SeLeNe — Self E-Learning Networks
IST-2001-39045

WP1 Deliverable 7

Report on the IST Workshop on Metadata Management in GRID and P2P Systems, and Key Transferable Outcomes of SeLeNe

The SeLeNe Consortium

Abstract

This deliverable reports on the *IST Workshop on Metadata Management in GRID and P2P Systems* that was held in London on 16th December 2003 as part of the SeLeNe project's dissemination and integration activities. It also discusses the main transferable outcomes of the SeLeNe project, and how these can be exploited by interested parties.

June 8, 2004

The SeLeNe Project

SeLeNe (Self e-Learning Networks) is an Accompanying Measure in the IST Action Line "Information and Knowledge Grids". The SeLeNe project is conducting a feasibility study into using semantic web technology for syndicating knowledge-intensive resources, such as learning objects. It is developing services for the discovery, sharing, and collaborative creation of learning resources, thus facilitating a syndicated and personalised access to such resources.

Executive Summary

This is a report on the *IST Workshop on Metadata Management in GRID and P2P Systems* that took place on 16th December 2003 in London.

The report begins by introducing the aims of the Workshop, then gives a detailed account of the proceedings of each session and ends with overall conclusions from the Workshop.

Also included is a section which highlights the main transferable outcomes of the SeLeNe project, and how these can be exploited by interested parties.

The appendices contain statistics on the speakers and participants, and the results of Workshop evaluation questionnaires completed by attendees.

Revision Information

Revision Date	Version	Changes
January 27th, 2004	0.1	First Draft Proposal
January 31st, 2004	1.0	First Submitted Version
June 8th, 2004	2.0	Revised Version following the Final Review Meeting

Contents

1 Introduction **4**
1.1 Workshop Programme 4

2 Summary of Workshop Sessions **6**
2.1 Session I - Metadata Management in Grid systems 6
2.2 Session II - Metadata Management in P2P systems 7
2.3 Session III - Applications I 8
2.4 Session IV - Applications II 10

3 General Conclusions **13**
3.1 Forthcoming 2nd MMGPS Workshop 13

4 Key Transferable Outcomes of SeLeNe **14**

References **16**

Appendices **20**

A Participation Statistics **20**
A.1 Number of Attendees 20
A.2 IST Participation 20
A.3 Breadth of European Representation 20
A.4 Presentation of IST results 21

B Responses to the Evaluation Questionnaires **21**
B.1 Morning Sessions 22
B.2 Afternoon Sessions 23
B.3 Workshop as a Whole 24
B.4 Comments and Suggestions 24

C Evaluation Questionnaires **25**
C.1 Morning 25
C.2 Afternoon 26

D List of Participants **27**

1 Introduction

As part of the dissemination of the SeLeNe project, and also as part of the integration role of the project as an Accompanying Measure, a Workshop was held in London on December 16th 2003. The target audience for the event were experts from the GRID, peer-to-peer and e-learning communities, as well as other application areas requiring GRID or peer-to-peer support.

The declared goal of the Workshop was "to identify recent technological achievements and open challenges regarding metadata management in novel applications requiring peer-to-peer information management in a distributed or GRID setting". The discussions in all sessions focussed on models, services and architectures from the perspectives of the metadata level in general and for specific application domains.

The event attracted 43 participants (including those from within the SeLeNe consortium) from 8 different European countries, and we believe that it was an important step in coordinating research activities in this area. In order to disseminate the findings of the Workshop as widely as possible, participants' presentations have been made publicly available via the Workshop website:

`http://www.ics.forth.gr/isl/ist_workshop/`

and also published on-line in a repository for Workshop proceedings at:

`http:CEUR-WS.org`

1.1 Workshop Programme

The Workshop programme was as follows. Details of the proceedings of each session can be found in Section 2.

Session I: Metadata Management in Grid systems
Chair: George Samaras (University of Cyprus)
1. Gavin McCance (CERN)
"Metadata Management in the EU DataGrid"
2. Savas Parastatidis (University of Newcastle Upon Tyne)
"Working with Grid Service and Resource Metadata Using Existing Tools and Specifications"
3. Mario Cannataro (University "Magna Græcia" of Catanzaro, Italy)
"Architecture, Metadata and Ontologies in the Knowledge Grid"

Session II: Metadata Management in P2P systems
Chair: George Loizou (Birkbeck)
1. Henrik Nottelmann (University Duisburg-Essen)
"Probabilistic Logics for Defining and Using P2P Service Descriptions"
2. Philippe Cudré-Mauroux (EPFL Distributed Information Systems Laboratory)
"The Chatty Web Approach for Global Semantic Agreements"
3. Wolfgang Nejdl (KBS/L3S)
"Data-Centric Networks and Peer-to-Peer Databases"

Session III: Applications I
Chair: Nicolas Spyratos (LRI, University de Paris-Sud)
1. Tamás Hauer (CERN)
"The Role of MetaData in Querying Grid-Resident Medical Images"
2. Bob Bentley (UCL)
"Management of Metadata for the Virtual Solar Observatory. Experiences from EGSO"
3. Theo Dimitrakos (CCLRC)
"Meta-data Management Issues Underpinning Emerging Solutions for Distributed Trust and Contract Management and Enforcement in Enterprise Grid and P2P Systems. Experiences from the GRASP, SWAD-Europe and CORAS projects"
4. Stavros Christodoulakis (MUSIC/TUC)
"Multimedia Metadata Management and t-Learning Applications"

Session IV: Applications II
Chair: Alex Poulovassilis (Birkbeck)
1. Klaus Jantke (DFKI)
"DaMiT - Peculiartities of an e-Learning System"
2. Kevin Keenoy and Vassilis Christophides (Birkbeck and ICS-FORTH)
"Personalisation in Self e-Learning Networks"
3. David Massart (European Schoolnet)
"Metadata Management in the Celebrate European Learning Network"
4. Zoltán Miklós (Vienna University of Economics)
"ELENA: Creating a Smart Space for Learning"

2 Summary of Workshop Sessions

2.1 Session I - Metadata Management in Grid systems

The first session of the Workshop addressed metadata management issues in Grid systems. The session contained three presentations on related projects and issues. The presentations complement each other in that they present mutually exclusive issues; what metadata we have in real Grid systems, how we can provide Grid services using existing technology and how metadata can be used in new Grid domains. In all three presentations the need for metadata management middleware was quite evident.

The first presentation, given by Gavin McCance of the University of Glasgow, presented how DataGrid, the European flagship Grid project, addresses issues of metadata management. DataGrid is a file-based system manipulating quantum-physical, Earth-observation and biomedical data. The DataGrid project utilises two types of metadata: (i) Grid internal metadata on files and application logical names, and (ii) general application-specific metadata. Internal metadata have a relatively simple structure and their main use is for indexing purposes. The internal metadata is used to solve the "replica location problem", answering the query "given a logical file identifier, how do we find all the replicas of that file on the Grid?". The system that has been built is composed of a Replica Metadata Catalog and a Replica Location Service. It is based on bloom filters and allows quite a flexible search facility. DataGrid provides a simple Grid-enabled front end, it is up to the end-user (from one of the supported domains — physics, Earth-observation, biomedicine) to provide the specific application and to request the relevant files to manipulate. There is no generic application-metadata-management middleware that can handle generic application-type metadata.

The second talk was given by Savas Parastatidis of North East e-Science Centre and focussed on how Grid services can access an organisation's resources. He identified the need for resource metadata that can be published outside the organisation's boundaries and argues that this can be achieved via existing specifications and tools without the need to change the existing infrastructure. He presented two possible solutions, one based on Service Data Elements (SDEs), where metadata about resources is exposed through SDEs, and one based on the Grid Resource Metadata document (GRM). The claim made is that when using SDEs a number of changes need to take place: WSDL and resource descriptions are affected, new SDE-aware tools are required, Web services need added semantics, and an SDE-specific interface is necessary. However, if we utilise the concepts in GRM no changes need to take place. This is so because GRM is based on XML-Schema and no additions to WSDL or any other specifications are needed, existing tools work, it does not add semantics to a Web service (as it is just a document) and it could be published into a registry. However, it requires a particular Web service interface (a management interface or another interface) but other solutions can be built around it (e.g. peer-to-peer, registries, etc.). The main conclusion is that the GRM document provides a functional equivalent to SDEs that allows the re-use of existing technologies and work, which is not the case for SDEs.

The third talk was given by Mario Cannataro of the University "Magna Græcia" of Catanzaro. The "Knowledge Grid" was presented as a software environment that integrates data-mining techniques and Grid resources to build Grid-aware data-mining applications. The clear claim was that metadata (for describing data mining tools, data sources, Grid resources and ontologies for semantic modelling of the application domain) is a must if the Grid is to be utilised in knowledge management and data mining. The need for two types of metadata was exposed, the first type of metadata being for data-mining software and the second type for classical data-source metadata. The first type of metadata is semantically related to data mining and is split into two parts: (a) a description part, that is used to classify software and (b) a usage part, that contains the information needed by clients to access and use the software. In the system presented, metadata is utilised extensively to define execution plans and other processes. The presentation then showed how metadata and ontologies are managed to build and execute distributed data mining applications on the Knowledge Grid. It finally argues that metadata and ontologies are major players when complex applications are developed in a Grid environment.

2.2 Session II - Metadata Management in P2P systems

The second session of the Workshop addressed metadata management issues in peer-to-peer systems. This session also contained three presentations on related projects and systems.

The first talk, given by Henrik Nottelmann (presenting work carried out jointly with Norbert Fuhr) of the University of Duisburg, focussed on the use of probabilistic logics for defining and using service descriptions in a peer-to-peer network with a large number of (Web) services. The goal of this work is to dynamically compute execution plans for services required to implement a given task. In this respect, the DAML-S upper ontology (a vocabulary for defining arbitrary services) is used for describing services, while probabilistic Datalog is used for match-making. A service is described in DAML+OIL by its profile, which is used for match-making, by its model, which describes how the service works, and by its grounding, which describes how to access the given service. A lower ontology for library services (e.g. search, schema mapping, or result modification services) is also introduced, containing definitions for generic search services and other query- and result-transformation services. Match-making rules were defined with probabilistic Datalog, which is a variant of predicate logic based on function-free Horn clauses and which defines probabilistic facts and rules. Match-making rules use facts derived from DAML-S directly to provide proper execution plans of services. By using the notion of chain (S1, S, S2) — i.e. a chain beginning with S1, ending with S2 and with S in between — two or more services with correct input/output can be chained to produce an execution plan. The probabilities are used as a primitive kind of cost estimation. The weight specifies the quality of the execution plan, so if its value is smaller than 0.5 then we have a quality loss, otherwise we have a quality improvement. The creation of a detailed lower ontology for library services, the introduction of composite services and the implementation of the project including grounding services are left as future work.

The second talk, given by Philippe Cudré-Mauroux (presenting work carried out jointly with Karl Aberer) of the EPFL and the Distributed Information Systems Laboratory, presented the Chatty Web approach for global semantic agreements. The problem posed is that of achieving semantic interoperability among heterogeneous data sources in a peer-to-peer data management system, without relying on pre-existing global semantic models. The solution given involves the use of local translations that enable global agreements. Semantic 'gossiping' is used for query forwarding and distribution through the system. The right peers are selected and the "Per-Hop Behaviours" are query-dependent. An analysis is done between original and transformed queries based on intrinsic (syntactic distances) and extrinsic (semantic distances) measures. For the semantic similarity, query cycles can be detected and analysis of the results with content-retrieval techniques can be used. Consequently, a self-repairing semantic network can be organised with the use of evaluations based on Chatty Web simulations and the automatic correction of erroneous mappings based on gathered evidence. Finally, results concerning the system's scalability and sensitivity to TTL were presented and discussed.

The third talk, given by Wolfgang Nejdl of the University of Hannover and the Learning Lab, Lower Saxony (L3S), addressed issues related to data-centric networks and peer-to-peer data management. Networks have evolved from host-centric to data-centric ones. In this context, the focus of research has shifted from the evaluation and optimisation of communication between network hosts to achieving worldwide physical and network data independence. Professor Nejdl presented the main results and challenges of several ongoing and past projects. REWERSE (Reasoning on the Web with Rules and Semantics) is an EU project addressing the issues of retrieving, protecting and integrating data. EDUTELLA has specified and implemented an RDF-based metadata infrastructure for peer-to-peer data networks. PROLEARN is a project that works towards innovative and interoperable e-learning resources and sustainable e-learning infrastructures and processes. The notion of schema and the use of RDF/S for describing distributed resources were identified in these projects as useful parts of a peer-to-peer data management system. Furthermore, RDF-QEL was presented as a Datalog-based Query Exchange Language that can be used to wrap other RDF and XML query languages. The HyperCup peer-to-peer topology and its broadcast algorithm were shown as an answer to the efficient routing problem in a peer-to-peer system. Super-peer networks and their routing indices offer a new network architecture for distributing workload in a more efficient manner. Access control and automated trust negotiation were also identified as an important problem in order to protect resources from unauthorised access using credentials and access control policies. All the above issues represent essential building blocks for forthcoming schema-based peer-to-peer networks and peer-to-peer-based data management infrastructures.

2.3 Session III - Applications I

The third session of the Workshop consisted of four presentations, each concerning meta-data and its use.

The first talk, from Tamás Hauer, concerned the role of metadata in querying Grid-resident medical images in the MammoGrid project, a EU-funded project aiming at a Grid solution for mammography and involving three hospitals among its partners. A federated system solution is proposed, whereby, with the help of shared metadata, a clinician can address a query to the system, which is then translated to remote local sub-queries whose results are then returned to the clinician. The specific characteristics of the medical domain (e.g. a user community which is heterogeneous and process-oriented rather than information-oriented, regional differences and a constantly changing knowledge base) necessitate flexibility and extensibility as well as management of domain information such as annotations of images. These and other considerations have led to a service-oriented architecture using Grid-middleware. In this context, metadata is used to describe service definitions. Each participating node is responsible for managing its own metadata and can change its service description on the fly, new sites can join seamlessly, and domain and service ontologies are defined independently.

The second presentation, from Bob Bentley, concerned metadata management in virtual solar observatories and reported experiences from the EGSO project, a Grid test-bed designed to improve access to solar data for the solar physics and other communities. The EGSO project addresses the generic problem of a distributed heterogeneous data set and a scattered user community. In such a setting, the availability of good quality metadata is important for searching. The architecture is defined in terms of roles: consumer, broker and provider. Resources are described by entries in a resource registry and managed by a broker. Brokers and registries are replicated to provide system resilience and permit load sharing. In order to provide an enhanced search capability, EGSO will improve the quality and availability of metadata, having a search registry describe all metadata available for search.

The third presentation, from Theo Dimitrakos, concerned metadata management issues underpinning emerging solutions for distributed trust and contract management and enforcement in enterprise Grid and peer-to-peer systems. Metadata management issues underlie a number of activities across e-science and information technology. In support of this claim, the speaker presented in some detail four projects where metadata play an important role:

- GRASP, Grid-based application service provision;
- CORAS, a CASE tool and method support for security-risk analysis;
- SWAD-Europe, Semantic Web technology development;
- PELLUCID, an agent-based platform supporting organisational mobility.

The final presentation, from Stavros Christodoulakis, also concerned metadata management, this time for audiovisual content to support intelligent video-content retrieval and e-learning services in digital TV (t-learning). These objectives are pursued in the context of TV-Anytime framework, MPEG-7 and SCORM. An ontology-driven framework for semantic metadata management was presented and the following points were stressed:

- TV-Anytime keywords are the only means to describe program segments;
- MPEG-7 semantic model is used to build domain specific ontologies;

- a coupling of OWL and MPEG-7 has been implemented;
- ontologies are used for filtering and retrieval of MPEG-7 multimedia content;
- semantic annotations are transformed into TV-Anytime segment keywords.

With respect to t-learning, two issues were discussed:
(1) providing interoperability for educational applications in different e-learning and digital TV environments;
(2) creation of metadata for digital TV for educational purposes in order to offer educational experiences exploiting usual TV programs.

2.4 Session IV - Applications II

The fourth session of the Workshop consisted of four presentations, all on e-learning applications that rely on some form of schema-based repository of Learning Object (LO) metadata.

The first talk, from Klaus Jantke, gave an outline of the DaMiT e-learning system developed to provide a tutorial on data mining, and went on to discuss particular problems posed by the development of the system and aspects that may distinguish it from other projects.

The system allows users to specify several of their preferences as a "profile", including the type of material preferred (example-oriented, theory-oriented, etc.), style (formal, informal) and language (although the system is mostly German-only at present). The e-learning content delivered by the system is stored as a collection of fragments of learning material, each described by associated metadata. DaMiT generates complete pages of learning material from these fragments "on the fly", adapting the page to the user's needs based on their profile. DaMiT also includes a fully-functional e-payment system that uses metadata about the user to keep track of their entitlements to access to the learning material (students at DFKI can access all material for free, but others must pay to use the resource).

Professor Jantke believes that future e-learning systems will include intelligent assistants that learn about, understand and adapt to the user, to provide them with a personalised learning experience. Metadata of several kinds will be needed to support this adaptive functionality, including the kinds used in DaMiT:
(1) metadata about each fragment of content;
(2) "storyboards" that aid in the structuring of content fragments;
(3) metadata on security and payment options for users.

The second presentation in the session was by Kevin Keenoy on the SeLeNe project. SeLeNe aims to provide techniques enabling the discovery, sharing and collaborative creation of LOs within a learning community. LOs to be shared within the community are described using an extended form of the IEEE-LOM metadata schema, encoded in RDF, and these descriptions together form a distributed repository of metadata that can be queried by users searching for learning material. The metadata entries make use of taxonomies of subject and topic domains, learning objectives and learning styles.

As well as metadata describing the LOs (LO descriptions), the system manages metadata describing the users of the system (user profiles), to enable adaptive personalisation services to be provided. The user profiles are also encoded using RDF, and make use of the same taxonomies as the LO descriptions. Both the LO descriptions and user profiles adapt and evolve over time, either automatically (e.g. the user profile records the history of LO accesses by the user) or as users add and amend the data manually (e.g. a LO author may add additional information to one of her LO descriptions).

There are three main areas of personalisation in SeLeNe:
(1) Personalised *views* of the LO descriptions and schemas;
(2) Generation of personalised *query results* based on the user profile;
(3) Personalised event and change *notification* services by means of event-condition-action (ECA) rules, allowing individuals to be informed of changes to the metadata repository of interest to them.

Several issues were identified as areas for future research, including peer-to-peer query processing over RDF, the evaluation of algorithms for personalised ranking of query results, the combination of ECA rules with transaction and consistency maintenance in RDF repositories, and the design of user interfaces enabling easy access to such advanced personalisation services.

The third presentation, from David Massart (of work carried out jointly with Frans Van Assche), described work done by the Celebrate (Context e-Learning with Broadband Technologies) project in developing a system providing metadata search and exchange of LOs, creating a European Learning Network (ELN) that is used by 500 schools across Europe.

The ELN is built around a brokerage system that manages exchanges between its members, enabling interoperability between e-learning systems by searching and exchanging LOs in their repositories. The IEEE-LOM schema is used for LO descriptions. Various communication protocols may be used between different ELN members, but the individual systems are shielded from this complexity by the "ELN client" that provides a simple communication API.

Unlike other virtual learning environment networks based either on a client-server or on a peer-to-peer architecture, Celebrate is based on a mixed approach where, although the brokerage system hosts a central metadata repository, each ELN member is also authorised to manage its own local metadata repository. Search requests are both handled centrally and propagated to local repositories.

The final presentation of the session, from Zoltán Miklós (of work carried out jointly with Bernd Simon), described work done in the ELENA project, "creating a smart space for learning".

ELENA, like both SeLeNe and Celebrate, is concerned with the sharing of learning resources that are contained in heterogeneous, distributed repositories. In Smart Spaces for Learning metadata repositories are connected in a peer-to-peer fashion, and Semantic Web techniques are used to achieve interoperability. Personal Learning Assistants support learners in selecting appropriate learning services from a large number of available sources.

The metadata for learning services is based on a common formal ontology, but learners do not need to be familiar with the domain ontology in detail as they can define views of the ontology using the TRIPLE language. These views allow queries to be formulated in terms of a user-specific ontology. A Simple Query Interface translates queries into an appropriate form, establishing interoperability among heterogeneous learning repositories.

The presentations in this session brought to light several core issues that arise for systems providing personalised e-learning. In systems that employ a user profile two main issues are *content* and *representation*: what information about the user is needed, and what format should it be stored in? There is also the issue of where the information for the profile comes from — is it user-supplied, teacher-supplied, automatically derived somehow, or some combination of all these three?

Once answers to questions about the generation and maintenance of the profile have been decided, the remaining issue (and perhaps the issue that has so-far been least explored) is the resolution of the profile with LO metadata — how can the profile be used to provide the most useful personalised LOs or personalised ranking? It seems that there is scope for much more work in empirically testing different matching algorithms in different learning situations — present methods generally seem to be "sensible" ad-hoc choices rather than empirically validated solutions.

Another area requiring future investigation is learning style taxonomies — which ones are most useful in the context of e-learning, and what the relationships are between different learning style taxonomies.

3 General Conclusions

From the above summaries of the four Workshop sessions, we see three sets of research challenges emerging in the area of metadata management for Grid and peer-to-peer systems:

1. Challenges stemming from the *distribution*, *autonomy* and *heterogeneity* of information and services, leading to the need for:

    - metadata describing information and services available at the nodes of the Grid/P2P system;
    - searching and matching techniques utilising this metadata for discovery of information and services;
    - metadata and techniques for controlling access to, and privacy of, information and services;
    - techniques for automatic service composition and orchestration;
    - metadata and techniques for translation of queries and query results in the absence of a controlled global schema;
    - metadata supporting replication and consistency of information;
    - automatic/semi-automatic extraction of metadata from large volumes of heterogeneous data.

2. Challenges arising from the *dynamicity* of Grid/P2P environments, e.g. changes in the network topology, information content at nodes, and service availability at nodes:

    - this leads to the need for the techniques being developed under (1) to be scalable, adaptive, extensible and fault tolerant.

3. Challenges arising from the *heterogeneity of users* accessing Grid/P2P systems:

    - this leads to the need for techniques that personalise content and presentation to different users' needs and preferences.

3.1 Forthcoming 2nd MMGPS Workshop

During the discussion at the conclusion of the Workshop it was generally agreed that the Workshop had been very beneficial and timely in bringing together these various common strands of research from the Grid, P2P and applications communities, and that a similar event in a year's time should be aimed for. We are therefore pleased to announce the *2nd Workshop on Metadata Management in GRID and P2P Systems* to be held on 17th December 2004, Senate House, University of London. Details of the programme committee and submissions procedure will be announced shortly.

4 Key Transferable Outcomes of SeLeNe

The major achievements of the SeLeNe project are as follows:

- a review of existing e-learning standards and the feasibility of expressing them using Semantic Web languages such as RDF/S, including identification of the pedagogical limitations of existing standards (see [SCKM03]);
- the specification of the functionality of a *Self e-Learning Network* (SeLeNe) — a system that exploits Semantic Web technology to provide learning communities with personalised and syndicated access to learning resources (see [KPP+03]);
- the identification of *composite learning objects* as a new paradigm for collaborative construction of educational material (see [KPP+03, RS03c]);
- an algorithm for *automatic generation of taxonomical descriptions of composite learning objects* (see [RS03c]);
- the specification of a *User Profile* for SeLeNe users, along with RDF schemas defining the profile structure. The profile combines elements from existing learner profile schemes and adds extra elements where these schemes are insufficiently expressive to adequately support SeLeNe's personalisation requirements (see [KLP03]);
- the specification of a *GRID service based architecture* supporting the SeLeNe functionality, together with three possible concrete service deployment scenarios each of which supports a different kind of learning community (see [SKC03, SeL04]);
- development of the *RVL view definition language*, which allows personalised views of learning object descriptions and schemas to be defined by the users of a SeLeNe (see [MCP03]);
- development of *XML and RDF event-condition-action rule languages*, which support SeLeNe's event notification and change detection/propagation services (see [PPW03b]).

The outputs of the project include a set of reports [SCKM03, KPP+03, SKC03, RS03c, KLP03, MCP03, PPW03b, SeL04], all of which are available from the project website at `www.dcs.bbk.ac.uk/selene`, where an introduction and summary of each report are given as well as the downloadable pdf file. All of the major outputs of the project have also been presented in numerous presentations at conferences and workshops [PPW03a, ACK+01, MTCP03, TASC03, RS03b, TCK+03, CKM+03, SKPC, KMA+03, RS03a, SKPC, GRS04a, BPPT04, KC04, GRS04b, MTCP04, KPC+04, PL03, PPW04, MP03] (see `www.dcs.bbk.ac.uk/selene/reports`).

The key transferable outcomes of the SeLeNe project are as follows:

- The survey of e-learning standards and discussion of the feasibility of expressing them using Semantic Web languages [SCKM03] is a useful resource for all other researchers in this area. The Resources section of SeLeNe's site `www.dcs.bbk.ac.uk/selene/` also contains an extensive collection of links to referenced work and related projects.

- The RDF/S bindings of the IEEE LOM e-learning standard and the ACM CSS taxonomy available from FORTH[1] are a useful resource for system developers wishing to employ these standards in semantic web applications.
- The requirements analysis and specification of the functionality of a Self e-Learning Network [KPP+03] can be used as the basis for the design of other e-learning systems, either as-is or incorporated into a modified requirements definition.
- The Grid service based architecture for SeLeNe specified and validated in [SKC03, SeL04] gives a possible architecture of services that can provide this functionality. This architecture can be re-used as-is or, alternatively, some parts of it can be incorporated into other systems.
- The upper layers of the architecture are specific to SeLeNe's e-learning application domain, while the lower layers are generically applicable to any domain which requires personalised and syndicated access to distributed repositories of RDF/S descriptions. The key services in this architecture are individually discussed in [RS03c, KLP03, MCP03, PPW03b].
- The three concrete deployment scenarios of the high-level architecture discussed in [SeL04] give examples of how this set of services can be used to support different kinds of learning communities.
- These three service placement alternatives also illustrate the feasibility of supporting a P2P, mediation-based and centralised systems with the same service architecture.
- The User Profile specification [KLP03] can be used as the basis for any RDF-based user profile, and the strategy of selecting only some elements from existing schemas can be used to extend or reduce the information held in such profiles as necessary. This will be useful not just for systems that manage learning material, but in any system providing personalised access to the Semantic Web.
- The paradigm of composite learning objects [KPP+03, RS03c] can be used in any system supporting the collaborative construction of educational material, or in fact the collaborative construction of any kind of documents. If such a system classifies the documents so produced, then the burden of classification will be much reduced by the application of SeLeNe's algorithm for automatic generation of taxonomical descriptions of composite learning objects [RS03c]. An experimental implementation of this algorithm, as used in the case study reported in [RS03c], is freely available from LRI[2].
- The RVL view definition language [MCP03] and the experimental XML and RDF event-condition-action rule languages [PPW03b] are fully specified, and so can be implemented for exploitation by anyone wishing to do so.

[1] `http://139.91.183.30:9090/RDF/Examples.html`

[2] Nicolas Spyratos (`Nicolas.Spyratos@lri.fr`) and Philippe Rigaux (`Philippe.Rigaux@lri.fr`)

- A prototype implementation of RVL is available upon request from FORTH[3] and an experimental implementation of the XML ECA language from Birkbeck[4].

- Several tools for the manipulation of RDF/S, including the RDF/S Semantic Web Portal Generator (SWPG), can be accessed from FORTH's RDFSuite website[5].

- The identification of issues relating to Grid metadata management and the proposal of an initial technique for the integration of SeLeNe services with the OGSA Information Service [SeL04] can inform future work on metadata management and services on the Grid.

- Finally, the discussion of routing and processing of RVL queries and RDF ECA rules in P2P architectures [MCP03, PPW03b] can inform future work on metadata management in P2P systems.

References

[ACK+01] S. Alexaki, V. Christophides, G. Karvounarakis, D. Plexousakis, and K. Tolle. The ICS-FORTH RDFSuite: Managing Voluminous RDF Description Bases. In *Proceedings of the 2nd International Workshop on the Semantic Web (SemWeb 2001)*, 2001.

[BPPT04] J. Bailey, G. Papamarkos, A. Poulovassilis, and P. T.Wood. An event-condition-action language for XML. In *Web Dynamics*. Springer, 2004.

[CKM+03] V. Christophides, G. Karvounarakis, A. Magkanaraki, D. Plexousakis, and V. Tannen. The ICS-FORTH semantic web integration middleware (SWIM). *IEEE Data Engineering Bulletin*, 26(4):11–18, 2003.

[GRS04a] B. Gueye, P. Rigaux, and N. Spyratos. Annotation automatique de documents XML. In *Proc. Nat. Conf. Extraction et gestion des connaissances (EGC'04)*, Clermont-Ferrand, France, Jan. 2004.

[GRS04b] B. Gueye, P. Rigaux, and N. Spyratos. Taxonomy-based annotation of xml documents: Application to e-learning resources. In *Proceedings of the Hellenic Conference on Artificial Intelligence (SETN'04)*, Samos, Greece, May 2004.

[3]Vassilis Christophides (`christop@ics.forth.gr`)

[4]George Papamarkos (`gpapa05@dcs.bbk.ac.uk`), Alex Poulovassilis (`ap@dcs.bbk.ac.uk`) and Peter Wood (`ptw@dcs.bbk.ac.uk`)

[5]`http://139.91.183.30:9090/RDF/`

[KC04] G. Kokkinidis and V. Christophides. Semantic query routing and processing in P2P database systems: The ICS-FORTH SQPeer middleware. In *Proceedings of the International Workshop on Peer-to-Peer Computing & DataBases (P2P&DB), in conjunction with EDBT 2004 Conference*, Heraklion – Crete, Greece, March 14 2004. `http://pc-erato2.iei.pi.cnr.it/meghini/p2p/`.

[KLP03] K. Keenoy, M. Levene, and D. Peterson. SeLeNe Deliverable 4.2: Personalisation and trails in Self e-Learning Networks, 2003. Available from `http://www.dcs.bbk.ac.uk/selene/reports/`.

[KMA+03] G. Karvounarakis, A. Magkanaraki, S. Alexaki, V. Christophides, D. Plexousakis, M. Scholl, and K. Tolle. RQL: A functional query language for RDF. In *The Functional Approach to Data Management : Modelling, Analysing and Integrating Heterogeneous Data*. Springer, 2003.

[KPC+04] K. Keenoy, A. Poulovassilis, V. Christophides, P. Rigaux, G. Papamarkos, A. Magkanaraki, M. Stratakis, N. Spyratos, and P. T. Wood. Personalisation services for Self e-Learning Networks. In *Proceedings of the International Conference on Web Engineering (ICWE'04)*, Munich, Germany, July 2004. Springer LNCS. To appear.

[KPP+03] K. Keenoy, G. Papamarkos, A. Poulovassilis, M. Levene, D. Peterson, P.T. Wood, and G. Loizou. SeLeNe Deliverable 2.2: Self e-Learning Networks – functionality, user requirements and exploitation scenarios, 2003. Available from `http://www.dcs.bbk.ac.uk/selene/reports/`.

[MCP03] A. Magkanaraki, V. Christophides, and D. Plexousakis. SeLeNe Deliverable 4.3: Views and structured querying in Self e-Learning Networks, 2003. Available from `http://www.dcs.bbk.ac.uk/selene/reports/`.

[MP03] P. McBrien and A. Poulovassilis. Defining peer-to-peer data integration using both as view rules. In *Proc. Workshop on Databases, Information Systems and Peer-to-Peer Computing, VLDB'03*, volume 2944 of *LNCS*, Berlin, Sept. 2003. Springer.

[MTCP03] A. Magkanaraki, V. Tannen, V. Christophides, and D. Plexousakis. Viewing the semantic web through RVL lenses. In *Proceedings of the Second International Semantic Web Conference (ISWC'03)*, Sanibel Island, Florida, USA, October 2003.

[MTCP04] A. Magkanaraki, V. Tannen, V. Christophides, and D. Plexousakis. Viewing the semantic web through RVL lenses. *Journal of Web Semantics*, 2004. To appear.

[PL03] D. Peterson and M. Levene. Trail records and navigational learning. *London Review of Education*, 1(3):207–216, 2003.

[PPW03a] G. Papamarkos, A. Poulovassilis, and P. T. Wood. Event-condition action rules languages for the semantic web. In *1st Workshop on Semantic Web and Databases, Berlin*, 2003.

[PPW03b] G. Papamarkos, A. Poulovassilis, and P. T. Wood. SeLeNe Deliverable 4.4: ECA rule languages for active Self e-Learning Networks, 2003. Available from `http://www.dcs.bbk.ac.uk/selene/reports/`.

[PPW04] G. Papamarkos, A. Poulovassilis, and P. T. Wood. RDFTL: Event-condition-action rule language for RDF. In *Proceedings of the 3rd International Workshop on Web Dynamics, at WWW'04*, New York, USA, May 2004.

[RS03a] P. Rigaux and N. Spyratos. Generation and syndication of learning object metadata. Technical Report 1371, Laboratoire de Recherche en Informatique, 2003.

[RS03b] P. Rigaux and N. Spyratos. Metadata management and learning object composition in a Self e-Learning Network. In *Proceedings of the 1st Franco-Japanese Workshop on Information Search, Integration and Personalization (ISIP'03)*, Sapporo, Japan, June 30–July 2 2003.

[RS03c] P. Rigaux and N. Spyratos. SeLeNe Deliverable 4.1: Generation and syndication of learning object metadata, 2003. Available from `http://www.dcs.bbk.ac.uk/selene/reports/`.

[SCKM03] M. Stratakis, V. Christophides, K. Keenoy, and A. Magkanaraki. SeLeNe Deliverable 2.1: E-Learning Standards, 2003. Available from `http://www.dcs.bbk.ac.uk/selene/reports/`.

[SeL04] The SeLeNe Consortium. SeLeNe Deliverable 5: An architectural framework and deployment choices for SeLeNe, 2004. Available from `http://www.dcs.bbk.ac.uk/selene/reports/`.

[SKC03] G. Samaras, K. Karenos, and E. Christodoulou. SeLeNe Deliverable 3: A Grid Service framework for Self e-Learning Networks, 2003. Available from `http://www.dcs.bbk.ac.uk/selene/reports/`.

[SKPC] G. Samaras, K. Karenos, E. Pitoura, and P.K. Chrysanthis. ViSMA: Extendible, mobile-agent based services for the materialization and maintenance of personalized and shareable web views. In *Proceedings of the 6th International Workshop.*

[TASC03] Y. Tzitzikas, A. Analyti, N. Spyratos, and P. Constantopoulos. An algebraic approach for specifying compound terms in faceted taxonomies. In *Proceeedings of the 13th European-Japanese Conference on Information Modelling and Knowledge Bases*, ita-Kyushu, Japan, 2003.

[TCK+03] V. Tannen, V. Christophides, G. Karvounarakis, I. Koffina, G. Kokkinidis, A. Magkanaraki, D. Plexousakis, and G. Serfiotis. The ICS-FORTH SWIM: A powerful semantic web integration middleware. In *Proceedings of the Workshop on Semantic Web and Databases at VLDB'03*, 2003.

Appendices

A Participation Statistics

A.1 Number of Attendees

We expected the number of participants external to the SeLeNe consortium to be around 75% of the total attendance, and so set a minimum of 30 external attendees out of 50 as a target.

	Number	Percentage
SeLeNe consortium members	11	26%
Non-SeLeNe participants	32	74%
Total	43	

Table 1: Attendance figures for the Workshop

As can be seen from Table 1, the target of 30 external attendees was met, despite there being slightly fewer participants overall than anticipated.

A.2 IST Participation

We set a target that a minimum of 50% of attendees should be from the IST community. In calculating the figures we have excluded members of the SeLeNe consortium, who also are obviously "IST participants".

	Number	Percentage
Attendees involved in IST projects	18	56%
Non-IST Attendees	14	44%
Total	32	

Table 2: IST participation in the Workshop

As can be seen from Table 2, we met our target for IST participation in the Workshop.

A.3 Breadth of European Representation

The targets we set for European representation were that participants would be from at least 6 different countries, and that about 60% of the attendees would be from outside the UK.

In fact, 8 different European countries were represented at the Workshop: Austria, Belgium, Cyprus, France, Germany, Greece, Italy and the UK, meaning that our target for

	Non-SeLeNe participants	All participants
United Kingdom	18 (56%)	25 (58%)
Non-UK	14 (44%)	18 (42%)
Total	32	43

Table 3: UK vs. non-UK participation in the Workshop

the number of countries represented was exceeded. However, as can be seen from Table 3, the overall percentage of non-UK participants fell short of our target. This is likely to be the result of higher expenses for participants from abroad (who had to meet their own travel costs).

A.4 Presentation of IST results

As this was an IST event, we set a target for a minimum of 50% of the speakers to be from, and to present results of, relevant IST projects.

	Number of Presentations	Percentage
Presentations of IST Results	9	64%
Non-IST Presentations	5	36%
Total	14	

Table 4: Number of IST presentations at the Workshop

As can be seen from Table 4, the proportion of talks presenting IST project results was above our target.

B Responses to the Evaluation Questionnaires

Participants at the Workshop were asked to complete two evaluation questionnaires to assess the quality of the event — one at the end of the morning and one at the end of the afternoon sessions. The afternoon questionnaire contained additional questions about the quality of the Workshop as a whole. Both questionnaire forms can be found in Appendix C.

The responses (only from participants who are not part of the SeLeNe consortium) are summarised in the sections below. The results are compared with the target that we set ourselves:

We expect to obtain an overall average score of at least 4 for each of the morning and afternoon sessions and for the Workshop as a whole.

B.1 Morning Sessions

A total of 26 evaluation questionnaires were returned completed at the end of the morning session. The responses were as follows:

1. Were the talks in the morning session stimulating?

	not at all 1	**a little** 2	**average** 3	**quite** 4	**very much so** 5
Number of respondants	0	0	5	16	5

Average score: 4

2. How relevant were the morning's talks to the aims of the Workshop?

	not at all 1	**a little** 2	**fair** 3	**quite** 4	**very relevant** 5
Number of respondants	0	0	2	14	10

Average score: 4.31

3. How informative were the morning's presentations?

	not at all 1	**weak** 2	**average** 3	**good** 4	**very informative** 5
Number of respondants	0	1	4	17	4

Average score: 3.92

4. How satisfied were you overall with the morning sessions?

	not at all 1	**somewhat** 2	**fairly** 3	**very** 4	**entirely satisfied** 5
Number of respondants	0	0	4	15	7

Average score: 4.11

The overall average score for the morning session (i.e. the average score for all questions) is 4.09, meeting our target of 4.

B.2 Afternoon Sessions

A total of 17 evaluation questionnaires were returned completed at the end of the afternoon session. The responses were as follows:

1. Were the talks in the afternoon session stimulating?

	not at all 1	**a little** 2	**average** 3	**quite** 4	**very much so** 5
Number of respondants	0	1	6	4	6

Average score: 3.88

2. How relevant were the afternoon's talks to the aims of the Workshop?

	not at all 1	**a little** 2	**fair** 3	**quite** 4	**very relevant** 5
Number of respondants	0	0	2	9	6

Average score: 4.23

3. How informative were the afternoon's presentations?

	not at all 1	**weak** 2	**average** 3	**good** 4	**very informative** 5
Number of respondants	0	3	4	5	5

Average score: 3.7

4. How satisfied were you overall with the afternoon sessions?

	not at all 1	**somewhat** 2	**fairly** 3	**very** 4	**entirely satisfied** 5
Number of respondants	0	1	6	6	4

Average score: 3.76

The overall average score for the afternoon session (i.e. the average score for all questions) is 3.90, which falls slightly short of our target of 4. This is slightly disappointing, but the afternoon has still been judged to be well above average.

B.3 Workshop as a Whole

The final two questions on the afternoon questionnaire concerned the Workshop as a whole. The responses (again, from 17 people) were as follows:

5. Was it a useful forum for the exchange of ideas?

	not at all 1	**weak** 2	**fair** 3	**good** 4	**very useful** 5
Number of respondants	0	0	0	10	7

Average score: 4.41

6. How do you rate the organisation of the Workshop?

	very weak 1	**weak** 2	**fair** 3	**good** 4	**very good** 5
Number of respondants	0	0	0	9	8

Average score: 4.47

The overall average score for the questions about the Workshop as a whole (i.e. the average score for both questions) is 4.44, exceeding our target of 4.

B.4 Comments and Suggestions

The final question, which was optional, asked for further comments and suggestions. Few people (in fact, four) gave comments, but where they did they were positive, describing the event as "...an excellent Workshop which promoted exchange of state-of-the-art knowledge and ideas" and "...a very useful occasion". There was a suggestion to post the presentations on the Website, which had always been our intention and has now been done. Another participant suggested that a printed version of the contributions should be produced, but we feel that as a small, informal Workshop the availability of all material on the Web is sufficient.

C Evaluation Questionnaires

C.1 Morning

IST Workshop on Metadata Management in Grid and P2P systems (MMGPS) Models, Services and Architectures

Morning Questionnaire

1. Were the talks in the morning session stimulating?

1	-	2	-	3	-	4	-	5
(not at all)		(a little)		(average)		(quite)		(very much so)

2. How relevant were the morning's talks to the aims of the Workshop?

1	-	2	-	3	-	4	-	5
(not at all)		(a little)		(fair)		(quite)		(very relevant)

3. How informative were the morning's presentations?

1	-	2	-	3	-	4	-	5
(not at all)		(weak)		(average)		(good)		(very informative)

4. How satisfied were you overall with the morning sessions?

1	-	2	-	3	-	4	-	5
(not at all)		(somewhat)		(fairly)		(very)		(entirely satisfied)

C.2 Afternoon

IST Workshop on Metadata Management in Grid and P2P systems (MMGPS) Models, Services and Architectures

Afternoon Questionnaire

The Afternoon session

1. Were the talks in the afternoon session stimulating?

1	-	2	-	3	-	4	-	5
(not at all)		(a little)		(average)		(quite)		(very much so)

2. How relevant were the afternoon's talks to the aims of the Workshop?

1	-	2	-	3	-	4	-	5
(not at all)		(a little)		(fair)		(quite)		(very relevant)

3. How informative were the afternoon's presentations?

1	-	2	-	3	-	4	-	5
(not at all)		(weak)		(average)		(good)		(very informative)

4. How satisfied were you overall with the afternoon sessions?

1	-	2	-	3	-	4	-	5
(not at all)		(somewhat)		(fairly)		(very)		(entirely satisfied)

The Workshop as a whole

5. Was it a useful forum for the exchange of ideas?

1	-	2	-	3	-	4	-	5
(not at all)		(weak)		(fair)		(good)		(very useful)

6. How do you rate the organisation of the workshop?

1	-	2	-	3	-	4	-	5
(very weak)		(weak)		(fair)		(good)		(very good)

7. Please put any further comments and suggestions on the reverse of this form. (Optional)

D List of Participants

Name	Country
Kevin Keenoy	UK
George Kokkinidis	Greece
Mark Levene	UK
George Loizou	UK
George Papamarkos	UK
Donald Peterson	UK
Alexandra Poulovassilis	UK
Phillipe Rigaux	France
George Samaras	Cyprus
Nicolas Spyratos	France
Peter Wood	UK

Table 5: SeLeNe Project Workshop Attendees

Name	Country	IST
Mohammed Alfreid	UK	N
Bob Bentley	UK	Y
Mario Cannataro	Italy	Y
Stavros Christodoulakis	Greece	Y
Simon Courtenage	UK	N
Philippe Cudré-Mauroux	France	N
Kevin Davis	UK	N
Theo Dimitrakos	UK	Y
Haya El-Ghalayini	UK	N
Hao Fan	UK	N
Tamás Hauer	UK	Y
Klaus Janke	Germany	Y
Maria Margetti	UK	N
David Massart	Belgium	Y
Gavin McCance	UK	Y
Carlo Meghini	Italy	Y
Zoltán Miklós	Austria	Y
Nektarios Moumoutzis	Greece	Y
Wolfgang Nejdl	Germany	Y
Henrik Nottelmann	Germany	Y
Eoghan O'Neill	Belgium	Y
Kevin O'Neill	UK	N
Savas Parastatidis	UK	N
Dmitry Rogulin	UK	Y
Eric Scharf	UK	N
Michel School	France	N
Faezeh Seyedarabi	UK	N
Tony Solomonides	UK	Y
Frans Van Assche	Belgium	Y
Pierangelo Veltri	Italy	Y
Stephen Williams	UK	N
Darioush Yarand	UK	N

Table 6: Non-SeLeNe Workshop Attendees