Semantic Grid E-Learning Platform For Education System

S. A. M. Rizvi¹, S. Z. Hussain² and Muqeem Ahmed³

¹,²,³Department of computer science Jamia Millia Islamia New Delhi 110025
E-mail: ¹Samsang_rizvi@yahoo.com, ²szhusain@rediffmail.com, ³Muqeem.ahmed@gmail.com

ABSTRACT

E-learning is shared scenario, educational resources, such as course documents, videos, test-bases, courseware, and teacher information etc., across different schools. Dart Grid is a semantic grid toolkit for data integration using technologies from Semantic Web and Grid. It is very challenging to integrate all the components of Semantic grid E-learning platform into a single framework because of interoperability issues. The implementation of Semantic grid E-learning will follow a bottom up approach to better understand and mask the heterogeneity issues of multiple tools and technologies.

In this paper, a Semantic Grid for E-learning based platform is introduced, and provides a Semantic-based distributed infrastructure for E-learning data resource sharing. Which evolve an end-to-end E-learning infrastructure from the integration of available technologies, specifically the semantic web, the grid, collaborative and personalization tools, and knowledge management techniques.

KEYWORDS
Semantic Grid E-Learning, Semantic Grid, Collaborative Learning, Personalization, knowledge, platform.

1. INTRODUCTION

The Semantic Web is a web of information which is used to process data directly and in directly by machine to improve the present Web by making Web resources machine-understandable by enriching present Web resources with machine understandable semantics. Tim Berners-Lee, director of the W3C, referred to the future of the present web as the "semantic web"—extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation[1]. It supports a common platform that allows information to be shared and reused across applications, enterprises, and community boundaries. It is based on the Resource Description Framework (RDF), which integrates a variety of applications using XML as syntax and URI for naming.

The Grid [2] tries to connect a wide variety of geographically distributed resources such as Personal Computers, workstations and clusters, storage systems, data sources, databases and special purpose scientific instruments and presents them as an integrated resource, and it is a technology that makes it possible for distributed computing resources to be shared, managed, coordinated, and controlled.

The Semantic Grid [3] is an Internet-centralized interconnection environment that can effectively organize, share, cluster, fuse, and manage globally distributed resources based on the interconnection semantics. In short, the Semantic Grid vision is to achieve a high degree of easy-to-use and seamless automation in an effort to provide flexible collaborations and computations on a global scale. It takes advantage of machine-understandable knowledge on the Grid. The Semantic Grid is an extension of the current Grid in which information and services are given well defined and explicitly represented meaning, better enabling computers and people to work in cooperation.

Internet and the World Wide Web gave the idea to introduce the concepts of E-learning and collaborative knowledge sharing, but due to largely unplanned and unanticipated growth, are now falling short of earlier promises. Lack of machine readable content coupled with data overload has put strains into the manually knowledge delivery model of WWW. The situation is especially serious in the E-learning domain where the success and usefulness directly correlates with the effectiveness of knowledge delivery in a dynamic setting. A large number of research efforts are hence focusing on a planned infrastructure development for E-learning [4, 5, 6].

In this paper we introduce a Semantic-grid based E Learning environment. Rather than proposing a strategy of development from scratch and ending-up with yet another monolithic structure with low integration capabilities, we suggest an integrated approach that involves minimal re-work of existing systems. We firstly introduce the key enablers for a realistic E-learning environment and map these to available technologies to establish a well-defined environment for the integration of these technologies so that the goals of effective E-learning can be achieved.

2. THE SEMANTIC GRID IN E-LEARNING

The basic architecture behind the World Wide Web is not capable of providing a seamless and artificially-intelligent environment required for a large scale effective and efficient E-learning implementation. Hence, several research works worldwide are focusing on resolving this issue.

The concept of grid computing support for E learning has long received criticism from various quarters. Critics hold the view that the modern day WWW architectures, tools and technology support nearly every feature required by E-learning and thus incorporating the grid is largely unnecessary for E-learning support. We are studying this assertion taking into consideration the evolved form of Computer Supported Collaborative Learning (CSCL) around the world supported by grid computing.

The support for dynamism in terms of resources, content and participants may be considered as a core grid based
architectural feature to support effective E-learning strategies. A closer look into any E-learning based infrastructure would identify the highly dynamic pattern of its key ingredients i.e. resources (both content and computation) and participants. The management of such dynamically changing environments, being a key task of grid computing, needs to be further extended seamlessly to E-learning environments. Since both the grid and the WWW hold certain strengths as far as E-learning support is concerned, we propose to create a synergy by incorporating the so-called ‘semantic grid’ in E-learning. The semantic grid merges the semantic web with grid computing. Also, incorporating semantic grid in E-learning will provide the best seamless support available through a merger of the best of both paradigms.

3. A SEMANTIC-GRID BASED E-LEARNING PLATFORM

The proposed semantic grid-based framework is presented in Figure 1. A service oriented semantic grid middleware lies at the core of Semantic grid E learning Environment. A Service Oriented Architecture will enable a loosely coupled interrelationship between Collaborative Partners (CPs) and provide a higher level of abstraction in the form of open interfaces. We propose a layered Semantic grid e learning environment stack at each Virtual Organization (VO) with two major segments corresponding to both E-learning and semantic grid applications/services. Such a layered approach gives a better understanding of interaction between the various components. This layered Service oriented architecture approach allows us to decouple the independent components of Semantic grid E learning platform. We present below a brief functional description of each of these layers.

3.1. E-LEARNING APPLICATIONS LAYER

The top layer in the Semantic grid E Learning Environment VO stack will carry various end-user applications such as group and courseware managers, search facilities, scheduling and tracking software etc. Of course all of these applications will be dependent on and controlled by the specific requirements of respective end users. Possible examples of such applications could fall in the domains of E-Teaching, E Training, E-Workshop and E-Conference.

3.2. E-LEARNING SERVICES LAYER

The development of applications will be facilitated by a set of generic application-level services such as collaboration tools, agents and personalization managers. A recent application of collaboration tools and services can be seen in CoAKTing [6], which provides services such as the status of collaborative partners, discussion minutes, meeting status, things to do list, project status etc. The benefit of decoupling applications and services into separate layers is twofold. First, it will minimize reworking and increased maintainability of applications by the result of high cohesiveness and loose coupling.

Second, it will also ensure compliance to some standardization criteria during application development.

Personalization Services (PS) may impart an important role by personalizing the individual centric information. That is, if someone is interested in lecture materials of some special domain, then the PS executing as a backend process will both keep track of such information based on the content usage and reduce the latency involved in information retrieval. Also, the PS can be deployed at the site level so that each individual’s required information can be kept up-to-date. It could also change the traditional learning processes from strong push delivery, lack of personalization and the linear/static learning processes to efficient, distributed, student-oriented, personalized, and non-linear/dynamic learning processes [14]. Readers interested in a detailed comparison of the characteristics of traditional learning process vis-à-vis E-learning process are referred to [14]. The PS will be based on specific policies and indexing approaches determined by the interest of users (either defined explicitly by the users or inferred through usage patterns). A common theme among E-learning applications in a collaborative environment is to provide intelligent search, matching and inference support. Our basic hypothesis in this regard is that a small set of generic inference services can cater for a large number of applications. A typology of the proposed inference services is shown in figure 2.

Resource and content matching becomes a major issue since the use of a description language still does not standardize services and content description to the extent where a direct string-based matching can be applied. We believe that the required matching fits nicely in to the problem of semantic data matching from the information theory paradigm. Also known as in-exact data matching, semantic data matching as the name suggests is matching on the basis of semantics or meanings of data rather than its character or literal formation. Probably the most successful application of semantic data matching is in the
web search engines where terms and documents are co-clustered on the basis of semantics.

**Figure 2: A typology of inferences services**

Latent Semantic Analysis (LSA) is a powerful semantic matching technique [13]. It is used to extract semantic similarity between pieces of textual information using statistical techniques. Similarly, there also exist other techniques, which involve graphical classification of the concepts and then semantic matching based on set theory and AI concepts [14]. A significant body of work could be found in the literature where such semantic or inexact matching techniques are proposed or applied to the grid domain [9, 10]. Although it is very challenging to integrate all the inference services together in one framework because of the interoperability issues, the advantages are that it will further strengthen the Semantic grid for information retrieval. To ensure appropriate support for the inference services each VO must maintain a knowledge base (e.g. for educational applications a repository of lectures, videos, tutorials, experiment designs and results and proof of experiments etc.). This may be done by deploying an artifact management system [8], which will maintain the documentation, processes, researcher or tutor profiles for future reuse. Since different ethnic groups from heterogeneous locations will be sharing their logical resources in a collaborative E-Learning environment, there must be some comparable standardized form of lectures, tutorials, videos etc. This could enable Collaborative Partners to be compliant within a shared environment. Some standards like the Educational Modeling Language (EML) from the Open University of Netherlands, IMS Enterprise Services for managed learning environments [8] already exist.

The following list may be considered as a minimal set of parameters required for an “exact” information search approximately equivalent to the IEEE Learning Object Metadata (LOM) model. These parameters could best be defined using ontologies from the semantic web domain.

- **Domain**: the major category of related material, whether it belongs to medical sciences, computer sciences, astronomy or any other domain.
- **Type**: the type of document e.g. ppt, pdf, doc, avi, wav
- **Author**: the author(s) who generated a resource.
- **Constraints**: the basic requirements like security restrictions or specific tools to execute a specific request.
- **Description**: the short description of a resource.
- **Size/Capacity**: the total size of a resource. If it is a document then this is measured in bytes or if it is some computational resource then its computation power in Hz.
- **Location**: the address/location of the resource.
- **Metadata**: the address of the resource.

3.3. THE SEMANTIC GRID AND SEMANTIC GRID E-LEARNING PLATFORM

The strongest and most innovative component of Semantic grid E Learning Platform is its core component, the semantic grid. When several VOs are participating, the problems of heterogeneity and low standardization limit the applicability of a conventional grid. The potential benefits of the semantic grid approach over conventional grids can be numerous. In a conventional grid infrastructure, VOs are required to agree upon a framework for resource description so that the grid services can share, locate and apply security measures [9, 10]. In a dynamically changing and evolving environment, such a requirement constrains the scalability of this framework. This is exactly where we anticipate that the semantic grid can help in overcoming the resource description problem using model-theoretic solutions from the semantic web domain. The semantic grid envisions a well-structured integration of semantic web and classical grid technologies.

In a recent work, Goble & De Roure provides three possible options to deal with dynamism and heterogeneity in a semantic grid environment [11]. These include knowledge management techniques, semantic grid services and Multi-Agent Systems (MAS). Among these options, MAS is probably the most innovative where intelligent software agents can negotiate and translate descriptions within the collaborating environment. Agents can also play a key role in extracting important information from required resources. Agents could also enable certain advance features such as masking-off the heterogeneity of structured information and support for autonomous asynchronous operations without affecting the normal workflow of other processes.

3.4. SECURITY INFRASTRUCTURE IN SEMANTIC GRID E-LEARNING PLATFORM

Security issues such as the authorized access to shared resources, the conservation of intellectual property rights, the confidentiality of contents, the authentication of individuals, and the auditing of resource usage play an important role in the smooth and controlled working of any distributed system. In a grid environment, it is difficult to manage the workflow of the complete system in a secure manner. Despite the abovementioned security requirements, several VOs which can become partners and form virtual markets in a single grid environment, might have different security policies and mechanisms to provide secure and controlled access to resources. Therefore a common mechanism which can mask
the heterogeneity of security policies is needed either in the form of wrappers (as in GRASP [15]) or common policy description languages (as in XACML [16]). Further more, we believe that in Semantic grid E Learning, we need a security infrastructure to support both the underlying core of grid (Grid Security Infrastructure – GSI) and its collaborating applications layer using security add-on components.

<table>
<thead>
<tr>
<th>Semantic grid E-learning Component</th>
<th>Tool/Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborative Tools</td>
<td>CoAKTing</td>
</tr>
<tr>
<td>Personalization Services</td>
<td>CHANDLER</td>
</tr>
<tr>
<td>Inference Service</td>
<td>OILED,JESS,Description Logic</td>
</tr>
<tr>
<td>Software Agents</td>
<td>SAGE[15],JADE[19],Aglets</td>
</tr>
<tr>
<td>Grid component and services</td>
<td>JClarens,GT3</td>
</tr>
<tr>
<td>Ontologies</td>
<td>Protégé etc</td>
</tr>
<tr>
<td>Semantic web Technologies</td>
<td>OWL,DAML-OIL</td>
</tr>
<tr>
<td>security</td>
<td>Assorted security</td>
</tr>
</tbody>
</table>

**TABLE 1:** Tools & technologies for Semantic E learning platform

4. ENABLING TECHNOLOGIES FOR SEMANTIC GRID E-LEARNING

As the popularity of the grid grows, a large number of open tools and technologies are being developed worldwide. The above Table specifies some open tools and technologies that we plan to use in the implementation phase. Not all of the tools or technologies are required for specific components; rather these are the multiple implementation options available. Figure 3 presents an abstract deployment model of Semantic Grid E Learning, showing the components needed at each site. The following abstract example explains the workflow of Semantic Grid Learning with the help of a suitable scenario. Assume that all participants from different sites have deployed all the layered components needed for Semantic Grid Learning. Suppose one user from site A intends to go through lectures/contents of some specific subject. His/Her knowledge query by using Semantic grid E Learning layers will be given to software agents as well as personalization services to keep track of user interests. Software agents by communicating with collaborative partner agents using inference services will resolve query for demanding resources based on the parameters outlined in section 3.2. Software agents and inference services will have to access the grid resource descriptions through access services, information monitoring and management services.

**Figure 3:** An abstract deployment model for Semantic grid E-Learning

This access to agents will also be provided based on the inference services specific to knowledge base of grid middleware, which will keep track of the resources descriptions and their usage status. It seems very trivial but it is very difficult to integrate all the technologies mentioned in figure 1 and table 2. Important and critical aspect of the Semantic grid is its adaptive learning, updating the knowledge bases for future requests and collaborative tools for meetings and participants’ information etc. In the proposed implementation scheme, we have purposely limited ourselves to open source technologies only. The biggest implementation challenge is to integrate various technologies in an efficient and scalable manner.

Precedents of such integration could be found in works such as the integration of web services with multi agent systems, the integration of collaboration tools in grid environment [6] and the incorporation of security features in JClarens (VOMS; Virtual Organization Membership Service). The Semantic grid E Learning implementation results will enable the suitability of Semantic grid E Learning Platform for CSCL to be evaluated.

5. RELATED WORK

With increasing intellectual and commercial collaborations across the globe, the E-learning domain is a rapidly developing field as demonstrated by the number of technologies that we have referred to in this paper. It is therefore important to locate the current work with respect to other research and it is in this light that we review some of the recent works which have informed and motivated the current research.

5.1. ONTOEDU: ONTOLOGY-BASED EDUCATION GRID SYSTEM FOR E-LEARNING

OntoEdu, a project of the University of Peking, China, is the most recent work wherein an ontology based grid is proposed for educational applications. Using educational technologies at
its crux, the OntoEdu architecture realized concept reusability with ontology, device and user adaptability with ubiquitous computing and automatic composition. Although the OntoEdu architecture is quite innovative and extensive in nature, its primary focus is oriented towards adaptability (personalization). Designers have not referred to the incorporation of some equally important areas such as special collaboration tools or services and intelligent search and matching agents.

5.2. COAKTING: COLLABORATIVE ADVANCED KNOWLEDGE TECHNOLOGIES IN THE GRID

CoAKTing provides a motivating example of the incorporation of collaborating technologies on top of a grid structure. Geared towards academic and intellectual collaborations, CoAKTing is a set of collaborating tools that enables enhanced process tracking and navigation of resources before, after, and during meetings in progress. These tools work through a shared ontology and could be integrated in an existing collaborative environments (such as the Access Grid).

Each of the CoAKTinG tools can be thought of as extracting structure from the collaboration process. The CoAKTing project has introduced tools such as Buddy Space for presence awareness, Compendium for keeping track of a bundle of ideas, issues and conceptual Inter relationships involved in projects, I-X Process Panels and Meeting Replay. Interested readers are referred to and for more details on working and features of these tools. The set of CoAKTing tools can be useful at collaboration services layer of semantic grid learning. We agree, with where they summarize that ‘the CoAKTing tools can be transposed into the Learning Grid’.

5.3. APPLE: A Novel P2P based e-Learning Environment

The APPLE (A novel P2P based e-Learning Environment) project emphasizes the importance of grid and P2P infrastructures for e-Learning applications instead of a static web. This work proposes the use of the grid for group-centric and P2P for individual-centric information retrieval. The designers of APPLE used WSRF.NET to develop and deploy a virtual classroom service. They integrated a P2P platform with the grid to exploit extensive resource potential from the grid. Despite being an extensive framework, a major limitation of APPLE seems to be its dependency on a proprietary Microsoft technology (WSRF.NET). Moreover, intelligent semantic matching structures, personalization and collaboration technologies have not been explicitly addressed in the original APPLE proposal. In a larger sense, the use of P2P (as in APPLE) or the Grid have similar final objectives — the pooling and coordinated use of large sets of distributed resources. These technologies work with the same approach to solving the problems but target different communities, resources and applications.

In an important paper, Foster and I amnitchi state that the complementary nature of the strengths and weaknesses of the two approaches suggests that the interests of the two communities (grid and P2P) are likely to grow closer over time. In the same spirit, the designers of APPLE incorporate the strengths of both technologies by adopting hybrid architecture. The Semantic grid E Learning environment philosophy further enhances this approach by introducing the semantic grid based underlying middleware with reasoning support for easy service discovery and request submission, software agents for intelligent negotiation and collaborating tools for the purpose of collaborative activities like meetings, things to do list etc.

5.4. Other Related Works

An exciting work in collaborative learning is the Access Grid project of Argonne National Labs. Currently deployed at 150 institutions around the world, the Access Grid is a multicast videoconference technology that enables its users to conduct real-time virtual conferences and maintain a wholesome online groupware. Boldyreff et. al. have explored the concept of shared artifacts over the grid. All the resources, such as documentation including architectural details, design documents, test cases, process definitions and details, researcher or partner profiles are considered to be the artifacts which can be shared for future reuse over the grid. Boldyreff et. al. also made an effective analogy between collaborative software development and collaborative learning by highlighting the significance of shared artifacts over the grid.

Some recent works have also been reported in the domain of the integration of semantic web technologies either in the form of deploying translators or using ontology. Ontology based matchmakers. The induction of semantics in the grid will further improve the collaborative efforts in different domains. Large scale projects including DILIGENT and BRICKS are underway for the integration of digital libraries for collaborative heterogeneous knowledge sharing within grid environments.

5. CONCLUSION

An effective, end-to-end and practical E-learning environment cannot be realized from a loose integration of available technologies or by starting the development from scratch. The former approach could lead to an unrealistic and non-scalable infrastructure, while the latter strategy might end up with wasteful rework. A rather efficient approach requires, (i) an understanding of key enablers behind the target E-learning infrastructure,(ii) a comparative analysis of available tools and technology on the basis of customizability, applicability and cost, (iii) a mapping of key E-learning enablers onto the available technology, and (iv) a detailed architecture that specifies the interaction among the technological solutions at various levels. In this paper, we have attempted to evolve an end-to end E-learning infrastructure from the integration of available technologies, specifically the semantic web, the grid, collaborative and personalization tools, and knowledge management techniques. We understand that it will be very challenging to integrate all the components of Semantic grid into a single framework because of interoperability issues. The
implementation of Semantic grid E Learning will follow a bottom up approach to better understand and mask the heterogeneity issues of multiple tools and technologies. A review of recent research and development work in this domain suggests the need for the development of such a large number of tools that may be used at various levels in the proposed Semantic grid E Learning architecture. Finally, the outcomes of this research at this initial stage may be regarded as a step forward to disseminate ideas on a proposed semantic and grid-based architecture for the effective understanding, integration, and deployment of E-learning applications based on a proposed framework for E-learning.

6. FUTURE SCOPE
The vision of the Semantic Grid as a future e-Science infrastructure in which there is a high degree of easy-to-use and seamless automation and in which there are flexible collaborations and computations on a global scale. In order to make the Semantic Grid a reality, a number of research challenges need to be addressed. These include (in no particular order):

- Service-Oriented Architectures. Research the provision and implementation of grid facilities in terms of service oriented architectures. Also research into service description languages is a way of describing and integrating the grid’s problem solving elements.
- Agent Based Approaches. Research the use of agent based architectures and interaction languages to enable e-Science marketplaces to be developed, enacted and maintained.
- Trust and Provenance. Further research is needed to understand the processes, methods and techniques for establishing computational trust and determining the provenance and quality of content in Grid systems. This extends to the issue of digital rights management in making content available.
- Metadata and Annotation. Whilst the basic metadata infrastructure already exists in the shape of RDF, metadata issues have not been fully addressed in current grid deployments. It is relatively straightforward to deploy some of the technology in this area, and this should be promoted. RDF, for example, is already encoding metadata and annotations as shared vocabularies or ontologies. However, there is still a need for extensive work in the area of tools and methods to support the design and deployment of e-Science ontologies. Annotation tools and methods need to be developed so that emerging metadata and ontologies can be applied to the large amount of content that will be present in Grid applications.
- Knowledge Technologies. In addition to the requirement for the research in metadata and annotation, there is a need for a range of other knowledge technologies to be developed and customised for use in e-Science contexts. These include knowledge capture tools and methods, dynamic content linking, annotation based search, annotated reuse repositories, natural language processing methods (for content tagging, mark-up, generation and summarisation), data mining, machine learning and internet reasoning services. These technologies will need shared ontologies and service description languages if they are to be integrated into the e-Science workflow. These technologies will also need to be incorporated into the pervasive devices and smart laboratory contexts that will emerge in e-Science.
- Integrated Media. Research into incorporating a wide range of media into the e-Science infrastructure. This will include video, audio, and a wide range of imaging methods. Research is also needed into the association of metadata and annotation with these various media forms.
- Content Presentation. Research is required into methods and techniques that allow content to be visualised in ways consistent with the e-Science collaborative effort. This will also involve customising content in ways that reflect localised context and should allow for personalisation and adaptation.
- E-Science Workflow and Collaboration. Much more needs to be done to understand the workflow of current and future e-Science collaborations. Users should be able to form, maintain and disband communities of practice with restricted membership criteria and rules of operation. Currently most studies focus on the e-Science infrastructure behind the socket on the wall. However this infrastructure will not be used unless it fits in with the working environment of the e-Scientists. This process has not been studied explicitly and there is a pressing need to gather and understand these requirements. There is a need to collect real requirements from users, to collect use cases and to engage in some evaluative and comparative work. There is also a need to more fully understand the process of collaboration in e-Science.
- Pervasive E-Science. Currently most references and discussions about grids imply that their primary task is to enable global access to huge amounts of computational power. Generically, however, we believe grids should be thought of as the means of providing seamless and transparent access from and to a diverse set of networked resources. These resources can range from PDAs to supercomputers and from sensor’s and smart laboratories to satellite feeds.
- E-Anything. Many of the issues, technologies and solutions developed in the context of e-Science can be exploited in other domains where groups of diverse stakeholders need to come together electronically and interact in flexible ways.
- Thus it is important that relationships are established and exploitation routes are explored with domains such as e-Business, e-Commerce, e-Education, and e-Entertainment.

7. REFERENCES
[1]. Huajun Chen, Zhaohui Wu, and Jiefeng Xu” KB-Grid: Towards Building Large-Scale Knowledge System in Semantic Web” Springer-Verlag Berlin Heidelberg 2003