INTEGRATING E-LEARNING OBJECTS IN A P2P SYSTEM

Gabriel André D. Dubois Brito
Military Institute of Engineering - IME/RJ - Engineering System Department - Rio de Janeiro - Brazil

Ana Maria de C. Moura
Military Institute of Engineering - IME/RJ - Engineering System Department - Rio de Janeiro - Brazil

Keywords: Peer-to-peer technology, Distributed and Parallel Applications, Data Semantics, Information Integration.

Abstract: ROSA is an e-learning system, which enables the creation, storage, reuse and management of Learning Objects (LOs). LO is a collection of reusable material used to support learning, education, or training. However, since ROSA is still a centralized system, it does not provide yet a complete integration of LOs created in local ROSAs of other institutions. This paper presents the evolution of ROSA into a peer-to-peer (P2P) system - the ROSA - P2P - and describes the integration process of LOs in this environment. It provides the required interoperability to execute queries throughout all ROSA - P2P peers, taking into account a strategic data integration system that includes queries rewriting based on their semantic meanings. Controlled vocabularies are also used to support the query rewriting process and the identification of relevant peers that are able to answer queries on a specific knowledge domain.

1 INTRODUCTION

ROSA (**R**epository of **O**bjects with **S**emantic **A**ccess) (Porto et al., 2004) is an e-learning centralized system used by academic professionals in educational area. It has been conceived to support the design phase of e-learning courses, where tutors want to share and search for course material.

However, in order to confer a real role of inter-institutional cooperative environment on ROSA, it is important that contents can be stored in different institutions, before being interchanged and integrated. This would be essential to provide global answers to queries submitted by users through local ROSAs.

Advances in distributing computing, fostered by the real need to exchange information on the Web, led to the development of standard interconnection specifications to support semantic data interoperability and integration. In this context, metadata, ontologies and P2P technologies raise as important research issues to support semantic heterogeneity integration worldwide. The former provides for data description concerning different schemas, including mappings, associations, source locations, etc., furnishing essential information for data integration. Ontologies are used to represent the semantics of a knowledge domain, and in conjunction with metadata and controlled vocabularies, are essential to ensure the correct query interpretation. Additionally, they are responsible for providing systems interchange, supplying them with more refining queries, and enabling more relevant and precise answers. P2P is a recent technology that aims to harness Internet-connected resources at a global scale, and can be self-organizing, ad-hoc and decentralized.

The purpose of this paper is to describe the ROSA transformation process into a P2P environment, generating the ROSA - P2P system version, and to develop a strategy to integrate e-learning objects in this system. In this context, a P2P architecture based on super-peers was developed (Brito, 2005), including specific strategies to provide: peers connection/disconnection into/from the P2P network, such as super-peers grouping based on knowledge domain and location; super-peers definition and election; peers balancing and redistribution in the system; and some fault tolerance issues. Another great contribution of this paper concerns the e-learning objects integration strategy

André D. Dubois Brito G. and Maria de C. Moura A. (2006).
INTEGRATING E-LEARNING OBJECTS IN A P2P SYSTEM.
In *Proceedings of WEBIST 2006 - Second International Conference on Web Information Systems and Technologies - Internet Technology / Web Interface and Applications*, pages 214-221
DOI: 10.5220/0001244502140221

defined for ROSA - P2P, which is completely described. Each query is only broadcasted to the relevant peers in the system, which is rewritten according to its domain semantics and executed based on ROSA algebra (Coutinho and Porto, 2004). Partial results returned from individual ROSA peers are then sent to the requested peer, which is responsible for the objects integration and presentation to the user.

The rest of this paper is organized as follows. Section 2 presents a brief description of ROSA system, with its main objectives and functionalities. Section 3 introduces P2P systems, focusing on the main types of architectures, generally used to implement these systems. Section 4 describes the ROSA - P2P system, giving emphasis to its architecture and strategies adopted for peers configuration and e-learning objects integration in this environment. Next, section 5, presents some related work, and finally, section 6 concludes the paper with additional comments and future work.

2 ROSA SYSTEM

ROSA main purpose is storing e-learning objects (LOs), and exploring their access according to the context in which they have been created. A LO is identified by a set of metadata descriptors established by an international metadata standard, such as LOM[1] (Learning Object Metadata). These metadata are organized into a hierarchy, providing information about identifier, title, keywords, idiom, version, aggregation level, etc. Indeed, LOs represent instructional contents, whose contexts are determined by semantic relationships between them. These relationships are expressed through a conceptual map, according to a well-defined model (Porto et al., 2004).

A conceptual map is represented by a directed graph where nodes correspond to LOs, identified by their names, and arcs refer to relationships between them, such as RDF[2] (Resource Description Framework) predicates. ROSA also provides an algebra and a query language, the ROSAQL (Porto et al., 2004), so that semantic queries such as "*which course material does an OO Database topic **comprehend***? and *"which subjects are **basis for teaching** Query Optimization?"* are supported by ROSA system, taking into account the predicate semantics. In these examples ***comprehend*** and ***basis for*** are part of a pre-defined predicate set that relates different LOs. These can be of two types: logical and physical LOs. A logical LO represents a collection of LOs, which may contain several physical LOs; and a physical LO corresponds to a stored LO, such as files (.jpg, .doc, .ppt, etc.). Questions related to synonyms, specific/generic, and associated terms are supported by a domain thesaurus that helps during query processing.

3 P2P SYSTEMS

P2P systems are characterized by the sharing process of computing resources and services through a direct and decentralized communication among systems (Ooi et al., 2003). They can be classified according to 3 basic types of architectures (Brito and Moura, 2005): **partially centralized**: contains a central server responsible for the search mechanism and infra-structure maintenance, leaving to the participant peers the task of sharing resources and services in a distributed way; **decentralized**: does not have a central peer, and the search mechanism and infra-structure maintenance, as well as services and information contents, are distributed throughout the network, in each participant peer; and ***super-peer***: is composed of a set of inter-linked peers with higher computing capacity, named super-peers. These are responsible for the management and sharing of resources, where each super-peer has other peers linked to it. Due to its characteristics, this architecture raises as the most adequate for developing and maintaining P2P systems, since it provides, besides other advantages: time reduction for research; fault tolerance; super-peers management; scalability; and a reasonable accepting confidence level (Brito, 2005).

Although it does not exist yet a consensus concerning a well-defined topology for data integration architectures in P2P systems, we adopted a data integration strategy exclusively dependent on the system objectives, architecture, functioning and specific characteristics according to ROSA - P2P system proposal. This strategy is very important in P2P data integration systems, so that data can be stored, filtered, accessed and integrated in an optimized way. However, to provide more flexibility in the schema mapping process, this strategy should incorporate a semantic connotation. This is accomplished using metadata, domain ontologies and controlled vocabularies in order to improve and facilitate the semantic interpretation and integration of objects stored throughout the peers network.

[1] http://ltsc.ieee.org
[2] http://www.w3.org/RDF

Due to the dynamic nature of the e-learning environment, P2P technology raises as one of the most appropriate infrastructure for developing this kind of system, since it encourages the creation of educational communities in an easy and cheap way; allows information to be shared and organized by didactic contents; increases the volume and quality of instructional resources; offers the sensation of data readiness at all moment to the user, still making it possible efficient search, since queries can be processed in parallel (Nguyen and Sanchez, 2004).

4 ROSA - P2P SYSTEM

ROSA - P2P aims at integrating e-learning objects in a P2P environment, where users will be able to submit queries using either a portal (Toledo, 2002), named ROSA portal, either a ROSA - P2P peer. Hence, when a query is submitted by a user through a super-peer, the latter will verify if it is able to answer the query. In affirmative case, it will rewrite the query to itself, storing the result in cache. Then it will resend the original query to its own peers and to other relevant super-peers, activating a clock that will control the time a super-peer will wait for having results. Nevertheless, when all the results will have been returned, or the corresponding waiting time will have elapsed, all results stored in cache will be integrated and the final query result will be returned to the user.

The system proposal will be presented according to its functionalities in the following sections.

4.1 Internal Architecture

Figure 1 illustrates the internal system architecture, showing its modules and components, as following:

- **Interoperability Module:** is composed of the interoperator component (P2P), which presents the necessary characteristics and functionalities to create and maintain a P2P network, such as connection establishment, routing indices maintenance, super-peers election and network balance;
- **Query Processing Module:** is composed of the user interface component, which is responsible for providing a more friendly communication environment between users; and the query processing component, which adopts a strategy classified in two phases, as presented in section 4.3.2;
- **Data Management Module:** consists of two components: controlled vocabularies, which support semantic interpretation during query execution. Indeed, it facilitates peers information interchange, providing for more precise searches and most relevant results; and data cache/integrator, responsible for temporally storing partial query results. Once received the partial results from the relevant peers and/or super-peers, the integration process is started, after which the final query result is returned to the user.

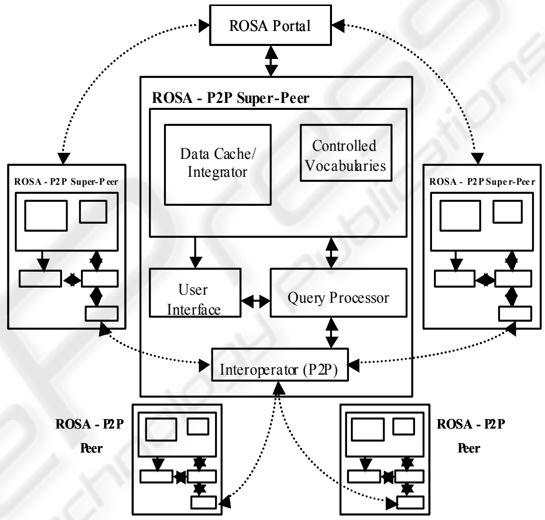

Figure 1: ROSA - P2P architecture.

4.2 P2P Environment

ROSA - P2P system has been developed according to the characteristics described in section 3, and it is based on a super-peer architecture.

Even though ROSA portal refers to super-peers, it does not take part in the P2P architecture as a whole. It is situated in a layer above, and it is used as a starting point for users to submit their queries through the Web, in case they do not have ROSA - P2P at their disposal, as well as an exit point to receive their query results. However, its hosting machine is used to store some important services provided by the system, such as the Directory Service (DS), which is responsible for making available the list of existing super-peers to the new peers that want to connect to the system for the first time; and the Controlled Vocabulary Delivery Service (CVDS), responsible for storing all controlled vocabularies (global, local and of keywords) in this machine, external to the P2P environment. Hence, when a peer connects to the system for the first time, it receives these vocabularies through the network. These services will be better described in sections 4.2.1 and 4.3.1 respectively.

4.2.1 Main Strategies

This system adopts a strategy based on peers grouping, similar to the one defined in Edutella system (Nejdl et al., 2003), referenced as peers aggregation. However, in this work we used a more comprehensive strategy, where this first idea has been extended. It consists in grouping these aggregations, called here super-peers grouping, as shown in Figure 2. Aggregations are created according to two important features: subject and geographical localization, whereas groupings are classified according to the subject dealt by the corresponding aggregations. This strategy ensures that peers with similar characteristics stay close to each other, facilitating its localization and optimizing query processing (Brito, 2005).

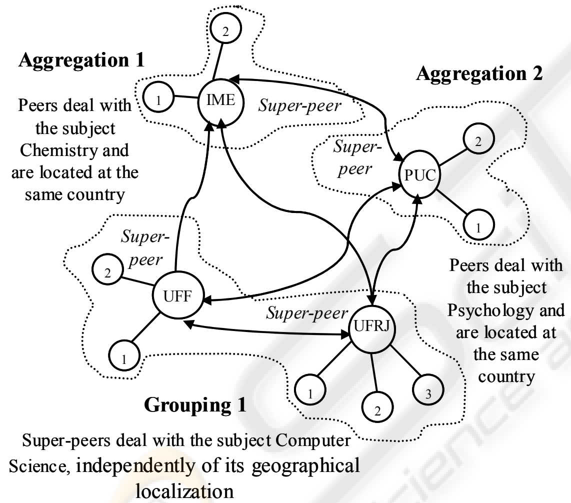

Figure 2: Aggregations and groupings in ROSA - P2P, where IME, UFF, UFRJ, PUC are universities in Rio de Janeiro.

There are two different ways to connect a peer to the system: i) when the peer connects to the system for the first time. In this case it sends a query to the DS, requesting for a list of available super-peers. Then it checks, among the existing super-peers, those that have the same subject and localization similar to its own (it is worthwhile mentioning that this information is provided whenever a ROSA peer is installed in the system). Once this information is obtained, it sends to the respective super-peer a connection request, which validates it, verifying if the new peer can indeed take part in the P2P network. This validation attests if ROSA - P2P system is present in the system. Once the connection is established with the super-peer, the latter has to provide information about its metadata, such as: machines' name, IP address, subject, origin country, if it wants to be a super-peer in the future, etc; ii) when a peer wants to be reconnected to the system. In this case, the peer has already a reference to its super-peer, and hence it can automatically connect to the system. In both cases, once connected, the peer is already able to share resources and submit queries.

In order to be a super-peer, a peer must provide some important physical characteristics to ensure a good system performance (Zhu et al., 2003). In ROSA - P2P, it was considered that every academic institution would be automatically a super-peer. Otherwise, some relevant information concerning physical characteristics of that peer must be provided by the user at the moment he/she installs the system.

The number of existing super-peers is dynamic, defined according to the maximum quantity of peers that a super-peer can support. This quantity is determined in function of the result time evaluation of queries submitted to similar hardware machines, situated in different locations (Brito and Moura, 2005). So, the ideal quantity of super-peers in the system will be indirectly balanced.

Election of super-peers occurs only among aggregations and groupings (instead of considering all peers in the system), i.e., a peer can only be a super-peer within its own aggregation or grouping.

System balancing happens whenever a super-peer has more peers than another of the same grouping. In this situation, there will be peers redistribution among not balanced supper-peers of each grouping, in order to ensure better performance to queries results.

According to Nejdl. (Nejdl et al., 2002), the use of super-peers indices minimizes significantly query redistribution time among relevant peers, i.e., those that are able to answer a specific query. Thus, the system adopts a strategy based on routing indices, such as in Edutella system, using two data structures, named routing tables. The first concerns communication between a super-peer and its respective peers (SP/P); and the other provides information of a super-peer and its super-peers (SP/SP). This information focuses on data and metadata of peers, such as: subject domain, location, peer status (online/offline) and physical characteristics, which are provided by each peer and super-peer. This information is used to optimize the query redistribution process among relevant peers. Due to ROSA - P2P dynamic characteristics, these tables must always be kept updated neither to prejudice system performance nor its reliability. This process is accomplished by triggers, which are able to detect all modifications occurrences and update them in the DS and routing tables.

The system also provides fault tolerance mechanisms. These procedures increase system reliability, since they are always available to solve possible faults, avoiding it becomes inoperative.

4.3 Data Integration System

Taking into account the architecture, functionalities and characteristics of ROSA - P2P system, it was possible to define a valid architecture for data integration. Each peer has its own data integration component, including an entry query point, which is also used for integrating results. A friendly interface has been provided to interact with the user, allowing him to submit queries in an easy way, where results are exhibited clearly (Brito, 2005). Data integration is supported by controlled vocabularies, responsible for providing some semantic value to data. They help finding data and relevant peers, rewriting queries, and solving semantic conflicts during all the integration process.

Data integration is still a great challenge, since it depends directly on the way the semantics of a concept is defined in a peer. This information is essential to build a global integrated view, yet a complex task to manage. Special attention has been given to optimization aspects, in order to provide simplicity, performance and reliability, as described next.

4.3.1 Controlled Vocabularies

As already mentioned, ROSA - P2P system uses controlled vocabularies to support data integration. These structures raise as powerful tools to facilitate semantic interpretation and information retrieval, whereas providing systems interoperation and enabling more refining searches, restricted only to relevant information. In fact, they become essential to: correctly locate relevant peers to answer a query; help peers in the query rewriting process; solve semantic conflicts; suggest options and associated paths related to the corresponding search, helping the user to reach his/her objectives; and in the automation of tasks that require reasoning. In order to reach this objective, the system uses three different vocabularies:

- **Global controlled vocabulary:** is used by all peers in the system. It is composed of a synonymous vocabulary according to existing predicates in ROSA system and some specific properties borrowed from the thesaurus approach, added of some LOs predicate properties, such as transitivity and symmetry. This vocabulary is very important in the navigation rewriting operation (section 4.3.2), enabling the query to be rewritten taking into account all relevant data, independently of the semantic used by each peer to describe a predicate;
- **Local controlled vocabulary:** specifies vocabularies according to each existing subject domain referred to in the system. Thus, the system will have as many local vocabularies as the number of subjects treated in the system. Differently from the global vocabulary, only peers concerned with a specific domain will be supplied of a corresponding local vocabulary. This vocabulary is based on a thesaurus structure, composed of equivalent, generic, specialized and associated terms. It is also very important to rewrite the selection operation (section 4.3.2), since it allows a query to be rewritten based on all its relevant concepts, independently of the semantic used by each peer to describe a LO.
- **Keywords controlled vocabulary:** consists of a vocabulary associated to each existing subject domain in the system. It makes it possible to detect a query subject in running time, so that the query is only sent to the peers that are able to answer it. It is composed of a set of semantic related terms on a specific knowledge domain.

In order to deliver these vocabularies to the users, the system adopts a specific strategy. It is managed by the service CVDS that is available in ROSA portal. This service consists in storing all the controlled vocabularies in this portal, external to the P2P environment, so that whenever a peer connects to the system for the first time it can receive them via network.

4.3.2 Query Processing

According to Arenas (Arenas et al., 2003), query processing is the most important service in a P2P network, consisting basically of the query distribution among peers. In the context of ROSA - P2P system, query execution uses the query execution machine named MEC ROSA (Coutinho and Porto, 2004). Being responsible for the ROSA algebra implementation, it is composed of a set of operators to manipulate ROSA data such as: select, project, browse, join and transitive closure.

However, as ROSA was initially built as a local system, MEC ROSA access was restricted to a single database, hence not allowing for the generation of a query distributed plan, which would be the most adequate to the ROSA - P2P distributed environment. Thus, it was necessary to define a strategy to use MEC ROSA in the system, which will be described later in this section.

As query process in ROSA - P2P takes into account a rich semantic context, it uses a particular query processing strategy, executed according to the following steps:

i) Sending queries

It consists in transmitting a query submitted by a peer to other peers in the system, making it possible this query can only be executed and answered by the peers concerning a query domain. Therefore, the main strategy adopted to optimize query processing was firstly locating the relevant super-peers related to the query domain.

Thus, when a query is submitted to the system, the query processing initially verifies if the metadata descriptors "title" and/or "keywords" are included in the query. If it is the case, these values are compared to the existing terms included in the keywords controlled vocabulary. Once matched, the corresponding subjects are then returned and the query is sent to the peers associated to that knowledge domain. In case these terms are not part of the query, or if their values are not located, not providing for the query domain identification, a message is sent to the user, asking him/her for one or more terms to include some additional semantic value to the query. Some terms examples are then exhibited, so that the user can suggest a term to identify the query domain, allowing it to be sent to the appropriate super-peers.

ii) Rewriting queries

It consists in rewriting queries through the selection and/or navigation operations, which are then rewritten according to the information stored in the global and local vocabularies. This way queries are able to encompass a more extensive data universe, making it possible that all possible answers can be retrieved, independently of the way data have been semantically stored.

Therefore, once the query processing verifies the query contains a select operation including the metadata "title", it will compare the title value with the other synonymous terms of the local controlled vocabulary. If it matches, equivalent terms will be retrieved, and the select operation will be rewritten, i.e., these new terms will be added to the select clause and linked through the disjunction operator (or). In case there exist other metadata in the select operation such as "aggregation level" (this metadata descriptor identifies if the object is a program, course or topic), they should also be included in the query, concatenated by the operator "or" or "and" and placed at the end of the sentence according to the *conjunctive normal form* (Coutinho and Porto, 2004) used by the select operation. The use of the operator "or" or "and" changes according to the select operation initially defined by the user. If the query does not contain the metadata descriptor "title", it is not rewritten.

The query rewriting process continues, this time for the navigation operation. Once it is included in the query, the query processing compares each of the predicates declared in the query with the synonymous stored in the controlled global vocabulary. For each query predicate, its corresponding equivalent predicates are retrieved from the vocabulary and rewritten, and hence forming a set of rewritten predicates, which are then concatenated through the disjunction operator (or). This operator, together with the "and" conjunction and the "." navigation operators are used to join all sets of rewritten predicates. Finally, the navigation (or *browsing*) operation rewriting will be complete when the rewriting of all predicate sets is joined within the same sentence. The following example illustrates these procedures. Suppose a query defined as: "Select the LOs titles generated by those that comprehend and fundament other LOs whose "title" is equal to *distributed database* and their "aggregation level" is equivalent to *course*".

Query: select|LOs@lom/general/title = distributed database and LOs@lom/general/aggregation_level = course
browsing|LOs@((comprehends and fundaments). generates)
project|LOs@lom/general/title

This query will be rewritten by the query processing into a semantically richer query, defined as:

Query: select|LOs@lom/general/title = distributed database or LOs@lom/general/title = DDBMS and LOs@lom/general/aggregation_level = course

browsing|LOs@((((((comprehends or consists_of) or (is_composed_of or is_formed_of)) or (has or includes)) or (encompasses or involves)) or contains) and ((is_prerequisite_for or fundaments) or (is base_for or is_condition_for))).((generates or creates) or (develops or produces)) project|LOs@lom/general/title

iii) Executing queries

One of the critical issues analyzed in this section refers to MEC ROSA. As already commented in the beginning of this section, it can only submit queries to a single database, meaning that an optimized distributed query plan, such as the one developed in (Nejdl et al., 2002), cannot be generated. Therefore, the solution adopted in this work considers that each peer processes a query similarly as the requestor peer, having the same autonomy over it. Hence, each peer will be able to rewrite the query and to manage its own processing, overcoming the lack of a distributed plan. The processing strategy considers two phases, as described next:

- **First phase:** consists in identifying, at query submission or reception, the relevant peers and/or super-peers able to answer it. Thus, from the point of view of the query submission made by a user, the query processing will analyze if the peer is a super-peer or not. In affirmative case, it will verify if it is able to answer the query. In positive case, it will rewrite the query to itself, storing the result in cache. Then it will resend the query to its own peers and will locate, among the super-peers to which it refers to, those that are relevant, in order to send it to them. Otherwise, when the peer is not a super-peer, the query processing will simply send the query directly to the corresponding super-peer, which will be responsible for resending the query in the system. However, from the point of view of the query reception made by a peer or super-peer, the query processing will also analyze if the peer is a super-peer or not. If it is the case, it will verify if the super-peer that originally sent the query belongs to its grouping. If it is true, it will not be necessary to resend the query to the other grouping super-peers, since this super-peer will indeed do this. However, it is necessary the super-peer resends the query to its own peers. In case the super-peer does not belong to its subject grouping, it needs to resend the query to its super-peers grouping and to its own peers. In both cases, at the moment when a query is resent by a peer or super-peer, two timers are activated: one will estimate the time limit the system will wait for the corresponding partial results; and the other refers to the expected number of results, according to the quantity of peers or super-peers to which the query has been sent to (Brito, 2005);
- **Second phase:** consists in rewriting and processing the query, returning the result to the requestor peer or super-peer. Thus, after submitting or receiving a query, the corresponding peer or super-peer will be responsible for rewriting it. Once rewritten, the query is processed by MEC ROSA, and its results, from the query submission point of view, will be kept in the respective peer or super-peer cache. From the query reception point of view (made by a peer or super-peer), results are sent to the requestor peer (or super-peer), also remaining in cache for future integration. These results will remain there until the time limit defined by the query processing is over, or until all results have been returned; afterwards results will be integrated.

4.3.3 Data Integration

It consists in integrating all cache query results returned from relevant peers and/or super-peers, making it possible a correct global answer to be returned to the user. As each peer and/or super-peer has to identify, rewrite and process a submitted query, the resolution of existing consistency problems in this phase is the responsibility of each peer, and not only of the query requestor peer. Hence, each of the partial results returned to the requestor peer (or super-peer) is already free of any inconsistence, becoming ready for the system to process their union. In fact, this union is equivalent to the query result integration.

5 RELATED WORK

In the literature there are many works related to P2P systems, each one defined according to its specific characteristics and requirements. *Peer*DB, Hyperion, Piazza, SeLeNe and Edutella are one of these systems.

*Peer*DB (Ooi et al. 2003) provides content based queries, mobile agents integration, and a schema mapping strategy based on descriptive words, which are the only available metadata information. The focus on the Hyperion (Arenas et al., 2003) project resides on the specification and management of metadata in order to enable data coordination and sharing between peers. Similarly to the others, the Piazza (Tatarinov et al., 2003) project defines data sharing between peers from its mapping schemas, besides having a specific results system that recursively expands any relevant mapping for a query, retrieving important data from peers. The SeLeNe (Self e-Learning Networks) project (Keenoy et al. 2004) is based on a GRID service architecture supported by metadata, which provides facilities for discovering and sharing e-learning resources. Edutella system (Nejdl et al. 2002) is characterized for providing a multiple platform to extend, specify and implement a metadata infrastructure in RDF for P2P network, which provides the sharing didactic resources among institutions.

In the context of ROSA - P2P, Edutella is the one that deserves more attention since, to the best of our knowledge, it is the only architecture based on super-peers. Hence, some of its characteristics were essential to define important issues in ROSA - P2P, such as routing indices and peers grouping. Furthermore, the main difference between ROSA - P2P and the other systems concerns the type of distributed architecture used to provide interoperability and data sharing on the Web. While the others use schema mapping tables, P2P, and GRID, ROSA - P2P profits from the computational capacity of P2P architecture based on super-peers, and a complex ontology based structure to process and integrate queries results. Nevertheless, ROSA - P2P presents some features not explored so far in the other projects, such as a specific strategy for grouping super-peers based on a subject domain, and instance integration instead of schemas, since ROSA does not provide a conceptual schema.

6 CONCLUSION

This paper presented the evolution of ROSA system that has been transformed from local system into a distributed one, named ROSA - P2P, now able to provide global answers to user's queries.

Throughout this paper two aspects have been carefully emphasized: the P2P environment and data integration, which specify among other characteristics: details about this environment, responsible for an adequate interoperability; controlled vocabularies to support semantic conflicts, broadcast and queries rewriting; a query processing strategy based on queries contents; and the strategy used to integrate ROSA - P2P LOs.

The system has been validated with intensive tests presenting results in satisfactory time, taking into account the number of peers and the domain knowledge diversity they encompassed (Brito, 2005). As future work, we intend to: continue the system evaluation work, using a more robust platform; perform communication protocols simulation using different network topologies and a larger number of peers to ensure the real system stability; and enrich the system with a query distributed optimized plan, which will provide more autonomy in query processing.

REFERENCES

Arenas, M. et al., 2003. The Hyperion Project: From Data Integration to Data Coordination. *ACM SIGMOD R.*

Brito, G. A. D. D. 2005. Learning Objects Integration in ROSA-P2P System (in Portuguese). Master Thesis, IME, May.

Brito, G. A. D. D. and Moura, A. M. de C. 2005. ROSA - P2P: a Peer-to-Peer System for Learning Objects Integration on the Web. *In Proceedings of 11th Brazilian Symposium of Multimedia System and Web, Poços de Caldas, MG.*, Brazil, December.

Coutinho, F. and Porto F., 2004. Query Processing in ROSA. *In Proceedings of 19th Brazilian Symposium on Databases.* Brazil, October.

Keenoy, K., Poulovassilis, A., et al., 2004 Personalization services for self *e-learning* networks. *In Proceedings of International Conference on Web Engineering.*

Nejdl, W. et al., 2002. Edutella: A P2P Networking Infrastructure Based on RDF. *In Proceedings of 11th International WWW Conference.*

Nejdl, W. et al., 2003. Super-Peer-Based Routing and Clustering Strategies for RDF-Based Peer-to-Peer Network. *In Proceedings of 12th International WWW Conference.*

Nguyen, T. and Sanchez, E., 2004. A distance e-Learning System in a P2P Network. Intl. Conf. RIVF'04.

Ooi, B., Shu, Y. and Tan, K., 2003. Relational Data Sharing in Peer-based Data Management Systems. *ACM SIGMOD Record, 32(3).*

Porto, F., Moura, A. M. and Coutinho, F., 2004. ROSA: a Repository of Objects with Semantic Access for e-Learning. *In Proceedings of 8th International Database Engineering & Applications Symposium*. Coimbra, Portugal, July.

Tatarinov, I. et al., 2003. The Piazza Peer Data Management Project. *ACM Record 32 (2003).*

Toledo, A., 2002. *Corporative Portals: A Strategic Tool for Supporting Knowledge Management (in Portuguese).* Master Thesis, UFRJ - Brazil.

Zhu, Y., Wang, H. and Hu, Y., 2003. Super-peer Based Lookup in Structured Peer-to-Peer Systems. *In Proceedings of 16th International Conference on Parallel and Distributed Computing Systems*. Nevada.