returned to the user in different ways, i.e. maps, illustrations and text directions. On the move, the planned trip may need revision because of external factors (e.g. traffic jams, bad weather, urgency), so query expression and evaluation is a

[1] This work is supported by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322.

dynamic, incremental and interactive process. This is illustrated in the following query example.

> ***Lausanne Tour**: After the conference, Shirley, a visitor, plans to visit some interesting places. Among top visiting sites, she decides to first visit the Olympic Museum. Shirley is also a history amateur. Getting out of the Olympic Museum, she has time to visit another museum nearby, but wants to have lunch before. Her query may now be: "Is there a Swiss and Vaudois restaurant nearby? How can I get there?" After lunch, she gets the new route to the Lausanne Historical Museum.*

Users are likely to look for different detailed information on museums and restaurants due to their diverse profiles. They may nevertheless be categorized into several fundamental communities according to motivation, constraints, familiarity and experience to the task, etc. This is similar to associating different views of knowledge with personality types [8] as discussed in Section 3.2.

3. RESEARCH CHALLENGES

Going through the scenario, several research challenges are very easily identifiable. Here we highlight four main hampers as follows.

3.1 Semantics in heterogeneous sources

With the rapid growth of information science, data has been accumulated, derived and created. Data may lack a well-formed organization. Generally, raw data has diverse formats in different data sources (e.g. text, table, database record, graphic, multimedia, and web pages). Descriptions of the same entity may exist at different levels of detail. In the 80's, distributed information systems were proposed to share information among multiple partners at different locations. In the 90's, the data warehouse approach was proposed as the most powerful technique to provide an integrated data store. Nowadays, exponential data exploration has brought more diverse data to be taken into consideration. Confronted to such a mass of data, superficial understanding of data values and data definitions is frequent, leading to losing some precision or hidden meaning, i.e. semantics. Currently, ontologies are proposed to extract more or all information behind the values, descriptions or visions. However, until now there is no completely satisfying solution or taxonomy yet to handle the diversity of required ontologies.

3.2 Users' diverse viewpoints

As user requirements are often unique in a particular situation, it is impossible for the system, even sometimes for users, to know exactly what users need. Factors participating in generating differences include:

- human features, e.g. career, age, hobby, education, budget, and language;
- cognition ability, e.g. familiarity, experience, preferences;
- external constraints, e.g. time constraints, geographic features (e.g. plain or mountain, road or district) [9], weather, traffic jam;
- uncertainty factors, e.g. urgency;
- social background, e.g. cultural events, local policies.

Most of the available data stores are targeted to fulfil specific functionality, e.g. traffic management, land management, tourism guide, history description, sport games. To make them interoperable interrelationship have to be built and constraints defined to navigate in the diverse parts and hierarchies. Data integration and data cleaning are major issues here.

3.3 Presenting the results

Data visualization and level of detail level should be personalized. Experts can get the necessary information through maps or professional material for they can make exact abstraction and generalization with their knowledge and experience. The local citizens possess comparatively strong cultural and geographical background, e.g. the main streets, shops, landmarks and holiday celebrations. What they need may be only a display of the area simply composed of relevant streets, landmarks and orientation. A new visitor, with no or little knowledge about the city, needs more details and appended information than city inhabitants. Many more alternatives exist.

3.4 Hardware constraints

To cope with limited display space, we should refine and transform the data for desk-PC. Table or frame formatting is to be avoided, and replaced with lists, and high-resolution maps replaced with simple illustrations or low-resolution maps. Interfaces should be designed for navigation. To use query history, caching techniques should be adapted. Another problem is transfer of web-pages, originally designed for large screen monitors, into formats compatible with small displays on handheld devices.

4. ARCHITECTURE

There is nowadays more information than what we can locate for a particular domain and task. Information is distributed in various sources with different formats, e.g. relational databases, GIS databases, multimedia databases, data warehouses, websites, XML files, etc. Hot issues in this new context include: How to locate relevant sources? How to navigate between the different sources for complementary content or details? How to describe the semantics of data in the sources? How to relate the sources with similar content or focus? If each source could be described in some formal and meaningful manner, it would be much easier to fully access to the relevant data and effectively process and interpret data in remote systems. Existing integration approaches have been discussed in [10]. Here we propose an approach for ontology-based information searching, assuming the architecture shown in Fig. 2. The elements in the architecture are:

- the global ontology,
- It is an acknowledged and comprehensive classification approach, e.g. WordNet or OntoLingua, that serves as reference ontology for the *local ontologies, shared ontologies* and *integrated ontology*,.
- the local sources and local ontologies.
- Each *local source* has its own *local ontology* that defines its content. These autonomous local ontologies may be derived from global ontology or other public ontologies according to its particular domain or format. While these local sources remain heterogeneous, they are grouped according to their sharing of ontologies.
- the shared ontologies.
- Shared ontologies are intended to exhibit the commonness and characteristics of every associated local source. They can be specified according to geographical scales, domains, formats or user profiles. The mediator could effectively access to the suitable group and further local source. Shared ontologies derive and evolve with the local ontologies, and its changes are propagated to the integrated ontology.
- the Mediator,
- It is used to respond users' queries and identify relevant shared ontologies. When the query is received, the mediator analyzes the query and locates relevant sources using the integrated ontology, and finally presents the answer that is merged, assembled or redefined. Furthermore, it may keep, fully or in part, the result of frequent queries.
- the integrated ontology.
- It is set up considering existing sources and users' viewpoints. It is derived from the shared ontology and created with the assistance of the global ontology. It

continuously reflects the semantics and hierarchy changes of the local sources, more than the global ontology does. Its ontology construction is more likely a comprehensive tree tailored to support the different users' viewpoints. Each branch is characterized by a semantic label, which can be derived from a simple tree for shared ontology.

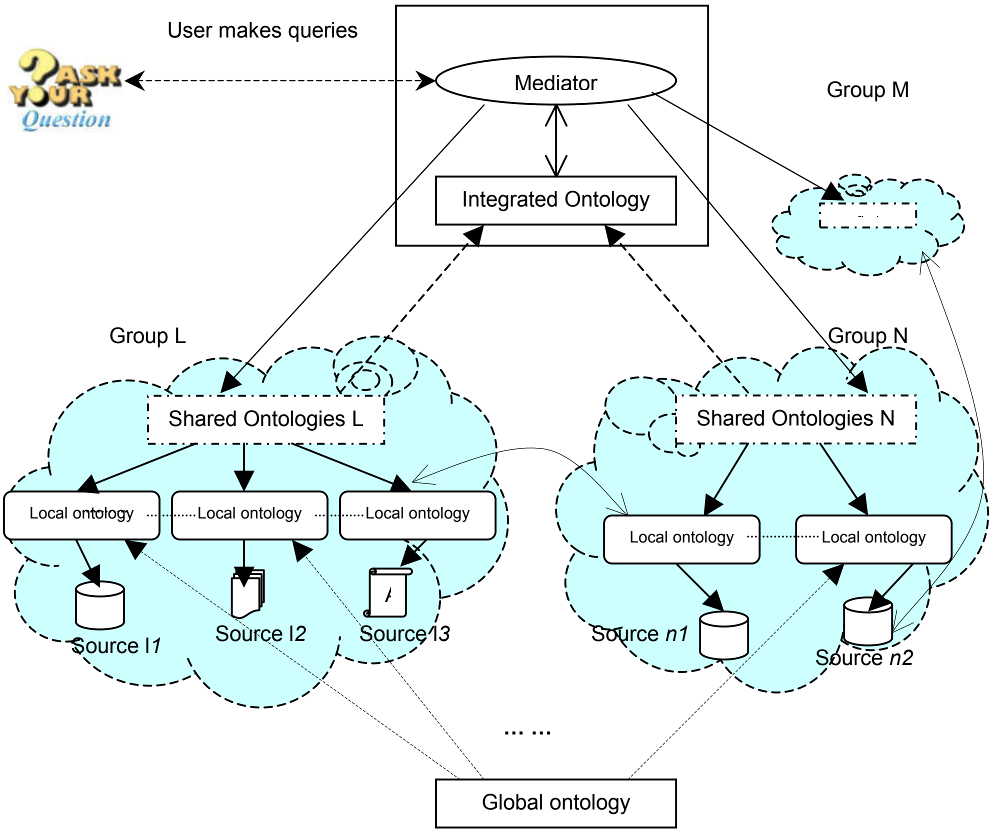

Fig. 2 An architecture for ontology-based information searching

Obviously, ontology mapping between related local sources is essential to the architecture. An ontology-mapping technique has been developed in e.g. [12]. Supported domain differences can be categorized into four types: terminology, scope, encoding and context, as described by Wiederhold [11]. We will explain how to materialize the architecture using the ontology trees in the next section.

5. THE PROPOSED APPROACH

From the users' points of view, the same real world entity or phenomenon can hold different contents and scenarios, as shown in Fig. 3.

The Olympic Museum can play a few roles in different domains, e.g. sports, museums, and top visiting places. As an unacquainted traveller, the user can care more about the museum' geographical position, available access routes, open hours, and current activities. The museum curator will focus on internal maintenance and management, and feedback from visitors. A transportation department officer will focus on the traffic status and noise pollution at various times.

Another critical fact is that most of existing information systems are closely related to certain foci. They often give the details in some aspects or provide services for a particular user-groups. How to make use of the relevant sources and relate them is discussed next.

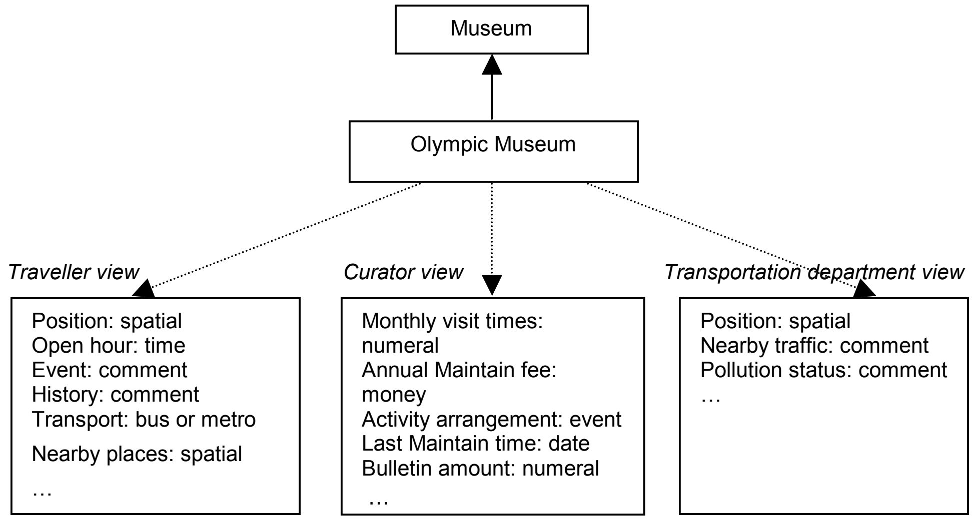

Fig. 3 The information of entity *Olympic Museum* from different users' viewpoints

5.1 Tree structures for information content

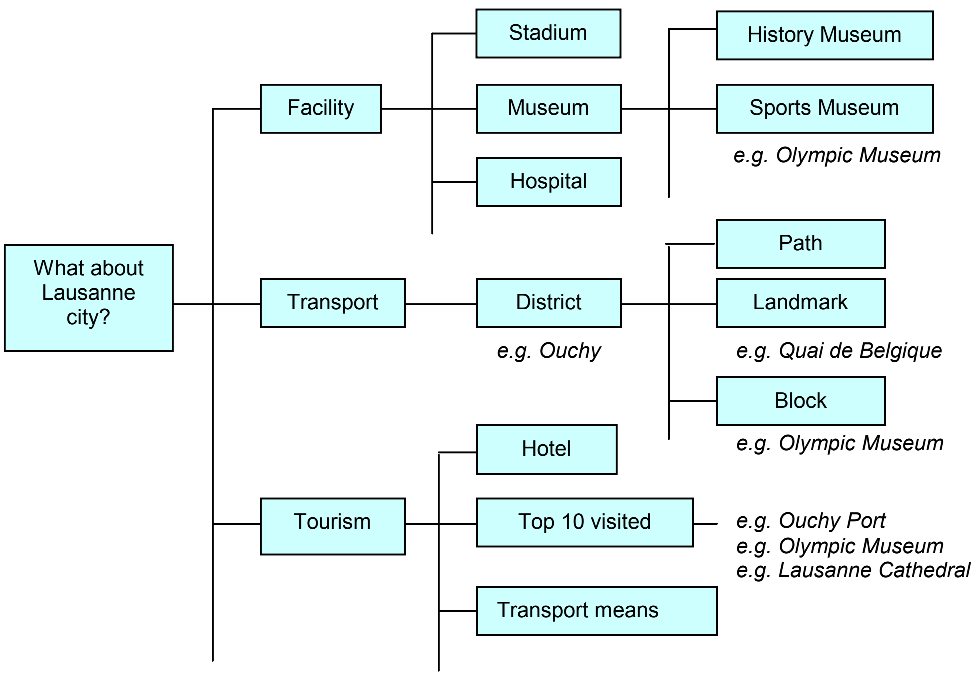

Fig. 4 The possible hierarchy of the city *Lausanne* from different viewpoints.

Hierarchical data structuring is a very common way to organize information. Fig. 4 shows a possible hierarchy for information in the city of Lausanne. Other hierarchies of descriptions about the city Lausanne may exist. We assume that existing information sources are grouped so that all sources in a group share the same ontology. This ontology may be defined from scratch or derived by analysis of existing schemas, keywords, tags, functionality, etc. Each information tree represents one possible hierarchical integrated ontology construction. For instance, Data Source l3 in Group L (in Fig. 2) may be a database about Lausanne City Facility Management, and represented in Figure 4 as the tree with root Facility. As the instance level, Olympic Museum can be described from the viewpoint of history, layout, and importance in the city development. The city transportation department is responsible for monitoring and canalizing the traffic, especially near famous landmarks or main streets, making reports or tables about the traffic status. The details about Olympic Museum's location, traffic, available buses are stored in Data Source n2 in Group N. As an important economic branch of city Lausanne, tourism information (shown in Fig. 4) can also be found from Data Source m1 in Group M. All data sources above contribute to the characterization of the city of Lausanne.

The tree is built to support multiple ontologies, separately for the curator's, the transportation department officer's and the common traveller's views. Moreover, it can be further developed according to user's profile, i.e. motivation, preference, constraints.

To nevertheless provide knowledge sharing, we dynamically set up and update the links and interactions between the relevant nodes, which could be in the different hierarchies. For example, Olympic Museum is an instance of both 'Sports Museum' and 'Landmark'. Obviously its functionality and emphasis are different in the two cases. The Path instance Quai de Belgique is the road along which the Olympic Museum lies. The traffic on the street Quai de Belgique at certain times will have an effect on some visitors and on the local transportation department. Even for instances in the same hierarchies, e.g. Ouchy Port and Olympic Museum, in addition to the fact that they belong to the same category 'Top 10 visited', there exists a special hidden relationship: They are close to each other. Such hidden relationships are an example of benefits expected from using a tree structure to improve functionality for the convenience of the unfamiliar visitors.

5.2 Using the tree

Generally, queries can be divided into sub-queries, especially when users preferences are taken into account. The mediator can make better decisions by identifying the matchings, based on terms in the ontology (keywords, schema or ontology), between user's query and system's ontology. The tree allows bridging the gap between the system architecture and the user query. The following advantages are identifiable:

- effectively locate the data resources;

  Each tree is constructed in a similar way within the group sharing ontologies. By developing an appropriate mapping between the groups and trees, the task of locating the information is made easier.

- easy to navigate between the heterogeneous sources;

  It is not necessary to provide all details considering efficiency and needs at a single node or level. Available information is distributed in different data sources. Each source has some certain emphasis that cannot be found in others, e.g. the transportation department keeps the traffic records about some regions but the tourist guide probably provides the information about parking in that region. On the other hand, it is impossible to overload the information to a node, which will result in the tree's redundancy and unbalance.

- easy to add or modify.

  The proliferation of data is dynamic and incremental with the evolution of society. When data is created, derived or evolves, the data structure needs to be updated. Using the tree, it is easy to add new branches or nodes or combine or modify the existing ones. The task at hand is to change the relevant semantic labels to provide the new accessibility to quit or join the particular group.

We assume that every branch in the tree is characterized by some semantic labels corresponding to the shared ontologies. When the query is received, it is analysed and decomposed, then recomposed using the ontologies recognized by the information systems (searching the top-level in the tree, finding the optimal branch, then going deep in the tree). When there is a difference or conflict with user's requirement, the process traces back to the higher level and goes to its relevant brother nodes. A variety of similar algorithms have been presented, e.g. SMART [13].

6. CONCLUSION AND FUTURE WORK

Providing location-based services is an increasingly demanding challenge in our free and convenient e-commerce era. To model and fully use the existing information is a necessary prerequisite to better services. We have described issues involved in supporting an ontology-based information searching approach. We presented our architecture of mobile information integration and the approach of navigating between the branches and nodes in the *ontology tree*. Building the mappings between ontologies and between the heterogeneous data sources and making precise inference reasoning regardless of the data sources appear as the major issues. Assessing our approach to support mobile users will be the focus of our future work.

REFERENCES

[1] UMTS (Universal Mobile Telecommunications System), 2002.

[2] New CASIO PRT-2 Watch integrates the GPS technology into the wristwatch, 2002.

[3] FOMA (Freedom Of Mobile multimedia Access) services using UMTS [1] standards in Japan. http://foma.nttdocomo.co.jp/

[4] C. S. Jensen, A. F-Christensen, T. B. Pedersen, D. Pfoser, S. Saltenis, N. Tryfona, Location-Based Services – A Database Perspective. *8th Scandinavian Research Conference on Geographical Information Science, ScanGIS2001 Norway, 2001*

[5] D. Pfoser, E. Pitoura, N. Tryfona, Metadata Modelling in a Global Computing Environment. *10th ACM International Symposium on Advances in Geographic Information Systems, ACM GIS 2002 McLean, VA, 2002.*

[6] F.T.Fonseca, M.J. Egenhofer, P.Agouris, Using Ontologies for Integrated Geographic Information systems, *Transactions in GIS 6(3), 2002.*

[7] Benjamin Kuipers, The Skeleton in the Cognitive Map: A Computational Hypothesis, *3rd International Space Syntax Symposium, Atlanta, 2001*

[8] G. Dueck, Views of knowledge are human views, *IBM System Journal, Vol. 40, No4, 2001*

[9] T. Tezuka, R. Lee, Y. Kambayashi, H. Takakura, Web-Based Inference Rules for Processing Conceptual Geographical Relationship. *1st International Workshop on Web Geographical Information Systems, WGIS'01, Kyoto, Japan, 2001*

[10] H.Wache, T.Vőgele, U.Visser, H.Stuckenschmidt, G.Schuster, H.Neumann and S.Hűbner, Ontology-Based Integration of Information—A Survey of Existing Approaches. *IJCAI 2001 Workshop on Ontologies and Information Sharing, 2001*

[11] Wiederhold, G. An Algebra for Ontology Composition. *1994 Monterey Workshop on Formal Methods. U.S. Naval Postgraduate School.*

[12] Omelayenko B. RDFT: A Mapping Meta-Ontology for Business Integration. *Workshop on Knowledge Transformation for the Semantic Web (KTSW 2002), 15-th European Conference on Artificial Intelligence, Lyon, France, 2002*

[13] Natalya Fridman Noy, Mark A. Musen, SMART: Automated Support for Ontology Merging and Alignment. *12th Banff Workshop on Knowledge Acquisition, Modelling, and Management, Banff, Alberta, Canada, 1999.*